HHS Public Access

Author manuscript

Methods Enzymol. Author manuscript; available in PMC 2023 August 23.

Published in final edited form as:

Methods Enzymol. 2022 ; 675: 1–32. doi:10.1016/bs.mie.2022.07.002.

Engineering protein and DNA tools for creating DNA-dependent protein switches

Harsimranjit Sekhon,

Jeung-Hoi Ha,

Stewart N. Loh*

Department of Biochemistry and Molecular Biology, State University of New York Upstate Medical University, Syracuse, NY, United States

Abstract

Switchable proteins are capable of changing conformations from inactive (OFF) to active (ON) forms in response to inputs such as ligand binding, pH or temperature change, or light absorption. A particularly powerful class of protein switches, exemplified by the Cas nucleases of CRISPR systems, are activated by binding of specific DNA or RNA sequences. The mechanism by which oligonucleotide binding regulates biological activity is complex and highly specialized in the case of Cas enzymes, but recent advancements in protein and DNA engineering have made it possible to introduce this mode of control into other enzymes. This chapter highlights recent examples of protein switches that combine these two fields of engineering for the purpose of creating biosensors that detect pathogen and other genomic sequences. One protein engineering method—alternate frame folding—has the potential to convert many proteins into ligand-activated switches by inserting a binding protein (input domain) into an enzyme (output domain). The steps for doing so are illustrated using GCN4 as a DNA recognition domain and nanoluciferase as a luminescent reporter that changes color as a result of DNA binding. DNA engineering protocols are included for creating DNA tools (*de novo* designed hairpins and modified aptamers), that enable the biosensor to be activated by arbitrary DNA/RNA sequences and small molecules/proteins, respectively. These methodologies can be applied to other proteins to gain control of their functions by DNA binding.

1. Introduction

A major goal of protein engineering research is to develop mechanisms by which protein activity can be switched on and off by a stimulus. A general approach is to fuse an input domain (which is responsible for recognizing the triggering signal) to an output domain (which establishes the biological response of the switch), in such a way that reception of the stimulus by the former induces a functional change in the latter (Fig. 1). The input domain is often a protein that binds a target ligand (e.g., nucleic acids, proteins, small molecules) with high affinity and selectivity or absorbs light. Output domains include enzymes (Karchin, Ha, Namitz, Cosgrove, & Loh, 2017; Nicholes et al., 2016; Wu et al., 2009; Zetsche, Volz,

*Corresponding author: lohs@upstate.edu.

& Zhang, 2015), binding proteins (Carrasco-López et al., 2020; Gil et al., 2020; Karchin et al., 2017), fluorescent proteins (Baird, Zacharias, & Tsien, 1999; John, Sekhon, Ha, & Loh, 2022; Nasu, Shen, Kramer, & Campbell, 2021), and luminescent proteins (Chang, Kim, Lindberg, & Winssinger, 2020; Engelen, Meijer, Saha, B., & Merkx, 2017; Quijano-Rubio et al., 2021; Sekhon & Loh, 2022). The modular nature of the design conceptually allows for mixing and matching of input/output domains to achieve combinatorial results. In practice, however, this is difficult to achieve because switching often involves coupled changes in protein conformation, and the mechanism developed for one input/output pair seldom translates to another.

In this chapter, we focus on one class of engineered switches—biosensors that detect specific nucleotide sequences (e.g., pathogen DNA/RNA and miRNA)—that can be adapted with relative ease to recognize new targets without redesign of conformational change mechanisms or extensive modification to the protein components. The sensors consist of a recognition domain that binds a DNA or RNA sequence joined to a luminescent or fluorescent protein to provide optical readout. Interest in these molecules has surged in recent years due to the COVID-19 pandemic, which has exposed the need for testing technologies that complement RT-PCR and can be employed in low-resource and point-of-care settings.

For the application of detecting custom DNA and RNA sequences, biosensor modularity can be achieved especially efficiently. If the input domain is an oligonucleotide, then modifying the sensor to bind a new target sequence simply entails changing the recognition oligonucleotide to its complement. Several of these detection platforms exist, including those based on CRISPR associated (Cas) enzymes (Jolany Vangah et al., 2020; Kaminski, Abudayyeh, Gootenberg, Zhang, & Collins, 2021). A very recent design employs a protein input domain that binds a fixed DNA sequence, and changing specificity involves designing short DNA hairpins to convert the target sequence to the fixed sequence (Sekhon & Loh, 2022). The common denominator is that the modifications are made only to the DNA components of the sensors—a process that's greatly facilitated by desktop computational tools that predict DNA structure and stability with high accuracy.

Here, we review recent protein switch designs in which DNA or RNA binding to the input domain turns on biological activities of various output domains. The main application is currently biosensing and for this purpose, luminescent or fluorescent output domains provide convenient readout. One of the sensing methods (alternate frame folding), however, can be applied to other proteins and enzymes to convert them to DNA-activated switches, and we outline the protein engineering steps for doing so. We provide protocols for designing DNA tools to convert an arbitrary DNA/RNA sequence (via short hairpin DNAs) and small molecule/protein (via modification of existing DNA aptamers) input signals to the DNA sequence that activates the switch.

1.1 Cas enzymes: Nature's RNA-activated switches

Many engineering efforts are inspired by natural proteins, and in terms of useful switches, few are as influential as the Cas nucleases. Their deceptively simple input/output functionality belies their transformative roles in gene editing and biosensing. As

biosensors, the best-known examples are Cas12 and Cas13 (Das, Goswami, Whyms, Sridhara, & Li, 2022; Kaminski et al., 2021; Nalefski et al., 2021; O'Connell, 2019). Both are nucleases that become activated in the presence of a guide RNA (gRNA), thus providing enzymatic amplification of the initial binding event. The complex of gRNA bound to its genomic complement activates Cas12 and Cas13 to cleave target DNA and RNA, respectively. Remarkably, following this primary cleavage event, these enzymes gain the ability to hydrolyze RNA (Cas13) or single-stranded DNA (ssDNA) and RNA (Cas12) in a non-sequence specific manner. This "collateral" nuclease activity forms the basis of diagnostic platforms such as SHERLOCK (Gootenberg et al., 2017) and DETECTR (Chen et al., 2018). Readout is established by synthetic oligonucleotide substrates derivatized with various functional groups (Kaminski et al., 2021). The simplest of these methods entails labeling the substrate with a fluorescent dye and a quencher. Cleavage produces a fluorescent turn-on as the two dissociate. The lateral flow readout system employs an oligonucleotide tagged with biotin at one end and a reporter group at the other. It is loaded onto a chromatographic strip containing streptavidin in one zone and an antibody against the reporter group in a second, downstream zone. When the strip is exposed to a mobile phase, the intact substrate accumulates in zone 1, whereas the cleaved product migrates beyond zone 1 and collects in zone 2 (Myhrvold et al., 2018).

The Cas enzymes illustrate the axiom that it is easier to engineer nucleic acids than it is to engineer proteins. Cas12 and Cas13 are large, multidomain enzymes (over 1100 amino acids), and although numerous structures have been determined by cryoEM and X-ray crystallography, the molecular basis of catalytic activation by oligonucleotide binding remains poorly understood. It is nonetheless evident that this mechanism is complex, highly specialized, and not readily transferrable to other enzymes. It has proven much more tractable to adapt Cas whereby it cleaves a desired site in the genome or reports on a presence of a specific pathogen DNA/RNA sequence. This is achieved by placing the desired antisense sequence in a variable region of the gRNA and not altering the constant sequence that is responsible for binding Cas.

1.2 Semisynthetic luciferase-based biosensors

For the remainder of this chapter, we focus on examples of proteins that have been engineered for activation by DNA binding, primarily for the purpose of biosensing. The semisynthetic class of sensors requires two components. Component 1 is a protein-nucleic acid hybrid molecule composed of a luminescent protein for readout (nanoluciferase, or nLuc; Box 1) chemically derivatized with a single-stranded DNA or peptide nucleic acid (PNA) in vitro. The purpose of the nucleic acid is to anchor component 1 to component 2 by means of base pairing, and more importantly, to establish the conformational change on target binding that changes luminescence of the nLuc domain. Component 2 is a synthetic ssDNA that binds to component 1 and contains a chemical compound that modulates nLuc luminescence by direct inhibition or bioluminescence resonance energy transfer (BRET; Box 1). Like Cas, nLuc is an enzyme and collecting one of its products (light) provides signal amplification. Luminescent sensors generally offer superior economics and sensitivity compared to their fluorescent counterparts because no excitation source is needed and there

is consequently no interfering light, which permits very low levels of luciferase activity to be detected using common devices such as cell phone cameras.

The protein-DNA component of the “BRET-beacon” sensor of Merkx and coworkers was nLuc to which a 21-mer ssDNA handle was attached (Fig. 2A) (Engelen et al., 2017). Component 2 was an ssDNA consisting of the anti-handle followed by a stem-loop structure (in which the target-binding sequence was embedded in the loop) and the fluorescent dye Cy3. The function of the stem-loop was to bring the Cy3 acceptor close to nLuc donor, which resulted in efficient BRET (orange emission at 578nm). Hybridization of the cognate ssDNA converted the stem-loop structure into a long double helix, lengthening the distance between Cy3 and nLuc and resulting in up to a 13-fold increase in blue-to-orange emission ratio.

Changing the specificity of the BRET-beacon sensor is simply a matter of altering the length and sequence of the loop in component 2 to complement that of the desired target and annealing that oligonucleotide to the DNA-nLuc hybrid molecule, without any modification to the latter. Moreover, by varying the length (and therefore stability) of the stem region the authors were able to tune the apparent affinity by 10–20-fold without altering sequence specificity.

Winssinger and colleagues recently added a twist to the above design intended to increase the intensity of the turn-on response at the expense of ratiometric output (Chang et al., 2020). They employed Halo-Tag technology to conjugate a PNA bearing the target recognition sequence to a Halo-Tagged nLuc construct (H-Luc; Hiblot et al., 2017) (Fig. 2B). Component 2 was a slightly shorter oligonucleotide that complemented most of the PNA sequence and contained a competitive nLuc inhibitor (Walker et al., 2017) which suppressed 97% of nLuc activity when component 2 was bound. Key to this design was the single-stranded overhang in the PNA sequence. This “toehold” region allowed the target DNA or RNA strand to invade the PNA/DNA duplex and oust component 2 by a mechanism known as toehold-mediated strand displacement (TMSD). This exchange reaction produced a 32-fold increase in blue light intensity. As in the BRET-beacon design, target specificity of this “inhibitor-displacement” sensor is easily changed by altering the sequence of the PNA strand and its complement, with the additional consideration noted below.

The ON/OFF ratio of TMSD-based switches is largely determined by the metastability of the TMSD duplex. The double-stranded region must be of an appropriate length and composition such that it affords a high degree of specificity and does not spontaneously dissociate, but not so stable as to resist displacement by the invader strand. The affinity of the inhibitor for nLuc factors into this calculation in the inhibitor-displacement sensor, but in switches that employ DNA-only TMSD complexes, metastability can be accurately designed by DNA software tools such as NUPACK (Zadeh et al., 2011), as discussed in Section 2.

In addition to functioning as luminescent biosensors, hybrid protein-nucleic acid molecules have been used to achieve enzymatic turn-on by directed localization. Here, the purpose of the nucleic acid components is to recognize DNA sequences that are present naturally or arrayed in engineered, self-assembling origami structures such as tubes and cages (Flory et

al., 2013; Rosier et al., 2020; Wijnands, Engelen, Lafleur, Meijer, & Merkx, 2018; Zhao et al., 2016). Protein activity is switched on or off by, for example, bringing together an enzyme and its cofactor or inhibitor (Gianneschi & Ghadiri, 2007; Liu et al., 2013; Saghatelian, Guckian, Thayer, & Ghadiri, 2003), two halves of a split enzyme (Furman et al., 2010; Sancho Oltra, Bos, & Roelfes, 2010), and two or more enzymes for substrate channeling (Fu et al., 2014). These approaches have been recently reviewed elsewhere (Madsen & Gothelf, 2019; Watson, Angerani, Sabale, & Winssinger, 2021).

1.3 Protein-only DNA-activated switches

Full genetic encoding of molecular switches facilitates their application in cells and eliminates the chemical coupling steps required by semisynthetic methods. Compared to protein-DNA hybrid sensors, protein-only switches present a fundamental design challenge as well as offer a substantial reward. In the semisynthetic sensors discussed above, nLuc is never actually switched off. Instead, a structural change in their DNA domains alters the distance between nLuc and a BRET acceptor or displaces an inhibitor, without affecting the structure or inherent function of the nLuc domain. These structural changes are straightforward to engineer when DNA is the recognition domain. By contrast, protein-only switches generally achieve turn-on/turn-off by coupled conformational changes, and as witnessed by the Cas-based sensors, allosteric mechanisms tend to be complex and idiosyncratic to a specific system. Development of simpler and more general mechanisms will open the functionality of the proteome to regulation by specific DNA and RNA sequences.

We recently applied a protein engineering approach (alternate frame folding, or AFF; Ha, Shinsky, & Loh, 2013; John et al., 2022; Stratton & Loh, 2011; Stratton, Mitrea, & Loh, 2008) to couple DNA binding to nLuc activation using protein conformational change (Sekhon & Loh, 2022). The AFF-mediated conformational change causes one set of residues in nLuc to be replaced by another set in the same molecule, extinguishing luminescence of the first form of nLuc and turning on luminescence of the second. To enable this transformation, we copied residues 6–50 of nLuc (designated 6*–50*) and appended it to its C-terminus using a short linker peptide (Fig. 3A). This duplication establishes two mutually-exclusive folding “frames”: the native (N) frame (residues 6–171) and a circularly permuted (CP) frame (residues 51–171 linked to residues 6*–50*). nLuc can adopt the N-fold or the CP-fold but not both. To report on the N ↔ CP fold shift, we employed the green enhanced nanolantern (GeNL) variant of nLuc (Suzuki et al., 2016). GeNL consists of the fluorescent protein mNeonGreen (mNG) fused to the N-terminus of nLuc, with deletion of nLuc residues 1–5 to increase BRET efficiency. The N-to-CP fold shift changes luminescence from green to blue according to the mechanism described below.

To trigger the N-to-CP fold shift by DNA binding, we inserted the 56-aa GCN4 binding domain into the surface loop between residues 50–51 of the N-frame (Fig. 3A). The resulting biosensor was designated nLuc-AFF. GCN4 consists of a C-terminal coiled-coil dimerization region and an N-terminal DNA binding region that is unstructured in the absence of the consensus 11-mer dsDNA sequence (AP-1). Thus, without AP-1, the GCN4 domain existed as a partially disordered, flexible surface loop that was well tolerated by

the N-fold. Luminescence of the N-fold was almost entirely green due to efficient BRET with mNG that abutted its N-terminus (Suzuki et al., 2016). Upon AP-1 binding, GCN4 folds into a ~75Å long helical rod (O'Shea, Klemm, Kim, & Alber, 1991) that stretches and unfolds the N-frame, forcing the nLuc domain to refold in the CP-frame (Fig. 3B). In the CP-fold, nLuc is separated from mNG by 118 aa and >75Å, which eliminates BRET and reverts luminescence to blue. The 6.3-fold increase in blue/green luminescence on AP-1 binding was quantified by cell phone camera and was evident by eye (Sekhon & Loh, 2022). We note that AFF switches can be created either by duplicating an N-terminal segment and linking it to the C-terminus (as we did to generate nLuc-AFF) or by duplicating a C-terminal segment and linking it to the N-terminus. Using Fig. 3A as an example, this latter protein would consist of mNG-(nLuc 51*–171*)-linker-(nLuc 6–50)-GCN4-(nLuc 51–171).

Synthetic ssDNA input domains are easily customized to bind sequences of choice. By contrast, proteinaceous DNA-binding domains such as GCN4 recognize only a single consensus sequence (with some exceptions such as TALE domains (de Lange et al., 2014; Morbitzer, Römer, Boch, & Lahaye, 2010; Rogers et al., 2015) and zinc finger domains (Klug, 2010; Liu, Segal, Ghiara, & Barbas, 1997; Moore, Klug, & Choo, 2001)). DNA engineering can be employed to transform different input signals (arbitrary DNA and RNA sequences, small molecules, and proteins) into the specific DNA sequence that activates the proteinaceous biosensor (Fig. 4). The principle is to hide the activating sequence in a DNA hairpin (Fig. 4A) or modified aptamer (Fig. 4B) where it is cryptic and unable to interact with the biosensor. Invasion of the target ssDNA or ssRNA strand into the hairpin (via TMSD), or binding of the small molecule/protein analyte to the aptamer, opens the hairpin and exposes the activating sequence for binding the biosensor. In Fig. 4B, the engineered aptamer undergoes a ligand-induced shift from the hairpin fold to the aptamer fold, which is essentially the DNA version of the AFF mechanism. Workflows for creating these DNA tools are presented in Section 3.

2. Design and cloning of AFF biosensors

The general approach for constructing a protein switch based on the AFF mechanism consists of three main steps. Step I is to select sites in the output domain (e.g., between residues 50–51 in nLuc (Fig. 3A)) that are likely to be compatible with both circular permutation and insertion of the input domain. Step II is to verify that the input and output domains retain their functions in the AFF construct. Step III to tune the thermodynamic stabilities of the two frames such that the N-fold (with the input domain inserted) is slightly favored over the CP-fold in the absence of ligand. This latter condition ensures that binding of the ligand at reasonable concentrations inverts the populations of N- and CP-folds, producing a robust output signal. The process can be significantly accelerated if a viable split point or CP site in the output protein is already known. If not, steps I and II can involve creating a small library of AFF constructs (<10 genes) consisting of the input gene inserted into candidate sites of the output gene (with the appropriate sequence duplicated and linked to either the 5′- or 3′-end). The expressed proteins can then be screened in cell lysates or in semi-purified condition for reasonable activity and solubility. The following protocols describe how to construct and optimize nLuc-AFF, but they are meant to provide guidance

for inserting GCN4 (and other DNA binding domains) into proteins other than nLuc, to create new DNA-activated protein switches.

2.1 General protocol for designing an AFF switch

1. To identify potential sites for insertion and circular permutation, analyze the structure of the output protein to identify solvent-exposed turns and loops containing 2–8 residues. The structure can be experimentally determined or predicted by algorithms such as AlphaFold (Jumper et al., 2021).
2. Measure the C_α-C_α distance between the two residues that define the boundaries of the loop and reject any loop in which this distance is equal to or greater than the N-C distance of the input domain (75Å for GCN4). There are ~11 potentially viable loops/turns in nLuc (PDB ID 5IBO). Reject loops that are between Cys residues involved in disulfide bonds.
3. Prioritize the loops for gene construction. Use an online computational tool such as SPELL (Dagliyan et al., 2018) or CPred (Lo et al., 2012) that predict viable sites for splitting and circularly permuting the output protein, respectively. These locations may be expected to tolerate insertions as well. Theory-based methods do not accurately predict effects on stability and especially solubility. The semi-rational approach of step 1 is therefore needed, but the computational data serve to identify the most promising sites. De-prioritize loops that are close to active sites.
4. In favorable cases, optimal sites(s) for insertion and permutation can be determined experimentally by NMR and/or mass spectrometry (Ha, Karchin, Walker-Kopp, Castañeda, & Loh, 2015; Ha, Presti, & Loh, 2019). However, steps 1–3 are generally more time efficient.
5. Link the input and output domains using 1–2 Gly residues at both junctions, provided that this does not make the C_α-C_α distance of the loop exceed the N-C length of the input domain (each Gly adds ~3Å to the former distance; Cutler, Mills, Lubin, Chong, & Loh, 2009). These linkages make the connection more flexible and help both domains retain their functions.
6. For tuning the stabilities of the two folds, keep the stability of the N-frame constant and change the stability of the CP-frame by varying the length of the linker peptide shown in Fig. 3A (see step 12 in Section 2.2). Short linkers exert a pinching force on circularly permuted proteins, destabilizing them (Butler, Mitrea, Mitrousis, Cingolani, & Loh, 2009).

2.2 Gene construction

Following the steps below, one person can construct a library of 10 AFF biosensor genes in less than 2 weeks including primer synthesis turnaround. Commercial synthesis of full-length sensor genes typically takes longer and is considerably more expensive. If genes are purchased, include a sequence encoding a 6 × HisTag at the C-terminus of the protein for purification. For constructing AFF switch genes in-house, the GCN4 gene is first inserted into the target protein gene at the desired site (e.g., between residues 50–51 of nLuc)

creating the N-frame gene. DNA encoding for the duplicate segment (e.g., residues 6*–50* of nLuc, preceded by an appropriate linker sequence) is then ligated to the 3′-end to complete the construction of the AFF switch gene.

Here, we assemble the GCN4 gene (174bp) from synthetic oligonucleotides. For larger input genes, one may construct the sensor using other cloning strategies such as Gibson assembly (Gibson et al., 2009), overlap extension PCR (Bryksin & Matsumura, 2010; Higuchi, Krummel, & Saiki, 1988), and megapriming/restriction-free cloning (Forloni, Liu, & Wajapeyee, 2019; Miyazaki, 2011; Ulrich, Andersen, & Schwartz, 2012). These methods employ PCR to amplify the output and/or input genes (or fragments thereof) and join them in one or a few steps. The same design considerations and screening protocols that we describe below apply to these other techniques as well.

1. Dissolve oligonucleotides to 100μM in either in DNase-free H_2O or TE buffer (10mM Tris-HCl, 1mM disodium EDTA, pH8.0). Dilute to 10μM for working stocks and freeze the rest at −20°C.
2. Set up PCR reactions according to Table 1, using Q5 polymerase and following to the manufacturer's instructions. The workflow of the PCR reactions is shown in Fig. 5A.
3. Perform PCR 1.0 and PCR 2.0 first and dilute the products directly into PCR 1.1 and PCR 2.1, respectively, without purification. PCR reactions 1 and 2 split the output protein gene and attach 3′-and 5′-portions of the GCN4 gene to the two pieces (Fig. 5A). These reactions are carried out in two steps to keep the primers small enough (60nt) for facile DNA synthesis.
4. Load the products of PCR 1.1 and PCR 2.1 on a 1% agarose gel (0.5g agarose in 50mL TAE buffer) and run at 7V/cm in TAE buffer for 20min. Stain the gel by submerging in 1μg/mL ethidium bromide solution in TAE and rocking for 20min.
5. Place the gel on a UV transilluminator and excise the bands corresponding to PCR 1.1 (960bp) and PCR 2.1 (470bp). Purify using the PCR cleanup kit following the manufacturer's instructions. Measure the concentration of purified products using a nanodrop. These are the templates for PCR 3.
6. For the PCR 3 reaction, add GC enhancer (included in the PCR kit) to 20% (V/V). Purify the product as described in steps 4–5 above. PCR 3 creates the N-frame gene (1400bp).
7. Digest the PCR 3 product in a 30μL reaction containing 25μL of the purified PCR product, 3μL of CutSmart buffer (NEB), and 1μL (10 units) each N*de*I and N*ot*I enzymes. Incubate at 37°C for 2h and purify both using the PCR cleanup kit. Digest 1μg pET21 vector (Novagen) the same way.
8. Mix the digested pET21 vector and the digested PCR 3 product in a 1:10 ratio (final volume 8μL) and add 1μL each of T4 ligase (20 units) and 10 × T4 ligase buffer (NEB). Ligate for 1h at room temperature.Step 8 inserts the N-frame gene into the expression vector.

9. Transform the above ligated product into Top-1 cells and spread the cells on a pre-warmed LB agar plate (50μg/mL ampicillin) and incubate at 37°C until colonies appear (overnight).

10. Screen for the correctly-sized N-frame genes by colony PCR prior to sequencing (optional). A representative gel is shown in Fig. 5B.

    a. Select 8 colonies and number them on the bottom of the plate. More colonies may be needed if the ligation was poor.

    b. Prepare a Taq polymerase PCR mix (Accuris) that contains the appropriate primers (mNG_SeqFwd and T7Terminator). One of the primers should bind inside the insert, and the second primer should bind the vector. Transfer 10 μL of the mix to 9 PCR tubes. One of these tubes will serve as a negative control.

    c. Gently touch each colony with a 10μL pipette tip and transfer a small fraction to the corresponding PCR tube.

    d. Transfer the tubes to the thermocycler. Set the initial denaturation phase to 10min and 95°C to ensure bacterial lysis and release of DNA.

    e. Run the PCR products on a 1% agarose gel and stain with ethidium bromide as in step 5 of this protocol. Verify that the negative control does not produce a product and identify colonies that produce the expected PCR product (870bp for the N-frame construct and 1040bp for the full AFF gene).

11. Pick colonies and prepare plasmid DNA for sequencing following established protocols. If screening (step 10) was performed, two positive colonies are sufficient.

12. Once the N-frame gene is fully sequenced and verified, perform PCR 4 to create the DNA specifying the CP-frame (encoding nLuc residues 6*–50*). The PCR 4 reaction in Table 1 specifies primers that encode for a 6-aa peptide (AAGASS) that links nLuc residue 171 to residue 6* in the CP-frame. This linker length was best for nLuc-AFF, but for new proteins we recommend that various linker lengths (e.g., 5 linkers from 2 to 15 amino acids, depending on N-C distance of the protein) be constructed in this step for tuning the stability of the CP-fold. Table 1 provides examples of primer redesign if a different linker length is desired (PCR 5-PCR 7).

13. Digest the product of PCR 4 (purified as in step 5) with *Not*I and *Xho*I following instructions in step 7. Also digest the plasmid containing the N-frame gene (from step 12) with *Not*I and *Xho*I. Purify both digested products using the PCR purification kit.

14. Assemble the complete nLuc-AFF gene. Mix the digested plasmid and the digested PCR 4 product (1:10 ratio). Ligate using T4 ligase for 1h at room temperature.

15. Transform competent Top-1 cells with the ligated products and plate them on an LB-agar dish supplemented with 50μg/mL ampicillin. Screen for insertion of the 6*–50* fragment (step 10 above; see Fig. 5B for example gel). Purify plasmid DNA from positive colonies and verify the gene sequence.

3. Design and validation of DNA tools

This section outlines the steps for creating DNA hairpins and modifying existing DNA aptamers to perform the adapter functions described in Fig. 4A and B, respectively. The protocols are specific for AP-1, but they are easily modified to accommodate different sequences for activating future biosensors that employ DNA recognition domains other than GCN4.

3.1 DNA hairpins

1. When choosing the DNA sequence to be detected, verify that it is free of stable secondary structures, either using NUPACK (Zadeh et al., 2011) or IDT OligoAnalyzer (Owczarzy et al., 2008). For example, the free energy of secondary structure formation calculated by NUPACK should be less than 5kcal/mol. The minimum length of the sequence to be detected is 23nt (equating to 5nt for the toehold plus 18nt for the stem of the hairpin). A good source of target sequences that meet these criteria are those for which PCR primers exist in literature, for instance the Centers for Disease Control 2019-Novel Coronavirus RT-PCR Diagnostic panel (CDC Catalog # 2019-nCoVEUA-01).
2. For each target sequence, construct two hairpins whose 5′-toehold (5–8nt) and stem (18–20nt) comprise the antisense recognition sequence and whose loops contain complementary AP-1 sequences. Complete the hairpin by adding the 3′-stem (without the toehold sequence). Refer to Fig. 4A.
3. Design the hairpins to dimerize when opened by TMSD but remain as stable monomers in the absence of target. Confirm this by analyzing the individual hairpins using NUPACK (maximum complex size of 2). The equilibrium fraction of monomeric hairpin should be ≥95% of all calculated structures. The palindromic AP-1 sequence in the loop contributes to dimerization efficiency. If the fraction of dimerized species is too high, disrupt base stacking between AP-1 and the hairpin by including a 1–3nt “bulge” at either end of the loop. Make the corresponding change in the matching hairpin to maintain complementarity.
4. Check for efficiency of hairpin opening in response to the target sequence by entering both the target and the hairpin sequences for each hairpin independently in NUPACK. At least half of the hairpin monomer should react with target to form the hairpin-target dimer. The palindromic AP-1 sequence can reduce the efficiency of this reaction by forming a larger, more stable hairpin. If the efficiency is too low, add a small bulge as described in the above step to disrupt base stacking between AP-1 and the hairpin stem.

5. Purify the synthesized hairpins following the protocol in Section 3.3. Validate the metastability, efficiency, and specificity of the hairpins using non-denaturing PAGE. A representative gel is shown in Fig. 6A.

    a. Resuspend hairpins individually in TE buffer to a final concentration of 100μM. Dilute to 2μM in 20mM Tris (pH8.5), 150mM NaCl.

    b. Heat at 95°C for 5min, snap cool to room temperature, and leave at room temperature for 30 min. This step monomerizes any dimerized hairpins. Snap-cooled hairpins can be stored at 4°C for at least 2days.

    c. Prepare the following reactions (0.5μM of each oligonucleotide in 20mM Tris pH8.5, 150mM NaCl): each hairpin alone (i and ii), both hairpins mixed (iii), target alone (iv), target added to each hairpin individually (v and vi), and target mixed with both hairpins (vii). Incubate at room temperature for 2h.

    d. Load the reactions on a 12% PAGE gel (19:1 acrylamide:bisacrylamide) and run in 1 × Tris-borate-EDTA (TBE) buffer at 160V for 1h. Stain the gel with ethidium bromide (2min) and soak in water (10min) with several rinses.

    e. Check for sufficient metastability by noting a single band in reactions (i–iii). If a second band is seen in (i) or (ii), optimize the bulge region according to step 3 of this section. If reaction (iii) produces two bands, elongate the hairpin stems to increase stability.

    f. Verify successful TMSD by observing a higher molecular weight species when either hairpin was mixed with the target (reactions v and vi). If TMSD is too inefficient (<50% of the hairpin reacted), elongate the toehold sequence up to 12nt and increase the bulge length to further destabilize the hairpin. We typically obtain an optimal balance of hairpin metastability and reactivity when the bulge region is ~3nt long.

    g. Confirm formation of the activated complex (duplex of hairpins hybridized via their AP-1 sequences) by verifying a new, slower-migrating band in reaction (vii). When the target is added to both hairpins in the same reaction, a synergistic effect occurs in which the AP-1 regions of the hairpins drive formation of the activated complex. This causes TMSD to be more efficient when both hairpins are present (reaction vii) compared to when only one is present (reactions v and vi).

3.2 DNA aptamers

1. Choose a DNA aptamer whose 3′- and 5′-ends form a duplex at least 6bp long. If the structure is not known, it can be predicted using NUPACK.

2. Using NUPACK, design two modified aptamers in which the AP-1 sequence and a 6–8nt clamp are added to either end, according to steps 2a and 2b below.

Refer to Fig. 4B. Only one of the aptamers is needed for the biosensor and either orientation will theoretically work. However, the simulations will typically predict that one will perform better by the criteria described below:

a. Add AP-1 to the 5′-end of the aptamer followed by a 6–8nt clamp. The clamp is complementary to the 5′-end of the aptamer immediately preceding AP-1. The goal is for NUPACK to predict that greater than 90% of molecules will fold into a stable hairpin consisting of the 5′-end/clamp as the stem and AP-1 embedded in the loop. No more than 10% should adopt other structures, including the aptamer form and multimeric states. If a large fraction of DNA adopts the aptamer fold, increase the clamp length by 1nt at a time so as not to make the stem too stable. If the hairpin fold is much more stable than the aptamer fold, excessive amounts of the aptamer ligand will be required for binding and conversion to the aptamer form.

b. Add AP-1 to the 3′-end of the aptamer preceded by a 6–8nt clamp. The clamp is complementary to the 3′-end of the aptamer immediately following AP-1. Repeat the NUPACK optimizations as described in step 2a above.

3. Using NUPACK, design the naked AP-1 strand. Make the initial sequence complementary to AP-1 plus (initially) 3–4nt of clamp, the order of which depends on which modified aptamer was chosen from step 2 above. NUPACK should predict no interaction between naked AP-1 and modified aptamer. If this complex is populated significantly (greater than 10%), decrease the length of the anti-clamp sequence in naked AP-1.

4. Purify synthesized aptamer and naked AP-1 according to Section 3.3. Validate metastability, ligand binding, and activation using non-denaturing PAGE. A representative gel is shown in Fig. 6B.

a. Resuspend modified aptamer and naked AP-1 oligonucleotides, heat, and snap cool as described in steps 5a and 5b of Section 3.1.

b. Prepare the following reactions (0.5μM of each oligonucleotide in 20mM Tris pH8.5, 150mM NaCl, 5mM $MgCl_2$): modified aptamer only (i), naked AP-1 only (ii), modified aptamer mixed with naked AP-1 (iii). Split reactions (i–iii) in half to generate reactions (iv–vi). Add ligand to reactions (iv–vi) (5mM ATP or 10μM serotonin in the examples shown in Figs. 6B and 7C, respectively) and an equal volume of buffer to reactions (i–iii). Incubate at room temperature for 2h.

c. Run samples on gel and stain as described in step 5d of Section 3.1.

d. Verify that the aptamer alone does not dimerize. Reaction (i) should produce a dominant faster-running band and a small fraction of slower-running band corresponding to the dimer. If the latter species is >10% of total aptamer, decrease clamp length as detailed in step 2a of this protocol. The fraction of monomeric species can also be increased by

introducing a bulge between AP-1 and clamp sequences (see step 3 of Section 3.1), but we generally find this to be unnecessary. If only a single species is observed, it can be assumed that there is no detectable dimer.

e. Ascertain that aptamer and naked AP-1 do not bind in the absence of ligand. Reaction (iii) with no serotonin or ATP added should closely approximate the sum of reactions (i) and (ii). If an intense, slower-migrating species is observed, decrease the stability of the heterodimeric complex as described in step 2 of this protocol.

f. Validate that the aptamer and naked AP-1 bind in the presence of ligand. Reaction (iii) with ligand added should produce a new, slower-migrating species corresponding to the activated complex. If this band is not apparent, reduce clamp length to destabilize the hairpin (see step 2a of this protocol) and facilitate aptamer folding. Note that the naked AP-1 must be correspondingly shortened.

3.3 PAGE purification of DNA oligonucleotides

All synthetic oligonucleotides must be purified to remove truncated products that will interfere with their intended reactions. Vendors offer HPLC and PAGE purification, but we have found that HPLC does not achieve acceptable levels of purity, and commercial PAGE incurs high cost and low yield. We therefore recommend ordering the oligonucleotides at 50nmol scale, without any purification, and purifying them using the protocol below. We routinely recover 40–60% yield using this method.

1. Prepare a 16.5cm × 19cm gel containing 8% acrylamide (19:1 acrylamide:bisacrylamide), 7M urea, 1 × TBE. Place the gel in an electrophoresis chamber filled with 1 × TBE and pre-heated to 55°C by a recirculating water bath.
2. Pre-run the gel (300V, ≥30min). While the gel is running, prepare 100μL samples containing 50μM oligonucleotide, 50% formamide, and 0.01% bromophenol blue.
3. Place the samples in a 95°C heating block for 5min and then quickly load them on the gel.
4. Run the gel at 200V for 2.5 h, continuing to warm the chamber to 55°C with the recirculating water bath.
5. Gently disassemble the gel and rinse it twice with ddH_2O. Stain for 20min with 2% methylene blue. Rinse the gel twice with ddH_2O.
6. The full-length oligonucleotide will appear as bright band with the truncated impurities running as a smear under it. Stable hairpins may not completely denature, resulting in two principal bands (monomer and dimer) with a smear underneath. In this case, select the lower (monomer) band.

7. Cut the full-length, monomeric band and transfer it to a 1.5mL centrifuge tube. Crush the gel plug by vigorously stabbing it with a 1mL pipette tip until it becomes sticky to the tube.
8. Add 5 × volume of 20mM Tris (pH7.5), 0.5mM NaCl, 1mM EDTA. Mix well and freeze/thaw twice by placing the tube in crushed dry ice for 20min followed by a 37°C bath for 5min. Incubate the tube at 37°C overnight with shaking.
9. Centrifuge the tubes at 16,000 RCF for 10min. Gently remove the solution from the tube, leaving the pelleted gel fragments undisturbed, and transfer to a new 1.5mL tube.
10. Add 800μL of 1-butanol followed by vigorous vortexing. Centrifuge at 16,000 RCF for 2min.
11. Discard the top organic phase and add 1/5 volume of 3M sodium acetate (pH5.2) followed by 2.5 volumes of 95% ethanol pre-chilled at −80°C. Place the tube at −80°C for 1h, or −20°C overnight.
12. Centrifuge at 16,000 RCF for 30min at 4°C. A small white pellet can be observed if the DNA concentration was high.
13. Discard the supernatant, taking special care not to dislodge the DNA pellet as it does not adhere tightly to the tube and is sometimes not visible.
14. Add 1mL of 95% ethanol (−80°C), and centrifuge for at 16,000 RCF (20 min, 4°C).
15. Carefully remove the supernatant and dry the pellet by placing the uncovered tube in a 50°C heating block for 20 min.
16. Resuspend the purified DNA in 50μL ddH_2O and measure its concentration by nanodrop.

4. Luminescence activity assays

When designing nLuc-AFF biosensor experiments, plan for a minimum sensor concentration of 8nM (≥30nM recommended). nLuc-AFF binds AP-1 with K_d = 14.3 ± 3.5nM; at least 100nM AP-1 is recommended for robust color change. Express nLuc-AFF in *E.* coli BL21(DE3) cells and purify using nickel-NTA chromatography as described (Sekhon & Loh, 2022).

1. Prepare DNA tools for use in biosensor assays. Dissolve all DNA oligonucleotides in 20mM Tris (pH8.5), 150mM NaCl.
    a. For the positive control, prepare AP-1 by mixing two ssAP-1 strands (50μM each). Anneal by heating at 95°C for 5min and cooling to room temperature over ~30min.
    b. For the hairpins in Fig. 4A, heat the hairpins individually (2μM) at 95°C for 5min and snap cool.

c. For the modified aptamers in Fig. 4B, heat the aptamers (2μM) at 95°C for 5min and snap cool.

2. Prepare the biosensor reactions. After mixing the components at the indicated concentrations (in 20mM Tris, 150mM NaCl, 0.1mg/mL BSA, pH8.5), incubate the solution at room temperature, overnight, and in the dark.

a. For the positive control, mix nLuc-AFF (30nM) and AP-1 (1μM).

b. For detecting a target DNA or RNA sequence, mix nLuc-AFF (30nM) and both hairpins (0.2μM each). Split into two tubes and add target ssDNA (or ssRNA) to one and an equal volume of buffer to the other.

c. For detecting small molecule/protein ligands, mix nLuc-AFF (30nM), modified aptamer (0.2μM), and naked AP-1 (0.2μM). Add $MgCl_2$ to 5mM. Split into two tubes and add target ligand to one and an equal volume of buffer or vehicle to the other.

3. Image the reactions. Add furimazine to a final concentration of 20μM to all reactions and aliquot 150μL samples onto an opaque, white 96-well plate. White plates offer ~fivefold greater sensitivity than black plates. Representative luminescence spectra and cell phone images are shown in Fig. 7.

a. For quantification by luminescence spectra, image using a scanning plate reader in luminescence mode. For the SpectraMax i3x, scan from 400 to 600nm with 1s integration time, 5nm interval, and a read height of 1mm.

b. For quantification by cell phone camera, obtain a box 4–5in. high and large enough to fit over the 96-well plate. Cut a hole in the bottom of the box just wide enough to allow a full view from the camera. Invert the box over the plate and place the phone over the hole. Manually focus on the wells. Turn off the room lights before taking the picture. Camera settings will vary depending on the phone. We use the following settings for a OnePlus 7 Pro: ISO 200, f/1.6, auto white balance (or manually set to 3500–4000K), and 10–15s exposure.

4. Quantify the biosensor readout

a. Quantify luminescence spectra by dividing the intensity at 460nm by the intensity at 520nm.

b. Quantify cell phone images using ImageJ (Schneider, Rasband, & Eliceiri, 2012). Split the image into RBG channels (Image ➔ Color ➔ SplitChannels). Manually draw a circle surrounding each well and calculate the average intensity (Analyze ➔ Measure), ensuring “Mean Gray Value” is selected in Analyze ➔ Set Measurements. Divide the intensity of the blue channel by that of the green channel. We generally find background subtraction to be unnecessary, as signal:background is typically ~70. If this value is low (e.g., in conditions of low target

concentration or high ISO setting), subtract the background (obtained from a buffer-only control well) from each channel before dividing.

5. Summary and conclusions

This chapter describes workflows for designing and constructing a luminescent protein switch, nLuc-AFF, that changes color upon an input from several DNA devices. The chief application of nLuc-AFF is for detecting clinically relevant biomarkers, such as small molecules and pathogenic DNA/RNA sequences. A biosensor's limit of detection (LoD) is typically dictated by its affinity for the target. The dissociation constant of nLuc-AFF/AP-1 DNA (as well as its LoD of a target sequence) is ~10nM. The semisynthetic nLuc sensors and Cas enzymes bind their nucleic acid targets more tightly by virtue of DNA-DNA base pairing and recognition of 3D RNA structures, respectively. These latter methods consequently exhibit lower LoD (50–100pM; Chang et al., 2020; Engelen et al., 2017; Kaminski et al., 2021), although they are still not sensitive enough to directly detect pathogen DNA/RNA that can exist in concentrations below 1fM (Aoki et al., 2021; Park, Roh, & Kim, 2022; Schiffer, Swan, Prlic, & Lund, 2018). Accordingly, many existing CRISPR-based diagnostics employ a DNA pre-amplification step such as loop-assisted isothermal amplification (LAMP; Notomi et al., 2000; Tomita, Mori, Kanda, & Notomi, 2008) or recombinase-polymerase amplification (RPA; Piepenburg, Williams, Stemple, & Armes, 2006) to increase pathogen DNA/RNA to requisite levels (Kaminski et al., 2021). The same principle may be applied to nLuc-AFF to enhance its sensitivity for use in clinical samples.

In addition to generating biosensors, this chapter is meant to provide guidance for creating novel DNA-activated proteins by inserting GCN4 (and other DNA binding domains) into output proteins and enzymes other than nLuc. The luminescent biosensors discussed above take advantage of phenomena specific to light-emitting proteins: resonance energy transfer and nLuc inhibition. By replacing one set of amino acids with another, the AFF mechanism establishes two mutually-exclusive folds of the protein whose biological activities can be made distinct. For example, a catalytic or ligand-binding residue can be mutated in one frame to switch enzymatic or binding activity on/off because of target binding.

An attractive feature of AFF switches is the ease by which modularity is achieved. To enable them to recognize new DNA/RNA or small molecule/protein targets, one needs only to modify the DNA hairpin or aptamer components; the protein stays the same. Protein engineering often involves trial-and-error experimentation, and Section 2 outlines an efficient workflow for creating a switchable protein. The DNA engineering steps in Section 3 can then readily generate the DNA tools that plug into the protein switch.

Acknowledgments

This work was supported by NIH grant R01GM115762 to S.N.L.

Abbreviations

AFF	alternate frame folding

BRET	bioluminescent resonance energy transfer
CP	circular permutant or circularly permuted
GeNL	green enhanced nanolantern
mNG	mNeonGreen
nLuc	nanoluciferase
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
TMSD	toehold-mediated strand displacement

References

Aoki K, Nagasawa T, Ishii Y, Yagi S, Okuma S, Kashiwagi K, et al. (2021). Clinical validation of quantitative SARS-CoV-2 antigen assays to estimate SARS-CoV-2 viral loads in nasopharyngeal swabs. Journal of Infection and Chemotherapy, 27(4), 613–616. 10.1016/j.jiac.2020.11.021. [PubMed: 33423918]

Baird GS, Zacharias DA, & Tsien RY (1999). Circular permutation and receptor insertion within green fluorescent proteins. Proceedings of the National Academy of Sciences, 96(20), 11241–11246. 10.1073/pnas.96.20.11241.

Biewenga L, Rosier BJHM, & Merkx M (2020). Engineering with NanoLuc: A playground for the development of bioluminescent protein switches and sensors. Biochemical Society Transactions, 48(6), 2643–2655. 10.1042/BST20200440. [PubMed: 33242085]

Bryksin AV, & Matsumura I (2010). Overlap extension PCR cloning: A simple and reliable way to create recombinant plasmids. BioTechniques, 48(6), 463–465. 10.2144/000113418. [PubMed: 20569222]

Butler JS, Mitrea DM, Mitrousis G, Cingolani G, & Loh SN (2009). Structural and thermodynamic analysis of a conformationally strained circular permutant of barnase. Biochemistry, 48(15), 3497–3507. 10.1021/bi900039e. [PubMed: 19260676]

Carrasco-López C, Zhao EM, Gil AA, Alam N, Toettcher JE, & Avalos JL (2020). Development of light-responsive protein binding in the monobody non-immunoglobulin scaffold. Nature Communications, 11(1), 4045. 10.1038/s41467-020-17837-7.

Chang D, Kim KT, Lindberg E, & Winssinger N (2020). Smartphone DNA or RNA sensing using semisynthetic luciferase-based logic device. ACS Sensors, 5(3), 807–813. 10.1021/acssensors.9b02454. [PubMed: 32124606]

Chen JS, Ma E, Harrington LB, Da Costa M, Tian X, Palefsky JM, et al. (2018). CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. Science (New York, N.Y.), 360(6387), 436–439. 10.1126/science.aar6245. [PubMed: 29449511]

Cutler TA, Mills BM, Lubin DJ, Chong LT, & Loh SN (2009). Effect of interdomain linker length on an antagonistic folding–unfolding equilibrium between two protein domains. Journal of Molecular Biology, 386(3), 854–868. 10.1016/j.jmb.2008.10.090. [PubMed: 19038264]

Dagliyan O, Krokhotin A, Ozkan-Dagliyan I, Deiters A, Der CJ, Hahn KM, et al. (2018). Computational design of chemogenetic and optogenetic split proteins. Nature Communications, 9(1), 4042. 10.1038/s41467-018-06531-4.

Das A, Goswami HN, Whyms CT, Sridhara S, & Li H (2022). Structural principles of CRISPR-Cas enzymes used in nucleic acid detection. Journal of Structural Biology, 214(1), 107838. 10.1016/j.jsb.2022.107838. [PubMed: 35123001]

de Lange O, Wolf C, Dietze J, Elsaesser J, Morbitzer R, & Lahaye T (2014). Programmable DNA-binding proteins from Burkholderia provide a fresh perspective on the TALE-like repeat domain. Nucleic Acids Research, 42(11), 7436–7449. 10.1093/nar/gku329. [PubMed: 24792163]

Engelen W, van de Wiel KM, Meijer LHH, Saha B, & Merkx M (2017). Nucleic acid detection using BRET-beacons based on bioluminescent protein–DNA hybrids. Chemical Communications, 53(19), 2862–2865. 10.1039/C6CC10032E. [PubMed: 28217801]

England CG, Ehlerding EB, & Cai W (2016). NanoLuc: A small luciferase is brightening up the field of bioluminescence. Bioconjugate Chemistry, 27(5), 1175–1187. 10.1021/acs.bioconjchem.6b00112. [PubMed: 27045664]

Flory JD, Shinde S, Lin S, Liu Y, Yan H, Ghirlanda G, et al. (2013). PNA-peptide assembly in a 3D DNA nanocage at room temperature. Journal of the American Chemical Society, 135(18), 6985–6993. 10.1021/ja400762c. [PubMed: 23521013]

Forloni M, Liu AY, & Wajapeyee N (2019). Megaprimer polymerase chain reaction (PCR)-based mutagenesis. Cold Spring Harbor Protocols, 2019(6). 10.1101/pdb.prot097824.

Fu J, Yang YR, Johnson-Buck A, Liu M, Liu Y, Walter NG, et al. (2014). Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm. Nature Nanotechnology, 9(7), 531–536. 10.1038/nnano.2014.100.

Furman JL, Badran AH, Ajulo O, Porter JR, Stains CI, Segal DJ, et al. (2010). Toward a general approach for RNA-templated hierarchical assembly of split-proteins. Journal of the American Chemical Society, 132(33), 11692–11701. 10.1021/ja104395b. [PubMed: 20681585]

Gianneschi NC, & Ghadiri MR (2007). Design of molecular logic devices based on a programmable DNA-regulated semisynthetic enzyme. Angewandte Chemie International Edition, 46(21), 3955–3958. 10.1002/anie.200700047. [PubMed: 17427900]

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, & Smith HO (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods, 6(5), 343–345. 10.1038/nmeth.1318. [PubMed: 19363495]

Gil AA, Carrasco-López C, Zhu L, Zhao EM, Ravindran PT, Wilson MZ, et al. (2020). Optogenetic control of protein binding using light-switchable nanobodies. Nature Communications, 11(1), 4044. 10.1038/s41467-020-17836-8.

Gootenberg JS, Abudayyeh OO, Lee JW, Essletzbichler P, Dy AJ, Joung J, et al. (2017). Nucleic acid detection with CRISPR-Cas13a/C2c2. Science (New York, N.Y.), 356(6336), 438–442. 10.1126/science.aam9321. [PubMed: 28408723]

Guo Z, Parakra RD, Xiong Y, Johnston WA, Walden P, Edwardraja S, et al. (2022). Engineering and exploiting synthetic allostery of NanoLuc luciferase. Nature Communications, 13(1), 789. 10.1038/s41467-022-28425-2.

Ha J-H, Karchin JM, Walker-Kopp N, Castañeda CA, & Loh SN (2015). Engineered domain swapping as an on/off switch for protein function. Chemistry & Biology, 22(10), 1384–1393. 10.1016/j.chembiol.2015.09.007. [PubMed: 26496687]

Ha J-H, Presti MF, & Loh SN (2019). A single protein disruption site results in efficient reassembly by multiple engineering methods. Biophysical Journal, 117(1), 56–65. 10.1016/j.bpj.2019.06.002. [PubMed: 31221439]

Ha J-H, Shinsky SA, & Loh SN (2013). Stepwise conversion of a binding protein to a fluorescent switch: Application to thermoanaerobacter tengcongensis ribose binding protein. Biochemistry, 52(4), 600–612. 10.1021/bi301105u. [PubMed: 23302025]

Hall MP, Unch J, Binkowski BF, Valley MP, Butler BL, Wood MG, et al. (2012). Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. ACS Chemical Biology, 7(11), 1848–1857. 10.1021/cb3002478. [PubMed: 22894855]

Hiblot J, Yu Q, Sabbadini MDB, Reymond L, Xue L, Schena A, et al. (2017). Luciferases with tunable emission wavelengths. Angewandte Chemie International Edition, 56(46), 14556–14560. 10.1002/anie.201708277. [PubMed: 28941028]

Higuchi R, Krummel B, & Saiki RK (1988). A general method of in vitro preparation and specific mutagenesis of DNA fragments: Study of protein and DNA interactions. Nucleic Acids Research, 16(15), 7351–7367. [PubMed: 3045756]

Huizenga DE, & Szostak JW (1995). A DNA aptamer that binds adenosine and ATP. Biochemistry, 34(2), 656–665. 10.1021/bi00002a033. [PubMed: 7819261]

John AM, Sekhon H, Ha J-H, & Loh SN (2022). Engineering a fluorescent protein color switch using entropy-driven β-strand exchange. ACS Sensors, 7(1), 263–271. 10.1021/acssensors.1c02239. [PubMed: 35006676]

Jolany Vangah S, Katalani C, Boone HA, Hajizade A, Sijercic A, & Ahmadian G (2020). CRISPR-based diagnosis of infectious and noninfectious diseases. Biological Procedures Online, 22(1), 22. 10.1186/s12575-020-00135-3. [PubMed: 32939188]

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583–589. 10.1038/s41586-021-03819-2. [PubMed: 34265844]

Kaminski MM, Abudayyeh OO, Gootenberg JS, Zhang F, & Collins JJ (2021). CRISPR-based diagnostics. Nature Biomedical Engineering, 5(7), 643–656. 10.1038/s41551-021-00760-7.

Karchin JM, Ha J-H, Namitz KE, Cosgrove MS, & Loh SN (2017). Small molecule-induced domain swapping as a mechanism for controlling protein function and assembly. Scientific Reports, 7(1), 44388. 10.1038/srep44388. [PubMed: 28287617]

Klug A (2010). The discovery of zinc fingers and their applications in gene regulation and genome manipulation. Annual Review of Biochemistry, 79, 213–231. 10.1146/annurev-biochem-010909-095056.

Liu M, Fu J, Hejesen C, Yang Y, Woodbury NW, Gothelf K, et al. (2013). A DNA tweezer-actuated enzyme nanoreactor. Nature Communications, 4(1), 2127. 10.1038/ncomms3127.

Liu Q, Segal DJ, Ghiara JB, & Barbas CF (1997). Design of polydactyl zinc-finger proteins for unique addressing within complexgenomes. Proceedings of the National Academy of Sciences, 94(11), 5525–5530. 10.1073/pnas.94.11.5525.

Lo W-C, Wang L-F, Liu Y-Y, Dai T, Hwang J-K, & Lyu P-C (2012). CPred: A web server for predicting viable circular permutations in proteins. Nucleic Acids Research, 40(W1), W232–W237. 10.1093/nar/gks529. [PubMed: 22693212]

Madsen M, & Gothelf KV (2019). Chemistries for DNA nanotechnology. Chemical Reviews, 119(10), 6384–6458. 10.1021/acs.chemrev.8b00570. [PubMed: 30714731]

Miyazaki K (2011). Chapter seventeen—MEGAWHOP cloning: A method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. In Voigt C (Ed.), Vol. 498. Methods in Enzymology (pp. 399–406). Academic Press. 10.1016/B978-0-12-385120-8.00017-6. [PubMed: 21601687]

Moore M, Klug A, & Choo Y (2001). Improved DNA binding specificity from polyzinc finger peptides by using strings of two-finger units. Proceedings of the National Academy of Sciences, 98(4), 1437–1441. 10.1073/pnas.98.4.1437.

Morbitzer R, Römer P, Boch J, & Lahaye T (2010). Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. Proceedings of the National Academy of Sciences, 107(50), 21617–21622. 10.1073/pnas.1013133107.

Myhrvold C, Freije CA, Gootenberg JS, Abudayyeh OO, Metsky HC, Durbin AF, et al. (2018). Field-deployable viral diagnostics using CRISPR-Cas13. Science (New York, N.Y.), 360(6387), 444–448. 10.1126/science.aas8836. [PubMed: 29700266]

Nalefski EA, Patel N, Leung PJY, Islam Z, Kooistra RM, Parikh I, et al. (2021). Kinetic analysis of Cas12a and Cas13a RNA-guided nucleases for development of improved CRISPR-Based diagnostics. IScience, 24(9). 10.1016/j.isci.2021.102996.

Nasu Y, Shen Y, Kramer L, & Campbell RE (2021). Structure- and mechanism-guided design of single fluorescent protein-based biosensors. Nature Chemical Biology, 17(5), 509–518. 10.1038/s41589-020-00718-x. [PubMed: 33558715]

Ni Y, Rosier BJHM, van Aalen EA, Hanckmann ETL, Biewenga L, Pistikou A-MM, et al. (2021). A plug-and-play platform of ratiometric bioluminescent sensors for homogeneous immunoassays. Nature Communications, 12(1), 4586. 10.1038/s41467-021-24874-3.

Nicholes N, Date A, Beaujean P, Hauk P, Kanwar M, & Ostermeier M (2016). Modular protein switches derived from antibody mimetic proteins. Protein Engineering, Design and Selection, 29(2), 77–85. 10.1093/protein/gzv062.

Notomi T, Okayama H, Masubuchi H, Yonekawa T, Watanabe K, Amino N, et al. (2000). Loop-mediated isothermal amplification of DNA. Nucleic Acids Research, 28(12), e63. 10.1093/nar/28.12.e63. [PubMed: 10871386]

O'Connell MR (2019). Molecular mechanisms of RNA targeting by Cas13-containing type VI CRISPR–Cas systems. Journal of Molecular Biology, 431(1), 66–87. 10.1016/j.jmb.2018.06.029. [PubMed: 29940185]

O'Shea EK, Klemm JD, Kim PS, & Alber T (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science (New York, N.Y.), 254(5031), 539–544. 10.1126/science.1948029. [PubMed: 1948029]

Owczarzy R, Tataurov AV, Wu Y, Manthey JA, McQuisten KA, Almabrazi HG, et al. (2008). IDT SciTools: A suite for analysis and design of nucleic acid oligomers. Nucleic Acids Research, 36(Suppl. 2), W163–W169. 10.1093/nar/gkn198. [PubMed: 18440976]

Park Y, Roh J, & Kim S (2022). Performance evaluation of the aptima assays in comparison with the cobas 6800 assays for the detection of HIV-1, HBV, and HCV in clinical samples. Annals of Laboratory Medicine, 42(4), 447–456. 10.3343/alm.2022.42.4.447. [PubMed: 35177565]

Piepenburg O, Williams CH, Stemple DL, & Armes NA (2006). DNA detection using recombination proteins. PLoS Biology, 4(7), e204. 10.1371/journal.pbio.0040204. [PubMed: 16756388]

Quijano-Rubio A, Yeh H-W, Park J, Lee H, Langan RA, Boyken SE, et al. (2021). De novo design of modular and tunable protein biosensors. Nature, 591(7850), 482–487. 10.1038/s41586-021-03258-z. [PubMed: 33503651]

Rogers JM, Barrera LA, Reyon D, Sander JD, Kellis M, Keith Joung J, et al. (2015). Context influences on TALE–DNA binding revealed by quantitative profiling. Nature Communications, 6(1), 7440. 10.1038/ncomms8440.

Rosier BJHM, Markvoort AJ, Gumí Audenis B, Roodhuizen JAL, den Hamer A, Brunsveld L, et al. (2020). Proximity-induced caspase-9 activation on a DNA origami-based synthetic apoptosome. Nature Catalysis, 3(3), 295–306. 10.1038/s41929-019-0403-7.

Saghatelian A, Guckian KM, Thayer DA, & Ghadiri MR (2003). DNA detection and signal amplification via an engineered allosteric enzyme. Journal of the American Chemical Society, 125(2), 344–345. 10.1021/ja027885u. [PubMed: 12517141]

Sancho Oltra N, Bos J, & Roelfes G (2010). Control over enzymatic activity by DNA-directed split enzyme reassembly. Chembiochem, 11(16), 2255–2258. 10.1002/cbic.201000517. [PubMed: 20941727]

Schiffer JT, Swan DA, Prlic M, & Lund JM (2018). Herpes simplex virus-2 dynamics as a probe to measure the extremely rapid and spatially localized tissue-resident T cell response. Immunological Reviews, 285(1), 113–133. 10.1111/imr.12672. [PubMed: 30129205]

Schneider CA, Rasband WS, & Eliceiri KW (2012). NIH Image to ImageJ: 25 years of image analysis. Nature Methods, 9(7), 671–675. 10.1038/nmeth.2089. [PubMed: 22930834]

Sekhon H, & Loh SN (2022). Engineering protein activity into off-the-shelf DNA devices. Cell Reports Methods, 2(4), 100202. 10.1016/j.crmeth.2022.100202. [PubMed: 35497497]

Stratton MM, & Loh SN (2011). Converting a protein into a switch for biosensing and functional regulation. Protein Science: A Publication of the Protein Society, 20(1), 19–29. 10.1002/pro.541.

Stratton MM, Mitrea DM, & Loh SN (2008). A Ca2+−sensing molecular switch based on alternate frame protein folding. ACS Chemical Biology, 3(11), 723–732. 10.1021/cb800177f. [PubMed: 18947182]

Suzuki K, Kimura T, Shinoda H, Bai G, Daniels MJ, Arai Y, et al. (2016). Five colour variants of bright luminescent protein for real-time multicolour bioimaging. Nature Communications, 7(1), 13718. 10.1038/ncomms13718.

Tomita N, Mori Y, Kanda H, & Notomi T (2008). Loop-mediated isothermal amplification (LAMP) of gene sequences and simple visual detection of products. Nature Protocols, 3(5), 877–882. 10.1038/nprot.2008.57. [PubMed: 18451795]

Ulrich A, Andersen KR, & Schwartz TU (2012). Exponential megapriming PCR (EMP) cloning—seamless DNA insertion into any target plasmid without sequence constraints. PLoS One, 7(12), e53360. 10.1371/journal.pone.0053360. [PubMed: 23300917]

Walker JR, Hall MP, Zimprich CA, Robers MB, Duellman SJ, Machleidt T, et al. (2017). Highly potent cell-permeable and impermeable NanoLuc luciferase inhibitors. ACS Chemical Biology, 12(4), 1028–1037. 10.1021/acschembio.6b01129. [PubMed: 28195704]

Watson EE, Angerani S, Sabale PM, & Winssinger N (2021). Biosupramolecular systems: Integrating cues into responses. Journal of the American Chemical Society, 143(12), 4467–4482. 10.1021/jacs.0c12970. [PubMed: 33565858]

Wijnands SPW, Engelen W, Lafleur RPM, Meijer EW, & Merkx M (2018). Controlling protein activity by dynamic recruitment on a supramolecular polymer platform. Nature Communications, 9(1), 65. 10.1038/s41467-017-02559-0.

Wu YI, Frey D, Lungu OI, Jaehrig A, Schlichting I, Kuhlman B, et al. (2009). A genetically encoded photoactivatable Rac controls the motility of living cells. Nature, 461(7260), 104–108. 10.1038/nature08241. [PubMed: 19693014]

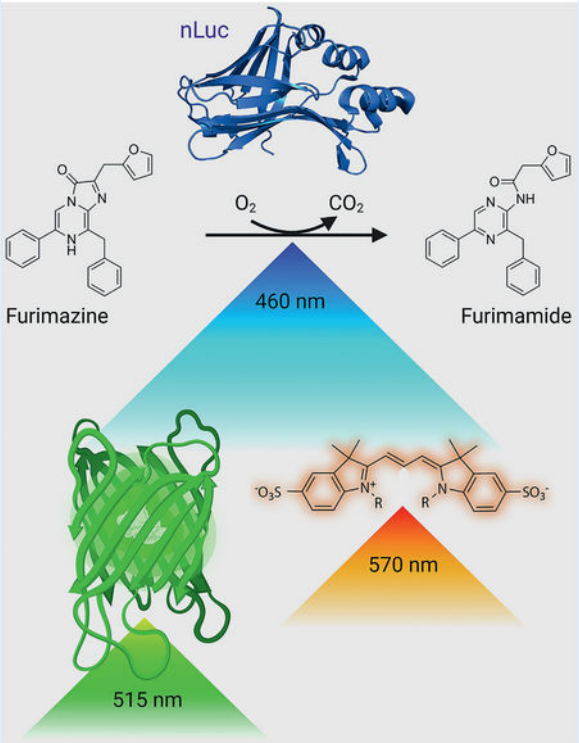
Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, et al. (2011). NUPACK: Analysis and design of nucleic acid systems. Journal of Computational Chemistry, 32(1), 170–173. 10.1002/jcc.21596. [PubMed: 20645303]

Zetsche B, Volz SE, & Zhang F (2015). A split-Cas9 architecture for inducible genome editing and transcription modulation. Nature Biotechnology, 33(2), 139–142. 10.1038/nbt.3149.

Zhao Z, Fu J, Dhakal S, Johnson-Buck A, Liu M, Zhang T, et al. (2016). Nanocaged enzymes with enhanced catalytic activity and increased stability against protease digestion. Nature Communications, 7(1), 10619. 10.1038/ncomms10619.

BOX 1

Nanoluciferase and bioluminescent resonant energy transfer. Nanoluciferase (nLuc; blue protein) is a luminescent protein from the deep-sea shrimp *Oplophorus gracilirostris*. nLuc is the brightest luciferase known (England, Ehlerding, & Cai, 2016), small (171 residues), and thermodynamically stable (Hall et al., 2012), making it amenable to engineering (Biewenga, Rosier, & Merkx, 2020; Guo et al., 2022; Ni et al., 2021). nLuc catalyzes the oxidation of furimazine to furimamide, generating blue light in the process. No ATP or other cofactors are needed. Proteins and dyes whose absorption spectra overlap with nLuc emission, e.g., mNeonGreen (green protein) and Cy3 (orange molecule), can be excited by nLuc via dipole-dipole coupling. They subsequently emit light at their characteristic wavelengths, transforming nLuc's blue emission to red-shifted colors. The efficiency of this conversion is inversely proportional to the distance between nLuc and the BRET acceptor and is also affected by their relative orientations.



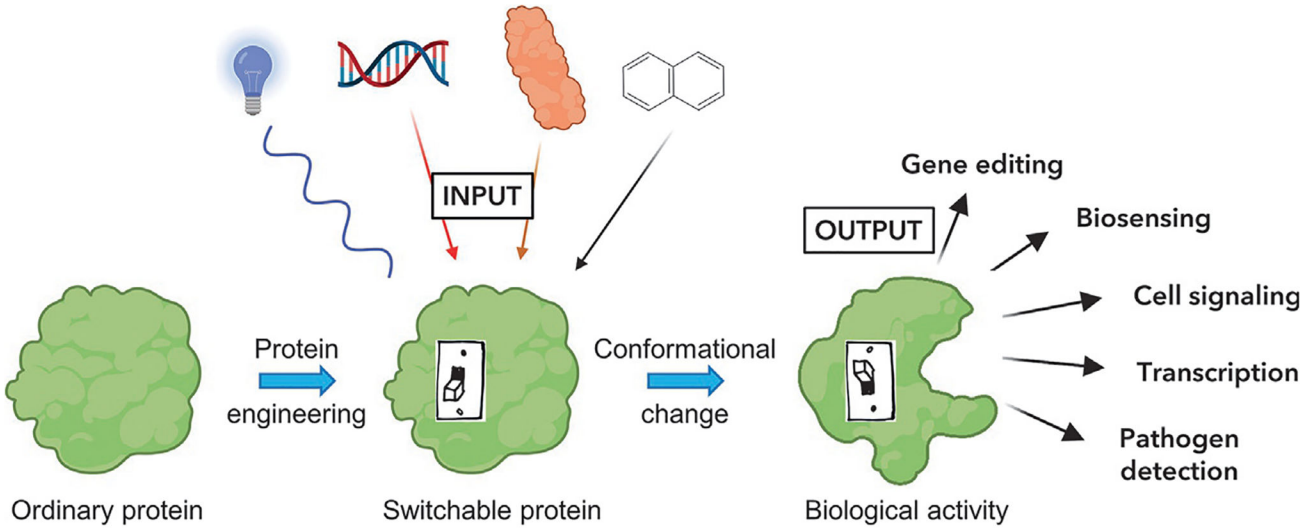


Fig. 1.

Protein and DNA switches convert input signals to biological output. The protein engineering steps typically entail fusing a binding/recognition protein (input domain) to an enzyme (output domain) such that ligand binding to the former induces a conformational change that alters activity of the latter. When the switch is designed to sense a specific DNA or RNA sequence, DNA engineering is employed to create DNA-based input domains as well as tools that convert small molecule and protein inputs to the DNA sequence that activates the biosensor.

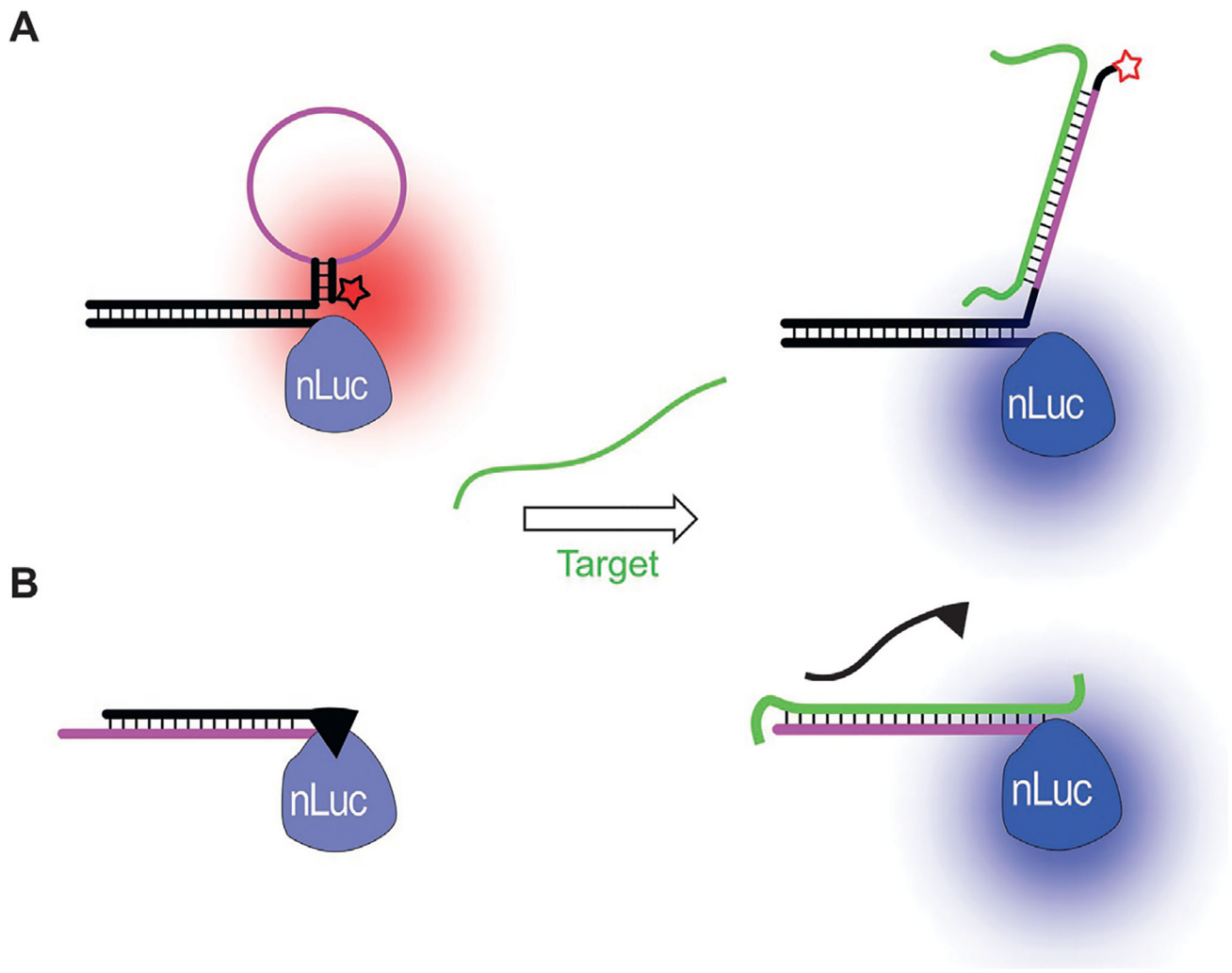


Fig. 2.

Semisynthetic, luminescent sensors for detecting DNA and RNA. The semisynthetic component of the BRET-beacon (A) and inhibitor-displacement (B) sensors consist of nLuc attached to ssDNA and ssPNA, respectively. The DNA-only components are ssDNA oligonucleotides derivatized with a BRET acceptor (Cy3; star in (A)) or an nLuc inhibitor (triangle in (B)). The DNA portions of the switch establish base recognition (pink region) and conformational change functionalities. Binding of the target ssDNA or ssRNA (green strand) to the pink regions increase the distance between nLuc and Cy3 and displace the inhibitor from nLuc, shifting luminescence from orange to blue by decreasing BRET efficiency (A) and turning on blue luminescence (B).

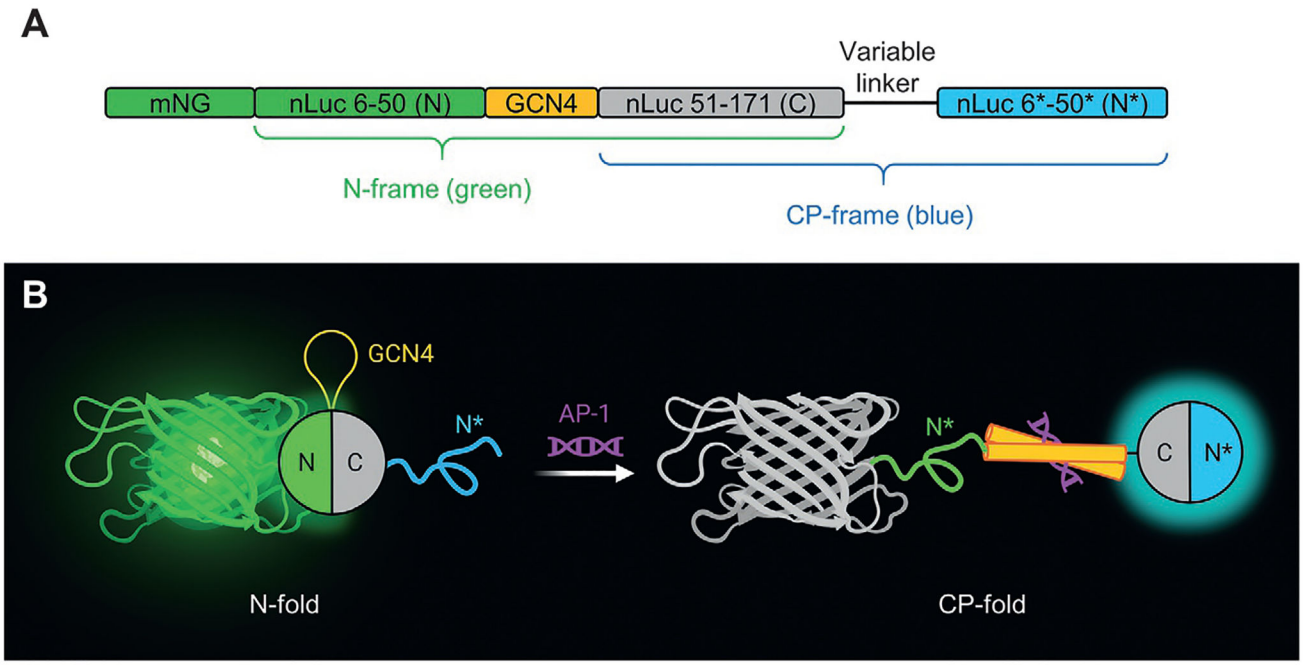


Fig. 3.

Creating protein conformational switches by alternate frame folding. (A) In this example, the nLuc-AFF biosensor was created by first inserting the GCN4 input domain (orange) between residues 50–51 in the nLuc domain of the GeNL protein (green-orange-gray), establishing the N-frame. The CP-frame was introduced by copying nLuc residues 6–50 and appending the duplicate sequence (residues 6*–50*; blue) to the C-terminus using a variable-length linker to adjust thermodynamic stability of the CP-fold. (B) DNA binding induces the GCN4 domain to fold into a long rod, stretching the N-frame and forcing nLuc to refold in the CP-frame. Luminescence changes from green to blue as the nLuc and mNG domains separate, reducing BRET efficiency.

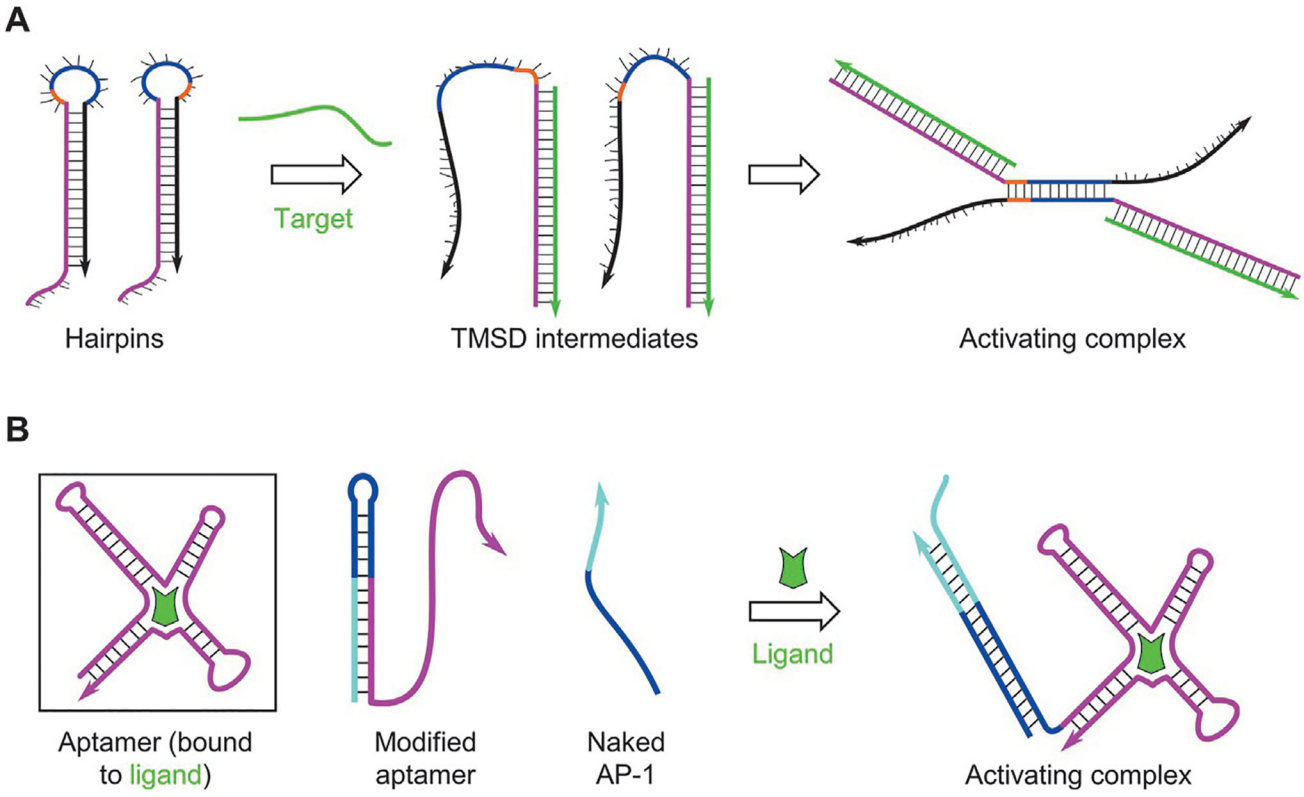


Fig. 4.

DNA tools for converting arbitrary DNA/RNA sequences (A) and small molecules/proteins (B) to AP-1 DNA for biosensor activation. (A) The DNA hairpins consist of the sequence complementary to the target (pink) as the stems and ssAP-1 and its complement (blue) as the loops. Toeholds extend from the 5′-ends. If necessary, a bulge region (orange) can be added between the stem and loop to offset the extra stability of the stem afforded by the palindromic nature of AP-1. Invasion of the target strand (green) into the two hairpins creates the TMSD intermediates which then anneal (via the ssAP-1 regions) to generate the activating species (duplex AP-1). (B) The parental DNA aptamer (pink), bound to a small molecule (green), is shown in the box. The modified aptamer is created by adding ssAP-1 (blue) and a variable-length clamp (cyan) to either the 5′- or 3′-end of the aptamer (5′-end shown). The modified aptamer folds into a hairpin with ssAP-1 comprising the loop, as in (A). The naked AP-1 strand consists of the complement to ssAP-1 and a short region of the clamp. On binding the small molecule, the modified aptamer refolds into the aptamer form, opening the hairpin. The exposed ssAP-1 then anneals with the naked AP-1 strand to form the activating complex.

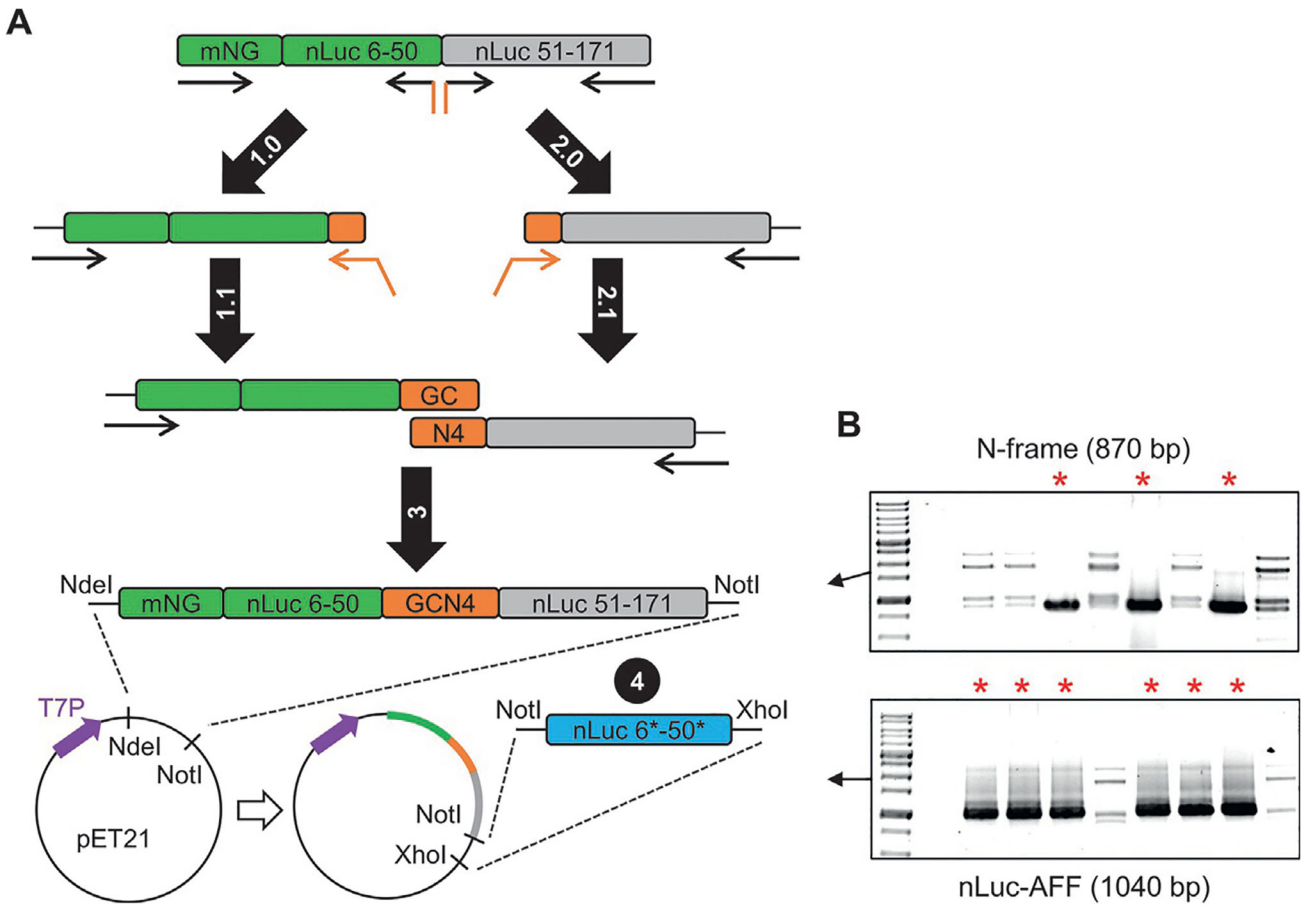


Fig. 5.

Workflow of biosensor gene construction from Section 2.2. (A) The nLuc-AFF gene is assembled by means the PCR reactions numbered in the black arrows and circle, and detailed in Table 1. PCR primers are shown as small arrows below the genes. (B) Colony screening for the N-frame gene (step 10 of Section 2.2) and the NLuc-AFF gene (step 14 of Section 2.2) reveals clones that contain inserts of the correct size (asterisked lanes).

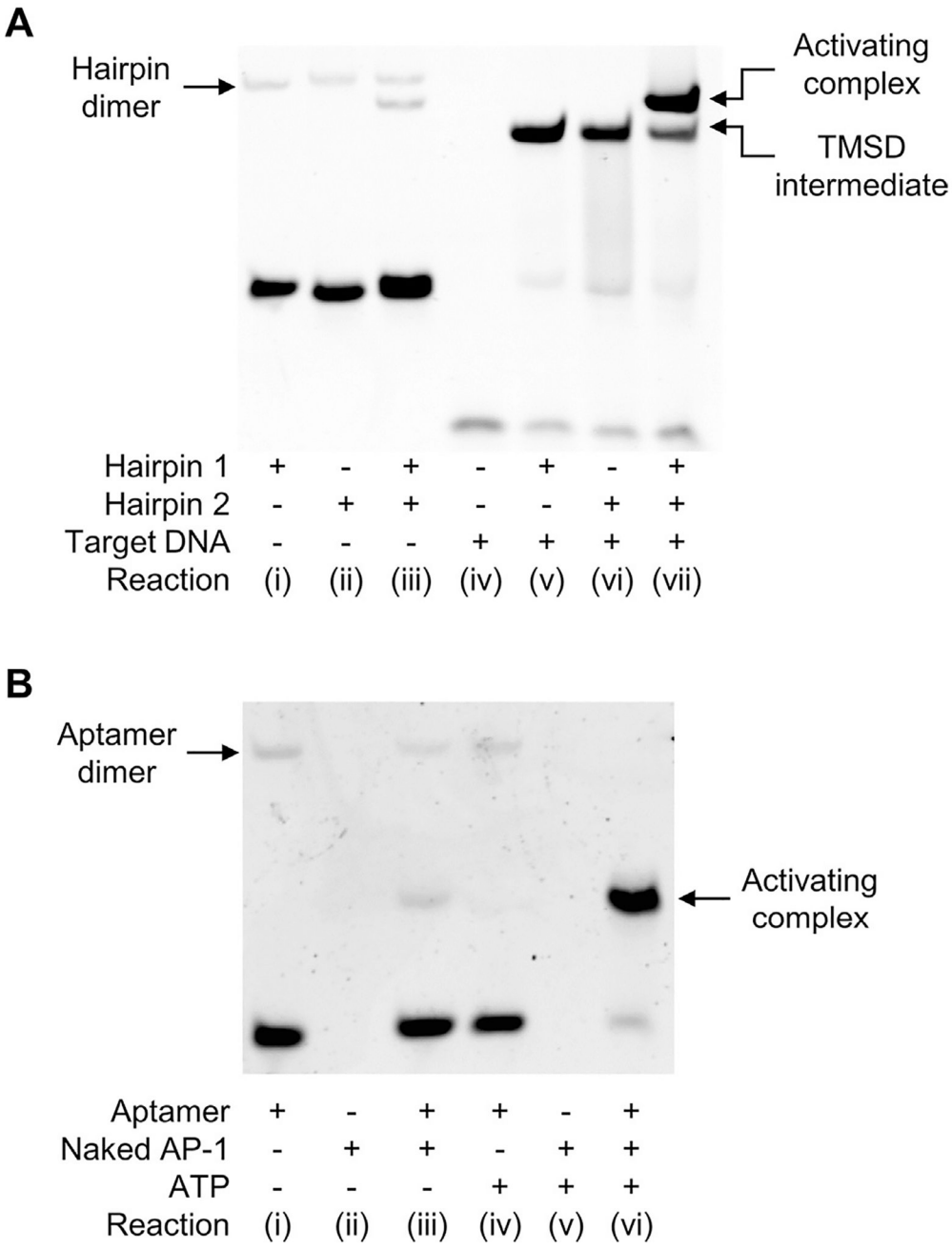


Fig. 6.

Validation of engineered DNA tools created in Section 3. (A) Hairpins from reactions (i)–(vii) in Section 3.1 are shown. The target DNA is the SARS-CoV-2N1 sequence from the CDC diagnostic panel. (B) The modified DNA aptamer (derived from an ATP aptamer; Huizenga & Szostak, 1995) and naked AP-1 strand from reactions (i)–(vi) in Section 3.2 are shown. A small amount of activating complex is formed in the absence of ATP ligand (reaction iii), indicating that AP-1 duplex formation alone can partially drive the hairpin-to-aptamer DNA fold shift. DNA sequences of hairpins, N1, modified aptamer, and naked AP-1 are listed in Table 2.

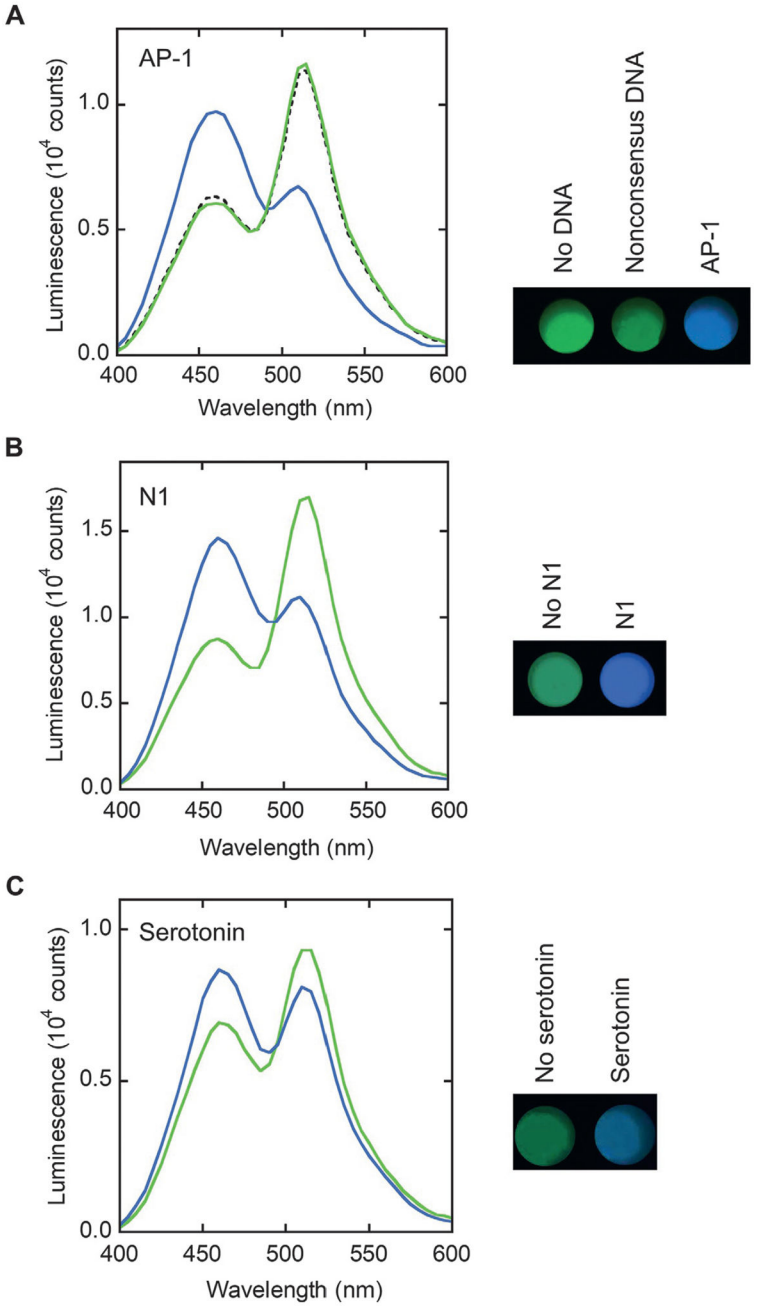


Fig. 7.

Luminescent biosensor imaging from Section 4. Ligands are 1μM AP-1 (A), 400nMN1 oligonucleotide from SARS-CoV-2 (B), and 10μM serotonin (C). Blue and green spectra from the plate reader indicate presence and absence of ligand. The dashed spectrum in (A) was recorded with 1μM nonconsensus DNA to demonstrate AP-1 sequence specificity. Raw cell phone images are shown at right.

Table 1

PCR reactions in Section 2.

PCR #	Forward primer	Reverse primer	Template
1.0	**mNG_NdeI_Fwd** GGAGACATATGGTGTCCAAGGGCGAAGAGG	**nLuc50_GCN4_Rev1 (0.02 μM)** GTTACGGGCACGTTTCAGCGCCGCCGGATC GCTGCTGCCACCGCCGGTACCTTCACCGCTCAGGACAAT **GCN4_rev1** CGTTGCAGTTTACGCGCACGAGAACGGCGG GCGGCTTCGGTGTTACGGGCACGTTTCAGC	GeNL
1.1	**mNG_NdeI_Fwd** GGAGACATATGGTGTCCAAGGGCGAAGAGG	**GCN4_Rev2** TTTCGACAGCAGTTCTTCAACCTTGTCTTC CAGTTGTTTCATGCGTTGCAGTTTACGCGC	Product of PCR 1.0 diluted 1:50
2.0	**nLuc51_GCN4_Fwd** TTGCCCGCCTGAAAAAACTGGTTGGCGAAC GCGGTGCTAGCAGCGGAGAAAATGGGCTGAAGATCGACAT	**nLuc_NotI_Rev** TCTGCGGCCGCCAGAATGCGTTCGC	GeNL
2.1	**GCN_Fwd** GTTGAAGAACTGCTGTCGAAAAATTATCAC CTGGAAAATGAAGTTGCCCGCCTGAAAAAA	**nLuc_NotI_Rev** TCTGCGGCCGCCAGAATGCGTTCGC	Product of PCR 2.0 diluted 1:50
3	**mNG_NdeI_Fwd** GGAGACATATGGTGTCCAAGGGCGAAGAGG	nLuc_NotI_Rev TCTGCGGCCGCCAGAATGCGTTCGC	Equimolar mix of PCR 1.1 and PCR 2.1
4	**nLuc_6AALink_For** TCTGGCGGCCGCAGGTGCAAGCTCTGAAGATTTCGTTGGGGACTGG	**nLucCP50_XhoI_Rev** GGTGGTGCTCGAGCCCATTTTCACCGCTCAGGAC	GeNL
5	**nLuc_2AALink_For** TCTGGCGGCCGCAGAAGATTTCGTTGGTGACTGG	**nLucCP50_XhoI_Rev** GGTGGTGCTCGAGCCCATTTTCACCGCTCAGGAC	GeNL
6	**nLuc_4AALink_For** TCTGGCGGCCGCAAGCTCTGAAGATTTCGTTGGTGACTGG	**nLucCP50_XhoI_Rev** GGTGGTGCTCGAGCCCATTTTCACCGCTCAGGAC	GeNL
7	**nLuc_10AALink_For** TCTGGCGGCCGCAAGCTCTGCGGGTGGTGC AAGCTCTGAAGATTTCGTTGGGGACTGG	**nLucCP50_XhoI_Rev** GGTGGTGCTCGAGCCCATTTTCACCGCTCAGGAC	GeNL
Screen	**mNG_SeqFwd** CAAGACCGAGCTGAAGCACTCC	**T7Terminator** GCTAGTTATTGCTCAGCGG	Bacterial clones

Underlined sequences are those that bind the template. Use 60°C as the annealing temperature and set the thermocycler lid to 100°C.

Concentrations of all oligonucleotides are 0.5μM in the reactions except where noted.

Table 2

Sequences of DNA tools and target DNA.

Hairpin 1	gcgcgacattccgaagaacgctgaatcGATGACTCATCtcagcgttcttcggaatg
Hairpin 2	gcgcgacattccgaagaacgctgaGATGAGTCATCgattcagcgttcttcggaatg
N1 target	tcagcgttcttcggaatgtcgcgc
Serotonin aptamer	gtcgtcccGATGACTCATCgggacgactggtaggcagataggggaagctgattcgatgcgtgggtcgtccc
Naked AP-1 (serotonin)	tagactGATGAGTCATCggga
ATP aptamer	acctgggggagtattgcggaggaaggtGATGAGTCATCaccttcct
Naked AP-1 (ATP)	aaggtGATGAGTCATCa

Nucleotides are colored according to their structures in Fig. 4. The AP-1 sequence is capitalized.