Published in final edited form as:

ACS Synth Biol. 2019 May 17; 8(5): 1010–1025. doi:10.1021/acssynbio.8b00450.

PERSIA for Direct Fluorescence Measurements of Transcription, Translation, and Enzyme Activity in Cell-Free Systems

Scott Wick¹, David I Walsh III¹, Johanna Bobrow¹, Kimberly Hamad-Schifferli², David S. Kong³, Todd Thorsen¹, Keri Mrosczyk¹, Peter A. Carr^{1,4,*}

¹.MIT Lincoln Laboratory, Lexington, MA, USA

².University of Massachusetts Boston, Department of Engineering, Boston, MA, USA

³.Media Lab, Massachusetts Institute of Technology, Cambridge, MA, USA

⁴.Synthetic Biology Center at MIT, Cambridge, MA, USA

Abstract

Quantification of biology's central dogma (transcription and translation) is pursued by a variety of methods. Direct, immediate and ongoing quantification of these events is difficult to achieve. Common practice is to use fluorescent or luminescent proteins to report indirectly on prior cellular events, such as turning on a gene in a genetic circuit. We present an alternative approach, PURExpress-ReAsH-Spinach In-vitro Analysis (PERSIA). PERSIA provides information on the production of RNA and protein during cell-free reactions by employing short RNA and peptide tags. Upon synthesis, these tags yield quantifiable fluorescent signal without interfering with other biochemical events. We demonstrate the applicability of PERSIA in measuring cell-free transcription, translation, and other enzymatic activity in a variety of applications - from sequence-structure-function studies, to genetic code engineering, to testing antiviral drug resistance.

Graphical Abstract

Keywords

In-Vitro Transcription/Translation; Cell-Free; Genetic Prototyping

Delivered to the U.S. Government with Unlimited Rights, as defined in DFARS Part 252.227-7013 or 7014 (Feb 2014). Notwithstanding any copyright notice, U.S. Government rights in this work are defined by DFARS 252.227-7013 or DFARS 252.227-7014 as detailed above. Use of this work other than as specifically authorized by the U.S. Government may violate any copyrights that exist in this work.

*Corresponding Author: Peter Carr, Bioengineering Systems & Technologies, MIT Lincoln Laboratory, 244 Wood St, Lexington, MA 02420.

Advances in nucleic acid synthesis have greatly facilitated investigation of both fundamental (“what is the function of this DNA sequence?”) and applied (“does my DNA-based design work?”) questions in transcription and translation. Cell-based extracts are now a critical tool in DNA design as they enable rapid assessment of sequences. For example, the use of S30 extracts in elucidating the codons of the genetic code and the role that tRNAs play in the process of protein synthesis^{1, 2}, to prototyping genetic circuits^{3–7}, or enzymes, such as for metabolic engineering.^{8–10} While mammalian cell lysates have been a common tool for exploring eukaryotic systems, there are current efforts to develop HeLa cell-free expression systems to study bottom-up mammalian synthetic applications.¹¹ For all of these applications, the ability to quantify transcription, translation, and protein activity of the DNA design is essential.

Cell-free systems are often used to produce preliminary data before *in vivo* experiments. To quantify protein products, the protein must be purified from the cell-free system and identified. This is predominantly achieved by appending the sequences with polypeptide tags, such as GST¹², 6x His¹³, and streptavidin.¹⁴ Other polypeptide tags provide a method of identifying the presence of the recombinant protein as it is expressed, with small peptides like FLAG¹⁵ and Hemagglutinin-HA.¹⁶ The use of these tags to identify the protein of interest involves immunological methods of affinity purification, enzyme linked immunosorbent assays (ELISA), and western blot immunoidentification which require substantial sample processing after the events of interest have occurred (i.e not in real time) and can take several hours to perform.

Real-time analysis is possible by using fluorescent proteins (FP). These proteins are most often used as an indirect reporter for prior events of interest (one or more previous transcription/translation events) and those events must be inferred from the FP fluorescent signal. Frequently, the FP is present as a fusion to a protein domain of interest. However, combining large protein sequences into one larger polypeptide, have the potential to interfere with protein function and characterization. Additional concerns can arise from misfolded, aggregated, or overexpression of fluorescent fusion proteins^{17, 18}, or steric hindrance (in an FP-fusion) which could inhibit the natural activity of either the protein of interest¹⁹, or loss of function by the fluorescent fusion.^{20, 21} Quantification and interpretation of the fluorescent signals can be further challenged by factors such as multimerization of FP domains, variable maturation of the fluorescent signal under different conditions, and delays between the actual events of interest and the production of the FP. In each case, FP production often requires substantial resources that may perturb the system. Alternatives to large fluorescent proteins such as fluorescent labeling kits enable tagging a protein of interest for quantitation and tracking (e.g., Thermo Fisher AlexaFluor dyes). Unfortunately, specific high-affinity quantifiable labeling is often only possible in purified protein solutions, where the fluorescently-tagged protein is later added to a complex mixture for tracking and analysis.

With these issues in mind, we sought to develop a method for directly monitoring transcription and translation as these events are occurring in an integrated cell-free system. For example, the tetracysteine (TC) amino acid tag (Cys-Cys-Pro-Gly-Cys-Cys)^{22, 23} can specifically bind the organoarsenic molecules FIAsh and ReAsH (Thermo Fisher, also sold

as the “Lumio” detection reagents) to change their fluorescent state even in complex biological solutions and living cells. Similarly, RNA aptamers such as the Malachite Green aptamer^{5, 24}, “Spinach”²⁵, “Spinach2”²⁶, “Broccoli”²⁷ and “Mango”²⁸ can specifically bind small molecules (Malachite Green, DFHBI, DHFBI-IT, and TO-1 Biotin, respectively) that change their fluorescent state; this fluorescence can be used to directly monitor transcription. Alternatively, molecular beacons can provide direct measurement of RNA species through the binding of synthetic complementary oligonucleotides with fluorescent tags and quenching molecules (e.g. molecular beacons).^{29, 30}

We integrate such commercially available reagents into the approach PERSIA (PURExpress-ReAsH-Spinach In-vitro Analysis). The core system combines RNA labeling and protein labeling into a single “one-pot” reaction to measure gene expression characteristics and protein quantitation (Figure 1), and the potential to interrogate the function of the produced protein by way of an additional enzymatic assay that produces a fluorescent by-product. PURExpress (New England Biolabs) cell-free expression provides a context where the functions encoded in DNA (whether synthetic or natural in origin, linear or circular in form) can be probed with minimal concern for contaminating nucleases or proteases acting upon the engineered nucleic acid or expressed protein. The Spinach RNA aptamer tag with the DFHBI small molecule label²⁵ provides an assessment of transcription rate and the amount of RNA produced. A small C-terminal tetracysteine peptide tag and the biarsenical ReAsH reagent fluorescently label the protein(s) of interest.^{22, 23} This reporter system allows direct interrogation of transcription, translation, and potentially enzymatic biochemical reactions in real-time, instead of requiring additional laboratory procedures downstream.

We developed PERSIA to expedite how we design, query, and explore functions encoded in DNA. In some cases these are designed DNA sequences, such as for exploring the effects of engineering genetic codes. In other cases the sequences are clinical in origin, such as from genetic variation comprising resistance to antiviral drugs. And in yet others we explore more fundamental research questions, such as employing mutation scanning to determine which amino acids in a protein are essential for proper folding and function. We implemented PERSIA at the scale of 96- and 384-well microtiter plates (microliters), and then demonstrated the potential for miniaturizing the approach as a high-throughput/low-cost technology at the microfluidic scale (nanoliters). We evaluated a range of conditions under which PERSIA is most effective, and optimized performance under these conditions.

Results and Discussion

Characterization and Optimization of PERSIA

The development of PERSIA required an optimization of reagent concentrations for best performance. Unlike whole cells which can continually synthesize more nucleotide triphosphates and aminoacyl-tRNAs, cell-free synthesis reactions have a finite amount of building blocks available to transcribe DNA into RNA and translate RNA into protein. Therefore, the amount of material obtained during *in vitro* Cell Free Synthesis (CFS) is restricted and dependent on the size of the genes analyzed. Within the PURExpress system, there is 0.3 mM of each amino acid available for protein synthesis. For a hypothetical

protein of 11 kD containing 100 amino acids that is assumed to have an consensus amino acid composition (for example, <http://web.expasy.org/protscale/pscale/A.A.Swiss-Prot.html>), a final concentration of 25 μM might be reached before amino acid depletion prevents further translation. Similarly, with 1-2 mM of each NTP available for RNA synthesis, the DNA encoding that same 11 kD protein would be able to generate a final concentration of $\sim 50 \mu\text{M}$ of RNA until NTP depletion prevents complete transcription of the template. Increases to the size of the template DNA in the PURExpress reaction will proportionally decrease the amounts of RNA and protein generated, in theory. However, other biological mechanisms will also influence the rate and amount of RNA and protein produced³¹, including the size of the protein being produced (see Supplemental Figure 3c), and other factors being detailed in the remainder of the manuscript.

To optimize detection and quantitation parameters for RNA's containing both the 3' coding region for the 6 amino acid tetracysteine peptide and 3' Spinach aptamer, the concentration of DNA was maintained at a constant level (34 nM) while we varied the amount of DFHBI added to the PURExpress solution. We determined that 50 μM DFHBI ensured that this reagent is not limiting (Figure 2a). To optimize detection parameters for tetracysteine-tagged proteins of interest, we varied the amount of ReAsH that was added to the PURExpress solution to determine the optimal concentration for protein quantitation. We determined that 5 μM ReAsH was effective for protein detection and quantitation, as this concentration provides strong fluorescent signals while being less likely to act as a limiting reagent. Adding more than 5 μM ReAsH substantially increased fluorescence background signal (Figure 2b).

Dilutions of purified Spinach-tagged ubiquitin mRNA and purified C-terminal TC-tagged (CCPGCC) ubiquitin protein were analyzed separately in the PERSIA master mix to establish standard curves and determine the limits of detection within our assay. The fluorescence of the Spinach-tagged mRNA showed a linear increase for mRNA concentrations up to 1.5 μM (Figure 3a), and a linear detection range of tagged ubiquitin protein up to 4.5 μM (Figure 3b). Also of note is while the representative curves (Figure 3a-b) extend only through $\sim 1.5 \mu\text{M}$ for transcription and $\sim 4.5 \mu\text{M}$ for translation, the PURExpress product has sufficient NTPs and amino acids to potentially produce $\sim 10 \mu\text{M}$ of a 1.5 kb transcript and $\sim 6 \mu\text{M}$ of a 55 kD translation product, respectively.

To determine the response time of the processes producing fluorescent signals for PERSIA, we again employed pre-purified samples of tagged mRNA and protein. One μg of purified Spinach-tagged ubiquitin RNA (414 nM final concentration) and 1 μg of purified C-terminal TC-tagged ubiquitin protein (6.5 μM final concentration) were spiked into separate PERSIA reactions and the fluorescence signal of each were measured as a function of time.

The RNA kinetics of DFHBI binding appear delayed. A large proportion of RNA binding to DFHBI appears to occur within the first 10 minutes, but there is an additional 50% increase in signal over the next 40 minutes, and a maximal value is not reached after 50 minutes total (Figure 4a). This observation could potentially be attributed to the time needed for the Spinach motif to re-arrange structurally while binding to DFHBI. This has been observed elsewhere.^{25, 32} Improved aptamer design and fluorophore development such as Broccoli²⁷

and Mango²⁸ may help produce faster response time in future PERSIA development. Because of these delays, the Spinach/DFHBI fluorescence is interpreted as a semi-quantitative measure of mRNA production. This information is useful for comparing relative expression levels between genes and gene variants and for troubleshooting, i.e. discriminating between problems with transcription, translation, or other processes in the cell-free reactions. More thoroughly characterized binding kinetics will be needed to accurately quantitate the levels of mRNA being produced in real time during PERSIA.

ReAsH-derived signal maintained a relatively stable level of fluorescence for 50 minutes (Figure 4b), indicating the binding of ReAsH to the CCPGCC tag was stable. We observed the reaction reach maximum fluorescence within 10 minutes (Figure 4b) when 6.5 μ M of protein was added to a PERSIA reaction (i.e., 5 μ M ReAsH should be limiting factor). The slight decrease in fluorescence intensity over time could be attributed to several factors. First, ReAsH binding occurs when the 4 cysteine residues of the amino acid binding sequence (CCPGCC) undergo a change in oxidation/reduction. Thus, gradual oxidation occurring during the PERSIA reaction could potentially affect the stability of the protein or the interaction of ReAsH with the TC tag. In standard PERSIA reactions which utilize dNTPs and amino acids as building blocks, it is plausible that depletion of amino acids and NTPs, as well as buffer degradation (including oxidation) could slightly influence the stability of the ReAsH-protein interactions. (But we note here that in the Figure 4 example protein was synthesized and purified beforehand, and protein synthesis is not occurring during the reaction shown.) Additionally, while care was taken in purifying the recombinant protein from bacteria, there is a possibility that a small amount of protease could have been co-purified, leading to slight degradation of the protein standard and thus reduction of the fluorescence signal.

As fluorescent proteins are commonly used for detecting and measuring real-time cell-free protein synthesis, we compared the performance of eGFP and ReAsH. To avoid any fluorescence spectral overlap between eGFP and Spinach/DFHBI, the linear DNA encoding eGFP did not include a Spinach-tRNA aptamer tag in its design. Since the DNA design incorporates both the eGFP and TC tag coding sequences, the onset time of ReAsH and GFP fluorescence can be compared for the same protein molecule. Three different concentrations of linear DNA (1 nM, 5 nM, and 25 nM) were examined in PERSIA reactions to compare eGFP fluorescence onset to ReAsH fluorescence onset. We defined onset threshold as 3 times the standard deviation of the background signal. TC/ReAsH fluorescence provided a slightly more rapid and initial fluorescence onset, reaching threshold 7 minutes earlier than for eGFP fluorescence (5 nM DNA template, Figure 5). Having observed the binding kinetics of TC-tagged proteins with ReAsH (Figure 4b), we speculate that this process is faster under these conditions than the processes of eGFP protein folding and oxidation to produce its fluorophore.³²

We conducted additional experiments to examine questions of general reproducibility and spectral overlap. In terms of possible spectral overlap or other cross-talk between components, Supplemental Figure 3a indicates virtually no difference in the observation of Spinach/DHFBI signal whether or not ReAsH is present. Supplemental Figure 3b demonstrates that if illuminated for excitation, the Spinach/DHFBI complex only contributes

a small fluorescent contribution (at the 615 nm wavelength used for TC/ReAsH detection). However, we note that typical PERSIA experiments do not employ simultaneous excitation/emission of both tag/label complexes, avoiding this concern.

In terms of reproducibility, while many of our pilot experiments employed only single samples for each data point, we have also found PERSIA to be highly consistent. Supplemental Figure 3a–d shows examples of PERSIA experiments measured as technical replicates (3 samples, error bars indicate standard deviation), with variation among replicates being modest compared to differences between distinct samples. When the same experiments are repeated on different days, the same trends are reproduced, though not the exact same values. Thus for purposes of calibration (see Figure 3) we recommend that the externally-synthesized calibrants be included in each 384-well plate (or other format) for greatest consistency.

We also examined other commercially available *E. coli* lysate-based transcription / translation kits besides PURExpress while developing PERSIA and generally experienced challenges integrating fluorescent quantitation of RNA and/or protein. Background levels of fluorescence were high, likely due to interactions of DFHBI and ReAsH reagents with cell components present in these less-purified extracts. Nevertheless, other CFS reaction systems have been employed to great utility^{8, 33, 34}, and may eventually be adaptable to the integration of one-pot multi-reaction assays in a similar vein as PERSIA.

We performed an initial investigation the effects of salts and other additives on PERSIA. In some cases, the addition of supplements in the form of salt additives or small biomolecules³⁵, passivants like PEG, BSA, trehalose³⁶, or chaperone proteins³⁷ have been shown to enhance protein translation levels and increase proper protein folding efficiencies. In other cases, specific metal ions or other cofactors are needed for enzymatic activity. We added a number of these materials to the PERSIA assay to determine how they affected fluorescent measurements of both RNA production and protein production, as well as the potential to use ReAsH and DFHBI for quantitative analysis (Table 1). In many cases, an upper threshold concentration was observed. These results are consistent with cautions provided in the PURExpress manual, which lists >50 mM NaCl, >1 mM EDTA, and >2 mM magnesium and potassium salts as potential inhibitors of transcription and translation.

PERSIA for Gene Expression Analysis – “Test-driving” Genetic Code Designs

We explored PERSIA’s direct transcription/translation readouts for multiple ongoing efforts, one of which is genetic code engineering.^{38, 39} Engineered genetic codes—as opposed to those found in nature—are expected to impart unusual properties to an organism, such as immunity to viruses⁴⁰. However, exploration of large libraries of engineered codes *in vivo* require extraordinary amounts of time and resources. PERSIA accelerated evaluation of genetic code designs, and could act as a screening tool for obvious failure of specific designs.

Previously Lajoie et al.⁴¹ tested one genetic code design (13 codons forbidden, i.e. not allowed for use in translation) by synthesizing variant versions of these genes. In that work, individual redesigned genes were tested by insertion into the *E. coli* genome, replacing the

wild-type copy. In addition, replacement of each instance of an allowed codon in a gene (with a synonymous equivalent) was employed to further gauge how amenable these genes are to being re-written.

To build on these results and others^{4, 42}, an initial PERSIA screen was applied to a panel of essential genes from *E. coli* to predict how genetic code designs would affect transcription and translation *in vitro*. We designed synthetic DNA encoding 14 of these essential genes (Supplemental Table 1) including T7 promoter, native RBS as found in the *E. coli* genome, native coding sequence, Spinach tag, TC tag, and T7 terminator (see Supplemental Figure 1). Coding sequences ranged from 192 bp (7 kD protein product) to 822 bp (30 kD protein product). We expressed these genes within our PERSIA reagent mixture, monitoring transcription and translation by green DFHBI fluorescence (500 nm) and red ReAsH fluorescence (615 nm), respectively. Figure 6 (with more detail given in Supplemental Table 1) shows a range of transcription and translation rates among these genes. Many of the transcripts are similar in size (see Supplemental Table 1) but demonstrate notably different transcription and translation rates. Each DNA construct yielded strong fluorescence distinct from control reactions lacking a DNA template, serving as a sufficient reference for following transcription and translation experiments with modified versions of these genes.

We expected that one factor affecting both production rate and final yields in PERSIA reactions would be resource limitations. In particular, the concentrations of amino acids determine an upper limit to how much protein can be synthesized by the PURExpress system. While we note that some of the smaller proteins were produced in the highest quantities (e.g. rplX, rpsS, rpmC) we did not observe strong correlations between fluorescence signals from mRNA product, protein product, and length of gene. These observations are consistent with the expectation that in addition to gene size, regulatory elements such as promoter and RBS (and their interactions with the gene sequence itself) would also contribute to variation. We explored RBS interactions with the coding sequence further in considering the impact that a novel genetic code would have on protein expression.

To demonstrate the potential of PERSIA for initial testing of genetic code designs, we chose rplX from the set of genes described above. In “recoding” this gene, the coding sequence was modified to reflect a reduced genetic code (Supplemental Table 2), i.e., with fewer than the natural number of 64 codons available for use. We tested a slightly modified code (forbidding use of the rare arginine codons AGA and AGG) and an extreme code (forbidding use of 43 codons, resulting in an absolutely minimal genetic code allowing one codon per amino acid plus one stop codon). For the simple code change, only one such rare codon was present (AGA for Arg86) out of the 104 amino acids encoded. We compared substitution with each of the four allowed Arg codons at this position (CGA, CGC, CGG, and CGT) to the wild-type sequence, as well as the extreme recoded version (roughly half of the original codons being replaced, altering 20% of the nucleic acid sequence). Translation results are shown in Figure 7a. As might be expected, the single codon substitutions (DNA designs R86a-d) resulted in effectively the same amounts of protein produced, and at the same rate. On the other hand, the extremely recoded version (“Min”) produced substantially less protein.

Multiple factors could result in the decreased translation observed for this extreme rplX variant. First, only a subset of the tRNA molecules present are being used for expression of the “Min” rplX due to its one codon – one amino acid design. In effect, the ribosome would experience lower concentrations of available tRNA. Second, the sequence restrictions of an extreme genetic code could result in a tendency toward mRNA secondary structures in the coding region that impede translation. Third, the revised coding sequences could give rise to new mRNA interactions with the RBS that inhibit translation initiation. We expect that other failure modes are possible as well. These observations are consistent with the hypothesis that reduced translation observed for rplX-Min was the result of interactions between the RBS and the modified coding sequence—interactions that were then avoided by implementing the new designed RBS sequences.

Since intramolecular mRNA interactions between the 5' untranslated region (including RBS), and the beginning of the recoded rplX reading frame could be disrupting translation initiation^{43–45}, we employed the RBS Calculator^{43, 44} to develop hypotheses of which genetic changes might be responsible for reduced translation. Indeed, a synthetic DNA construct restoring the first 36 bp of wild-type rplX also restored wild-type levels of cell-free translation (“Hybrid” in Figure 7b). Within that 36 bp region, five codons had been changed for the rplX-Min design. We used the RBS calculator to predict translation initiation rates for sequences with different combinations of those five codon changes, (Supplemental Table 2). Reversion of the first three modification (Ala2, Ile5, and Glu9) to wild type sequence (Min-rev123) gave the same predicted translation rate as the wild type sequence. Synthesizing this DNA and testing with PERSIA confirmed that restoring only those three modifications was sufficient to restore robust translation (Min-rev123 in Figure 7b).

The remaining question was whether this extreme minimum genetic code design could still be practicable in spite of the reduced translation levels observed. If a given genetic code has severe effects on one essential gene, that code almost certainly could not be implemented throughout the entire genome of an organism. We used the RBS Calculator⁴⁴ in design mode to devise replacement RBS sequences predicted to restore high translation levels (Supplemental Table 2). We synthesized and tested one of these designs (Min-RBS-6k) generated by setting a target translation initiation rate to 6000 (arbitrary units), approximately the rate predicted for the original *E. coli* RBS sequence. PERSIA experiments with Min-RBS-6k indicated that robust translation was restored (Figure 7b). As translation levels for Min-RBS-6k and Min-rev123 were in fact higher than for the wild type sequences, further tuning of the RBS may be needed for an optimal *in vivo* implementation of this recoded gene.

No one has yet determined the limit of extreme genetic code engineering. Thus we were encouraged that the translation of this gene, with this extreme genetic code, could be rescued through a combination of simple computational and *in vitro* experiments with PERSIA. These results do not show that this minimal code will prove viable for an entire *E. coli* genome, or that this rescue approach will be sufficient in all cases. Rather, this example serves as a starting point for developing a larger testing pipeline of genetic code design and testing, from *in silico* to *in vitro* to *in vivo*.

PERSIA with Additional Enzymatic Assay

The “one pot” nature of PERSIA can be extended to monitor even more simultaneous biochemical processes, in addition to transcription and translation, using the same complex reaction mixture. We explored applications of this kind using the enzyme HIV-1 protease (HIVpr hereafter). This 99 amino acid protein is normally produced as a cleavage product of the larger *pol*-encoded polyprotein. Homodimer formation is required for enzyme activity. Protease activity can be measured *in vitro* using a variety of peptide cleavage assays. One commercially available kit is the SensoLyte 520 HIV Protease Assay Kit (AnaSpec). This reagent is a FRET-quenched peptide in its intact form. Cleavage of the peptide by HIVpr separates the QXL520 quencher from the HiLyte 488 fluorophore, allowing monitoring of enzyme activity over time (excitation at 488nm / emission at 528nm). Due to spectral overlap with Spinach-DFHBI fluorescence, we employed two parallel PERSIA reactions – one to measure protease activity via SensoLyte, the other to quantify RNA via the PERSIA spinach-DFHBI interaction. Both reactions allowed observation of the amount of protein expressed (with ReAsH) to compare and normalize results between the 2 data sets.

To demonstrate PERSIA’s ability to interrogate protein function, we designed synthetic linear versions of the HIVpr gene (Supplemental Figure 5, in the SBOL format of Supplemental Figure 1). The sequences express an encoded HIVpr sequence (CA126802) from the HIV Drug Resistance Database at Stanford University (<https://hivdb.stanford.edu/>)^{46, 47}, with C-terminal TC tag and Spinach tag to allow RNA and protein quantitation with DFHBI and ReAsH-EDT₂, respectively.

Protein production measured by TC/ReAsH fluorescence indicated additives to the PERSIA reaction (DMSO, SensoLyte) had minimal, if any, effect on transcription and RNA synthesis, as well as the translation and protein folding process (Figure 8 and Supplemental Figure 6). Protein production peaks approximately 1.5 to 2 hours after initiating the assay (Figure 8a). Fluorescence onset from protease activity lagged behind transcription and translation, appearing approximately 1.5 hours after initiation of the PERSIA reaction, possibly indicative of slow protein folding or rearrangement before achieving full activity, or consequences of our instruments’ limit of detection once unquenched peptide was produced (Figure 8b). Protease activity was observed in the unmodified HIVpr version (CA126802), as expected. Controls with no DNA, DNA expressing a negative control ubiquitin protein standard, and an inactive variant of HIVpr (with mutation D25N) also behaved as expected, with no SensoLyte fluorescence above background (Figure 8b).

PERSIA with Enzymatic Assay for Probing Sequence/Function Relationships

As an example of investigating relationships between amino acid sequence and protein function, we adapted the concept of scanning mutagenesis to PERSIA. We performed the equivalent of alanine-scanning mutagenesis⁴⁸ with DNA samples encoding variant genes produced by commercial *de novo* synthesis. For each amino acid position in HIVpr (CA126802), a synthetic DNA construct was designed with that position changed to alanine. Ninety-eight synthetic constructs were tested in PERSIA reactions employing ReAsH and SensoLyte to simultaneously monitor protein production and enzymatic activity. To provide an estimate of each enzyme’s activity, SensoLyte fluorescence values (representing the

amount of cleavage product) were scaled by their corresponding ReAsH fluorescence (representing the amount of HIVpr present). The resulting profile of protease activities is shown in Figure 9, shown relative to the activity estimated for the unmodified version of HIVpr (CA126802). This graph indicates which amino acid positions are intolerant of alanine substitution, as observed by disruption of protease activity. Some of these interfere with catalytic activity, such as the conserved active site amino acids Asp25-Thr26-Gly27, while others are important for proper structure, such as amino acids Gly49-Ile50-Gly51-Gly52 in the flap region required for dimerization.

A conventional alanine-scanning mutagenesis experiment to obtain these observations would have once been very effort-intensive. After a researcher constructed a base plasmid with the wild-type coding sequence, 97 site-directed mutagenesis reactions would be required (either singly or in multiplexed batches) with 97 synthetic oligos. This step would be followed by cloning, transformation, solid agar culture, colony picking, liquid culture, plasmid prepping, sequence verification, a second transformation, second liquid culture, and protein purification. This process would provide the proteins needed for an enzymatic assay. Instead, the early cumbersome steps are consolidated by the operations of the commercial DNA synthesis provider, and the later ones obviated by PERSIA. Starting with the synthetic DNA, all variants and controls were analyzed by PERSIA, requiring roughly two hours to prepare, and 3-4 hours to run.

Employing PERSIA also provided information for potential troubleshooting and calibration. Spinach/DFHBI and TC/ReAsH fluorescence profiles confirmed that both transcription and translation were occurring. ReAsH fluorescence was generally within a factor of two compared to the unmodified control. In a few cases (P1A, G86A, G94A) the fluorescence intensity was much lower (14-16% of wild-type) but still clearly discernable. An ubiquitin-expressing DNA construct was used as a negative control for enzymatic activity, yielding an apparent activity 1.6% of that for reference HIV protease construct—a lower bound estimate of the noise present in this application of PERSIA. A complete lack of fluorescent signal for N98A (far below background fluorescence seen in the other PERSIA reactions) likely indicated a flawed fluid transfer. While we have not identified a comparable alanine-scan study, these results are consistent with previous analyses of HIV-1 protease genetic variability. For example, substitution-intolerant regions in Figure 9 (residues 8-11, 23-27, 49-52, and 81-89) are also low-variability regions seen by Vergne et al. in an alignment of consensus sequences across HIV-1 groups.⁴⁹

PERSIA with Enzymatic Assay to Evaluate Drug Resistance with HIV Protease

In pairing the PERSIA system with the HIVpr biochemical assay, we can foresee potential applications for personalized drug screening. For many viral infections, an individual patient harbors a large number of viral genetic variants, such with HIV, HSV, Ebola, influenza, and HCV. The majority of these variants go undetected by conventional Sanger sequencing that detects only the most abundant 1-3 genotypes. In an envisioned future application, such genetic diversity would be assessed from a patient sample at the site of care by high-throughput sequencing. This data would be used to synthesize genetic material encoding these drug targets, followed by cell-free expression and assay with candidate drugs. The

resulting drug resistance profiles would be used to craft a drug regimen for that individual, customized against their personal, unique constellation of viral variants. We explored this concept with our HIVpr constructs in PERSIA/SensoLyte assays.

Currently, two of the standards in testing for drug-resistant HIV in a clinical setting are the phenotypic “Phenosense” cell culture assay⁵¹ and genotypic “Genosense” analysis^{46, 52}. The Phenosense and Genosense assays (Monogram Bioscience) use clinical HIV blood samples to recover and sequence the viral protease and transcriptase genes encoding the primary targets for anti-HIV therapeutics. By sequencing, cloning, and expressing these genes in a luciferase reporter assay against a library of drugs, they are able to devise a customized pharmaceutical treatment regimen best suited to combat an individuals’ HIV strain combination. The amount of time and effort required to obtain this information, along with certain requirements associated with sample collection (i.e., patient titers must be more than 500 copies per mL of blood), beckons for the development of new techniques that are more sensitive and rapid.

We employed HIVpr gene sequences determined from clinical samples in conjunction with the PERSIA + SensoLyte assay to expedite the manner in which treatment options could be tested for a specific individual’s HIV infection. The clinical HIVpr sequences (HIV Drug Resistance Database) correspond to resistance against specific HIVpr inhibitors. Synthetic linear DNAs (with the design format in Supplemental Figure 1) encoded with these HIVpr sequences were added to the PERSIA + SensoLyte assay with varying doses of the antiviral drugs Atazanavir, Darunavir, and Lopinavir to observe changes in protease activity. These drugs are used clinically in the treatment of HIV-1 infection. The drug concentration range of 0.5 μ M to 10 μ M used in this study is clinically relevant. Drug concentrations for the inhibitors chosen have been measured circulating in the blood stream at 7.6 μ M, 9 μ M, and 9.4 μ M, and our lower experimental concentrations are near the minimally effective doses of 0.38 μ M, 0.27 μ M, and 1.6 μ M for Atazanavir⁵³, Darunavir^{54, 55}, and Lopinavir^{56, 57}, respectively.

As for the DMSO and SensoLyte additives above, TC-tag/ReAsH fluorescence (Supplemental Figure 6) indicated that the antiviral drugs added to the PERSIA reaction had minimal, if any, effect on the fluorescent signal. Protein production peaked approximately 2 hours after initiating the assay. Protease activity was observed for 2 hours after peak protein production by observing increases in green fluorescence due to the cleavage and unquenching of the SensoLyte peptide substrate (Supplemental Figure 6). Protease activity was observed in the absence of drug in the wild-type HIVpr version (N1802.1; HIVpr-var1-WT), clinical sequence CA126802 (HIVpr-var2), and clinical sequence CA50384 (HIVpr-var3) as expected. Controls with no DNA, DNA expressing a negative control ubiquitin protein standard, and an inactive variant of HIVpr (with mutation D25N) also behaved as expected, with no SensoLyte fluorescence above background.

In its current form, this assay was not amenable to precise IC₅₀-like standard kinetic measurements of enzyme activity. Although the protein levels appears to have stabilized after 2 hours of PERSIA, the potential of continued protein production due to the presence of PURExpress reagents was still possible. To compare protease activity among the variants,

the average concentration of protease was extrapolated from the TC-tag/ReAsH standard curve for each time point (Figure 3), then averaged for time points between 2.5 and 4.5 hours. The total amount of SensoLyte cleaved during the same time period was also extrapolated from a standard curve generated the HiLyte488 fluorescence reference standard supplied in the Sensolyte kit (Supplemental Figure 7). An enzyme activity estimate was calculated by dividing apparent nM of cleaved product by the apparent μ M protein present. “Percent activity” shown in Figure 10 represents these enzyme activity values scaled according to the “no drug” value. We employed 40% activity as a scoring threshold for the drug resistance observed with each HIVpr variant (black horizontal line, Figure 10). For each drug concentration allowing greater than 40% activity, a “+” rating was added to a HIVpr variants score. For example, a single “+” score would be given for HIVpr variants demonstrating >40% activity for only the 0.5 μ M concentration. Similarly, a “+++” score would be given for HIVpr variants demonstrating >40% activity for the 0.5 μ M, 1 μ M, and 2.5 μ M concentrations.

The expressed wild-type HIVpr sequence (N1802.1) showed little to no protease activity in the presence of 3 drugs known to inhibit HIVpr – Darunavir, Lopinavir, and Atazanavir (Figure 10). Clinical variant CA126802 (HIVpr-var2) is known to be highly resistant to all three of these antiviral drugs, according to the HIV Drug Resistance Database (database scores of 120, 175, and 200). Clinical variant CA50384 (HIVpr-var3) is known to have strong resistance to Lopinavir (database score of 125), moderate resistance to Atazanavir (score of 85), and meager resistance to Darunavir (score of 20) according to the HIV Drug Resistance Database. In our PERSIA/SensoLyte assay, the two clinical variants of drug resistant HIVpr yielded drug resistance profiles qualitatively comparable to those established through the Geno/Phenosense assays (Figure 10). We anticipate that these approaches may eventually be able to replace more cumbersome, costly, and time-intensive procedures that are currently considered the gold standard. However, extensive further development, optimization, and validation will be required to establish utility in a clinical setting. If an assay platform such as PERSIA were combined with on-site sequencing at clinics and rapid on-demand DNA synthesis of genetic variants, decisions for custom patient drug regimens might be accelerated substantially.

Not all assays will be adaptable to PERSIA in this manner. For example, another effort in our group has been examining biochemical activity of the organophosphate hydrolase (OPH) enzyme⁵⁸ and its variants. The activity of OPH mutants was analyzed by its ability to break down Demeton-S as a substrate, resulting in a 2-ethylthioethane thiol by-product. The PURExpress system is rich in thiol-containing reagents and buffers, so our difficulty in observing OPH enzyme activity (data not shown) may be the result of assay reagent reacting with both the thiol by-product and the other thiols in the PURExpress mixture.

Integration of PERSIA into a Microfluidic / Multi-Gene Assay

The utility of the PERSIA system can be expanded through integration into a microfluidic assay, ultimately scaling up to large numbers of samples and incorporating other features such as protein purification. As these reagents can be costly, microfluidics provides an attractive avenue to reduce reagent volumes, provide robust and repeatable data through

automation, and perform hundreds of assays simultaneously in a high-throughput manner.^{34, 59, 60} For example, the mechanically induced trapping of molecular interactions (MITOMI) device developed by Gerber *et al.* utilizes arrayed protein samples with microfluidic architecture aligned above to study protein-binding interactions.⁶¹ Our device (Supplemental Figure 7) was designed with 6 banks containing 2×8 sets of assay channels, and rows of valves used to automate and control fluid flow through the device (Figure 11a). In principle, each of the 96 resulting channels could be used as a reactor to express and test a different DNA design. Our proof-of-concept aimed to determine if PERSIA would be compatible with nanoliter volumes (reduced signal) and with the PDMS device material which has shown biofouling challenges historically.⁶² Our early investigations found a gradient of signal weakening down the channel length which we hypothesize is due to nonspecific binding of PERSIA reagents to the walls of the polydimethylsiloxane (PDMS) channels. We were able to mitigate biofouling effects by first passing excess PURExpress reagents through the microfluidic channels before a microfluidic experiment.

Repeated 6-plex experiments were completed to compare on-chip performance against a traditional well-plate. Five constructs (HIVpr-var1WT, HIVpr-var2, HIVpr-var3, Native ubiquitin, and Ubiquitin standard) and a no DNA control were tested in each bank. The results of the Spinach/DFHBI and TC-tag/ReAsH signals are shown in Figures 11b and 11c, respectively. Spinach/DFHBI fluorescence (Figure 11b) indicated linear mRNA production in the process, then slowing down, comparable to what we have observed in well plate reactions. TC-tag/ReAsH fluorescence (Figure 11c) peaked slightly more quickly, again consistent with previous observations in standard well plates. We did however observe higher relative levels of background fluorescence in the microfluidic devices. Future work will seek to employ arrayed DNA samples to take advantage of the device's 96 assay channels as individual reactors rather than large banks, and to incorporate additional enzymatic assays such as described above for HIVpr.

Conclusions

We developed PERSIA to accelerate how we ask and answer questions with DNA. Some of these are basic research questions, such as “how do specific DNA sequences affect transcription and translation?” or “how do changes in the amino acid sequence affect the function of a protein?” Others have a more applied intent, e.g. “will this genetic code work?” or “what drug regimen should we try first?” The pilot experiments described here demonstrate that PERSIA can be applied to a broad range of these questions. While performing these experiments we found that an already complex cocktail of cell-free expression reagents was compatible not only with the fluorescent labels DFHBI and ReAsH, but also with a range of enzymatic assay components, inhibitors, and additives.

The examples presented above each demonstrate key capabilities in quantifying transcription and translation. Initial testing and optimization of PERSIA included examples of how the Spinach-tag/DFHBI and TC-tag/ReAsH signals can be calibrated in terms of real-world concentrations and not simply arbitrary fluorescence units. However, the ideal standards and standardization approaches are yet to be determined. We now have a way to quickly test new genetic code designs in specific genes, revealing some—but certainly not all—potential

flaws, with an opportunity to troubleshoot. For maximum effect, though, that approach needs to be integrated into a larger pipeline, coupling to upstream computational design and downstream *in vivo* testing, enabling more effective design-build-test cycles.^{63, 64} We have the seeds of an approach for testing drugs and drug candidates against the protein targets of specific pathogens. But generalizability to many pathogens, targets, and drugs will need to be explored.

A major advantage of PERSIA is that GFP and similar fluorescent reporters are not required, thus obviating the need for including these larger DNA sequences into synthetic constructs or adding via cloning. The chemical resources of a cell-free reaction do not have to be allotted to producing a GFP reporter when the real question is centered on a specific protein of interest. In addition, employing tags at the 3'-terminus of RNA molecules and the C-terminus of proteins provides a general strategy for confirming that the complete biomolecule has been synthesized before a fluorescent signal is generated. These features allowed expedient assessment of multiple genes and gene variants. Cell-free systems provide a simplified environment in which protein functions can be observed, absent complications from other cellular interactions. This can include multimeric protein complexes that might otherwise be interfered with by fusion to bulkier fluorescent protein (which themselves have some tendency to multimerize).

Potential applications of PERSIA include testing genetic code designs, optimizing codon usage for protein expression, and rapidly screening drug candidates against their protein targets. PERSIA is compatible with automated miniaturization within microfluidic systems, facilitating future implementation in high throughput formats while consuming only trace amounts of valuable cell-free reagents. We are especially eager to apply these tools to challenges in personalized medicine (e.g. HIV therapy) and toward developing countermeasures against new and emerging pathogens

Methods

Synthetic DNA Design

Synthetic DNAs were ordered and obtained from IDT and Gen9. All DNAs were designed with universal M13 forward (GTAAAACGACGGCCAGTG) and M13 reverse (CATGGTCATAGCTGTTTCC) primers at the N-terminus and C-terminus, respectively, to allow for PCR amplification. Unless otherwise noted, immediately following the M13 forward sequence was a spacer sequence (AGAGGTAGCACATCTCGATGCCGCGAAAT), a universal T7 promoter sequence (TAATACGACTCACTATAGG), a 45 bp intervening sequence (GAGACCACAACGGTTTCC CTCTAGAAATAATTTTGTTAACTTT) and ribosome binding site (wtRBS; sequence GAAGGAGATATACC) based on the pUC57 plasmid. Downstream of the gene of interest and upstream of the M13 reverse primer, the DNA sequence encoding the ReAsH binding motif Cys-Cys-Pro-Gly-Cys-Cys (TGTTGCCCCGGGTTGCTGTA), purification tags for 6x His (CATCACCATC ACCATCAC), a stop codon (TAA), the Spinach tRNA scaffold sequence (GACGCGACCGAAATGGTGAAGGACGGGTCCAGTGCTTCGGCACTGTTGAGTAGA GTGTGAGCTCCGTAAC TGGTCGCGTCA), and T7 terminator

(AACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTT) were present. The generic schematic of the synthetic DNAs can be found in Supplemental Figure 1.

When designated, synthetic DNAs were amplified by PCR to obtain more gene to add into the assays. The M13 forward and reverse primers were used at 1 μ M concentrations with Phusion master mix (New England Biolabs, part # M0531L) to amplify the synthetic DNA for 30 PCR cycles (95 °C for 30 sec, 57 °C for 30 sec, 72 °C for 1 min/kb of DNA) and purified via PCR cleanup kit (Qiagen). Fragment concentrations were made by DNA absorbance at A_{260/280} using a NanoDrop spectrophotometer.

Protein Standard

A linear DNA (IDT gBlock) encoding previously described genetic circuitry (see Synthetic DNA Design), human Ubiquitin with a C-terminal ReAsH tag (CCPGCC) followed by a C-terminal 6xHis tag was PCR amplified with forward (gatcgatccatATGCAGATCTTCGTGAAGACCC) and reverse (gactgactgctagcAAGCTGACGCGACCAGTTACGG) primer sequences using the Phusion amplification mix (NEB). The PCR was cycled 98° C/57° C/72° C at 45 seconds each temperature for 30 cycles. The PCR fragment was purified using a PCR cleanup kit (Qiagen). The ubiquitin PCR fragment and pET100 plasmid were both digested with BamHI / EcoRI enzymes for 1 hour at 37° C in CutSmart buffer. The pET100 DNA was also treated with calf intestinal alkaline phosphatase for 30 minutes at 37° C. The digests were purified by agarose gel electrophoresis followed by a gel extraction kit (Qiagen) and the eluted fragments ligated together at 16° C overnight. The ligation was transformed into chemically competent DH5 *E. coli* (NEB) using a standard heat shock protocol and plated on LB + ampicillin (100 μ g/mL) to grow at 37° C overnight. Colonies were inoculated into 5 mL of LB + ampicillin (100 μ g/mL) and grown overnight at 37° C before processing through Spin 250 miniprep columns (Qiagen) to recover plasmid DNA. DNAs were digested with BamHI / EcoRI for 1 hour at 37° C, then analyzed by electrophoresis through a 1% agarose TBE gel. A clone with the correct restriction digest pattern was sequenced by the Sanger method to confirm the sequence identity.

The pET100 + Ubiquitin plasmid was transformed into chemically competent Rosetta BL21(DE3) pLysS *E. coli* (Novagen / Millipore part# 70956-4) using a standard heat shock method and plated on LB + ampicillin (100 μ g/mL) + chloramphenicol (34 μ g/mL) to grow at 37° C overnight. A colony was inoculated into 50mL of LB + ampicillin (100 μ g/mL) + chloramphenicol (34 μ g/mL) and grown overnight at 37° C. The following day, the 50 mL culture was inoculated into 1 L of LB + ampicillin (100 μ g/mL) + chloramphenicol (34 μ g/mL) and grown at 37° C until the OD600 reached 0.6. IPTG was added (final concentration of 0.5 mM) to induce protein expression for 4 hours. The bacteria were harvested by centrifugation at 5000 rpm for 30 minutes and the media discarded.

To purify the recombinant ubiquitin protein product, the pellet of induced *E. coli* was resuspended in 10 mL of 1x PBS + 0.1% Triton X-100 + 20 mM imidazole, pH 7.5. The cells were lysed by sonication while incubating on ice. The lysate was clarified by centrifugation at 10,000 rpm for 20 minutes. The lysate flowed by gravity through 5 mL of Ni-NTA resin (Thermo Fisher part # 901-10) in column format to bind the ubiquitin protein

by its C-terminal 6xHis tag. Ten column volumes of 1x PBS + 40 mM imidazole, pH 7.5 was used to wash the resin. Five column volumes of 1x PBS + 300 mM imidazole, pH 7.5 was used to elute the protein from the column resin. 1 mM DTT was added to the protein solution, then the solution was concentrated using an Amicon-10 filtration unit at 2500rpm. The protein was dialyzed into 25 mM Hepes + 100 mM NaCl + 1 mM DTT, pH7.5 at 4 °C overnight. The protein concentration was determined to be 9 mg/mL by Bradford assay and by densitometric analysis after electroporation of the ubiquitin protein and albumin standards through a 10-20% SDS-PAGE gel, staining with Coomassie Brilliant Blue G-250. The protein was diluted with glycerol to a final concentration of 50 % and stored at -20° C at a final protein concentration at 4.5 mg/mL (439 μ M).

RNA Standard

PCR amplified ubiquitin linear DNA (see Protein Standard section) was transcribed using the Megascript T7 Transcription kit (Thermo Fisher). One microgram of DNA, 10 μ L of 10x reaction buffer, 10 μ L of enzyme mix, and 10 μ L each of ATP, CTP, GTP, and UTP (7.5 mM each final concentration) were mixed in a 100 μ L reaction and incubated for 3 hours at 37 °C. One microliter of TURBO DNase was added to degrade the template DNA for 15 minutes at 37 °C. The RNA was purified from the reaction using 2 MegaClear RNA purification columns. The 100 μ L reaction was mixed with 100 μ L of elution buffer, 700 μ L of Binding Solution, and 500 μ L of ethanol prior to adding the solutions to the 2 columns. RNA was bound by centrifugation at 10,000 rpm for 15 seconds, followed by two 500 μ L washes and a 30 second drying centrifugation, each for 30 seconds at 10,000 rpm. A 100 μ L aliquot of elution buffer was added to each tube, and then the tube was heated for 5 minutes at 65° C to elute the RNA. The RNA was recovered by centrifugation at 10,000 rpm for 1 minute. The RNA was pooled then quantitated to be 1191 ng/ μ L (7.5 μ M) by DNA absorbance at A₂₈₀ using a NanoDrop spectrophotometer.

PURExpress Master Mix Solution

Components of the PURExpress cell-free synthesis kit (New England Biolabs, part #E6800L) were thawed on ice. Fluorescent molecule stock tubes for ReAsH-EDT₂ labeling reagent (2 mM in DMSO; Thermo Fisher, part # T34562), and SensoLyte assay reagent (Components A and B in DMSO; AnaSpec, part #AS-71127) were thawed at room temperature. DFHBI powder (Lucerna, part #410-10 mg) was resuspended in DMSO to a final stock concentration of 10 mg/ml. Stocks of linear DNA both PCR-amplified and original gBlocks manufactured by IDT were thawed on ice.

DNA stock solutions were quantitated by DNA absorbance at A₂₈₀ using a NanoDrop spectrophotometer and diluted in water to obtain a 10x DNA solution (see manuscript for amounts used in each individual experiment). DNA stock solutions were maintained on ice. Master mixes of PERSIA reaction components were assembled on ice according to Supplemental Table 3.

PERSIA Plate Reader Assay

For a given PERSIA reaction, 1.5 μ L of a 10x Linear DNA stock was added to the bottom of each well of a clear-bottom, black sided 384-well plate, followed by the addition of 13.5 μ L

of PURExpress + fluorophore master mix solution (Supplemental Table 3). DNA and reaction solution were added to the plate as it rested in ice. A sheet of clear adhesive film was used to seal the wells and prevent evaporation. The plate was placed into an M5 plate reader (Molecular Devices) at 30° C and excitation/emission readings were made. The plate was mixed for 2 seconds by the plate reader briefly at the beginning of the analysis, and for 2 seconds prior to each subsequent reading taken throughout the analysis. Fluorescence excitation was made into the top of the plate, with fluorescence measurements made through the clear bottom of the plate at the minimum kinetic interval allowed by the M5 plate reader order of operations (i.e., number of samples, calibration between reads, number of wavelengths being analyzed dictated the interval in between well reads).

For DFHBI, excitation was made at 472 nm and emission read at 500 nm, with a 495 nm filter cutoff. For the ReAsH-EDT₂, excitation was made at 590 nm and emission read at 615 nm with a 610 nm filter cutoff. SensoLyte and eGFP was excited at 495 nm and emission read at 528 nm with a 515 nm cutoff.

Determination of Fluorescent Substrate Binding Kinetics

One µg of purified ubiquitin-ReAsH-Spinach tRNA messenger RNA transcript was added into a PURExpress reaction with a 15 second delay between addition of RNA and initiation of the fluorescence measurement process, a time limitation imposed by the M5 plate reader's pre-calibration process. The CFS reaction was read by exciting the DFHBI at 472 nm and reading the emission at 500 nm.

One µg of purified Ubiquitin-Reash-6xHis tagged protein was added into a PURExpress reaction with a 15 second delay between addition of protein and initiation of the fluorescence measurement process. The CFS reaction was read by exciting the ReAsH at 590 nm and reading the emission at 615 nm.

Analysis of Ribosomal Gene Sequences and rplX Variant Synthetic DNA

Linear synthetic DNA constructs for ribosomal gene sequences were obtained from a commercial vendors (IDT) and resuspended in water to 50 ng/µl. These DNAs encoded the same T7 promoter, TC tag, Spinach tag, and T7 terminator as previously described in Synthetic DNA Design section unless otherwise noted in the manuscript. For each construct, the ribosome binding site (RBS) used matched that of the wild-type gene in the reference strain of *E. coli* MG1655. Min-RBS-6K had a modified RBS sequence of: AACGAACAGAGGAGGCAGACAGACCA. Min-RBS-max had a modified RBS sequence of: ACGGCTCAAAATATTAGCAAGGAGGGGGTAGGG.

Linear DNAs were PCR amplified with M13 forward and M13 reverse primer sequences using the Phusion amplification mix (NEB) and purified using a PCR cleanup kit (Qiagen) as described (see Synthetic DNA Design). Final DNA concentration in the PERSIA reactions was 5 nM (32 ng).

PERSIA Analysis of HIVpr variants and Alanine Scan mutagenesis

Linear synthetic DNA constructs for wild type HIVpr version (N1802.1; HIVpr-var1-WT), clinical sequence CA126802 (HIVpr-var2), and clinical sequence CA50384 (HIVpr-var3) were obtained from commercial vendors IDT (clinical variants var1WT, var2, var3) and Gen9 (alanine scan HIVpr variants and Ubiquitin). These DNAs encoded the same T7 promoter, RBS, TC tag, Spinach tag, and T7 terminator as previously described in Synthetic DNA Design section. Clinical HIVpr sequences (HIV Drug Resistance), correspond to resistance against specific HIVpr inhibitors.

DNAs were resuspended on arrival to 5 ng/μl (12.5 nM). Two separate reactions were assembled as per protocol (see “PERSIA Plate Reader Assay”). One reaction contained 5 μM ReAsH and 50 μM DFHBI to quantitate protein concentration and RNA concentration, respectively. The second reaction contained 5 μM ReAsH and 1:100 dilution of SensoLyte, to quantitate protein concentration and HIVpr enzyme activity, respectively. Fluorescence measurements on M5 plate reader were performed for 4.5 hours as described (see “PERSIA Plate Reader Assay”).

Microfluidic PERSIA

The microfluidic device was developed using photolithographic and soft lithographic methods described previously and designed with the footprint of a standard 25 mm x 75 mm glass microscope slide.⁶⁵ The schematics of the two layer device is shown in Supplemental Figure 7. The device is composed of two layers where the top control layer acts as pneumatic valves capable of blocking flow in the flow layer underneath where CFS reactions take place. To run the device, first the upper control layer is plumbed and connected to a programmable pneumatic pump, similar to the one utilized by ⁶⁶. Once connected, the microfluidic control valve lines and water jacket were dead-end filled by forcibly diffusing air out of the PDMS and replacing it with distilled water at a pressure of 15 psi to prevent evaporation. After all valves had filled, the PERSIA master mix, with and without DNA added, was introduced into the six banks on the device and immediately placed into a Nikon Eclipse Ti confocal microscope for automated imaging of each of the banks. While the device is being set into the microscope, the programmable pump then manipulates the fluid into each of the banks to degas the master mix and remove any bubbles. The confocal microscope with programmable stage and 10x objective is then set to image full z-stacks of channels in each bank over the course of 3 hours in 10 minute intervals for the first half of the time course and 30 minute intervals for the second half. A 488 nm (eGFP) and 561 nm (TexasRed) laser and pairing filter set was used with our confocal camera to quantify both Spinach and ReAsH signal respectively for each image over the time course. Once imaging was complete, the NIS-Elements software created a maximum intensity projection for each z-stack to condense all the z-stacks from the bottom to the top of every channel into a single image. The software then sampled fluorescence from multiple points on each channel and exported time, location, and fluorescence raw data to Excel. A custom MATLAB script was then used to process the data into graphs (Figure 11).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

Research reported in this publication was also supported by the National Institute of General Medical Sciences of the US National Institutes of Health under grant number P50 GM098792 and the United States National Science Foundation under grant numbers 1124247, 1522074, 1521925 and 1521759.

Abbreviations

PERSIA	<u>P</u> URE <u>Ex</u> press <u>Re</u> As <u>H</u> <u>S</u> pinach <u>I</u> n-vitro <u>A</u> nalysis
CFS	Cell Free Synthesis

References

- [1]. Khorana HG (1968) Nucleic acid synthesis in the study of the genetic code, Nobel Lectures: Physiology or Medicine (1963–1970), 341–369.
- [2]. Nirenberg MW, and Matthaei JH (1961) The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides, *Proc. Natl. Acad. Sci. U. S. A* 47, 1588–1602. [PubMed: 14479932]
- [3]. Pardee K, Green AA, Ferrante T, Cameron DE, DaleyKeyser A, Yin P, and Collins JJ (2014) Paper-based Synthetic Gene Networks, *Cell* 159, 940–954. [PubMed: 25417167]
- [4]. Sun ZZ, Yeung E, Hayes CA, Noireaux V, and Murray RM (2014) Linear DNA for rapid prototyping of synthetic biological circuits in an *Escherichia coli* based TX-TL cell-free system, *ACS Synth. Biol* 3, 387–397. [PubMed: 24303785]
- [5]. Siegal-Gaskins D, Tuza ZA, Kim J, Noireaux V, and Murray RM (2014) Gene Circuit Performance Characterization and Resource Usage in a Cell-Free “Breadboard”, *ACS Synth. Biol* 3, 416–425. [PubMed: 24670245]
- [6]. Kim J, Khetarpal I, Sen S, and Murray RM (2014) Synthetic circuit for exact adaptation and fold-change detection, *Nucleic Acids Res.* 42, 6078–6089. [PubMed: 24728988]
- [7]. Yeung E, Dy AJ, Martin KB, Ng AH, Del Vecchio D, Beck JL, Collins JJ, and Murray RM (2017) Biophysical Constraints Arising from Compositional Context in Synthetic Gene Networks, *Cell Syst.* 5, 11–24.e12. [PubMed: 28734826]
- [8]. Chappell J, Jensen K, and Freemont PS (2013) Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology, *Nucleic Acids Res.* 41, 3471–3481. [PubMed: 23371936]
- [9]. Dudley QM, Karim AS, and Jewett MC (2015) Cell-Free Metabolic Engineering: Biomanufacturing beyond the cell, *Biotech. J* 10, 69–82.
- [10]. Karim AS, and Jewett MC (2016) A cell-free framework for rapid biosynthetic pathway prototyping and enzyme discovery, *Metab. Eng* 36, 116–126. [PubMed: 26996382]
- [11]. Majumder S, Wang S, Emery NJ, and Liu AP (2018) Simultaneous monitoring of transcription and translation in mammalian cell-free expression in bulk and in cell-sized droplets, *Synth. Biol* 3.
- [12]. Smith DB, and Johnson KS (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase, *Gene* 67, 31–40. [PubMed: 3047011]
- [13]. Hochuli E, Döbeli H, and Schacher A (1987) New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues, *J. Chromatogr. A* 411, 177–184.
- [14]. Schmidt TGM, and Skerra A (1993) The random peptide library-assisted engineering of a C-terminal affinity peptide, useful for the detection and purification of a functional Ig Fv fragment, *Protein Eng., Des. Sel* 6, 109–122.

- [15]. Hopp TP, Prickett KS, Price VL, Libby RT, March CJ, Cerretti DP, Urdal DL, and Conlon PJ (1988) A short polypeptide marker sequence useful for recombinant protein identification and purification, *Nat. Biotechnol* 6, 1204.
- [16]. Wagner E, Plank C, Zatloukal K, Cotten M, and Birnstiel ML (1992) Influenza virus hemagglutinin HA-2 N-terminal fusogenic peptides augment gene transfer by transferrin-polylysine-DNA complexes: toward a synthetic virus-like gene-transfer vehicle, *Proc. Natl. Acad. Sci. U. S. A* 89, 7934–7938. [PubMed: 1518816]
- [17]. Hanazono Y, Yu J-M, Dunbar CE, and Emmons RVB (1997) Green Fluorescent Protein Retroviral Vectors: Low Titer and High Recombination Frequency Suggest a Selective Disadvantage, *Hum. Gene Ther* 8, 1313–1319. [PubMed: 9295126]
- [18]. Liu H-S, Jan M-S, Chou C-K, Chen P-H, and Ke N-J (1999) Is Green Fluorescent Protein Toxic to the Living Cells?, *Biochem. Biophys. Res. Commun* 260, 712–717. [PubMed: 10403831]
- [19]. Hanson MR, and Köhler RH (2001) GFP imaging: methodology and application to investigate cellular compartmentation in plants, *J. Exp. Bot* 52, 529–539. [PubMed: 11373302]
- [20]. Feilmeier BJ, Iseminger G, Schroeder D, Webber H, and Phillips GJ (2000) Green Fluorescent Protein Functions as a Reporter for Protein Localization in *Escherichia coli*, *J. Bacteriol* 182, 4068–4076. [PubMed: 10869087]
- [21]. Waldo GS, Standish BM, Berendzen J, and Terwilliger TC (1999) Rapid protein-folding assay using green fluorescent protein, *Nat. Biotechnol* 17, 691. [PubMed: 10404163]
- [22]. Andresen M, Schmitz-Salue R, and Jakobs S (2004) Short tetracysteine tags to beta-tubulin demonstrate the significance of small labels for live cell imaging, *Mol. Biol. Cell* 15, 5616–5622. [PubMed: 15469986]
- [23]. Griffin BA, Adams SR, and Tsien RY (1998) Specific covalent labeling of recombinant protein molecules inside live cells, *Science* 281, 269–272. [PubMed: 9657724]
- [24]. Grate D, and Wilson C (1999) Laser-mediated, site-specific inactivation of RNA transcripts, *Proc. Natl. Acad. Sci. U. S. A* 96, 6131–6136. [PubMed: 10339553]
- [25]. Paige JS, Wu KY, and Jaffrey SR (2011) RNA Mimics of Green Fluorescent Protein, *Science* 333, 642. [PubMed: 21798953]
- [26]. Strack RL, Disney MD, and Jaffrey SR (2013) A superfolder Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA, *Nat. Methods* 10, 1219–1224. [PubMed: 24162923]
- [27]. Filonov GS, Moon JD, Svensen N, and Jaffrey SR (2014) Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution, *J. Am. Chem. Soc* 136, 16299–16308. [PubMed: 25337688]
- [28]. Dolgosheina EV, Jeng SC, Panchapakesan SS, Cojocaru R, Chen PS, Wilson PD, Hawkins N, Wiggins PA, and Unrau PJ (2014) RNA mango aptamer-fluorophore: a bright, high-affinity complex for RNA labeling and tracking, *ACS Chem. Biol* 9, 2412–2420. [PubMed: 25101481]
- [29]. Niederholtmeyer H, Xu L, and Maerkl SJ (2013) Real-Time mRNA Measurement during an In Vitro Transcription and Translation Reaction Using Binary Probes, *ACS Synth. Biol* 2, 411–417. [PubMed: 23654250]
- [30]. Kawakami T, Ogawa K, Goshima N, and Natsume T (2015) DIVERSE System: De Novo Creation of Peptide Tags for Non-enzymatic Covalent Labeling by In Vitro Evolution for Protein Imaging Inside Living Cells, *Chem. Biol* 22, 1671–1679. [PubMed: 26687484]
- [31]. Nagaraj VH, Greene JM, Sengupta AM, and Sontag ED (2017) Translation inhibition and resource balance in the TX-TL cell-free gene expression system, *Synth. Biol* 2.
- [32]. Jackson SE, Craggs TD, and Huang J. r. (2006) Understanding the folding of GFP using biophysical techniques, *Expert Rev. Proteomics* 3, 545–559. [PubMed: 17078767]
- [33]. Garamella J, Marshall R, Rustad M, and Noireaux V (2016) The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology, *ACS Synth. Biol* 5, 344–355. [PubMed: 26818434]
- [34]. Niederholtmeyer H, Stepanova V, and Maerkl SJ (2013) Implementation of cell-free biological networks at steady state, *Proc. Natl. Acad. Sci. U. S. A* 110, 15985–15990. [PubMed: 24043836]
- [35]. Leibly DJ, Nguyen TN, Kao LT, Hewitt SN, Barrett LK, and Van Voorhis WC (2012) Stabilizing additives added during cell lysis aid in the solubilization of recombinant proteins, *PLoS One* 7, e52482. [PubMed: 23285060]

- [36]. Wu P, Castner DG, and Grainger DW (2008) Diagnostic devices as biomaterials: a review of nucleic acid and protein microarray surface performance issues, *J. Biomater. Sci. Polym. Ed* 19, 725–753. [PubMed: 18534094]
- [37]. Li J, Gu L, Aach J, and Church GM (2014) Improved cell-free RNA and protein synthesis system, *PLoS One* 9, e106232. [PubMed: 25180701]
- [38]. Isaacs FJ, Carr PA, Wang HH, Lajoie MJ, Sterling B, Kraal L, Tolonen AC, Gianoulis TA, Goodman DB, Reppas NB, Emig CJ, Bang D, Hwang SJ, Jewett MC, Jacobson JM, and Church GM (2011) Precise Manipulation of Chromosomes in Vivo Enables Genome-Wide Codon Replacement, *Science* 333, 348–353. [PubMed: 21764749]
- [39]. Lajoie MJ, Rovner AJ, Goodman DB, Aerni H-R, Haimovich AD, Kuznetsov G, Mercer JA, Wang HH, Carr PA, Mosberg JA, Rohland N, Schultz PG, Jacobson JM, Rinehart J, Church GM, and Isaacs FJ (2013) Genomically Recoded Organisms Expand Biological Functions, *Science* 342, 357–360. [PubMed: 24136966]
- [40]. Lajoie MJ, Söll D, and Church GM (2016) Overcoming challenges in engineering the genetic code, *J. Mol. Biol* 428, 1004–1021. [PubMed: 26348789]
- [41]. Lajoie MJ, Rovner AJ, Goodman DB, Aerni H-R, Haimovich AD, Kuznetsov G, Mercer JA, Wang HH, Carr PA, Mosberg JA, Rohland N, Schultz PG, Jacobson JM, Rinehart J, Church GM, and Isaacs FJ (2013) Genomically Recoded Organisms Expand Biological Functions, *Science* 342, 357–360. [PubMed: 24136966]
- [42]. Borkowski O, Bricio C, Murgiano M, Rothschild-Mancinelli B, Stan G-B, and Ellis T (2018) Cell-free prediction of protein expression costs for growing cells, *Nat. Commun* 9, 1457. [PubMed: 29654285]
- [43]. Salis HM, Mirsky EA, and Voigt CA (2009) Automated design of synthetic ribosome binding sites to control protein expression, *Nat. Biotechnol* 27, 946. [PubMed: 19801975]
- [44]. Salis HM (2011) Chapter two - The Ribosome Binding Site Calculator, In *Methods Enzymol.* (Voigt C, Ed.), pp 19–42, Academic Press.
- [45]. Espah Borujeni A, Channarasappa AS, and Salis HM (2014) Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites, *Nucleic Acids Res.* 42, 2646–2659. [PubMed: 24234441]
- [46]. Rhee SY (2003) Human immunodeficiency virus reverse transcriptase and protease sequence database, *Nucleic Acids Res.* 31, 298–303. [PubMed: 12520007]
- [47]. Shafer RW (2006) Rationale and Uses of a Public HIV Drug-Resistance Database, *J. Infect. Dis* 194, S51–S58. [PubMed: 16921473]
- [48]. Cunningham BC, and Wells JA (1989) High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis, *Science* 244, 1081. [PubMed: 2471267]
- [49]. Vergne L, Peeters M, Mpoudi-Ngole E, Bourgeois A, Liegeois F, Toure-Kane C, Mboup S, Mulanga-Kabeya C, Saman E, Jourdan J, Reynes J, and Delaporte E (2000) Genetic Diversity of Protease and Reverse Transcriptase Sequences in Non-Subtype-B Human Immunodeficiency Virus Type 1 Strains: Evidence of Many Minor Drug Resistance Mutations in Treatment-Naive Patients, *J. Clin. Microbiol* 38, 3919. [PubMed: 11060045]
- [50]. Venkatakrishnan B, Palii M-L, Agbandje-McKenna M, and McKenna R (2012) Mining the Protein Data Bank to Differentiate Error from Structural Variation in Clustered Static Structures: An Examination of HIV Protease, *Viruses* 4, 348. [PubMed: 22590675]
- [51]. Petropoulos CJ, Parkin NT, Limoli KL, Lie YS, Wrin T, Huang W, Tian H, Smith D, Winslow GA, and Capon DJ (2000) A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1, *Antimicrob. Agents Chemother* 44, 920–928. [PubMed: 10722492]
- [52]. Rhee SY, Taylor J, Fessel WJ, Kaufman D, Towner W, Troia P, Ruane P, Hellinger J, Shirvani V, Zolopa A, and Shafer RW (2010) HIV-1 protease mutations and protease inhibitor cross-resistance, *Antimicrob. Agents Chemother* 54, 4253–4261. [PubMed: 20660676]
- [53]. Busti AJ, Hall II RG, and Margolis DM (2004) Atazanavir for the Treatment of Human Immunodeficiency Virus Infection, *Pharmacotherapy* 24, 1732–1747. [PubMed: 15585441]
- [54]. Back D, Sekar V, and Hoetelmans RM (2008) Darunavir: pharmacokinetics and drug interactions, *Antiviral Ther.* 13, 1–13.

- [55]. Boffito M, Jackson A, Amara A, Back D, Khoo S, Higgs C, Seymour N, Gazzard B, and Moyle G (2011) Pharmacokinetics of Once-Daily Darunavir-Ritonavir and Atazanavir-Ritonavir over 72 Hours following Drug Cessation, *Antimicrob. Agents Chemother* 55, 4218–4223. [PubMed: 21709075]
- [56]. Jackson A, Hill A, Puls R, Else L, Amin J, Back D, Lin E, Khoo S, Emery S, Morley R, Gazzard B, and Boffito M (2011) Pharmacokinetics of plasma lopinavir/ritonavir following the administration of 400/100 mg, 200/150 mg and 200/50 mg twice daily in HIV-negative volunteers, *J. Antimicrob. Chemother* 66, 635–640. [PubMed: 21172791]
- [57]. Lopez-Cortes LF, Ruiz-Valderas R, Sánchez-Rivas E, Lluch A, Gutierrez-Valencia A, Torres-Cornejo A, Benmarzouk-Hidalgo OJ, and Viciano P (2013) Lopinavir plasma concentrations and virological outcome with lopinavir-ritonavir monotherapy in HIV-1-infected patients, *Antimicrob. Agents Chemother* 57, 3746–3751. [PubMed: 23716055]
- [58]. Liberman V, Hamad-Schifferli K, Thorsen TA, Wick ST, and Carr PA (2015) In situ microfluidic SERS assay for monitoring enzymatic breakdown of organophosphates, *Nanoscale* 7, 11013–11023. [PubMed: 26041657]
- [59]. Karzbrun E, Tayar AM, Noireaux V, and Bar-Ziv RH (2014) Programmable on-chip DNA compartments as artificial cells, *Science* 345, 829. [PubMed: 25124443]
- [60]. Swank Z, Laohakunakorn N, and Maerkl SJ (2019) Cell-free gene-regulatory network engineering with synthetic transcription factors, *Proc. Natl. Acad. Sci. U. S. A.*, 201816591.
- [61]. Gerber D, Maerkl SJ, and Quake SR (2009) An in vitro microfluidic approach to generating protein-interaction networks, *Nat. Methods* 6, 71–74. [PubMed: 19098921]
- [62]. Wong I, and Ho C-M (2009) Surface molecular property modifications for poly(dimethylsiloxane) (PDMS) based microfluidic devices, *Microfluid. Nanofluid* 7, 291. [PubMed: 20357909]
- [63]. Moore SJ, MacDonald JT, Wienecke S, Ishwarbhai A, Tsipa A, Aw R, Kylilis N, Bell DJ, McClymont DW, Jensen K, Polizzi KM, Biedendieck R, and Freemont PS (2018) Rapid acquisition and model-based analysis of cell-free transcription–translation reactions from nonmodel bacteria, *Proc. Natl. Acad. Sci. U. S. A* 115, E4340. [PubMed: 29666238]
- [64]. Niederholtmeyer H, Sun ZZ, Hori Y, Yeung E, Verpoorte A, Murray RM, and Maerkl SJ (2015) Rapid cell-free forward engineering of novel genetic ring oscillators, *eLife* 4, e09771. [PubMed: 26430766]
- [65]. Kong DS, Carr PA, Chen L, Zhang S, and Jacobson JM (2007) Parallel gene synthesis in a microfluidic device, *Nucleic Acids Res.* 35, e61–e61. [PubMed: 17405768]
- [66]. Kong DS, Thorsen TA, Babb J, Wick ST, Gam JJ, Weiss R, and Carr PA (2017) Open-source, community-driven microfluidics with Metafluidics, *Nat. Biotechnol* 35, 523. [PubMed: 28591125]

Figure 1.

Schematic of PERSIA. PERSIA employs biomolecular tag sequences and fluorescent small-molecule labels during experiments using cell-free extracts. This approach avoids the need for further time- and resource-consuming sample processing steps after the cell-free experiment. While the core elements of this method are monitoring transcription and translation, PERSIA can also accommodate additional assays, such as detecting the products of enzymes generated in the cell-free reaction. Depending on available fluorescent tools, the additional assay for enzymatic products can either employ a different spectral region (color) than used for detecting mRNA and protein, or replace one of these. (See PERSIA with Additional Enzymatic Assay, below.)

Figure 2.

Optimizing fluorophore concentrations for PERSIA. (A) Increasing amounts of DFHBI were added to the PURExpress reaction to determine an effective concentration for measuring mRNA present through DFHBI binding to the Spinach RNA tag. 50 μ M was chosen as the standard amount of DFHBI to be used in future reactions due to a combination of low background and high signal. (B) Increasing amounts of ReAsH-EDT₂ were added to the PURExpress reaction to find an effective concentration to quantitate the amount of protein present through ReAsH-EDT₂ binding to the tetracysteine (TC) tag. 5 μ M was chosen as the standard amount of ReAsH to be used in future reactions.

Figure 3.

Representative standard curves generated in PERSIA samples. (A) Increasing concentrations of Spinach-tagged ubiquitin mRNA purified from a prior T7 transcription reaction are added to a PERSIA reaction to determine the amount of fluorescence produced by a specific concentration of mRNA. Standard curves are produced during each experiment to accurately assess the response of the PERSIA reaction. (B) Increasing concentrations of TC-tagged ubiquitin protein (purified from prior expression in *E. coli* cells) are added to a PERSIA reaction to determine the amount of fluorescence produced by a specific amount of protein. Standard curves such as these are produced during each PERSIA experiment.

Figure 4.

Binding kinetics of fluorescent molecules to biological domains of interest. (A) DFHBI demonstrates delayed binding to 1 μg of purified Spinach-tagged mRNA in a PERSIA reaction, as the RNA must establish intramolecular base pair binding to establish the proper structure necessary for DFHBI fluorescence. (B) ReAsH-EDT₂ demonstrates rapid binding to 1 μg of purified TC-tagged ubiquitin protein.

Figure 5.

Comparison of fluorescence onset time for (A) TC/ReAsH fluorescence and (B) eGFP protein fluorescence from a TC-tagged eGFP (no 3' Spinach tag) being produced in a PERSIA reaction. Concentrations of DNA template (0, 1, 5, or 25 nM) are indicated. For each template concentration, the threshold time was determined as the point where (baselined) fluorescent signal reached 3 times the standard deviation of the background (0 nM DNA sample). The onset time for the 5 nM concentration is indicated by a dashed line.

Figure 6.

Gene expression analysis of *E. coli* ribosome proteins using PERSIA. (A) Transcriptional and (B) translational comparison of genes under transcriptional control of the same T7 promoter sequence shows variation in the amount of Spinach-tagged RNA and tetracycline-tagged protein, respectively, produced from identical concentrations of starting dsDNA template.

Figure 7.

The effect of genetic code variations on transcription and translation rates on rplX. (A) Changes from the wild type Arg86 codon (AGC) to four other arginine codons (CGX, versions R86a through R86d) has little effect on the translation rate or final yield. Implementing an extreme, minimal genetic code—with only one codon is allowed per amino acid—severely diminished the translation rate of rplX (version “Min”). (B) In spite of extreme changes in codon usage which reduce the translation rate (version “Min”; line (i)), restoring the first 30 bp of coding region (variant “Hybrid”) to the wild type sequence can restore high levels of translation (line ii). Further, three base changes in the first 5 codons (Min-rev123), or separately engineering the ribosome binding sequence (version Min-RBS-6k) also restored high levels of translation (line iii). (C) Transcriptional output measured by Spinach-DFHBI (same PERSIA reactions as 7A) show minor variations, but do not correlate with decrease in translational output of extreme codon limitations (“Min”), demonstrating that drastic changes in codon usage can directly affect translation with minimal consequences for transcription.

Figure 8.

Biochemical assay analysis coupled to PERSIA. (A) ReAsH fluorescence demonstrating the production of TC-tagged HIVpr. (B) SensoLyte520 fluorescence indicating HIV-1 protease activity during the same PERSIA reactions. The onset of fluorescence is a result of unquenching of the fluorophore upon proteolytic peptide cleavage. (C) Spinach/DHFBF fluorescence indicating mRNA production in a parallel reaction. Spinach/DHFBF and SensoLyte520 were not used in the same reactions due to spectral overlap of these reagents.

HIV Protease Crystal Structure: Courtesy of National Institute of Health

Figure 9.

Sequence-function analysis of an HIV-1 protease variant sequence (HIVpr) by alanine scanning synthesis. 97 variants of the protein sequence were designed for DNA synthesis, each containing a different amino acid substitution to alanine to observe potential disruptions of enzyme function. Non-native flanking sequences (N-terminal MSG- and C-terminal - GSCCPGCCGSHHHHH) were not modified and are not shown. Positions that already possess alanine residues in the sequence (22 and 28) are displayed with the activity value of the unmodified sequence (denoted by the dotted line). For variants T91A and F99A, DNA synthesis by the vendor was unsuccessful and omitted. Values for variants G17A and E34A (off-scale) are 8.4 and 10.9, respectively. Positions flanking the catalytic D25 amino acid (23-27) are shown in red and the tip of the flap region involved in dimerization (49-52) is shown in yellow. HIVpr protein structure from ⁵⁰.

Figure 10.

PERSIA analysis of clinical HIV protease variants. Protein concentration quantitation is made through extrapolating TC-tag/ReAsH fluorescence using a protein standard curve, while Sensolyte cleavage quantitation is made through serial dilution of a fluorescent standard included with the commercial assay. The rate of Sensolyte cleavage by the amount of HIVpr variants produced is calculated in the presence, or absence, of drugs known to inhibit HIV protease enzyme activity. Activity rates in the absence of drug are designated as having 100% activity for each genetic variant of HIVpr. All other rates are calculated relative to that value. HIV Drug Resistance Database scores included for comparison derive from a combination of clinical drug data and amino acid variants known to result in resistance to specific drugs. Wild-type and clinical variants analyzed in the presence of HIV antiviral drugs targeting the protease (A) Atazanavir, (B) Darunavir, and (C) Lopinavir.

Figure 11.

Integration of PERSIA into microfluidic device. (A) Schematic of device with 6 banks and representative images of ReAsH and Spinach signal. (B) Spinach/DFHBI and (C) TC/

ReAsH fluorescence of several DNA sequences ($n = 3$, error bars are standard error of the mean) normalized by highest single measurement and then background subtracted.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Effects of additives to the PERSIA reaction.

Additive	Effect on Spinach-DFHBI signal	Effect on protein ReAsH signal	Details regarding ReAsH signal
Zinc Chloride	No effect Up to 100 μ M	Reduced with additions >10 μ M	Reduced 1/3 when 20 μ M added No signal when 100 μ M added
Cobalt Chloride	No effect up to 100 μ M	Reduced with additions >10 μ M	Reduced by 1/3 when 20 μ M added No signal when 100 μ M added
Sodium Chloride	No effect up to 100 μ M	Reduced with additions >50 mM	No signal when 150 mM added
Bovine Serum Albumin	No effect up to 4 mg/ml	High background when no DNA present	Large # of surface Cys-Cys pairs; BSA may naturally bind ReAsH
Dodecylmaltoside	No effect up to 0.5%	Modest increase up to 0.5%	Background increases 2x @ 0.1%; increases 3x @ 0.5%;
Pluronic F127	No effect up to 0.5%	Modest increase up to 0.5%	Background increases 1.5x @ 0.1%; increases 2x at 0.5%