nature chemical biology

ARTICLE

PUBLISHED ONLINE: 13 FEBRUARY 2017 | DOI: 10.1038/NCHEMBIO.2299

Evolution of a split RNA polymerase as a versatile biosensor platform

Jinyue Pu[1,2], Julia Zinkus-Boltz[1,2] & Bryan C Dickinson[1]*

Biosensors that transduce target chemical and biochemical inputs into genetic outputs are essential for bioengineering and synthetic biology. Current biosensor design strategies are often limited by a low signal-to-noise ratio, the extensive optimization required for each new input, and poor performance in mammalian cells. Here we report the development of a proximity-dependent split RNA polymerase (RNAP) as a general platform for biosensor engineering. After discovering that interactions between fused proteins modulate the assembly of a split T7 RNAP, we optimized the split RNAP components for protein–protein interaction detection by phage-assisted continuous evolution (PACE). We then applied the resulting activity-responsive RNAP (AR) system to create biosensors that can be activated by light and small molecules, demonstrating the 'plug-and-play' nature of the platform. Finally, we validated that ARs can interrogate multidimensional protein–protein interactions and trigger RNA nanostructure production, protein synthesis, and gene knockdown in mammalian systems, illustrating the versatility of ARs in synthetic biology applications.

Diverse areas of chemical biology and biotechnology, including directed evolution, synthetic biology, and bioengineering, require methods to link chemical and biochemical processes to defined genetic outputs[1–4]. RNA is a particularly useful output owing to the availability of technologies to drive cellular responses on the basis of computed nucleic acid signals[5–7]. In nature, transcription factors detect target activities and drive genetic responses, and many naturally occurring transcription factors, such as the Lac and Tet repressors, have been repurposed as gene expression control elements[4,8]. However, because the scope of detection of such natural systems is limited, several engineered alternative technologies have been developed.

n-hybrid systems, such as those for detection of protein–DNA (one-hybrid), protein–protein (two-hybrid), protein–RNA (three-hybrid), and protein–small-molecule (three-hybrid) interactions, are among the approaches most commonly deployed for genetic sensor design[9–12]. The development of programmable DNA binding domains such as transcription activator-like effectors (TALEs) and nuclease-deficient Cas9 (dCas9) have revolutionized one- and two-hybrid approaches[8,13–15]. However, a challenge with *n*-hybrid systems is that they must be carefully tuned and optimized for each new interaction, and more complex multicomponent systems such as three-hybrids often lack sensitivity and have low signal-to-noise ratios[16,17]. Although one-hybrids (and two-hybrids, in some cases) can be used for synthetic biology purposes, these methods are generally not suitable for applications that require a high level of control and dynamic range. A primary alternative to *n*-hybrid approaches is riboswitches, RNA-based elements that drive translational outputs on the basis of an aptamer's interaction with a target ligand[18,19]. Riboswitches have become an important tool for designing systems that respond to target inputs, which are usually small molecules but can also be proteins[20]. Although riboswitches can work well with targets for which aptamers can be created, protein-based biochemical activities are generally beyond their scope of detection, and translating these tools in mammalian systems often results in diminished performance[17]. Outside of these general strategies, a host of other synthetic biology parts have been developed for highly specialized activities[2,21]. Therefore, there is a need for a general method to transduce endogenous chemical and biochemical information into DNA or RNA for subsequent storage or integration with engineered regulatory systems.

Engineered polymerases have been proposed as a strategy to respond to and measure endogenous biochemical processes[22,23]. Inspired by this concept, we developed protease-responsive RNAPs (PRs) to respond to protease activity by production of defined RNA outputs[24]. We showed that PRs can encode multidimensional protease activities in defined sequences of RNA in both prokaryotic and mammalian cells. In principle, RNAPs provide a new platform for biosensor creation, but engineering such complex enzymes is challenging, and the inhibitor-based design strategy used for PRs is fundamentally limited to protease activities.

We envisioned a new biosensor system based on recently reported 'split' T7 RNAPs[25,26], in which N- and C-terminal components of the T7 RNAP spontaneously assemble to form a functional RNAP enzyme. We aimed to engineer ARs in which assembly is not spontaneous but instead depends on fused interaction partners, similarly to other protein fragment complementation technologies (PFCs)[27]. PFC involves tethering potential binding partners to complementary halves of a split protein marker, such as a fluorescent protein[28]. Generally, the protein binding partners must be expressed at much higher levels than endogenous conditions to achieve a detectable signal, and only two or fewer signals can be monitored at once[29]. More importantly, the fluorescence output in PFC cannot be integrated into downstream synthetic control systems. We reasoned that if a proximity-dependent split RNAP could be developed, genetic biosensors could be created through a 'plug-and-play' approach in which target interacting domains can be swapped in, resulting in a GFP-like platform for biosensor engineering (**Fig. 1a**).

In this study we engineered a proximity-dependent split T7 RNAP sensor using continuous molecular evolution. We then demonstrated the versatility and ease of use of the platform by creating robust light- and small-molecule-responsive genetic sensors. To illustrate the power of polymerase-based biosensors, we showed that multidimensional protein–protein interactions (PPIs) can be monitored using ARs. Finally, we confirmed that ARs can trigger RNA nanostructure formation, protein synthesis, and gene knockdown in mammalian cells using a small-molecule-triggered AR biosensor. The AR platform greatly simplifies and expands genetic circuit

[1]Department of Chemistry, The University of Chicago, Chicago, Illinois, USA. [2]These authors contributed equally to this work. *e-mail: dickinson@uchicago.edu

creation and opens up new opportunities in protein engineering, synthetic biology, and bioengineering.

RESULTS

Biophysical feasibility of proximity-dependent split RNAPs

Two key biophysical requirements for our proposed AR strategy are that fused protein domains do not sterically interfere with the split RNAP and that interactions of fused domains can influence the RNAP assembly process. We chose to deploy T7 RNAP split at position 179 because (i) the N-terminal half is small, (ii) structural data indicate that this position is solvent exposed and removed from the DNA-binding face of the protein, and (iii) mutations that influence DNA promoter specificity are C terminal to this position[26,30,31]. First, we validated that the two RNAP halves spontaneously assemble using an *Escherichia coli* luciferase reporter system (**Fig. 1b** and **Supplementary Results**, **Supplementary Fig. 1** and **Supplementary Table 1**). We then fused leucine zipper peptides that form a tight interaction with one another (ZA and ZB)[32,33] to the split RNAP halves. Fusion of ZA or ZB to only the N-terminal or C-terminal RNAP, respectively, did not dramatically affect spontaneous RNAP assembly, indicating that the split RNAP can tolerate fusions. However, fusion of both RNAP halves to the interaction partners resulted in a ~5-fold enhancement in transcription, indicating that additional pendant interactions enhance split RNAP assembly (**Fig. 1c**). A triple mutant of ZB (ZBneg) that weakens the interaction confirmed that the enhancement in transcription was due to the fused PPI. We note that T7 RNAP undergoes large conformational changes during the course of its enzymatic activity[34], including dramatic structural changes in the N-terminal RNAP half. It is therefore not obvious that the RNAP can tolerate fusions or that the interactions between fusions can modulate the assembly process while maintaining enzymatic viability. However, because these preliminary data confirmed that the split enzyme is amenable to controlled assembly, we turned our attention to engineering the split RNAP to be more dependent on the fused interaction partners.

Development of an evolution to optimize split RNAPs

For split RNAPs to work as a platform for biosensor design, the RNAP assembly process needs to be more dependent on fused interaction partners. Achieving this involves tuning the assembly of the RNAP halves while maintaining all other aspects of RNAP enzymatic function, including DNA binding, nucleotide binding, and RNA synthesis. Such a protein engineering problem presents substantial challenges, but these can in principle be overcome by molecular evolution. We chose to deploy PACE, a rapid evolution system[35] that has been used to evolve RNAP promoter specificity, protein–DNA interactions, protease activities, and PPIs[3,30,35–40]. Briefly, PACE involves providing an evolving gene of interest to M13 bacteriophage, linking the life cycle of the phage to an activity of interest to be evolved in the target gene, and then propagating the virus until the activity evolves. Expression of *gIII*, a required phage gene, is the basis of the life cycle link.

We envisioned a PACE system for the evolution of selective assembly for split RNAPs using leucine zipper peptides as a model PPI. In this system, the phage would carry an evolving N-terminal RNAP variant fused to ZA, and the host *E. coli* cells would express two different C-terminal variants, each with orthogonal DNA promoter specificity, fused to either ZB or ZBneg. Assembly of the evolving phage-carried ZA-fused N-terminal RNAP with the ZB-fused C-terminal RNAP in the host cells would result in enhanced phage propagation, whereas assembly with the ZBneg-fused C-terminal RNAP would decrease phage propagation. We postulated that this simultaneous positive–negative selection would result in the most robust evolutionary outcome. We would use relatively long (6–8 amino acids), unstructured linkers to tether the fusion proteins

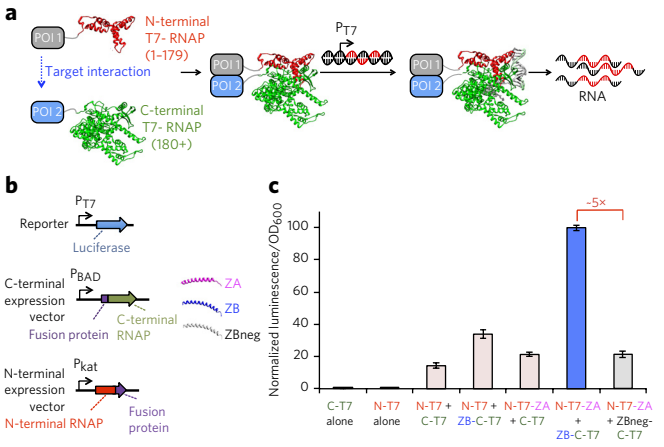

Figure 1 | Design and biophysical feasibility of ARs based on proximity-dependent split RNAPs. (**a**) Schematic of AR design. Split T7 RNAP assembles into a functional RNAP when proteins of interest (POIs) fused to each half interact, resulting in transcription of a user-defined sequence of RNA from a supplied DNA substrate. P_{T7}, T7 promoter. (**b**) Vectors designed to test split RNAPs *in vivo*. N-terminal split RNAP (red) and C-terminal split RNAP (green) were fused to ZA (pink), ZB (blue), or ZBneg (gray). P_{kat}, kat promoter. (**c**) Transcriptional output of split RNAPs with fusion proteins assayed in *E. coli* using the vectors shown in **b**. Cells were induced for 2 h with arabinose then analyzed for luminescence. Error bars, s.e.m., $n = 4$ biological replicates. Data are normalized to signal from the N-T7-ZA–ZB-C-T7 interaction.

to the RNAP halves, enforcing the evolution of a mechanism for proximity-dependent RNAP assembly that is less dependent on geometry and linker composition, which we hypothesized would result in a more versatile biosensor platform.

To develop this new PACE system, we engineered M13 phage by replacing *gIII* with N-terminal T7 RNAP fused to ZA, which is the target of the evolution (**Fig. 2a**). *E. coli* cells were engineered with a 'positive accessory' plasmid (posAP) that expresses a ZB-fused C-terminal T7 RNAP variant (C-term CGG RNAP) containing seven point mutations that allow it to act selectively on the CGG promoter over the T7 promoter[25,31] and CGG promoter–driven *gIII* (**Supplementary Fig. 1i,j**). We also engineered a negative accessory plasmid (negAP), which expresses wild-type C-terminal T7 RNAP fused to ZBneg and a dominant-negative form of *gIII* under control of the T7 promoter[37] (**Supplementary Fig. 1k**). Therefore, if an evolving N-terminal RNAP variant assembles efficiently with the T7 and CGG C-terminal RNAP halves regardless of the fusion protein, phage production will be blocked. However, phage encoding N-terminal variants that selectively assemble with the ZB-fused CGG C-terminal half will replicate more efficiently and continue to mutate until the interaction is optimized (**Fig. 2b**). The only differences between the positive and negative selection are whether or not the fused peptides interact, which is based on three point mutations in ZBneg, and the seven mutations that alter DNA binding of the CGG C-terminal RNAP, which are not at the protein–protein interface and are not expected to alter RNAP assembly.

Evolution of a proximity-dependent split RNAP using PACE

After cloning and validating the system components, we initiated PACE. We modified the positive and negative selection pressures by carefully tuning the system components (**Fig. 2c** and **Supplementary Table 2**) and monitored the progress of the evolution by activity-dependent plaque assays and genetic analysis of the evolving phage (**Supplementary Table 3**). Specifically, we altered the concentrations of the on-target and off-target interactions and the strength of

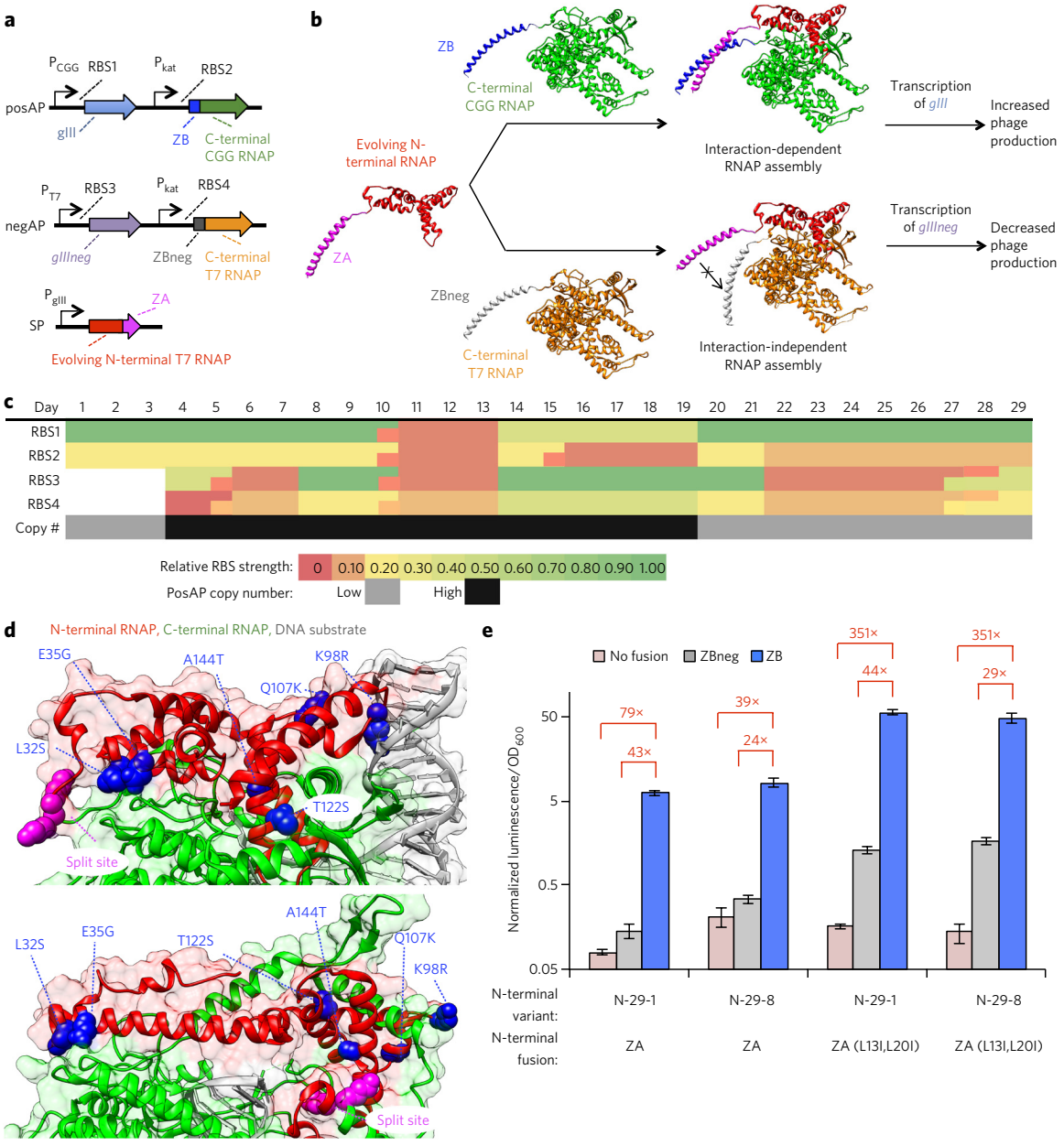

Figure 2 | Evolution of a proximity-dependent split RNAP for PPI detection. (**a**) Vectors for PACE system for proximity-dependent RNAPs. P_{CGG}, CGG promoter; P_{T7}, T7 promoter; P_{gIII}, *gIII* promoter, P_{kat}, kat promoter. (**b**) Mechanism of PACE system for proximity-dependent RNAPs. Phage carry an evolving N-terminal RNAP fused to ZA, which can assemble with either a C-terminal RNAP variant fused to ZB to allow phage replication or to the noninteracting ZBneg, which hinders phage production by producing a dominant-negative form of *gIII* (*gIIIneg*). (**c**) Schematic of evolution parameters used during PACE. The RBS strengths[50] controlling expression of the C-terminal RNAPs, RBSs controlling proteins on vectors shown in **a**, and posAP copy number were carefully tuned during a 29-d evolution. (**d**) Mapping the mutations of N-29-1 onto T7 RNAP crystal structures. Top, initiation complex (PDB 1QLN); bottom, elongation complex (PDB 1H38). (**e**) Transcriptional reporter assay of N-29-1 and N-29-8 fused to ZA or the evolved ZA double mutant, interacting with either the C-terminal RNAP alone (pink) or fused to ZB (blue) or ZBneg (gray). Error bars, s.e.m., $n = 4$ biological replicates. Data are plotted on a logarithmic scale to show background.

selection of the RNA output by tuning the ribosome-binding sites (RBSs) controlling each system component. After 3–4 d of PACE on a given target, we would use activity-dependent plaque assays to choose the subsequent evolutionary targets. This process continued for 29 d, after which the N-terminal RNAP converged on two main variants, one with six-mutations (N-29-1) and one with seven (N-29-8), with several of the mutations near the interface between the halves of the RNAP (**Fig. 2d**). Several mutations that are prevalent in the phage population are present in either solvent-exposed regions of the structure or, less predictably, at the protein–DNA interface[34,41,42]. This suggests epistatic interactions between mutations that tune the protein–protein interface with key mutations elsewhere in the protein, further illustrating why our unbiased directed evolution strategy is optimal for tuning a complex molecular machine such as an RNAP.

Assays of the two primary variants that emerged from PACE in the luciferase transcription reporter system revealed that the background level of transcription with the ZBneg control was much lower than with ZB (**Fig. 2e**). Further genetic analysis revealed that ZA, which we assumed was already fully optimized for interaction

with ZB, also evolved during PACE, converging on two leucine-to-isoleucine substitutions (L13I and L20I). Fusion of this ZA variant into N-29-1 and N-29-8 resulted in a dramatic enhancement in assembly of the split RNAP with the ZA and ZB partners but maintained low levels of background with the ZBneg control, with variant N-29-1 showing a 44-fold increase in RNA synthesis based on the interaction (**Fig. 2e**). This observation demonstrates that our split RNAP PACE system can be deployed to optimize biomolecular interactions similarly to recent two-hybrid PACE systems[36].

We hypothesized that the actual background of the evolved split RNAP was lower than that measured with the original ZA and ZB owing to enhancement of the interaction of ZA (L13I, L20I) with both ZB and ZBneg as a result of the isoleucine substitutions. To test this, we assayed the transcriptional output of N-29-1 and N-29-8 fused to ZA (L13I, L20I) with an unfused C-terminal RNAP. As expected, we found a dramatically lower background signal, indicating a measurable affinity for ZA (L13I, L20I) and ZBneg. Therefore, the actual dynamic range of ZA (L13I, L20I)–ZB-driven N-29-1 is >350-fold higher than that of the background assembly without a fused PPI (**Fig. 2e**). These results indicate that the assembly of the evolved RNAP is not only dependent on fused interaction partners but is also dynamically sensitive to the affinity of the interaction. Most notably, PACE yielded the N-terminal RNAP variant N-29-1 for the AR strategy.

ARs as light and small-molecule biosensors

Having optimized and validated the AR system for PPIs, we next sought to explore the generality of the approach by developing inducible PPI systems, and therefore targeted light- and small-molecule-activated AR sensors. To create a light-activated RNAP, we appended the light-oxygen-voltage 2 (LOV2)–SsrA fusion variant to N-29-1 and SspB from the improved light-induced dimer (iLID-nano) system[43] to the C-terminal RNAP (**Fig. 3a**) without additional optimization of linkers, geometry, or concentrations, and assayed the fusions in *E. coli*. Illumination with blue LED light, which induces dimerization of the iLID-nano system, resulted in a 26-fold enhancement in transcriptional output of the light-activated AR, whereas control fusions did not show a light response (**Fig. 3b**). The dynamic range of this non-optimized system is notable because the reported light-induced difference in affinity of the iLID-nano system proteins is only 36-fold[43].

Next, to further explore the versatility of ARs to detect a three-hybrid-like interaction, we sought to engineer a small-molecule-activated RNAP. We chose the rapamycin-induced dimerization of the proteins FKBP-rapamycin binding domain (FRB) and FK506 binding protein (FKBP), a workhorse small-molecule-induced dimerization system for many applications[44], as the next target for our split RNAP system. There are methods to control RNA synthesis with small molecules in prokaryotic systems, but few such methods function well in mammalian systems[45,46], and many have issues with signal-to-noise ratios and background signal. We replaced the peptide fusions on the *E. coli* expression vectors for the C-terminal RNAP and N-29-1 with FKBP and FRB, respectively, again without optimization of linkers, concentration, geometry, or any other system component (**Fig. 3c**). We observed a dramatic, dose-responsive increase in RNA synthesis, assessed in the *in vivo* luciferase transcription reporter assay, upon treatment with rapamycin, with a 340-fold enhancement in RNA synthesis and essentially undetectable background (**Fig. 3d**).

Compared with traditional *n*-hybrid strategies that require extensive linker and geometric optimization[36], the ARs appeared to be much less dependent on linkers. Using the rapamycin-inducible AR as a prototype, we experimentally assessed the effect of linker length on the split RNAP assembly process by varying the linker lengths of the fusion proteins from 2 to 14 amino acids. Transcriptional assays revealed minimal linker-length dependency, and the longer linkers

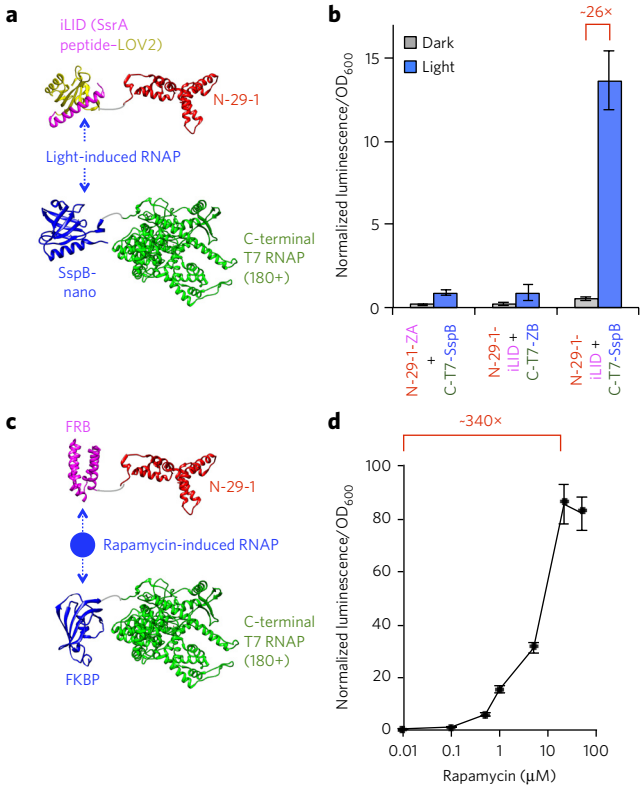

Figure 3 | Small-molecule- and light-responsive ARs. (**a**) Light-activated RNAP design using the iLID-nano system. (**b**) Transcription response of the light-activated RNAP system in *E. coli*. Cells were transformed with expression vectors for the halves of the light-inducible RNAP and a reporter vector then either kept in the dark or illuminated with blue LED light for 3 h before transcriptional analysis. (**c**) Small-molecule-responsive RNAP design using FRB and FKBP. (**d**) Transcription response of the rapamycin-inducible RNAP system in *E. coli*. Cells were transformed with expression vectors for the halves of the small-molecule-inducible RNAP and a reporter vector then induced with rapamycin for 3 h before transcription analysis. Error bars, s.e.m., $n = 4$ biological replicates (**b**,**d**).

performed slightly better (**Supplementary Fig. 2**). This might be due to the large conformational changes that occur at the split site (**Fig. 2d**), as they may be better accommodated by longer linkers. Regardless of the mechanism, these data demonstrate the ability to swap new binding domains into the AR system with minimal optimization to create biosensors. Finally, we cloned two split GFP sensors[33,47] into the rapamycin AR vector system, which, under the same experimental conditions, revealed no detectible signal, further illustrating the advantages of ARs (**Supplementary Fig. 3**).

ARs can monitor multidimensional PPI networks

One core advantage of an RNAP-based biosensor for analyzing endogenous molecular interactions is that the signals are encoded in an output RNA. Fluorescent-based split biosensors have problems with spectral overlap, differential binding affinities, and linker dependencies, limiting their capability for multidimensional analysis[48]. RNA, however, potentially permits highly multidimensional analysis. To explore this possibility, we next sought to test whether ARs with orthogonal RNA outputs could be used to monitor dynamic, multidimensional PPIs. We engineered a well-controlled, synthetic trimolecular PPI network with an inducible change in interactions to validate the concept. We first engineered N-29-1 fused to both ZA and FRB (FZ-N), which we could then deploy with a C-terminal CGG RNAP variant fused to ZB (Z-C_G)

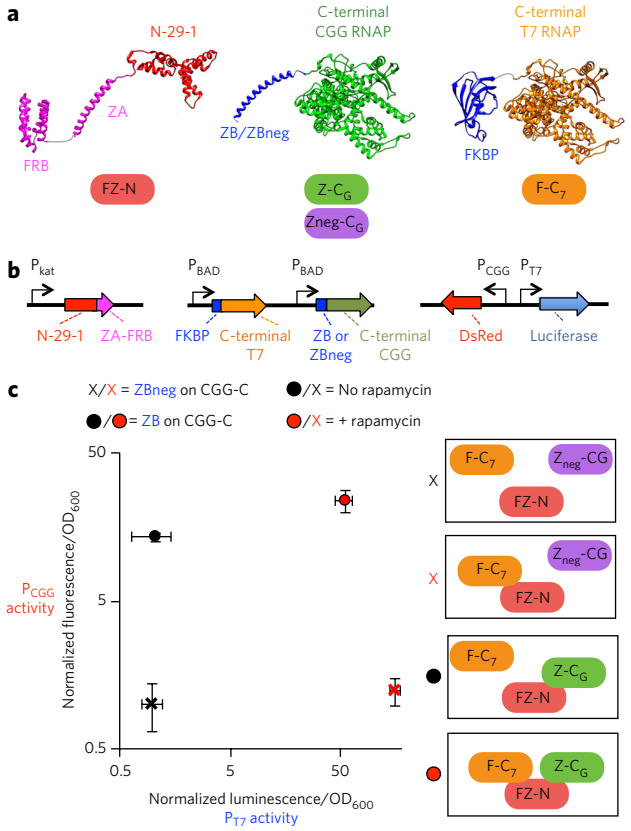

Figure 4 | Multidimensional PPI detection by ARs. (**a**) Design of a synthetic trimolecular protein interaction network using FZ-N and Z-C_G or Z_{neg}-C_G and F-C_7. In the absence of rapamycin, FZ-N–Z-C_G should be the dominant PPI, driving a CGG-promoter output. In the presence of rapamycin, FZ-N–F-C_7 should also be present. (**b**) Vectors used to monitor the two PPIs simultaneously. (**c**) Simultaneous monitoring of PPIs in the same cells. *E. coli* cells were transformed with expression vectors, induced with DMSO or 10 μM rapamycin for 5 h, then analyzed for luminescence and DsRed fluorescence. Error bars, s.e.m., n = 4 biological replicates.

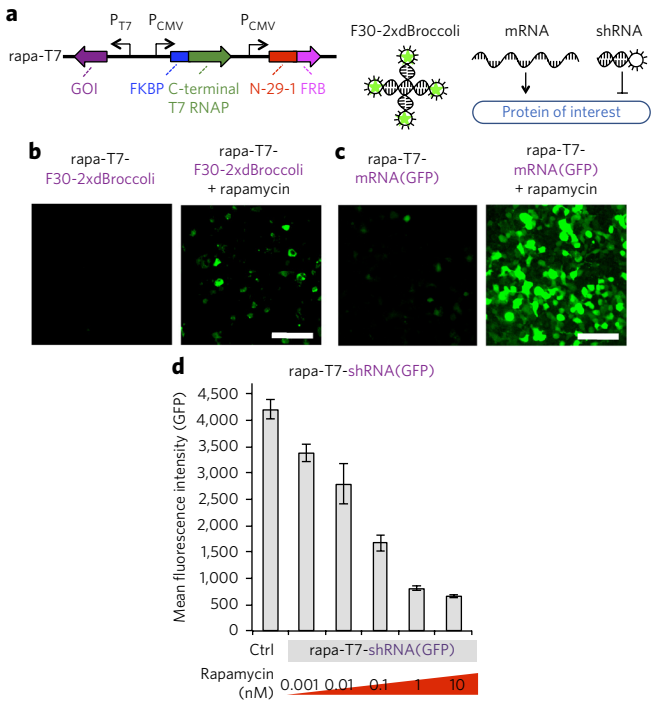

Figure 5 | ARs can trigger a variety of outputs in mammalian cells. (**a**) Design of the rapa-T7 vectors for rapamycin-induced transgene expression in mammalian systems. (**b**) Validation of the rapa-T7 vector with a fluorescent RNA aptamer as the output. HEK293T cells were transfected with rapa-T7-F30-2xdBroccoli and induced with 0 or 100 nM rapamycin for 30 min in the presence of 20 μM DFHBI-1T then analyzed by fluorescence microscopy. (**c**) Validation of the rapa-T7 vector with mRNA as the output. HEK293T cells transfected with rapa-T7-mRNA(GFP) were induced with 0 or 10 nM rapamycin overnight then analyzed by fluorescence microscopy. (**d**) GFP fluorescence (arbitrary units) analyzed by flow cytometry to validate rapa-T7 with shRNA as the output. HEK293T cells were transfected with a GFP expression vector and a rapa-T7-shRNA(GFP) vector and induced with varying concentrations of rapamycin for 44 h. Error bars, s.e.m., n = 3 biological replicates. Scale bars, 100 μm (**b**,**c**).

or ZBneg (Z_{neg}-C_G) and C-terminal T7 RNAP fused to FKBP (F-C_7) (**Fig. 4a**). In this design, interaction between FZ-N and Z-C_G should produce an RNA signal from the CGG promoter, whereas interaction between FZ-N and F-C_7 should produce an RNA signal from the T7 promoter (**Fig. 4b**).

We first validated the engineered trimolecular model using our luciferase reporter system with one interaction at a time, demonstrating that each interaction can be monitored and selectively acts only on the prescribed promoter (**Supplementary Fig. 4**). Next, to simultaneously monitor both interactions in the same cells, we redesigned the vector system to express both F-C_7 and Z-C_G or Z_{neg}-C_G along with FZ-N and produce DsRed from the CGG promoter and luciferase from the T7 promoter, allowing both transcriptional outputs to be monitored. We found that the FZ-N–Z-C_G interaction is dependent on the zipper peptides and is unresponsive to rapamycin, whereas the FZ-N–F-C_7 interaction is induced upon rapamycin addition (**Fig. 4c**). These data confirm that orthogonal C-terminal variants can be used along with N-29-1 to monitor multidimensional PPI networks in live cells.

ARs can control RNA synthesis in mammalian cells

Finally, we assayed the ability of the ARs to function in mammalian cells using the rapamycin-inducible system as an exemplar. We generated rapa-T7, a vector that expresses the rapamycin-inducible AR and contains a T7-promoter-driven gene of interest (GOI) output circuit (**Fig. 5a**). To confirm that RNA was being generated and measure the kinetics of AR activation, we first deployed a fluorescent aptamer (F30-2xdBroccoli)[49] as the GOI output of the rapa-T7 vector (rapa-T7-F30-2xdBroccoli) to enable visualization of RNA synthesis with fluorescence microscopy. Treatment of HEK293T cells transfected with the rapa-T7-F30-2xdBroccoli vector with 100 nM rapamycin for 30 min resulted in robust enhancement of intracellular fluorescence compared to cells not treated with rapamycin (**Fig. 5b** and **Supplementary Fig. 5**), demonstrating the fast kinetics of a T7 RNAP-based biosensor. Next, to assess whether we could trigger protein production and to assay processivity of the evolved split RNAP, we set IRES-driven GFP mRNA (mRNA(GFP)) as the GOI on the rapa-T7 vector (rapa-T7-mRNA(GFP)). Again, the background fluorescence in cells transfected with rapa-T7-mRNA(GFP) was low, but addition of 10 nM rapamycin resulted in a strong enhancement in GFP fluorescence (**Fig. 5c** and **Supplementary Fig. 6**). Finally, to test dose responsiveness and whether ARs could trigger genetic changes to the cell, we tested whether RNA interference (RNAi) is a viable output. For this, we set small hairpin RNA (shRNA) targeting GFP as the GOI in the rapa-T7 vector (rapa-T7-shRNA(GFP)), cotransfected

cells with both a constitutive GFP expression vector and the rapa-T7-shRNA(GFP) vector, and analyzed GFP production by flow cytometry. Induction with rapamycin resulted in a dose-dependent knockdown of GFP signal (**Fig. 5d**). Collectively, these results demonstrate that ARs function in mammalian cells and can trigger a variety of outputs via the RNA signal.

DISCUSSION

Here we present the design, optimization, and deployment of a new split RNAP as a versatile biosensor platform. After discovering that fused PPIs can modulate the assembly of T7 RNAP split at position 179, we developed a PACE system to evolve an optimal proximity-dependent split RNAP. PACE yielded a variant with a 350-fold dynamic range based on fused leucine zipper peptide PPIs. We then explored the versatility of the split RNAP system by generating light- and small-molecule-activated RNAPs through swapping in different combinations of fusion proteins. Without any optimization, the resulting sensors had between 26- and >300-fold dynamic range. Moreover, we used a model trimolecular PPI system to demonstrate that multidimensional PPI networks can be monitored using orthogonal RNAP sensors with different RNA outputs. Finally, we demonstrated the versatility of RNAP-based sensors by showing that a small-molecule-driven RNA output can be used to synthesize fluorescent RNA nanostructures, generate proteins, or knock down genes in mammalian cells.

A challenge with traditional *n*-hybrid approaches is that the linker lengths, compositions, and geometries of the parts need to be carefully tuned for each new interaction to be interrogated, measured, selected, or evolved. In practice, this often means months of cloning, screening, and careful optimization, and for some targets, steric or geometric concerns preclude detection. With this in mind, we sought to create the ARs in a manner that is less dependent on these variables, which we postulated would streamline the process of engineering new sensors. From a first approximation, tuning the assembly of a split RNAP to be more dependent on fusion proteins would involve weakening the affinity of the two RNAP halves. However, such a thermodynamic design is likely to result in a high dependency on linkers and composition of the fusion proteins, which would make up the loss of binding energy. Therefore, we deliberately made the evolution of the ARs more challenging by tethering the fusion zipper peptides to the RNAP halves with long (6–8 amino acids), flexible linkers. In this way, the RNAP assembly process would be likely to evolve altered thermodynamics and kinetics of assembly, because the flexible linkers would not allow sufficient thermodynamic gain in binding energy between the fusion proteins. Further studies delineating the mechanism of controlled assembly of the evolved proximity-dependent RNAPs could test this hypothesis and improve engineering efforts. However, our results showing that the AR system worked well with diverse new binding partners suggest that it functions as we expected, whatever the evolved mechanism of RNAP assembly is. Moreover, our observation that linker length had little effect on RNAP activity suggests that thermodynamics is not the primary driving force.

Although PACE is a powerful directed evolution platform, a fundamental limitation of the technology, and all *in vivo* evolution systems, is the challenge of designing genetic circuits that link target activities to fitness[3]. ARs provide a new, robust mechanism to link target activities to gene expression and fitness for applications in directed evolution. It will be interesting to explore whether other target protein-interaction domains can replace the leucine zipper peptides for engineering by continuous evolution. Even more broadly, our system can also be adopted to other three-hybrid-like approaches, where different types of 'baits', such as small molecules or RNA, are displayed on the C-terminal RNAP half. Finally, because this new PACE system utilizes continuous and simultaneous positive and negative selection pressures, the approach should have advantages in terms of evolving selectivity in target proteins, as off-target interactions can be displayed on the orthogonal C-terminal RNAP and drive negative selection.

The AR system provides a new approach to monitor and respond to molecular interactions, molecules, and external cues in live cells. Combined with the ability to generate ARs with orthogonal DNA promoter specificity, these tools open the possibility of performing highly multidimensional interaction network analysis, presaging a new approach to cell analysis using high-throughput sequencing[23,24]. ARs should have advantages in terms of sensitivity owing to the signal amplification of the RNA output made possible by PCR. Unlike fluorescence-based technologies, which are solely for analysis, the RNA outputs of the ARs can be engineered to store information or drive cell fate changes on the basis of measured events. Engineering AR-based gene circuits that simultaneously measure and compute both pathologically relevant endogenous events and external cues provides a new strategy to develop 'smart' genetic therapies. Although we demonstrated PPI-, small-molecule-, and light-activated systems here, in principle any existing fluorescent protein-based probes could be immediately transported into the AR system for integration with synthetic biology applications. We anticipate that ARs will provided a simplified and more robust strategy for engineering gene circuits for applications in screening, directed evolution, and synthetic biology.

Received 2 June 2016; accepted 5 December 2016;
published online 13 February 2017

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

References

1. Ruder, W.C., Lu, T. & Collins, J.J. Synthetic biology moving into the clinic. *Science* **333**, 1248–1252 (2011).
2. Lienert, F., Lohmueller, J.J., Garg, A. & Silver, P.A. Synthetic biology in mammalian cells: next generation research tools and therapeutics. *Nat. Rev. Mol. Cell Biol.* **15**, 95–107 (2014).
3. Packer, M.S. & Liu, D.R. Methods for the directed evolution of proteins. *Nat. Rev. Genet.* **16**, 379–394 (2015).
4. Zhang, J., Jensen, M.K. & Keasling, J.D. Development of biosensors and their application in metabolic engineering. *Curr. Opin. Chem. Biol.* **28**, 1–8 (2015).
5. Benenson, Y. RNA-based computation in live cells. *Curr. Opin. Biotechnol.* **20**, 471–478 (2009).
6. Green, A.A., Silver, P.A., Collins, J.J. & Yin, P. Toehold switches: de-novo-designed regulators of gene expression. *Cell* **159**, 925–939 (2014).
7. Brophy, J.A. & Voigt, C.A. Principles of genetic circuit design. *Nat. Methods* **11**, 508–520 (2014).
8. Copeland, M.F., Politz, M.C., Johnson, C.B., Markley, A.L. & Pfleger, B.F. A transcription activator-like effector (TALE) induction system mediated by proteolysis. *Nat. Chem. Biol.* **12**, 254–260 (2016).
9. Putz, U., Skehel, P. & Kuhl, D. A tri-hybrid system for the analysis and detection of RNA--protein interactions. *Nucleic Acids Res.* **24**, 4838–4840 (1996).
10. Fields, S. & Song, O. A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245–246 (1989).
11. SenGupta, D.J. *et al.* A three-hybrid system to detect RNA-protein interactions in vivo. *Proc. Natl. Acad. Sci. USA* **93**, 8496–8501 (1996).
12. Baker, K. *et al.* Chemical complementation: a reaction-independent genetic assay for enzyme catalysis. *Proc. Natl. Acad. Sci. USA* **99**, 16537–16542 (2002).
13. Konermann, S. *et al.* Optical control of mammalian endogenous transcription and epigenetic states. *Nature* **500**, 472–476 (2013).
14. Polstein, L.R. & Gersbach, C.A. A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.* **11**, 198–200 (2015).
15. Nuñez, J.K., Harrington, L.B. & Doudna, J.A. Chemical and biophysical modulation of Cas9 for tunable genome engineering. *ACS Chem. Biol.* **11**, 681–688 (2016).
16. Martin, F. Fifteen years of the yeast three-hybrid system: RNA-protein interactions under investigation. *Methods* **58**, 367–375 (2012).

17. Church, G.M., Elowitz, M.B., Smolke, C.D., Voigt, C.A. & Weiss, R. Realizing the potential of synthetic biology. *Nat. Rev. Mol. Cell Biol.* **15**, 289–294 (2014).
18. Yen, L. *et al.* Exogenous control of mammalian gene expression through modulation of RNA self-cleavage. *Nature* **431**, 471–476 (2004).
19. Winkler, W.C. & Breaker, R.R. Regulation of bacterial gene expression by riboswitches. *Annu. Rev. Microbiol.* **59**, 487–517 (2005).
20. Culler, S.J., Hoff, K.G. & Smolke, C.D. Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins. *Science* **330**, 1251–1255 (2010).
21. Feng, J. *et al.* A general strategy to construct small molecule biosensors in eukaryotes. *eLife* **4**, e10606 (2015).
22. Lee, J.H. *et al.* Highly multiplexed subcellular RNA sequencing in situ. *Science* **343**, 1360–1363 (2014).
23. Zamft, B.M. *et al.* Measuring cation dependent DNA polymerase fidelity landscapes by deep sequencing. *PLoS One* **7**, e43876 (2012).
24. Pu, J., Chronis, I., Ahn, D. & Dickinson, B.C. A panel of protease-responsive RNA polymerases respond to biochemical signals by production of defined RNA outputs in live cells. *J. Am. Chem. Soc.* **137**, 15996–15999 (2015).
25. Segall-Shapiro, T.H., Meyer, A.J., Ellington, A.D., Sontag, E.D. & Voigt, C.A.A. A 'resource allocator' for transcription based on a highly fragmented T7 RNA polymerase. *Mol. Syst. Biol.* **10**, 742 (2014).
26. Shis, D.L. & Bennett, M.R. Library of synthetic transcriptional AND gates built with split T7 RNA polymerase mutants. *Proc. Natl. Acad. Sci. USA* **110**, 5028–5033 (2013).
27. Shekhawat, S.S. & Ghosh, I. Split-protein systems: beyond binary protein-protein interactions. *Curr. Opin. Chem. Biol.* **15**, 789–797 (2011).
28. Kerppola, T.K. Visualization of molecular interactions using bimolecular fluorescence complementation analysis: characteristics of protein fragment complementation. *Chem. Soc. Rev.* **38**, 2876–2886 (2009).
29. Hu, C.D. & Kerppola, T.K. Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat. Biotechnol.* **21**, 539–545 (2003).
30. Dickinson, B.C., Leconte, A.M., Allen, B., Esvelt, K.M. & Liu, D.R. Experimental interrogation of the path dependence and stochasticity of protein evolution using phage-assisted continuous evolution. *Proc. Natl. Acad. Sci. USA* **110**, 9007–9012 (2013).
31. Ellefson, J.W. *et al.* Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nat. Biotechnol.* **32**, 97–101 (2014).
32. Ghosh, I., Hamilton, A.D. & Regan, L. Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein. *J. Am. Chem. Soc.* **122**, 5658–5659 (2000).
33. Magliery, T.J. *et al.* Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. *J. Am. Chem. Soc.* **127**, 146–157 (2005).
34. Steitz, T.A. The structural changes of T7 RNA polymerase from transcription initiation to elongation. *Curr. Opin. Struct. Biol.* **19**, 683–690 (2009).
35. Esvelt, K.M., Carlson, J.C. & Liu, D.R. A system for the continuous directed evolution of biomolecules. *Nature* **472**, 499–503 (2011).
36. Badran, A.H. *et al.* Continuous evolution of *Bacillus thuringiensis* toxins overcomes insect resistance. *Nature* **533**, 58–63 (2016).
37. Carlson, J.C., Badran, A.H., Guggiana-Nilo, D.A. & Liu, D.R. Negative selection and stringency modulation in phage-assisted continuous evolution. *Nat. Chem. Biol.* **10**, 216–222 (2014).
38. Dickinson, B.C., Packer, M.S., Badran, A.H. & Liu, D.R. A system for the continuous directed evolution of proteases rapidly reveals drug-resistance mutations. *Nat. Commun.* **5**, 5352 (2014).
39. Hubbard, B.P. *et al.* Continuous directed evolution of DNA-binding proteins to improve TALEN specificity. *Nat. Methods* **12**, 939–942 (2015).
40. Leconte, A.M. *et al.* A population-based experimental model for protein evolution: effects of mutation rate and selection stringency on evolutionary outcomes. *Biochemistry* **52**, 1490–1499 (2013).
41. Cheetham, G.M. & Steitz, T.A. Structure of a transcribing T7 RNA polymerase initiation complex. *Science* **286**, 2305–2309 (1999).
42. Tahirov, T.H. *et al.* Structure of a T7 RNA polymerase elongation complex at 2.9 A resolution. *Nature* **420**, 43–50 (2002).
43. Guntas, G. *et al.* Engineering an improved light-induced dimer (iLID) for controlling the localization and activity of signaling proteins. *Proc. Natl. Acad. Sci. USA* **112**, 112–117 (2015).
44. Rivera, V.M. *et al.* A humanized system for pharmacologic control of gene expression. *Nat. Med.* **2**, 1028–1032 (1996).
45. Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. USA* **89**, 5547–5551 (1992).
46. Paulmurugan, R. *et al.* A novel estrogen receptor intramolecular folding-based titratable transgene expression system. *Mol. Ther.* **17**, 1703–1711 (2009).
47. Blakeley, B.D., Chapman, A.M. & McNaughton, B.R. Split-superpositive GFP reassembly is a fast, efficient, and robust method for detecting protein-protein interactions in vivo. *Mol. Biosyst.* **8**, 2036–2040 (2012).
48. Kerppola, T.K. Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. *Annu. Rev. Biophys.* **37**, 465–487 (2008).
49. Filonov, G.S., Kam, C.W., Song, W. & Jaffrey, S.R. In-gel imaging of RNA processing using broccoli reveals optimal aptamer expression strategies. *Chem. Biol.* **22**, 649–660 (2015).
50. Ringquist, S. *et al.* Translation initiation in *Escherichia coli*: sequences within the ribosome-binding site. *Mol. Microbiol.* **6**, 1219–1229 (1992).

Acknowledgments

We thank I. Chronis (University of Chicago), Y. Koh (University of Chicago), and D. Ahn (University of Chicago) for technical assistance, and D. Liu (Harvard University), B. McNaughton (Colorado State University), A. Deiters (University of Pittsburgh), B. Glick (University of Chicago), M. Glotzer (University of Chicago), Y. Krishnan (University of Chicago), J. Thornton (University of Chicago), and Y. Weizmann (University of Chicago) for supplying equipment and materials. This work was supported by the University of Chicago, the Cancer Research Foundation, the National Institute of General Medical Sciences of the National Institutes of Health (R35 GM119840) to B.C.D., the University of Chicago Medicine Comprehensive Cancer Center (P30 CA14599), and the National Center for Advancing Translational Sciences of the National Institutes of Health (UL1 TR000430). J.Z.-B. was supported by a Chemical Biology Training Grant from the US National Institutes of Health (T32GM008720).

Author contributions

J.Z.-B., J.P., and B.C.D. cloned and validated all materials and performed transcription assays. J.P. and B.C.D. performed PACE experiments. J.Z.-B. and J.P. performed cell culture experiments. J.Z.-B., J.P., and B.C.D. designed experimental strategies and wrote the paper.

Competing financial interests

The authors declare competing financial interests: details accompany the online version of the paper.

Additional information

Any supplementary information, chemical compound information and source data are available in the online version of the paper. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to B.C.D.

ONLINE METHODS

Cloning. All plasmids were constructed by Gibson Assembly[51] from PCR products generated using Q5 Hot Start DNA Polymerase (NEB) or Phusion Polymerase. Phage were cloned by Gibson Assembly of the split N-terminal RNAP-ZA fusion into a previously optimized SP phage backbone[38] and transformed into 1059 cells[37], which supply *gIII* in an activity-independent manner. After overnight growth in medium, the supernatant was isolated by centrifugation and plaque assays were performed on 1059 cells. Single plaques were selected for overnight growth and sequencing to identify clonal phage samples with the correct insert. All plasmids and phage were sequenced at the University of Chicago Comprehensive Cancer Center DNA Sequencing and Genotyping Facility. All new vectors are described in **Supplementary Table 1** and maps for each plasmid are shown in **Supplementary Figure 1**.

Sequence of split RNAP fusions. Listed below are the structures and sequences of the leucine zipper peptide fusions, iLID-nano light-induced dimerization fusions, and rapamycin-induced dimerization constructs used in this study. Linkers are enclosed in square brackets. The three point mutations between ZB and ZBneg are underlined and were obtained from previous studies[33], and the two mutations in ZA that evolved during PACE are also underlined.

N-ZA: RNAP(1-179)–[GGSGSGSS]–ALKKELQANKKELAQLKWELQALKKELAQ

ZB-C: MASEQLEKKLQALEKKLAQLEWKNQALEKKLAQ–[TSGGSG]–RNAP(180+)

ZBneg-C: MASEQLEKELQALEKELAQLKWKNQALEKKLAQ–[TSGGSG]–RNAP(180+)

N-ZA-(L13I,L20I): N-28-1(1-179)–[GGSGSGSS]–ALKKELQANKKEIAQLKWEIQALKKELAQ

N-term-iLID: N-28-1(1-179)–[GGSGSGSS]–iLID

SspB-C-term: SspB nano–[TSGGSG]–RNAP(180+)

N-term-FRB: N-28-1(1-179)–[GGSGSGSS]–FRB

FKBP-C-term: FKBP–[TSGGSG]–RNAP(180+)

N-ZA-FRB: RNAP(1-179)–[GGSGSGSS]–ZA (L13I,L20I)-[GGSAGGSG]-FRB.

***In vivo* transcription assays of split RNAPs.** N- and C-terminal halves of the RNAP were cloned into expression vectors, with the N-terminal RNAP under a constitutive promoter and the C-terminal RNAP under the arabinose-inducible promoter. S1030 cells[37] were transformed by electroporation with three plasmids: (i) an N-terminal RNAP expression plasmid, (ii) a C-terminal expression plasmid, and (iii) a reporter plasmid that encodes luciferase under control of the T7 promoter. The transformed cells were then plated onto agar plates (15 g/L in LB) with 50 μg/mL carbenicillin, 50 μg/mL spectinomycin, 33 μg/mL chloramphenicol, and 10 mM glucose. Single colonies were grown to saturation overnight at 37 °C, and each well of a 96-well deep-well plate containing 0.54 mL of LB with antibiotics and 10 mM arabinose was inoculated with 60 μL of the overnight culture. After growth with shaking at 37 °C for 2 h, 150 μL of each culture was transferred to a 96-well black-wall, clear-bottom plate (Costar), and luminescence and OD_{600} was measured on a Synergy H4 Hybrid Reader (BioTek). The data were analyzed by dividing the luminescence values by the background-corrected OD_{600} value, then subtracting out the background from the reporter vector alone. All values were then normalized to the wild-type split RNAP fused to ZA and ZB (**Fig. 1c**), which was assigned an arbitrary value of 100, allowing the values from each luminescence plot to be compared to one another. For the light-activated system, the experiment was performed identically except upon outgrowth for 3 h, cells in the light condition were illuminated with a homemade blue LED lightbox and cultured at 25 °C, and cells in the dark condition were cultured at 37 °C; this was done to correct for heat output from the light source to maintain similar temperatures in both conditions. For the rapamycin-inducible system, the experiment was performed identically, except upon outgrowth, rapamycin was added for 3 h, then luminescence was analyzed. Sample size (n = 4 biological replicates for each condition) was determined on the basis of previous work using a similar *in vivo* luciferase reporter system and provided excellent reproducibility both between biological replicates on a given day and between days of experimental replicates.

Phage-assisted continuous evolution (PACE). PACE was carried out to evolve the split T7 RNAP variants using a modified version of previously described methods[24]. *E. coli* strain S1030 was transformed by electroporation with combinations of the positive accessory plasmid (posAP), negative accessory plasmid (negAP), and mutagenesis plasmid (MP)[38]. 5-mL starter cultures were grown overnight in LB supplemented with antibiotics and 10 mM glucose. Chemostats (100 mL sterile bottles) containing 80 mL of Davis rich medium[37] were inoculated with 2 mL starter culture and grown at 37 °C with magnetic stir-bar agitation. At approximately OD_{600} 1.0, fresh Davis rich medium was pumped in at 60–80 mL h^{-1}, with a waste needle set at 80 mL. 10 μL phage was used to seed a fresh lagoon (25 mL flask with a rubber septum). To initiate the evolution, a monoclonal phage population was used. Waste needles were set to maintain the lagoon volume at 15–20 mL, and host cell cultures were flowed in at 15–17 mL h^{-1}. Arabinose (10% w/v in water) was added directly to lagoons via syringe pump at 1.0 mL h^{-1} to induce mutagenesis. A lagoon sample was taken from the waste withdrawal line every 24 h and centrifuged, and the supernatant was stored at 4 °C. The complete evolutionary protocol is described in **Supplementary Table 2**. After the completion of each leg of the evolution, activity-dependent plaque assays were used to select the next evolutionary target, and the PACE experiment was again initiated as described, using 10 μL phage from the previous endpoint of the evolution. The strength of the positive and negative selection pressures were varied by altering the ribosome binding sites (RBSs)[50] controlling the expression of each of the C-terminal target fusions, *gIII* and a dominant-negative form of *gIII* (*gIIIneg*). Mixed selection pressures from multiple posAP and negAP sets at a given time point were used as appropriate to enhance the likelihood of successful evolution[24,30,40].

Sequence and activity analysis of variants from PACE. Phage samples were boiled for 10 min to lyse the phage and release the genomes. PCR was then used to amplify the DNA library containing the N-terminal RNAP variants, which was then subcloned into vector p3-7. Single colonies were picked from the transformation and subjected to analysis by Sanger sequencing. The results of the sequence analysis during the course of the evolution are shown in **Supplementary Table 3**. Variants N-29-1 and N-29-8 cloned into vector p3-7 were subjected to analysis by the luciferase assays as shown in **Figure 2e**. In order to sequence potential mutations that occurred in the peptide fusion in the phage, single plaques from an activity-independent plaque assay were picked, grown overnight, boiled, and analyzed by Sanger sequencing.

***In vivo* split GFP assay.** S1030 cells were transformed by electroporation with two plasmids: (i) a constitutive N-terminal GFP–FRB fusion and (ii) an arabinose inducible FKBP–C-terminal GFP fusion. The transformed cells were then plated onto agar plates (15 g/L in LB) with 50 μg/mL spectinomycin, 33 μg/mL chloramphenicol, and 10 mM glucose. Single colonies were then grown to saturation overnight at 37 °C, and then each well of a 96-well deep-well plate containing 0.54 mL of LB with antibiotics, 10 mM glucose or 10 mM arabinose, and varying concentrations of rapamycin (0 nM, 1 nM, 10 nM, 0.1 μM, 1 μM, 20 μM, or 100 μM) were inoculated with 60 μL overnight culture. After growth with shaking at 37 °C for 3 h, 6 h, or 30 h, 150 μL culture was transferred to a 96-well deep-well plate. The cultures were centrifuged (10 min, 25 °C, 2,000 RCF) and washed with 1 mL PBS three times before suspension in 150 μL PBS and transfer to a 96-well black-wall, clear-bottom plate (Costar). GFP fluorescence (excitation 485 ± 20 nm, emission 516 ± 20 nm) and OD_{600} were measured on a Synergy Neo2 Microplate Reader (BioTek). The data were analyzed by dividing the background-corrected GFP fluorescence values by the background-corrected OD_{600} value. All values were then normalized to the 0 nM rapamycin conditions. (sample size n = 5 biological replicates for each condition).

Dual reporter PPI *in vivo* detection assays. FRB and ZA peptide were fused to the N-terminal RNAP (N-29-1) in an *E. coli* expression vector to form pJin200. FKBP was fused to the C-terminal of T7 RNAP and ZB or ZBneg was fused to C-terminal of CGG RNAP in one vector to construct the dual C-terminal vectors. Then the CGG promoter–driven DsRed-Express2 circuit was added to p2-22, which contained a T7 promoter–driven luciferase gene to generate pJin216 as the dual reporter vector. To test the multidimensional system,

doi:10.1038/nchembio.2299

pJin200, pJin216, and either pJin207 or pJin208 were transformed into S1030 cells. The transformed cells were then plated onto agar plates (15 g/L in LB) with 50 μg/mL carbenicillin, 50 μg/mL spectinomycin, 33 μg/mL chloramphenicol, and 10 mM glucose. Single colonies were cultured overnight in LB liquid medium with the same antibiotics and 10 mM glucose. The next day, 50 μl overnight cultures were transferred to a deep 96-well plate containing LB with same antibiotics, 10 mM arabinose with 0 or 20 μM rapamycin. After shaking at 37 °C for 5 h, 150 μL culture was transferred to a 96-well deep-well plate. The samples were centrifuged, washed and prepared in the same way as for the split GFP assay. Then the OD_{600}, luminescence and DsRed-Express2 fluorescence (excitation 555 ± 15 nm, emission 590 ± 15 nm) were measured on a Synergy Neo2 Microplate Reader (BioTek). The data were analyzed by dividing the background-corrected luminescence values and DsRed-Express2 fluorescence background-corrected values by the background-corrected OD_{600} value. All luminescence values were normalized to the 0 μM rapamycin conditions with the ZB-fused C-terminal variant, and DsRed-Express2 fluorescence values were normalized to 0 μM rapamycin conditions with the ZBneg-fused C-terminal variant.

Cell culture. HEK293T cells (ATCC) were maintained in DMEM (high glucose, L-glutamine, phenol red, sodium pyruvate; obtained from Gibco (11995-065) or Hyclone (SH30081.01)) supplemented with 10% FBS (Gibco/Life Technologies, qualified US origin) and 1% penecillin–streptomycin (P/S, Gibco/Life Technologies). As HEK293T cells are listed in the database of commonly misidentified cell lines maintained by ICLAC (http://iclac.org/databases/cross-contaminations/), we obtained fresh cells from ATCC, which were frozen down at an early passage (passage 5) in individual aliquots. The cells were then used for <25 passages for all experiments. Multiple biological replicates were performed with cells from different passages and freshly thawed aliquots. There was no testing for mycoplasma infection or further authentication because early-passage cells were used for all experiments.

Imaging mammalian AR activation by fluorescence microscopy. HEK293T cells cultured in DMEM (high glucose, L-glutamine, phenol red, pyruvate; Gibco/Life Technologies) supplemented with 10% FBS (Gibco/Life Technologies, qualified US origin) were plated on an 8-well coverglass slide (Labtek) and transfected with 600 ng rapa-T7 vector (pJin141 or p6-8) using 1.5 μL of Lipofectamine 2000 (Thermo Fisher Scientific) using the standard protocol. For the rapa-T7-F30-2xdBroccoli (pJin141) experiments, 100 nM rapamycin or DMSO control was added along with 20 μM DFHBI-1T for 30 min before imaging. For the rapa-T7-mRNA(GFP) (p6-8) experiments, 10 nM rapamycin or DMSO control was added to the sample 20 h after transfection, and then incubated for an additional 24 h. The cells were imaged on an Olympus BX53 microscope using a GFP filter set and a ×10 objective. Each image for a given condition was processed using identical conditions to adjust brightness and contrast to a level where background fluorescence was observed for control samples in ImageJ (NIH).

Flow cytometry. HEK293T cells were cultured in DMEM (high glucose, L-glutamine, phenol red, pyruvate; Gibco/Life Technologies) supplemented with 10% FBS (Gibco/Life Technologies). The day before transfection, cells were passaged and plated at ~50,000 cells per well in a 48-well plate (NEST Biotechnology). After 19 h, 50 ng RFP plasmid (p3-62), 200 ng GFP plasmid (p1-53), and 400 ng rapa-T7-shRNA(GFP) vector (pJin140) were transfected into cells using 1.5 μL Lipofectamine 2000 (Thermo Fisher Scientific) using the standard protocol. 30 min after transfection, DMEM supplemented with FBS and either DMSO or rapamycin was added to the wells so the final rapamycin concentration was 0 nM, 0.001 nM, 0.01 nM, 0.1 nM, 1 nM, or 10 nM. 29 h after transfection, the medium was replaced with fresh medium including the previously added concentration of rapamycin. 44 h after transfection, the cells were trypsinized and suspended in DMEM supplemented with FBS and rapamycin, then analyzed on a LSR-Fortessa 4-15 (BD digital instrument, 488 nM laser, 530 ± 30 nM filter for GFP, and 610 ± 20 nM filter for RFP). Mean GFP fluorescence intensity was calculated for HEK293T cells expressing RFP using FlowJo. Reported values are average mean GFP fluorescence intensity values taken from 3 separate replicate samples. Sample size (n = 3 biological replicates for each condition) was determined by initial trial experiments to find the spread in the data.

Data availability. All data generated or analyzed during this study are included in the published article (and its supplementary information) or are available from the corresponding author on reasonable request.

51. Gibson, D.G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).