# Toehold-Mediated Strand Displacement in RNA-Based Molecular Switches and Gene Regulation

Dissertation at Physics of Synthetic Biological Systems E14, School of Natural Sciences, Technische Universität München

Tianhe Wang

TECHNSCHE UNIVERSITÄT MÜNCHEN
TUM SCHOOL OF NATURAL SCIENCES

# Toehold-Mediated Strand Displacement in RNA-Based Molecular Switches and Gene Regulation

Tianhe Wang

Vollständiger Abdruck der von der TUM School of Natural Sciences der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigten Dissertation.

Vorsitz: Prof. Dr. Martin Zacharias

Prüfer der Dissertation:

1. Prof. Dr. Friedrich C. Simmel
2. Prof. Dr. Matthias Rief

Die Dissertation wurde am 25.10.2023 bei der Technischen Universität München eingereicht und durch die TUM School of Natural Sciences am 08.12.2023 angenommen.

# Table of Content

# Summary

Nucleic acids, as key biopolymers, are not only responsible for the storage and propagation of genetic information but also play crucial roles in gene regulation, synthesis, and metabolism. Strand dynamic interactions, such as dynamic binding and strand displacement (SD), are essential processes to achieve these biological functions. SD occurs when a single strand invades a double-strand duplex and forms a new heteroduplex, and it is prevalent in various biological contexts, such as homologous recombination, rolling circle replication, and CRISPR guide RNA binding, among others. These processes are commonly found in DNA repair, replication, and post-transcriptional regulation, contributing to the maintenance of genetic integrity and the regulation of gene expression.

Beyond in essential biological function, strand displacement has also emerged as a powerful tool in DNA nanotechnology for developing dynamic nanodevices. Toehold-mediated strand displacement (TMSD) reaction is one of the most commonly used SD topologies in DNA nanotech. It is driven by a toehold, which is a single-stranded overhang domain on the target strand that inhibits the detachment of the Invader strand during a branch migration process. Previous studies have commonly utilized bulk fluorescence measurements to study the kinetics of TMSD. However, these measurements only provide effective, bulk-averaged reaction rates and limit our ability to fully understand the underlying heterogeneity and fluctuations that might exist at the single-molecule level.

In our first chapter, we explored of the dynamic branch migration process of a hairpin TMSD a single-molecule resolution via single molecule force spectroscopy (SMFS). The strategic Introduction of mismatches within the invader stand enabled us to quantitatively evaluate the equilibrium between invasion/reinvasion during the branch migration process and estimate the force-biased single step transition time under certain force conditions. The observation of sequence dependent effects on the intermediate state during toehold hairpin unzipping and branch migration process, which is instructional to the design of RNA riboregulators. This study of force-biased branch migration process can be useful for understanding of the above-mentioned enzyme-based SD process which influences DNA or RNA strands under similar force condition. Notably, our investigation unveiled the sequence dependent effects on intermediate states during the unzipping of the toehold hairpin and branch migration process. These findings hold strong instructional value for the rational design of RNA riboregulators. Furthermore, our comprehensive analysis of force-biased branch migration holds implications for the future study, offering insights into enzymatic strand displacement processes that occur under analogous force conditions.

Moreover, TMSD has garnered significant attention due to its wide-ranging applications in the development of riboregulators for gene regulation within living cells. In the following chapters, we initially employed a previously established toehold switch to detect the expression levels of

endogenous mRNAs involved in stress responses within *Escherichia*. *coli*. However, we didn't observe an expected reaction profile under either starvation or SOS response due to the leakage from switch RNA. To extend the application of TMSD in gene regulation, we developed a set of riboswitch-inspired riboregulators in *Escherichia coli* that combine the concept of TMSD with the switching principles of naturally occurring transcriptional and translational riboswitches. Specifically, for translational activation and repression, we sequestered anti-anti-RBS or anti-RBS sequences, respectively, inside the loop of a stable hairpin domain, which is equipped with a single-stranded toehold region at its 5' end and is followed by regulated sequences on its 3' side. A trigger RNA binding to the toehold region can invade the hairpin, inducing a structural rearrangement that results in translational activation or deactivation. We also demonstrate that TMSD can be applied in the context of transcriptional regulation by switching RNA secondary structure involved in Rho-dependent termination. Our designs expand the repertoire of available synthetic riboregulators by a set of RNA switches with no sequence limitation, which should prove useful for the development of robust genetic sensors and circuits.

Additionally, we utilized the above design principles discussed earlier to engineer switchable fluorescent light-up aptamers (FLAPs). These FLAPs, can be activated or repressed by trigger RNA or small metabolites, contain guanine (G) quadruplexes forming a fluorophore binding pocket where critical nucleotides can be sequestered by anti-FLAP sequences, rendering an inactive conformation for binding with the fluorophore. Several designs exhibit minimal leakage and high ON/OFF fluorescence ratios. Purine aptamers were also modified to sequester anti-FLAP sequences, controlling fluorogen-binding conformation. Thus, FLAP fluorescence can be activated or deactivated by guanine or adenine. Combining switching modules generates FLAPs with fluorescence responding to multiple inputs and various input logics. Our switchable FLAPs show promising potential for applications in *in vitro* RNA diagnosis, offering the advantages of rapid and cost-effective detection methods.

# Zusammenfassung

Nukleinsäuren sind als wichtige Biopolymere nicht nur für die Speicherung und Verbreitung genetischer Informationen verantwortlich, sondern spielen auch eine entscheidende Rolle bei der Genregulation, Synthese und dem Stoffwechsel. Dynamische Strangwechselwirkungen wie dynamische Bindung und Strangverdrängung (Strand displacement, SD) sind wesentliche Prozesse zur Erreichung dieser biologischen Funktionen. SD tritt auf, wenn ein Einzelstrang in einen Doppelstrang-Duplex eindringt und einen neuen Heteroduplex bildet, und kommt in verschiedenen biologischen Kontexten vor, wie unter anderem bei der homologen Rekombination, der Rolling-Circle-Replikation und der CRISPR-Guide-RNA-Bindung. Diese Prozesse finden sich häufig bei der DNA-Reparatur, -Replikation und der posttranskriptionellen Regulierung und tragen zur Aufrechterhaltung der genetischen Integrität und zur Regulierung der Genexpression bei.

Über ihre wesentlichen biologischen Funktionen hinaus hat sich die Strangverdrängung auch zu einem leistungsstarken Werkzeug in der DNA-Nanotechnologie für die Entwicklung dynamischer Nanogeräte entwickelt. Die Toehold-vermittelte Strangverdrängung (Toehold-mediated strand displacement, TMSD) ist eine der am häufigsten verwendeten SD-Topologien in der DNA-Nanotechnologie. Es wird von einem 'Toehold' angetrieben, einer einzelsträngigen Überhangdomäne am Zielstrang, die während eines Verzweigungsmigrationsprozesses das Ablösen des Invader-Strangs verhindert. Frühere Studien verwendeten üblicherweise Massenfluoreszenzmessungen, um die Kinetik von TMSD zu untersuchen. Allerdings liefern diese Messungen nur effektive, massengemittelte Reaktionsgeschwindigkeiten und schränken unsere Fähigkeit ein, die zugrunde liegende Heterogenität und Schwankungen, die auf Einzelmolekülebene existieren könnten, vollständig zu verstehen. In unserem ersten Kapitel befassten wir uns mit der Untersuchung des dynamischen Zweigmigrationsprozesses des TMSD auf Einzelmolekülebene mithilfe der Single molecule force spectroscopy (SMFS). Die strategische Einführung von Nichtübereinstimmungen innerhalb des Eindringlingsstandes ermöglichte es uns, das Gleichgewicht zwischen Invasion/Reinvasion während des Zweigmigrationsprozesses quantitativ zu bewerten und die kraftabhängige Übergangszeit in einem Schritt unter bestimmten Kraftbedingungen abzuschätzen. Die Beobachtung sequenzabhängiger Effekte auf den Zwischenzustand während des Entpackens der Zehenhaarnadel und des Astmigrationsprozesses, die für das Design von RNA-Riboregulatoren von Bedeutung sind. Diese Untersuchung des kraftabhängigen Zweigmigrationsprozesses kann für das Verständnis des oben erwähnten enzymbasierten SD-Prozesses nützlich sein, der DNA- oder RNA-Stränge unter ähnlichen Kraftbedingungen beeinflusst. Insbesondere enthüllte unsere Untersuchung die sequenzabhängigen Auswirkungen auf Zwischenzustände während der Entpackung des Haarnadel- und Zweigmigrationsprozesses. Diese Erkenntnisse sind von

großem Lehrwert für das rationale Design von RNA-Riboregulatoren. Darüber hinaus bietet unsere umfassende Analyse der kraftbedingten Zweigmigration Implikationen für die zukünftige Studie und bietet Einblicke in enzymatische Strangverdrängungsprozesse, die unter analogen Kraftbedingungen ablaufen.

Darüber hinaus hat TMSD aufgrund seiner weitreichenden Anwendungen bei der Entwicklung von Riboregulatoren für die Genregulation in lebenden Zellen große Aufmerksamkeit erregt. In den folgenden Kapiteln verwendeten wir zunächst einen zuvor etablierten Zehenschalter, um die Expressionsniveaus endogener mRNAs zu ermitteln, die an Stressreaktionen in *E. coli* beteiligt sind. Aufgrund des Austretens von Switch-RNA beobachteten wir jedoch weder unter Hunger noch unter SOS-Reaktion das erwartete Reaktionsprofil. Um die Anwendung von TMSD in der Genregulation zu erweitern, haben wir eine Reihe von Riboschaltern inspirierter Riboregulatoren in Escherichia coli entwickelt, die das Konzept von TMSD mit den Schaltprinzipien natürlich vorkommender Transkriptions- und Translations-Riboschalter kombinieren. Insbesondere für die Aktivierung und Repression der Translation haben wir Anti-Anti-RBS- bzw. Anti-RBS-Sequenzen innerhalb der Schleife einer stabilen Haarnadeldomäne sequestriert, die an ihrem 5'-Ende mit einer einzelsträngigen Halteregion ausgestattet ist und verfolgt wird durch regulierte Sequenzen auf seiner 3'-Seite. Eine Trigger-RNA, die an die Toehold-Region bindet, kann in die Haarnadel eindringen und eine strukturelle Neuordnung auslösen, die zu einer translatorischen Aktivierung oder Deaktivierung führt. Wir zeigen auch, dass TMSD im Kontext der Transkriptionsregulation angewendet werden kann, indem die RNA-Sekundärstruktur, die an der Rho-abhängigen Termination beteiligt ist, verändert wird. Unsere Entwürfe erweitern das Repertoire verfügbarer synthetischer Riboregulatoren um eine Reihe von RNA-Schaltern ohne Sequenzbeschränkung, die sich für die Entwicklung robuster genetischer Sensoren und Schaltkreise als nützlich erweisen dürften.

Darüber hinaus nutzen wir die oben besprochenen Designprinzipien, um schaltbare “fluorescent light-up aptamers” (FLAPs) zu entwickeln. Diese FLAPs können durch Trigger-RNA oder kleine Metaboliten aktiviert oder unterdrückt werden und enthalten Guanin (G)-Quadruplexe, die eine Fluorophor-Bindungstasche bilden, in der kritische Nukleotide durch Anti-FLAP-Sequenzen abgesondert werden können, wodurch eine inaktive Konformation für die Bindung mit dem Fluorophor entsteht. Mehrere Designs zeichnen sich durch minimale Leckage und hohe AN/AUS-Fluoreszenzverhältnisse aus. Purin-Aptamere wurden auch so modifiziert, dass sie Anti-FLAP-Sequenzen binden und so die Fluorogen-bindende Konformation steuern. Somit kann die FLAP-Fluoreszenz durch Guanin oder Adenin aktiviert oder deaktiviert werden. Durch die Kombination von Schaltmodulen werden FLAPs mit Fluoreszenz erzeugt, die auf mehrere Eingaben und verschiedene Eingabelogiken reagieren. Unsere schaltbaren FLAPs zeigen vielversprechendes Potenzial für Anwendungen in der In-vitro-RNA-Diagnose und bieten die Vorteile schneller und kostengünstiger Nachweismethoden.

# Publications

Some ideas and figures have appeared previously in the following publications:

T. Wang and F.C. Simmel, Riboswitch-inspired toehold riboregulators for gene regulation in *Escherichia coli Nucl. Acids Res. 50, 4784-4798 (2022)*

T. Wang, H. Hellmer, F. C. Simmel, Genetic switches based on nucleic acid strand displacement, Curr. Op. Biotechnol. 79, 102867 (2023)

T. Wang and F.C. Simmel, Switchable fluorescent light-up aptamer based on riboswitch architecture Angew. Chem. Int. Ed. E202302858 (2023)

# Chapter I

## Single-Molecule Force Spectroscopy of Toehold-Mediated Strand Displacement

**Key words:** Riboregulators • toehold mediated strand displacement • branch migration • single molecule force spectroscopy

We performed the collaboration work for this project with E22, TUM School of Natural Sciences, Technical University of Munich, Germany, and Prof. Petr Šulc's lab in School of Molecular Sciences and Center for Molecular Design and Biomimetics, the Bio-design Institute, Arizona State University. The experiment designs and measurements were carried out together with Andreas Walbrun. Data analysis and OxDNA simulations were mainly performed by Andreas Walbrun, Michael Matthies respectively. 1D random walk model is conducted by Prof. Simmel.

1

# Abstract

Toehold-mediated strand displacement (TMSD) has been a fundamental concept in DNA nanotechnology, serving as a key element for numerous dynamic DNA devices. Previous studies have typically relied on bulk fluorescence measurements to investigate the kinetics of TMSD, which only provide effective, bulk-averaged reaction rates, and do not resolve the process on the level of individual molecules or even base-pairs. In this chapter, we addressed this limitation by exploring the dynamics of strand displacement (SD) processes at the single-molecule level using single-molecule force spectroscopy (SMFS) with an optical trap. By introducing mismatches on the invader strand, we elucidated the equilibrium between invasion and reinvasion during branch migration, allowing subsequently estimates the force-dependent single-step time under specific force conditions.

The observed influence of sequence-dependent effects and potential secondary structure formation on the invader strand attribute to intermediate states during the unzipping of the toehold hairpin and the branch migration process, which provides valuable insights for the design of RNA riboregulators. Furthermore, we present a pioneering exploration of the force-biased SD between DNA invader and RNA hairpin, enabling the estimation of the single step times during branch migration under force. This investigation not only advances our understanding of force-biased branch migration but also offers potential insights into enzyme-based strand displacement processes, particularly when DNA or RNA strands experience similar force conditions.

# Introduction

Nucleic acid SD is a common phenomenon occurring when two DNA or RNA strands of similar or identical sequences attempt to bind with a complementary target strand. This mechanism holds immense significance in various biological processes such as DNA recombination[1], CRISPR-based target recognition[2], and RNA-based gene regulation[3]. During a typical SD reaction, an incumbent strand is initially bound within a target duplex. Temporal fraying of the duplex can expose unbound nucleotides at its ends, allowing a single-stranded invader to attach. Binding of invader strand forms a three-stranded intermediate complex. Subsequently, the invader and incumbent strands engage in a competition for binding to the target, facilitated by a process known as branch migration[4]. Introducing a short single-stranded region, often referred to as a "toehold," into the sequence of the target strand can facilitate the strand displacement reaction in one direction. This toehold region enables the invader to initiate binding and subsequent displacement. Toehold-mediated strand displacement (TMSD) reactions entail the use of an invader strand that is comparatively longer than the incumbent. This elongation facilitates the formation of a longer base-pair with the target. As a result, the incumbent strand is consistently displaced by the invader in TMSD reactions.

Toehold-mediated strand displacement (TMSD) has found wide-ranging applications in the realm of nucleic acid nanotechnology, which extends to modulating the conformation of DNA-based molecular assemblies, propelling molecular machines, and executing computations within intricate chemical reaction networks[5,6]. Combining concepts from both DNA nanotechnology and synthetic biology, TMSD mechanisms have been harnessed to govern the behavior of functional RNA entities in living organisms. This innovation has led to the creation of conditional CRISPR guide RNAs[7-10] and programmable RNA-responsive riboregulators[11-14], exhibiting remarkable ON/OFF ratios.

The kinetics underlying strand displacement reactions have predominantly been explored using both bulk methodologies and computational modeling. Early studies performed the examination of branch migration phenomena within recombination intermediates present in bacteriophage T4[15], as detailed by Broker. Initial endeavors to understand the kinetics of single-strand branch migration hinged on the utilization of extensively radioactively labeled DNA fragments[16-18]. Subsequently, this investigation was advanced through Förster resonance energy transfer (FRET) analyses, facilitated by the incorporation of fluorescently labeled oligonucleotides[19].

Bulk FRET assays have found widespread application within the field of DNA nanotechnology. These experiments have been utilized to determine the impact of variables like toehold length sequence mismatches, and secondary structure effects on the kinetics of TMSD reactions[20-22]. Nonetheless, it's important to acknowledge that bulk measurement can only provide insights into the overall behavior of the reaction, including toehold binding, and are not suitable for directly visualizing the strand displacement process itself.

Single-molecule techniques provide unique insights into individual molecular reactions, offering a level of detail that is inaccessible through bulk measurements, where observations are inherently averaged. Notably, single-molecule force spectroscopy (SMFS), employing optical or magnetic tweezers, has enriched our comprehension of the folding dynamics of biomolecules, encompassing proteins and nucleic acids. Furthermore, SMFS has found extensive utility in elucidating the force-related behaviors of molecular machines, including polymerases[23] and ribosomes[24,25], while they interact with these biomolecular structures. SMFS has been extensively employed to investigate the sequence-dependent folding of RNA hairpins[26-28] and riboswitch aptamers[29], offering intricate insights into their folding energy landscapes and the influence of mismatches during folding. In addition, single-molecule supercoiling experiments conducted with magnetic tweezers have been instrumental in examining R-loop formation by CRISPR-Cas nucleases as they interact with DNA target duplexes[30,31]. These experiments have enabled a quantitative understanding of the target recognition process. Moreover, magnetic tweezers have contributed to our understanding of the kinetics and mechanics of strand displacement, shedding further understanding on how short oligonucleotides are displaced from a closing DNA hairpin under mechanical strain[32].

In this study, we deployed an optical trap setup to investigate the toehold-mediated invasion of DNA or RNA molecules into the stem of nucleic acid hairpin structures containing a complementary sequence. In contrast to traditional TMSD experiments involving three distinct strands (invader, incumbent, and target strands), in our setup, incumbent and target sequences are covalently linked within the hairpin loop (Fig. 1). The invasion of the hairpin stem results in the unfolding of the hairpin structure but does not lead to the release of the incumbent strand. This structure allows us to observe multiple invasions, unfolding, and refolding processes within the same molecule.

Toehold-mediated invasion of a hairpin, as explored in this study, forms the basis of the mechanism behind recently developed toehold switch riboregulators[12-14,33]. These riboregulators feature an unpaired toehold sequence (14 nt long) at the 5’ end, followed by a stable hairpin-like secondary structure that effectively sequesters a ribosome binding site (RBS) of mRNA within loop region. In the absence of specific RNA invader molecules, referred to as "triggers" in this context, translation of such an mRNA is suppressed. However, when trigger RNA is introduced, it can bind to the toehold region, initiating a strand displacement process. This process unfolds the RNA hairpin structure, revealing the RBS and consequently enabling mRNA translation. Numerous design attributes of toehold switches, including stem length, loop size, and the presence of mismatches or bulges, have been methodically investigated and were observed to exert a significant impact on the functionality of these engineered gene regulators[34,35].

In our work, we investigated toehold hairpin structures featuring the same 14 nt toehold sequence as characterized in an original toehold switch[11]. To enable direct observation of the invasion process in SMFS, we made specific modifications to the hairpin sequence. These modifications involved removing the interior bulge, introducing well-defined mismatches in the trigger strand,

and extending the stem region of the hairpin. Through a combination of coarse-grained simulations using OxDNA[36], our research offers detailed insights into the strand invasion process, achieving close-to-single-nucleotide resolution for DNA-DNA, RNA-RNA, and RNA-DNA strand invasions. This approach allows us to uncover the significant sequence-dependent dynamics associated with strand invasion.

# Results

Force-induced unfolding of DNA and RNA hairpins observed by SMFS

To visualize the process of strand invasion within nucleic acid hairpins through SMFS, we engineered molecular constructs that enabled the attachment of these hairpins to 1 µm-diameter silica beads. These silica beads could be trapped and manipulated using two infrared laser beams within a commercial optical trapping setup (Fig. 1, C-Trap® LUMICKS; for detailed methods, refer to Materials and methods). Our toehold hairpins contain 110 nt DNA or RNA sequences, which are taken the original toehold switch riboregulatory[11]. They consisted of a 14 nt single-stranded toehold segment linked to a 52 bp long stem via a 6 nt loop region (refer to Appendix Fig. S1). As depicted in Fig. 1 a, the 5' and 3' ends of these hairpins were joined to two 185 nm long double-stranded DNA handles that assisted their connection to the silica beads. This connection was achieved through biotin/streptavidin and digoxigenin/anti-digoxigenin linkers. Given that the DNA handles featured overhangs on the 5' end, a single-stranded adapter strand (grey in Fig. 1 a) was employed to connect the 3' end of the hairpin to the handle. Single-molecule experiments were conducted within a microfluidic flow chamber configuration, which allowed a single molecule suspended between the two traps to be exposed to different mobile phase by manipulating the traps within distinct laminar flow (Fig. 1, 2 c). In a typical measurement, two beads were confined within the bead channel, while the toehold hairpin was suspended, creating a dumbbell assay in the buffer channel. Subsequent measurements were carried out either in a buffer-only channel or in the trigger channel containing 100 nM trigger strand.

To understand the mechanical behavior of the hairpin, we initiated single molecule unzipping experiments on both the DNA and RNA hairpin configurations in the absence of trigger molecules. The results, obtained from stretch/relax cycles conducted at a pulling velocity of 0.2 µm/s, are illustrated in (Fig. 2 a & b). During the stretching phase, we observed rapid transitions representing the progressive unfolding of the hairpin structure. Several intermediate states (I1, I2, and I3) were identified on the pathway from the fully folded (Fol) state to the fully unfolded (Unf) state. To determine the number of nucleotides involved in each transition, we fitted the experimental data using a worm-like chain (WLC) polymer model[37] (colored lines in Fig. 2 a and b, Table S1). Notably, while RNA exhibited significantly higher unfolding forces compared to DNA, the length changes associated with transitioning between each intermediate were virtually identical for both DNA and RNA hairpins within the resolution of our experiment. Further details regarding the sequence characteristics of the folded and unfolded segments of these intermediates can be found in Fig. S2. Upon relaxation, the molecule promptly reverted to the fully folded state (light grey traces in Fig. 2 a & b).

To obtain a more detailed kinetic and energetic characterization of the unfolding-refolding equilibrium transitions, we also performed so-called passive mode experiments where we kept the distance between the lasers constant while observing the fluctuations of the molecule through its intermediate states over tens of seconds (Fig. S2). For DNA, a zoom into the data shows 5 different populated levels corresponding to the intermediates mentioned above (Fig. 2 a). Assignment and coloring of the states was done using hidden-Markov-modeling (HMM)[38] (See Data analysis). From passive mode data, we calculated the folding free energy of the hairpin from the ratio of the population probabilities of the folded and unfolded state, correcting for energetic contributions from stretching the linkers and spring energies from beads deflected from the trap center (see Data analysis for details). We find a folding free energy of -92.8 $k_{BT}$ (-54.9 kcal/mol) which is in reasonable agreement with the values predicted by nucleic acid thermodynamics software packages (NUPACK (DNA: -118.88 $k_{BT}$ (-70.50 kcal/mol), mFold: -118.93 $k_{BT}$ (-70.53 kcal/mol)). Deviations can be explained by systematic errors in the force calibration of the tweezers. The kinetics of RNA folding/refolding is significantly slower as compared to DNA (Fig. S2 D). This slow kinetics precludes observation of RNA unfolding at equilibrium. While equilibrium transitions can be observed between intermediates FoI, I1, I2 and I3, the construct stays permanently unfolded as soon as the fully opened state is reached. Table S2 summarizes free energy values for states accessible in our experiments and from software packages.

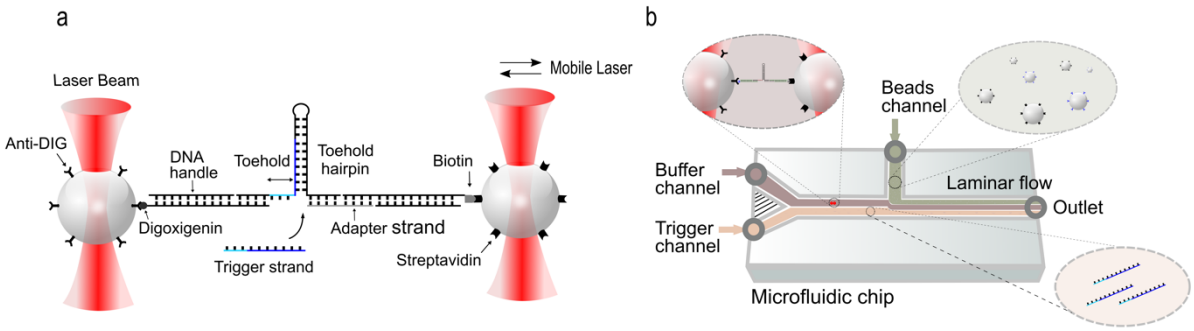


Figure 1. Measurement setup and measurement principle. a), A dual-beam optical trap setup is utilized to measure the force-dependent hairpin opening promoted by a trigger strand in a toehold-mediated strand displacement (TMSD) process. DNA (or RNA) hairpin molecules are anchored between two beads using 545 base pair (bp) long DNA handles, as indicated (see legend for details). b), Schematic of microfluidics setup employed for the TMSD measurements is illustrated. DNA (or RNA) constructs with attached beads are incubated and subsequently pumped into the microfluidic device. The bead mix is pumped into the 'Beads channel' along with the running buffer channel (containing 20 mM $MgCl_2$, 300 mM KCl, and 50 mM HEPES) and the 'Trigger channel', which also contains the same buffer but with the addition of 100 nM trigger strand. Laminar flow separates the running buffer from the buffer containing the trigger strand. The bead pairs are trapped within the 'Beads channel' and then transferred to the buffer channel to perform 'ditching (Binds strand to bead),' where two beads are brought into close proximity to create a tether between the DNA handle and anti-Digoxigenin beads. Once the tether is formed, the beads are moved to the trigger channel to carry out TMSD measurements.

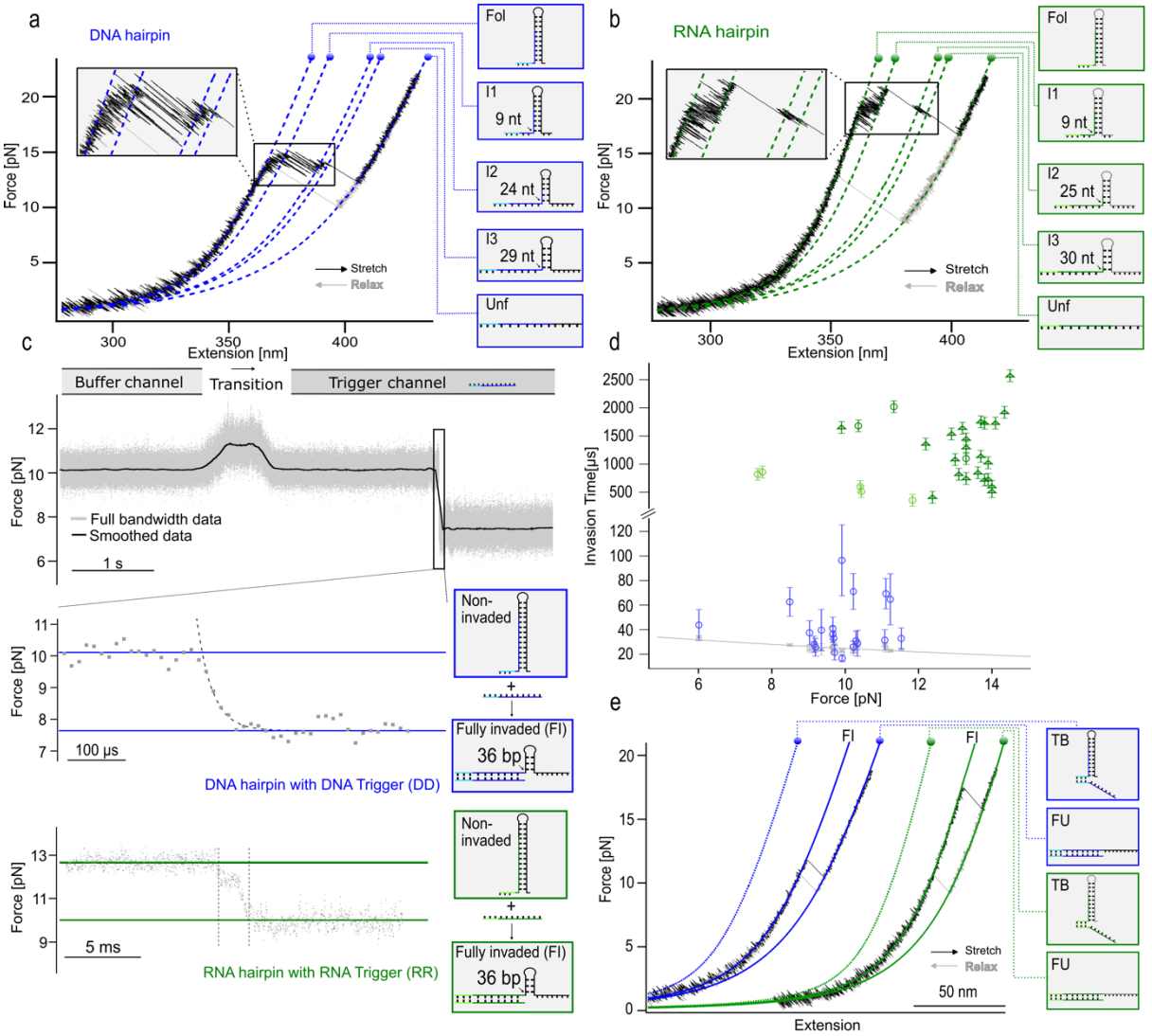


Figure 2. Toehold hairpin unzipping and strand invasion. a), Representative force-extension curves (stretching: black, relaxation: grey) are shown for hairpin unzipping in the absence of trigger strands, using a pulling velocity of 0.2 µm/s and [$Mg^{2+}$] concentration of 20 mM. The colored curves are fits based on a worm-like chain (WLC) model, identifying different intermediate states of the hairpin's conformation. The inset shows transitions between the different conformations in more detail, b), Similar to (a), but for RNA hairpin. c), Passive mode traces for a DNA (upper trace with zoom) and RNA hairpin (lower trace) when transitioning from the buffer to the trigger channel ([trigger] = 0.1 µM). Trigger binding occurs after a lag time of approximately 2 seconds, followed by toehold-mediated strand invasion, visible as a force drop (from $\approx$ 10 pN to 8 pN for DNA and $\approx$ 13 pN to 10 pN for RNA). The exponential fit (decay time $\tau$ = 26 µs) provides an upper limit for the time required for the branch migration process. d, Invasion time plots for various toehold hairpin constructs. Grey cross: DNA toehold hairpin in absent of trigger invasion, with an exponential fit illustrating invasion times as a function of applied force. Blue circle: DNA hairpin with trigger invasion. Dark green circle: RNA hairpin with trigger. Light green circle: RNA hairpin with trigger invasion in a 300 mM KCl environment. Dark green triangle: RNA hairpin with two proximal mismatches with the trigger. Error bars represent the standard deviation (s.d.) from exponential fit. e, Representative force-extension curves for the trigger-hairpin complex after TMSD. Colored curves correspond to fits of the WLC model to the data between the toehold bound (TB) and fully invaded state (FU).

Single-molecule observation of toehold-mediated strand displacement (TMSD)

To conduct single-molecule TMSD experiments using our setup, we leveraged the microfluidic capabilities of the optical trap. We applied a constant force, gradually balancing it to keep the hairpin in a folded state yet sufficiently high to yield a detectable signal during strand invasion. As the molecule transitioned from the buffer channel to a channel containing 100 nM of trigger strand molecules (Fig. 1 b), we observed a distinct change in trapping force. This abrupt reduction in force was attributed to the binding of a trigger strand to the toehold hairpin, followed by the subsequent invasion of the hairpin stem. The invasion process led to the elongation of the hairpin and was reflected in this sudden force drop (Fig. 2 c). In cases where the trigger strand was fully complementary to the hairpin sequence, binding and invasion events occurred nearly simultaneously, making it challenging to distinguish between the two processes.

For DNA hairpin, the strand invasion transitions induced by an invader happened rapidly and cooperatively, typically taking place within a timeframe of $\approx$ 10 – 100 µs (mean: 41.9 ± 4.7 µs) (refer to the zoomed-in trace in Fig. 2 c, middle trace). In contrast, RNA exhibited a significantly longer invasion process, featuring a distinct intermediate state of approximately 2 ms in length (mean: 1.28 ± 0.11 ms) (see Fig. 2 c, bottom). The relationship between measured strand invasion times and the applied force is illustrated in Fig. 2 d.

In the case of DNA (represented by blue circles), the measured transition times closely approximated the response time of our instrument (indicated by grey symbols and the fitted line), as determined by analyzing bead relaxation in water through autocorrelation analysis[39] (see Data analysis). This suggests that, under the forces applied in our experiment, the entire strand invasion process, spanning 36 base pairs, occurs in less than 41.9 µs. Given the high forward biasing forces, the likelihood of reinvasion (backward) steps is minimal. Consequently, our findings propose an upper limit of 1.16 µs (41.9 µs / 36) per invasion step at forces around 10 pN.

For RNA, the invasion times average to approximately 1.28 ± 0.11 ms, as indicated by the dark green symbols, when in a buffer containing 20 mM magnesium chloride. These measured invasion times are primarily influenced by the presence of an intermediate state (sample traces are provided in Fig. S3). In the absence of $MgCl_2$, the intermediate state is still present, albeit with shorter lifetimes, leading to an average total invasion time of 630 ± 107 µs.

The dark green triangle symbols represent measurements conducted with a trigger strand containing 2 proximal mismatches (RRp2), specifically an RNA hairpin with an RNA trigger that has two proximal mismatches. This contrasts with experiments involving the fully complementary sequence. The introduction of these mismatches was done to rebalance the equilibrium towards reinvasion, thereby increasing the number of data points that could be obtained with a single molecule. It's worth noting that the duration of the invasion event is not significantly affected, as it still involves passing through 34 out of the 36 base pairs. Interestingly, we observed that invasion times remained largely independent of force, as depicted in Fig. 2 f. This suggests that the intermediate state pausing the invasion process may be attributed to secondary structure

formation in the trigger strand, which is not influenced by mechanical force. The observation of potential secondary structure in the trigger strand was unexpected, especially considering that the sequence had been optimized to minimize secondary structure formation using NUPACK.

Direct observation of TMSD kinetics under force

For both DNA and RNA, the hairpin stayed permanently open after the invasion and even a drastic reduction of force would not lead to reversal of the invasion. Consequently, stretch/relax cycles in the buffer channel with the trigger-bound complex (Fig. 2 e) show the molecule always in the fully invaded state (FI) indicating that re-invasion does not happen at non-zero force values. The additional unfolding transition we observe at ca. 10 pN for DNA merely reflects unfolding of the remaining hairpin from the FI to the fully unfolded state (FU).

To observe repeated invasion and re-invasion steps close to thermodynamic equilibrium, we intentionally introduced sequence mismatches into the center of the branch migration domain of the invader strand, denoted by the red base in Fig. 3 a[40,41]. This mismatch was designed to slow down the transition between the toehold-bound (TB) and the fully invaded (FI) state by imposing an energy penalty on invasion, while leaving reinvasion unaffected. Consequently, this energy penalty introduced an additional intermediate state at the mismatch position (IM), allowing for reversible strand invasion and reinvasion to proceed towards the FI state.

Compared to a fully complementary trigger strand (as shown in Fig. 2 e and Fig. 3, 1st trace), when using a trigger sequence with a single G→T mismatch at position 19 of the branch migration domain b' (designated as of DNA-DNA with 1 center mismatch, that is DDc1, for sequence details, refer to Fig. S1), we observed the expected rapid invasion and re-invasion equilibrium occurring at forces around 3-4 pN (Fig. 3 b, 2nd trace, marked in blue). The analysis of the contour length changes corroborated the structural interpretation of the various states (please see Table S3, S4 in the SI for further details). Furthermore, introducing an additional mismatch on the trigger strand adjacent to position 19 (denoted as DDc2, for sequence details, refer to Fig. S1) shifted the force fluctuations observed from approximately 3-4 pN to around 5 pN (Fig. 3 b, 3rd trace, marked in purple).

Additionally, we explored the impact of a similar mismatch (G→U mismatch at position 19) in an RNA trigger strand during its invasion of an RNA toehold hairpin, denoted as RRc1 (for sequence details, please refer to Fig. S1). This mismatch caused the initiation of invasion to occur at even higher forces, around 10 pN. Furthermore, the kinetics of invasion and re-invasion were significantly slowed down to the extent that multiple invasion and re-invasion events were seldom observed within the timescale of our pulling experiments (Fig. 3 b, 4th trace, pulling velocity 0.2 µm/s, marked in green). As a consequence of this slower kinetics, the pulling and relaxation traces for RRc1 exhibited a pronounced hysteresis. The structure remained in the fully invaded (FI) state until the force was reduced to values as low as 3 pN. Traces for an RNA construct analogous to DDc2 (denoted as RRc2) can be found in Fig. S3. These traces demonstrated an

even more substantial hysteresis, with pulling forces as high as 14 pN required to induce branch migration (Fig. S7 B).

To quantitatively assess the kinetics of this repeated branch migration process, we conducted passive mode measurements at various forces for all three systems. Figure 3 c displays sample traces near F1/2 forces, where the intermediate (IM) and fully invaded (FI) states were each populated to 50%. It's evident that DDc2 experiences higher forces compared to DDc1, and RRc1 exhibits significantly slower kinetics. Sample traces for other forces are available in Fig. S6.

A plot of the invasion and re-invasion rates measured at different forces (Fig. 2 d, Fig. S7) allows extrapolation to zero force (Fig. 3 d, dashed lines). The zero-force rates reveal that the double mismatch slows down the invasion of DDc2 by an order of magnitude compared to DDc1 (0.0533 ± 0.015 $s^{-1}$ vs. 1.57 ± 0.12 $s^{-1}$), while re-invasion rates are less affected (1490 ± 390 $s^{-1}$ vs. 900 ± 60 $s^{-1}$). This pronounced effect on invasion rate be interpreted since the invader strand has to overcome a higher barrier when two base-pairs need to be broken invasion can move forward, as opposed to just one base-pair. The slightly lower extrapolated rates for re-invasion in the case of DDc1 may reflect the re-invasion for the double mismatches has one base-pair less to compete with the invader and hence occurs faster. A summary of all measured and extrapolated rates can be found in Table S5.

The extrapolated value for re-invasion at zero force provides an upper limit estimation for the speed of branch migration per base pair. We find re-invasion rates of approximately 1000/s (mean: 780 ± 30 $s^{-1}$ (DDc1) and 1480 ± 220 $s^{-1}$ (DDc2)) for branch migration across 17 (DDc1) or 16 (DDc2) bases. This suggests an upper limit of approximately 59 µs per single migration step. It's important to note that this calculation assumes linear scaling with the number of base pairs rather than diffusive scaling with $N^2$, which would result in significantly smaller values (see more details in the Appendix).

Comparing DDc1 and RRc1 (Fig. 3 d top vs. bottom), we observe that both extrapolated zero-force values for invasion (1.57 ± 0.12 $s^{-1}$ vs. 0.00905 ± 0.00319 $s^{-1}$) and re-invasion (900 ± 60 $s^{-1}$ vs. 226 ± 85 $s^{-1}$) are slower for the RNA construct. As discussed earlier, slower invasion rates can be explained by the larger energy barrier associated with an unpaired RNA base compared to a DNA base. The lower value obtained by extrapolating the re-invasion branch directly demonstrates that branch migration in RNA is considerably slower than in DNA[42], in agreement with the previous work. This estimate indicates that RNA branch migration is approximately 4.4 times slower (around 260 µs). For RRc2, conformational transitions recorded in passive mode experiments were too slow to allow an accurate estimate of the corresponding kinetic rates (Fig. S3 F). As the above mentioned RRp2, our SMFS experiments indicate that proximal mismatches, located at positions 15 and 16 close to the toehold (as shown in Fig. S1 and S4 b), do not appear to significantly affect the branch migration process between the intermediate (IM) and fully invaded (FI) states. This observation contrasts with previous studies on branch migration kinetics and related processes, such as R-loop formation by Cas proteins[30,31]. This effect could potentially

be attributed to the fact that proximal mismatches tend to destabilize the initial toehold complex between the invader and substrate, particularly when the toehold region is relatively short.

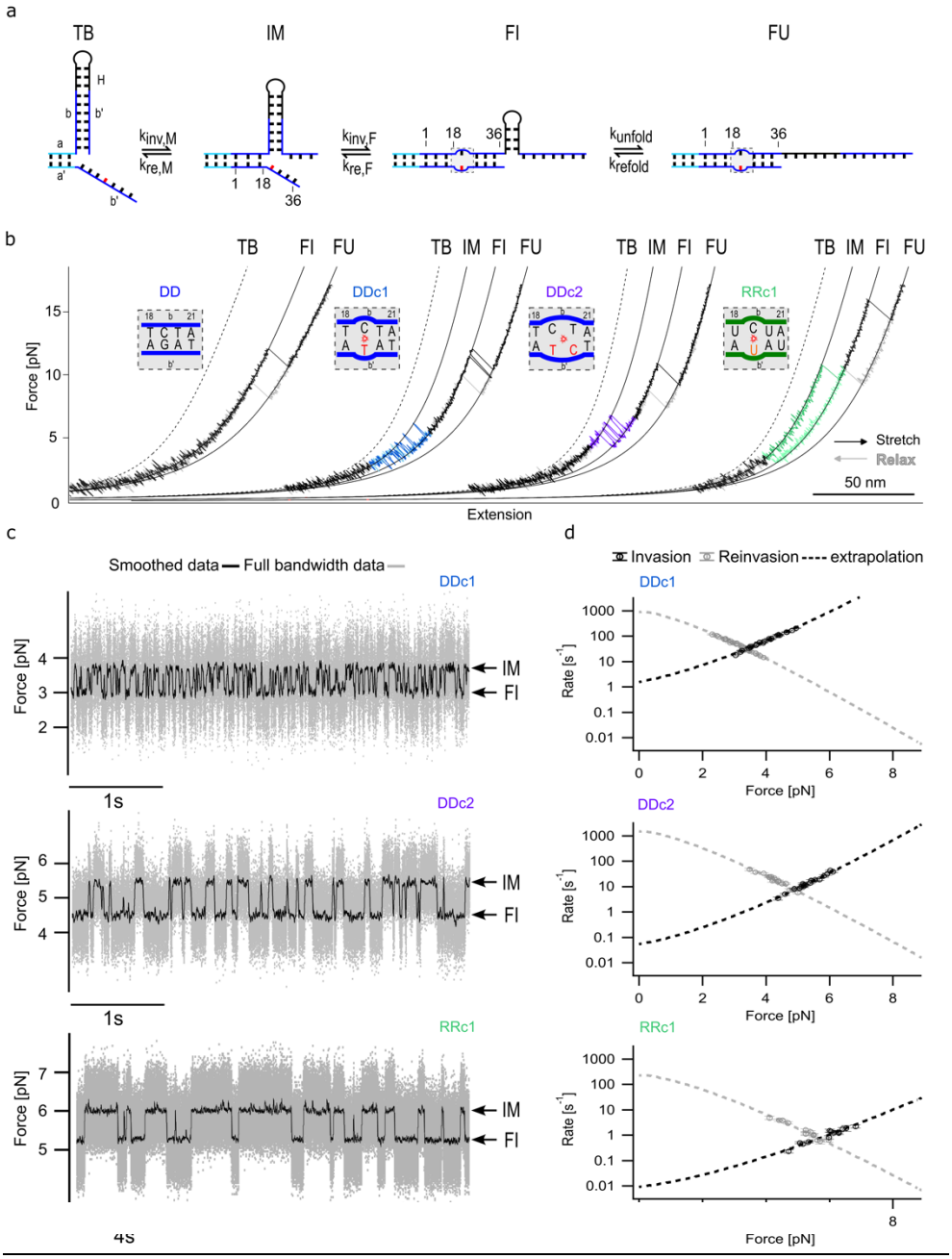


Figure 3. a), Conformational states during the invasion of a mismatched trigger strand into a toehold hairpin is illustrated. The states include TB (toehold-bound state), IM (intermediate state) where the invader is bound up to the mismatch position, FI (fully invaded state), and FU (fully unfolded state). The mismatch is highlighted in red (for DNA) and orange (for RNA). b), Force-extension traces show the behavior of a DNA toehold hairpin with a fully matched DNA trigger (DD, first trace), DNA trigger strands with one or two center mismatches (DDc1, second trace, and DDc2, third trace, respectively), and an RNA toehold hairpin with an RNA trigger strand with one center mismatch (RRc1, fourth trace). The insets provide details about the mismatch positions and sequences. c), Force-versus-time traces recorded at $F_{avg\ 1/2}$ for the mismatched trigger strands DDc1, DDc2, and RRc1, displaying several invasion/re-invasion transitions between the IM and FI states due to strand displacement. d), this graph

illustrates the force dependence of forward and backward transition rates between the IM and FI states. The dotted lines represent extrapolations of the data based on a theoretical model described in the Data analysis

## Simulation of force−extension curves and energy landscapes

To facilitate the interpretation of our experimental findings regarding force induced TMSD and to gain insights into the impact of mismatches in the invader sequence, we conducted simulations using the oxDNA model[43,44]. oxDNA is a coarse-grained DNA model that represents each nucleotide as a single rigid body and is parameterized to replicate the structural, thermodynamic, and mechanical properties of single-stranded and double-stranded DNA. This model has been extensively utilized in simulating strand displacement processes and has demonstrated good agreement with experimental results in various studies[40,45,46]. Additionally, it accurately reproduces how DNA responds mechanically to tension, making it a valuable tool for our investigations[47,48].

Our simulation process began by constructing DNA entities, which consisted of DNA handles, an adapter, DNA hairpin, and trigger strand, using the oxDNA package. Initially, we conducted molecular dynamics (MD) simulations without the presence of trigger DNA. In this phase, tension was applied to both ends of the DNA handle, pulling on the toehold hairpin at a constant velocity of 0.14 mm/s (Fig. 4 a, left). It's worth noting that the pulling velocity in the simulations was higher than the experimental rate, as achieving the exact experimental pulling rates was not feasible within MD simulations. Nevertheless, the force-extension curve calculated through our simulations exhibited hairpin unfolding transitions at approximately 22 pN, closely mirroring our experimental findings (Fig. 2 b). The snapshots generated from the simulated DNA molecule provided a visual representation of the intermediate states attained during the stretching process.

In our pulling simulations, we initiated the process with a fully complementary trigger strand (DD) bound to the toehold. As the trigger strand promoted strand invasion, the unfolding process was facilitated, resulting in a brief unfolding transition occurring at approximately 19 pN (Fig. 3 a, right). By examining the force-extension associated with each base pair along the trajectory, we identified the potential intermediate state that closely resembled the intermediate state observed in our experiments (Fig. 3 c, Fig. S10). Additionally, during the branch migration process in our MD, we observed the formation of secondary structures on the trigger strand (Data not shown). These secondary structures might contribute to the slower branch migration rate and the formation of intermediate states. We hypothesized that the presence of more stable secondary structures on RNA triggers could explain the slower branch migration rate observed during RNA SD compared to DNA SD (Fig. 2 d).

We then conducted MD simulations for the passive mode, applying a constant force ranging from 0 pN to 10 pN on the DNA handles (see Materials and Methods). For each pulling force, we performed at least 7 replicas of the system. Simulations with different trigger strand variants demonstrated that, in line with our experimental findings, increasing pulling forces generally

enhanced the overall kinetics of the branch migration process and favored complete displacement (Fig. S13). For instance, when a pulling force of 10 pN was applied, a two proximal mismatch, DDp2 could be crossed more easily compared to 1 lower forces.

The single-nucleotide resolution of the oxDNA simulation enables a more detailed qualitative understanding of the free-energy landscape when subjected to an externally applied force. However, performing free-energy sampling simulations is computationally intensive. Therefore, we simplified the system by reducing the length of both the toehold hairpin and the trigger strand (Fig. S10). We focused on a three-stranded system with a total of only 30 base pairs (Fig. 4 b) and applied constant forces ranging from 0 to 5 pN to the strands (As well as Fig. S12). We derived the free-energy landscapes from simulations, which depict the number of base pairs formed by the target strand with the trigger and the incumbent as a function of different force biases (Fig. 4 b). We utilized Virtual Move Monte Carlo (VMMC) algorithm and umbrella sampling to generate these landscapes[49] (see Materials and Methods).

From the free-energy landscapes, it's evident that as the force applied increases, states with more base pairs formed between the invader and the substrate become more favorable. We plotted a 2D free-energy landscape, representing the number of base pairs formed between the invader and substrate, as well as between the incumbent and substrate, reveals that higher applied forces favor states where the trigger is bound to the incumbent (Fig. S11).

We also examined the influence of mismatches on the free-energy landscape at different applied forces. Initially, we simulated a proximal mismatch which introduces a barrier to displacement (Fig. S12), and the presence of two mismatches further increases this barrier. However, with a high pulling force (5 pN), the barrier decreases, resulting in a more favorable free-energy landscape at higher forces. Additionally, the presence of a distal mismatch (analogues to center mismatch) alters the energy landscape, as shown in Fig. 4 b. Interestingly, the free-energy barrier due to a distant mismatch decreases faster than that for a proximal mismatch under the same high force compared to other force conditions.

It's worth noting that a small local minimum in the free-energy landscape (Fig. 4 b) at around 8 base pairs formed with the trigger strand. This corresponds to longer waiting times observed in the MD simulations and is likely a result of sequence-dependent effects in poly-A stacked regions of the trigger strand.

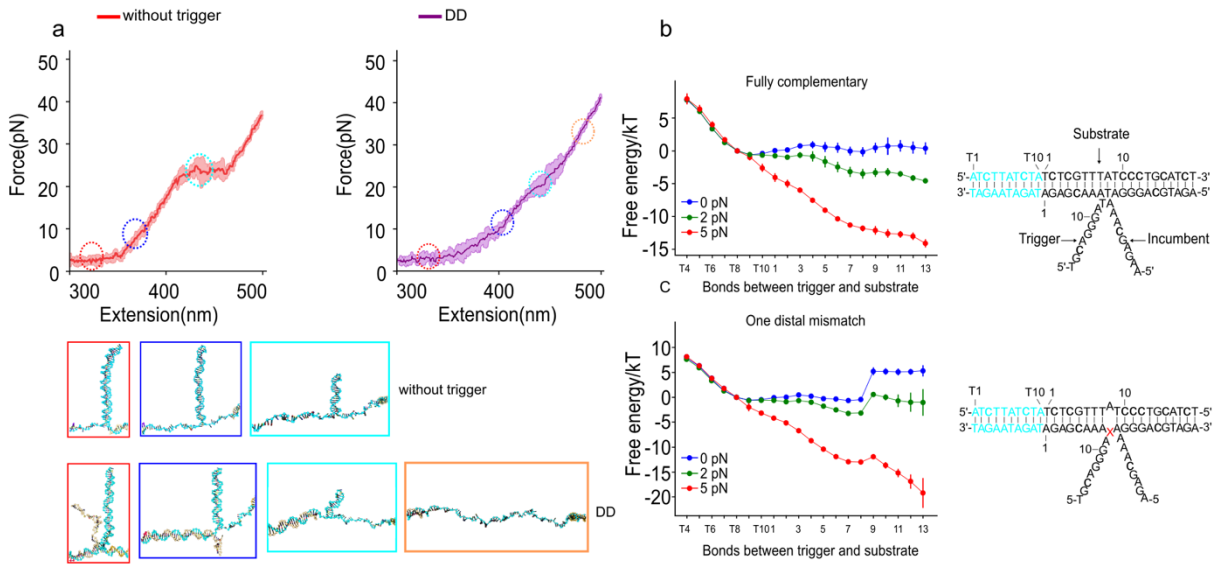


Figure 4. Simulation of force-extension of toehold hairpin and free energy landscape of TMSD with mismatched sequences using oxDNA. a), averaged force-extension curve obtained from pulling the DNA construct at a constant speed of 0.14 mm/s, comparing the toehold hairpin (left, red line) and with a fully complementary trigger strand (DD, right, purple line). The simulation has 7 replicates to ensure reliability. Snapshots of the DNA constructs are shown to depict the intermediate states observed during the unfolding process. b), Free-energy landscapes for different trigger strands: fully complementary, one proximal mismatch, and one distal mismatch. The DNA constructs' sequences are depicted on the right side, with the mismatches (A → X) and (T → X) highlighted in red. The coordinates represent the specific pairing interactions between the target sequence and the trigger sequences in each state of the system. The free-energy landscapes are shown for various force conditions, distinguished by different colors. Error bars represent standard deviation from 3 independent simulations.

## Force induced TMSD in a DNA/RNA hybrid

Compared to B-form DNA duplexes, double-stranded RNA and RNA-DNA hybrids typically adopt an A-form conformation[50], which remain a lower energy state. The double-stranded RNA is notably more stable than RNA-DNA hybrids, as shown in Fig. 5 a (middle vs. right), making invasion into an RNA stem by a DNA invader energetically unfavorable[51-53]. Without the application of force (as depicted by the black energy landscape in Fig. 5 b), this process would rarely occur spontaneously because diffusion would need to occur against an uphill energy barrier.

In SMFS experiments, tilting the energy landscape by applying an external pulling force can equilibrate invasion and reinvasion processes, enabling the observation of strand displacement by a DNA trigger (as illustrated by the grey energy landscape in Fig. 5 a, b). Fig. 5 displays a passive mode experiment in which a fully complementary DNA trigger invades the RNA hairpin (RD) at forces around 11 pN. In contrast to the DD and RR constructs measured in Fig. 2, repeated invasion/reinvasion transitions occur close to equilibrium in this experiment. These transitions involve pronounced intermediates TB, RD1, RD2, RD3, FI. Pulling/relaxation cycles also show these intermediates populated near equilibrium (Fig. S7).

It's crucial to note that all these intermediates occur at significantly lower forces compared to the intermediates populated during the unzipping of the RNA hairpin in the absence of a trigger. Furthermore, these intermediates also occur at different positions, except for RD1 (8 vs. 9 base pairs invaded). The presence of several intermediates for the RD construct strongly suggests that sequence variation plays a pivotal role in invasion, and a base-pair-by-base-pair view does not adequately describe the system. Relative free energy differences between each pair of intermediate states were computed by Boltzmann inversion of the population probabilities, confirming that, at an applied pulling force of 10.6 pN, all states were roughly at the same free energy (Fig. 5 c, grey). Incorporating the measured transition rates between the various states allowed for the extraction of transition state energies and the construction of a schematic energy landscape, assuming an Arrhenius pre-factor[54] of 3 x $10^6$ /s and transition state positions in the middle between the states (energy landscape in Fig. 5). Force-dependent rates and extrapolations to zero load are shown in Fig. S7. Transformation to zero load yields the black energy landscape in Fig. 5 c. Under a load of $F_{avg\ 1/2}$, all lifetimes between transitions occur on the order of milliseconds. However, when the load is reduced, reinvasion rates will quickly become dominant over invasion rates, causing the equilibrium to strongly shift towards the TB state. This shift indicates that the presence of force strongly influences the kinetics of the system, favoring either invasion or reinvasion depending on the applied load.

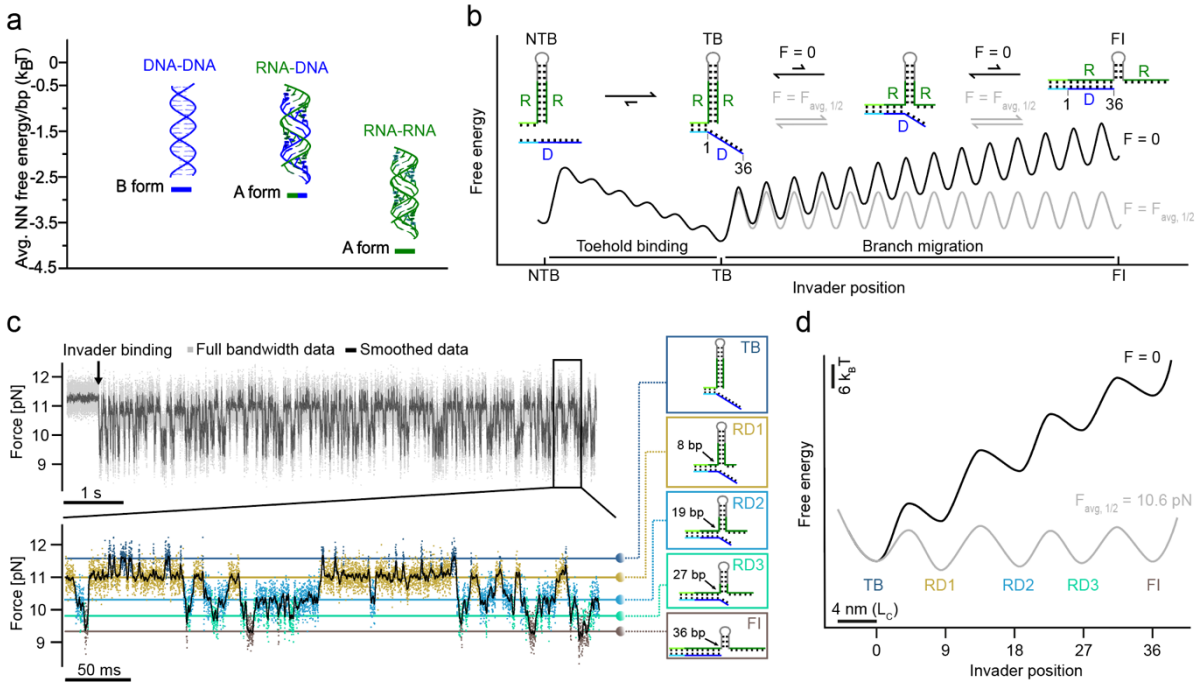


Figure 5. Kinetic profiles of force induced DNA-RNA hybrid TMSD. a), Average free-energy per base-pair among double strand DNA, RNA and DNA/RNA hybrids. DNA trigger has extremely low invasion to RNA hairpin due to the energy barrier between RNA-RNA stem and DNA-RNA strand. b), schematic of force-biased strand displacement process and corresponding free energy landscape. Branch migration cannot occur spontaneously due the energy barrier while remains equilibrium under $F_{avg\ 1/2}$. d), force versus time trace of DNA-RNA hybrid complex. Each intermediate state during branch migration process is distinguished based on calculation of contour length. FI: fully invaded, DR1, 2, 3: intermediate invasion/reinvasion 1, 2, 3, TB: toehold bound. d), Free energy profile shows different free energies between each two states. Transition rates between each two different

intermediate states under different force conditions are shown. Forward and backward rates were calculated using a model described in Data analysis.

# Discussion

The investigation of strand displacement processes has predominantly relied on bulk kinetic measurements, as evidenced by previous studies[18,20,46,55-57]. However, only a handful of single-molecule investigations have been conducted, mainly employing techniques like single-molecule FRET[58] or magnetic tweezers[59]. Our study, utilizing optical tweezers based SMFS, as presented here, offers a unique advantage by providing an exceptionally detailed view of the process, with a resolution approaching that of individual base pairs. While our single-molecule TMSD experiments consistent with the findings of previous bulk studies, they also introduce unprecedented insights into sequence-specific characteristics that exert a significant influence on the TMSD process. In the following discussion, we spotlight several of these remarkable discoveries, which have been gleaned from our SFMS measurements.

*Time-scales of TMSD*:

In terms of time scales, our experimental setup offers a notable advantage over other recent SMFS studies[32,60], with a response time nearly an order of magnitude faster. This enables us to capture branch migration rates that closely approximate the true kinetics of the process. In passive mode experiments conducted at a force bias of 10 pN, we directly observed strand invasion, resulting in strand displacement times for a 36-base-pair domain of approximately 42 µs for fully complementary DNA triggers. It's worth noting that this value serves as an upper bound, as our instrumental limitation is around 30 microseconds.

This observation of rapid SD is surprisingly fast when compared to estimates derived from step times obtained in bulk FRET experiments (approximately 28 microseconds per step, as reported in Seidel et al.). According to these step times, the full invasion process covering 36 base pairs would take approximately 19 milliseconds, almost three orders of magnitude slower than our experimental findings. Several factors contribute to this apparent discrepancy: First, at zero force, SD follows a diffusive process, scaling with the square of the number of steps involved. However, at around 10 pN of force, strand displacement becomes strongly biased and scales linearly with the number of steps. Additionally, force can significantly affect the rate of branch migration by reducing the barrier for each individual step, thereby accelerating the displacement process. Another factor contributing to faster displacement under force could involve a change in the displacement mechanism. In the presence of force, it's possible that more than one base pair is broken at a time, leading to larger effective step sizes. Our simulations support the idea that faithful, step-by-step invasion remains the primary process at 10 pN, although parallel pathways with larger step sizes begin to compete. These findings on significantly faster branch migration under load may have important implications in contexts where DNA machinery applies forces to

DNA while another strand invades, as seen in processes such as CRISPR, T7 RNA Polymerase, and the Ribosome.

The measured times for TMSD in RNA under loads of 10 pN are significantly longer than those for DNA, with times of approximately 1.28 milliseconds for RNA compared to 42 microseconds for DNA (as shown in Fig. 2). However, these longer times are primarily due to the presence of two intermediate states that are populated during the branch migration process (refer to Fig. S5 for sample traces of RNA and Fig. S4). As a result, these longer times cannot be used to calculate average step times for RNA branch migration. The nature of these intermediate states warrants further investigation. Analysis of the distribution of invaded nucleotides during individual strand invasion events (as shown in Fig. S4 D) suggests two potential intermediate candidates, one at approximately 8 base pairs and another at 15 base pairs. Several possibilities could explain the presence of these intermediates. One possibility is that there may be sequence-related energy minima at the positions where these intermediates are found. Another possibility is the formation of secondary structures in the invading strand, particularly in regions not subjected to mechanical load in our experiments. Several lines of evidence support the idea that these intermediates are associated with structure formation in the trigger strand. Firstly, the duration of the intermediates exhibits weak dependence on force (as seen in Fig. 2 d, Fig. S3 E), suggesting that the structure formation occurs in the part of the strand that is not under mechanical load, i.e., the trigger strand. Furthermore, analysis of traces from the RRc2 construct (as shown in Fig. S3 A & B) reveals that reinvasion can extend beyond the expected intermediate IM at the position of the two mismatches, reaching a new intermediate labeled as BI. This behavior aligns with the notion that the formation of secondary structures in the invader strand can block base-pairing with the hairpin, allowing re-invasion to proceed beyond what would typically be expected based solely on the presence of a mismatch. It's worth noting that the observation of transient secondary structure formation in the trigger strand is unexpected, given that the trigger strand was designed specifically to minimize such secondary structure formation. Even though these secondary structures are only populated on a millisecond timescale, they may still be important to consider in the design of future RNA riboregulators. Additionally, the intermediates were observed even in the absence of magnesium, although with shorter lifetimes (approximately 600 microseconds compared to 2 milliseconds). Experiments conducted at lower salt concentrations and in the absence of magnesium also resulted in faster kinetics (as shown in Fig. S5). These findings indicate that salt concentration and the presence of magnesium can modulate the kinetics of RNA invasion processes.

*Invasion of DNA into an RNA duplex:*

The application of force to bias the invasion process also enabled the induction of strand displacement between a DNA trigger and an RNA hairpin. In bulk experiments[42], such a displacement process between DNA and RNA would typically exhibit extremely slow kinetics due to the substantial free-energy difference between DNA/RNA and RNA/RNA duplexes. However,

in our experiments conducted under force, we observed three distinct intermediate states between the toehold-bound and the fully invaded strand, even in the absence of any mismatches on the trigger strand. This observation suggests that the application of force can significantly alter the kinetics and pathways of strand displacement reactions, allowing for the observation of intermediate states that might not be readily apparent in bulk experiments.

The kinetics of strand invasion into RNA double strands differ significantly from those of DNA invasions due to the unfavorable free energy balance in favor of RNA double strands. Invasion into a long (36 bp) fully complementary RNA sequence occurs rapidly (average times of 69 ms) only under substantial mechanical forces (10.6 pN in Fig. 5). In the absence of force, such an invasion would not occur spontaneously, taking an extremely long time ($5.08 \times 10^{13}$ s or 1.61 million years, calculated using the mean first passage time analysis for a 1D random walk in the Theory and Data analysis). However, by using shorter sequences, the necessary time for invasion can be significantly reduced. For instance, invasion into the first 8 bp from the toehold-bound (TB) state to the RD1 intermediate state takes only 0.741 seconds even at zero force (see Fig. S7). This adaptability in sequence length offers a way to finely tune the desired kinetics of the invasion process.

Using DNA as the invader strand presents advantages, such as a reduced likelihood of secondary structure formation compared to RNA invader strands. Additionally, introducing interior bulges into the stem of an RNA toehold hairpin provides another means of adjusting invasion kinetics by altering the energy balance. This can result in a shift that allows a fully complementary DNA invader to quickly invade an RNA toehold hairpin, even in the absence of mechanical load. To exemplify this concept, we show that DNA invasion into RNA can indeed shift the equilibrium towards the open state of the toehold hairpin using the original toehold switch structure (Fig. S9).

A significant difference between RNA/DNA hybrids and non-hybrids is the pronounced sequence effect, leading to the observation of three clear intermediates during hybrid strand displacement. This is distinct from non-hybrid systems where intermediate states may not be as apparent. Our analysis rules out secondary structure formation as the cause of these intermediates. A simple calculation of the energy landscape helps to elucidate the intricacies of this sequence-specific behavior (Fig. S7 F & G). Individual step times in this process can be complex, as upward steps may differ from downward ones, and the kinetics may vary for each base pair, depending on the neighboring sequence. While our simulations show a sequence effect for DNA/DNA branch migration, this effect is much more pronounced for RNA/DNA hybrids. A coarse estimate of an upper bound for downward steps is 10 microseconds, which, although faster than expected, is similar to the upper bound for DNA/DNA branch migration due to the higher free energy gain during the invasion process.

*Effect of mismatches on branch migration kinetics*:

The introduction of sequence mismatches is a well-established technique in DNA nanotechnology to modulate branch migration kinetics. In our study, we strategically placed sequence

mismatches on the trigger strands, allowing us to repeatedly observe invasion and reinvasion events under mechanical load. By analyzing the kinetics under load (Fig. 3 d), we could extrapolate these measurements to estimate the kinetics at zero load. In the case of fully complementary constructs, invasion occurred almost instantaneously after the toehold had bound, but at zero load, the invasion rates were notably slower for DDc1, DDc2, and RRc1 (1/s at zero load for DDc1 and 0.01/s for DDc2, see Table S5 and Fig. 3 d).

The rate plots in Fig. 3 d directly illustrate a result of the change in free energy induced by the mismatches compared to the fully complementary sequences. The calculated free energy differences are summarized in Table S5. Notably, the values we obtained are within 18% of the predictions made by the nearest neighbor model NUPACK for the DNA mismatches and within 27% for the RNA mismatch. It's worth mentioning that the systematic deviation of our experimental results towards lower values as compared to the predictions could be attributed to potential systematic errors in our force calibration. For a more detailed exploration of this aspect, please refer to the Appendix.

Previous research has highlighted that the position of a mismatch (whether it's proximal, central, or distal) can have a significant impact on displacement kinetics[61], and this phenomenon has been observed in bulk measurements[62,63]. There are two key reasons why this position-dependent effect is evident in bulk experiments: Firstly, when a mismatch is positioned distally (further away from the toehold), it tends to have a less drastic impact on the displacement rate. This is because the last few steps of branch migration can be influenced or accelerated by the spontaneous dissociation of the incumbent strand[46]. In other words, when the mismatch is far from the toehold, the kinetics of strand displacement are less affected. Secondly, when a mismatch is positioned proximally (close to the toehold), its effect can be weaker, especially if the toehold is short and not saturated with binding. In this case, a waiting time at the mismatch position could increase the probability of the invader strand dissociating, affecting the second-order rate of binding to the toehold. When the mismatch is positioned further downstream, the complex formed between the invader and target is more stable because it contains more base pairs. As a result, the dissociation of the invader becomes less likely.

Our experimental design stands out for its independence from certain effects and, fundamentally, enables the determination of displacement kinetics that are not influenced by specific positional factors. This robustness can be attributed to two key factors: Firstly, the clever use of our toehold hairpin design, along with the inclusion of an additional remaining stem even after full invasion, effectively precludes the possibility of spontaneous dissociation of the incumbent strand. This intrinsic stability ensures that once the invader binds, it remains firmly attached throughout the entire experiment. This stability is pivotal for the precise examination of strand displacement kinetics, free from the complexities associated with spontaneous dissociation. Another crucial aspect is our choice of a 14-nucleotide toehold length, mirroring the length used in a toehold switch design in Ref. 11. This specific toehold length facilitates prolonged and stable binding of the invader to the toehold, ensuring a reliable interaction over the entire duration of our

experiment. Even in scenarios where the toehold might be shorter and detach during the measurement, our setup remains capable of detecting these events. Toehold binding and unbinding can be observed separately, particularly in systems like RRp2 (Fig. S8). In this case, toehold binding is characterized by an increase in force, as the stretched double-stranded toehold-invader complex has a shorter extension compared to the stretched single-stranded toehold. The invasion event is momentarily delayed due to the presence of proximal mismatches directly following the toehold binding domain (represented as $\tau_{TB}$ Fig. S8). After approximately 100 milliseconds, the invasion proceeds, and this transition is marked by a drop in force as the invader opens the hairpin.

*OxDNA simulations*: To verify our experimental findings and explore our hypotheses from a different perspective, we turned to oxDNA simulations to investigate how mismatches impact the free energy landscape and strand displacement process. Our simulations yielded several key insights: Initially, we observed similar behavior in the force-extension curves when using a constant pulling velocity, mirroring our experimental results. However, it's important to note that the pulling speeds in our simulations were orders of magnitude faster than those used in the experiments. As a result, the simulations were strongly out of equilibrium, making it challenging to identify well-defined intermediates. The free energy landscape simulations indicated that the energy barrier created by the presence of mismatches could be alleviated by the energy activation induced by an applied force. In other words, when force was applied, it helped overcome the energy barrier posed by the mismatches, facilitating the strand displacement process. Our simulations with double proximal mismatches revealed a significant reduction in the branch migration process at low forces (1, 2, 5 pN). This aligns with the observed slow effect of proximal mismatches in bulk experiments. We hypothesized that this slow effect primarily arises from the spontaneous detachment of the trigger strand from the trigger-bound state. This detachment event reduces the overall reaction rate, resulting in slower kinetics of the branch migration process. Interestingly, in the case of a single mismatch, we didn't observe significant differences in the energy barrier between different mismatch positions under the same force conditions. This suggests that the variation in branch migration kinetics is not primarily governed by thermodynamic properties but rather by kinetic factors. In summary, our oxDNA simulations support and complement our experimental observations and provided valuable insights into how mismatches influence the strand displacement process and the role of applied force in overcoming energy barriers created by these mismatches.

*Significance for nucleic acid nanotechnology and synthetic biology:*

Our findings hold strong implications for advancing dynamic nucleic acid nanotechnology and its applications in synthetic biology across multiple dimensions. Prior to our study, researchers in DNA nanotechnology primarily delved into the kinetics of TMSD processes through bulk fluorescence investigations. These studies meticulously dissected the components of bulk kinetic

data to extract effective branch migration rates, distinguishing between the initial toehold binding and the subsequent strand displacement process. Consistently, bulk studies underscored the substantial influence of mismatches within the branch migration domain, impacting both the incumbent and the invader strands. This knowledge has already been harnessed for diverse applications, such as the detection of single nucleotide polymorphisms in an array of DNA sensors[64] and even *in vivo* applications involving RNA sensors (SNIPRs). Our single-molecule studies not only validate the pivotal role of mismatches but also offer a unique vantage point: direct observation of strand displacement occurring within individual molecules under an applied load. These experiments further unveil an intriguing aspect—the process displays remarkable variability along the branch migration domain. This observation strongly implies a profound sequence dependence that has yet to be comprehensively understood or characterized. In essence, our research bridges a critical gap in our understanding of strand displacement kinetics and opens up exciting possibilities for harnessing these insights in the advancement of nucleic acid nanotechnology and synthetic biology applications.

To realize the molecular circuitry within living cells, the utilization of RNA molecules as inputs or substrates for strand displacement processes is paramount. These scenarios often involve RNA invaders intruding into RNA or DNA duplexes, or DNA itself infiltrating RNA duplexes or RNA-DNA hybrids. In this context, our study offers a critical insight: RNA's invasion into RNA occurs significantly more slowly than DNA's invasion into DNA duplexes. Furthermore, the emergence of previously undocumented intermediate states, attributed to secondary structure formation, may further decelerate branch migration when RNA is invading RNA. Intriguingly, our research unveils that DNA can expedite its invasion into RNA duplexes under the influence of mechanical force. Particularly, SD processes holds pivotal significance in both biological and synthetic biological contexts. Instances include the invasion of CRISPR-guide RNA complexes into DNA duplexes, the dynamic switching of riboregulators, or the conditional guide RNAs employed in various applications. While we possess a general comprehension of the underlying mechanisms, it has often been perplexing to interpret the highly variable efficiencies of sequences that, superficially appear quite similar. This complexity has led to the widespread application of machine learning approaches to assist in the design of functional RNA molecules[65]. Our biophysical insights have the potential to provide valuable context for interpreting the predictions generated by such computational networks. This, in turn, could lead to the formulation of refined yet understandable design principles. Ultimately, the deliberate selection of specific branch migration sequences, including the strategic introduction of mismatches and bulges, can serve as a means to finely tune the performance of gene regulatory switches tailored for specific applications.

Lastly, it's worth noting that the forces in the range of pN are prevalent within the biological milieu, where strand displacement processes come into play. Remarkably, molecular motors like RNA and DNA polymerases have exhibited the capability to generate forces as high as 35 pN[66-68]. Our discovery that lowers forces lead to a significant acceleration of strand displacement processes implies that TMSD within cells might operate with kinetics considerably distinct from those

observed in controlled laboratory settings. For example, the forces exerted by RNA polymerases during the transcription of a riboregulator or a ribosome navigating through the secondary structure of an RNA molecule could either hinder or facilitate strand invasion by a trans-acting RNA effector. This dynamic interplay underscores the relevance of our findings in real biological scenarios.

# Conclusion

In this chapter, our primary objective was to meticulously explore the kinetics of hairpin TMSD through the utilization of a microfluidic-based single-molecule assay. Our approach was distinct in that we deliberately introduced base-pair mismatches to modulate the branch migration process, allowing us to obtain a precise estimation of the force-biased kinetic rate. Our inspiration for this novel approach stemmed from the realm of riboregulators, particularly 'toehold switches'[11]. In our experimental design, we engineered a toehold hairpin structure and applied force simultaneously to both ends of the molecule, ensuring the incumbent strand's continued association throughout the process. Unlike traditional 3-way branch migration, our setup enabled systematic measurement of both invasion and reinvasion kinetics of branch migration in numerous repetitions in a passive mode. As a result, our methodology enabled a more comprehensive statistical analysis, providing a clearer and more detailed characterization of the kinetic events at the single-molecule level. To maintain alignment with the original toehold switch design, we adopted a 14-nucleotide toehold length in both our DNA and RNA hairpin structures. Importantly, this choice, corroborated by previous studies, optimizes the strand displacement rate and effectively prevents the dissociation of the trigger strand from the toehold, ensuring the stability of our experimental setup.

Our experimental investigations, coupled with the insights gained from molecular dynamic simulations, provide a comprehensive understanding of the intricate mechanisms underlying strand displacement during TMSD. These observations delve into the dynamic intricacies of TMSD at the level of individual base pairs, shedding light on a process reminiscent of RNA-protein interactions within cellular contexts. These findings carry significant implications, particularly in the realm of designing and fine-tuning dynamically responsive RNA riboregulators. Moreover, they hold promise for advancing fields such as CRISPR-mediated genome editing and gene therapy. We've discovered that the kinetics of the branch migration process are distinctly influenced by the mechanical forces applied to DNA or RNA strands, mirroring conditions akin to translation or RNA chaperone binding processes. Our discoveries are poised to deepen our understanding of the dynamic interplay involved in enzyme-mediated DNA/RNA interactions, encompassing facets such as RNA folding and strand displacement, all at the single-molecule level. Consequently, this study yields valuable insights into the kinetics and thermodynamics governing strand displacement, offering a clearer picture of the factors influencing branch

migration rates. The knowledge gleaned from this investigation holds significant potential for practical applications in the design and optimization of DNA/RNA-based molecular devices. These applications span a wide spectrum, encompassing areas such as biosensing, diagnostics, and therapeutics, ultimately contributing to advancements in these fields.

# Materials and Methods

Culture media

We employed LB medium (Carl Roth) and Turbo® (NEB) cloning strain to culture our cells. The medium was supplemented with 100 µg/ml kanamycin to maintain selective pressure.

Plasmid construction and cloning process

All DNA oligonucleotides were obtained from Eurofins Genomics, Ebersberg, Germany and biomers.net GmbH. Toehold hairpin and trigger sequences were constructed using a combination of overlap extension PCR, restriction ligation, and blunt-end ligation[69]. Toehold hairpin and trigger sequences were constructed using a combination of overlap extension PCR, restriction ligation, and blunt-end ligation. First, we amplified each part of the sequence using two steps PCR and annealed the reverse primer (*SpeI*, *PstI*) in Q5® Master Mix (NEB) at a calculated annealing temperature (NEB Tm calculator (https://tmcalculator.neb.com/#!/main). DNA polymerase was used to extend the 3' ends and fill up the gaps. We then added forward primer, including restriction site sequences (*EcoRI*, *XbaI*), to the PCR mix and amplified the target strand. The PCR products were purified using the Monarch® PCR & DNA Cleanup Kit (NEB) and their concentration and quality were assessed using a Nanodrop 8000 spectrophotometer (Thermo Fisher).

Next, toehold hairpin and trigger sequences were ligated with vector using restriction ligation. Cloning vector plasmids (modified pet28b) were also digested with *EcoRI* and PstI and gel-purified to remove the digested strands. Finally, all three parts (with a ratio of inserts to vector of 1:3) were ligated using T4 ligase (NEB) with the standard protocol. The ligation products were then transformed into chemically competent cells (Turbo®, NEB) or electroporation-competent cells (Turbo®, NEB) using a standard protocol. The cells were plated on LB agar plates containing 100 µg/ml kanamycin and incubated overnight at 37°C. A single colony was picked and checked using colony PCR. The selected colony was inoculated in 5 mL LB medium and incubated overnight at 37 °C. After overnight culture, cells were collected, and plasmids were purified using miniprep kits (QIAprep Spin Miniprep Kit).

Blunt-end ligation was used to optimize the toehold hairpin and trigger sequence. We amplified toehold hairpin and trigger constructs and the vector using primers that included a portion of the optimized sequence. Next, we *in vitro* phosphorylated the PCR products using T4 Polynucleotide

Kinase (NEB) following the standard protocol. The phosphorylated PCR products were then ligated using T4 ligase at room temperature for 2 hours and digested with DpnI (NEB) using the standard protocol to remove any remaining original plasmid DNA. The final products were transformed into chemically competent cells (Turbo®, NEB). A list of all plasmids used in this study can be found in the DNA construct sequence list.

Cell culture

Bacterial strains were grown in LB media using 5 mL culture each in 50 mL centrifuge tubes at 37°C while shaking at 250 rpm (Innova$^{TM}$ 44 incubator).

DNA handle preparation and *In vitro* transcription

We PCR-amplified the DNA handle strands (511 bp) from Lambda phage DNA (NEB) using modified primers (see Sequence and Primers). Specifically, the forward primer was labeled with two dT-Biotin or dT-Digoxigenin molecules at the 5' end, while the reverse primer included a stable abasic-site to preserve a single-stranded overhang for binding to the target molecule.

All *in vitro* gene transcription experiments with toehold hairpin and trigger RNAs were performed using a homemade *in vitro* transcription mix including a homemade T7 RNA polymerase. The T7 RNA polymerase with a 6xHis tag was expressed in *E. coli* BL21 DE3, followed by cell lysis using lysis buffer (1 mM Benzamidine, 1 mM PMSF, 1:2000 (1mU) dilution Turbo Dnase from Ambion, 1 mg/ml Lysozyme of chicken egg white) and sonication, followed by purification using an ÄKTA pure Chromatography System. Next to the T7RNAP, the TX mix contained transcription buffer (50mM HEPES, 22mM $MgCl_2$, 100mM KCl, pH 7.8) and Murine Rnase inhibitor (NEB) in a 20 µl or 100 µl reaction. Linear transcription templates for toehold hairpins and trigger RNAs were first amplified using PCR and purified using a Monarch® PCR Cleanup Kit (NEB). The concentration and quality of purified DNA templates were quantified via their 260/280 and 260/230 ratios using a Nanodrop 8000 spectrophotometer (Thermo Fisher). The molar concentration of each DNA template was calculated via (1):

$$\frac{Concentration\ (ng/\mu l) \times 10^6}{Molecular\ weight(g/mol)} = Concentration\ (nM) \quad (1)$$

Agarose gel electrophoresis and Urea-PAGE purification

The DNA handle used for attaching the target molecule and silica beads (0.5 kb) was initially PCR amplified (primer sequence see Sequence and primers) and purified through agarose gel (2% wt, from CARL ROTH). However, we encountered a false priming issue that resulted in an additional 200 bp junk strand on the PCR product, which could significantly impact the subsequent folding process with the target molecule. The target bands were cut and purified using Gel Purification Kit from QIAGEN.

After *in vitro* transcription, the toehold hairpin and trigger RNAs were initially digested for 30 minutes using Dnase I (NEB) to remove the original DNA template. Next, we added 0.5 M EDTA to chelate the remaining $Mg^{2+}$ from the samples and denatured them at 65 °C. After denaturation, the RNA samples were purified using a 10% Urea-PAGE gel (consisting of Urea 4.8g, 40% Acryl (29:1) 2.5 ml, 30% APS 50 µL, TEMED 10 µL, and 10xTBE 1mL, all from Carl Roth). The gel electrophoresis was performed using the Owl™ gel system. Following gel electrophoresis, the target bands were cut from the gel and the RNA was extracted using the ZR small-RNA™ PAGE Recovery Kit (from Zymo Research). The RNA concentrations were also measured using a Nanodrop 8000 spectrophotometer (from Thermo Fisher).

Toehold hairpin constructs preparation

After agarose and PAGE gel purification, we measured and calculated the molarity of each component, including DNA handles, toehold hairpins (DNA or RNA) and adapter strand. Toehold hairpin strands and adapter strands are firstly mixed with 1:1 molarity ratio of dT-Biotin-DNA handles (40 mM) and dT- Digoxigenin-DNA handles (40 mM) respectively. Next, the samples were dried using Concentrator 5301 Eppendorf and resuspended in folding buffers (1M NaCl, 50 mM HEPES, pH, 7.8 or 20 mM $MgCl_2$, 50 mM HEPES, pH 7.8). We incubated samples under different annealing temperature cycles (Table.1). We finally checked the folding constructs on agarose gel.

Table 1 Folding of toehold hairpin and adapter strand with DNA handles

| STEP | TEMP | TIME |
|---|---|---|
| Initial Denaturation | 70°C | 1 min |
| Folding | 68°C 30s<br>65 °C 20 min | 30 cycles |
| Final | 25°C | 5 minutes |
| Hold | 4-10°C | ∞ |

Folding of toehold hairpin-DNA handles with adapter-DNA handles

| STEP | TEMP | TIME |
|---|---|---|
| Initial Denaturation | 68°C | 1 min |
| Folding | 68°C 30s<br>63°C 20 min | 30 cycles |
| Final | 25°C | 5 minutes |
| Hold | 4-10°C | ∞ |

Preparation of Dumbbell Assay and data analysis

To minimize multi-binding of the DNA handle, we diluted the final constructs to a concentration of 0.4 nM and incubated them for 10 minutes with 1 µm-sized streptavidin-coated beads (Bangs Laboratories, Inc.) in 14 µL of running buffer (20 mM $MgCl_2$, 300 mM KCl, 50 mM HEPES, pH 7.2) at room temperature. In the meantime, we prepared the mobile phases in 500 µL of running buffer by adding

an additional oxygen scavenger system (final concentrations: 26 U/ml glucose oxidase (SIGMA-ALDRICH), 17 000 U/ml catalase (SERVA), and 0.65 % glucose (SIGMA-ALDRICH)). The trigger phase included 0.1 μM of purified trigger strand, while the bead phase consisted of a mixture of streptavidin-coated beads and anti-digoxigenin-coated beads incubated together in 300 μL of the same scavenger system. All the phases were added to the syringe pump of the C-Trap® Optical Tweezers – Fluorescence & Label-free Microscopy. We used a commercial microfluidic chip with multiple inlets that allowed us to generate a laminar flow to separate different phases (see Fig. 1B). In the Bead channel, we trapped the two different kinds of beads, one in the fixed beam and the other one in the mobile beam. In the Buffer channel, the construct was tethered between the two beads by moving the mobile trap towards the fixed trap, therefore bringing the bead surfaces in proximity. We maintained the laminar flow of buffer and trigger phase by applying ~ 0.35 bar to the syringes, leading to a flow velocity of ~ 20 μm/s, during the experiment to inhibit trigger diffusion into the buffer channel even for measurement times of up to an hour. The trap stiffness used in all measurements was between 0.25 pN/nm and 0.40 pN/nm. The sampling rate was 78.125 kHz and was down sampled by a factor of 3 for data analysis except for the transition path time analysis and autocorrelation analysis used for data shown in Fig. 2 c, E. Measurement temperatures were ~ 25°C.

Stretch and relax cycles (Active mode)

Once the tether was formed, we performed several stretch & relax cycles (forward and backward) at a constant velocity of 0.2 μm/s by moving the traps apart to a distance where the toehold hairpin was fully unfolded, and then reducing the trap distance to allow the hairpin to refold. To increase resolution, we also used slow constant pulling velocities of 50 nm/s. From these experiments, we generated force-extension traces as shown in Fig. 2 b and S3. Force-extension traces were used to determine the unfolding intermediates of the toehold hairpin including the associated contour length gains for calculating the opened base pairs. The data was recorded and followed by the analyzing process in next section.

Passive-mode

We further carried out constant distance measurements by applying a constant force bias to the toehold hairpin while observing the fluctuations among intermediate folding states. During this process, the force varied as the hairpin transitioned between completely opened or folded states while the trap remained at a fixed distance, causing changes in the contour length of the tether. This approach enabled us to obtain free energies, transition rates, and contour lengths of these states, as demonstrated in Fig. 1B, c, Fig. 2B, and Fig. 4c. The raw data was smoothed using a running average of 20 points and is shown as black lines. To integrate the raw data points into the different intermediate states, we used a Hidden-Markov-Model (HMM) analysis.

OxDNA simulations

To create the DNA handle and toehold hairpin structure, we utilized Oxview online tools in the OxDNA2 version of the model for DNA simulations[42-44], employing sequence-dependent parametrization. For relaxation of DNA constructs and pulling simulations, we employed molecular dynamics (MD) simulations.

To simulate the DNA system being pulled, we conducted molecular dynamics simulations with a time step of 15.15 fs and utilized an Andersen-like thermostat at 20°C. The simulations involved applying constant force on the last two nucleotides of both the left and right DNA handles, and we also used a harmonic trap with stiffness k = 1 acting on the center of mass of the nucleotides in the handles. During the pulling simulation, we moved the harmonic traps in opposite directions at a constant speed of 0.14 mm/s. The force was measured as the stiffness constant k multiplied by the distance between the position of the trap and the center of mass of the nucleotide it acts on. It's important to note that due to the highly coarse-grained nature of the model, establishing a direct correspondence between the simulated time and experimental time is not straightforward. For example, in the model, we artificially increased the diffusion coefficient, leading to approximately 100x faster diffusion than what is observed experimentally[70].

To obtain the free-energy profiles of DNA systems under different tensions, we employed a simplified three-strand system based on the DNA constructs sequence. To perform this simulation, we utilized the Virtual Move Monte Carlo (VMMC) algorithm with umbrella sampling. Due to the large size of the full system with DNA handles used in experiments, it is not feasible to sample using the VMMC algorithm. Hence, we focused on a truncated system consisting of three strands: invader, substrate, and incumbent (shown in Fig. S10). Following the approach described in Reference 41, we sampled the free-energy landscape in terms of the number of bonds between the invader strand and substrate, as well as between the incumbent strand and the invader strand. To achieve this, we applied an umbrella sampling potential with iteratively adjusted weights, allowing us to sample different states based on the bonds between the invader/incumbent strands and the substrate. We employed the Weighted Histogram Analysis Method (WHAM) to combine simulations from various windows, each exploring different force conditions applied to the substrate strand. Specifically, opposing forces of magnitude 1, 2, 5, and 10 pN were applied to a pair of nucleotides in the system (as schematically shown in Fig. S12).

# Theory and Data analysis

The fundamental dumbbell assay involves the application of cyclic loads at a consistent speed (Fig. 6 a). This is achieved by moving one laser beam in a triangular wave pattern while keeping the other trap stationary. This controlled manipulation stretches the molecule, gradually increasing the forces applied, allowing the experimenter to identify the force range associated with unfolding events. Typically, the data is presented as the force on the beads plotted against the tether extension. This controlled manipulation involves gradually stretching the molecule while

incrementally increasing the applied forces. This process enables the experimenter to pinpoint the force range associated with unfolding events by precisely tracking the position of the beads in relation to the laser, a parameter termed deflection x (Fig. 6 a). The force-extension trace can then be utilized to calculate the contour length using the WLC model (Fig. 6 b). For a more comprehensive exploration of the energy landscape at equilibrium, a passive mode measurement is employed. In this mode, two traps maintain a fixed distance to apply tension to the molecule, enabling the examination of different conformations, including folding and unfolding, under equilibrium conditions. The force is recorded at a consistent sample rate over a specified duration, as illustrated in Fig. 6 c.

To understand how energy landscapes change under different pre-tensions and to distinguish unfolding or unbinding events from overlapping data in force vs. time traces, we utilized HMMs. These models work under the assumption that a molecule's transition through various states is a memoryless process, meaning each step only depends on its current state. At discrete time intervals, the system has a consistent transition probability ($T_{ij}$) of either staying in the current state or moving to another (as shown in Fig. 6 d). And the emission probabilities are defined as the probabilty of observing a data point at a specific force (F), given that the molecule is in a hidden state (i). The algorithm calculates the probabilities ($p_{forward}$ and $p_{backward}$) that the system is in a certain state (i) at time (t), considering the force values from time 0 to t, alongside emission and transition probabilities. Finally, to assign the most probable state to each data point, we multiply these probabilities (using the forward-backward algorithm).

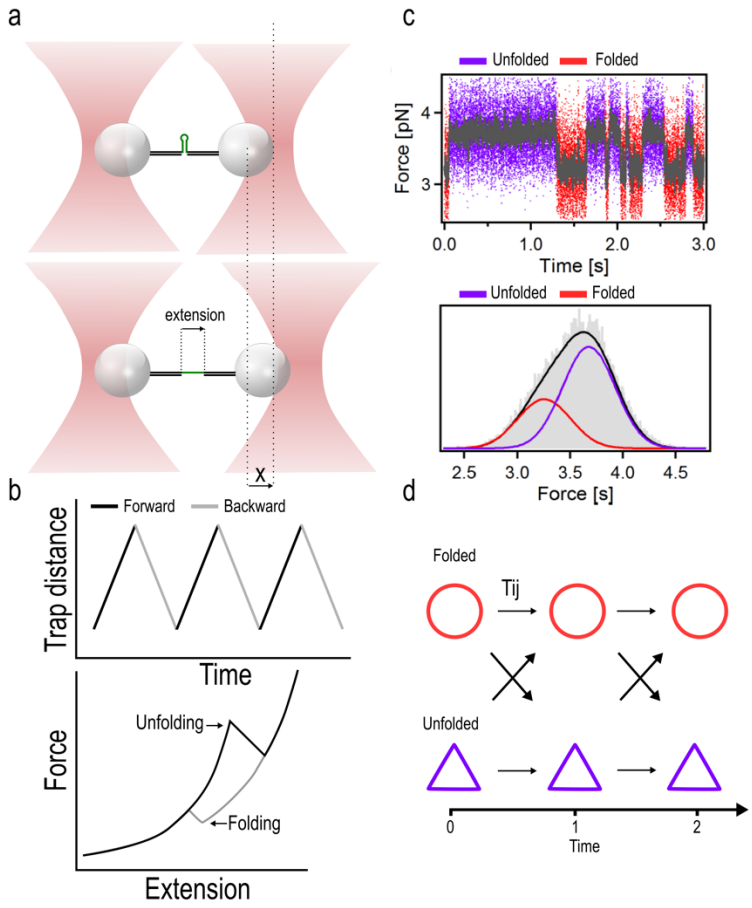


Figure 6. Dumbbell Assay Modes and Hidden Markov Models (HMMs). a), In the active mode of the dumbbell assay, a molecule (green) is positioned between two DNA handles (black) and subjected to unfolding using dual-beam optical tweezers. As the molecule unfolds, the tether's extension increases, causing the bead deflection x to decrease, resulting in a decrease in the measured force. b), the assay involves moving one trap at a constant velocity, both forward (black) and backward (grey) (above). The graph depicting force vs. extension (below) illustrates unfolding and refolding events, characterized by changes in both force and extension. c), an example of HMM analysis applied to raw data to distinguish between folded (red) and unfolded (purple) states (above). The force distribution of the trace is represented as emission probabilities (below), with double-Gaussian fits (purple and red) and their combined curve (black). d), schematic illustrating the principle of HMMs in the discrete-time case. At each time step, the system maintains constant transition probabilities ($T_{ij}$) for staying in the current state or transitioning to another state.

In the absence of a trigger strand, we initiated the tether formation process. Subsequently, we conducted multiple stretch and relax cycles at a constant velocity of 0.2 µm/s. These cycles involved moving the traps apart to a distance at which the toehold hairpin was completely unfolded, followed by a reduction in trap distance to facilitate hairpin refolding. Analysis of the force (F) versus extension (e) traces allowed us to identify unfolding intermediates within the toehold hairpin, along with the associated gains in contour length, which were instrumental in calculating the number of opened base pairs.

Stretch and relax cycles (active mode):

Modeling polymer elasticity in stretch-relax cycles. The force-extension curves for stretching of the DNA only, can be modeled by using the extensible worm-like-chain (eWLC) model[59,71] (2),

and If parts of the RNA are unfolded, the elastic behavior can be described by the eWLC in series with a standard WLC[37] (3).

$$F_{eWLC}(e) = \frac{k_B T}{p_{dsDNA}}\left[\frac{1}{4}\left(1 - \frac{e}{L_{dsDNA,linker}} + \frac{F}{K}\right)^{-2} - \frac{1}{4} + \frac{e}{L_{dsDNA,linker}} - \frac{F}{K}\right] \quad (2)$$

Where $k_B T$ is the thermal energy, $p_{dsDNA}$ the DNA persistence length, $L_{dsDNA,}$ linker the DNA contour length and $K$ the elastic stretch modulus.

$$F_{WLC}(e) = \frac{k_B T}{p_{ssDNA/ssRNA}}\left[\frac{1}{4}\left(1 - \frac{e}{L_{ssDNA/ssRNA}}\right)^{-2} - \frac{1}{4} + \frac{e}{L_{ssDNA/ssRNA}}\right] \quad (3)$$

The persistence length of the unfolded single-stranded DNA/RNA toehold hairpin, denoted as $p_{\text{ssDNA/ssRNA}}$, and the contour length of the unfolded single-stranded DNA/RNA toehold hairpin, referred to as $L_{\text{ssDNA/ssRNA}}$, play crucial roles in our analysis. It's worth noting that the standard Worm-Like Chain (WLC) model only considers entropic effects. In our fits, which incorporate the unfolded ssDNA/ssRNA (via $F_{eWLC}$ in series with $F_{WLC}$), certain parameters were held fixed. These included $L_{dsDNA,\text{linker}}$, $p_{dsDNA}$, and $K$, which were obtained from a previous fit related to the folded state. Additionally, we used fixed values of $p_{ssDNA}$ (1.0 nm) and $p_{ssRNA}$ (0.9 nm).To determine the number of unfolded nucleotides based on the increase in contour length, we followed a two-step process: First, we calculated the conversion ratio between unfolded contour length and unfolded nucleotides. This calculation considered the diameter of the nucleic acid and assumed that the first base pair was already opened due to fraying. Second, we divided the unfolded contour length of each intermediate state by this ratio and added this value to the two frayed nucleotides. This yielded the number of unfolded nucleotides associated with each intermediate state.

In the presence of trigger strand, as illustrated in Fig. S8 (with a detailed contour length explanation), the extension increases as the toehold-bound (TB) state transitions to the fully invaded state (FI). This extension gain is attributed to two components: the double-stranded invader-target complex (invader-5' toehold hairpin) and the single-stranded incumbent (3' toehold hairpin) (refer to the schematics in Fig. S8). Given the increase in both single-stranded and double-stranded extension, fitting the force-extension trace with a series of eWLC and WLC models while keeping the double-stranded part $L_{dsDNA}$ fixed becomes impractical. To address this, we express the additional double-stranded contour length ($L_{\text{ds,bm}}$) in terms of the additional single-stranded contour length ($L_{\text{ss,bm}}$). This approach eliminates the need for an extra fitting parameter. Notably, a fixed constant ratio of $L_{\text{ss,bm/nt}}$ (contour length per nucleotide) to $L_{\text{ds,bm/bp}}$ (contour length per base pair) is required to establish the connection between these two variables. In our calculations, we employed values of 0.59 nm/nt for single-stranded DNA, 0.6 nm/nt for single-stranded RNA, 0.34 nm/bp for double-stranded DNA-DNA, 0.30 nm/bp for DNA-RNA, and 0.28 nm/bp for RNA-RNA. These values were used to determine the theoretical values of the opened

contour length. To provide context for these values, it's worth noting that the total unfolded contour length was in good agreement with an unfolding involving 110 nucleotides. This calculation takes into account a conversion factor of 0.6 nm/nt for single-stranded RNA and 0.59 nm/nt for single-stranded DNA, while also considering the nucleic acid diameter. When fitting, for example, the FI state of RRp2, we once again employ a series of eWLC and WLC models. However, in this case, we tailor them to suit the invasion process, making necessary adjustments as outlined in (4):

$$L_{ssDNA/ssRNA} \rightarrow L_{ss,bm}$$
$$L_{dsDNA,linker} \rightarrow L_{dsDNA,linker} + \frac{L_{ds,bm/bp}}{L_{ss,bm/nt}} \cdot L_{ss,bm} \quad (4)$$

Where, $L_{\mathrm{dsDNA,linker}}$, $p_{\mathrm{dsDNA,linker}}$, and K are held constant, and their specific values are determined through fitting the TB state using the eWLC model. The sole fitting parameter in this scenario is $L_{\mathrm{ss,bm}}$.

Passive mode:

In absence of trigger strand. In passive mode experiments, we converted the previously state-assigned force data into a contour length by reversing eq. (2) and (3). This transformation employed the elastic parameters obtained during the stretch and relax cycles[72]. Subsequently, we followed the same two-step process used in the stretch and relax cycles to calculate the number of unfolded nucleotides for each intermediate. This approach also enabled us to extract free energies and kinetic rates for these states, as illustrated in Figure 2 b & c, Figure 3 b, and Figure 5 c. In the presence of trigger strand: In the case of an invasion/reinvasion transition, as shown in Fig. 3C, state assignment was performed in the same manner as for unfolding/refolding transitions. To determine the opened contour length from the hairpin, we once again applied the contour length transformation from previous work4. However, this time, we used the eWLC and WLC models with adjusted parameters derived from Eq. (4)

For each trace at a given trap distance, we calculated the population probabilities of the various states at equilibrium by accumulating the total time spent in each state. We were able to compute a 1D folding energy landscape along the hairpin contour length coordinate based on knowledge of the free energies of all folding intermediate states and their respective transition states. The population probabilities of the different states, extracted from each trace at a specific trap distance, can be described by the Boltzmann distribution, which relates these probabilities to the free energy differences between the folded and unfolded states. This relationship can be expressed as $\Delta G_{ij}\ (F_i,\ F_j) = G\ (F_j) - G\ (F_i)$, where $P_i\ (F_i)$ represents the probability of the system being in state $i$ under the force $F_i$.

$$\Delta G_{ij} = -k_B T ln \frac{P_j}{P_i} \quad (5)$$

In the absence of trigger strand, to calculate the free energies at zero load for each constant distance step, we employed a model previously introduced for protein folding under force, which

comprehensively incorporates all energetic contributions from the dumbbell assay (as illustrated in Figure 1 and 6, dumbbell assay)[73]. The system, consisting of beads-dsDNA-ssDNA/ssRNA-dsDNA-beads, is simplified to an equivalent bead-dsDNA-ssDNA/ssRNA system. In this simplified system, we consider an effective trap stiffness given by $k_{eff}^{-1} = k_1^{-1} + k_2^{-1}$, a combined contour length that encompasses both dsDNA linkers, and an effective bead deflection $x_{eff} = |x_1| + |x_2|$. Because passive mode traces are not constant force measurements, we must account for changes in extension and force during folding/refolding transitions in our free energy calculations. The total energy $G_i(F_i)$ stored in the $G_{device}$ includes bead-dsDNA-ssDNA/ssRNA system at force $F_i$ can be divided into several components: the energy stored in the displacement of the beads $G_i^{beads}(F_i)$, the energy stored in the stretched dsDNA linkers $G_i^{dsDNA,linker}(F_i)$, and the energy associated with the unfolded ssDNA/ssRNA $G_i^{ssDNA/ssRNA}(F_i)$, along with the free energy $\Delta G_{0,i}^{DNA/RNA}(F_i)$ of the DNA/RNA toehold hairpin in state i:

$$G_i(F_i) = \Delta G_{0,i}^{DNA/RNA}(F_i) + G^{device}(F_i) = G_i^{beads}(F_i) + G_i^{dsDNA,linker}(F_i) + G_i^{ssDNA/ssRNA}(F_i)$$

The calculations for each of these individual terms can be derived using equations (5-7):

$G^{bead}(F) = \frac{1}{2}k_{eff}^{-1}F^2$ (5) is the energy of the bead displacement in the harmonic trap potential, the energy of the DNA handle includes dsDNA handle and adapter strand, which can be determined by integrating $G^{dsDNA,linker}(F) = \int_0^{xeWLC(F)} F_{eWLC}(x')dx'$ (6) and the elastic unfolded DNA hairpin. $G^{ssDNA,linker}(F) = \int_0^{xeWLC(F)} F_{WLC}(x')dx'$(7).

We computed the free energy differences *ΔGij* (*Fi, Fj*) for a transition from state *i* to *j* and forces bias *Fi*, *Fj* is then given by

$$\Delta G_{ij}(F_i, F_j) = G_j(F_j) - G_i(F_i) = -k_B T \cdot ln\left(\frac{P_j(F_j)}{P_i(F_i)}\right) \quad (8)$$

By applying the Boltzmann equation, we can calculate the free energy differences between various DNA/RNA toehold hairpin intermediate states at zero load as shown in (9):

$$\Delta G_{0,ij} = -k_B T \cdot ln\left(\frac{P_j(F_j)}{P_i(F_i)}\right) - \Delta G^{beads}(F_i, F_j) - \Delta G^{dsDNA,linker}(F_i, F_j) - \Delta G^{ssDNA/ssRNA}(F_i, F_j). \quad (9)$$

The HMM assigned state probabilities were used to obtain $P_i(F_i)$ and $P_j(F_j)$.

In the presence of trigger, for both invasion and re-invasion processes, the extension increase is once more divided into single-stranded and double-stranded components. Consequently, it is necessary to adjust the parameters as illustrated in (10)

$$\Delta G^{ssDNA/ssRNA} \rightarrow \Delta G^{ss,bm}$$
$$\Delta G^{dsDNA,linker} \rightarrow \Delta G^{dsDNA,linker+ds,bm} \quad (10)$$

Where, the contour length in the WLC integrals is adjusted as in (4).

Calculation and Extrapolation of transition rates:

We calculated the force dependent rate constants between different intermediate states, which not only serve as an alternative method for free energy difference calculation but also provide a quality control to verify the state assignment from the HMM analysis. To characterize the transition rate from each state, we firstly plotted integrated probability versus time length, called dwell time of each state.

We expected a single-exponential distribution of the dwell time for the first-order reactions,

$$p(t) = \frac{\exp(-kt) - \exp(-k\tau_{min})}{\exp(-k\tau_{max}) - \exp(-k\tau_{min})} \tag{11}$$

In this case, the maximum value, τ max, for the total length of the time trace was set, while the minimum value, τ min, was set to 0.2 ms. The fit returns the force-dependent rate constant k = $k_i$ ($Fi$) of state *i* at force $F$. As a result, the transition rate constant $kij$ can be calculated by utilizing the probability p$ij$, which is obtained by counting the corresponding events in the passive mode trace.

$$k_{ij} = k_i p_{ij} \tag{12}$$

In situations where the kinetics are exceptionally rapid and the contour length undergoes only minor changes, the traditional HMM may struggle to accurately assign data points. To address this challenge, we employed an alternative approach by utilizing the discrete unnormalized autocorrelation function (11). Due to the rapid kinetic process of toehold binding and branch migration, we used autocorrelation function to verify the maximum response rate of the system. Regarding the time duration of toehold binding and branch migration, we fitted the data with simple exponential function $F(t) = ae^{-\tau t}$ (Fig. 2 C) and compared with the response rate from the results of above autocorrelation function.

$$x^*\left(\tau = \frac{j}{\delta}\right) = \frac{1}{N}\sum_{i=0}^{N-1}(x_i - \mu_x)(x_{i-j} - \mu_x) \tag{13}$$

The autocorrelation function was fitted with a double-exponential function, considering the time lag τ of a trace x(t) and the mean μx,

$$f(\tau) = A_0 \exp(-\lambda_0\tau) + A_1 \exp(-\lambda_1\tau) \tag{14}$$

Where $\lambda_0$ corresponds to the fast kinetics resulting from Brownian diffusion of the beads. While $\lambda_1 = k_{ij}(F) + k_{ji}(F)$ can be utilized to determine the transition rates of toehold hairpins or branch migration at a specific force after fitting auto-correction function. Additionally, the detected rate increases with force as the stiffness of the dumbbell system intensifies.

In the presence of the trigger strand, we analyzed the passive mode traces using the HMM-assigned states. We constructed dwell time distributions, which were represented as normalized integrated histograms (see Fig. S2). These distributions were fitted using a single exponential function, accounting for the experimental time frame and finite time resolution[74]. This allowed us to determine rate constants, denoted as $k_{ij}(F_i)$, at the force Fi corresponding to the starting state i of the transition. Since a folding/refolding transition involves step-by-step zipping/unzipping events under applied force, the contour length increase serves as a well-defined reaction coordinate. However, it's important to note that the transition state position depends on the applied force, and a simple Bell model isn't applicable in this context. Instead, we adapted a model commonly used for folding/refolding of globular proteins[75] and coiled coils[73] under force to describe the folding/refolding of nucleic acids. This model considers energy changes of dsDNA linkers, springs, and unfolded ssDNA/ssRNA within an energy barrier $\Delta G_{iT}^{\#}$. In a two-state system, the free energies of the folding intermediate states $Gij$ were calculated from the force-dependen transition rates are then described by equation (15):

$$k_{ij}(F) = k_{0,i}\exp\left(\frac{-\Delta G_{iT}^{\#}(F_i=F,F_T)}{k_BT}\right) \quad (15)$$

The folding rate constant at zero load, denoted as $k_{0,i}$, was determined as a fit parameter (y-intercept in Fig. 3 d) using the equation provided in (16), which defines the force-dependent energy difference $\Delta G_{iT}^{\#}$ from state *i* to the transition state *T*.

$$\Delta G_{iT}^{\#}(F_i,F_T) = \Delta G_{iT}^{beads}(F_i,F_T) + \Delta G_{iT}^{dsDNA,linkers}(F_i,F_T) + \Delta G_{iT}^{ssDNA/ssRNA}(F_i,F_T) \quad (16)$$

The second fit parameter that defines the slope of the fits in Fig. 3d is the contour length difference, denoted as $\Delta L_{iT}^{\#}$, which represents the change in contour length from state *i* to the transition state *T*, where the force acting on the transition state T between two intermediate toehold hairpin states *i* and j is denoted as $F_T$.

With the trigger strand present, we followed a similar approach as in the previous section for extracting free energies. In the invasion/re-invasion transition, it was necessary to adapt the free energy contributions of both the double-stranded and single-stranded regions, as outlined in (10).

The Berkemeier-Schlierf model[75] is used here to evaluate the force-dependent transition rate. We fitted and extrapolated the force dependent transition rates $k_{ij}$ $(Fi)$ from state $i$ to state $j$ (e.g., Fig. 3 d), also considering the force dependent energy difference $\Delta G_{iT}$ $(Fi,\ FT)$ between state $i$ and the transition state TS.

$$k_{ij}(F_i) = k_{ij}(F_i=0)\exp\left(-\frac{\Delta G_{iT}(F_i,F_T)}{k_BT}\right) \quad (17)$$

The elastic parameters and contour length of the system determine the curvature on state *i*, leaving the zero-force rate constant $k_{ij}$ ($Fi$ = 0) and the contour length difference $\Delta LiT$ from state $i$ to the transition state $k_{ij}$ as the only free fitting parameters.

Transition path time analysis and autocorrelation analysis:

The transition path time $\tau$ for the invasion transition in Figure 2 b was fitted using an exponential function, as shown in (18):

$$F(t) = y_0 + A\ exp(-(t_0 - t)/\tau) \quad (18)$$

To determine the system's response time, we employed a technique based on the autocorrelation function of the force vs. time signal in passive mode traces, as previously described[39].

Strand displacement and mean first passage time for a 1D random walk:

The strand displacement process can be effectively described using a master equation, which models it as a one-dimensional random walk among N+1 states.

These states range from state 0 to state N. In this framework, each state, denoted as 'j', is capable to transition to its neighboring states j-1 and j+1 with respective rate $k_j^-$ and $k_j^+$.

The time evolution of the probability, $P_{(j,t)}$, which represents the likelihood of the system being in state j at time t, can be described as follows:

$$\frac{dP(j,t)}{dt} = k_{j-1}^+ P(j-1,t) + k_{j+1}^- P(j+1,t) - \left(k_j^- + k_j^+\right)P(j,t). \quad (19)$$

The Mean First Passage Time (MFPT) for reaching the absorbing boundary at N- (with state 0 as the initial state) can be calculated using the following expression. This expression is derived from the solution of the backward Kolmogorov equation[76,77].

$$T_0 = \sum_{m=0}^{N-1} \sum_{l=0}^{m} \frac{1}{k_l^+} \prod_{j=l+1}^{m} \frac{k_j^-}{k_j^+} \quad (20)$$

$k_l^+$ is the transition rates between states *j* and *j*+1, are related to the energies $\epsilon_j$, $\epsilon_{j+1}$ via:

$$\frac{k_j^+}{k_{j+1}^-} = e^{(-\beta(\epsilon_{j+1} - \epsilon_j))} = e^{(-\beta \Delta \epsilon_j)} = \omega_j \quad (21)$$

Where $\beta = (k_B)^{-1}$. We make the assumption that all energy differences are equal, resulting in all $\omega_j$ values being identical.

We further assume that all individual forward rates are uniform, denoted as r (i.e., $r \coloneqq k_j^+$ for all j). Consequently, $r = k_j^+ = \omega k_j^-$. This assumption simplifies the expression for $T_0$ to:

$$T_0 = \sum_{m=0}^{N-1} \sum_{l=0}^{m} \frac{1}{r} \prod_{j=l+1}^{m} \frac{1}{\omega}$$

(22)

When all energies are same: When all energies are equal, we get $\omega = 1$ and therefore:

$$T_0 = \sum_{m=0}^{N-1}\sum_{l=0}^{m}\frac{1}{r}\prod_{j=l+1}^{m}1 = \frac{1}{r}\sum_{m=0}^{N-1}\sum_{l=o}^{m}1 = \frac{1}{r}\sum_{m=0}^{N-1}(m+1) = \frac{1}{r}\frac{N(N+1)}{2}$$

(23)

i.e., $T_0 \approx N^2/2$ is multiplied by with the 'step time', $\tau_0 = \frac{1}{r}$. In this scenario, the process exhibits the anticipated diffusive scaling of approximately $N^2$.

When strand displacement is driven by an energy difference $\Delta\epsilon_j<0$, resulting in $\omega>1$, the forward rates are accelerated compared to the backward rate. In this scenario, $T_0$ can be expressed as follows:

$$T_0 = \frac{1}{r}\sum_{m=0}^{N-1}\sum_{l=0}^{m}\frac{1}{\omega^{m-l}} = \frac{1}{r}\sum_{m=0}^{N-1}\frac{w-(1/\omega)^m}{\omega-1}$$

(24)

Which results in:

$$T_0 = \frac{1}{r}\frac{\omega}{(\omega-1)^2} = [N\omega - N - 1 + (1/\omega)^N] \quad (25)$$

When $\omega>1$ and $N$ is sufficiently large, we can neglect the term $(\frac{1}{\omega})^N$. The expression for MFPT then becomes

$$T_0 \approx \frac{1}{r}\frac{\omega}{(\omega-1)^2} = [N\omega - N - 1] \approx \frac{1}{r}\frac{\omega}{(\omega-1)^2}N(\omega-1) = \frac{N}{r}\frac{\omega}{\omega-1}$$

(26)

Under this moderate assumption, we observe that $T_0$ scales approximately with $N$. If we were to straightforwardly calculate $T_0$ and divide it by $N$, the result would be:

$$T_0/N \approx \frac{1}{r}\frac{\omega}{\omega-1} \quad (27)$$

Now, in the scenario where $\omega \gg 1$, we can derive from Eq. (25):

$$T_0 \approx \frac{1}{r}\frac{\omega}{\omega^2} \times N\omega = \frac{N}{r} \quad (28)$$

Alternatively, we can derive it directly from Eq. (22) above. In this case, we have a simple relationship: $T_0/N=1/r$, which we can identify as the step time, $\tau_0$.

When considering $\omega$ the influence of force, we can, in a simplified view, associate the lower energy of state $j$ at position $x_j = x_0 + j \cdot \Delta x$ with a force $f$ as follows:

$$\epsilon_{j+1}(f) = \epsilon_{j+1}(0) - f \cdot (j+1)\Delta x = \epsilon_j(f) - f \cdot \Delta x \quad (29)$$

When all energies in the absence of force are equal, $\Delta\epsilon(f)$ can be expressed as $\Delta\epsilon(f) = -f \cdot \Delta x$, leading to $\omega = e^{(\beta f \cdot \Delta x)}$. The choice of $\Delta x$ is crucial in this context. For instance, if $\Delta x$ is on the order of 0.5 nm, then at a force f = 10 pN, the term $f \cdot \Delta x$ is $\approx k_B T$. Consequently, $\omega \approx e$, which is not significantly large. In this scenario, we would observe a scaling of $T_0 \sim N$, but the ratio $T_0/N$ would be larger than $\tau_0$.

# Chapter II

Intracellular sensing of endogenous gene expression in *Escherichia coli* using toehold switch riboregulators

**Key words:** Endogenous mRNAs • toehold riboregulatory • RNA sensing • diauxic growth

2

# Abstract

In the previous chapter, we delved into the intricacies of TMSD of both DNA and RNA toehold hairpins using SMFM. From this chapter, we shift our focus to explore the diverse applications of toehold riboregulator in RNA sensing, gene regulation, and RNA switches. RNA serves not only as the main vector of genetic information but also as a universal mediator underlying the diverse cellular conditions and functions. Despite the significant advancements in RNA sequencing technologies and transcriptome datasets in recent years, real-time sensing of *in vivo* RNA expression profiles under various cell states remains a challenging task. In this chapter, we present a modified toehold mRNA sensor built upon the previously developed toehold switch. This programmable RNA-sensing tool has the capability to bind to diverse endogenous mRNA targets and trigger mCherry expression in *E. coli*. We conducted tests on the expression of various target genes, including *rpsR*, *yfiA, cspD*, *ihfA*, and *lacZ etc.*, which are known to be upregulated in different cellular states and play crucial roles in stress responses. Albeit some of our toehold mRNA sensors effectively respond to overexpressed mRNAs with an ON/OFF ratio above 10, the presence of leakage in the toehold switch configuration diminishes the sensitivity for detecting the precise timing of endogenous mRNA expression.

# Introduction

RNA-based gene regulatory mechanisms have been elucidated across diverse biological domains, spanning prokaryotes, eukaryotes, and even viruses[78-80]. These mechanisms encompass essential processes like CRISPR and RNA interference (RNAi), which are mediated by RNA-protein complex.

They require the involvement of non-coding RNA molecules, such as cr/tracrRNA and miRNA, to modulate gene expression at both the transcriptional and translational levels[81,82]. In prokaryotes, many mRNA molecules carry complex folded domains in their 5' -untranslated regions, which are known as riboswitches[83]. Additionally, small regulatory RNAs (sRNAs) play a pivotal role in post-transcriptional gene regulation through partial base-pairing with target mRNAs. Unlike their counterparts in eukaryotes, most of riboswitches and sRNAs in prokaryotes often operate independently of additional protein factors. Small metabolites binding to riboswitches or directly to target mRNAs elicit allosteric conformational changes that modulate transcription termination efficiency or translation initiation within the mRNA molecules[84,85].

According to our previous discussion of the toehold switch in the preceding chapter, it's worth noting that the toehold switch can effectively bind to small trigger RNA strands, subsequently activating a fluorescence reporter *in vivo*. This capability has been experimentally confirmed by detecting induced endogenous sRNA, *ryhB*[11], within *E. coli*. Nevertheless, it demonstrated a relatively modest ON/OFF ratio of $\approx$ 2. In this chapter, we investigated their application of toehold switch as intracellular sensors to monitor the expression profiles of particular genes in *E. coli*, specifically those linked to several cellular stringent response and growth phases (Fig. 7 b). Firstly, ribosomal RNAs (rRNAs) engage in binding interactions with small ribosomal subunit proteins to fortify the structural integrity of the ribosome's 30S subunit platform. It is noteworthy that *rpsR*[86], a specific rRNA, forms a heterodimeric complex with the bS6 subunit. Importantly, the expression of *rpsR* remains consistently elevated under normal growth condition, especially during the logarithmic growth phase, making it a valuable indicator reflective of sustained growth conditions.

Secondly, the *lacZ* gene[87], which encodes β-galactosidase, is regulated by the lac operon when lactose is absent in the cell as the sole carbon source. This enzyme actively facilitates the hydrolysis of β-galactosides into monosaccharides during glucose-lactose diauxic growth16. This gene is located within the *lac* operon of the *E. coli* genome and experiences significant upregulation in response to lactose when glucose is depleted in the medium[88].

During the stationary phase, certain proteins, including cold shock protein YfiA and the single-strand DNA/RNA binding protein CspD, are known to be highly expressed. YfiA, encoded by the *yfiA* gene, binds to the 30S subunit of the ribosome during the late log phase or at the initiation of the stationary phase[89]. This binding induces compositional and conformational changes in the 70S ribosomes, ultimately repressing global translation during growth arrest. Likewise, the expression of the *cspD* gene is significantly upregulated during growth arrest, leading to the

inhibition of both DNA replication and elongation, critical for long-term survival[90]. Additionally, the integration host factor (IHF), encoded by the *ihfA* and *ihfB* genes[91], is involved in various cellular processes, such as DNA replication, gene transcription regulation, and site-specific recombination during the growth arrest[92,93]. These regulations result in a substantial increase in gene expression and protein production, approximately 5-fold and 3 to 4-fold, respectively, compared to the exponential phase[94,95]. In extreme conditions, such as during the DNA damage response, cells upregulate genes associated with the SOS response, such as *sulA*, to prolong cell survival by inhibiting cell division[96].

Differing from the aforementioned *in vivo* sensing of sRNA using the toehold switch based on endpoint values, our objective was to achieve a real-time monitoring of the fluctuation of fluorescence signals throughout the growth curve under various growth phases or stress responses. We expected that the naturally triggered upregulation of these genes will result in an increase in the fluorescence signal from our toehold mRNA sensor, as evidenced by a shift in the normalized fluorescence curve when compared to the controls.

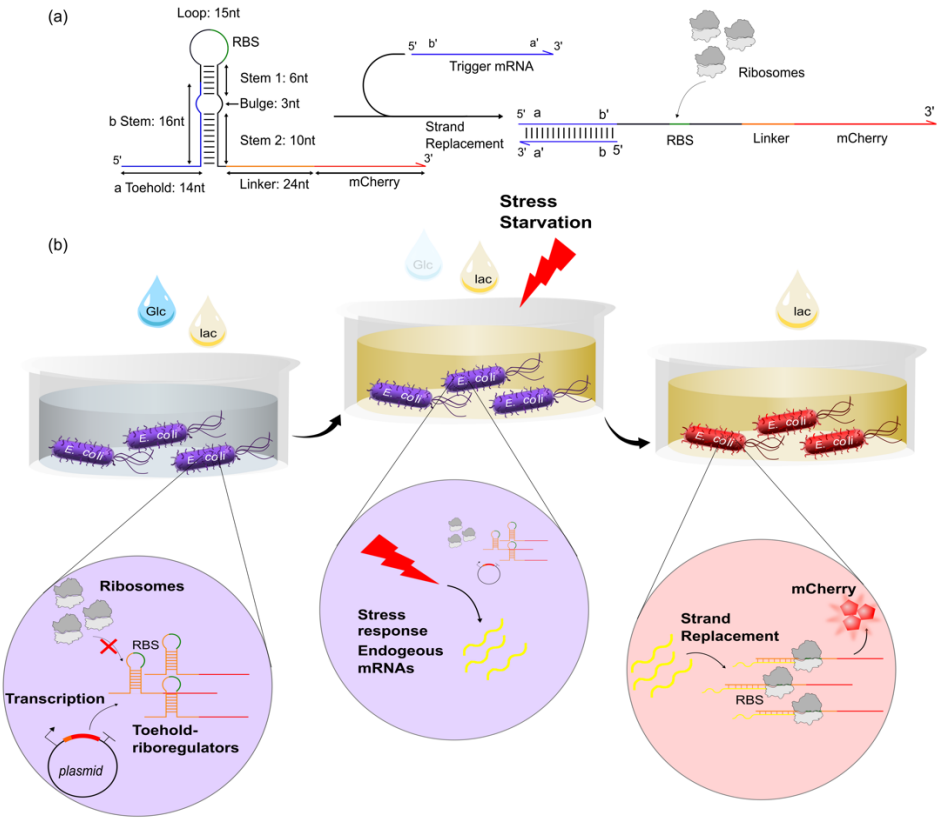


Figure 7. Schematics of toehold switch's architecture. a), the toehold switch features a hairpin structure located at the 5' end of the reporter gene. This hairpin structure acts as a regulatory element, modulating translation strength by sequestering the ribosome binding site (RBS) sequence within a 15 nt loop. Trigger RNAs, carrying a 30 nt complementary sequence to the toehold binding region (14 nt) of the hairpin, can bind to this region. This binding initiates a strand replacement reaction that opens the hairpin structure. This conformational change exposes the RBS region, facilitating translation initiation. b), toehold-mRNA sensors are transcribed from a recombinant plasmid in *E. coli*. These sensors consist of mRNA molecules with a toehold hairpin structure at their 5' end, effectively suppressing translation initiation under normal growth conditions. However, during stress or starvation conditions (e.g., exposure to mixtures of carbon sources, glucose, and lactose), specific stringent

response genes (highlighted in yellow) become upregulated. The toehold-mRNA sensors function as detectors of mRNA triggers and undergo toehold-mediated strand displacement (TMSD), resulting in the unwinding of the hairpin structure. This process enables the translation of the mCherry reporter gene, facilitating the monitoring of fluorescence signals

# Results

Design and characterization of the toehold switch mRNA sensors

In this chapter, we have developed a collection of toehold switch mRNA sensors based on the earlier design (Fig. 7 a) elucidated in the aforementioned toehold switch paper[11]. To detect the presence of endogenous mRNAs, our initial step involved identifying unpaired regions within the previously mentioned stress response genes using the nucleic acid analysis tool NUPACK[97]. Subsequently, these candidate sequences were utilized to fine-tune the toehold region and stem sequence (Fig. 7 a, blue region) to enable binding to the target mRNAs *in vivo*. In addition, we replaced the GFP reporter gene with mCherry to reduce the background fluorescence signal.
To characterize the dynamic range of our toehold mRNA sensors, we initially conducted co-transformation experiments. Recombinant plasmids containing the toehold switch mRNA sensors, along with their corresponding trigger RNAs regulated by the *ptet* and *pt7* promoters, were co-transformed into *E. coli* BL21 DE3. We measured the endpoint fluorescent signals for each strain after an overnight culture using microplate reader (See Material and Methods). As illustrated in Fig. 7 b-f, our *in vivo* experiments with *E. coli* BL21 DE3 transformed with plasmids containing the toehold switch sensors and their corresponding triggers revealed varying ON/OFF ratios ranging from $\approx$ 2.6 to $\approx$ 10.8. Specifically, the *rpsR*, *infA*, and *cspD* sensors exhibited a substantial dynamic range (ON/OFF > 8), similar to that of the original toehold switch under identical conditions (Appendix, Fig. S30). In contrast, the *lacZ* sensor displayed the lowest ON/OFF ratio of $\approx$ 2.6, possibly due to the longer trigger strand (around 3 Kb). In addition, we chose the *bfp* gene as the reporter for the *yfiA* sensor, which exhibited a relatively higher leaky signal in the OFF state. These results validate that the toehold switch mRNA sensors effectively respond to overexpressed trigger RNAs within an acceptable dynamic range in *E. coli* compared to the original toehold switch configuration.

*In vivo* real-time characterization under growth arrest

After characterizing the dynamic range of our toehold mRNA sensors, we proceeded by transforming recombinant plasmids containing only the toehold switch sensors into *E. coli* BL21 DE3 strain. To monitor the growth curve and fluorescence signal of the bacterial strain, we initiated the process by pre-culturing a single colony until it reached a certain cell density. Subsequently, we inoculated the pre-cultured bacteria into a 96-well plate for overnight culture under a microplate reader (as detailed in the Materials and Methods section).

As depicted in Figures S14 and S20, when *E. coli* carrying the toehold sensors for *rpsR*, *ihfA*, *cspD*, and *yfiA* were cultured in either M9 or LB medium, we observed noticeably similar kinetic patterns in relative fluorescence signals (measured as fluorescence/absorbance $OD_{600}$) when compared to the positive control. The positive control solely carried the same fluorescent reporter gene, devoid of any toehold switch regulation but driven by the same promoter. Notably, the fluorescence signals for *rpsR*, *ihfA*, and *cspD* sensors exhibited a continuous increase during cell growth, initiating from the exponential phase and stabilizing upon entering the stationary phase in M9 medium. Remarkably, we detected no significant fluctuations in fluorescence signals beyond the stationary phase, even with concurrent upregulation or downregulation of the corresponding gene expressions of *ihfA* and *rpsR*.

Regarding the *cspD* and *yfiA* sensors, we noted a faster growth rate in *E. coli* carrying the same toehold sensors plasmid (e.g., doubling time of 66 minutes vs. 47 minutes for the *cspD* sensor in M9 and LB medium respectively). However, they displayed similar kinetic patterns in relative fluorescence signals when compared to their corresponding positive controls, as we observed in M9 medium. The fluorescence signal began to increase as the growth entered the stationary phase for both strains. Due to the rapid cell division and growth of the cell population in LB medium, the expression of the fluorescent reporters was repressed. Consequently, it became challenging to discern whether the increasing fluorescence signal from the toehold mRNA sensor was influenced by the upregulation of the *cspD* and *yfiA* genes during the stationary phase.

In addition, we conducted tests on our toehold mRNA sensors in response to starvation during diauxic growth using M9 medium with a mixed glucose-lactose carbon source (see Materials and methods). The results revealed a clear diauxic lag time[98] (for instance, ≈ 13.9 minutes for *E. coli* carrying the *yfiA* sensor (Fig. 9 c) – which allowed bacteria cells to adapt during the diauxic shift from one carbon source to another. However, much like the previously observed kinetic patterns in *E. coli* cells grown in M9 medium with a single carbon source, we did not detect significant fluctuations in fluorescence signals beyond the diauxic lag time, even with concurrent upregulation or downregulation of the corresponding gene expressions of *rpsR*, *yfiA*, and *cspD* (Fig. 9). We also tested the toehold *lacZ* sensor in the *E. coli* same strain, which retains the *lacI* gene on the genome to repress the constitutive expression of *lacZ* in the absence of lactose. The results exhibited similar diauxic behaviors as observed with the aforementioned toehold sensors in the same M9-glu-lac medium, although each replicate exhibited large differences in growth patterns under the same conditions (Fig. S19). Similar to the previous sensors, the toehold *lacZ* sensor also displayed a comparable kinetic pattern in line with the corresponding positive controls, albeit one replicate exhibited a significant increase in relative fluorescence signal production following the diauxic shift. This behavior distinguished it from both the other replicates and the positive control within the same time duration.

Lastly, apart from the characterization of starvation response, we also aimed to investigate the SOS response in *E. coli* cells using our toehold *sulA* sensor. To trigger the SOS response, we added $H_2O_2$ into the 200 µl culture, reaching a final concentration of 1.5 µM after a 4-hour pre-

culture, during which the absorbance reached approximately 0.4 (see details in Materials and methods). The results, depicted in Fig. S15, revealed that $H_2O_2$ induced a growth arrest, reminiscent of the diauxic shift observed during starvation. While the sensor group exhibited a more extended diauxic lag time ($\approx$ 93 minutes) compared to the corresponding negative controls ($\approx$ 24 minutes), we still observed similar kinetic patterns in comparison to the positive controls after recovering from the growth arrest. For both the *E. coli* cells carrying the *sulA* sensor and those with only the *mCherry* gene, the fluorescence signal began to exponentially increase following $H_2O_2$ induction, although with slower fluorescence signal production from the *sulA* sensor due to the extended diauxic lag time.

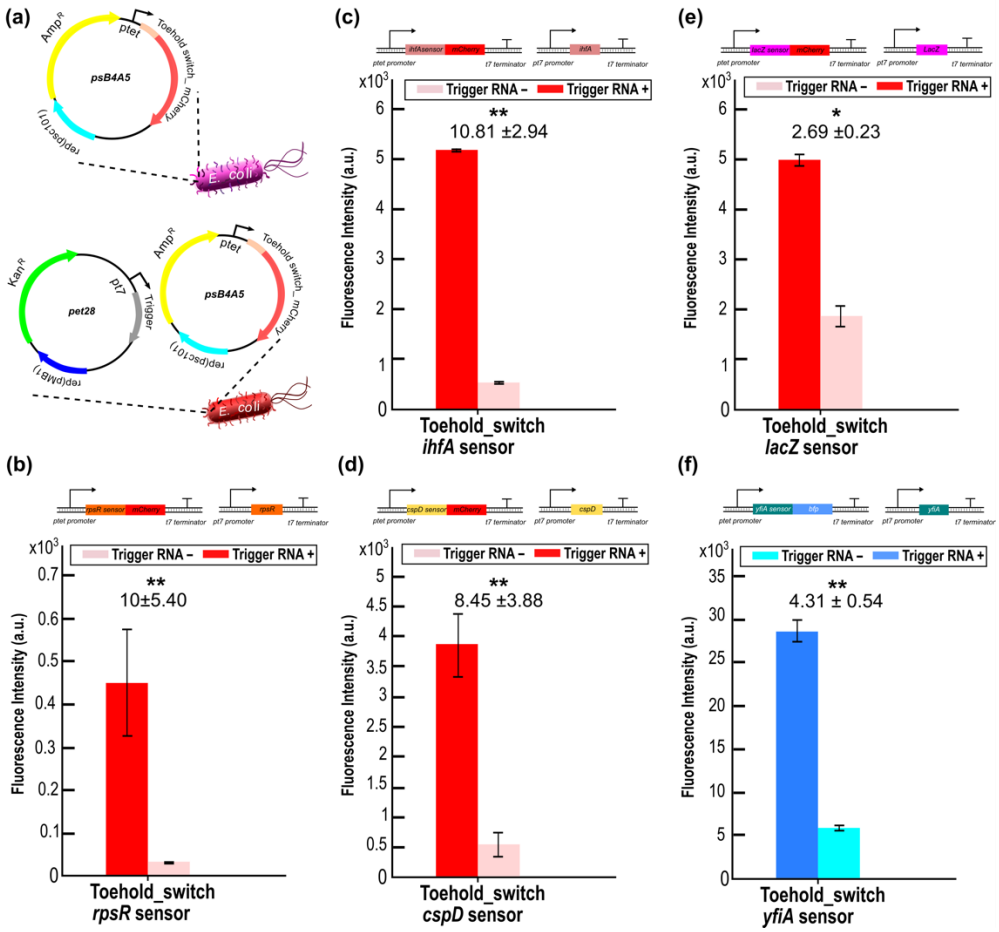


Figure 8. Characterization of the dynamic range of toehold-switch mRNA sensors in *E. coli*. a) Transformation of recombinant plasmids containing toehold-switch sensors and triggers into *E. coli*. b) Relative fluorescence intensities of mCherry from *in vivo* measurements in the ON and OFF states for the toehold switch *rpsR* sensor, with the No/Off ratio highlighting the relative dynamic range. c), Relative fluorescence intensities of mCherry obtained from *in vivo* measurement with in the ON and OFF state for toehold switch *ihfA* sensor, No/Off ratio indicates the relative dynamic range. d), Relative fluorescence intensities of mCherry obtained from *in vivo* measurement with in the ON and OFF state for toehold switch *cspD* sensor, No/Off ratio illustrates the relative dynamic range. e), Relative fluorescence intensities obtained from *in vivo* measurement with in the ON and OFF state for toehold switch *lacZ* sensor, No/Off ratio shows the relative dynamic range. f), Relative fluorescence intensities of BFP obtained from *in vivo* measurement with in the ON and OFF state for toehold switch *yfiA* sensor, No/Off ratio demonstrates the relative dynamic range. For both the relative fluorescence/OD and fluorescence intensity data, Welch's t-tests were performed for each construct; *P < 0.05, indicating conditions where the fluorescence/OD and fluorescence intensity for the trigger RNA + condition is statistically significantly different

from that of the trigger RNA- condition. Error bars represent the standard deviation (s.d.) for biologically independent samples.

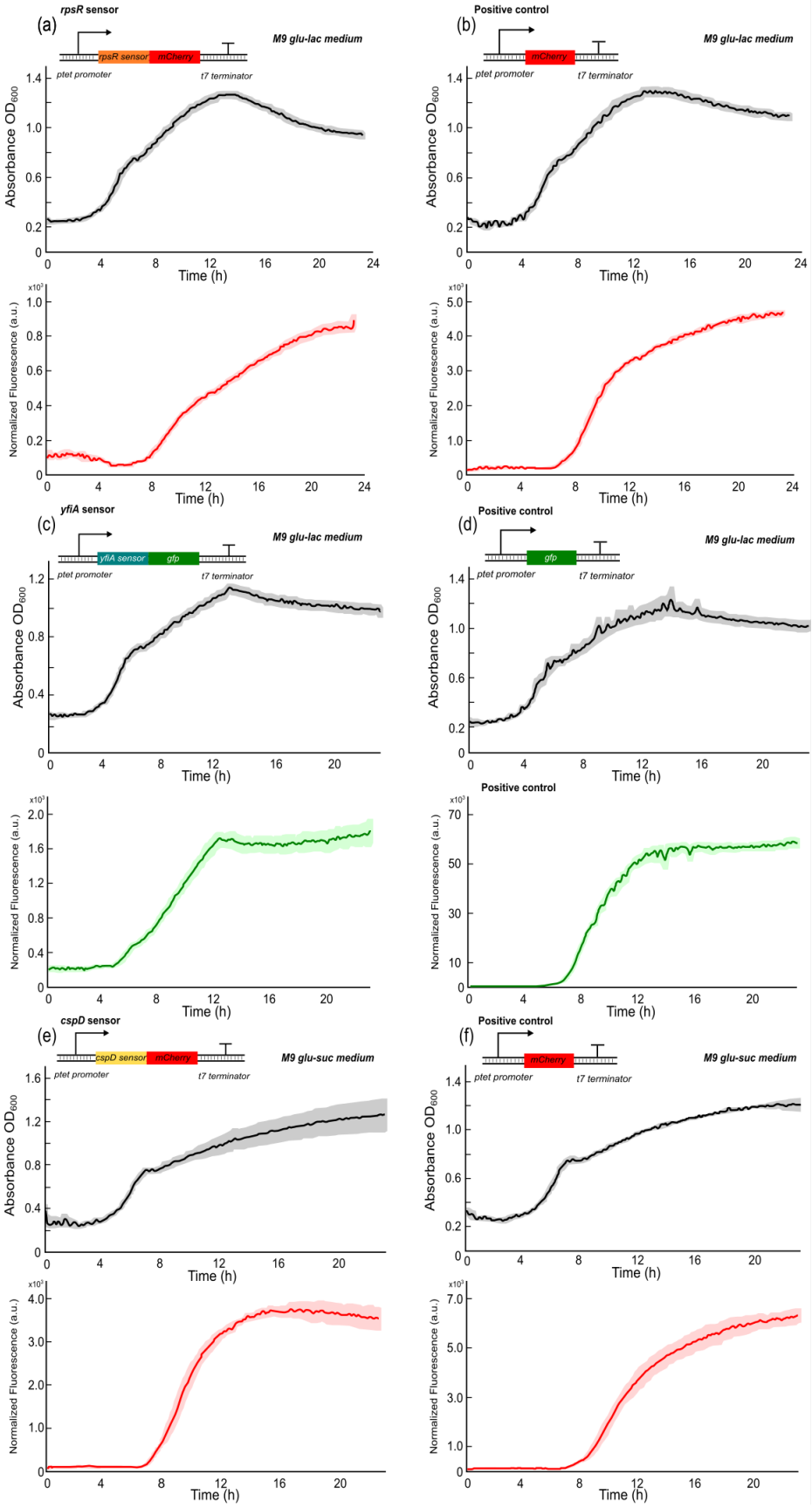

Figure 9. Growth Curves and Fluorescence-Based Kinetic Curves of *E. coli* BL21 DE3 and DH5α Z1 with the toehold mRNA sensors responding to starvation. a), *rpsR* Sensor with mCherry (M9-glu-lac Medium, 37°C): In the top graph, we observe the growth curve of *E. coli* BL21 DE3 strains carrying the *rpsR* sensor fused to the *mCherry* gene. The bottom graph presents the kinetic curve, showing the relative fluorescence compared to absorbance ($OD_{600}$). These experiments were conducted under conditions of M9-glu-lac medium at 37°C. b), Positive Control with mCherry (M9-glu-lac Medium, 37°C): In the top graph, we display the growth curve for *E. coli* BL21 DE3 carrying only the *mCherry* gene as a positive control. The bottom graph illustrates the kinetic curve of relative fluorescence to absorbance. The conditions for these cultures matched those of the *rpsR* sensor experiments, using M9-glu-lac medium at 37°C. c), *yfiA* Sensor with GFP (M9-glu-lac Medium, 37°C): The top graph represents the growth curve for bacterial strains carrying the *yfiA* sensor fused to the *gfp* gene. In the bottom graph, we present the kinetic curve of relative fluorescence to absorbance. These experiments were conducted in M9-glu-lac medium at 37°C. d), Positive Control with GFP (M9-glu-lac Medium, 37°C): In the top graph, we illustrate the growth curve for *E. coli* BL21 DE3 carrying only the *gfp* gene as a positive control. The bottom graph displays the kinetic curve of relative fluorescence to absorbance. The culture conditions matched those of the *yfiA* sensor experiments, using M9-glu-lac medium at 37°C. e), *cspD* Sensor with mCherry (M9-glu-suc Medium, 37°C): The top graph shows the growth curve for bacterial strains carrying the *cspD* sensor fused to the *mCherry* gene. In the bottom graph, we present the kinetic curve of relative fluorescence to absorbance. These experiments were conducted in M9-glu-suc medium at 37°C. f, Positive Control with mCherry (M9-glu-suc Medium, 37°C): The top graph illustrates the growth curve for *E. coli* BL21 DE3 carrying only the *mCherry* gene as a positive control. The bottom graph displays the kinetic curve of relative fluorescence to absorbance. The culture conditions matched those of the *cspD* sensor experiments, using M9-glu-suc medium at 37°C. The relative fluorescence signals/absorbance values for the sensors are presented as mean values of background-subtracted fluorescence levels. Shaded error bars represent the standard deviation (s.d.) for three replicates from individual colonies, with the positive controls data in (d), and (f) based on two replicates.

# Discussion

*In vivo* endogenous mRNA sensing holds great promise, offering not only a valuable tool for monitoring the dynamic states of living cells but also the potential to assist in the regulation of complex gene networks. In contrast to advanced techniques like MERFISH[99] or other single-cell RNA sequencing[100], *in vivo* mRNA sensing can provide a unique insight into dynamic cellular processes and extends its applicability even to tissue architecture, beyond the boundaries of traditional transcriptome profiling. To date, several notable efforts, such as CellREADR[101] and CRISPR-based TIGER[102], have leveraged both transcriptional and translational mechanisms in mammalian cells. These approaches enable the detection of specific cellular RNAs and the subsequent activation of downstream genes for monitoring and manipulating cellular behavior. However, these methods often rely on a combination of RNA Watson–Crick base-pairing and effector proteins, such as ADARs and CRISPR base editors. While effective, the RNA-protein interactions can introduce complexities into the monitoring system and increase the metabolic burden on the cells.

In this chapter, we aimed to achieve real-time endogenous mRNA sensing in *E. coli* using only RNA-RNA interactions, building upon our previously developed toehold switch riboregulators. To target specific mRNA expressions, we focused on genes that exhibit distinct upregulation or downregulation patterns during growth arrest, including the stationary phase, periods of starvation, or SOS responses. Our results from real-time growth curves and fluorescence measurements revealed that the expected fluorescence signal profiles were not as anticipated,

primarily due to leaky expression from the toehold switch. For genes like *yfiA*, *cspD*, and *ihfA*, which are primarily upregulated during the stationary phase and growth arrest, we observed different expression profiles when bacteria were grown in LB and M9 media. Our toehold mRNA sensors exhibited some leakage signals during the exponential phase when these stationary phase genes were not expressed, which reduced the sensitivity and response range of our toehold mRNA sensors (see Fig. S14 and S20). Consequently, while we observed a diauxic shift during cell growth, the expression profiles remained similar to the positive controls.

Similarly, we detected leakage from the original toehold switch sequence in the absence of corresponding trigger RNA in BL21 DE3 during cell growth (see Fig. S16). This leakage from the toehold switch could be attributed to ribosome invasion of the RBS within the loop region. By measuring with different loop sizes of the toehold *cspD* sensor (see Fig. S17 and S18) in *E. coli*, we observed a reduction in leaky expression with a decrease in the size of the toehold hairpin loop, both during the exponential phase and at the final equilibrium fluorescence intensity after overnight culture. However, reducing the loop size may also affect the hairpin TMSD process and further narrow the dynamic range of the toehold mRNA sensor. We will explore other potential factors, such as transcription-translation coupling and target RNA concentration, in more detail in the next chapter. Additionally, when *E. coli* grows in a rich medium, such as LB medium, due to its fast growth rate, cells primarily produce proteins necessary for growth (R and P-class proteins) rather than unnecessary proteins like fluorescence reporters[103]. Consequently, we observed an increase in fluorescence signals for both toehold mRNA sensors and positive controls when the cells entered the stationary phase (refer to Fig. S14 and S16). These observations exhibited pseudo-positive expression profiles, seemingly activated by the upregulation of *cspD* gene during the stationary phase.

# Materials and Methods

Culture media

We utilized LB medium (Carl Roth) and M9 medium (M9 minimal salts, 5X, Sigma-Aldrich, 1 mM thiamine hydrochloride, 0.2% Casein H, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 20 mM glucose, all from Carl Roth) for cell cultures, depending on the purpose. Both mediums are utilized to culture the cloning strain (Turbo®), and the strain (BL21 DE3® and DH5α Z1, NEB) for gene expression regulated by the toehold mRNA sensor. The antibiotics carbenicillin (100 μg/ml) and kanamycin (100 μg/ml) were added to the media. For diauxic growth conditions, we supplemented the M9 medium with two different carbon sources: glucose 0.1% and lactose 0.4% or succinate 0.4% weight percentage.

Plasmid construction and cloning process

All DNA oligonucleotides were purchased from Eurofins Genomics, Ebersberg, Germany. All toehold mRNA sensor sequences were constructed by overlap extension PCR, restriction ligation and blunt-end ligation. We used overlap extension PCR to amplify the toehold switch sequence. Two DNA oligos that share complementary overhang sequences were then annealed in Phusion® Master Mix (NEB) at an appropriate annealing temperature, calculated with the NEB Tm calculator. DNA polymerase is then used to extend the 3' ends and fill up the gaps. Then we amplified the target sequence and added restriction sites as we described in the previous chapter. In the next step, toehold switch sequences were ligated with a reporter gene using restriction ligation. The reporter gene templates (*gfp* and *mCherry*) were obtained from the iGEM parts registry (http://parts.igem.org/Main_Page). Each reporter template was amplified, and restriction sites (*XbaI*, *PstI*) were added on both ends through primers (cf. Primer List). Each endogenous gene was amplified from the *E. coli* genome using primers contain restriction sites. Cloning vectors (modified pet28b, psb4A5) plasmids were also digested with *EcoRI* and *PstI* and gel purified to remove the digested strands. Last, all three parts (a ratio of inserts to vector of 1:3 was used) were ligated using T4 ligase (NEB) with the standard protocol. The following transformation and miniprep process has been described in previous chapter.

Cell culture

Bacterial strains for cloning were grown in both medium using 5 mL culture as we described in the previous chapter. For expression strains, which were used to characterize the expression of the toehold mRNA sensors, we picked up three colonies each from an LB agar plate, followed by a small-scale preliminary culture in M9 medium (500 µl) for 4 hours. This step helps the bacteria to adapt to the new growth conditions in minimal medium. After preliminary culture, we added 10 µl of the pre-culture into a final volume of 5 mL LB or M9 medium for large-scale culture in 200 mL shake flask. The production of T7 RNA polymerase by the BL21 DE3 cells was induced with 1 mM IPTG (Carl Roth) at 0.4-0.5 $OD_{600}$ after 4 hours of growth. For DH5α Z1, we added additional 20nM aTc to the medium to induce the transcription from the *ptet* promoter. Fluorescence measurements on the cell cultures with several replicates were obtained after overnight culture. For growth curve measurement, we added 1 µl of the pre-culture into a final volume of 200 µl M9 medium in 96-well plate (Plate 96 Well black, ibidi). The bacteria were shake cultured in microplate reader (CLARIOstar®, BMG LABTECH) with 200 rpm (double orbital) at 37 °C. To facilitate gas exchange, we created a small puncture hole in the film covering each well using a syringe needle with a diameter of 0.1 mm.

Fluorescence and absorbance measurements with microplate reader

After overnight culture, cell cultures (250 µl) were transferred to a 96-well plate and characterized via fluorescence and absorbance ($OD_{600}$) measurements using a microplate reader with the following settings (Excitation/Emission wavelength: 570-20/630-40 nm; gain value: 1000; focus

height: 2.4 mm). For growth curve measurement, data were measured every 5 min for both absorbance and fluorescence in each cycle up to 24 hours and further analyzed using MARS data analysis software (BMG LABTECH). $OD_{600}$ and fluorescence values for each replicate were first corrected by subtracting the values of a blank measurement obtained with culture medium (15)

$$F = \frac{F_0 - F_M}{OD_{600} - OD_{600,M}} \tag{30}$$

where F is the fluorescence of cells, FM is the fluorescence of media. $OD_{600}$ is the optical density of cells and $OD_{600,\,M}$ is the optimal density of media. A Welch's t-test was applied to determine the statistical significance (P <0.05 or 0.01) of the results obtained under different conditions.

To calculate the bacterial doubling time from the growth curves, we determined the exponential growth phase of the bacteria from a log-plot of the OD against time and fitted an exponential function to the corresponding data:

$$N(t) = N_0 e^{\lambda t} \tag{31}$$

$$t_d = \frac{ln_2}{\lambda} \tag{32}$$

Here $N_0$ represents the cell density at time t, $N_0$ is the initial cell density, is the growth rate, and is the doubling time.

To calculate the diauxic lag time[98], we determined the of the different growth phase of bacteria from a log-plot of the OD against time. Phase 1: cell growth on glucose until it is exhausted, phase 2: cells stop growing, phase 3: resumption of growth on lactose. We fitted each phase a linear function to the corresponding data. The diauxic lag time is determined by extrapolating the time (t) between two points where two straight lines intersect ($X_i$ and $X_f$) on a logarithmic plot of bacterial growth (measured at $OD_{600}$) against time.

$$X_i = \frac{b_2 - b_1}{a_1 - a_2}, X_f = \frac{b_3 - b_2}{a_2 - a_3} \tag{33}$$

$$t = X_i - X_f \tag{34}$$

Where $a_{1\text{-}3}$, $b_{1\text{-}3}$ are constants of linear function, the b as the intercept.

The SOS response was triggered by adding $H_2O_2$ at a sublethal concentration of 1.5 μM[95] after the bacteria had grown in a 96-well plate until they reached an $OD_{600}$ of 0.4. This step was undertaken to ensure that a specific cell density was maintained for further culturing. During this process, the measurement cycle was momentarily paused and then promptly resumed immediately after the addition of $H_2O_2$.The ratio of the relative fluorescence intensities (Fluorescence/$OD_{600}$) was then calculated for each replicate. The mean relative fluorescence intensities values were calculated from these replicates, error bars shown in the Figure S15

represent the corresponding standard deviation (s.d.). The ON/OFF ratio for each toehold mRNA sensor was calculated by dividing the relative fluorescence intensities obtained from a toehold mRNA sensor in the presence of trigger RNA (ON state) by the relative fluorescence intensities in the absence of trigger RNA (OFF state).

# Chapter III

## Riboswitch-inspired toehold riboregulators for gene regulation in *Escherichia coli*

**Key words:** Riboregulators • riboswitch architecture • toehold mediated strand displacement • gene regulation

3

# Abstract

Regulatory RNA molecules have also been widely investigated as components for synthetic gene circuits, complementing the use of protein-based transcription factors. Among the potential advantages of RNA-based gene regulators are their comparatively simple design, sequence-programmability, orthogonality, and their relatively low metabolic burden. In this chapter, we continued develop a set of riboswitch-inspired riboregulators in *Escherichia coli* that combine the concept of TMSD with the switching principles of naturally occurring transcriptional and translational riboswitches, aiming to minimize the undesired leakage. Specifically, for translational activation and repression, we sequestered anti-anti-RBS or anti-RBS sequences, respectively, inside the loop of a stable hairpin domain, which is equipped with a single-stranded toehold region at its 5' end and is followed by regulated sequences on its 3' side. A trigger RNA binding to the toehold region can invade the hairpin, inducing a structural rearrangement that results in translational activation or deactivation. We also demonstrate that TMSD can be applied in the context of transcriptional regulation by switching RNA secondary structure involved in Rho-dependent termination. Our designs expand the repertoire of available synthetic riboregulators by a set of RNA switches with no sequence limitation, which should prove useful for the development of robust genetic sensors and circuits.

# Introduction

Building upon our discussion in the previous chapter, it's noteworthy that riboswitches find their primary utility in bacteria. These riboswitches are categorized as cis-acting regulatory RNA elements that involve complex folded RNA domains in the 5 -untranslated regions (UTR) of mRNA molecules, which can change their conformation in response to the presence of small metabolites [104]. Most riboswitches contain an aptamer region as the recognition domain for the metabolite, which is followed by an expression platform that controls gene expression either by terminating transcription prematurely or by sequestering the RBS sequence to inhibit translation initiation. When small metabolites such as guanine[105,106], adenine[85], or vitamin B12[107] bind to the aptamer region, they trigger a rearrangement of the aptamer structure, which allosterically induces a conformational change in the expression platform, thus altering the expression of the downstream mRNA sequence[107] (Appendix Fig. S21 a).

Inspired by such small-molecule dependent RNA regulators, over the past decade a range of synthetic RNA-based regulatory systems have been developed. Such regulators have already been applied in metabolic pathway engineering[108], in the construction of synthetic gene circuits[109], for the development of biosensors[110,111], *in vivo* sensors[112,113] and regulators for inducible gene expression[114,115]. RNA based gene regulation provides several advantages, which make it particularly interesting for such synthetic applications. Due to the inherent sequence programmability of RNA secondary structural elements, RNA switches can be rationally designed and optimized for orthogonality[111,116], which is further supported by the availability of a wide range of computational tools for sequence design[117,118] prediction of thermodynamic properties and even molecular dynamics[119,120]. Compared to protein-based regulators, RNA regulators also represent a relatively low metabolic burden for the host organism[121,122], which is advantageous when scaling to larger systems.

We continued studying the toehold switch riboregulators mentioned earlier, which utilizes the mechanism of TMSD to trigger mRNA conformational changes and activate gene translation. Compared to a different approach towards RNA dependent riboregulators developed earlier[123], utilization of the toehold strategy allowed to increase the dynamic range of the switches by more than an order of magnitude and also enabled the development of large sets of mutually orthogonal regulators (Fig. S21 b). This, in turn, enabled the development of multi-input logic gates based on toehold switches[13]. Apart from translational activation, TMSD recently was also successfully applied to the development of translational repressors[13]. Recently, prediction of toehold-switch performance was also shown to be amenable to a deep learning methodology[35].

In an alternative approach, Chappell et al. developed RNA transcriptional activators (termed small transcription activating RNAs, or STARs), which exploit anti-sense RNAs to regulate the formation of an intrinsic transcriptional terminator located upstream of a reporter gene through a strand invasion process[124]. Both approaches - toehold switches and STARS - can be combined to create heterogeneous RNA regulators that exert control over gene expression both at the

transcriptional and translational level, and thus potentially increase the overall dynamic range of the regulation process[125].

In contrast to most naturally occurring riboswitches, toehold switches and STARS take RNA molecules as their input and are thus 'sequence-programmable'. This property has been used for the development of *in vitro* biosensors for the detection of viral nucleic acids[12,126], single nucleotide variations[127], and other analytes[128]. Toehold switches could even be shown to respond to the expression of the endogenous sRNA *RyhB* in *E. coli* [11]. Apart from the detection of nucleic acids, RNA-responsive elements naturally lend themselves as components for genetic circuits[33], as they can be 'wired up' rationally based on their nucleotide sequences, and accordingly a large variety of sequence-orthogonal elements can be designed[11,33].

Current designs for synthetic, RNA-triggered riboregulators still exhibit certain shortcomings such as leaky translation in the OFF state and sequence constraints for the trigger RNAs, respectively. These issues compromise the dynamic range and orthogonality of these components, which poses a challenge for the further development of genetic circuitry composed of larger numbers of RNA regulators, and also their potential use as sensors of endogenous RNA molecules. In the present work, we therefore sought to further expand the repertoire of rationally designed toehold regulators whose RNA inputs can be freely chosen. We applied the TMSD principle to a variety of mechanisms of translational and transcriptional regulation, which were inspired by the architecture of naturally occurring riboswitches, and thus combined functional elements of synthetic and natural gene regulators.

The overall design of our riboregulators is characterized by a toehold hairpin structure at the 5' UTR, which includes an anti-sense or anti-anti-sense sequence within the loop that is complementary to a functional sequence domain further downstream, followed by an expression platform. Binding of a trigger RNA at the toehold induces an allosteric rearrangement of the switch domain via TMSD, which either leads to premature termination of transcription or to sequestration of the RBS and hence repression of translation initiation (Fig. S21 c). The switching process is thus similar to that found in riboswitches, but it is induced by an RNA input rather than by a small molecule. We find that our design strategy indeed provides a viable approach to reduce the sequence constraints on the trigger RNA molecules, as the sequence for the toehold hairpin structure responsible for the initial allosteric rearrangement can be chosen freely.

Next to translational activators and repressors, we realized transcriptional activators based on interference with the formation of an intrinsic terminator. Of note, we were also able to demonstrate RNA regulators that utilize a Rho dependent mechanism, which also allowed the realization of a transcriptional repressor. Combination of translational and transcriptional control on a single transcript enabled the implementation of a genetic NOR gate with an overall enhanced ON/OFF ratio. At present, however, operation of our switches required the relatively high RNA levels provided by transcription with T7 RNA polymerase, limiting their application to synthetic gene circuitry

# Results

Riboswitch inspired control of translation initiation

Studies of the secondary structure of naturally occurring translational riboswitches have consistently shown[105,129,130] that in the absence of ligands, the RBS of the expression platform is either completely free or completely sequestered by an anti-RBS sequence within a duplex, thus allowing or precluding binding of the ribosome, respectively. In the absence of its ligand, the aptamer module of a riboswitch masks the corresponding anti-RBS or anti-anti RBS sequence, resulting either in an 'OFF' or an 'ON' riboswitch. In an OFF switch, binding of a ligand to the aptamer induces a refolding process that releases the anti-RBS, enabling interactions with the complementary RBS and thus switching the expression platform into the OFF state. In an ON switch, ligand binding releases the anti-anti RBS, which in turn sequesters the anti-RBS and thus makes the RBS available for ribosome binding. The utilization of anti-sense RBS sequences in natural riboswitches contributes to the relatively low leak expression of the controlled genes, which usually play an important role in cellular metabolism.

For our synthetic riboregulators, we replaced the riboswitch aptamer domain by a toehold hairpin whose switching via TMSD induces refolding of the expression platform. The toehold hairpin contains an unpaired cis-acting regulatory sequence (either an anti-RBS or anti-anti RBS), which avoids any sequence constraints for the RNA trigger input.

Our designs for translational activators (ON switches) and repressors (OFF switches) are shown in Figures 10 and 11, respectively (cf. Appendix Fig. S22 a & b). The translational activators consist of a 5' toehold hairpin with an anti-anti-RBS sequence in the loop, followed by different anti-RBS hairpin (AARH) loops. Our AARH design termed anti-RBS stem 1 is derived from the *E. coli thiM* riboswitch[131], which sequesters the RBS in the OFF state of the switch. We varied the 'natural' design to include different loop sizes and stem structures (anti-RBS stem 2–4). In each case, activating trigger RNA molecules can bind to the 14 nt long toehold and break up the first hairpin via TMSD. The anti-anti-RBS sequence exposed by this process binds to the anti-RBS sequence, releasing the RBS and thus facilitating translation initiation of the downstream *mCherry* reporter. A short unpaired region was added between toehold hairpin and anti-RBS hairpin to support the refolding of the anti-anti-RBS hairpin.

Thermodynamically, the TMSD-driven refolding process of our toehold riboregulators is driven by the free energy difference ($\Delta G_s$) between the conformations attained in the ON and OFF states. To ensure switch ability of our RNA constructs, we adjusted the folding free energies of the RNA structural motifs present in the absence and presence of a trigger RNA molecule. For example, translational activators in the OFF state (in the absence of a trigger RNA) include a toehold hairpin ($\Delta G_{TH}$ = −22.30 kcal/mol) and an anti-RBS hairpin ($\Delta G_{ARH}$ = −7.80 kcal/mol). The sum of the free energies of these motifs has to be lower than the alternatively folded anti-anti-RBS hairpin ($\Delta G_{AARH}$ = −15.00 kcal/mol) (Fig. S22 a). Binding of trigger RNA binding followed by

TMSD disrupts the toehold hairpin to form a stable double stranded region ($\Delta G_{ds}$ = −47.38 kcal/mol < $\Delta G_{TH}$ + Δ $G_{ARH}$), which prompts the formation of the anti-anti-RBS hairpin.

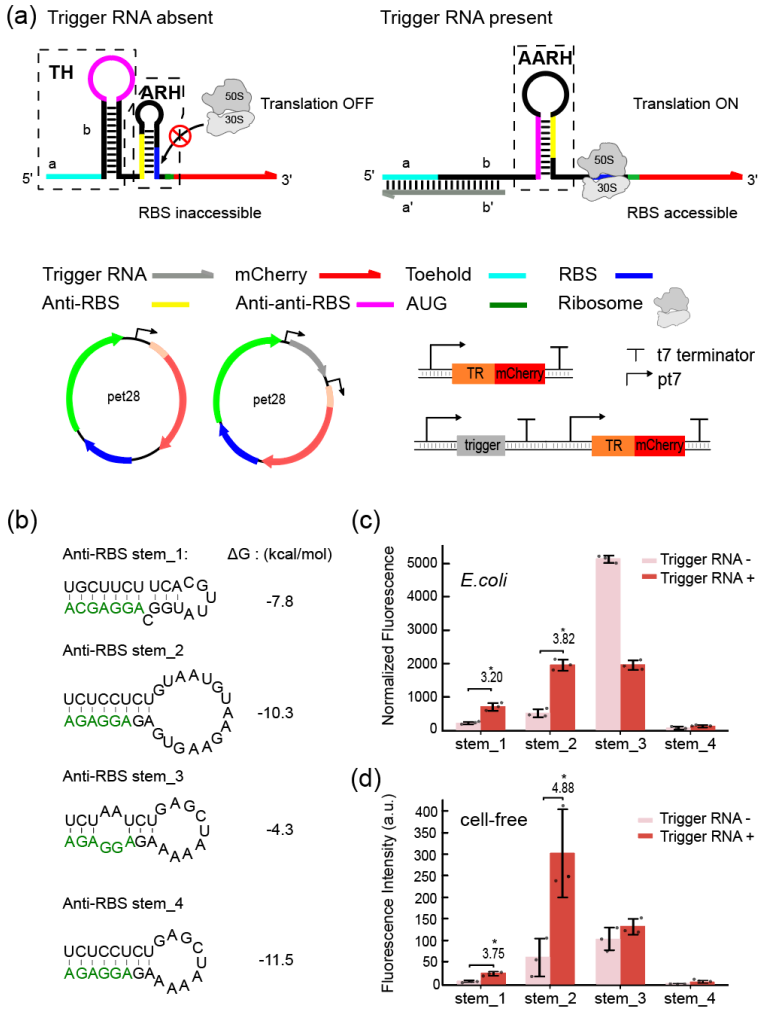


Figure 10. Design and characterization of riboswitch-inspired toehold-riboregulator controlling translational activation. a), Scheme of a toehold riboregulator that activates translation initiation in response to trigger RNA input. In the absence of a trigger RNA (grey), the toehold hairpin (TH) confines an anti-anti-RBS sequence (purple) within its loop region. The RBS (blue) is sequestered within an anti-RBS hairpin (ARH) by an anti-RBS sequence (yellow), which prevents binding of the ribosome. Trigger RNA can initiate a TMSD process at the toehold (light blue), which releases the anti-anti-RBS sequence. The RBS sequestration hairpin is unfolded by the released anti-anti-RBS sequence and forms an anti-anti-RBS hairpin (AARH), which in turn exposes the RBS to the ribosome and allows translation of the mCherry readout (red) to proceed. b), Predicted secondary structure and total free energy of each anti-RBS hairpin - the RBS sequence is highlighted in green. c), Relative fluorescence intensities from in vivo measurements in the ON and OFF state for each anti-RBS hairpin, respectively. d), Relative fluorescence intensities obtained in cell-free experiments with riboregulators in the ON and OFF state, respectively. For both the relative fluorescence/OD and fluorescence intensity data, Welch's t-tests were performed for each construct; *P < 0.05, indicating conditions where the fluorescence/OD and fluorescence intensity for the trigger RNA + condition is statistically significantly different from that of the trigger RNA-condition. Error bars represent the standard deviation (s.d.) for biologically independent samples.

For consistency, we designed the toehold hairpins of all our riboregulators with the same loop size (15 nt) to ensure similar behavior during the initial TMSD process. By contrast, the stem lengths and loop sizes of the anti-RBS, anti-terminator and anti-rut sequence domains depend on

the downstream target RNA sequences and were adjusted individually to meet the design rules and ensure an overall G < 0.

In order to assess the *in vivo* performance of the translational activator, its components were cloned into a recombinant plasmid and studied in experiments with *E. coli* BL21 DE3 in M9 medium. Both trigger RNA and toehold riboregulators were put under the control of T7 RNA polymerase. To activate the riboregulators, IPTG (1mM) was added to an *E. coli* cell culture, which resulted in expression of T7 RNA polymerase (RNAP) by the *E. coli* BL21 DE3 bacteria. In fact, we found that transcription by T7 RNAP was crucial for the performance of the switches. Experiments with a weaker *E. coli* promoter led to worse performance or even loss of function (cf. Appendix Fig. S36 and Discussion).

Upon activation, the activator derived from the *thiM* riboswitch displayed a relatively low leak and an ON/OFF ratio of ≈ 3.2, while our best-performing design (anti-RBS stem 2) had a higher leak, but an ON/OFF ratio of ≈ 3.8 (Fig. 10 c) (See Methods). Interestingly, anti-RBS stem 3 containing two mismatches in the stem shows a very strong leak and reduced activity in the presence of trigger rather than activation, while the most stable AARH anti-RBS stem 4 had a very low leak, but could not be activated. In addition, we also tested our translational activator (anti-RBS stem 1) in LB medium and the performance (ON/OFF ratio ≈ 3.5, Appendix Fig. S36) is similar as in minimal media.

We also tested our translational riboregulators in an *E. coli*-derived cell-free expression system (PURExpress®) to compare their performance with the *in vivo* results. To this end, linear DNA templates containing toehold riboregulators and triggers were added to the cell-free system, and fluorescence intensities were measured with a microplate reader (See Appendix - *In vitro* protein expression in cell-free system). The *in vitro* results were in line with the *in vivo* experiments, but tended to show a slightly better performance in terms of ON/OFF ratios. The best performing AARH (anti RBS stem 2) exhibited an ON/OFF ratio of ≈ 4.9 (Fig. 10 d).

We wish to note that the measured ON/OFF ratios are highly dependent on the experimental details. When assessed using the same experimental workflow, the performance of our translational activators is comparable to that of other previously developed riboregulators[11] (cf. Discussion). In the case of our translational repressors, we utilized an anti-RBS to bind the RBS after refolding of the switch with the aim to prevent undesired ribosome invasion in the OFF state and thus improve translational repression. In the 'ON state' of the repressor (Fig. 11 a), the anti-RBS sequence is initially located in the loop of the toehold hairpin, leaving the RBS freely accessible for ribosome binding, and thus allowing translation of the *mCherry* reporter. Upon binding of the trigger RNA and strand invasion into the toehold hairpin, the anti-RBS sequence is released, followed by sequestration of the RBS (Fig. 11 b) and thus repression of translation initiation of the *mCherry* reporter. To ensure proper switching, the free energy of the toehold hairpin ($\Delta G_{TH}$ = −23.10 kcal/mol) was designed to be lower than the alternatively folded anti-RBS hairpin ($\Delta G_{AR}$ = −15.90 kcal/mol). In the translational ON state, i.e. in the absence of trigger, the anti-RBS sequence is thus safely sequestered within the loop of the toehold hairpin (Fig. S22 b).

Corresponding in vivo experiments for the translational repressor resulted in ON/OFF ratios of ≈ 2.7 (Fig. 11 c, while *in vitro* experiments with the repressor resulted in a slightly better ON/OFF ratio of ≈ 4 (Fig. 11 d).

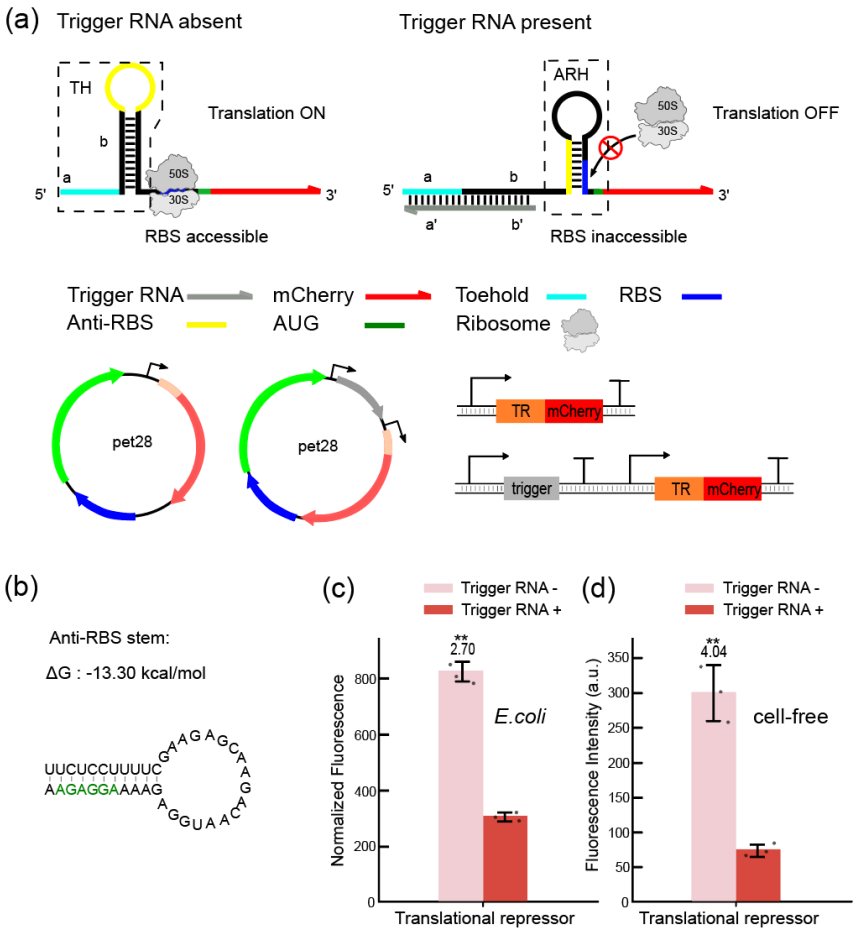


Figure 11. Design and characterization of a riboswitch-inspired toehold-riboregulator controlling translational repression. a), In the absence of a trigger RNA (grey), the toehold hairpin (TH) constrains an anti-RBS sequence (yellow) within its loop region, the RBS is freely accessible and translational initiation is enabled. In the presence of trigger, toehold (light blue)-mediated invasion of the hairpin stem releases the anti-RBS, which leads to formation of anti-RBS hairpin (ARH) and sequestration of the RBS and thus translational repression. b), Predicted secondary structure and total free energy of the anti-RBS hairpin. The RBS sequence is highlighted in green. c), Relative fluorescence intensities in the ON and OFF state of the translational toehold repressor measured *in vivo*. d), *in vitro* relative fluorescence intensities of the translational toehold repressor in its two states measured in a cell-free expression system. For both relative fluorescence/OD and fluorescence intensity data, Welch's t-tests were performed on each construct; *P < 0.05 and **P < 0.01 indicate conditions where the fluorescence/OD and fluorescence intensity for the Trigger RNA + condition is statistically significantly different from that of the trigger RNA- condition. Error bars in c, d represent the s.d. from at least three biologically independent samples

## Intrinsic terminator dependent transcriptional termination

A variety of naturally occurring riboswitches are based on the control of transcriptional termination. We therefore sought to apply our approach also to the development of toehold-mediated transcriptional terminators or anti-terminators (Fig. 12, Fig. S23 a & b). The structures of the transcriptional activators (Fig. 12 a & e) each comprise a toehold hairpin and an intrinsic terminator whose sequence is derived from the late terminator *t22* from phage P22[129]. The t22

terminator releases an RNA transcript at a critical guanine nucleotide located eight nucleotides away from the terminator stem. Mutating the critical guanine was previously found to result in increased transcriptional readthrough.

In our first design, termination can be interrupted by an anti-terminator sequence, which initially is sequestered within the loop of the toehold hairpin (Fig. 12 a). When an activating trigger RNA invades the toehold hairpin via TMSD, the anti-t22 sequence (the sequence on the anti-*t22* stem 1) is exposed in the nascent RNA and can thus hybridize to a 5 subsequence of the *t22* terminator (Fig. 12 b) and form an anti-*t22* hairpin (ATH). In our second design, anti-*t22* stem 2 also includes the complementary sequence of the *t22* sequence near its 3' end, which contains the critical guanine nucleotide of the *t22* terminator (Fig. 12 f). Binding of trigger RNA results in a refolding of the RNA structure, after which the critical guanine is also sequestered within the stem of the anti-*t22* hairpin (Fig. 12 e). In each case, transcriptional elongation is expected to proceed after refolding.

As shown in Figure 12 c and g, *in vivo* experiments with *E*. coli BL21 DE3 carrying plasmids with the components of the two transcriptional activators exhibited ON/OFF ratios of ≈ 2.3 and ≈ 3.1, respectively, indicating that the 'critical G' design indeed performs better than the simpler anti-*t22* stem 1 design. As an alternative means of estimating the switching efficiency, quantitative PCR (qPCR) was used to characterize the steady-state level of all toehold-*mCherry* mRNAs in the presence and absence of trigger RNAs in *E. coli* (Fig. 12 d & e, Methods, Experimental procedures, Quantitative PCR). In terms of fully transcribed mRNAs, our transcriptional activators showed ON/OFF ratios of 2.0 and 2.2, confirming that control is exerted at the transcription level.

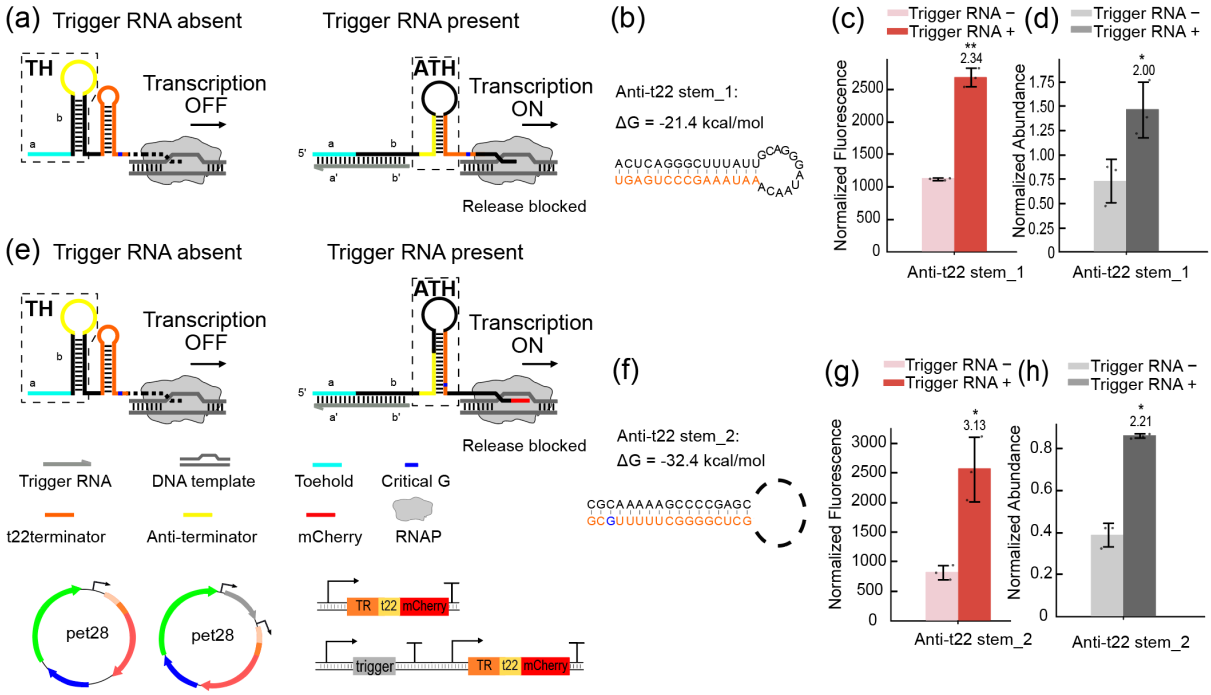


Figure 12. Design and characterization of riboswitch-inspired toehold riboregulator mechanisms controlling transcriptional activation via an intrinsic terminator. a), Anti-*t22* stem 1 design: In the absence of a trigger RNA (grey), a transcriptional anti-terminator (yellow) is constrained within the loop of the toehold hairpin (TH), and thus

transcription is terminated by the following intrinsic *t22* terminator (orange). In the presence of trigger RNA, trigger binding and TMSD opens the TH, and the released anti-terminator sequence sequesters the terminator in an alternative structure (anti-terminator hairpin, ATH), which allows downstream gene transcription to proceed. b), Predicted secondary structure and total free energy of the anti-t22 stem, a subsequence of the *t22* terminator is highlighted in orange. c), Relative mCherry fluorescence intensities measured with the transcriptional activator in the OFF and ON state, respectively. d), Normalized abundance of mRNA transcripts characterized by qPCR in the transcriptional OFF and ON state of the anti-*t22* stem 1 activator. e), Anti-*t22* stem 2 design: In this design TMSD induced refolding of the RNA structure leads to a sequestration of the critical guanine nucleotide (indicated in blue) within the anti-anti-*t22* terminator hairpin (ATH) stem, which allows downstream gene transcription to proceed. f), Predicted secondary structure and total free energy of anti-*t22* stem 2, with the sequestered *t22* terminator subsequence highlighted in orange. g), Relative mCherry fluorescence intensities and h), abundance of mRNA transcripts measured by qPCR for the two states of the transcriptional activator. For both fluorescence/OD values and qPCR quantification, Welch's t-tests were performed on each construct; *P < 0.05 and **P < 0.01, indicate that the trigger RNA + condition is statistically significantly different from that of the trigger RNA- case.

Rho-dependent transcriptional termination.

Next to transcriptional riboregulators that target intrinsic terminators, we also attempted to control Rho-dependent termination processes. In *E. coli*, about half of the factor-dependent termination processes are Rho-dependent, and they are typically associated with genes involved in metabolism and metabolic control[132-134]. Termination factor Rho[135,136] is a homo-hexametric RNA chaperone that binds to nascent RNA by recognizing the Rho utilization sequence (*rut* site) and aborts transcription by pulling the RNA away from the RNAP and DNA template.

A well-studied example of a Rho-dependent termination process is found in the *E. coli* tryptophanase (*tna*) operon which encodes tryptophanase and permease for tryptophan metabolism[136]. At high cellular levels of tryptophan, the ribosome stalls during translation of the tnaC peptide at the *tnaC* stop codon that is adjacent to the rut site. The stalled ribosome blocks access of the Rho factor and therefore allows transcriptional elongation to proceed (Appendix Fig. S24). At low tryptophan levels, the ribosome is not stalled, resulting in Rho-dependent termination after completion of tnaC synthesis.

Based on this mechanism, we first attempted to use a toehold riboregulator to activate access of Rho to the rut site directly. In the OFF state of this riboregulator, in the absence of trigger RNA an anti-rut sequence is sequestered within the toehold hairpin, making the downstream *rut* sequence accessible for the Rho factor, which leads to transcription termination (Fig. 13 a, Appendix Fig. S25). Addition of an activating trigger RNA opens the toehold hairpin and releases the anti-rut sequence, which consequently binds rut within the anti-rut hairpin (ARH) and thus prevents Rho-dependent termination.

Rather than utilizing a regulatory sequence for Rho binding, we also designed a toehold transcriptional repressor, in which we controlled the access of the ribosome to the *tnaC* sequence in a similar manner as in the translational riboregulators described above (cf. Fig. 11 a, and Appendix Fig. S25). In the ON state, i.e. in the absence of trigger RNA, an anti-RBS sequence is confined within a toehold hairpin, and therefore the RBS is accessible for ribosome binding and translation of the tnaC peptide encoded on the downstream sequence. In consequence, ribosome

stalling during tnaC translation blocks the rut site for the Rho factor and thus transcription proceeds (Fig. 13 e). Binding of an activating trigger RNA followed by TMSD opens the toehold hairpin, enabling sequestration of the RBS sequence by the anti-RBS sequence. As a result, Rho factor binds to the – now free – *rut* site and terminates transcription elongation.

In order to assess the performance of the rut-dependent riboregulators, recombinant plasmids coding for the regulators in the presence and absence of trigger RNA were transformed and expressed in *E. coli* BL21 DE3 with additional tryptophan in the medium. Comparison of the resulting mCherry expression end levels resulted in ON/OFF ratios of up to 3.50 for the transcriptional activator (Fig. 13 c), and up to 2.33 for the transcriptional repressor (Fig. 13 g). We also used qPCR to quantify the steady-state level of all toehold-mCherry mRNAs in the presence and absence of trigger RNAs in *E. coli* (Figure 13 d & h), which were found to be in close agreement with the fluorescence data.

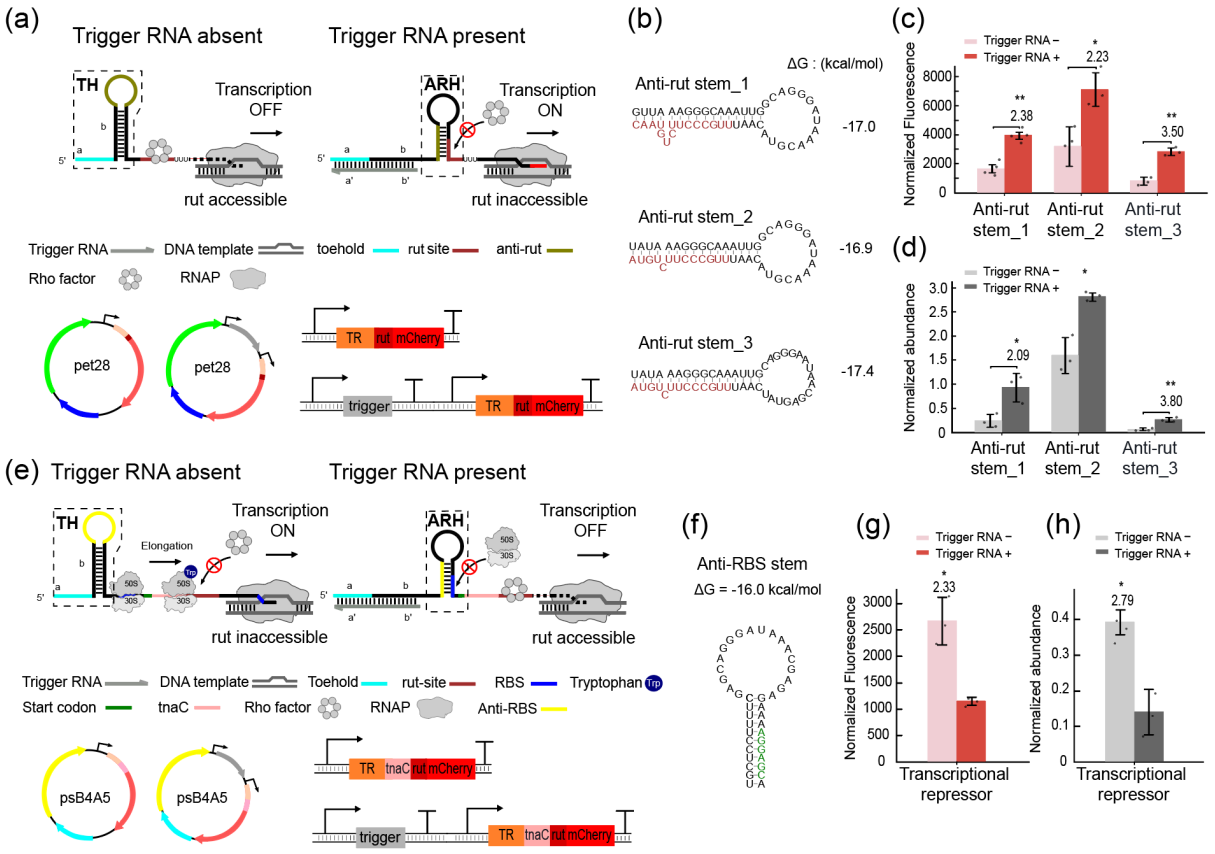


Figure 13. Design and characterization of transcriptional toehold riboregulators based on Rho-dependent termination. a), Principle of a transcriptional activator: In the absence of trigger RNA (grey), the toehold hairpin (TH) confines anti-rut sequence (olive) within its loop, while the rut site (brown) is exposed to Rho factor, which terminates transcription. Upon invasion of the toehold hairpin (TH) by trigger RNA, the anti-rut sequence is released, and anti- transcriptional elongation is switched ON. The design of the corresponding plasmids and gene circuits are also shown. b), Predicted secondary structure and free energies of several variants of the anti-rut stem, where the rut sequence is highlighted in brown. c), Relative fluorescence intensities of the transcriptional activators measured *in vivo* in the OFF and ON state. d), Corresponding normalized abundance of mRNA transcripts measured by qPCR. e), Scheme of transcriptional repression by a toehold riboregulator, which is based on the *tna* operon. In the absence of trigger RNA (grey), ribosomes can bind to translate the tnaC peptide, followed by stalling at the rut site (brown). This prevents Rho from binding and thus allows transcription to proceed. In the presence of trigger, translation of tnaC is disabled and Rho factor can bind to the exposed rut site and thus terminate transcription. f), Predicted secondary structure and free energy of the anti-RBS hairpin. The

RBS sequence is highlighted in green. g), Relative fluorescence intensities and (H) mRNA abundance in the ON and OFF state of the transcriptional repressor and in the presence of 5 mM tryptophan. Based on Welch's t-tests, *P < 0.05 and **P < 0.01, indicate conditions where the fluorescence/OD and qPCR quantification for the trigger RNA + condition is statistically significantly different from that of the trigger RNA- condition.

## A logic NOR gate based on combined transcriptional and translational toehold repression

We finally investigated whether it is possible to combine transcriptional and translational control within a single transcript. To this end, we fused the *tna* operon-based transcriptional repressor with a translational repressor, which is expected to generate a logic NOR gate that only results in active gene expression in the absence of the two corresponding trigger RNAs (Fig. 14 a, Appendix, Fig. S27). Genetic constructs containing the NOR gate were co-transformed into *E. coli* BL21 DE3 with plasmids coding for the different combinations of trigger RNAs, followed by bacterial culture and quantification of the fluorescence output with a microplate reader (Fig. S27) or via flow cytometry (Fig. 14 b & c). The fluorescence data demonstrate NOR gate performance as desired. Notably, the ON/OFF ratio for the 'transcriptional part' of the switch is on the order of 7 according to flow cytometry (Fig. 14 b), whereas the ratio for the translational part or for the combined inputs is > 15 (see Discussion of this chapter).

In principle, tryptophan could be interpreted as a third input, which controls the transcriptional module of the gate. In the absence of tryptophan, transcription in the absence of trigger A would be reduced by a factor of ≈ 2 compared to the operation with 5 mM tryptophan. Correspondingly, Trp concentration could be used to modulate the maximum output level of the NOR gate.

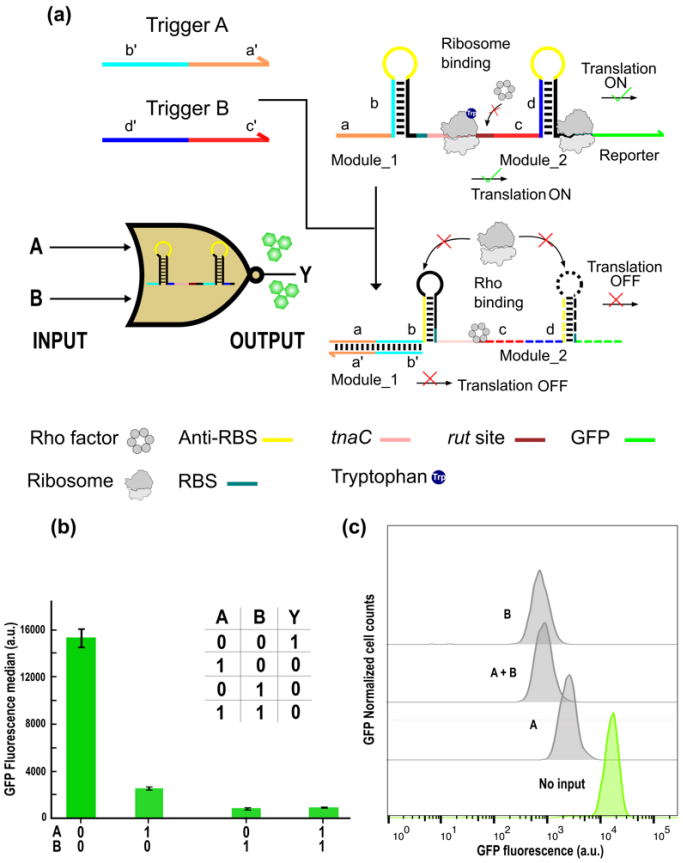

Figure 14. A NOR logic gate based on riboswitch-inspired transcriptional and translational toehold switches. a), The switch is composed of two modules which each control translation from an RBS. The function of the first module is to switch translation of the tnaC peptide, which results in blockage of the rut site by the ribosome (cf. Fig. 13). This in turn prevents termination by Rho factor and thus allows transcription to proceed. The second module is a translational toehold repressor for the GFP reporter gene (cf. Fig. 11). Binding of the two trigger RNAs lead to transcriptional termination and translational repression, respectively. b), GFP fluorescence output of the NOR gate for different input combinations, measured by flow cytometry in the presence of 5 mM tryptophan (required for the tna operon). Given are the mean values from three biologically independent samples, error bars represent their standard deviation. c), GFP fluorescence histograms for the NOR gate toehold riboregulator in the absence and presence of inputs obtained from a single flow cytometry run.

# Discussion

We designed and characterized a series of riboswitchinspired riboregulators which are switched by RNA inputs via a toehold-mediated strand displacement (TMSD) process. These riboregulators combine structural features of natural riboswitches, and the switching principle of recently developed synthetic riboregulators. Instead of using a natural, small molecule-responsive RNA aptamer to control the interaction of a cis-regulatory sequence with an expression platform, we placed a regulatory RNA sequence (e.g., an anti-RBS sequence) into the loop region of a toehold hairpin which could be switched by TMSD, inducing refolding of the riboregulator and thus controlling translational initiation or transcriptional termination. Various factors contribute to the observed performance of the translational and transcriptional switches, which relate to the kinetics and thermodynamics of the switches as well as the details of their experimental assessment.

Quantification of the performance of the riboregulators

In the present work, the performance of the riboregulators was quantified by comparing the endpoint of the fluorescence of bacterial cultures normalized by the density (or $OD_{600}$) in the respective ‘ON’ and ‘OFF’ states of the switches. This enabled us to assess and rank different riboregulator designs with respect to each other. We wish to note, however, that the actual value of the ON/OFF ratio was found highly dependent on the details of the experimental approach. For instance, when investigating previously described toehold riboregulators[11,33] with reported fold changes of more than 600 using our experimental workflow, we obtained ON/OFF ratios of only ≈ 7 (Appendix Fig. S29). We obtained different values for this measure when using RFP or GFP as a fluorescent reporter, or when quantification was performed with Fluorescence/OD, RTqPCR or via flow cytometry. In the following, we assume that the overall trends observed when varying the design of the riboregulators are robust with respect to the choice of quantification method.

Influence of promoter choice

We found that our riboregulators performed best when both trigger and riboregulator were transcribed by T7 RNAP. For comparison, we tested two of our riboregulators using a weaker promoter (the *E. coli* promoter BBa_J23119) for transcription of either trigger or riboregulator, or both (Appendix Fig. S36). When riboregulator anti-RBS stem 1 was transcribed from the strong T7 promoter, but trigger RNA was generated from the weaker *E. coli* promoter, we found a strongly reduced ON/OFF ratio for the switch (Fig. S36 a). Transcription of translational activator anti-RBS stem 2 from the *E. coli* promoter led to a strong leak expression even in the absence of trigger, which is in line with results from a previous study on toehold switches. In fact, transcription of both riboregulator and trigger led to no switching at all (Fig. S36 b). These findings suggest two distinct effects of the choice of promoter on the performance of the system: First, fast transcription is required to enable fast co-transcriptional folding of the riboregulator and thus prevent binding of ribosomes to the nascent switch[137], which would result in leaky translation. Second, transcription of trigger by T7 RNA polymerase generates higher intracellular concentrations, which are required for switching the riboregulators via toehold-mediated strand displacement (cf. also the discussion of intracellular RNA concentrations and sensing below). With an eightfold higher maximum transcription speed for T7 RNAP than for *E. coli* polymerase[137], and transcription from a high copy plasmid number (pET28b has copy number ≈ 40), trigger and riboregulator concentrations are expected well in the M range. As overexpression of genes in *E. coli* using T7 RNAP is known to inhibit cellular growth[138,139], we also tested the effect of the generation of trigger RNA by T7 RNAP in *E. coli* BL21 DE3 on bacterial growth. Compared to a negative control, in which only T7 RNAP is expressed, but no T7 promoter is present, we found an increase in doubling time from ≈ 30 to ≈ 60 mins (Appendix Fig. S34).

Translational activators

Translational activators are supposed to be translationally inactive in the OFF state. The natural strategy to suppress translation found in riboswitches is the use of an anti-RBS/RBS hairpin stem, in which the RBS is sequestered. Our experiments indicate that the hairpin stem has to be stable enough to prevent invasion by the ribosome, but not too stable to prevent switching altogether (Fig. 10 b & c). Further, sufficiently fast transcription of the riboregulators appears to be required to prevent binding of ribosomes to the nascent RNA (cf. discussion on the use of T7 RNAP above).

We initially designed and characterized several toehold translational activators whose anti-RBS/RBS sequences were rationally chosen (Anti-RBS stem 2,3,4) and formed hairpin stems of different stability (Fig. 10 b). Indeed, the most stable hairpin (anti-RBS stem 4) resulted in the lowest leak, but also did not lead to a strong increase in gene expression in the presence of trigger. The slightly less stable anti-RBS 2 structure performed best in terms of ON/OFF ratio, but

with the trade-off of a higher leak expression. We also replaced the in silico designed anti-RBS/RBS hairpin by the corresponding structure taken from a natural *thiM* riboswitch from *E. coli* (Anti-RBS stem 1). This naturally occurring anti-RBS hairpin comprises an RBS (aggagc) of lower efficiency compared to that (aggaga) of the other anti-RBS/RBS hairpin stems (RBS calculator[140], Table. S8), and contains a bulge formed by an unpaired cytosine next to the RBS. Even though the resulting structure had a lower predicted thermodynamic stability than Anti-RBS 2, it had a lower leak *in vivo*, and the second-best ON/OFF ratio obtained for our design variations. The comparatively high leak expression from anti-RBS stem 2 presumably is caused by the large size of the hairpin loop, which potentially allows undesired ribosome binding and translation initiation. Our results thus suggest that ribosome invasion (generating leak) can be prevented by appropriate secondary structure, and potentially is more severe *in vivo* than in the cell-free context. The latter may be the consequence of the ≈ 10-fold lower ribosome concentration in a cell-free system. In agreement with previous findings[10,46], we also found that strategically placed bulges in the toehold stem can improve TMSD efficiency by introducing a forward bias into the strand invasion process (Fig. S31).

For our riboregulators, we adopted the sequence of the toehold and its adjacent hairpin loop from a previous study on toehold riboregulators, which had been tested for input orthogonality with 12 other trigger RNAs with random sequences. To confirm the specificity of the design, we tested the performance of translational activator anti-RBS stem 1 with three additional trigger RNAs with scrambled sequences. None of them was found to activate expression appreciably (Fig. S35).

Translational repressors

For translational repressors, kinetic considerations become more important than for the activators. In the presence of trigger RNA, translational repressors need to be switched into a translationally inactive state quick enough to avoid leaky translation from an accessible RBS. Transcription of the riboregulator at a speed of ≈ 10–100 nt/s defines a time window of ≈ 1 s for trigger binding to the toehold and disruption of the toehold hairpin (or prevention of its formation). For our repressors, we therefore designed a relatively strong anti-RBS hairpin that is expected to sequester the RBS following toehold-mediated switching of the toehold hairpin (Fig. 11 b). While our design is shown to work, in principle, we observed relatively strong leakage in the OFF state (in the presence of trigger RNA) (Fig. 11 c), indicating that TMSD and formation of the anti-RBS hairpin does not occur fast enough to prevent binding of ribosomes to the RBS. Compared to the *in vivo* case, cell-free experiments show a slightly reduced leak in the OFF state of the repressor, which again might be a consequence of the lower ribosome concentration in the cell-free reaction. Due to the reduced leak, the cell-free translational repressor also exhibits a better ON/OFF ratio than *in vivo*.

Besides *E. coli*, we extended our testing of the toehold translational activator to *Vibrio natriegens* (Vmax™ X2, Telesis Bio), which exhibits more favorable switching behaviors (ON/OFF ratio: 9.58 ± 3.16) in Vmax X2 compared to BL21 DE3 as shown in Figure S40. This observed divergence in

behavior likely stems from the discrete cellular microenvironments of the two organisms. Given that *V. natriegens* is a marine bacterium, it likely acclimates to higher intracellular salt concentrations relative to *E. coli*[141,142]. This variance in reaction conditions could potentially contribute to the observed variations in RNA folding and TMSD energy profiles. Our results highlight the compatibility our toehold riboregulators across different organisms.

Transcriptional riboregulators

The function of transcriptional riboswitches, which are based on termination or anti-termination, critically depends on the kinetics of transcription, ligand binding and refolding of the expression platform[143]. Natural riboswitches respond to small metabolites, which have to be present at relatively high concentrations (in the M to mM range) for switching, and whose binding to the riboswitch's aptamer domain is associated with a relatively low change in free energy. By contrast, our transcriptional riboregulators are challenged with responding respond to trigger RNAs, which bind with a much higher binding free energy (ΔG), but which are present at typically lower concentrations (in the nM to M range, see below). A sufficiently high concentration and a correspondingly high on-rate of the triggers is crucial for the functioning of the switches, however.
In a previously developed type of transcriptional regulators termed STARs[124], small trans-acting RNA molecules were utilized that contained an anti-terminator sequence to regulate the formation of the intrinsic terminator and thus control downstream gene transcription. While this strategy results in efficient transcriptional activation, the trigger RNA necessarily includes a part of the complementary sequence of the intrinsic terminator and thus cannot be chosen without constraints. By contrast, our design strategy (Fig. 12) leaves the anti-terminator sequence unpaired and confined within the loop of the toehold hairpin, which avoids any sequence limitations for the trigger RNA. An interesting outcome of our experiments is the sensitivity of the performance of the riboregulators to sequence details in the switching domains. For instance, we confirmed that the termination process from the *t22* terminator can be suppressed efficiently, when its 'critical' guanine nucleotide is included in the anti-terminator/terminator stem, which is in line with the conclusions of previous work[144,145].
We also demonstrated transcriptional toehold activators and repressors that interfere with Rho-dependent termination (Fig. 13), for which we modified the Rho utilization (rut) sequence of the naturally occurring transcriptional switch of the *tna* operon. Experiments with the wildtype rut site alone showed substantial transcriptional readthrough and a correspondingly leaky expression of the reporter sequence (Fig. S28). Several factors may contribute to the observed leak. As the Rho factor has multiple cellular functions other than transcriptional termination, its recruitment to the rut site depends on its availability under the given cellular context. Further, secondary structure close to the rut site might reduce its accessibility for Rho. When we removed some of the original sequence context upstream of the *tna* operon's rut site, leaky expression was reduced (Fig. S28). We also found that the termination efficiency of Rho factor was enhanced by insertion of a transcriptional pausing site (U7) right after the rut site (Fig. S28 & S29). Based on

these insights, we were able to realize TMSD-based anti-rut regulators that controlled rut accessibility and thus Rho dependent termination with a relatively low leak transcription and ON/OFF ratios of up to 3.5 (Fig. 13 a-d).

Previous work on Rho factor binding[132] found that most of the riboswitches and sRNAs that modulate Rho-dependent termination are based on ribosome stalling, which is similar to the mechanism found in the *tna* operon (Fig. S24). Our transcriptional repressors (Fig. 13 e-h) were thus designed to control ribosome binding (and thus stalling) in the same way as the translational activators discussed above. Several processes have to play together co-transcriptionally to make the switch work: in the transcriptional ON state (in the absence of trigger), ribosomes need to bind to the RBS quick enough to be able to block the rut site before Rho can bind. In the presence of trigger RNA, however, refolding of the toehold switch needs to take place fast enough to prevent undesired binding of ribosomes and thus facilitate binding of Rho to rut – our experiments suggest that the kinetic competition between these processes results in an appreciable leak, but still displays a decent ON/OFF ratio of ≈ 2.3.

Kinetic considerations and potential for sensing of endogenous RNA molecules

An exciting potential application for RNA-triggered riboregulators is the detection of endogenous RNA species such as mRNAs or small regulatory RNAs (sRNAs). However, at this point our designs do not appear to be sensitive enough for such sensing tasks. The range of concentrations expected for endogenous RNAs in *E. coli* lies in the range from 1 nM to ≈ 10 M. The lower bound corresponds to a single copy of the molecule in the bacterial cell, while the most highly expressed RNAs are the ribosomal RNAs, which are present at concentrations of ≈ 20 M[146]. mRNA concentrations in *E. coli* have been estimated from single molecule FISH experiments, and were found to lie in the range from ≈ 0.1 nM to 100 nM[147], and the total mRNA concentration is ≈ 1.4 M[146]. Similar values (several 10 nM) can be estimated for sRNAs[148]. Whether such small concentrations can be sensed, critically depends on several factors: efficient hybridization of the target with the sensor riboregulator, low leak, and sufficient ON/OFF ratio.

In the case of translational ON-switches, we found that a strong anti-RBS hairpin reduces leaky translation, but also diminishes the switchability of the structure. All other designs – the translational repressors and the transcriptional regulators – depend on a kinetic competition, e.g. either between trigger and ribosome binding, or between trigger binding and transcriptional termination (other factors such as the speed of transcription and folding of the secondary structure also play a role[143]).

In the first case, the target molecules would have to be present at similar concentrations or higher than the concentration of free ribosomes, i.e. those not bound to other mRNAs and engaged in translation, which has been estimated to be on the order of 500 nM[149]. In the second case, hybridization of the trigger and the nascent mRNA has to take place in the time window between transcription start and formation of the terminator hairpin, which we estimate to be on the order of 1 s[143]. With an estimated RNA association rate in *E. coli* of ≈ $3 \times 10^5$ $M^{-1}s^{-1}$ [148], efficient switching

would require trigger concentrations above 3 M. In addition, hybridization rates are strongly dependent on secondary structure and the presence of RNA binding molecules, which may further reduce the efficiency of the target binding[150].

All these considerations indicate that our current designs only allow detection of RNA species, which are present at relatively high copy numbers, corresponding to concentrations in the M range. This is consistent with our finding that the riboswitch-inspired riboregulators performed well only when transcribing the triggers and switches from a T7 promoter (see above).

We therefore also sought to improve the efficiency of a transcriptional activator by utilizing the bacterial RNA chaperone Hfq. Binding of Hfq to RNA containing a specific Hfq-binding motif protects the RNA from degradation (increasing the cellular concentration) and promotes RNA hybridization reactions *in vivo*[151]. However, modification of trigger RNA with an Hfq-binding hairpin did not result in an appreciable improvement in performance (Fig. S38). As Hfq is thought to promote rather weak RNA interactions, it is likely to not have an effect in the case of our 'optimized', secondary structure-free triggers and toehold hairpins.

Realization of cellular logic computation

In principle, RNA-based regulatory mechanisms are ideal for the implementation of cellular computing circuits, as the sequence-specificity of RNA interactions allows a rational design and 'wiring' of the different components of the circuits. In this respect, our riboswitch-inspired riboregulators have the benefit that the RNA trigger sequences (the 'input') can be independently chosen from the more or less fixed sequences required for gene regulation (i.e. RBS/anti-RBS or terminator/anti-terminator sequences).

Input sequences can be chosen independently and orthogonally to trigger their respective toehold hairpins – e.g., the sequences for triggers A and B in the NOR gate demonstrated in Fig. 14 had been previously used in the context of toehold switches - they were chosen to be orthogonal and have no biological meaning. As also exemplified by the NOR gate, combining transcriptional and translational regulation within a single switch leads to a comparatively compact design, and also yields an improved ON/OFF ratio compared to a single switch.

However, we also found that combining several switches in the 5 untranslated region of a single transcript appears to be less modular than naively expected. We attempted to realize a range of other logic gates with this strategy, but in most cases the switches were non-functional or showed poor performance. For instance, an IMPLY gate can be constructed by fusing a translational activator and a translational repressor (Fig. S39). For this gate, one would expect a low output only in the presence of the trigger for the translational repressor. Due to strong leakage, the IMPLY gate did not show a clear 'Boolean' behavior.

Putting several switches in a row appears to be challenging, and our results indicate that the different components do not act independently and thus cannot be combined simply in a modular fashion. Potentially, optimization of such logic gates (or more complex functions) could be

achieved with a screening approach combined with machine-learning methodology, as recently demonstrated for toehold switches[35].

# Materials and methods

RNA structure design

We designed the riboswitch-inspired toehold riboregulators by combining a modified toehold switch structure with rationally designed regulatory sequences. Initially, we adopted sequences for the toehold and stem region (30 nt in total), which had previously shown excellent *in* vivo performance (with ON/OFF ratios of 665 and 557, respectively, and proven input orthogonality). We then modified the original toehold hairpin structure by removing the uppermost base-pairs in the stem (which were not opened through TMSD in the original design) to increase the refolding efficiency. We then put a sequence domain targeting one of the regulatory downstream sequence elements (i.e. RBS, *t22* terminator, or *rut*) into the loop region and adjusted this sequence to prevent stacking within the loop region. The loop sequences were typically chosen to be partially complementary to the targeted sequence and to additional 2–3 nt on their 5 sides. These sequence domains were included to act as internal toeholds to increase refolding efficiency. We calculated the free energy of each structural domain of the riboregulator to guarantee the thermodynamic favorability of a secondary rearrangement induced by TMSD, including the free energy of toehold hairpin ($\Delta G_{\text{toehold hairpin}}$), anti-hairpin ($\Delta G_{\text{anti-hairpin}}$), and anti-anti-hairpin ($\Delta G_{\text{anti-anti-hairpin}}$). Sequences for the hairpin were designed to follow the order: $\Delta G_{\text{toehold hairpin}} < \Delta G_{\text{anti-anti-hairpin}} < \Delta G_{\text{anti-hairpin}}$ and $\Delta G_{\text{toehold hairpin}} < \Delta G_{\text{anti-hairpin}}$. All RNA structures were designed and simulated using NUPACK[97] and RNAfold from ViennaRNA Web Services[152]. All G terms were calculated using NUPACK at 37◦C with the default parameter set (Serra and Turner, 1995). We adjusted the length of the stem of toehold hairpin (12 nt- 16 nt) and the free toehold region (14 nt-18 nt) to maintain the stability of the toehold hairpin.

Culture media

As previously mentioned in the earlier chapter, we employed LB medium and M9 medium for our experiments. Additionally, for the culture of *vibrio natriegens* (Vmax™ X2, Telesis Bio), we utilized LBv2 medium, which is LB supplemented with v2 salts (v2 salt: 204 mM NaCl, 4.2 mM KCl, and 23.14 mM $MgCl_2$). LBv2 medium was supplemented with 200 µg/mL kanamycin, as required for specific experiments.

Plasmid construction and cloning process

All DNA oligonucleotides were purchased from Eurofins Genomics, Ebersberg, Germany. All toehold regulator sequences were constructed by overlap extension PCR, restriction ligation and blunt-end ligation as we described in the previous chapters. The *tna* operator sequence used in some of the switch designs was obtained from the *E. coli* genome using PCR with primers. Blunt end ligation was used to fine-tune the anti-stem sequence of the riboregulators.

Fluorescence and absorbance measurements and Flow cytometry

After transforming the recombinant plasmids that encode toehold riboregulators into electro or chemically competent cells (BL21 DE3 NEB, VmaxTM X2), we conducted fluorescence and absorbance ($OD_{600}$) measurements and growth curve measurement using a same microplate reader (CLARIOstar®, BMG LABTECH) as detailed in the previous chapter. For GFP expression measurements, the following settings were used: Excitation/Emission wavelength: 570-20/630-40 nm; gain value: 1000; focus height: 2.4 mm. A Welch's t-test was calculated to determine the statistical significance ($P < 0.05$ or 0.01) of the results obtained under different conditions.

Flow cytometry of riboregulator logic gates (Fig. 14) was performed using a BD FACSMelody™ cell sorter. Cells were sampled from overnight culture and diluted by a factor of 500 into phosphate buffered saline (PBS). Each sample containing a logic gate regulator and a combination of RNA inputs was measured using $GFP^+$ as a fluorescent reporter. The forward scatter signal (FSC) was used for thresholding, and 10,000 individual events were measured in forward scatter, side scatter (SSC) and GFP fluorescence (488 excitation /536 emission) using a flowrate setting of 1.

Flow cytometry data were analyzed using FlowJo 10.1r1 software (FLOWJO). We set a gate to remove background events resulting from PBS buffer only. Specifically, we first plotted the histogram of SSC for each experimental group (toehold logic gate RNA with inputs) and the blank (background events) (Fig. S33 a). The *E. coli* population had unimodal distributions in SSC and partially overlapped with the blank.

We set the subset gating threshold by removing the overlapping part from the SSC histogram. The remaining events were used to generate a density plot with FSC-H against SSC-H (Fig. S33 b). We further create a second subset by removing the outlier events, which had a low FSC-H ($<10^2$). Then we applied a new subset gating to analyze the relative fluorescence intensities (Median of the GFP-H) of each experimental group. Figure S33 c shows the GFP fluorescence histogram for *E. coli* cells containing toehold gate RNA and inputs. Error levels for the fluorescence measurements are calculated from the s.d. of measurements from at least three biological replicates.

*In vitro* protein expression in cell-free system

All *in vitro* gene expression experiments with toehold riboregulators were performed using a commercial *in vitro* protein synthesis kit (PURExpress®, NEB) according to the manufacturer's protocol. Linear transcription templates for toehold riboregulators were first amplified using PCR

and purified using Monarch® PCR Cleanup Kit (NEB). The concentration and quality (via their 260/280 and 260/230 ratios) of purified DNA templates were quantified using a Nanodrop 8000 spectrophotometer (Thermo Fisher). The molar concentration of each DNA template was calculated via (1). A final concentration of 10 nM DNA template was used in triplicate. 25 µl of each cell-free reaction mixture was transferred to a black 384-well plate (BRAND®), covered with a plate seal (Microseal®, BIO-RAD) and placed on plate reader with the excitation/emission wavelength set to 570-20/630-40 nm, a gain value of 1000 and focus height 2.4 mm. The temperature was controlled at 37 °C, fluorescence intensity was measured every 5 min for 5h.

Total RNA extraction

Total RNA for quantitative PCR (qPCR) experiments was extracted from *E. coli* by using a TRIzol reagent (Invitrogen™) RNA isolation protocol. For each biological replicate, a single colony was picked from an LB agar plate of an overnight transformation and precultured in 500 l of M9 medium containing the appropriate antibiotics for several hours until the $OD_{600}$ reached a value of 0.5. Volumes of 10 µl each of the precultured cells were added to Conical Tubes containing 4985 l (1:500 dilution) of antibiotic containing M9 medium and grown for 4 to 5 h under the same incubation condition until the $OD_{600}$ reached the value 0.5. Then to each culture 5 µl of 1M IPTG was added (1:1,000 dilution) to induce the expression of T7 RNA polymerase, followed by further incubation for 3h. After culture, cells were collected from a 1.5 mL volume by centrifugation at 1,000g. for 5 min. After discarding the supernatant, the remaining cell pellet was suspended in 1 mL of TRIzol reagent and homogenized by pipetting, followed by incubation at room temperature for 5 min. Cell debris was removed by centrifugation at 12,000 g for 2 min. Samples were then transferred into new 2 mL tubes, 200 µl chloroform (Carl Roth) was added, and the samples were mixed for 20 s and incubated at room temperature for 3 min. After incubation, the samples were centrifuged for 10 min at 12,000 g at 4 °C, and 400 µl of the aqueous layer containing the RNA was transferred into a fresh 1.5 mL tube. 500 µl isopropanol (Carl Roth) were added to the aqueous phase, the sample was inverted and incubated at room temperature for 10 min and then centrifuged for 10 min at 14,000 rpm at 4 °C. After centrifugation, the supernatant was discarded, the remaining pellets were properly washed in 1 mL of pre-cooled 70% ethanol (Carl Roth) and centrifuged for 2 min at 14,000 rpm at 4 °C. The remaining supernatant was discarded, and the samples were dried at room temperature. 40 µl of RNase-free ddH2O were added to resuspend the pellets for further digestion.

DNase treatment of total RNA extracts

Purified total RNA samples were treated with DNase I (NEB) with reaction buffer for 1h to remove the remaining plasmid and genome DNA. After digestion of the DNA, 0.5 M EDTA solution (Invitrogen™) were added to samples (1:100 dilution) to prevent $Mg^{2+}$ dependent RNA hydrolysis.

DNase I was denatured by heating at 75C for 10 minutes. The RNA samples were further purified using an RNA-clean up kit (NEB). The concentration and quality of the purified total RNA samples were quantified via the 260/280 and 260/230 ratios using a Nanodrop 8000 spectrophotometer (Thermo Fisher).

Normalization of total RNA, reverse transcription, and qPCR measurements

The concentrations of the purified samples were determined with the Nanodrop 8000 spectrophotometer, and then diluted to 500 ng/l of total RNA in 10 µl RNase-free ddH2O. 1 µl of diluted total RNA, 0.5 µl of 10 M reverse transcription primer, 2 µl of 10 mM of dNTPs (NEB) 5X Reverse Transcriptase Buffer (biotech rabbit) and RNase-free ddH2O (up to 18.5 µl) were incubated for 5 min at 65 °C and cooled on ice for 5 min. 1 µl of RevertUP II reverse transcriptase (biotech rabbit), 0.5 µl of Murine RNAse inhibitor (NEB) were then added, and the solution was incubated at 55 °C for 1 h, 80 °C for 5 min and then stored at – 20 °C. qPCR was performed using 5 µl of Lunar qPCR master mix (NEB), 1 µl of cDNA and 0.5 µl of 0.5 M mCherry qPCR primers (Primer list) and up to 10 µl RNasefree ddH2O (dilute cDNA if necessary). An iQ™ 5 real-time PCR machine (BIO-RAD) was used for data collection using the following PCR program: 95◦C for 2 min, followed by 30 cycles of 95◦C for 15 s and 60◦C for 35 s. All the measurements were followed by melting curve analysis. Strips of 8 Thermo-Tubes in White & Clear Caps (Thermo Fisher) were used for all measurements. Results were analyzed using iQ™ 5 software (BIO-RAD). To quantify the relative abundance of cDNA concentration, a 5-point standard curve covering a 10,000-fold range of quantified linear DNA concentrations was measured and used to determine the relative mCherry cDNA abundance in each sample (Appendix, Fig. S32). Non-template controls were run in parallel to each measurement to check contamination and nonspecific amplification or primer dimers. Additionally, qPCR was performed on total RNA samples to confirm that no DNA plasmid was detected under the same conditions. Melting curves were recorded to confirm that only a single product was amplified.

# Chapter IV

## Switchable Fluorescent Light-Up Aptamers Based on Riboswitch Architectures

**Keywords:** Fluorescent light-up aptamers • toehold-mediated strand displacement • purine aptamer • sensors

4

# Abstract

Fluorescent light-up RNA aptamers (FLAPs) such as Spinach or Mango can bind small fluorogens and activate their fluorescence. Here, we adopt the similar design principle that were discussed earlier in the above chapter by utilizing the switching mechanism typically found in riboswitches to engineer switchable FLAPs that can be activated or repressed by trigger oligonucleotides or small metabolites. The fluorophore binding pocket of the FLAPs comprises guanine (G) quadruplexes, whose critical nucleotides can be sequestered by corresponding anti-FLAP sequences, leading to an inactive conformation, and thus preventing association with the fluorophore. In this chapter, we build upon the design principles discussed in the previous chapter to further advance the field of FLAPs by incorporating designed toehold hairpins which are strategically engineered to carry either an anti-FLAP or an anti-anti-FLAP sequence within their loop region. The addition of an input RNA molecule triggers a toehold-mediated strand invasion process that refolds the FLAP into an active or inactive configuration. Several of our designs display close-to-zero leak signals and correspondingly high ON/OFF fluorescence ratios. We also modified purine aptamers to sequester a partial anti-FLAP or an anti-anti-FLAP sequence to control the formation of the fluorogen-binding conformation, resulting in FLAPs whose fluorescence is activated or deactivated in the presence of guanine or adenine. We demonstrate that switching modules can be easily combined to generate FLAPs whose fluorescence depends on several inputs with different types of input logic.

# Introduction

Biomolecular fluorescent reporters are powerful tools with a wide range of applications in cell biology, molecular biology, and biomedicine. Over the past decades, in particular the green fluorescent protein (GFP) and its derivatives have been broadly utilized in bio-labeling and imaging[153], biosensing[154], the exploration of signaling pathways[155] and the characterization of gene circuits[156]. With the recent development of RNA-based fluorescent light-up aptamers (FLAPs) such as Spinach[157], Mango[158], Corn[159], RhoBAST[160], and others[161], the scope of fluorescent reporters has been considerably widened. The Spinach aptamer can bind the small fluorogen DFHBI (3, 5-difluoro-4-hydroxybenzylidene imidazolinone) and thereby increase its fluorescence quantum yield by roughly thousandfold (with an emission wavelength of $\lambda_{em}$=501 nm), while Mango binds to the thiazole-orange derived TO1-biotin fluorogen, leading to a similar increase in fluorescence at $\lambda_{em}$=535 nm. Compared to fluorescent proteins, FLAPs only need to be transcribed from a DNA template and do not require translation or maturation, which makes their dynamics potentially much faster. As functional RNA molecules, they are ideally suited for applications in RNA sensing or as readouts for RNA-based gene circuits. However, they typically suffer from a lower stability and signal intensity *in vivo*.

Based on crystallographic studies[162], FLAPs such as Spinach, Mango or Corn all share three basic structial motifs: i) two G-quadruplex platforms, ii) a triplex lid, and iii) an in-plane guanine, which together assemble a binding pocket for the fluorogen[161] (Fig. 15 d). In addition, a potassium ion is required to stabilize each G-quadruplex. The fluorogen is bound to the binding pocket through non-covalent interactions between its functional groups and the RNA residues. Given the importance of the G-quadruplex motif for the FLAPs, various studies previously attempted to develop switchable FLAPs as biosensors by controlling the folding of the G-quadruplex and the overall fluorogen-binding conformation. For instance, FLAPs can be designed as fusion aptamers (Fig. S41), combined with other sensor units such as RNA aptamers[163], riboswitches[164,165], reverse complementary sequences[166] and antigens[167], and coupled to in situ amplification methods such as the hairpin chain reaction[168]. Upon binding to their cognate targets (small metabolites, RNA, proteins), a structural rearrangement is induced that brings the FLAPs into the stabilized fluorogen-binding conformation. Following a different strategy, folding of split-FLAPs into a fluorescently active conformation can also be facilitated by ligand binding[169].

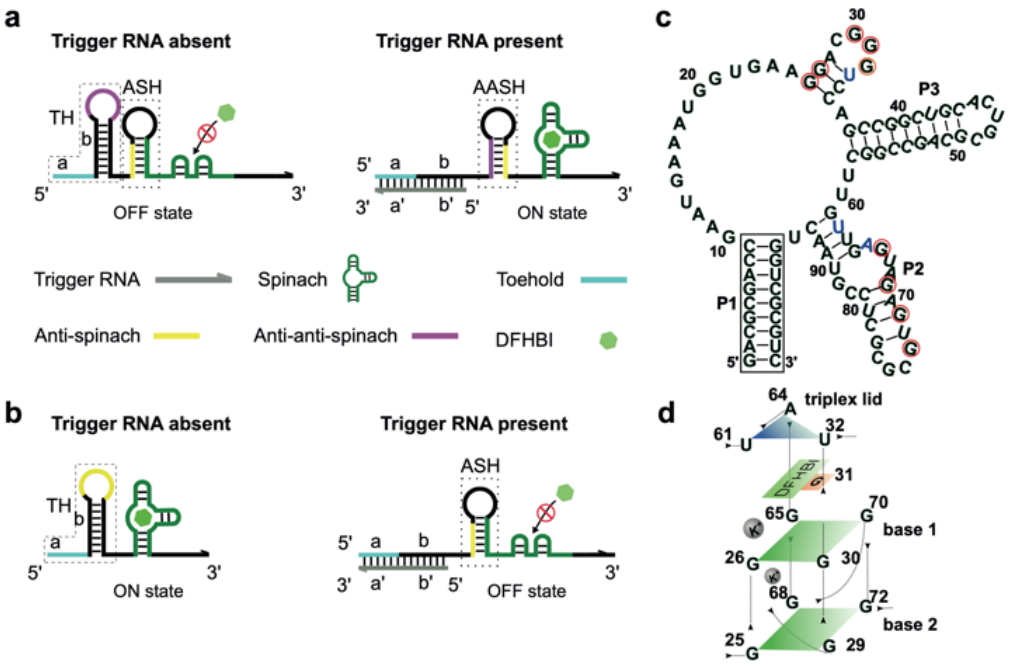


Figure 15. Overview of fluorescent light-up aptamers switched by RNA inputs via toehold-mediated strand invasion. a), Scheme of the switching process of a toehold-Spinach activator. The activator comprises a toehold hairpin (TH) and an anti-Spinach hairpin (ASH), which prevents folding of the Spinach aptamer into an active conformation. In the presence of trigger RNA, an anti-anti-Spinach hairpin (AASH) is formed, and Spinach can fold correctly. b), conversely, a toehold-Spinach repressor comprises a toehold hairpin and a Spinach aptamer. In the presence of trigger RNA, an ASH is formed, preventing formation of the Spinach aptamer. c), Secondary structure of the Spinach aptamer. Red circles indicate critical guanines involved in G-quadruplex formation, while blue circles indicate nucleotides involved in triplex lid formation. The box marks the lower stem of the aptamer. d) Schematic image of the binding pocket highlighting G-quadruplexes and triplex lid, bases are numbered as in c).

In the present work, we sought to utilize switching mechanisms derived from naturally occurring riboswitches to develop FLAPs from fluorogenic aptamers such as Spinach and Mango that can be switched using short RNA triggers and small metabolites[85] (adenine and guanine) as inputs. They are composed of two functional domains - the metabolite-binding domain (an RNA aptamer structure) and the “expression platform”, which can switch between two alternative conformations in the free and metabolite-bound states[170].

According to the discussion from the previous chapters, we incorporated the anti-RBS sequence within the expression platform of the riboswitch in our riboregulators. This sequence forms base pairs with the RBS and experiences a conformational change upon RNA binding. As a result, the RBS and start codon become accessible to the ribosome, thereby initiating protein expression from the downstream open reading frame (see Fig. S21). We adopted this switching principle to realize switchable FLAPs, in which the folding of the FLAP aptamer is controlled by the presence of an adjacent anti-FLAP sequence. We found that anti-FLAP sequences that sequester the critical guanine nucleotides of the FLAPs can prevent the folding of the fluorogen-binding conformation very efficiently, resulting in an extremely low leak signal and thus a very high dynamic range. In the case of ON-switchable FLAPs, an anti-anti-FLAP sequence is initially sequestered in an alternative secondary structure. Refolding of the structure can be promoted by an RNA trigger molecule, which releases the anti-anti-FLAP through a TMSD process[171,172]

similar as in gene regulatory toehold switches[11]. Upon re-folding, the anti-FLAP sequence is bound by the anti-anti-FLAP, allowing the FLAP to fold into its fluorogen-binding conformation. Alternatively, the anti-FLAP sequence can be hidden within the secondary structure of a purine aptamer[173-175], which allows to activate the FLAP by the presence of either adenine or guanine. OFF-switchable FLAPs are designed accordingly by initially sequestering the anti-FLAP sequence and releasing it upon RNA or ligand binding. In our experimental setting, the FLAPs are transcribed from their corresponding gene templates in the absence or presence of triggers, and fluorescence is monitored during *in vitro* transcription (IVT). We also demonstrate switchable FLAPs with two-input logic, whose fluorescence output depends on the presence of two different triggers. Several of our designs of RNA switchable FLAPs display essentially zero intrinsic leak and can be switched by freely choosable input RNAs, and are thus well-suited to act as reporters for *in vitro* RNA circuits and nucleic acid amplification schemes.

# Results

RNA-switchable FLAPs based on TMSD

Our initial goal was to engineer switchable FLAPs with high ON/OFF signal ratios that are activated through binding to small trigger RNAs. To this end, we designed switchable toehold aptamers byaugmenting the original FLAP with a toehold hairpin (TH) at its 5' end containing a 14 nt long, single-stranded toehold region and a 16 bp long stem (Fig. 15). Inspired by the structure of riboswitches, we placed a regulatory sequence within the loop region of the toehold hairpin, which was designed to interact with a specific downstream sequence. The designs of the switchable Spinach and Mango aptamers are completely analogous, and we here first focus on the Spinach aptamer. In the case of the Spinach activator (Fig. 15 a, Fig. S42), in the absence of a trigger RNA the toehold hairpin sequesters an anti-anti-Spinach (AAS) domain (generally, an "anti-anti-FLAP" domain), while an anti-Spinach (AS) sequence base-pairs with critical nucleotides of the Spinach aptamer to form the stem of the anti-Spinach hairpin (ASH). For the design of the FLAP sequences, we adjusted the folding free energies of critical secondary structural elements using the nucleic acid analysis tool NUPACK[97]. The folding free energy of the ASH ($\Delta G_{ASH}$) is designed to be lower than the free energy of the lower stem P1 of the Spinach aptamer ($\Delta G_{Spinach\text{-}P1}$ = -17.48 kcal/mol). In the absence of trigger, the formation of the G-quadruplexes and thus binding of DFHBI is disabled (OFF state). In the presence of trigger RNA, the toehold hairpin is opened via TMSD, releasing the AAS sequence, which can base-pair with the AS sequence and thus form a stable anti-anti-Spinach hairpin (AASH). The free energy of the AASH ($\Delta G_{AASH}$) is chosen intermediate between the free energies of ASH and TH ($\Delta G_{TH} < \Delta G_{AASH} < \Delta G_{ASH}$). The conformational rearrangement restores the G-quadruplexes, which facilitates binding of the DFHBI fluorogen and thus activates fluorescence (ON state). We also designed toehold-Spinach repressors (Fig. 15 b), which use a similar switching process to inhibit

the formation of the binding pocket for the fluorogen. In this case, the TH loop is used to sequester the anti-Spinach sequence. In the absence of trigger RNA, the Spinach aptamer correctly folds. In the presence of trigger RNA, the TH is unfolded, releasing the AS, which in turn leads to the formation of the anti-Spinach hairpin and thus inhibition of fluorogen binding (OFF state).

To fine-tune the switching behavior of the toehold-Spinach activators, we designed and tested five different AS sequences that sequester different sub-sets of the nine critical guanine nucleotides of the Spinach aptamer (Fig. 16 a). In order to assess their performance, we transcribed both switch RNA and trigger RNA in the presence of 20 µM of fluorogen using the same DNA template concentration (10 nM) using an IVT mix (cf. Experimental Section and Supporting Information) and simultaneously measured the fluorescence signal. Anti-Spinach sequence #1 can base-pair with the first eleven nucleotides of the lower stem sequence of the Spinach aptamer (in this case none of the critical guanines is targeted), while anti-Spinach #2 and #3 target the loop region of the Spinach aptamer (nucleotides 10 to 34), including 5 critical guanines and a uracil that participate in the formation of the G-quadruplex bases and the triplex lid (cf. Fig. 15 d). In addition, anti-Spinach #4 and #5 can bind to the critical nucleotides in the range from 61 to 86, which is close to the 3' end of the aptamer. As shown in Fig. 2a, in the OFF state – in the absence of trigger – fluorescence is tightly suppressed for all versions. Notably, for guanine-sequestering versions #2, #3, #4 and #5 the signal is indistinguishable from the fluorescence of a blank measurement containing only transcription mix and DFHBI. Among these versions, #3 provided the highest ON signal in the presence of trigger RNA, corresponding to an ON/OFF ratio of ≈ 22 under the conditions of this co-transcriptional experiment. FLAP #3 sequesters all three critical guanines on the 5' side of the Spinach aptamer, which appears to be favorable for refolding upon activation. FLAP #4, sequestering four guanines closer to the 3' end of Spinach appears to be less performant in comparison. This may be attributed to the larger distance between the anti-anti-Spinach sequence in the TH at the 5' end and the ASH. We also studied the switching behavior of FLAP #3 using other experimental settings, either by co-transcribing switch and trigger RNA or using purified RNA components, and found ON/OFF ratios ranging between 20 and over 200 (Fig. S43, note that the theoretical maximum would be ≈ 1000, which is given by the increase in quantum yield when DFHBI binds to Spinach[161]), indicating that experimental conditions can be further optimized for sensor applications.

For the toehold-Spinach repressors, we similarly designed five variants with different AS sequences (Fig. 16 b). Anti-Spinach #1 and #2 were designed to sequester the sequence of lower stem P1 close to the 5' end of the Spinach aptamer and thus prevent its folding. Anti-Spinach #1 also includes a part of the linker sequence between the toehold hairpin and the aptamer, and resulted in the highest ON/OFF ratio among the five studied AS sequences. Anti-Spinach #2 also targets the lower stem P1 but exhibits higher leakage in the OFF state than anti-Spinach #1, leading to a lower ON/OFF ratio. Other variants such as anti-Spinach #3 (targeting two critical guanine nucleotides) and anti-Spinach #5 (targeting the 3' end of lower stem P1)

resulted in lower ON/OFF ratios or, as in the case of anti-Spinach #4 (targeting five guanines), did not perform at all. The reduced performance of the repressors compared to the activators is expected and is similar as observed for riboregulator switches. Whereas the activators will always have a very low OFF state due to the co-transcriptional sequestration of a critical subsequence of the fluorogenic aptamer, repressors initially fold into an active fluorescent state, unless they are de-activated by trigger RNA co-transcriptionally. The latter process is expected to be inefficient, however, for typical RNA trigger concentrations, and thus a considerable leak signal is unavoidable.

We analogously applied our design approach to the Mango aptamer[158] which also folds into a G-quadruplex topology, in which it binds to its specific fluorogen TO1-biotin (Fig. S44). As for the Spinach aptamer, we designed five toehold-Mango activators comprising different anti-Mango (AM) sequences to target sub-sequences of the mango aptamer (Fig. S45). Unlike the Spinach aptamer, AM variants #1 and #5, which interacted with the stem of the Mango aptamer near its 5' or 3' end, exhibited superior performance compared to variants #2 to #4, which targeted the loop region containing the crucial nucleotides. We also designed five variants of a toehold-Mango aptamer repressor, of which again only those targeting the Mango stem (#1, #5) resulted in switching. The RNA-switchable Mango aptamers generally performed less than the Spinach versions, which we attribute to the lower $K_D$ for the complex of Mango with TO1-biotin (≈ 3 nM) compared with the Spinach-DFHBI complex (≈ 300 nM), presumably leading to a higher leak signal in the OFF state.

Switchable FLAPs based on purine aptamers

Purine aptamers consist of a three-way junction structure in which three stems (P1-P3, cf. Fig. 17 c) surround a central core that contains several critical nucleotides for ligand binding[173,176]. The P2 and P3 stems can interact via long-range loop-loop interactions even in the absence of ligand, which along with the core forms a pre-organized binding pocket that enables rapid ligand binding (Fig. 17 d)[175].

In the context of a riboswitch, once the aptamer binds its ligand, quadruplex interactions within the central core and the P1 stem are further stabilized, preventing the disassembly of the aptamer and rearranging the secondary structure of the downstream sequence. In this work, we modified the P1 stem of the guanine and adenine aptamers to achieve switching of FLAPs in the presence of purines (Fig. 17 a & b, Fig. S46). The stability of the P1 stem affects the free energy of formation of the ligand-binding pocket. We therefore designed a series of structures, in which the P1 stem was modified to contain different AS or AAS sequences, each comprising six bases complementary to specific downstream sequences. The critical nucleotides of the guanine or adenine aptamers that participate in loop-loop interaction and binding pocket formation were left untouched. In the case of the guanine-Spinach activator (Fig. 17 a), in the absence of guanine the AS sequence base-pairs with the Spinach aptamer to form a stable ASH, which prevents folding of the Spinach G-quadruplexes and fluorogen binding (OFF state).

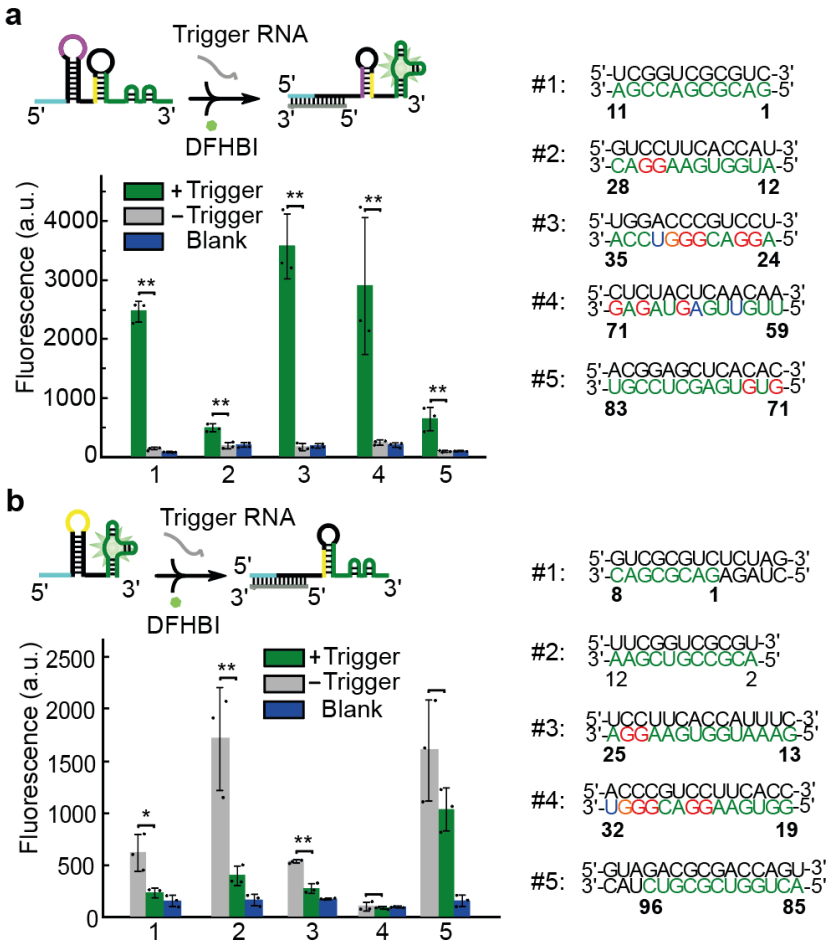


Figure 16. Relative fluorescence intensities obtained from different toehold-Spinach activators a), and repressors b), that utilize different anti-Spinach sequences in the presence or absence of trigger RNA respectively. Sequestered critical guanine nucleotides of the Spinach aptamer are highlighted in red (cf. Fig. 1c). The heights of the bars are obtained as the mean of three independent replicates, the error bars represent their standard deviation (s.d.). In a), activators #3 and #5 displayed zero fluorescence above background in the absence of trigger RNA, and therefore the ON/OFF ratio was not calculated. The statistical significance of the difference between the trigger RNA+ and the trigger RNA- condition was determined via Welch's t-test, ** indicates a p-value < 0.01.

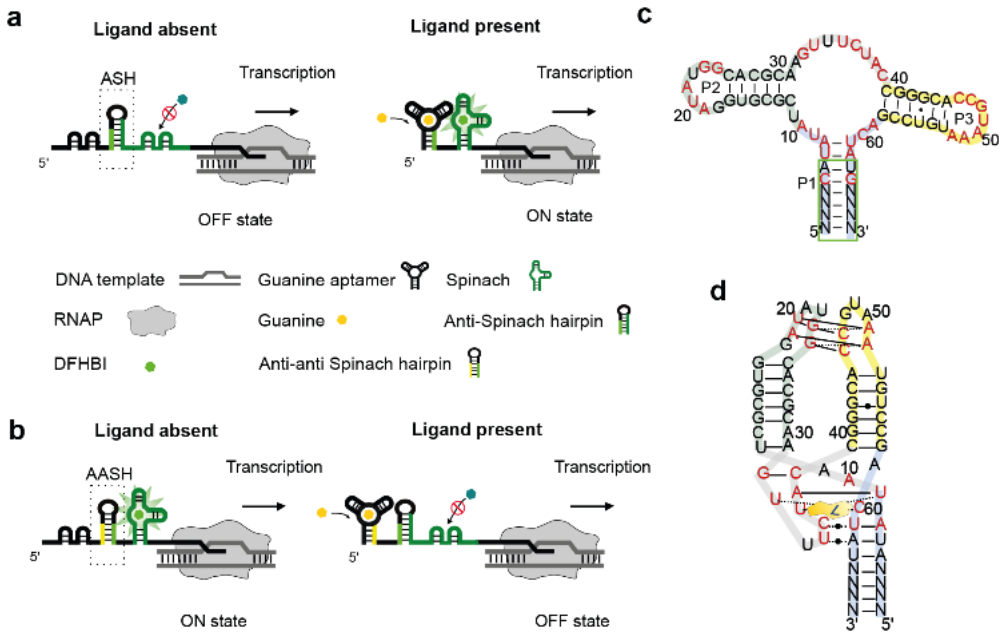


Figure 17. Guanine-dependent Spinach switches. a), Scheme of the switching process for the guanine-Spinach activator. In the OFF state, critical nucleotides of the Spinach aptamer are sequestered in the anti-Spinach hairpin (ASH). In the presence of guanine, the binding pocket of the guanine aptamer is stabilized, and the anti-Spinach sequence is sequestered in its P1 stem. Hence, the Spinach aptamer can fold, bind to DFHBI and activate its fluorescence. The transcription process is also indicated as its kinetics determines the time window during which the ligand can bind, which is given by the time between the formation of the guanine binding pocket (see panels c) & d)) and the formation of the anti-spinach hairpin. b), in the guanine-Spinach repressor, in the absence of guanine the anti-Spinach sequence is bound in the anti-anti-Spinach hairpin (AASH). In the presence of guanine, the AS sequence sequesters critical nucleotides of the Spinach aptamer and thus de-activates it. As in panel a), binding of guanine is assumed to occur co-transcriptionally. c), Secondary structure of the guanine aptamer with the P1, P2 and P3 stems indicated. Nucleotides critical for the function of the aptamer are colored in red. The "N" nucleotides in the P1 stem (highlighted in the green box) were varied for the design of the guanine-switchable FLAPs. d), Scheme of the tertiary structure of the guanine ligand binding pocket involving loop-loop interaction between the P2 and P3 stems.

To ensure proper switching, the folding free-energy of the ASH was designed to be lower than that of the pre-organized, ligand-free guanine aptamer structure ($\Delta G_{(G\text{-free})}$) and of the P1 stem of the Spinach aptamer (i.e., $\Delta G_{ASH} < \Delta G_{(G\text{-free})}$ and $\Delta G_{ASH} < \Delta G_{(spinach\text{-}P1)}$). Upon ligand binding, the guanine aptamer is stabilized ($\Delta G_{(G\text{-bound})} < \Delta G_{ASH}$), suppressing the formation of the ASH and thus activating Spinach folding and fluorescence (ON state). Conversely, in the case of the guanine-Spinach repressors, we inserted an AAS into the P1 stem of the guanine aptamer, which can form a stable AASH and thus – in the absence of guanine – allow the formation of the Spinach G-quadruplexes (ON state). As above, the folding free energy of the AASH is designed to be lower than $\Delta G_{(G\text{-free})}$ and $\Delta G_{ASH}$. Ligand binding stabilizes the aptamer and the AAS sequence is sequestered within its P1 stem, which in turn results in the formation of an anti-Spinach hairpin (OFF state). For all guanine-Spinach switch designs, we utilized the anti-Spinach sequence #1 developed for the toehold-Spinach switches and fine-tuned the free-energy of the guanine aptamer P1 stem to optimize the switching performance. As shown in Figure 17 a, we designed four P1 versions for the guanine-Spinach activators, of which two (#1 and #2) were comprised of 8 bp (resulting in a folding free energy > -10 kcal/mol), while versions #3 and #4

comprised an additional base-pair in the stem (and a free energy below –10 kcal/mol). We again tested the guanine-Spinach activators in IVT reactions using ligand concentrations of 20 µM (DFHBI) and 0 µM or 50 µM of guanine, respectively. We found that with co-transcriptional ON/OFF ratios of ≈ 12 and ≈ 23 the short stem versions #1 and #2 performed considerably better than the two other versions, which exhibited relatively high leak signals. The leak is likely caused by the higher stability of the P1 stem in versions #3 and #4, whose formation competes with the folding of the anti-spinach hairpin and might stabilize the active Spinach conformation also in the absence of guanine. In the case of the guanine-Spinach repressors, four design versions with different P1 stem sequence were investigated (Fig. 17 b). Repressor version #2 with an internal mismatch in the P1 stem performed worst in terms of leak signal in the OFF state but had the highest fluorescence signal in the ON state, resulting in an ON/OFF ratio of ≈ 15. All other designs with more stable stems showed a lower leak, but displayed a reduced ON-fluorescence, indicating the mismatch is required to facilitate refolding of the secondary structure. We applied the same design principle to develop ligand-dependent spinach switches based on an adenine aptamer derived from a riboswitch found in many Gram-positive bacteria[177], resulting in several performant adenine-sensing activators and repressors (Fig. S46).

Logic gate construction by using switchable FLAPs

We next sought to combine several of the toehold- and ligand-controlled FLAPs to create two-input switches with various types of input logic. To this end, we developed a design strategy, in which the folding of the Spinach aptamer structure was influenced by introducing switchable domains at both of its 5’ and 3’ ends.

To achieve AND gate behavior, we combined the Spinach aptamer with two toehold-Spinach activator domains (Fig. 18 a). Each activator includes an ASH which sequesters critical nucleotides of the Spinach aptamer and prevents folding of its G-quadruplexes. Only in the presence of the two cognate RNA triggers as inputs, each module is activated by TMSD, the G-quadruplexes are restored and fluorescence is activated. Similarly, we also combined two toehold-Spinach repressor modules with Spinach aptamer to sequester its critical nucleotides either from 5’ end or 3’ end, resulting in NOR gate behavior (Fig. 18 b).

Accordingly, the combination of an activator and a repressor module on the same FLAP platform results in a logic NIMPLY (= A AND (NOT B)) gate (Fig. 18 c). Notably, in this case we observed an appreciable leak signal in the presence of both trigger RNAs, indicating that inhibition by the repressor module at the 3’ end was incomplete. As fluorescence was monitored during *in vitro* transcription of the switchable FLAP, it is likely that folding of a fraction of the Spinach aptamers was promoted by the activating trigger before the repressing trigger could bind to the repressor module. We also created an AND gate by controlling the Spinach aptamer with both a guanine- and an RNA-dependent activator (Fig. 18 d). Such or similar gates may play a role in the evaluation of diagnostic rules, which can be formulated as logical expressions, such as the detection of cancer-related miRNA patterns[178,179].

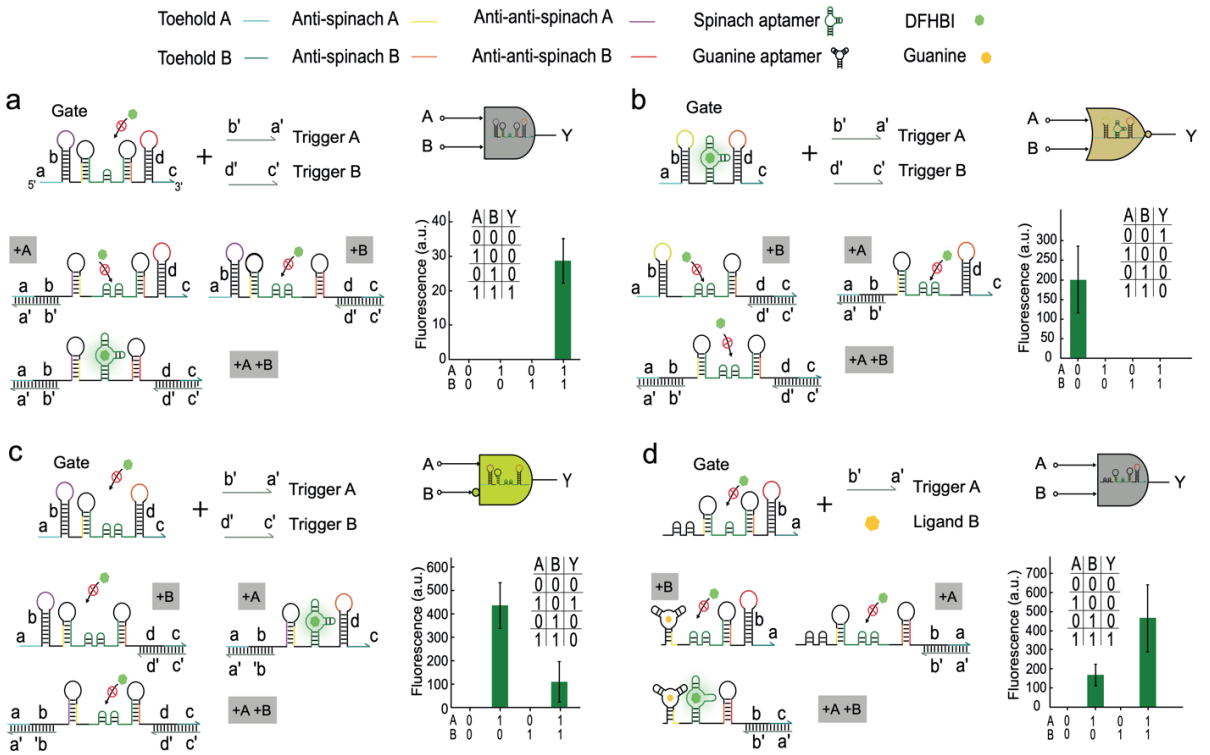


Figure 18. Design and characterization of two-input logic gates based on toehold-Spinach and guanine-Spinach switches. a), A two-input AND gate is realized with two toehold-Spinach activators at the 5' and 3' termini of the Spinach aptamer. Each activator module is composed of a toehold hairpin and an ASH. Binding of input Triggers A and B opens their respective toehold hairpins and restores the Spinach aptamer. b), A two-input NOR gate is composed of two toehold-Spinach repressors at its both ends. Binding of either Trigger A or B leads to the disruption of the Spinach aptamer. c), A NIMPLY gate combines an activating and a repressing input module, leading to activation of fluorescence only in the presence of Trigger A as indicated. d), An AND gate with hybrid input composed of a guanine-Spinach activator and a toehold-Spinach activator. The presence of guanine stabilizes the purine aptamer at the 5' end, while RNA Trigger A leads to a refolding of the structure at the 3' end, resulting in AND gate input logic as indicated. The fluorescence outputs of all gates are shown as the mean values of background subtracted fluorescence levels for three independent measurements, error bars represent their s.d

# Discussion

In summary, we have successfully developed and characterized switchable fluorescent light-up aptamers (FLAPs), which can be switched using trigger RNAs or purine ligands. Our switchable FLAPs combine the structural features of natural riboswitches with the switching mechanism of synthetic riboregulators. This approach allows the realization of both ON and OFF switches with a performance that compares favorably with previously developed switchable Spinach aptamers[180,181]. In our designs, we employed a similar strategy as described in the previous chapter. We inserted a regulatory RNA sequence into the loop region of a toehold hairpin which could be switched by toehold-mediated strand displacement, inducing the formation of a secondary structure in which the fluorogenic aptamer is either activated (for ON switches) or de-

activated (OFF switches). Importantly, the sequence of the toehold-region and the stem – and thus the trigger RNA - can be freely chosen and is not constrained by the aptamer sequence itself. In order to demonstrate the modularity and orthogonality of our design, we studied the crosstalk between eight toehold-Spinach activators with orthogonal trigger RNAs. As shown in Fig. S47, the switches exhibit highly orthogonal switching behavior, indicating the potential for application in the detection of natural RNA sequences.

Using a similar strategy, we have also shown that FLAPs can be rendered into riboswitch-inspired aptamer switches, in which the binding of a small molecule ligand influences the formation of the fluorogenic aptamer. In contrast to recent work, in which the adenine riboswitch was re-purposed into an allosteric light-up aptamer using a SELEX approach[165], we re-modeled the variable sequences on the P1 stem of the guanine or adenine aptamers into specific AS and AAS sequences, resulting in strong fluorescence suppression in the OFF state and improved ON/OFF ratios. Purine aptamers have a characteristic preorganized ligand-binding structure. Previous studies revealed that several nucleotides contributing to the formation of the ligand-binding site have a major influence on the $K_D$ of ligand binding[105]. Aside from these nucleotides, the free energy of folding of the preorganized structure affects the stability of the ligand-binding site, and thereby influences the refolding process of the downstream structure. We found that when the free energy of the purine aptamer domain of our switch is below -18.00 kcal/mol, it exhibits the ability to sequester the AS sequence even in the absence of its ligand. We hence fine-tuned the free energy of the P1 stem to influence the refolding of the ASH and the Kd of ligand binding. Interestingly, for both guanine and adenine Spinach switches, the ON/OFF ratios did not have a clear correlation with the free energy of the P1 stem, suggesting a more complex competition between the formation of the ASH and the ligand-binding pocket. This suggests a more intricate interplay between the formation of the ASH and the ligand-binding pocket in determining the switch's behavior. In our design, we have not performed calculations to determine the binding free energy of the ligand-bound aptamers, both for guanine/adenine and spinach switches. This absence of parameter creates ambiguity when calculating the overall free energy of the entire aptamer between the ON and OFF states. Consequently, the uncertainties in predicting the actual free energy differences could lead to variations in the performance of the P1 stem sequences. Previous work by Salis and colleagues showcased an automated computational design method for developing synthetic riboswitches, which involves calculating an energy model using RBS and Riboswitch calculators [140,182]. However, due to the limited availability of experimental data for the ligand binding free energy of aptamers, the computational predictions are currently restricted to a few well-studied aptamers, such as Theophylline[182], Fluoride[183] and TMR[184]. Accordingly, further experimental characterization is crucial to expand the predictive capabilities and application of computational design methods for designing novel riboregulators and customized RNA switch with various functionalities.

Our design approach also allows for control over the folding of the FLAP using two distinct toehold hairpins attached to the 5’ and 3’ end, resulting in two-input logic control of fluorescence

activation by two independent RNA inputs. We also developed a two-input switch that is activated by a small ligand and a trigger RNA, using combination of the guanine-Spinach switch and the toehold-Spinach switch. While the AND gate activators with two RNA inputs displayed almost ideal behavior, logic gates involving repressor modules or guanine as an input were slightly leaky, as would be expected from the behavior of the individual switches. In principle, our general approach should be applicable also to other fluorogenic aptamers containing “critical nucleotides”. A potential approach to extend our two-input gates to larger numbers of inputs could be the utilization of multi-arm junctions as input modules, as previously demonstrated for translational toehold riboregulators[13].

With potential *in vivo* sensing applications in mind, we also engineered the Broccoli aptamer into a switch, which is known to display greater *in vivo* stability in bacteria. Such a switch indeed appears to be functional in *E. coli*, albeit with a moderate ON/OFF ratio of ≈ 4 (Fig. S48).

Isothermal detection methods, such as NASBA, LAMP, and paper-based diagnostics using toehold switches, offer several advantages over qPCR detection. These include the simplicity of the device, cost-effectiveness, and reduced detection time. To further explore the potential applications of diagnostic RNA sensing, we devised toehold FLAPs capable of responding to *orf7* mRNA sequences from the SARS-CoV-2 genome. These partial mRNA sequences were pre-amplified using the isothermal amplification method NASBA[12,185] (Further details are provided in the Appendix, Fig. S49). Our experimental findings revealed that with pre-amplification, we successfully detected as low as 100 fM of target RNA within a short time frame of 20 minutes.

In comparison to the aforementioned isothermal detection methods or paper-based diagnostics, our *in vitro* transcribed toehold FLAP offers distinct advantages, including lower cost and shorter reaction time. The results obtained from our experiments emphasize the potential suitability of these toehold FLAPs for clinical diagnosis purposes, where rapid and cost-effective detection is of utmost importance.

Furthermore, our design strategies can be readily extended to incorporate other small regulatory RNAs, including CRISPR guide RNAs (as depicted in Appendix, Fig. S50). In this context, we elected to engineer the Cas12a guide RNA (gRNA), applying a toehold hairpin at its 5' end to exert control over the structural configuration of the Cas12a handle. Concomitantly, in the presence of the trigger strand, the anti-handle sequence within the gRNA forms base-pair interactions with the Cas12a handle, which results in the obstruction of Cas12a binding. This in turn leads to the repression of the Cas12a-gRNA complex formation and subsequent inhibition of the target DNA strand cleavage process.

This versatility enables these RNAs to effectively recognize and interact with random sequences, paving the way for their potential application in *in vivo* gene circuits for sensing and self-regulation. Such capabilities hold promise expanding the range of RNA-based regulatory tools available for various applications.

# Conclusion

In conclusion, we have demonstrated that the molecular architecture of riboswitches and riboregulators can be adopted to realize switchable fluorescent light-up aptamers which are controlled by one or several RNA or small molecule inputs. The resulting switches show very low OFF signals, which in several cases is indistinguishable from background. Notably, our design allows a complete decoupling of the inputs from the FLAP sequence, which is essential for the realization of sensor or bio-computing applications without any sequence constraints. Some of the features of our switches can be easily understood – e.g., leaky signals in OFF states in cases where the RNA structures are too weak, or the loss of "switchability" in cases, where one of the alternative structures becomes too strong. Nevertheless, rational design of switches for low OFF signals and high dynamic range remains challenging[186]. We anticipate that further optimization of such switchable RNA structures may benefit from recent machine learning approaches[113], and incorporation of "community-based" knowledge[29,187].

# Materials and methods

Culture media and Cell culture

We employed LB medium (Carl Roth) and Turbo® cloning strain to culture our cells. The medium was supplemented with 100 µg/ml kanamycin. Bacterial strains were grown in LB media using 5 mL culture each in 50 mL centrifuge tubes at 37°C while shaking at 250 rpm.

Plasmid construction and cloning process

All DNA oligonucleotides were obtained from Eurofins Genomics, Ebersberg, Germany. FLAPs and aptamer sequences were constructed using a combination of overlap extension PCR, restriction ligation, and blunt-end ligation as we described in the previous chapters.

*In vitro* transcription

All *in vitro* gene transcription experiments with switchable FLAPs and trigger RNAs were performed using a homemade *in vitro* transcription mix including a homemade T7 RNA polymerase. As we described in the previous chapter.

Orthogonal experimental design

We performed orthogonality tests for the toehold-Spinach activator using 9 different toehold sequences to illustrate the possibility to choose arbitrary, orthogonal RNA inputs. Each activator was co-transcribed with 9 different trigger RNA templates in 10 µl transcriptional mixture.

Pipetting was performed using an I.DOT non-contact liquid handler (Dispendix GmbH) to reduce pipetting time and errors.

NASBA amplification

We used the NASBA wit kit from AMSBIO to perform the reaction. Firstly, reaction buffer (6.6 µL) and nucleotide mix (3.3 µL) were aliquoted for each single reaction and stored at -80 °C. Next, primers (250nM) and RNA template (10 pM) were diluted in water to their final concentrations in 20 µL. The buffer and nucleotide mix were then added to the dilution and gently mixed for a few seconds, followed by an annealing process (heating for 2 minutes at 65°C and cooling for 10 minutes at 41°C). Enzyme mix (5 µL) was added to the dilution and incubated at 41°C for 90 minutes. The RNA product was analyzed using Urea Page gel electrophoresis. After, amplification, we added 3 µL of product into *in vitro* transcription mix and followed the above protocol for fluorescence measurement.

Urea-PAGE purification

After *in* vitro transcription, the trigger RNAs were purified using Urea-PAGE gel and ZR small-RNATM PAGE Recovery Kit as we described in the previous chapter.

Fluorescence microscopy

For Fig. S48, we observed the cells under a fluorescence microscope (IX81, iXon3 DU88) using a 100x objective with an exposure time of 500 ms. Images were processed with ImageJ. Fluorescence images in Fig. S48 d were Gaussian blurred (σ = 2 pixels), background subtracted (Rolling ball radius 50.0) and then scaled to the same range of pixel values for comparison.

Fluorescence measurements with microplate reader

In triplicate experiments, we used a final concentration of 10 nM DNA template for each. Corresponding fluorogens (stocks solutions were 400 µM DFHBI from Sigma-Aldrich®, 66 µM TO1-biotin obtained from abm®, or 400 µM BI from Lucerna Technologies) were added and diluted to their final concentrations from stock solutions. To test guanine and adenine Spinach aptamers, 10 mM guanine or adenine powder were dissolved in 100 mM NaOH in 1ml and diluted in 20 µl of each transcription mixture. The samples were transferred to a black 384-well plate to perform fluorescence measurements under CLARIOstar® plate reader with excitation/emission wavelength set to 447/501 (DFHBI), 505/540 nm (TO1-biotin), a gain value of 1000, and focus height 2.4 mm. The temperature was controlled at 37°C, and the fluorescence intensity was measured every 2 minutes for 5 hours. The data were analyzed using MARS data analysis software. We report averaged fluorescence values of the replicate measurements and compare them with those of blank measurements obtained with a transcription mix also

containing the fluorophore. A Welch's t-test was calculated to determine the statistical significance ($P < 0.05$ or 0.01) of the results obtained under different conditions.

As alternative, we also used a Cary Eclipse Fluorescence Spectrometer to measure the fluorescence intensity of the PAGE gel-purified toehold-spinach RNA. In each experiment, 0.5 μM RNA samples were mixed with transcription buffer, Murine RNase inhibitor, and BI (0.5 μM) in a 40 μl volume. The mixture was then transferred to a fluorescence cuvette and measured at excitation/emission wavelengths of 470/505. The temperature was maintained at 37°C, and the fluorescence intensity was measured every 0.1 minutes over a period of 20 minutes.

To image the bacterial cells, we first grew a pre-culture from a single colony in 5 mL of LB medium until the $OD_{600}$ reached 0.5. Then, we induced RNA switch expression by adding 1 mM IPTG to the culture and incubated it for 1 hour. The cells were centrifuged and washed twice with 1x PBS, and then re-cultured in M9 medium for 30 minutes. Next, we diluted and stained the cells using 200 μM DFHBI-1T for 20 minutes, followed by drying them on a glass slide and covering them with a gel pad made of low melt agarose (C).

CRISPR *in vitro* digestion and ssDNAse assay

We used the above co-transcription approach to synthesize toehold-sgRNA and trigger RNAs, using DNA templates with a same total concentration of 10 nM in the in vitro transcription (IVT) mixture. After a 4-hour incubation at 37°C, the transcribed products were subjected to a 10x dilution under the same buffer conditions. Subsequently, we combined these transcripts with a digestion mix for *in vitro* digestion, using 30 nM EnGen® Lba Cas12a (NEB) and 30 nM target DNA at 37°C for 30 minutes. The resulting digestion products were assessed through agarose electrophoresis, with undigested target DNA serving as the negative control. To mitigate the influence of electroosmotic flow, the digestion buffer was added into the DNA ladder (1 Kb, NEB) to maintain consistent ion conditions. For the ssDNAase assay, an additional 100 nM FQ-labeled (FAM-TTATT-BHQ1, IDT) single-stranded DNA probe was introduced to the digestion mix. Fluorescence measurements were performed at 37 °C using a microplate reader for 1 hour. The fluorescence values were normalized by subtracting the blank signal and plotted against time.

# Outlook

In the preceding chapters, we delved into the application of SMFS to investigate the intricacies of branch migration equilibrium under specific force conditions. This scenario closely mirrors various RNA-protein interactions within cellular environments and bears considerable significance. As previously explored, our discoveries hold the potential to make substantial contributions to the design and fine-tuning of dynamic RNA riboregulators, as discussed in chapter I. In future research, our experimental setup can be utilized to investigate the kinetics of strand displacement under conditions resembling RNA-protein interactions, specifically focusing on the intricate coupling between ribosome binding and the invading process during translation initiation (Fig. 19 a). As previously discussed regarding the issue of leaky expression in the original toehold switch in Chapter II, it was observed that ribosomes could interact with the loop region, potentially leading to the unzipping of the toehold hairpin even in the absence of a trigger strand. Our dumbbell assay is designed to effectively detect both the binding and unzipping of the toehold hairpin, particularly under low force conditions. Furthermore, our experimental setup offers the possibility to investigate the kinetics of branch migration under the influence of ribosome binding and invasion via employing the same mismatch settings on the trigger strand (Fig. 19 b).

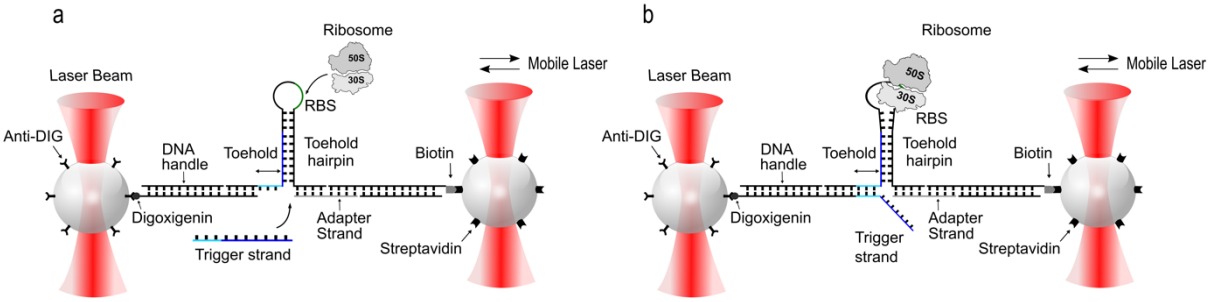


Figure 19. Dumbbell assay of toehold switch with ribosome. a), a dual-beam optical trap setup is used to measure force-dependent hairpin opening promoted by a trigger strand in a toehold-mediated strand displacement (TMSD) process. The experimental setup utilizes a dual-beam optical trap system to quantify force-induced hairpin opening resulting from the interaction of a trigger strand and ribosomes. RNA hairpin molecules are tethered between two beads using 500 base pairs long DNA handles as indicated (see legend for details). Ribosomes are introduced to initiate the translation process from the toehold switch. b), the ribosome selectively binds to the RBS located within the loop region and subsequently initiates the unwinding of the toehold hairpin, intricately connected with TMSD process.

In the upcoming chapters concerning toehold riboregulators and RNA switches, we have designed these riboregulatory systems based on the calculation of free energy for each substructure during TMSD and secondary structure rearrangement. We adhered to the design rule ensuring that $\Delta G_{TB} < \Delta G_{TH} < \Delta G_{AAH} < \Delta G_{AH}$ (TB: trigger-bound region, TH: toehold hairpin,

AAH: anti-anti hairpin, AH: anti-hairpin). This criterion was utilized to guarantee that the free energy difference (ΔΔG) was sufficiently large to facilitate the secondary structure rearrangement between different states. This design concept has some limitations, particularly when predicting *in vivo* switching behaviors, especially under varying ion conditions and interactions with proteins or small metabolites. For instance, in Chapter IV, we relied on the assumption of free energy for the ligand-bound aptamer calculated through NUPACK, which involved adding additional base-pairs to the P1 stem. However, this approach may not offer a comprehensive understanding of the variations in free energy among different riboswitch and RNA switch conformations. An alternative approach has been studied using the previously developed RBS[140] and riboswitch calculators[182] to developed rational designed riboswitch riboregulators. Vezeau et al[188]. Initially characterized the binding free energies of several RNA aptamers to their corresponding small metabolites or protein and incorporated these values into their energy model for riboswitch design, which also accounts for the ribosome binding free energy in the ON state. This design concept takes into account the potential effects of proteins and small metabolites, offering a more comprehensive approach. This design concept can be employed in future endeavors to refine toehold switch design, with a particular emphasis on accurately calculating the free energy associated with ribosome invasion.

To enhance the design efficiency and accuracy of synthetic riboregulators and RNA switches in the future, it would be ideal to incorporate a combination of direct evolutionary approaches, high-throughput screening, and machine learning models. One of the primary advantages of employing RNA aptamers and TMSD for RNA regulation, as opposed to protein-based designs, is their sequence programmability. This feature enables a relatively straightforward and rapid design-build-test-learn (DBTL) cycle (Fig. 19).

Initially, in the 'design' step, powerful software tools like NUPACK[97,189], ViennaRNA[152], Mfold[190] other packages such as CONTRAfold[191] and EternaFold play a pivotal role. These tools aid in a preliminary predicting of nucleic acid structures, thermodynamics, and the strength of ribosome binding (RBS calculator)[140]. Additionally, tools such as Multistrand[192] and OxDNA[193] enable the analysis of strand displacement kinetics at various levels of coarse graining. Moving on to the 'build' step, oligo sequences are synthesized and assembled into gene templates[194]. These templates can be further transformed into plasmids using standard cloning techniques. In the 'test' step, genetic switches and circuits undergo evaluation in cell-free, bacterial, or eukaryotic expression systems. Performance assessment often involves monitoring the expression of fluorescent proteins, aptamers, or enzymatic processes. This is achieved through the utilization of laboratory equipment like microplate readers, flow cytometry devices, or droplet microfluidics setups. Typically, the relationship between sequence and function in RNA regulation proves to be more intricate than initially envisaged through software-assisted 'rational design'. In such scenarios, the exploration of alternative variants may become imperative. Fortunately, it is relatively uncomplicated to create sequence libraries and systematically evaluate their performance through high-throughput techniques. This allows for the identification of successful

library members by employing sequencing methods. Finally, leveraging the wealth of results and data encompassing expression profiles and their associated sequences or secondary structures, in the ultimate 'learn' stage, we can enhance design principles and integrate them into the DBTL cycle. Through the comprehensive characterization of a library exceeding 90,000 toehold switch riboregulators, a deep neural network approaches were recently used to predict the performance of toehold switches These approaches also yield valuable biological insights into notable sequence and structural attributes[35,65].

Lastly, numerous parameters related to the expression profiles of riboregulators and their corresponding sequences and secondary structures remain unexplored. These factors, including cellular ion conditions, cell growth conditions, and interactions with cellular proteins and RNAs, could impact the accuracy of the learning process. For instance, as previously mentioned, the force biased TMSD influenced by ribosome binding has yet to be considered and incorporated into the learning model.

Furthermore, regarding with recently developed protein design strategies[195], applying analogous methods to RNA design[196] can expand the potential applications of RNA switches and riboregulators within the diagnose and gene therapy. This approach enables the precise engineering of desired and functional tertiary structures in RNA molecules, enhancing their versatility and efficacy in pharmaceutical applications. Also, by further advancing our understanding of RNA behaviors in cellular contexts and integrating these parameters, it is promising to achieve more effective and dependable riboregulators in the future.

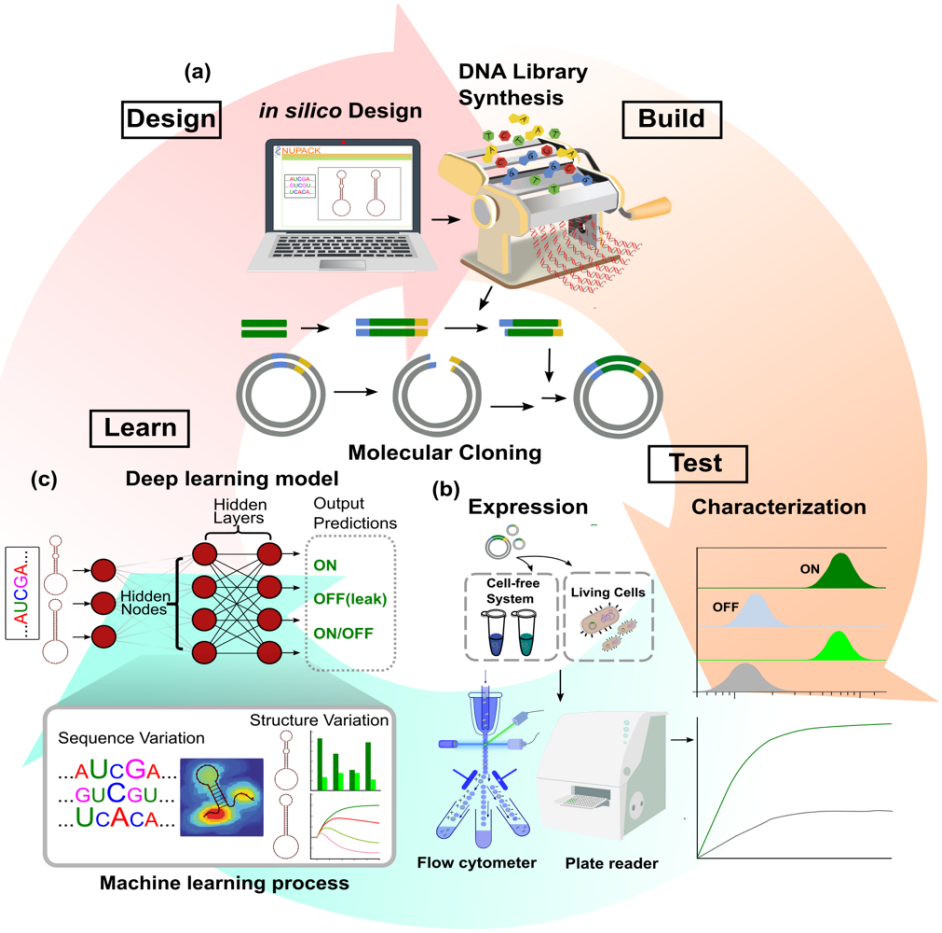


Figure 20: Development of RNA-based gene regulatory switches through a design-build-test-learn cycle (DBTL cycle). a), the process initiates with RNA structure design, guided by advanced computational tools. Subsequently, these designed sequences are synthesized efficiently using contemporary oligo synthesis and gene assembly techniques (Build). b), following this, the regulatory functionality of the RNA components is comprehensively characterized in high-throughput experiments conducted across various platforms, including flow cytometry, plate readers, and sequencing, both in *in vivo* and in cell-free expression systems (Test). c), the insights gathered from these experiments are then harnessed to train machine-learning models (Learn), facilitating further refinement and enhancement of computational design. It's worth noting that previous studies have successfully demonstrated the utilization of machine-learning pipelines for toehold switches. The Learn panel, as depicted here, is adapted from Angenent-Mari et al., showcasing the integration of machine-learning techniques into the design process, ultimately enabling the superior performance and optimization of RNA-based gene regulatory switches.

# Appendix

The values in brackets next to the mean value represent the standard error of the mean (SEM). All errors indicate 1σ. If the first digit of the error is one or two, then two significant digits of the error value are provided; otherwise, only one significant digit is indicated.

Table S1: Unfolded nucleotides of DNA and RNA toehold hairpin without trigger strand from constant velocity (CV) pulling and passive mode (PM) experiments (cf. Fig. 5).

| Nucleic acid | Measurement mode | Intermediate states (unfolded nucleotides) | | | | |
|---|---|---|---|---|---|---|
| | | Fol | I1 | I2 | I3 | Unf |
| DNA | CV | 2 | 18.1 (0.7) | 50 (4) | 60 (4) | 110 |
| | PM | 2 | 17.373 (0.006) | 47.21 (0.13) | 58.60 (0.22) | 110 |
| RNA | CV | 2 | 19.0 (0.9) | 52 (4) | 62 (5) | 110 |
| | PM | 2 | 17.7 (0.6) | 49.0 (1.2) | 59.6 (2.8) | 110 |

For the schematics, we calculated the number of unfolded base pairs as a weighted average of CV and PM data, then divided by two, rounding to the nearest whole number. This yielded 9, 24, and 29 unfolded base pairs for I1, I2, and I3 in the case of DNA, and 9, 25, and 30 for RNA. To determine the number of unfolded nucleotides, we employed the contour length calculation method described in Chapter I. It's worth noting that we assumed the first base pair to be already open due to fraying[45], which accounted for 2 unfolded nucleotides in the folded state, as shown in Table S1. The total unfolded contour length closely matched an unfolding of 108 nucleotides, considering a conversion rate of 0.6 nm/nt for ssRNA and 0.59 nm/nt for ssDNA, while also factoring in fraying and the nucleic acid's diameter.

Table S2: Free energies of intermediates compared to different predicting packages.

| Nucleic acid | $-\Delta G_0$ between intermediate states [$k_BT$] at 25 °C | | | |
|---|---|---|---|---|
| | Fol <-> I1 | I1 <-> I2 | I2 <-> I3 | I3 <-> Unf |
| DNA (experiment) | 15.236 (0.014) | 30.777 (0.016) | 10.798 (0.025) | 35.96 (0.21) |
| DNA (NUPACK) | 22.2 | 38.9 | 10.6 | 47.8 |
| DNA (mFold) | 22.1 | 37.8 | 11.4 | 47.6 |
| DNA (RNAstructure) | 21.9 | 37.6 | 12.1 | 46.9 |
| RNA (experiment) | 24.075 (0.024) | 50.90 (0.09) | 15.06 (0.23) | N.A. |
| RNA (NUPACK) | 30.9 | 64.5 | 15.4 | 66.3 |
| RNA (RNAstructure) | 31.5 | 59.5 | 17.4 | 70.5 |
| RNA (RNAsoft BLstar) | 26.8 | 51.6 | 15 | 58.9 |
| RNA (Vienna RNA at 37**°C**) | 26.8 | 51.8 | 15.2 | 58.9 |

| Nucleic acid | $-\Delta G_0$ between intermediate states [$k_BT$] at 25 °C | | | |
|---|---|---|---|---|
| | FoI <-> I1 | I1 <-> I2 | I2 <-> I3 | I3 <-> Unf |
| DNA (experiment) | 15.236 (0.014) | 30.777 (0.016) | 10.798 (0.025) | 35.96 (0.21) |
| DNA (NUPACK) | 22.2 | 38.9 | 10.6 | 47.8 |
| DNA (mFold) | 22.1 | 37.8 | 11.4 | 47.6 |
| DNA (RNAstructure) | 21.9 | 37.6 | 12.1 | 46.9 |
| RNA (experiment) | 24.075 (0.024) | 50.90 (0.09) | 15.06 (0.23) | N.A. |
| RNA (NUPACK) | 30.9 | 64.5 | 15.4 | 66.3 |
| RNA (RNAstructure) | 31.5 | 59.5 | 17.4 | 70.5 |
| RNA (RNAsoft BLstar) | 26.8 | 51.6 | 15 | 58.9 |
| RNA (Vienna RNA at 37**°C**) | 26.8 | 51.8 | 15.2 | 58.9 |

The systematic deviations observed in our results, compared to the predicted values from NUPACK, can be attributed to systematic errors in the force calibration process. Specifically, we estimated that approximately 10% of the deviation can be attributed to an underestimation of the viscosity of the buffer and scavenging system mixture. To validate this hypothesis, we conducted measurements on a control hairpin using a custom-built optical tweezer setup[197], which employs an advanced calibration method[198]. This control experiment yielded free energy values approximately 20% higher than our initial measurements and was within a 1% margin of the predicted NUPACK free energy. This advanced calibration approach eliminates the need for assumptions regarding viscosity and bead radius, thus enabling a more precise force calibration. It provides a plausible explanation for the observed deviations in our data. It's worth noting that we couldn't obtain a value for the transition between I3 and Unf since this transition occurred too slowly to be observed multiple times within the time frame of our experiment.

Table S3: Opened contour length in TMSD processes.

| Hairpin + trigger system | Intermediate states (opened contour length [nm]) | | | | | | |
|---|---|---|---|---|---|---|---|
| | TB | IM | IM (theo. 18 bm steps) | FI | FI (theo. 36 bm steps) | FU | FU (theo. 36 bm steps + 38 nts unfolding) |
| DD | 0 | N.A. | N.A. | 32.6 (0.3) | 33.5 | 53.42 (0.24) | 54.9 |
| RR | 0 | N.A. | N.A. | 32.5 (0.3) | 31.7 | 53.5 (0.4) | 53.3 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| DDc1 | 0 | 16.0 (0.4) | 16.7 | 32.20 (0.26) | 33.5 | 53.09 (0.19) | 54.9 |
| DDc2 | 0 | 16.1 (0.3) | 16.7 | 32.4 (0.6) | 33.5 | 53.1 (0.6) | 54.9 |
| RRc1 | 0 | 14.71 (0.16) | 15.8 | 28.81 (0.08) | 31.7 | 48.83 (0.06) | 53.3 |

In this section, we present the additional contour length increased due to branch migration. The parameters required for calculating the theoretical values corresponding to a specified number of branch migration steps (bm steps) and/or unfolded nucleotides can be found in Chapter I. The mean values obtained from our fits closely align with the theoretical values, with deviations of less than 10%. It's important to note that we are only considering statistical errors in this comparison. Slight systematic deviations, resulting in slightly lower calculated means, can be attributed to several factors. Firstly, possible fraying of one base pair at the branch migration junction could contribute to these deviations. Secondly, the drift in the position of the laser beams in the optical tweezer's setup may also play a role. Lastly, any assumptions made regarding fixed parameter values such as $p_{ssRNA}$, $p_{ssDNA}$, $L_{ssRNA/nt}$, $L_{dsRNA/bp}$, $L_{ssDNA/nt}$, and $L_{dsDNA/bp}$ can contribute to these systematic deviations.

Table S4: Opened contour length in TMSD processes with branch migration intermediates (BI).

| **Hairpin + trigger system** | **Intermediate states (opened contour length [nm])** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | TB | BI | IM | IM (theo. 18 bm steps) | FI | FI (theo. 36 bm steps) | FU | FU (theo. 36 bm steps +38 nts unfolding) |
| RRc2 | 0 | 11.97 (0.25) | 15.85 (0.17) | 15.8 | 30.3 (0.3) | 31.7 | 50.2 (0.4) | 53.3 |
| | TB | IM | BI | N.A. | FI | FI (theo. 36 bm steps) | FU | FU (theo. 36 bm steps +38 nts unfolding) |
| RRp2 | 0 | 0 | 5.6 (0.4) | N.A. | 29.31 (0.08) | 31.7 | 49.58 (0.18) | 53.3 |

In summary, our findings exhibit a notable overall agreement with theoretical values, despite minor systematic deviations towards lower values, which can be attributed to factors mentioned earlier (as explained in Table S3). The analysis of RRc2 reveals that there are approximately 13.6 branch migration steps between the TB and BI states, starting from the TB state. For clarity in our schematic (see Fig. S3), we rounded this value to 14.

Regarding RRp2, the contour length of the branch migration intermediate BI corresponds to approximately 9.3 branch migration steps. Utilizing an alternative method, where constant velocity traces of RRp2 were contour length transformed[72] and analyzed via histograms fitted with

a double Gaussian, we observed a prominent intermediate with approximately 7.91 (0.09) branch migration steps and a less populated intermediate with roughly 14.62 (0.29) branch migration steps. It's important to note that the value obtained from the constant velocity fits can be considered an average of these two intermediates, as it was challenging to distinguish between them using this particular analysis method. Additionally, the opened contour length of the IM state of RRp2 is 0, as the mismatches are situated directly after the toehold domain of the trigger strand. Consequently, states TB and IM are essentially identical for RRp2.

Table S5: Extrapolated rates and free energies at zero load for DDc1, DDc2 and RRc1. The contour length distances to the transition state $\Delta L_{C,IM-TS}$ and $\Delta L_{C,TS-FI}$ are determined by fitting it to a model explained in Chapter I .

| **Hairpin + trigger system** | **$\Delta L_{C,IM-TS}$ [nm]** | **$\Delta L_{C,TS-FI}$ [nm]** | **$k_0^{inv}$ [1/s]** | **$k_0^{re}$ [1/s]** | **$-\Delta G_0$ [$k_BT$]** | **$-\Delta G_0$ (Nupack) [$k_BT$]** |
|---|---|---|---|---|---|---|
| DDc1 | 8.4 (0.3) | 8.76 (0.14) | 1.75 (0.16) | 780 (30) | 6.56 (0.11) | 7.94 |
| DDc2 | 8.6 (0.4) | 8.7 (0.4) | 0.052 (0.008) | 1480 (220) | 9.76 (0.10) | 10.01 |
| RRc1 | 7.0 (0.9) | 8.4 (0.9) | 0.009 (0.003) | 230 (90) | 8.64 (0.14) | 11.8 |

In the analysis of contour lengths and rates, a weighted average approach was employed, taking into account multiple molecules with varying numbers of constant distance steps and, consequently, differing individual errors. Notably, the contour length differences observed from IM to FI are in excellent agreement, showing less than a 5% deviation from the theoretically expected values, assuming 18 branch migration steps from IM to FI (refer to values in Table S4 under "IM (theo. 18 bm steps)"). The position of the transition state is relatively symmetric between IM and FI, with only a slight asymmetry observed for RRc1 (7.0 vs. 8.4). Furthermore, it's worth noting that these values fall within 18% of the predictions made by the nearest neighbor model NUPACK for DNA mismatches and 27% for RNA mismatches, under comparable salt and temperature conditions. However, it's essential to acknowledge that the systematic deviations of our experimental results towards lower values, compared to the predictions, are likely attributed to systematic errors inherent in our force calibration process.

Table S6: Opened contour length in DNA-RNA hybrid TMSD.

| **Hairpin + trigger system** | **Intermediate states (opened contour length [nm])** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | TB | RD1 | RD2 | RD3 | FI | FI (theo. 36 bm steps) | FU | FU (theo. 36 bm steps + 38 nts unfolding) |
| RD | 0 | 6.71 (0.17) | 14.82 (0.24) | 21.3 (0.3) | 28.6 (0.4) | 32.4 | 49.67 (0.03) | 54 |

The values for the toehold-bound (TB), RD1, RD2, RD3, and fully invaded (FI) states were primarily obtained from passive mode experiments. It's worth noting that distinguishing between

the RD2 and RD3 states in constant velocity experiments posed challenges, hence the preference for passive mode data. The FU state, representing the final unfolding step of the remaining hairpin after full invasion, was determined by averaging the FI state value with the last unfolding step recorded during constant velocity experiments.The observed systematic deviations, which consistently showed lower values (~10% lower) compared to the theoretical predictions, can be attributed to the factors discussed in Tables S3 and S4. To calculate the number of branch migration steps corresponding to the RD1, RD2, and RD3 intermediates, we used the following formula:

(Number of steps) = (Contour length of one state / Contour length of FI state) * Total number of branch migration steps

For example, for RD1:

(Number of steps for RD1) = (6.71 nm / 28.6 nm) * 36 bms = 8.45 bms

The resulting values for these three intermediates were then rounded to the nearest whole number and are depicted in Figure 5.

Table S7: Extrapolated rates and free energies at zero load for RNA-DNA hybrid TMSD. The contour length distances to the transition state are determined by fitting it to a model explained in Chapter I.

| Hairpin + trigger system | Transition (i <-> j) | $\Delta L_{c,i\text{-}TS}$ [nm] | $\Delta L_{c,TS\text{-}j}$ [nm] | $k_0^{ij,inv}$ [1/s] | $k_0^{ji,re}$ [1/s] | $-\Delta G_0$ [$k_BT$] |
|---|---|---|---|---|---|---|
| RD | TB <-> RD1 | 4.2 (1.6) | 3.8 (1.8) | 1.4 (1.2) | 34000 (36000) | 8.34 (0.15) |
| RD | RD1 <-> RD2 | 4.7 (5.7) | 3.7 (9.3) | 0.05 (0.15) | 15000 (82000) | 12.9 (0.4) |
| RD | RD2 <-> RD3 | 4.0 (5.3) | 3.3 (8.1) | 1.6 (4.7) | 50000 (23000) | 9.2 (0.5) |
| RD | RD3 <-> FI | 3.3 (10) | 3.6 (3.5) | 5 (29) | 33000 (66000) | 9.54 (0.26) |

The high errors (as indicated in brackets) are a result of the extensive extrapolation required, spanning from 10 pN down to 0 pN. This extrapolation relies on a limited data range, typically falling within the small range of approximately 8-12 pN. Additionally, the statistical power is relatively weak due to the limited dataset, with only three molecules for the TB-RD1 transition and just one molecule for the other transitions. Analyzing the contour length differences between states *i* and *j*, we observe that, similar to the values obtained from contour length-transformed passive mode traces, the most significant contour length difference exists between state RD1 and RD2. However, it's essential to note that these values fall within a similar range as those extracted from passive mode experiments. Nevertheless, due to the high error values associated with these extrapolations, these results may be considered less statistically significant compared to the values extracted from passive mode experiments.

Table S8: Results of RBS calculator of mCherry translation

| **Predicted mCherry mRNA translation initiation rate based RBS sequence_1 (5'-aggagc-3')** |
|---|
| start_position [CDS Start Position (nt)],"ORF_number [Open Reading Frame]","tir [Translation Initiation Rate (au)]","dG_total [Total Gibbs Free Energy Change (kcal/mol)]","dG_mRNA_rRNA [Gibbs Free Energy Change of mRNA:rRNA complex (kcal/mol)]","dG_spacing [Gibbs Free Energy Penalty for Non-optimal Spacing (kcal/mol)]","dG_stacking [Gibbs Free Energy Change of Stacked |

| |
|---|
| Nucleotides in Spacer Region (kcal/mol)]","dG_standby [Gibbs Free Energy Penalty for Ribosome Binding to Standby Site (kcal/mol)]","dG_start [Gibbs Free Energy Change of mRNA:tRNA complex (kcal/mol)]","dG_mRNA [Gibbs Free Energy Change of mRNA folded complex (kcal/mol)]","warnings [Warnings and Errors Issued]" |
| 94,1,271.6098003539922,3.3622490386842117,-28.241351876843133,0.288,0,0,-2.76,-34.290000915527344, |
| **Predicted mCherry mRNA translation initiation rate based RBS sequence_1 (5'-aggaga-3')** |
| start_position [CDS Start Position (nt)],"ORF_number [Open Reading Frame]","tir [Translation Initiation Rate (au)]","dG_total [Total Gibbs Free Energy Change (kcal/mol)]","dG_mRNA_rRNA [Gibbs Free Energy Change of mRNA:rRNA complex (kcal/mol)]","dG_spacing [Gibbs Free Energy Penalty for Non-optimal Spacing (kcal/mol)]","dG_stacking [Gibbs Free Energy Change of Stacked Nucleotides in Spacer Region (kcal/mol)]","dG_standby [Gibbs Free Energy Penalty for Ribosome Binding to Standby Site (kcal/mol)]","dG_start [Gibbs Free Energy Change of mRNA:tRNA complex (kcal/mol)]","dG_mRNA [Gibbs Free Energy Change of mRNA folded complex (kcal/mol)]","warnings [Warnings and Errors Issued]" |
| 94,1,3596.173203489391,-2.377752639782585,-28.241351876843133,0.288,0,0,-2.76,-28.549999237060547, |
| Note: the start position represents the position of the start codon (AUG) of the given ORF. And the Translational initiation Rate represents the predicted translation rate based on the given upstream RBS sequence |

Table S9: DNA sequence of toehold riboregulators

**A. Sequence and information for the set of toehold riboregulators**

| Toehold-riboregulators | Switch sequence | Trigger sequence | Reporter gene | Switch Plasmid | Plasmid Origin/Resistance |
|---|---|---|---|---|---|
| ld translational activator_st | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCGGGGAATAACGTGAAGGCAGGGAAAAACGAGATGCTTC TTCACGTTATGGCAGGAGCAACTAGA***ATG*** | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mCherry | pet28b-modified | pBR322/kanamycin |
| ld translational activator_st | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCGGGGAGATTGAAATAGCAGGGATAAACGTTTCAATCTC CTCTGTAATGTAAGAAGTGAGAGGAGAAGTAGTAGA***ATG*** | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mcherry | pet28b-modified | pBR322/kanamycin |
| ld translational activator_st | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCCCTG CCAGGAGATGAGACAAGTGCAGGGATAAACGACTTGTCTAATCTGAGCT AAAAAGAGGAGAAGTACTAG***ATG*** | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mcherry | pet28b-modified | pBR322/kanamycin |
| ld translational activator_st | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCCCTG CCAGGAGGGGAAAAGTGCAGGGATAAACGACTTTTCTCCTCTGAGC TAAAAGAGGAGAAGTAGTAG***ATG*** | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mcherry | pet28b-modified | pBR322/kanamycin |
| ehold translational activator | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCGATGCTCCTTTTCGAGCAGGGATAAACGAGAGAAAAGG AGCACTAGATG | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mcherry | pet28b-modified | pBR322/kanamycin |
| ehold translational activator | TAATACGACTCACTATAGGATCTATTACTACTTACCATTGTCTT GCTCTGATTCTCCTTTTCGAAGAGCAAGACAATGGAGAAAAGG GAGAACAACAATG | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA AGAGCAAGACAATGGTAAGTAGTAATAGATATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mcherry | pet28b-modified | pBR322/kanamycin |
| ptional activator_stem1_int | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCACTCAGGGCTTTATTGCAGGGATAATACAAATAAAGCC CTGAGTTTAACCGCTCGGGGCTTTTTGCGTTTACTAGAGAGAC CACAACGATCCTAGAGAAGCAATAATTTGTTAACTTTAGAGGA GATACTAG***ATG*** | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mcherry | pet28b-modified | pBR322/kanamycin |
| ptional activator_stem2_int | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCCGCAAAAAGCCCCGAGCAGGGATAAACGTACAAATAA AGCCCTGAGTTTAACCGCTCGGGGCTTTTTGCGTTTACTAGAG AGACCACAACGATCCTAGAGAAGCAATAATTTGTTAACTTTAG AGGAGATACTAG***ATG*** | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mcherry | pet28b-modified | pBR322/kanamycin |
| al activator_stem1_ Rho-d | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCGTTAAAGGGCAAATTGGCAGGGATAAACGTACAATTTG CCGTTCTGTAGCCATCACCTTTTTTTACTAGAGAGACCACAAC GATCCTAGAGAAGCAATAATTTGTTAACTTTAGAGGAGATACT AG***ATG*** | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mcherry | pet28b-modified | pBR322/kanamycin |
| al activator_stem2_ Rho-d | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCTTTATAAAGGGCAAATTGGCAGGGAAAAACGAACAATT TGCCCTTCTGTAGCCATCACCTTTTTTTACTAGAGAGACCACAA CGATCCTAGAGAAGCAATAATTTGTTAACTTTAGAGGAGATAC TAG***ATG*** | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mcherry | pet28b-modified | pBR322/kanamycin |
| al activator_stem3_Rho-dep | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCTTTATAAAGGGCAAATTGGCAGGGAATAACGAGTATCC AATTTGCCCTTCTGTAGCCATCACCTTTTTTTACTAGAGAGACC ACAACGATCCTAGAGAAGCAATAATTTGTTAACTTTAGAGGAG ATACTAG***ATG*** | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mcherry | pet28b-modified | pBR322/kanamycin |
| tional repressor Rho-depe | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCGATGCTCCTTTTCGAGCAGGGATAAACCGAGAGAAAAGG GAGCACTAGAATGAATATCTTACATATATGTGTGACCTCAAAA TGGTTCAATATTGACAACAAAATTGTCGATCACCGCCCTTGATT TGCCCTTCTGTAGCCATCACCTTTTTTTACTAGATGATGATAAC AACGATCCTAGAGAAGCAATAATTTGTTAACTTTAGAGGAGAT ACTAG***ATG*** | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mcherry | pet28b-modified | pBR322/kanamycin |

## Table S10: DNA sequence of switchable FLAPs

**A. Sequence and information for the set of toehold riboregulators**

| Toehold-FLAP switch | Switch sequence | Trigger sequence | FLAP | Switch Plasmid | Plasmid Origin/Resistance |
|---|---|---|---|---|---|
| toehold spinach activator #1 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCACAGTCCATCATTCAGGCAGGGATAACTGAATGATGGA CCCGTCCTACTAGAGACGCGACCGAATGAAATGGTGAAGGA CGGGTCCAGCCGGCTGCACTGCGCAGCCGGCTTGTTGAGTA GAGTGTGAGCTCCGTAACTGGTCGCGTC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | spinach aptamer | pet28b-modified | pBR322/kanamycin |
| toehold spinach activator #2 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCACAAAGGATCATTCAGGCAGGGATAACTGAATGAGTCC TTCACCATACTAGAGACGCGACCGAATGAAATGGTGAAGGAC GGGTCCAGCCGGCTGCACTGCGCAGCCGGCTTGTTGAGTAG AGTGTGAGCTCCGTAACTGGTCGCGTC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | spinach aptamer | pet28b-modified | pBR322/kanamycin |
| toehold spinach activator #3 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCCCTG CACAGTCCATCATTCAGGCAGGGATAACTGAATGATGGACCCGTCC TACTAGAGACGCGACCGAATGAAATGGTGAAGGACGGGTCCAGCCG GCTGCACTGCGCAGCCGGCTTGTTGAGTAGAGTGTGAGCTCCGTA ACTGGTCGCGTC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | spinach aptamer | pet28b-modified | pBR322/kanamycin |
| toehold spinach activator #4 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCCCTG CACAGCCAGTCATTCAGGCAGGGATAACTGAATGACTCTACTCAACA ACCACTAGAGACGCGACCGAATGAAATGGTGAAGGACGGGTCCAGC CGGCTGCACTGCGCAGCCGGCTTGTTGAGTAGAGTGTGAGCTCCG TAACTGGTCGCGTC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | spinach aptamer | pet28b-modified | pBR322/kanamycin |
| toehold spinach activator #5 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCACACTCCGTTCATTCAGCAGGGATAAATGAATGAACGG AGCTCACACACTAGAGACGCGACCGAATGAAATGGTGAAGG ACGGGTCCAGCCGGCTGCACTGCGCAGCCGGCTTGTTGAGT AGAGTGTGAGCTCCGTAACTGGTCGCGTC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | spinach aptamer | pet28b-modified | pBR322/kanamycin |
| toehold spinach repressor #1 | TAATACGACTCACTATAGGATCTATTACTACTTACCATTGTCTT GCTCTCTCGTCGCGTCTCTAGCAGGGATAAACTAGAGACGCG ACCGAATGAAATGGTGAAGGACGGGTCCAGCCGGCTGCACT GCGCAGCCGGCTTGTTGAGTAGAGTGTGAGCTCCGTAACTGG TCGCGTC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | spinach aptamer | pet28b-modified | pBR322/kanamycin |
| toehold spinach repressor #2 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCTTATCTTCGGTCGCGTCGCAGGGATAAACTAGAGACGCG ACCGAATGAAATGGTGAAGGACGGGTCCAGCCGGCTGCACT GCGCAGCCGGCTTGTTGAGTAGAGTGTGAGCTCCGTAACTGG TCGCGTC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | spinach aptamer | pet28b-modified | pBR322/kanamycin |
| toehold spinach repressor #3 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCTAAACAACAACCCGTCCTGCAGGGATAAACTAGAGACG CGACCGAATGAAATGGTGAAGGACGGGTCCAGCCGGCTGCA CTGCGCAGCCGGCTTGTTGAGTAGAGTGTGAGCTCCGTAACT GGTCGCGTC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | spinach aptamer | pet28b-modified | pBR322/kanamycin |
| toehold spinach repressor #4 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCCACCCGTCCTTCACCGCAGGGATAAACTACTAGAGACG CGACCGAATGAAATGGTGAAGGACGGGTCCAGCCGGCTGCA CTGCGCAGCCGGCTTGTTGAGTAGAGTGTGAGCTCCGTAACT GGTCGCGTC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | spinach aptamer | pet28b-modified | pBR322/kanamycin |
| toehold spinach repressor #5 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCGTAGACGCGACCAGTGCAGGGATAAACTAGAGACGCG ACCGAATGAAATGGTGAAGGACGGGTCCAGCCGGCTGCACT GCGCAGCCGGCTTGTTGAGTAGAGTGTGAGCTCCGTAACTGG TCGCGTC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | spinach aptamer | pet28b-modified | pBR322/kanamycin |
| toehold mango activator #1 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCACTCGGCACGTACTCAGCAGGGATAAACGAATGAGTAC GTGCCACTAGAGGCACGTACGAAGGGACGGTGCGGAGAGG AGAGTACGTGC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mango aptamer | pet28b-modified | pBR322/kanamycin |
| toehold mango activator #2 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCTGAAGGCGAAGGGAGTGCAGGGATAAACGAATGACTC CCTTCGGTACTAGAGGCACGTACGAAGGGACGGTGCGGAGA GGAGAGTACGTGC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mango aptamer | pet28b-modified | pBR322/kanamycin |
| toehold mango activator #3 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCTGAAGACGGTGCGGGTGCAGGGATAAACGAATGACCC GCACCGTACTAGAGGCACGTACGAAGGGACGGTGCGGAGA GGAGAGTACGTGC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mango aptamer | pet28b-modified | pBR322/kanamycin |
| toehold mango activator #4 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCTGAAGAAGAGGAGAGTGCAGGGATAAACGAATGACTC TCCTCTCCACTAGAGGCACGTACGAAGGGACGGTGCGGAGA GGAGAGTACGTGC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mango aptamer | pet28b-modified | pBR322/kanamycin |
| toehold mango activator #5 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCTGAAGGAGAGTACGGTGCAGGGATAAACGAATGACCG TACTCTCACTAGAGGCACGTACGAAGGGACGGTGCGGAGAG GAGAGTACGTGC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mango aptamer | pet28b-modified | pBR322/kanamycin |
| toehold mango repressor #1 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCACATTAAGTGCCTGAAGCAGGGATAAACGATCTTCAGG CACGTACGAAGGGACGGTGCGGAGAGGAGAGTACGTGC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mango aptamer | pet28b-modified | pBR322/kanamycin |
| toehold mango repressor #2 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCACATTACCCTTCGTACGCAGGGATAAACGATCTTCAGG CACGTACGAAGGGACGGTGCGGAGAGGAGAGTACGTGC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mango aptamer | pet28b-modified | pBR322/kanamycin |
| toehold mango repressor #3 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCAGATTTCTCCGCACCGGCAGGGATAAACGATCTTCAGG CACGTACGAAGGGACGGTGCGGAGAGGAGAGTACGTGC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mango aptamer | pet28b-modified | pBR322/kanamycin |
| toehold mango repressor #4 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCAGATTCTCCTCTCCGCGCAGGGATAAACGATCTTCAGG CACGTACGAAGGGACGGTGCGGAGAGGAGAGTACGTGC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mango aptamer | pet28b-modified | pBR322/kanamycin |
| toehold mango repressor #5 | TAATACGACTCACTATAGGTCTTATCTTATCTATCTCGTTTATCC CTGCACATTAGCACGTACTCGCAGGGATAAACGATCTTCAGG CACGTACGAAGGGACGGTGCGGAGAGGAGAGTACGTGC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | mango aptamer | pet28b-modified | pBR322/kanamycin |
| toehold broccoli activator | ACATCGGGACATTCATGCAGGGATAAATGAATGTACCCGACC GTCTCACTAGAGTATGACGCGACCGAGACGGTCGGGTCCAG ATATTCGTATCTGTCGAGTAGAGTGTGGGCTCGGTCGCGTC | TAATACGACTCACTATAGGGACTGACTATTCTGTGCAATAGTCAGTAAA GCAGGGATAAACGAGATAGATAAGATAAGAATGGAACCATTAGCATAAC CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG | Broccoli aptamer (extended) | pet28b-modified | pBR322/kanamycin |

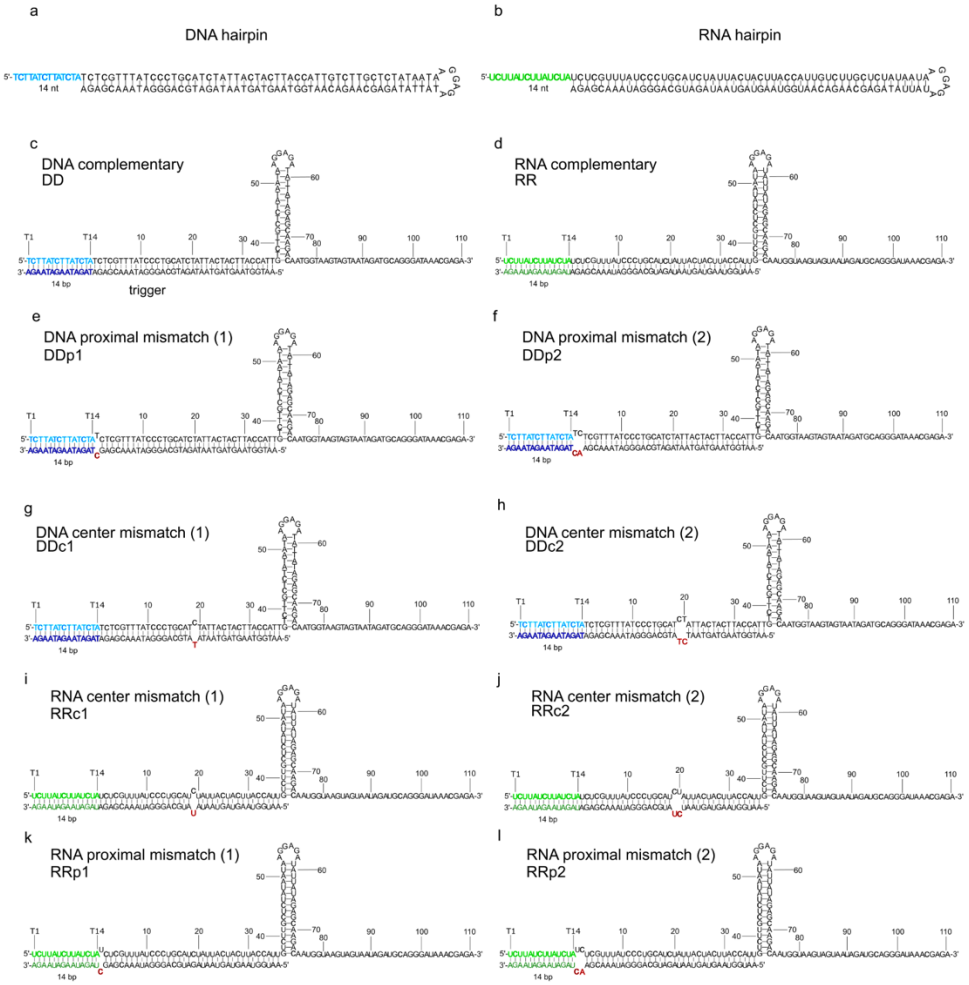


Figure S1. Schematic representation of the toehold hairpins employed in this study, along with their corresponding sequence details. a), b), both the DNA and RNA toehold hairpins are composed of 124 nucleotides, featuring a 14-nucleotide long toehold (T1-T14), a 52-base pair stem, and a 6-nucleotide loop. In specific cases, mismatches were introduced at different positions for investigation. c), d), the fully complementary trigger strand spans 50 nucleotides and invades the toehold hairpin up to position 36. e), k), 1 proximal mismatch scenarios - DDp1 and RRp1 (C→T, indicated in red) at position 1 following the toehold region. f), l), 2 proximal mismatches scenarios - DDp2 and RRp2 (AG → CA, shown in red) at positions 1 and 2, immediately after the toehold sequence (highlighted in light blue and green). g, i, 1 central mismatch scenarios - DDc1 and DDc2 (C→T, indicated in red) at positions 19. h), j), 2 central mismatches scenarios - DDc2 and RRc2 (GA → TC, shown in red) at positions 19 and 20.

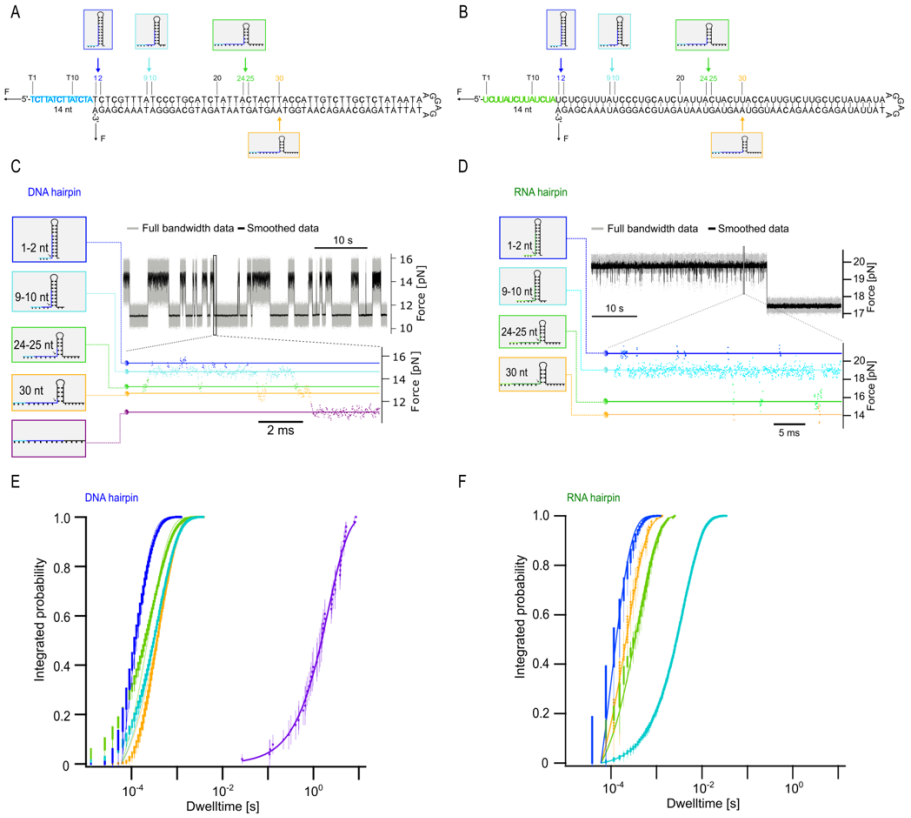


Figure S2. Intermediate state of toehold hairpins from passive mode. A), B), schematic depiction of estimated unzipping intermediate positions for both DNA and RNA toehold hairpins. Each intermediate state (from position 1 to 30) is distinguished by different colors. Displayed are force vs. time traces for both DNA and RNA toehold hairpins using passive mode measurements. This approach involves maintaining a constant distance in the laser traps while recording the applied force on the beads. Conformational transitions between intermediate states are evident as shifts between distinct force levels. Each color signifies a specific unzipping intermediate position on the hairpin, with purple indicating the fully opened state. E), F), Dwell time integrated probabilities of each intermediate state corresponds to the force vs. time traces.

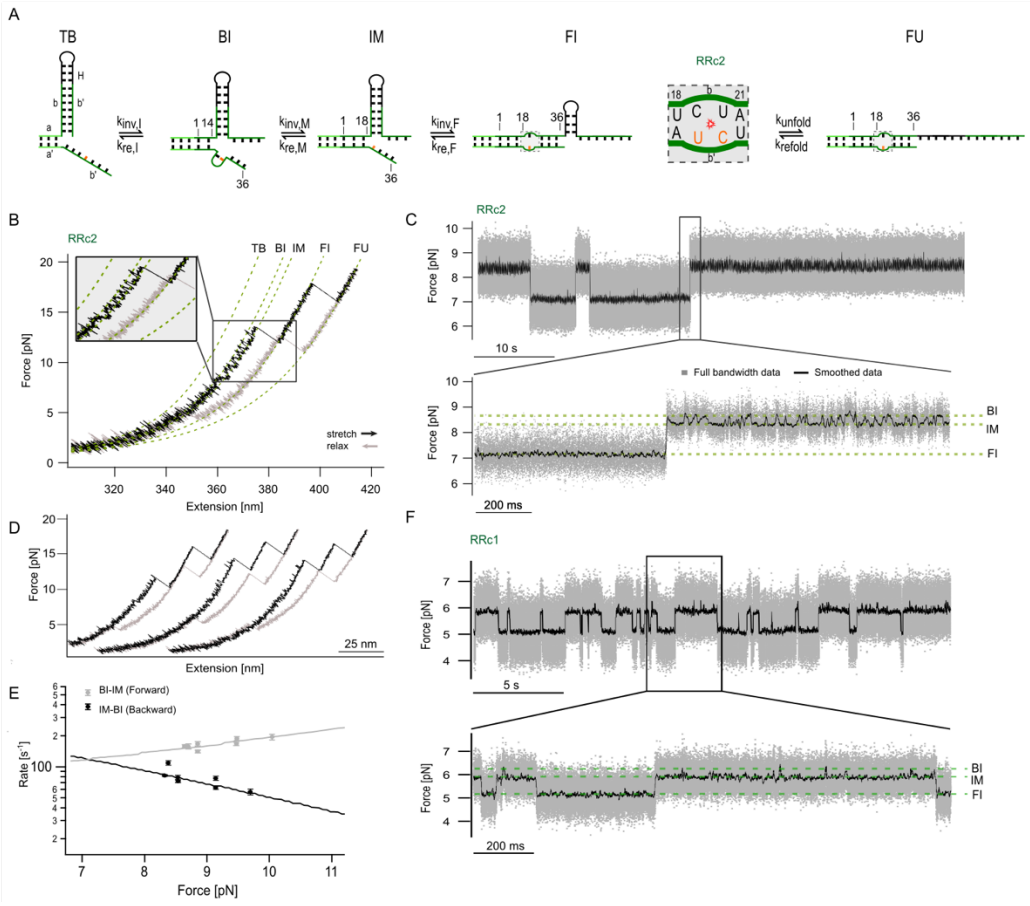


Figure S3. Dynamics of RNA toehold hairpin with RRc2 Trigger Strand. A), Illustration of the RNA toehold hairpin exhibiting 5 distinct states during the force-dependent folding and unfolding process triggered by the RRc2 strand: toehold bound (TB), Branch migration intermediate before mismatch at position 14 (BI), invasion until mismatch at position 18 (IM), fully invaded (FI), and fully unfolded (FU). The mismatched base on the trigger strand is highlighted in orange at position 19 and 20. B), Force-extension traces of the RNA toehold hairpin binding to trigger strands with RRc2. Stretch (black) and relax (grey) cycle using a constant pulling velocity of 0.2 μm/s. Each trace corresponds to different contour lengths, and intermediate states are indicated. The transition details between the BI and IM states, influenced by the mismatch, are shown in the Inset. C), Force-versus-time traces of the toehold hairpin with the trigger strand RRc2, maintaining a constant trap distance. Slow transition kinetics are observed and states BI, IM, and FI are labeled. D), Additional force-extension traces of the RNA toehold hairpin with the trigger strands with RRc2. E), extrapolated forward and backward transition rates between the BI and IM states. F), force versus time traces of toehold hairpin with trigger strand RRc1 while the trap keeps at a constant distance. State BI, IM and FI are labelled to represent the transition kinetics.

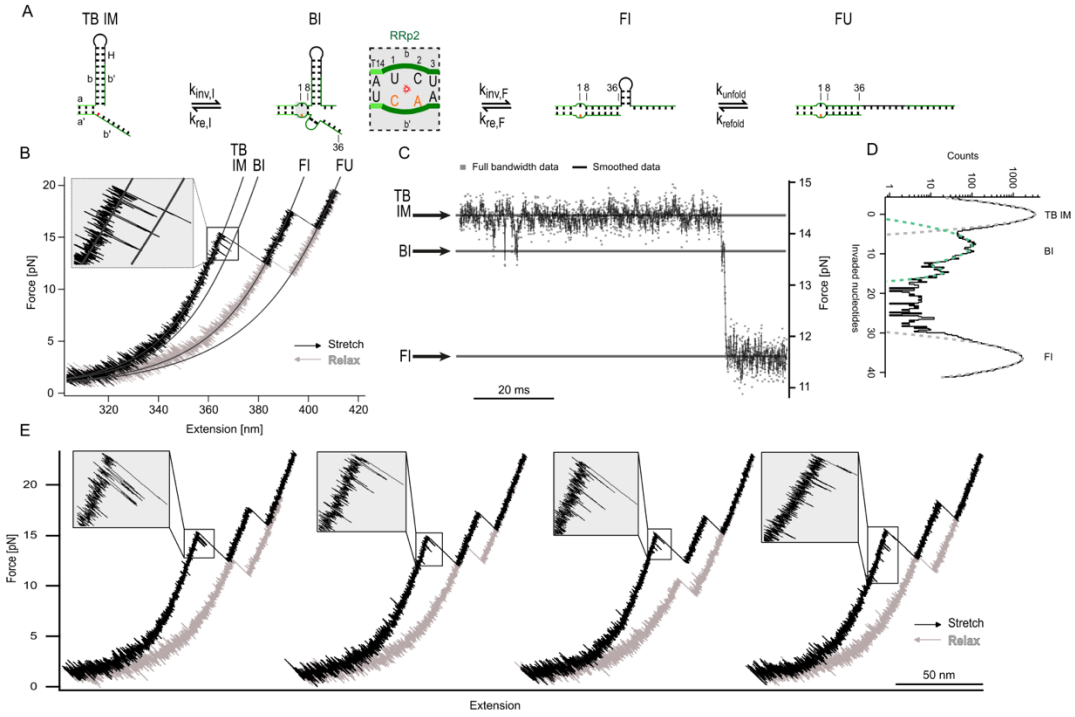


Figure S4. Dynamics of RNA toehold hairpin with RRp2 trigger strand. A), Illustration of the RNA toehold hairpin exhibiting 4 distinct states during the force-dependent folding and unfolding process triggered by the RRp2 strand: toehold bound, and invasion stop before mismatch (TB-IM), Branch migration intermediate slightly crossed mismatch at position 8 (BI), fully invaded (FI), and fully unfolded (FU). The mismatched base on the trigger strand is highlighted in orange at position 1 and 2. B), Force-extension traces of the RNA toehold hairpin binding to trigger strands with RRp2. Stretch (black) and relax (grey) cycle using a constant pulling velocity of 0.2 µm/s. Each trace corresponds to different contour lengths, and intermediate states are indicated. The transition details between the BI and TB-IM states, influenced by the mismatch, are shown in the Inset. C), Force-versus-time traces of the toehold hairpin binding the trigger strand RRp2 and strand invasion, maintaining a constant trap distance. Transition kinetics are observed and states TB-IM, FI are labeled. D), Double-gaussian fits of invaded nucleotides from Force-versus-time traces for each state (green and grey). E, Additional force-extension traces of the RNA toehold hairpin with trigger strands RRp2. The transition details between the BI and IM states, influenced by the mismatch, are shown in the Inset.

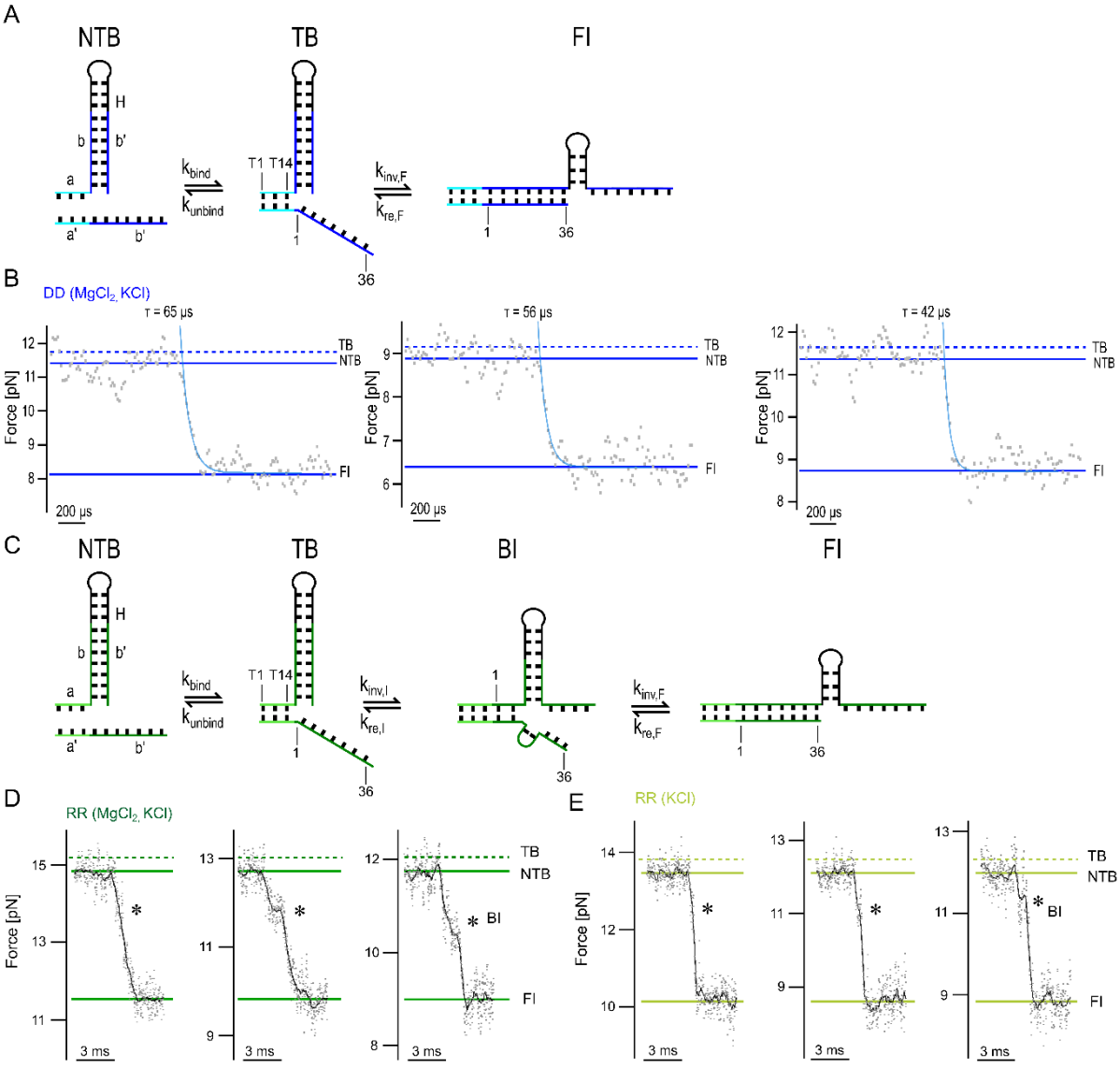


Figure S5. Toehold hairpins with fully complementary trigger. A), Schematics of the toehold binding and strand invasion of DNA toehold hairpin with fully complementary strand DD, including three different states: non-toehold bound (NTB), toehold bound (TB) and fully invaded (FI). B), Force-versus-time traces of the DNA toehold hairpin binding the trigger strand DD and strand invasion while the trap maintains at a constant distance under certain salt condition (20mM $MgCl_2$, 100mM, KCl). Each state at different force conditions have shown on the traces. Estimated TB state is marked with dash lines. The exponential fits show different decay times. C), Schematics of the toehold binding and strand invasion of RNA toehold hairpin with fully complementary strand RR, including four different states: non-toehold bound (NTB), toehold bound (TB), Branch migration intermediate (BI) and fully invaded (FI). D), Force-versus-time traces of the RNA toehold hairpin binding the trigger strand RR and strand invasion while the trap maintains at a constant distance under different salt condition (20mM $MgCl_2$, 100mM KCl, and 100mM KCl only). Each state from different force conditions have shown on the traces. Estimated TB state is marked with dash lines. BI states are shown on the traces with *.

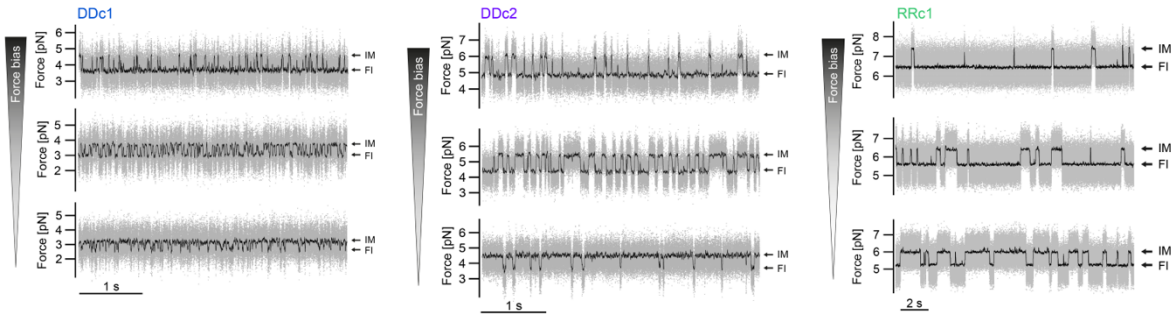


Figure S6. Force versus time traces of all DDc1, DDc2 and RRc1 complex captured at different constant trap distances. The force bias on the complex intensifies as the trap distance increases. Notably, the equilibrium between states IM and FI responds to the applied force, leading to a shift in transition towards state IM with increasing force. The states IM and FI are labeled in the traces to highlight the distinct transition kinetics observed.

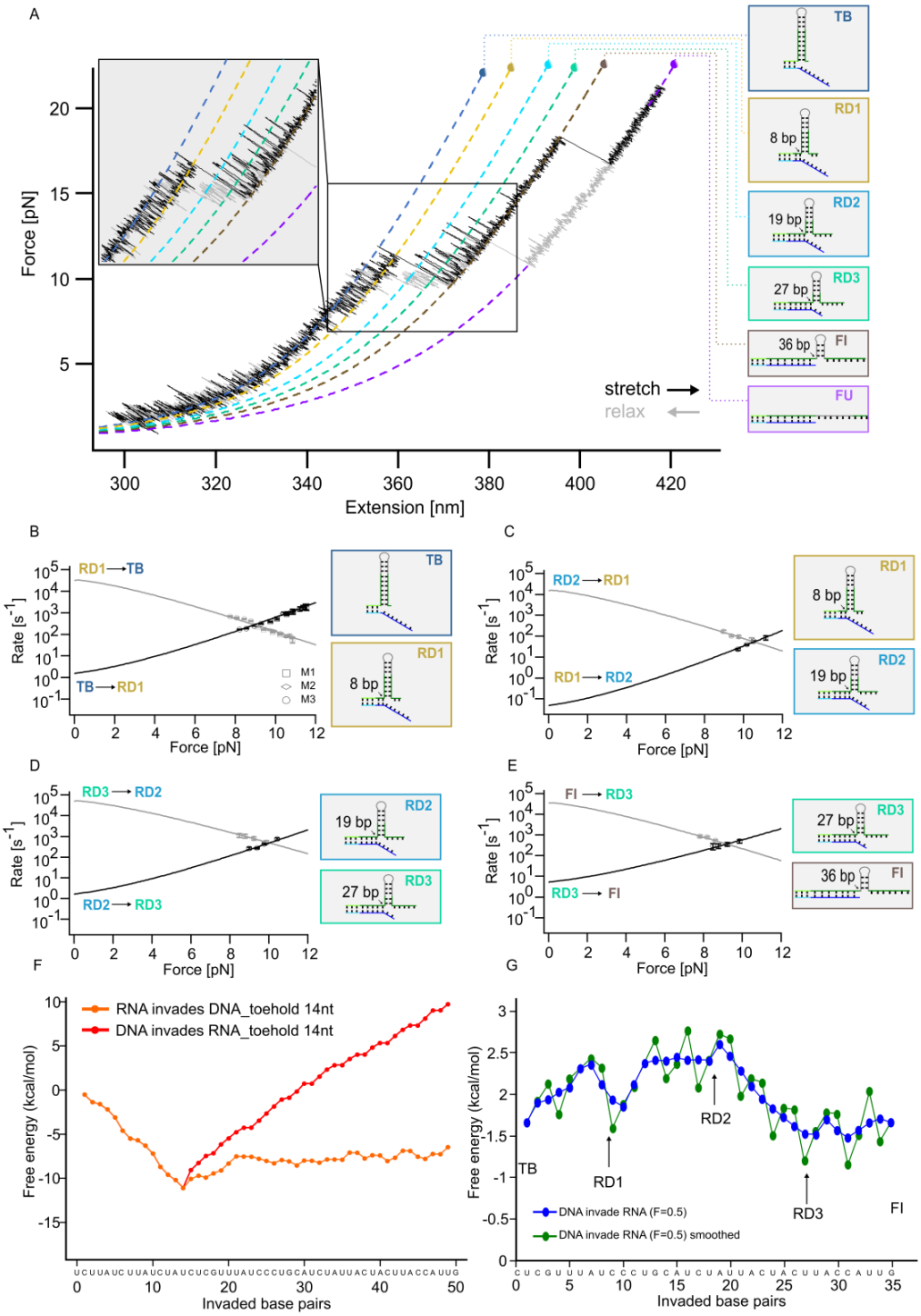


Figure S7. A), Force extension traces of RNA toehold hairpin binding to a fully complementary DNA trigger strand RD. Intermediate states are indicated by distinct colors for different contour length fits: Toehold bound (TB), RNA-DNA hybrid intermediates (RD1, RD2, RD3), fully invaded (FI), and fully unfolded (FU). Transition details between state TB and FI are illustrated in the inset. Numbers of nucleotides indicate the opened stem base-pairs. B), Force dependence of forward and backward transition rates between each two intermediate states. The dotted lines represent extrapolations of the data based on a theoretical model described in the Chapter I. F), Free-energy profiles for the three strand displacement systems of the branch migration region from our toehold hairpin sequence, calculated using the Nearest Neighbor model. G), Force corrected Free-energy landscape depicting the transition of the DNA invading RNA strand at $F_{avg\ 1/2}$ from TB to FI state, with potential intermediate states indicated along the corresponding sequence.

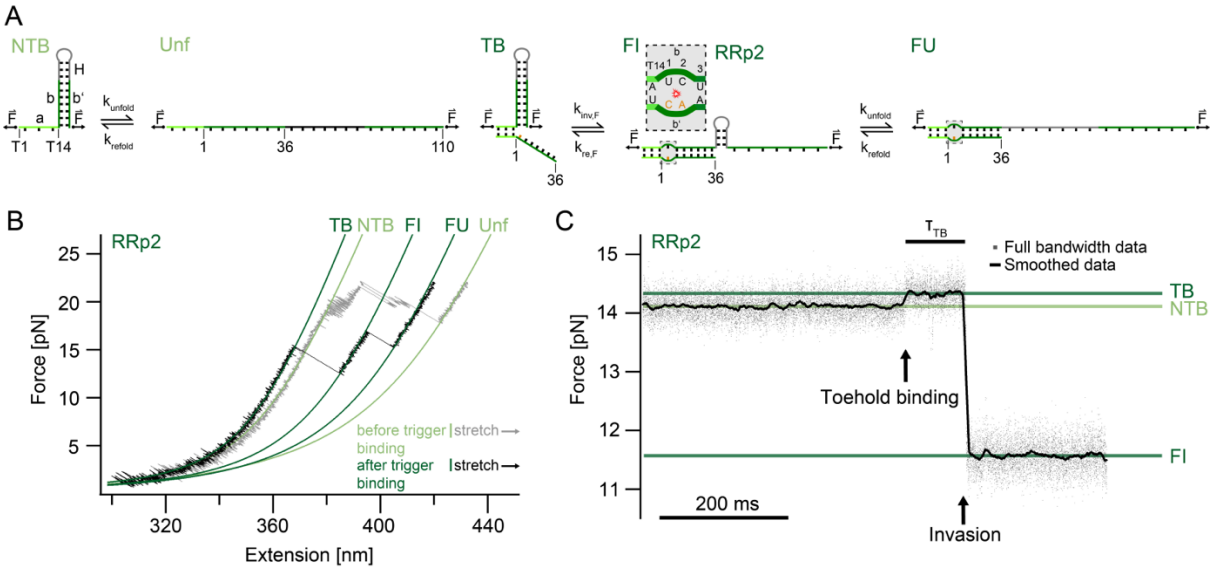


Figure S8. Observation of toehold bound (TB) state from trigger strand RRp2. A), Schematic depicting the change in contour length of single-stranded RNA due to the application of force. In the absence of a trigger strand, the single-stranded RNA is stretched as indicated by the difference in contour length from the double-stranded RNA configuration. The toehold hairpin transitions from a folded non-toehold bound (NTB) state to an unfolded state (Unf) through unzipping. Binding of the trigger strand (TB) reduces the contour length and results in an increase in force. B), Force-extension traces of the RNA toehold hairpin (light green) before and after binding to an RNA trigger RRp2 (green). Intermediate states are identified by distinct colors corresponding to different contour length fits. C), Force-versus-time traces of the RNA toehold hairpin binding to the trigger strand RRp2 and undergoing strand invasion, while the trap maintains a constant distance. Different states under various force conditions are highlighted on the traces. The stages of toehold binding and TB state dwell are indicated.

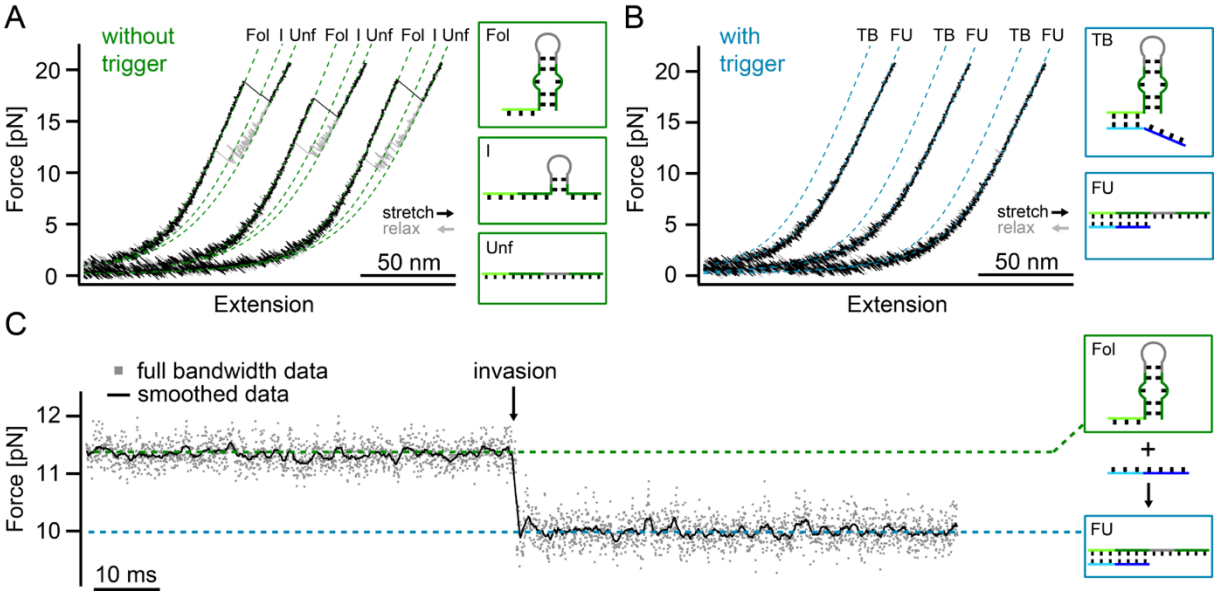


Figure S9. Toehold switch with bulges on the stem. A), original toehold switch hairpin folding and TMSD of the with fully complementary trigger (30nt). We firstly performed the active mode in the buffer channel. The passive mode trace shows the signals of trigger binding and partial invading until state STriA1 (smoothed trace (black), full bandwidth data (dark gray)).


Figure S10. A schematic illustration of a three-strand complex as represented in the oxDNA model. The incumbent (red) is bound to the substrate (blue), and the invading strand is shown in green. The constant applied force is applied on the nucleotides shown with black arrow.

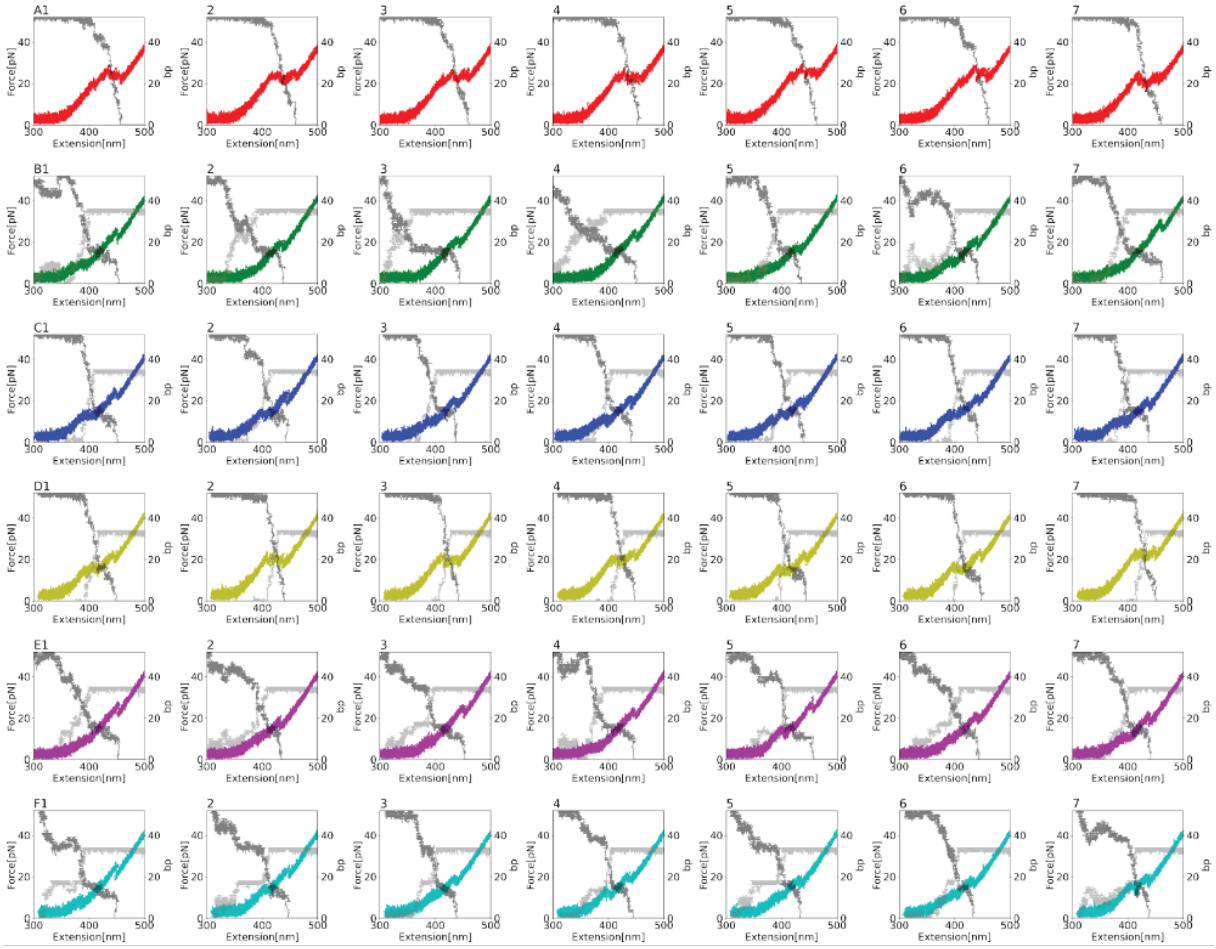


Figure S11. Individual trajectories used to compute the average plot in Figure 3. (Pulling the handles with 0.14 mm/s constant speed pulling) A), no trigger system B), fully complementary invader C), one basal mismatch for the invader. D), two basal mismatches. E), one mismatch at the center of the invader. F), two mismatches at the center of the invader. Black curves denote the number of base pairs in the stem of the hairpin. Grey curves denote the number of base pairs formed by the invader.

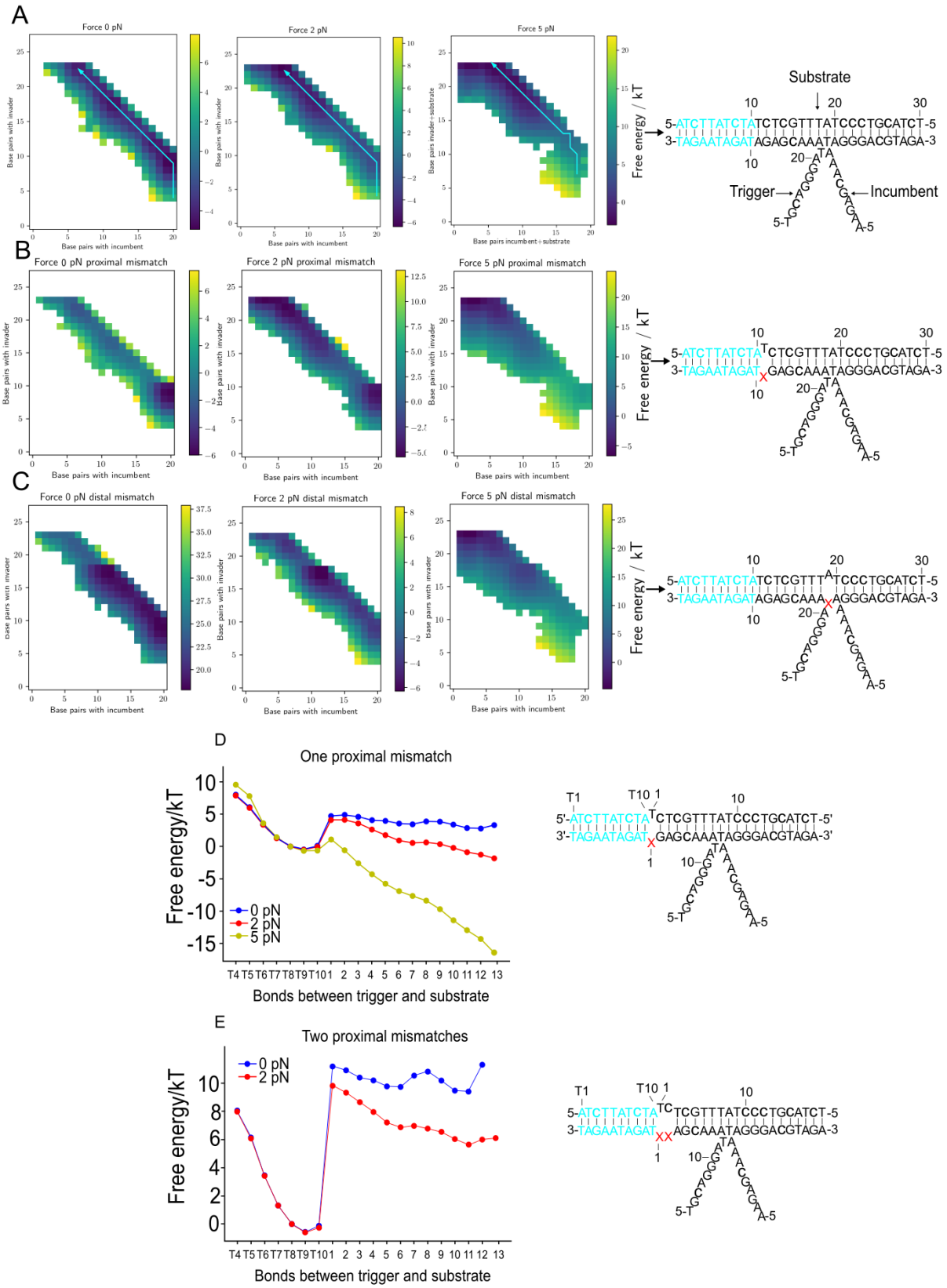


Figure S12. Free-energy landscape of truncated DNA construct as a function of bonds between the substrate and the trigger/incumbent strand under different force condition plotted along the two-dimensional path. A), TMSD with full complementary trigger under pulling force from 0 – 10 pN. B), Trigger has one proximal mismatch under pulling force, 0, 2, 5 pN. C, Trigger has one distal mismatch under pulling force, 0, 2, 5 pN. The free energy of the state with the trigger bound to the toehold but with no displacement (trigger base pairs = 10, incumbent base pairs = 20) under each force condition is set to zero. Empty spaces are unsampled. Blue arrows indicate representative reaction pathways from toehold initiation to the end of displacement. Most states with low free energy consist of a

total of 30 base pairs, which reduces to 20 after the formation of the mismatch. States in which not all target bases are paired (< 30 bp) or the total base pairs exceeds the number of target bases are (> 30 bp) are located around the reaction pathways of each landscape; they correspond to spontaneous base melting/fraying of incumbent strand or the additional base pairs between trigger and incumbent strands. C), one-dimensional path for trigger has one proximal mismatch under pulling force, 0, 2 and 5 pN. D), one-dimensional path for trigger has two proximal mismatches under pulling force, 0, 2 pN.

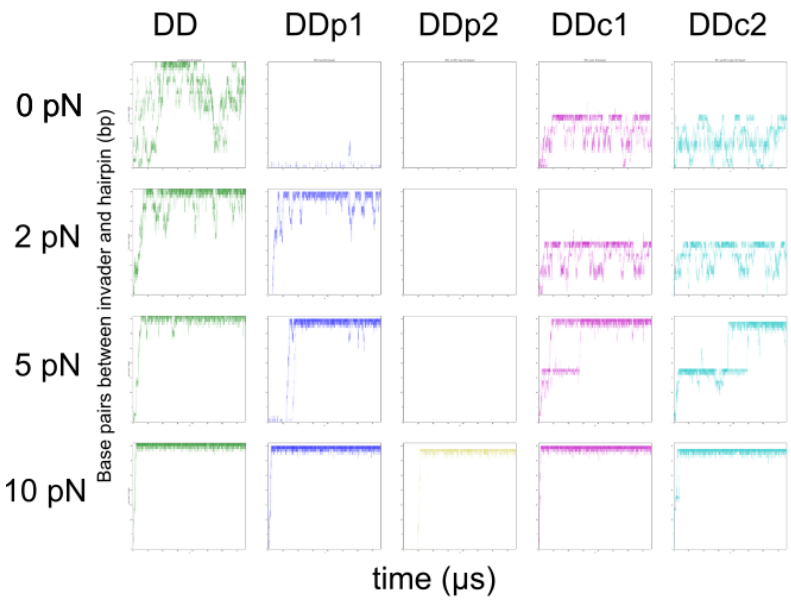


Figure S13. Passive mode simulations of DNA hairpin with different trigger using oxDNA. a), branch migration trajectory of DNA hairpin with fully complementary trigger strand. b), c), branch migration trajectory of DNA hairpin with trigger DDc1 at 1, and 2 pN condition. Histograms shows dwell time distribution at different branch migration intermediate B1 and B2. c), d), branch migration trajectory of DNA hairpin with trigger DDc2 at 1, and 2 pN condition. Histograms shows dwell time distribution at different branch migration intermediate B1 and B2. Each simulation has 7 replicas.

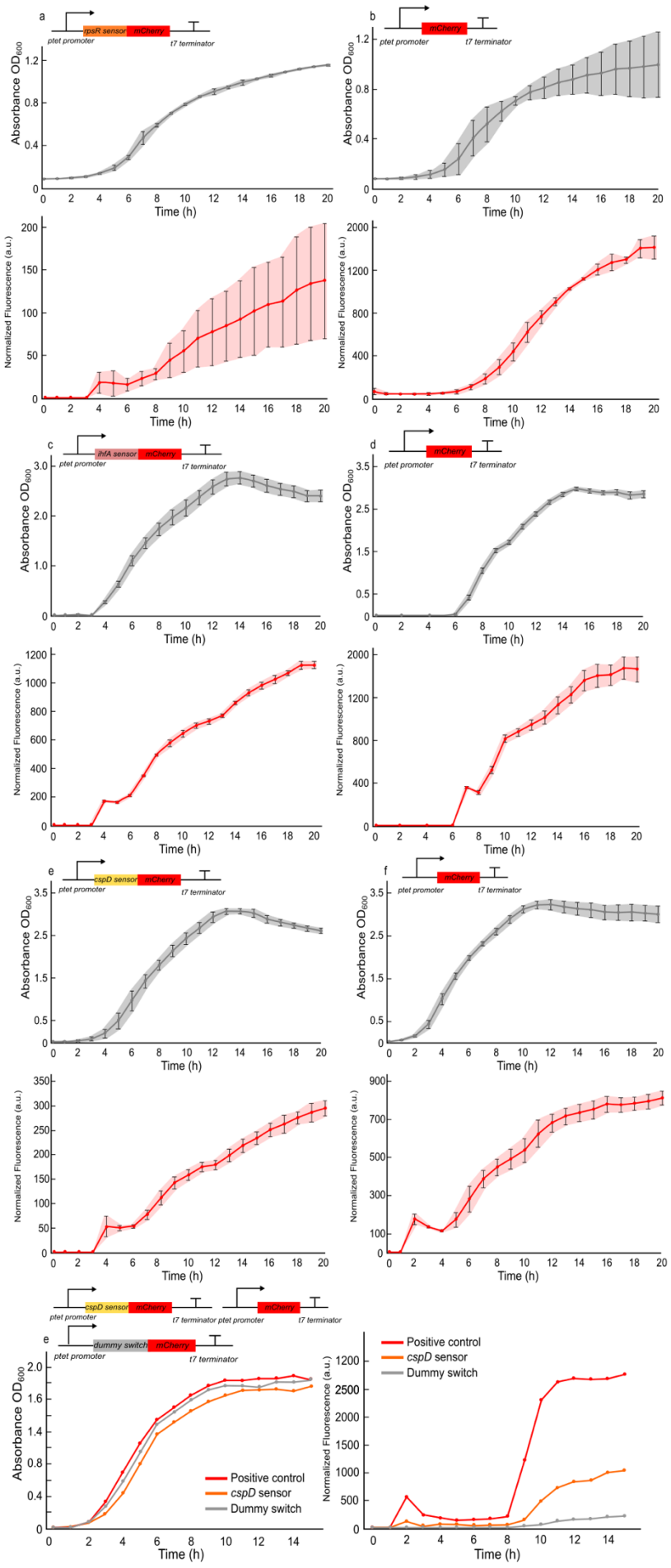
a
gpsR sensor
mCherry
ptet promoter
t7 terminator
Absorbance OD600
Time (h)
Normalized Fluorescence (a.u.)
b
mCherry
c
ihfA sensor
d
e
cspD sensor
f
dummy switch
Positive control
cspD sensor
Dummy switch

Figure S14. Growth Curves and Fluorescence-Based Kinetic Curves of *E. coli* BL21 DE3 and DH5α Z1 with toehold *rpsR*, *ihfA* and *cspD* sensors. a), growth curve (above) and kinetic curve (below) showing the relative fluorescence to absorbance ($OD_{600}$). These data refer to *E. coli* BL21 strains carrying the *rpsR* sensor. Experiments were conducted in M9 medium at 37 °C. b), growth curve (above) and kinetic curve (below) illustrating the relative fluorescence to absorbance ($OD_{600}$) for bacteria only carrying *mCherry* gene as positive control. Cultures were grown in the same M9 medium at 37 °C. c), the growth curve (above) and kinetic curve (below) present the relative fluorescence to absorbance ($OD_{600}$) for DH5α Z1 carrying the *ihfA* sensor. These experiments were conducted in M9 medium at 37 °C. d), illustrating the growth curve (above) and kinetic curve (below), this section displays the relative fluorescence to absorbance ($OD_{600}$) for *E. coli* cultures containing only the *mCherry* gene as a positive control. These cultures were grown in the same M9 medium at 37 °C. e), the growth curve (above) and kinetic curve (below) illustrating the relative fluorescence to absorbance ($OD_{600}$) for BL21 DE3 carrying the *cspD* sensor. These experiments were conducted in M9 medium at 37 °C. f. growth curve (above) and kinetic curve (below) presenting the relative fluorescence to absorbance ($OD_{600}$) for *E. coli* cultures containing only the *mCherry* gene as a positive control. These cultures were grown in the same M9 medium at 37 °C. g), growth curve (left) and kinetic curve (right) displaying the relative fluorescence to absorbance ($OD_{600}$) for *E. coli* cultures carrying *cspD* sensor, positive control, and dummy switch as negative. These cultures were grown in the 200 ml LB medium in shake flasks at 37 °C. The expression of *ptet* promoter was induced by aTC 20nM. The relative fluorescence signals/absorbance of the sensors are shown as the mean values of background subtracted fluorescence levels for three replicates from individual colonies, error bars represent their s.d, in which the *rpsR* sensor in (a), only includes only 2 replicas.

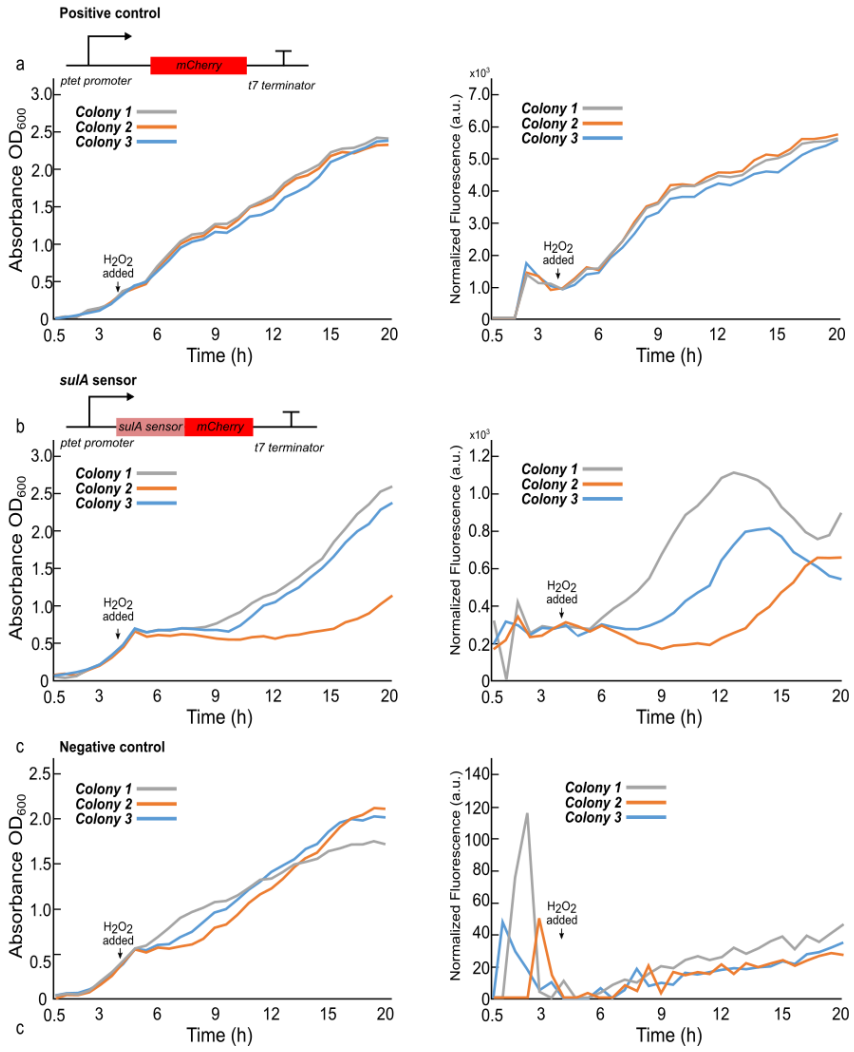


Figure S15. Growth Curves and Fluorescence-Based Kinetic Curves of *E. coli* DH5α Z1Containing the *sulA* toehold switch sensor and controls with Induction of $H_2O_2$. a), growth curve (left) and kinetics curve (right)

depicting the relative fluorescence to absorbance ($OD_{600}$) ratio for the positive control. The positive control contains only the *mCherry* gene, and the *E. coli* culture is grown in M9 medium. b), the growth curve (left) and kinetics curve (right) of relative fluorescence to absorbance ($OD_{600}$) for *E. coli* DH5α Z1 carrying the *sulA* sensor, *E. coli* culture is grown in M9 medium. c), the growth curve (left) and kinetics curve (right) representing the relative fluorescence to absorbance ($OD_{600}$) ratio for the negative control. The negative control consists of a non-fluorescent sequence, and the *E. coli* culture is grown in M9 medium. For all three groups, the expression of *ptet* promoter was induced by aTC 20nM, and SOS response was induced by adding 50 μg/ml of $H_2O_2$ after approximately 4 hours incubation at 37°C, at an OD600 reading of around 0.4. Each group includes three replicates from individual colonies

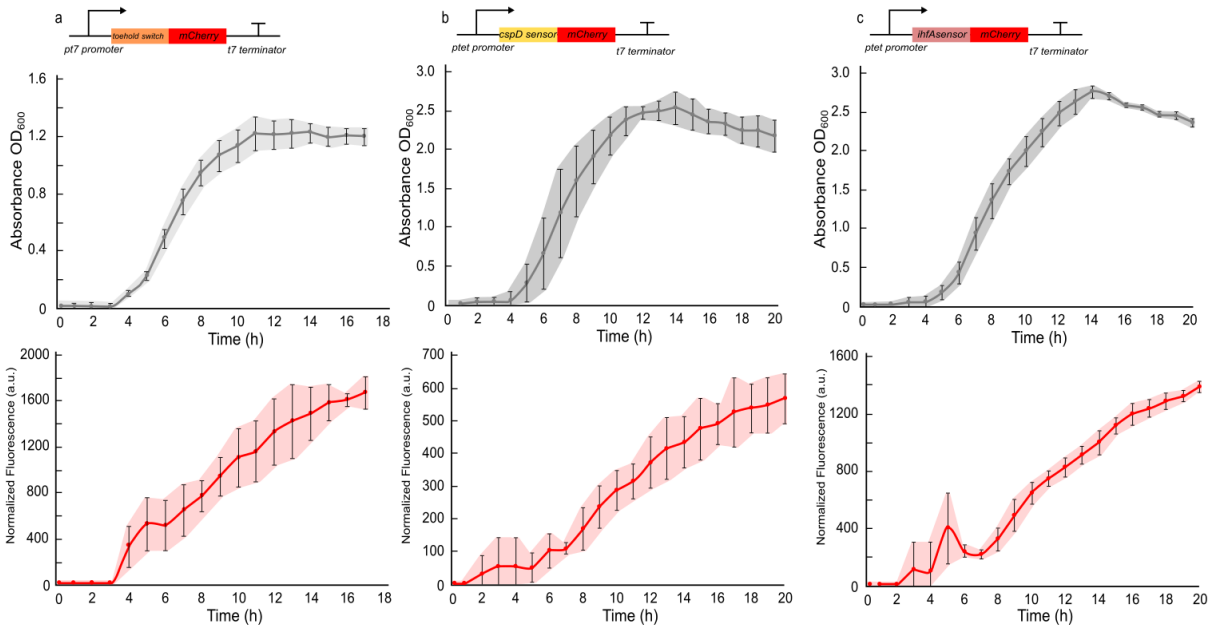


Figure S16. Growth Curves and Fluorescence-Based Kinetic Curves of *E. coli* BL21 DE3 containing different toehold switch or toehold mRNA sensors. a), growth curve (above) and kinetic curve (below) depicting the relative fluorescence to absorbance ($OD_{600}$) for BL21 DE3 carrying the original toehold switch (Ref. ) with *mCherry* gene in absence of corresponding trigger strand, and the *E. coli* culture is grown in M9 medium at 37°C. b), the growth curve (above) and kinetics curve (below) of relative fluorescence to absorbance ($OD_{600}$) for *E. coli* BL21 DE3 carrying the *cspD* sensor, *E. coli* culture is grown in M9 medium at 37°C. c), the growth curve (above) and kinetics curve (below) representing the relative fluorescence to absorbance ($OD_{600}$) ratio for *E. coli* BL21 DE3 carrying *ihfA* sensor. *E. coli* culture is grown in M9 medium at 37°C. The relative fluorescence signals/absorbance of the sensors are shown as the mean values of background subtracted fluorescence levels for three replicates from individual colonies, error bars represent their s.d.

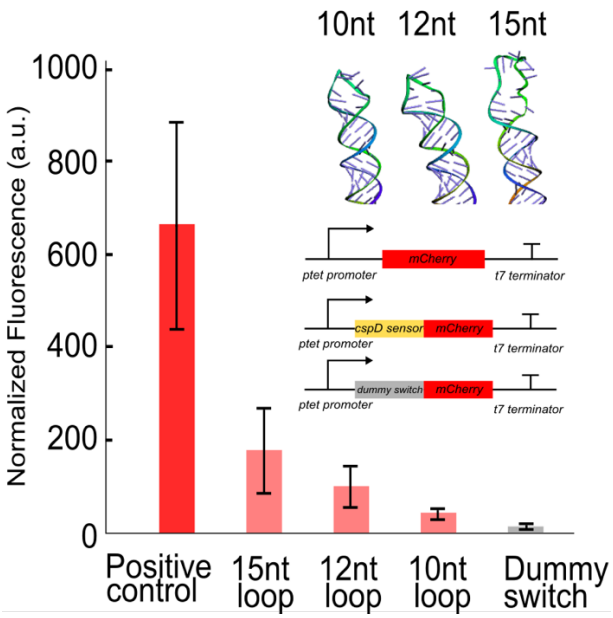


Figure S17. Relative fluorescence intensities of the *cspD* toehold sensor controlling the expression of mCherry in *E. coli*, BL21 DE3 and controls. The loop size of the toehold hairpin varies in a range from 10nt to 15nt, as depicted in the predicted tertiary structure[199]. Dummy switch as negative control includes a 15nt loop and nonhomologous toehold sequence. *E. coli* were cultured in LB medium at 37 °C, overnight. Error bars represent the standard deviation over three biological replicates.

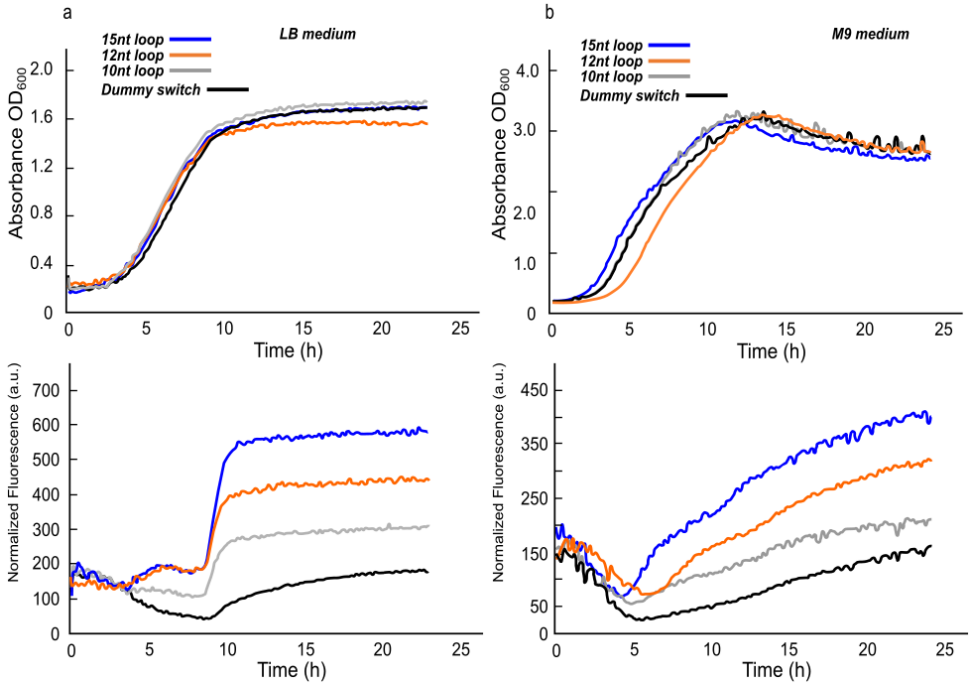


Figure S18. Growth Curves and Fluorescence-Based Kinetic Curves of *E. coli* BL21 DE3 with *cspD* toehold mRNA sensors featuring different loop sizes. a), growth curve (above) and kinetic curve (below) depicting the relative fluorescence to absorbance ($OD_{600}$). These data refer to *E. coli* BL21 DE3 strains carrying the *cspD* sensor, with loop sizes ranging from 10 nt to 15 nt. A dummy switch was included as a negative control. The experiments were conducted in LB medium at 37°C. b), growth curve (above and kinetic curve (below) illustrating the relative fluorescence to absorbance ($OD_{600}$) for BL21 DE3 carrying the *cspD* sensor with different loop sizes from 10 nt-15 nt and dummy switch grow in M9 medium at 37°C.

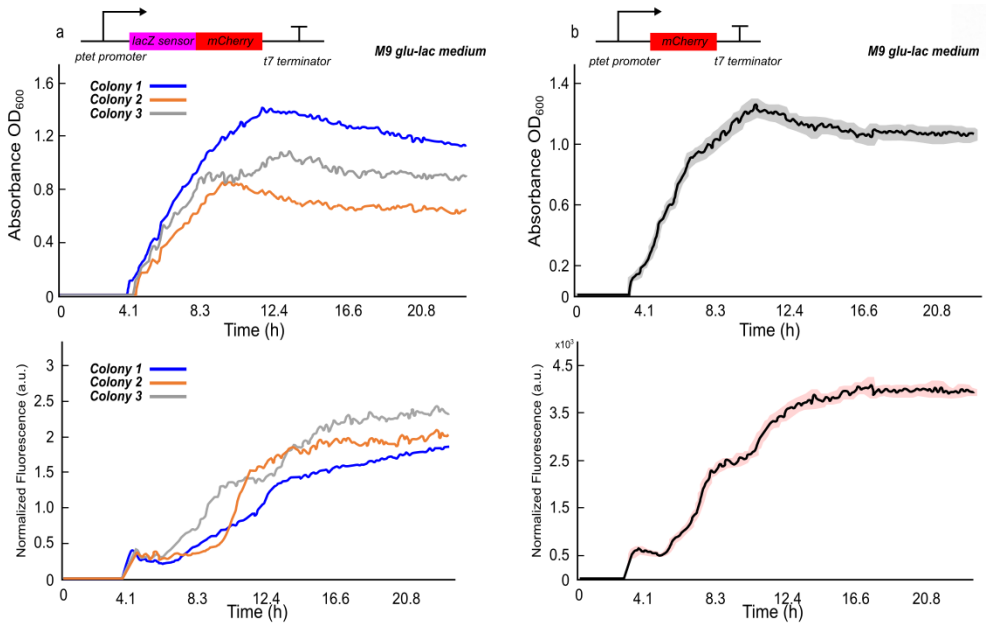


Figure S19. Growth Curves and Fluorescence-Based Kinetic Curves of *E. coli* BL21 DE3 with the toehold *lacZ* sensor responding to Starvation. a), growth curve (above) and kinetic curve (below) showing the relative fluorescence to absorbance ($OD_{600}$). These data refer to *E. coli* BL21 DE3 strains carrying the *lacZ* sensor. For each culture, bacterial experience different growth conditions under same culture condition, resulting in varying relative fluorescence signals. Experiments were conducted in M9-glu-lac medium at 37°C. b), growth curve (above) and kinetic curve (below) illustrating the relative fluorescence to absorbance ($OD_{600}$) for BL21 DE3 only carrying *mCherry* gene as positive control. Cultures were grown in the same M9-glu-lac medium at 37°C. The relative fluorescence signals/absorbance of the sensors are shown as the mean values of background subtracted fluorescence levels for three replicates from individual colonies, shaded error bars represent their s.d.

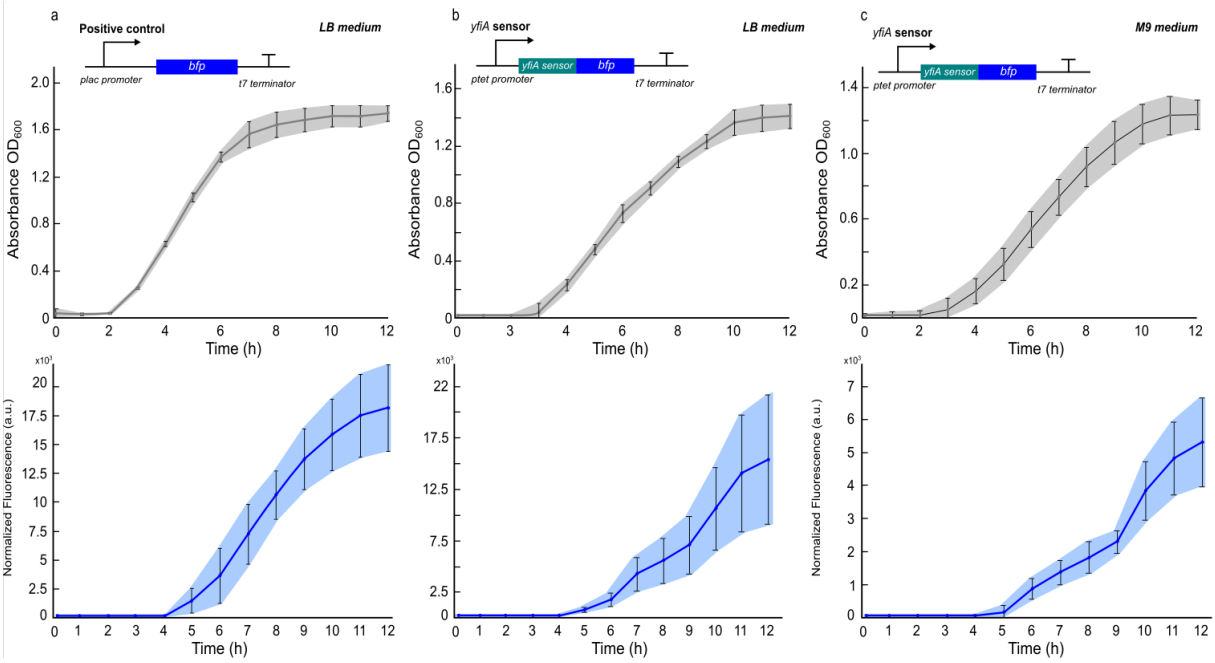


Figure S20. Growth Curves and Fluorescence-Based Kinetic Curves of *E. coli* DH5α Z1 with the toehold *yfiA* sensor in LB and M9 mediums. a), Depicting the growth curve (above) and kinetic curve (below), it shows the relative fluorescence to absorbance ($OD_{600}$) for *E. coli* cultures containing only the *bfp* gene as a positive control. These cultures were grown in LB medium at 37 °C. b), The growth curve (above) and kinetic curve (below)

illustrate the relative fluorescence to absorbance ($OD_{600}$) for *E. coli* DH5α Z1 strains carrying the *yfiA* sensor. These experiments were conducted in LB medium at 37 °C. c), growth curve (above) and kinetic curve (below) showing the relative fluorescence to absorbance ($OD_{600}$). These data refer to *E. coli* DH5α Z1 strains carrying the *yfiA* sensor. Experiments were conducted in M9 medium at 37°C. The expression of ptet promoter was induced by aTC 20nM. The relative fluorescence signals/absorbance of the sensors are shown as the mean values of background subtracted fluorescence levels for three replicates from individual colonies, shaded error bars represent their s.d.

## Design rationale for riboswitch-inspired toehold riboregulators

The riboregulators investigated in this study were inspired by the previously developed synthetic toehold switches and naturally occurring riboswitches (Fig. S21 a & b). Natural riboswitches consist of an aptamer located at the 5' end, which, upon binding to a specific small metabolite, triggers an allosteric effect. This allosteric rearrangement of the RNA structure subsequently influences a downstream "expression platform" containing gene regulatory elements, resulting in metabolite-dependent control of gene expression. In contrast to riboswitches, the toehold switches exploit toehold-mediated strand displacement (TMSD) by an RNA trigger molecule to induce structural rearrangement and thus regulate gene expression.

To combine the functional features of natural riboswitches with those of toehold switches, we utilized a toehold hairpin that contains riboswitch-inspired regulatory elements – anti-RBS, anti-anti-RBS or anti-terminator sequences - within the loop region. Binding of the trigger RNA is thus expected to induce a conformational rearrangement that results in the control of the downstream translational initiation or transcriptional termination or (Fig. S21 c).

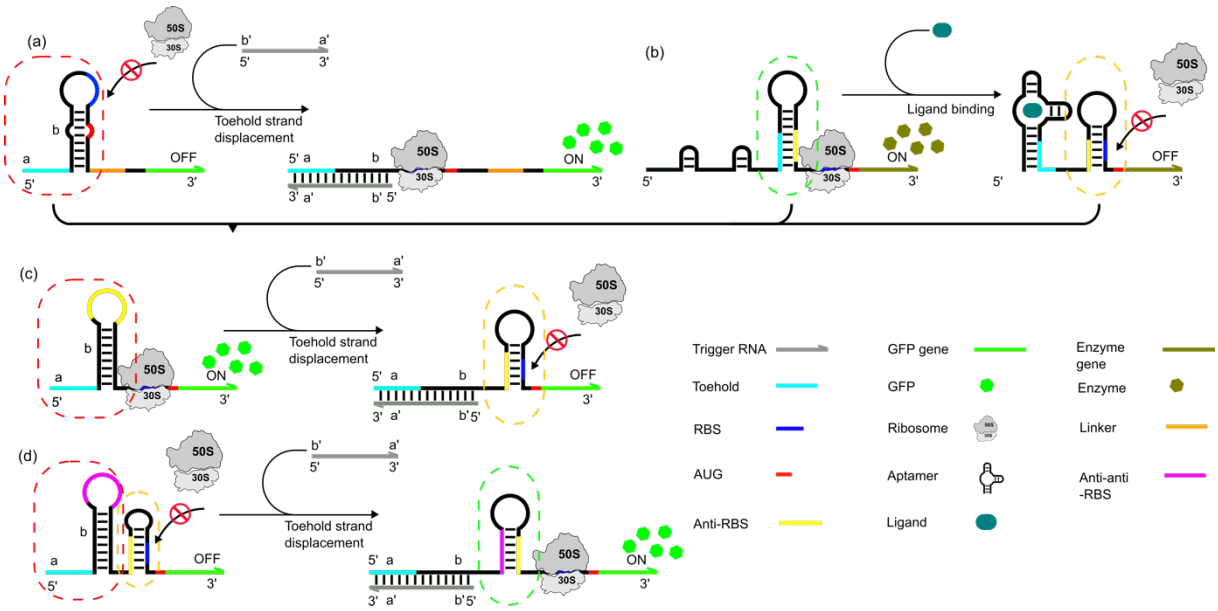


Figure S21. Comparison of the regulatory mechanisms of the toehold switch, riboswitches, and the riboswitch-inspired toehold riboregulators controlling translational initiation and repression investigated in this work. a), Regulation of translational initiation by a toehold switch. In the absence of trigger RNA (grey), the toehold hairpin (within the red box) includes a free toehold region (light blue) at the 5' end, a loop, and bulge in which the RBS (blue) and start codon (red) are sequestered, preventing the binding of the ribosome to the RBS and inhibiting translational initiation. Trigger RNA binding to the toehold region leads to toehold-mediated strand displacement (TMSD) and unfolding of the toehold hairpin, exposing the RBS and thus enabling translational initiation. b),

Regulatory mechanism of a natural riboswitch controlling translational repression. In the absence of ligands (cyan), the riboswitch contains an aptamer sequence and an anti-RBS hairpin (within the green box) in which the anti-RBS (yellow) sequence is masked, the RBS is accessible and downstream gene translation is enabled. When the aptamer is bound to a ligand, however, the riboswitch adopts a conformation in which the anti-RBS sequence can bind to and block the RBS (within the yellow box), and thus inhibit translation. c), Schematic representation of a riboswitch-inspired toehold riboregulator controlling translational repression. In the absence of trigger RNA, the riboregulator is in the translational ON state, in which a toehold hairpin (within the red box) sequesters the anti-RBS sequence (yellow) within its loop region. Upon binding of trigger RNA to the toehold region, TMSD results in unfolding of the toehold hairpin, releasing the anti-RBS sequence and thus allowing the formation of the anti-RBS hairpin (within the yellow box), which masks the RBS and inhibits translation. d), Schematic representation of a riboswitch-inspired toehold riboregulators that controls translational activation. In the absence of trigger RNA, the toehold hairpin (within the red box) confines an anti-anti-RBS sequence (purple) within its loop, which is followed by a hairpin that contains an anti-RBS bound to the RBS in its stem (within the yellow box). Binding of trigger RNA unfolds the toehold hairpin via TMSD, releases the anti-anti-RBS and thus allows formation of the anti-anti-RBS hairpin (within the green box). The RBS is then accessible to the ribosome and translation is initiated.

### Anti-RBS translational toehold riboregulators

We first constructed a series of anti-RBS toehold riboregulators, which control the accessibility of the RBS for ribosome binding and translational initiation. We developed two types of translational regulators (Fig. S22) which can either activate or repress gene translation through trigger binding and TMSD. The anti-RBS toehold activator is composed of a toehold hairpin and an anti-RBS hairpin (Fig. S22 a). The toehold hairpin comprises a 14nt unpaired toehold region at the 5' end and a stem-loop, which confines the anti-anti RBS sequence (11 nt) within the loop region. The sequence of the anti-RBS hairpin (ΔG = -7.80 kcal/mol) is derived from the natural *thiM* riboswitch. Trigger RNA binds to the toehold region and initiates TMSD, which unwinds the toehold hairpin and thus releases the anti-anti RBS sequence. The trigger and toehold riboregulator form an intermediate complex, from which the anti-anti-RBS sequence can invade the anti-RBS hairpin. The resulting formation of an anti-anti-RBS hairpin (ΔG = -15.0 kcal/mol) exposes the RBS, resulting in translational initiation. In the case of the anti-RBS toehold repressor, the toehold hairpin comprises a free toehold region (15nt) and an anti-RBS sequence within the loop region (Fig. S21 b). The RBS is accessible for ribosome binding, allowing the translation process to proceed. Trigger RNA binding and TMSD unwind toehold hairpin and free anti-RBS sequence, leading intermediate state and refolding process. After secondary structure rearrangement, anti-RBS blocks the RBS region, resulting in the formation of anti-RBS hairpins and repression of ribosome binding.

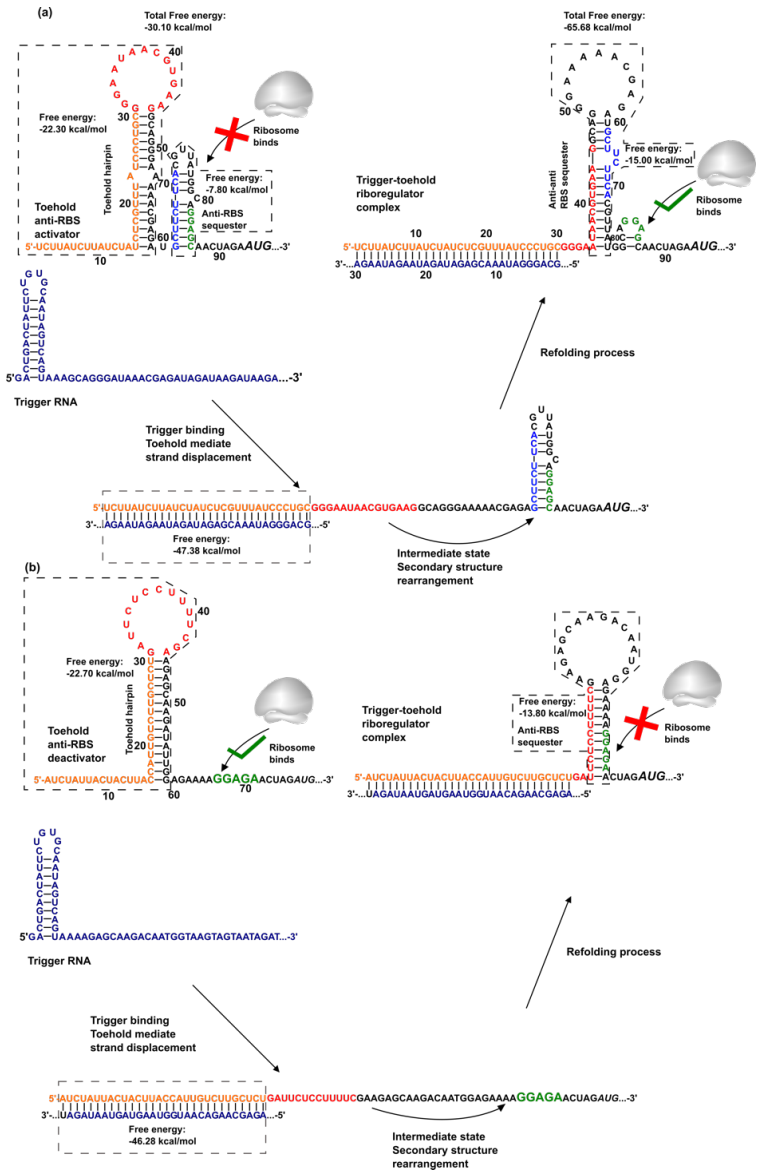


Figure S22. Schematic of translational toehold riboregulators with RNA sequences. a), Regulatory mechanism of a translational activator. The toehold hairpin includes a free toehold region, incumbent sequence (orange), and unpaired anti-anti RBS sequence (red). Anti-RBS hairpin includes an anti-RBS (blue) and RBS (green). Trigger RNA (dark blue) binds to the toehold riboregulator, forming an intermediate complex which promotes refolding of the structure. b), Regulatory mechanism of a toehold translational repressor. The toehold hairpin has a free toehold region, an incumbent sequence (orange), and an unpaired anti-RBS sequence (red). The RBS (green) is unpaired and accessible for ribosome binding. Trigger RNA (dark blue) binds to the toehold riboregulator, which induces refolding of the riboregulator to a translationally inactive state. The trigger RNAs are protected from degradation via a 5' hairpin (shown) and the 3' terminator hairpin (not shown).

## Anti-terminator transcriptional toehold riboregulators

The following designs are toehold riboregulators, which control transcriptional termination through an intrinsic (rho-independent) terminator. The toehold transcriptional activators I and II each consist of a toehold hairpin and the *t22* intrinsic terminator (Fig. S23 a & b). The *t22* terminator hairpin (ΔG = -18.60 kcal/mol) derives from the phage P22 late terminator. The toehold hairpin comprises a 17nt unpaired toehold region at the 5' end and a stem-hairpin which confines the

anti-terminator sequence (15 nt) in its loop region. Trigger RNA binds to the toehold region and initiates TMSD, uwinding the toehold hairpin and releasing the anti-terminator sequence. In the resulting intermediate state, the anti-terminator sequence is designed to base-pair either with the releasing site of the *t22* terminator (GCG) (activator I) or with the stem sequence of the terminator (activator II), preventing the release of the nascent RNA and thus allowing transcriptional elongation to proceed.

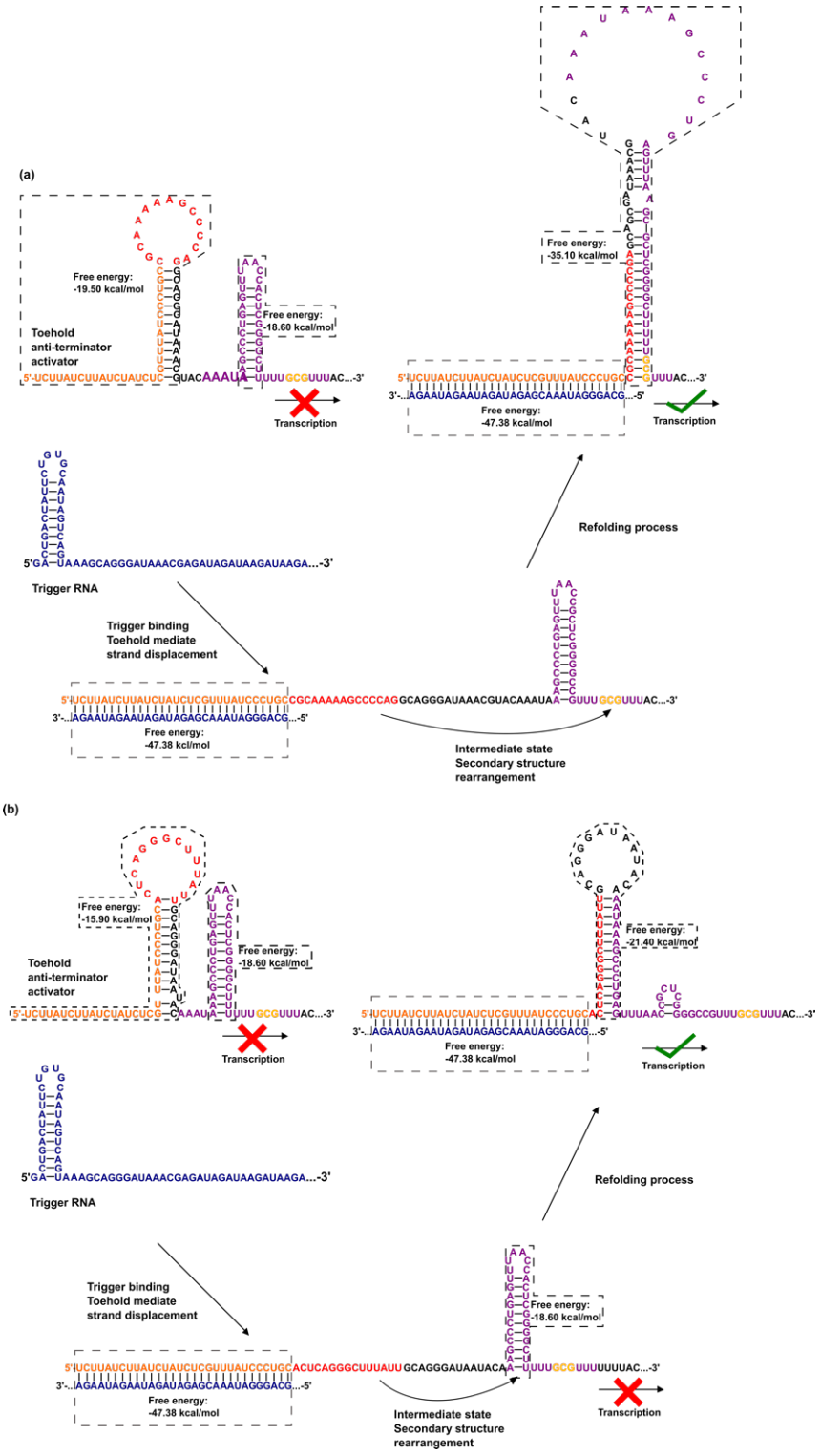

Figure S23. Schematic representation of transcriptional toehold riboregulators with RNA sequences. a), Mechanism of transcriptional toehold activator I. The toehold hairpin includes a free toehold region, an incumbent sequence (orange), and an unpaired anti-terminator sequence (red). The releasing site (yellow) of the *t22* terminator (purple) is initially unpaired. Trigger RNA (dark blue) binds to the toehold riboregulator, forming a new complex, which includes a double-stranded region and anti-terminator hairpin, in which the releasing site is sequestered. b), Mechanism of transcriptional toehold activator II. In this case the initially sequestered anti-terminator sequence (red) base-pairs with the first half of the stem sequence of the *t22* terminator (purple). As before, binding of trigger results in refolding and thus disruption of the terminator hairpin.

## The *tna* operon

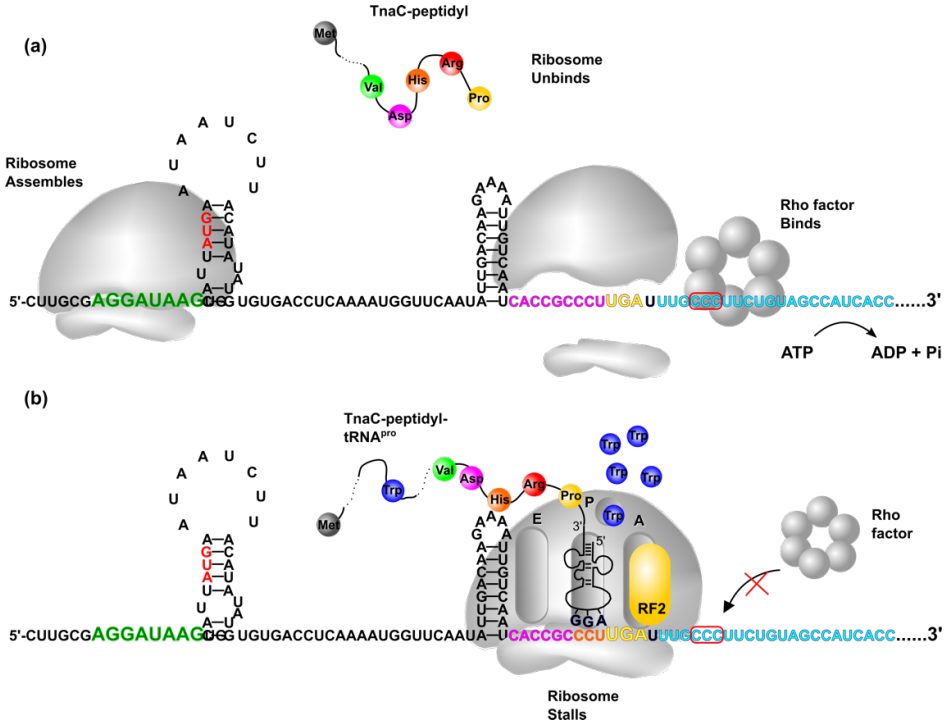


Figure S24. Schematic representation of the mechanism of the *tna* operon in *E. coli* including the *tna* operon mRNA sequence and secondary structures. a), The ribosome initially assembles in the RBS region (in green) and starts translational elongation of the 24-residue leader peptide TnaC (chain of circles) until it reaches the *tnaC* stop codon, UGA (in yellow). At low levels of tryptophan, translation terminates at the *tnaC* stop codon and the ribosome releases, while the Rho factor binds to the *rut* (rho utilization) site (in light blue) and terminates transcriptional elongation. b), High tryptophan levels inhibit the action of release factor RF-2 and thus the TnaC peptide remains covalently linked to $tRNA^{pro}$. In conseqence the ribosome stalls at the stop codon. The presence of the stalled TnaC-peptidyl-$tRNA^{Pro}$ ribosome complex prevents binding of the Rho factor to the boxA (in purple) and *rut* sites and therefore prevents transcriptional termination.

## Anti-*rut* toehold riboregulators

As an alternative to transcriptional toehold switches based on an intrinsic terminator, we designed an anti-*rut* toehold riboregulator (Fig. S25), which controls the accessibility of the *rut* site for the binding of the termination factor Rho. Anti-*rut* toehold riboregulators can activated through trigger binding and TMSD.

The anti-*rut* activator is composed of a toehold hairpin and an anti-rut hairpin (Fig. S25). The toehold hairpin comprises a 16nt unpaired toehold region at the 5' end and a hairpin that sequesters the anti-*rut* sequence (15 nt) in its loop region. In this state, the *rut* site is accessible for the Rho factor, which binds and leads to transcriptional termination. Trigger RNA can bind to the toehold region, invade the toehold hairpin, and thereby release the anti-*rut* sequence. Refolding of the structure sequesters the *rut* site in the stem of an anti-*rut*/*rut* hairpin (ΔG = -16.90 kcal/mol), allowing transcription elongation to proceed.

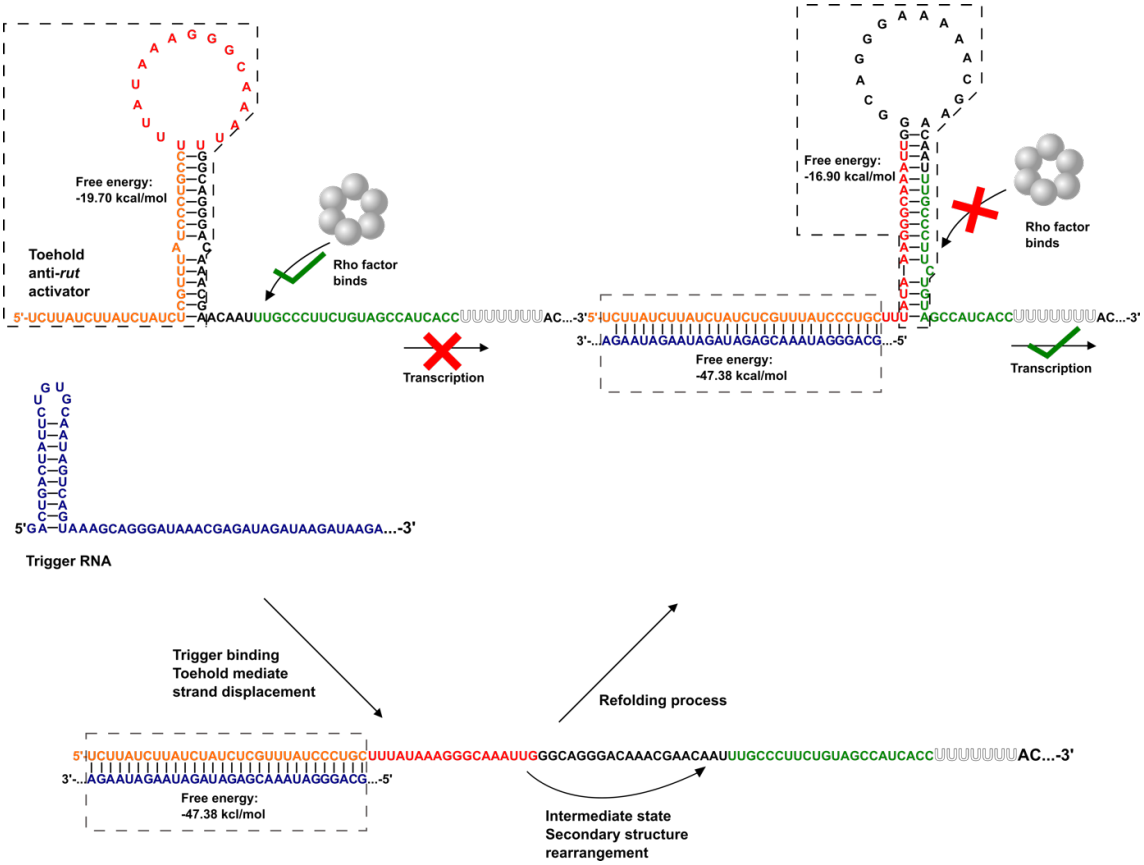


Figure S25. Schematic of toehold anti-*rut* riboregulators with RNA sequences. Regulatory mechanism of a transcriptional toehold activator. An unstructured toehold region precedes the toehold hairpin, which includes an incumbent sequence (orange) and an unpaired anti-rut sequence (red) in the loop region. The rut site (green) for Rho factor binding is accessible, followed by a transcriptional pausing site ($U_7$). Trigger RNA (dark blue) can bind to the toehold and invade the hairpin. Refolding of the mRNA results in an alternative anti-rut hairpin, which sequesters the rut site.

## A toehold riboregulator based on the tna operon

Rather than using an anti-*rut* sequence to control the accessibility of the *rut* site, we also utilized ribosome stalling to regulate binding of the Rho factor, which is similar in approach as in the natural *tna* operon (cf. Fig. S24). Riboregulators that switch the accessibility of an anti-*tna* sequence can be repressing. Our anti-*tna* toehold repressor (Fig. S26) comprises a 15 nt unpaired toehold region at the 5' end and a toehold hairpin that sequesters an anti RBS sequence (15 nt) within its loop region. In the OFF state (in the absence of trigger RNA), the ribosome binds to RNA and stalls at the at the stop codon between boxA sequence *rut* site,

allowing transcriptional elongation to proceed. In the presence of trigger RNA, refolding of the riboregulator sequesters the RBS by the anti-RBS and forms a hairpin stem (ΔG ≈ -12.50 kcal/mol), effectively rendering the *rut* site accessible for Rho factor and therefore resulting in transcription termination.

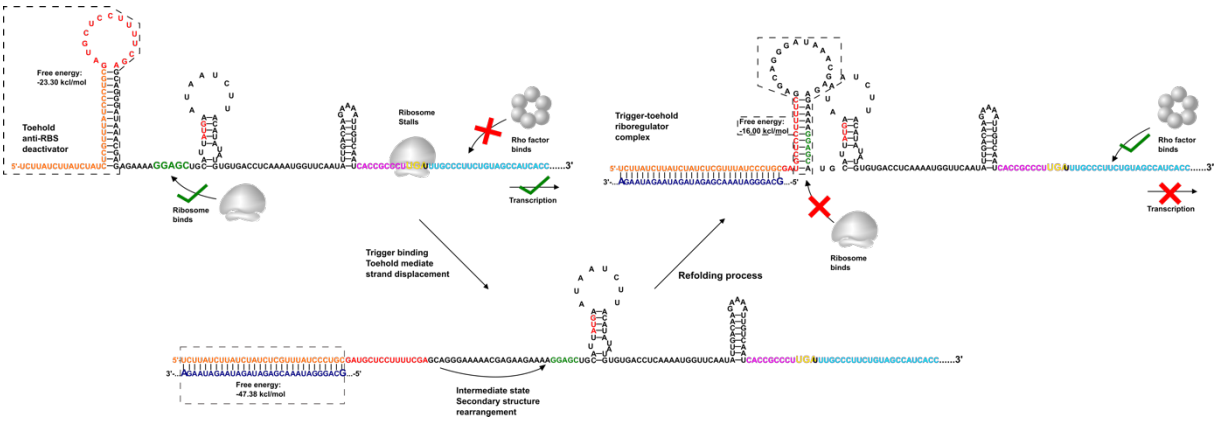


Figure S26. Schematic of riboregulators derived from the *tna* operon with RNA sequences. Regulatory mechanism of an anti-*tna* toehold repressor. The toehold hairpin is comprised of a free toehold region, an incumbent sequence (orange) and an unpaired anti RBS sequence (red), and the RBS (green) is accessible for ribosome binding. In the OFF state, the ribosome can bind and translate the tnaC peptide starting from the start codon (red AUG) until it reaches the stop codon (yellow) situated between the boxA (purple) and *rut* site (light blue). Translational stalling prevents Rho-dependent termination, allowing transcription elongation. Trigger RNA activates the riboregulator by TMSD into the toehold hairpin. Refolding of the anti-RBS hairpin sequesters RBS from ribosome binding, Rho factor binds to the *rut* site (light blue), which leads to transcriptional termination.

## A NOR gate based on combined transcriptional and translational riboregulation

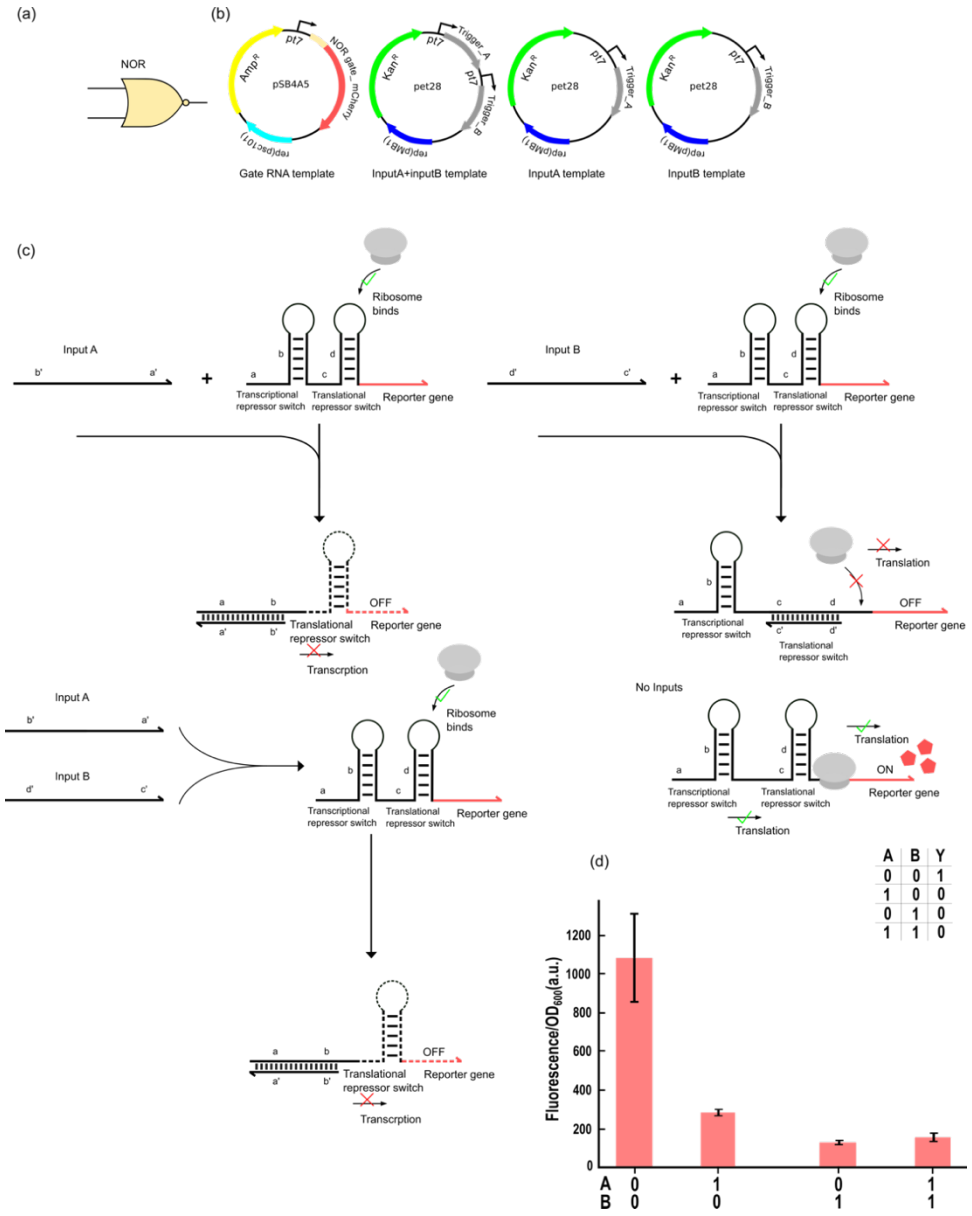


Figure S27. A two-input logic NOR gate device that combines transcriptional and translational repression. a), Logic table for the NOR gate. b), Plasmid schemes. The NOR gate RNA coding for mCherry is under the control of a *pt7* promoter on a pSB4A5 low copy number plasmid. The different trigger RNA combinations are each transcribed from a *pt7* promoter on a pet28 high copy number plasmid. c), the NOR gate RNA consists of a toehold-transcriptional repressor based on the *tna* operon (as in Fig. S24) and a translational toehold repressor (as in Fig. S21 b). In the absence of triggers, ribosomes can bind and translate the reporter gene. In the presence of input, A, transcription is terminated prematurely in a Rho-dependent manner as described above. In the presence of input B alone, transcription is not terminated, but translation is repressed. Addition of both triggers also leads to a low output of the reporter protein mCherry. d), Relative fluorescence intensities under ON and OFF state of toehold translational activator respectively. The error bars for the ON and OFF states are from the standard deviation (s.d.) for biologically independent three samples.

## Leak expression under Rho-dependent transcription termination

We observed leaky expression from the *tna* operator *in vivo*. First, we tested a wildtype and a modified *tna* operator, which controls the transcription of a downstream *mCherry* gene (Fig. S28). The wild type *tna* operon includes a sequence coding for the tnaC peptide and the *rut* binding site,

followed by a non-coding sequence before the downstream gene. We modified the *tna* operon by replacing the wild-type non-coding sequence with 7 thymidine, which in the RNA transcript mimics the poly-U of the intrinsic terminator to generate an additional transcriptional pause. Both genetic constructs are inserted into a high copy number plasmid and expressed in *E. coli* BL21 DE3. With the induction of tryptophan in the culture medium, binding of the tryptophan to the ribosome results in ribosome stalling at the stop codon of the *tnaC* sequence, which blocks the binding of Rho facto to *rut* site and thus allows transcription of the downstream gene.

Next, we removed the tnaC peptide sequence from the operon and retained only the *rut* site to control downstream *mCherry* gene expression to check the influence of the tnaC peptide sequence. The results still show leaky expression of mCherry under the regulation of *rut* site (Fig. S28). However, in the absence of the tnaC peptide sequence, the additional transcriptional pausing site enhanced the Rho-dependent termination compared to the wild-type sequence.

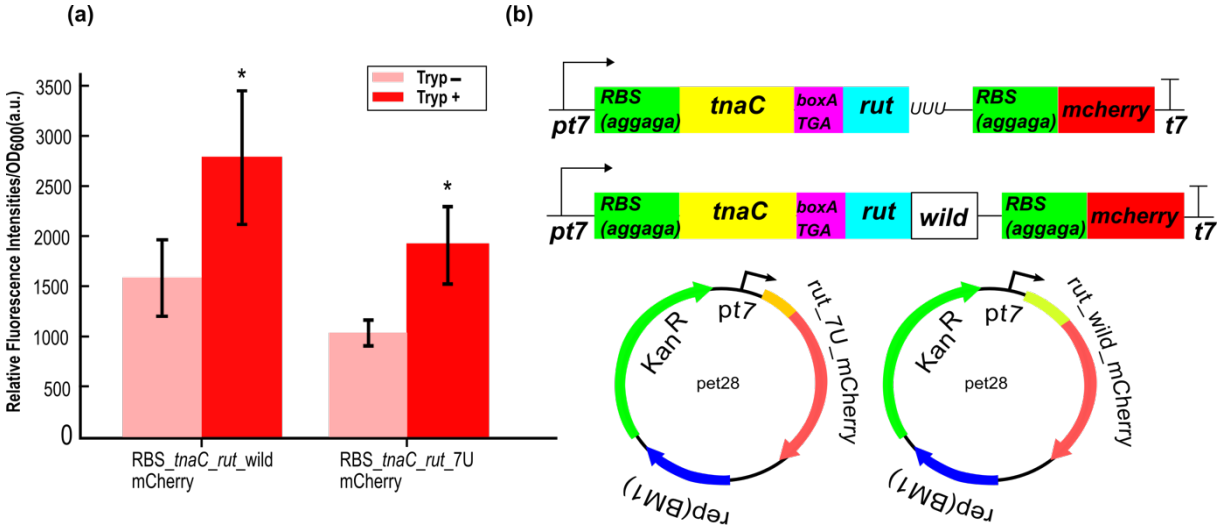


Figure S28. Relative fluorescence intensities measured for mCherry expression under the control of a wild-type and a modified *tna* operator in *E. coli*, BL21 DE3 with and without induction with 5mM tryptophan. T7 RNA polymerase is induced by 1mM IPTG. a), *in vivo* mCherry expression for the different *tna* operators and induction states. Error bars represent standard deviation over three biological replicates. b), Modified tna operators with *rut*-7U and *rut*-wild. For the modified operator, an additional pausing (7 U) site is inserted right after rut site. Genetic constructs are under the control of a t7 promoter and terminator and coded on a high copy number plasmid pet28.

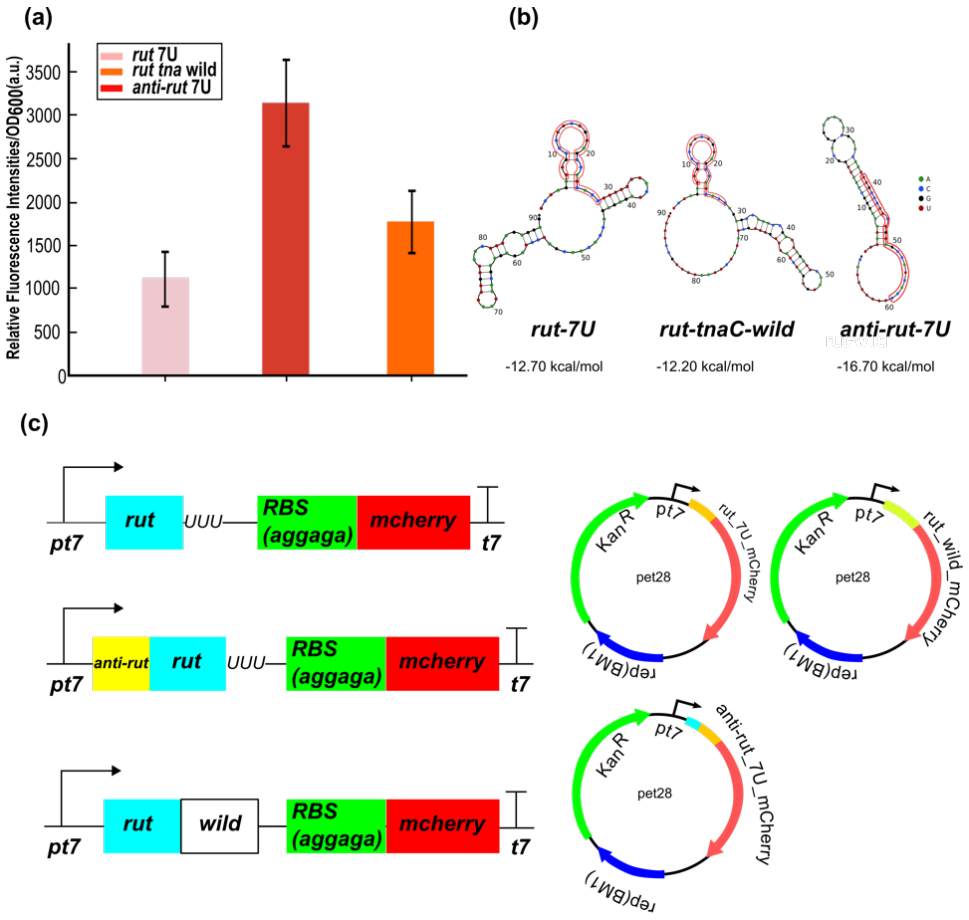


Figure S29. Expression of mCherry in *E. coli*, BL21 DE3 from modified *tna* operators. a), *in vivo* mCherry expression levels for modified *tna* operator rut7U, rut tna wild type and anti-rut 7U. Error bars represent the standard deviation from three biological replicates. b), Prediction of the RNA secondary structures and free energies of the *rut* site including part of the downstream sequence using NUPACK. The *rut* sites of *rut*-7U and rut wild are marked with red frames. c), Scheme of the corresponding genetic constructs. Gene cassettes are under the control of a *t7* promoter and terminator on high copy number plasmid.

## mCherry expression controlled by the original toehold switch

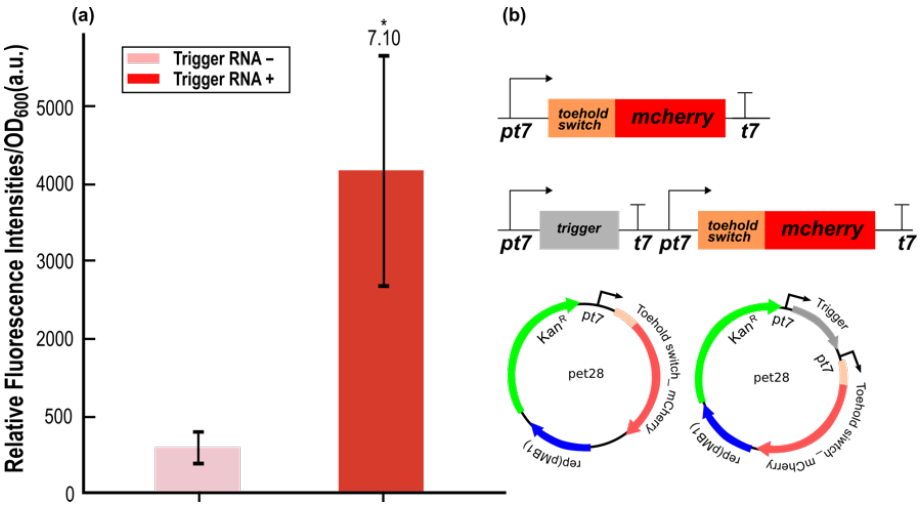

Figure S30. Relative fluorescence intensities of the best performing toehold switch from Green et al. (Ref. 11) controlling the expression of mCherry in *E. coli*, BL21 DE3, where the expression of T7 RNA polymerase is induced with 0.1 mM IPTG. The original toehold switch was reported to have an ON/OFF ratio of 660 in flow cytometer measurements when controlling the expression of GFP. Under our experimental conditions, this ratio is ≈ 7. a), *in vivo* mCherry expression measured as the fluorescence end level/$OD_{600}$ for the OFF and ON state of the toehold switch, i.e., expressed in the absence or the presence of trigger RNA. The ON/OFF ratio of the switch is shown, error bars represent the standard deviation over three biological replicates. b), Genetic constructs for the toehold switch and trigger RNA. Toehold switch-*mCherry* and trigger RNA are each under the control of a *pt7* promoter and terminator on the same high copy number plasmid and are thus expected to be generated at similarly high expression levels.

## Improving TMSD efficiency via mismatches in the toehold hairpin stem

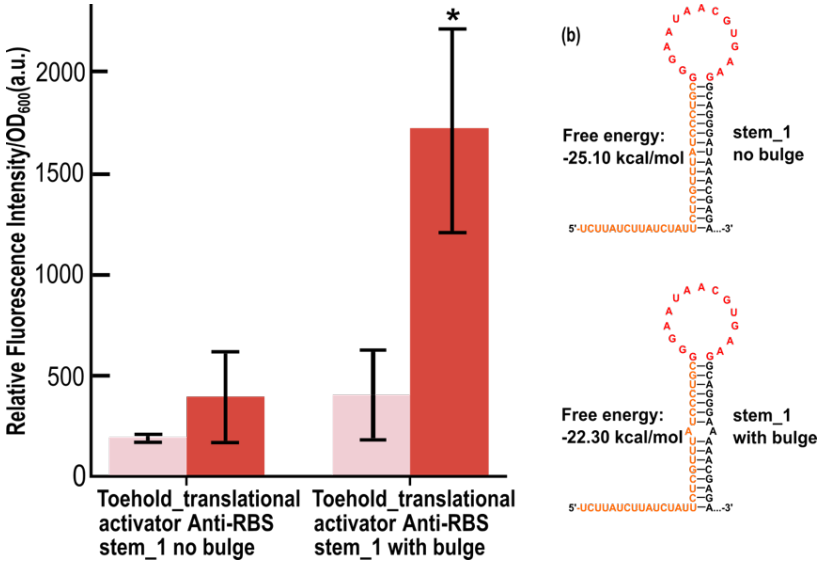


Figure S31. Introduction of mismatches in the toehold hairpin stem improves TMSD. a), *in vivo* mCherry expression levels (fluorescence/OD600 end level) under the control of translational toehold activator stem_1 with and without bulge in the stem, in the absence and presence of trigger RNA. Expression is measured in *E. coli*, BL21 DE3 with induction of the T7 RNA polymerase by 1mM IPTG. Error bars represent the standard deviation over three biological replicates. b), Prediction of mRNA secondary structures and free energies of the toehold hairpins (with and without bulge) using NUPACK.

## qPCR calibration curve

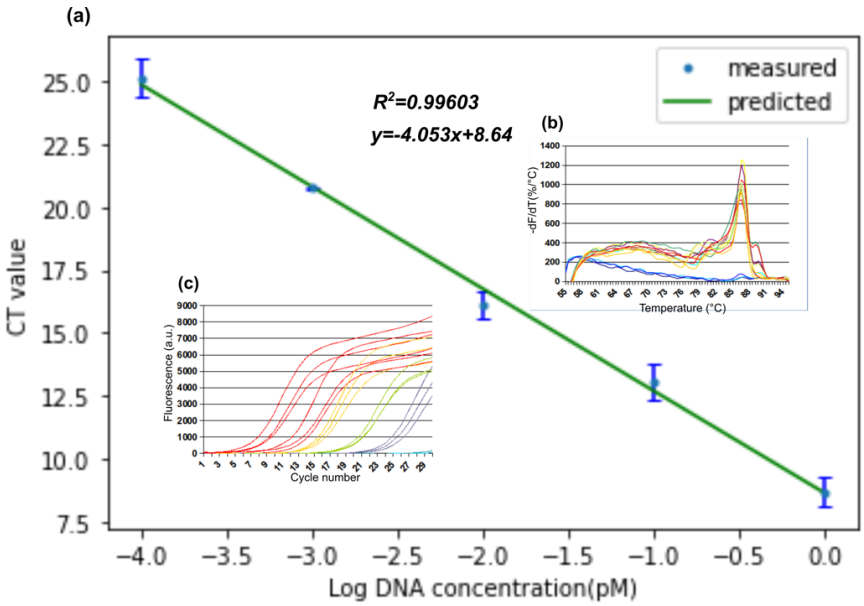


Figure S32. Linearity of the qPCR assay. a), Standard curve of threshold cycle number plotted against the log of the purified *mCherry* DNA template concentration in picomolar (pM) units. The ten-fold serial dilutions are linear over five orders of magnitude. Data represent three replicates of each dilution. $R^2$= 0.99603. b), melting curve of dF/dT value of purified *mCherry* DNA template plotted against the increasing temperature. No significant unspecific amplification was observed in control reactions that did not contain DNA template. c), Amplification plot of the same experiment showing the ten-fold dilution series of DNA concentrations ($10^{-4}$-10 pM) replicated 3 times.

## Flow cytometry measurements

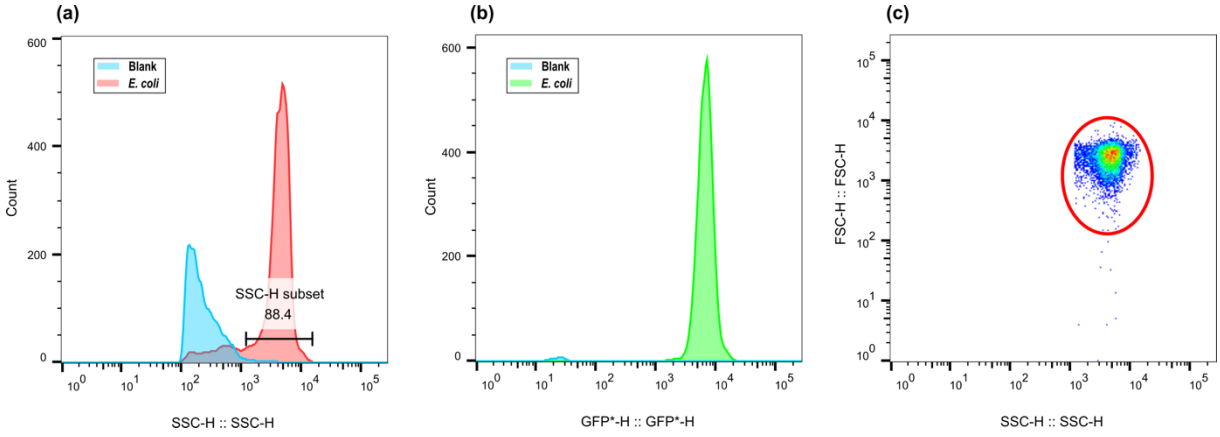


Figure S33. Flow cytometry measurements. a), Histogram of SSC-H for the *E. coli* population and blank. b), Histogram of median of GFP-H for the *E. coli* population and blank. c), SSC-H vs FCS-H density plot. Each dot or point on the plot represents an individual event that passed through the laser of the flow cytometer.

## The effect of trigger RNA expression on cell growth

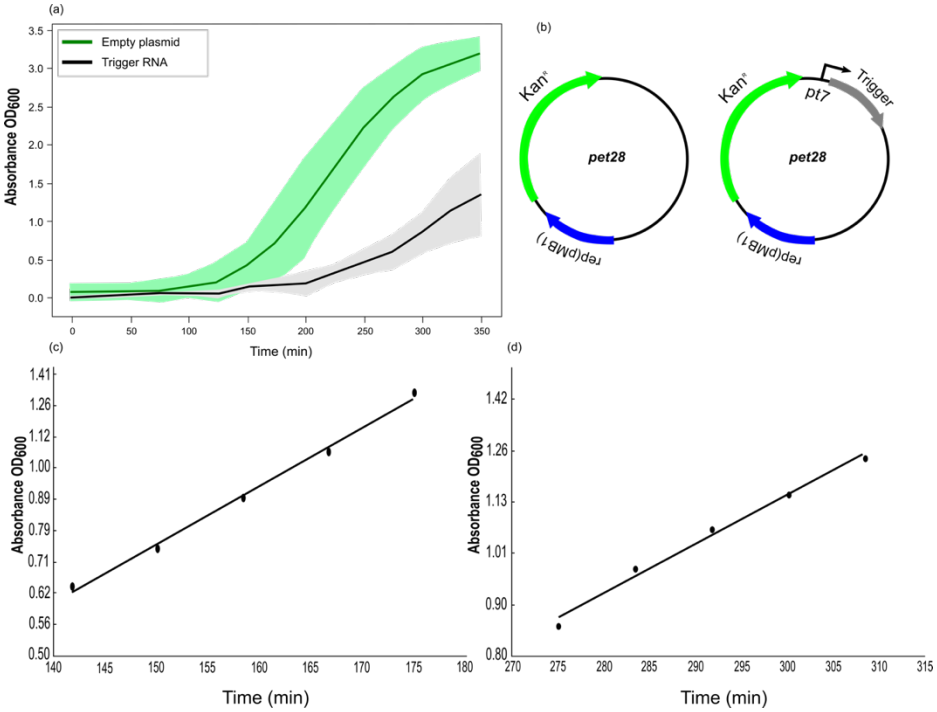


Figure S34. Effects on bacterial growth. a), Growth curves of *E. coli* BL21 DE3 expressing trigger RNA grown at 37°C, 500 rpm for 6 h. The optical densities at 600 nm ($OD_{600}$) were measured. Shaded error bars represent the s.d. of three replicate experiments. b), Genetic constructs for the trigger RNA. Trigger RNA is under the control of a pt7 promoter and terminator on the high copy number plasmid, an empty plasmid is used for a negative control. c and d, Absorbance in the exponential growth phase of *E. coli* BL21 DE3 that c), carry the empty control plasmid (doubling time ≈ 30 min), and d), that express trigger RNA (doubling time ≈ 65 min).

## Orthogonality test with non-cognate RNA triggers and toehold-translational activator

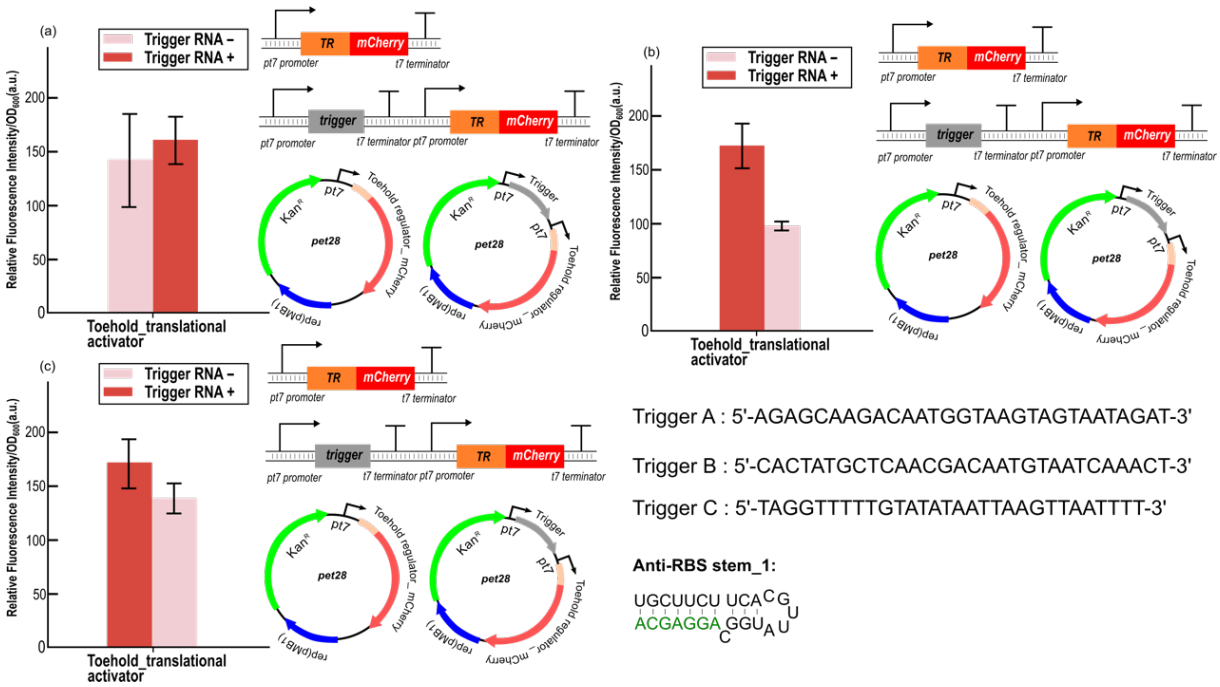


Figure S35. Relative fluorescence intensities of mCherry under the control of toehold-translational activator with non-cognate trigger RNAs. a – c), shows *in vivo* mCherry expression levels (fluorescence/OD600 end level) under the control of translational toehold activator stem_1 in the absence and presence of the non-cognate triggers A, B,

and C, respectively. Expression is measured in *E. coli*, BL21 DE3 with induction of the T7 RNA polymerase by 1mM IPTG. Error bars represent the standard deviation over three biological replicates.

## Switching behavior of a translational activator transcribed from a constitutive promoter

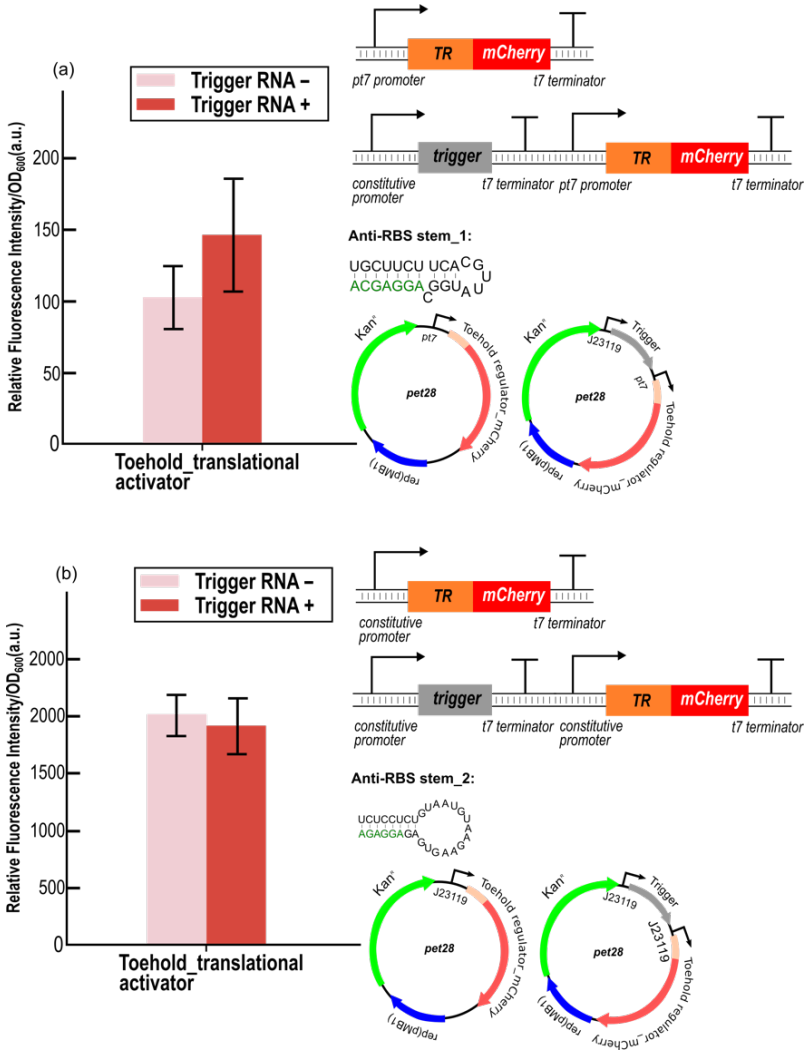


Figure S36. Relative fluorescence intensities of mCherry under the control of toehold-translational activator with constitutive promoter in *E. coli*, BL21 DE3. a), *in vivo* mCherry expression levels (fluorescence/$OD_{600}$ end level) under the control of translational toehold activator stem_1 in the absence and presence of trigger RNA. Toehold-translational activator is transcribed from a constitutive promoter (BBa_J23119), while trigger RNA is under the control of a pt7 promoter. b), *in vivo* mCherry expression levels (fluorescence/OD600 end level) under the control of translational toehold activator stem_1 in the absence and presence of trigger RNA. Here, both the translational activator and the trigger RNA are controlled transcribed from a constitutive promoter (BBa_J23119). Expression is measured in *E. coli*, BL21 DE3 with induction of the T7 RNA polymerase by 1mM IPTG. Error bars represent the standard deviation obtained from three biological replicates.

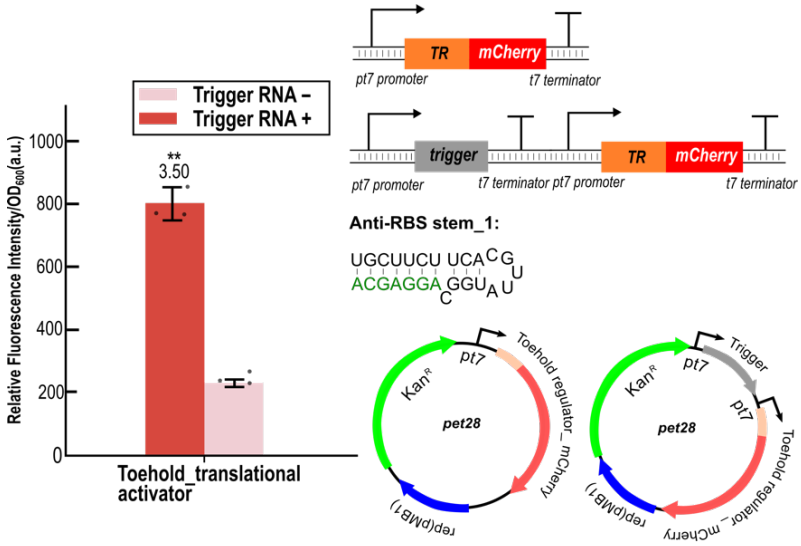


Figure S37. Test of the switch in LB medium. The bar plot shows the relative fluorescence intensities of mCherry (fluorescence/$OD_{600}$ end level) under the control of toehold-translational activator stem in *E. coli*, BL21 DE3 cultured in LB medium, in the absence and presence of trigger RNA. T7 RNA polymerase is induced by the addition 1mM IPTG. Error bars represent the standard deviation of three biological replicates.

## Addition of an Hfq binding site to the trigger RNA

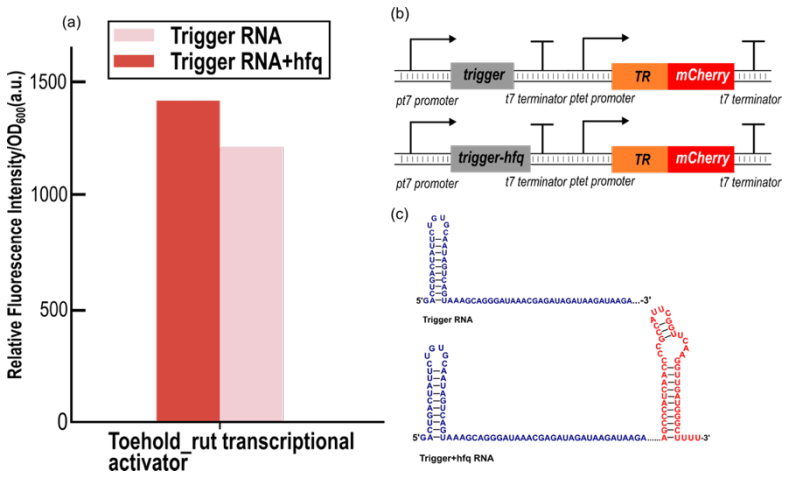


Figure S38. Effect of the RNA chaperone Hfq: Relative fluorescence intensities of mCherry under the control of toehold-anti-rut transcriptional activator in *E. coli*, BL21 DE3 cultured in M9 medium. a), *in vivo* mCherry expression levels (fluorescence/$OD_{600}$ end level, excitation/emission: 570/620 nm) under the control of transcriptional toehold activator stem_1 in the presence of a standard trigger RNA and a trigger RNA with an *hfq* hairpin that is recognized by Hfq. Hfq is expected to promote RNA-RNA interactions in *E. coli*, but only shows a minor effect in this experiment. b), Structure of the gene templates: toehold-activators were controlled by a *ptet* promoter and trigger RNAs were under the control of a pt7 promoter. c), Prediction of mRNA secondary structures of the trigger RNA and trigger RNA with *hfq* hairpin (red) using NUPACK. Expression is measured in *E. coli*, BL21 DE3 with induction of the T7 RNA polymerase by 1mM IPTG.

## A gate implementing IMPLY logic

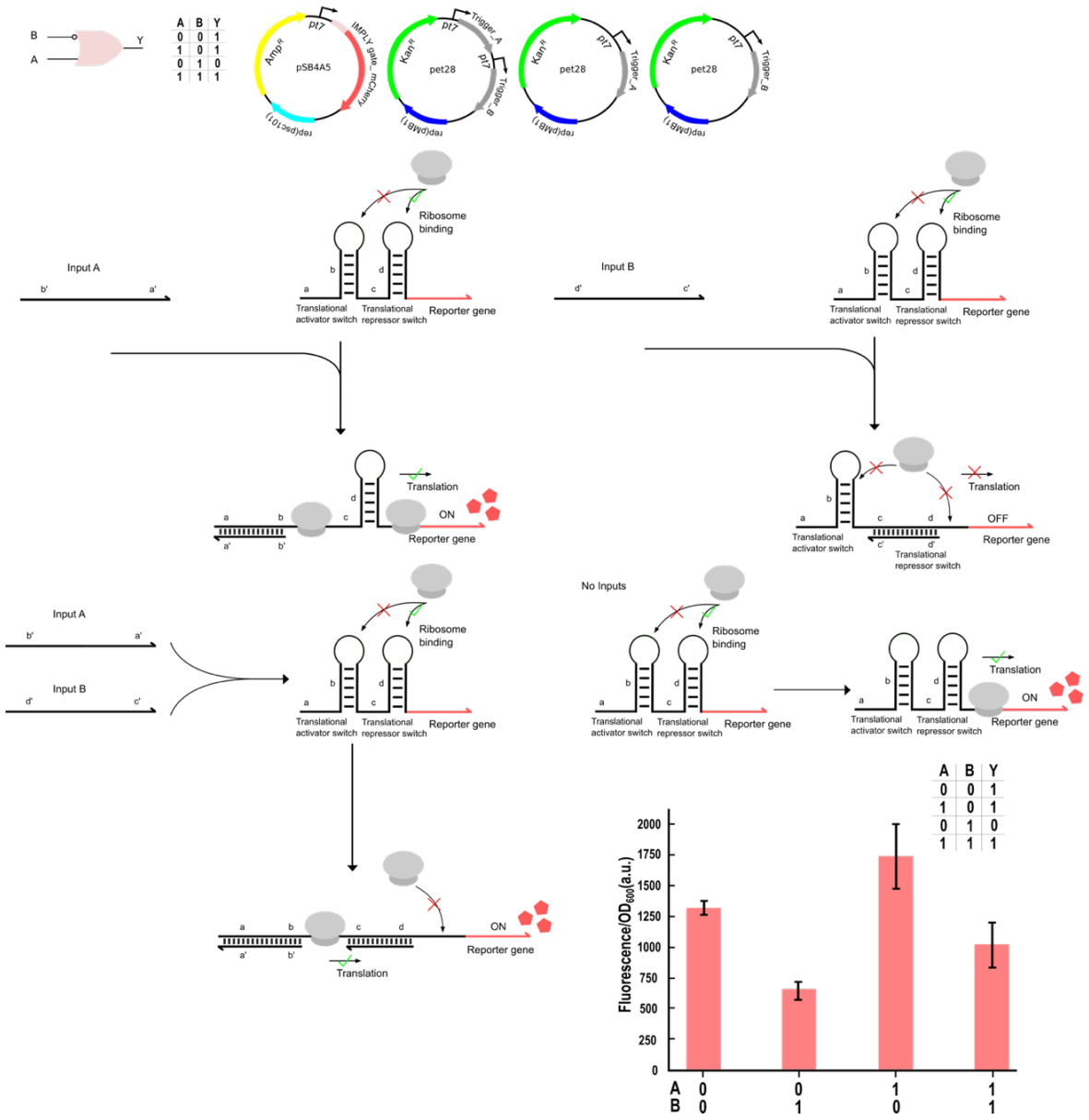


Figure S39. Schematic of a logic gate implementing an IMPLY logic function (B IMPLY A = (NOT B OR A)) by combining a translational activator and inhibitor. As indicated, the IMPLY gate RNA with a downstream mCherry sequence is under the control of a *pt7* promoter on a pSB4A5 low copy number plasmid, whereas trigger RNAs are transcribed from a *pt7* promoter on a pet28 high copy number plasmid with different trigger combinations. The IMPLY gate is comprised of a toehold-translational activator in series with a toehold-translation repressor. In the absence of trigger RNAs, the RBS of the translation activator is sequestered by the anti-RBS sequence, while the RBS of translation repressor is available for ribosome binding, resulting in translation of the downstream reporter gene. In the presence of trigger-A alone, translation is activated, in the presence of trigger-B alone, translation is inhibited as indicated in the schemes. When both triggers are present, both toehold hairpins are opened. Ribosomes initiating at the RBS exposed by the translation activator can move through the downstream duplex (5'-cd-3'/5'-d'c'-3') and translate the reporter gene. The normalized fluorescence end levels show that the gate behaves as an IMPLY gate, in principle, but there is a relatively strong leak in the presence of trigger B, which is probably the result of the strong individual leak of the translational repressor.

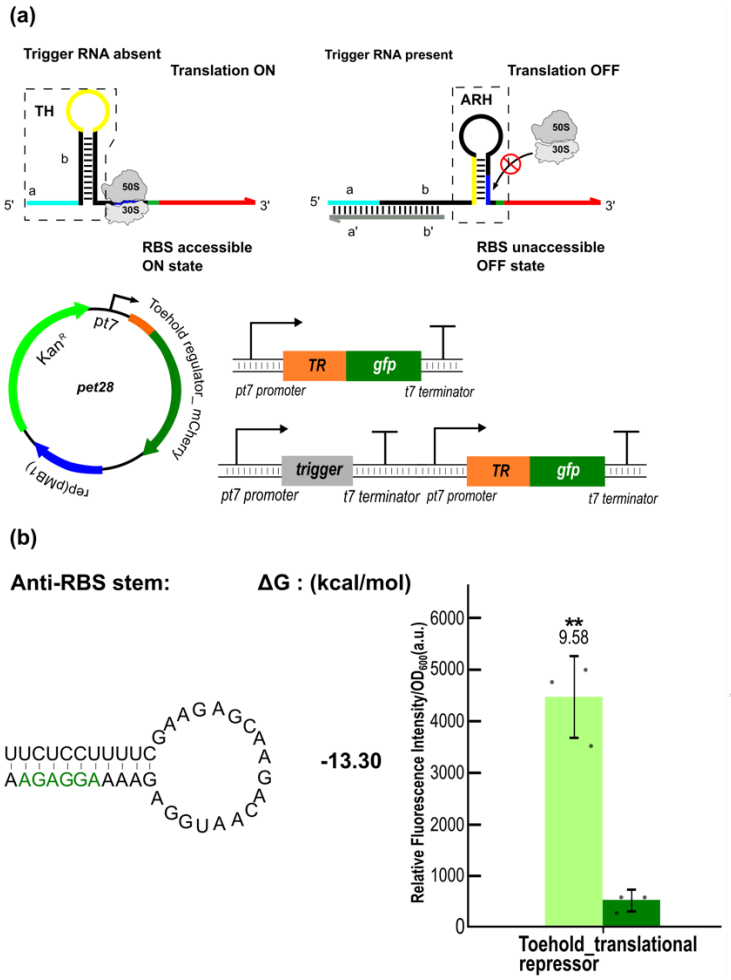


Figure S40. Relative fluorescence intensities of GFP under the control of toehold-translational repressor with *pt7* promoter in *Vibrio natriegens*, Vmax™ X2. a), in the absence of a trigger RNA (grey), the toehold hairpin (TH) constrains an anti-RBS sequence (yellow) within its loop region, the RBS is freely accessible and translational initiation is enabled. In the presence of trigger, toehold (light blue)-mediated invasion of the hairpin stem releases the anti-RBS, which leads to formation of anti-RBS hairpin (ARH) and sequestration of the RBS and thus translational repression.b, predicted secondary structure and total free energy of the anti-RBS hairpin. And *in vivo* GFP expression levels (fluorescence/$OD_{600}$ end level) under the control of translational toehold repressor stem in the absence and presence of trigger RNA. For both relative fluorescence/OD and fluorescence intensity data, Welch's t-tests were performed on each construct; *P < 0.05 and **P < 0.01 indicate conditions where the fluorescence/OD and fluorescence intensity for the Trigger RNA + condition is statistically significantly different from that of the trigger RNA- condition. Error bars in b), represent the s.d. from at least three biologically independent samples

## Strategies for the realization of switchable Spinach aptamers

Switchable aptamers similar to the ones investigated in this study have been previously developed using various switching mechanisms (Fig. S41). Our approach for switchable Spinach aptamers is same as in previous chapter, which combines toehold switches and naturally occurring riboswitches (Fig. S21).

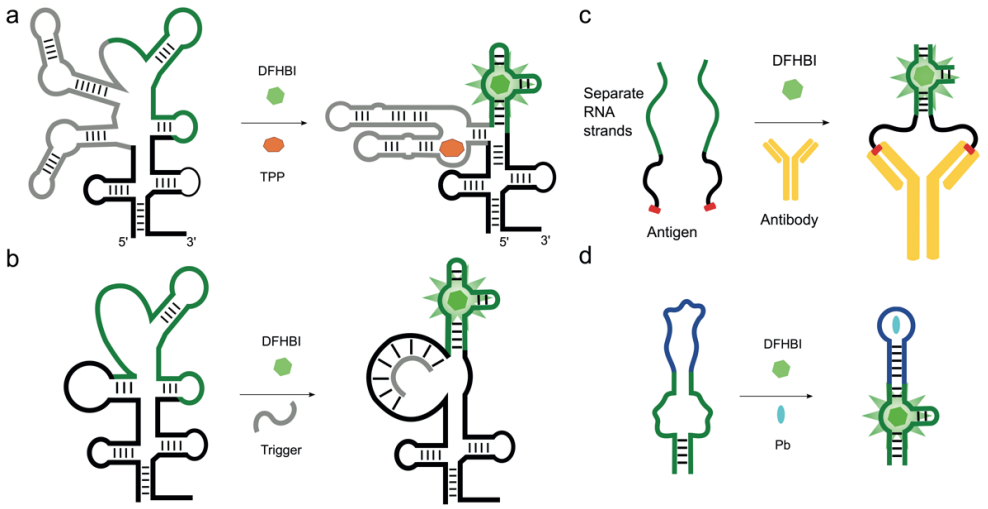


Figure S41. Comparison of the regulatory mechanisms of previously developed switchable Spinach aptamers: a) Switchable Spinach aptamer combined with the TPP riboswitch aptamer: In the presence of the ligand TPP, the TPP aptamer folds into an alternative conformation and induces a secondary structural rearrangement of the Spinach aptamer. This allows for the binding of DFHBI and the production of a fluorescence signal[200]. b), Switchable Spinach aptamer using small RNA as an input. In the presence of the small RNA, the aptamer folds into an alternative, allowing for the binding of DFHBI[181]. c) Switchable Spinach split aptamer: The Spinach aptamer is separated into two strands, each labeled with an antigen. In the presence of the corresponding antibody, the antigens bind to the antibody and promote the hybridization of the separate strands, allowing for the binding of DFHBI[167]. d) Allosteric Spinach aptamer combined with an RNA aptamer: In the presence of a ligand (here Pb), ligand-binding stabilizes the RNA aptamer, supports folding of the lower Spinach aptamer into an active conformation[163].

### Switchable Spinach aptamers based on TMSD

We constructed a series of switchable Spinach aptamers, which can either activate or repress the folding of the Spinach G-quadruplexes through trigger binding and TMSD. The toehold-Spinach activator comprises a toehold hairpin and an anti-Spinach hairpin (Fig. S42 a). The toehold hairpin consists of a 14 nucleotide long unpaired toehold region at the 5' end and a stem-loop, which confines the anti-anti-Spinach sequence (12 nucleotides) within the loop region. The sequence of the anti-Spinach hairpin (ΔG = -19.60 kcal/mol) base-pairs with critical nucleotides of the G-quadruplex. When trigger RNA binds to the toehold region, TMSD is initiated, which unwinds the toehold hairpin and releases the anti-anti-Spinach sequence. The trigger RNA and toehold riboregulator form an intermediate complex, from which the anti-anti-Spinach sequence invades the anti-Spinach hairpin. This results in the formation of an anti-anti-Spinach hairpin (ΔG = -19.70 kcal/mol) and the restoration of the Spinach aptamer, allowing for fluorogen binding.
In the case of the toehold-Spinach repressor, the toehold hairpin comprises a free toehold region (16 nucleotides) and an anti-Spinach sequence within the loop region (Fig. S42 b). The Spinach aptamer is initially accessible for fluorogen binding and thus is in a “fluorescence ON” state. When trigger RNA binds and TMSD unwinds the toehold hairpin, the anti-Spinach sequence is released, leading to a secondary structure rearrangement. This blocks critical nucleotides of the Spinach G-quadruplexes, resulting in the formation of an anti-Spinach hairpin and repression of fluorogen binding.

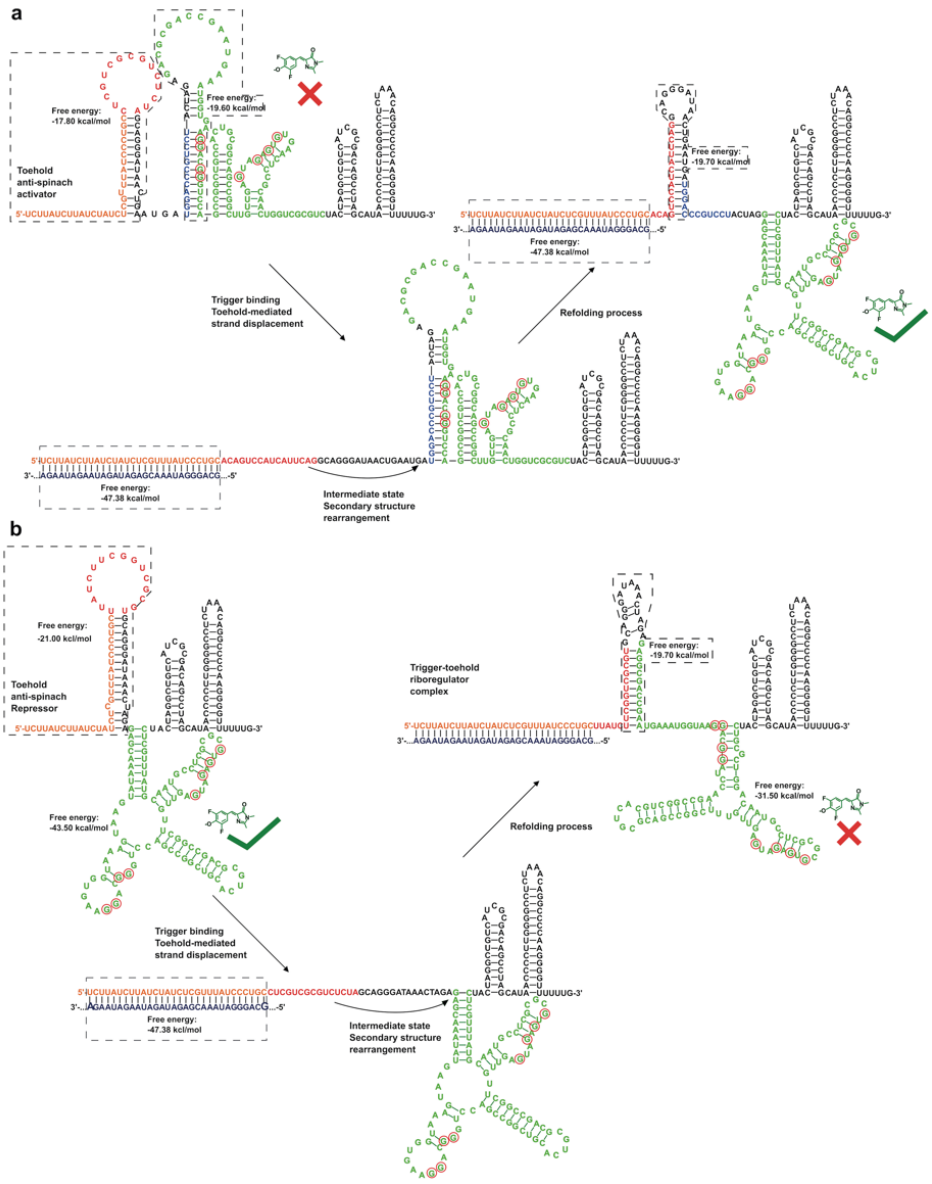


Figure S42. Scheme of toehold-Spinach switches with RNA sequences shown. a) The proposed mechanism of a toehold-Spinach activator shows a toehold hairpin that includes a free toehold region (light blue), incumbent sequence (orange), and unpaired anti-anti Spinach sequence (red). The anti-Spinach hairpin includes an anti-Spinach (blue) and Spinach aptamer (green) and the critical nucleotides for G-quadruplex are marked with red circles. When trigger RNA (dark blue) binds to the toehold-Spinach switch, an intermediate complex is formed which promotes refolding of the structure, leading to activation of the aptamer. b) Proposed mechanism of the toehold-Spinach repressor, which consists of a toehold hairpin with a free toehold region (light blue), an incumbent sequence (orange), and an unpaired anti-Spinach sequence (red). The transcription terminator at the 3' end is shown in black. The Spinach aptamer (green) is initially folded and accessible for fluorogen binding. When trigger RNA (dark blue) binds to the toehold riboregulator, the aptamer folds into an inactive state. The trigger RNA is protected from degradation via a 5' hairpin and a 3' terminator hairpin (which are not shown).

## Comparison of the performance of co-transcribed vs. purified toehold-Spinach switches

To evaluate experimental factors that influence the ON/OFF ratio determined for our switches, we performed additional tests with our best performing toehold Spinach switch containing the anti-spinach #3 sequence (see Fig. 16 a). We conducted these tests in a fluorescence

spectrophotometer by either co-transcribing the toehold-Spinach switch and its trigger RNA, or by using PAGE gel-purified toehold-spinach switch and trigger RNA and mixing (see Fig. 16 a).

For the co-transcriptional experiments, we added pyrophosphatase (NEB®) to remove any pyrophosphate generated during transcription, as insoluble magnesium pyrophosphate results in a cloudy suspension that yields unreliable fluorescence intensities (see Experimental procedures). The use of pyrophosphatase allowed us to accurately estimate the ON/OFF ratio of the switch by subtracting the blank signal (containing no pyrophosphate) from the fluorescence intensities obtained in the presence of trigger RNA (ON state) from the intensities obtained in the absence of trigger RNA (OFF state). Specifically, we co-transcribed toehold-spinach and trigger RNAs with the same concentrations of fluorogen and DNA templates and added pyrophosphatase (5 units/mL) to the IVT mix. After 4 hours of incubation, we observed an ON/OFF ratio of ≈ 200 from anti-spinach #3 (see Fig. S43 a).

Notably, when we tested the response of the gel-purified toehold-Spinach switch to the presence of trigger RNAs (0.5 µM each) in the presence of 0.5 µM fluorogen (BI) using fluorescence spectrophotometry in our transcription buffer (as described above for *in vitro* transcription), we observed an ON The significant difference in the ON/OFF ratios between the co-transcribed and gel-purified toehold-Spinach switch suggest that its extraction and purification results in a reduced switching performance. This observation can be attributed to several factors. On the one hand, co-transcriptional folding of nascent RNA limits the number of alternative folding pathways[201,202], which stabilizes the formation of local structures, such as the toehold hairpin and anti-spinach hairpin, as shown in (Fig. S43 a). This may reduce the generation of leaky fluorescence that is due to the competing ligand-bound conformation of the switch. Furthermore, the binding of trigger RNA and subsequent TMSD can occur during transcription, altering the original folding pathway and facilitating long-range conformational rearrangement, resulting in the formation of an anti-anti-spinach hairpin. On the other hand, the gel-purified toehold-spinach RNA undergoes denaturation during the purification process, and subsequently refolds in the presence of fluorogen, leading to a larger population of ligand-bound aptamers in the OFF state (Fig. S43 b). Maximum ON/OFF ratios for our toehold-Spinach switches are thus obtained under co-transcriptional conditions. /OFF ratio of only ≈ 20 after 10 minutes of incubation (Fig. S43 b).

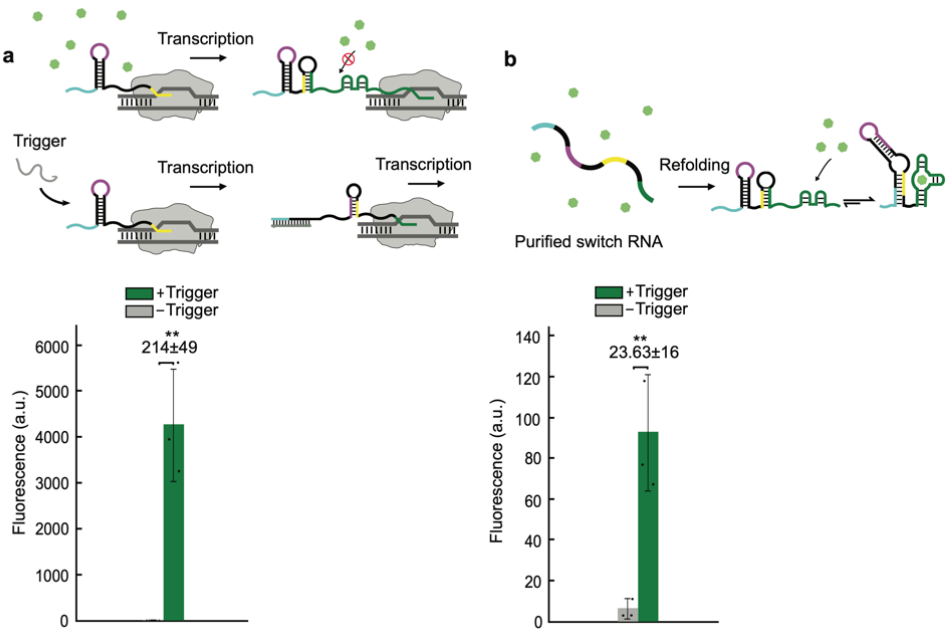


Figure S43. Depicts the fluorescence intensities obtained for the *in vitro* co-transcriptional and gel-purified toehold-spinach switch RNA. a), shows a schematic of the toehold-spinach activator RNA folding co-transcriptionally and the corresponding fluorescence values obtained in the presence and absence of trigger RNA. b) shows a schematic of the alternative folding path of toehold-spinach RNA and the corresponding fluorescence values obtained from the gel-purified toehold-spinach RNA. The fluorescence values shown in a), and b), are the mean values of three independent replicates, with error bars representing their standard deviation (s.d.). Welch's t-tests were performed for each construct, and *P < 0.05, **P<0.01 indicate conditions where the fluorescence intensity for the trigger RNA+ condition is statistically significantly different from that of the trigger RNA- condition.

## Switchable Mango aptamers based on TMSD

Toehold-Mango activators consist of a toehold hairpin and an anti-Mango hairpin in the OFF state (Fig. S44 a & b). The toehold hairpin comprises a 16 nt unpaired toehold region at the 5' end and a stem-hairpin which confines the anti-anti Mango sequence (16 nt) in its loop region. When trigger RNA binds to the toehold region, it initiates TMSD which unwinds the toehold hairpin (TH) and releases the anti-anti Mango (AAM) sequence. This allows the anti-anti Mango sequence to base pair with the anti-Mango hairpin (AMH), forming an anti-anti Mango hairpin (AAMH) and restoring the Mango aptamer to its active state. In the case of the toehold-Mango repressor, the toehold hairpin comprises a free toehold region (16 nt) and an anti-Mango sequence within the loop region (Fig. S44 b). When trigger RNA binds to the toehold region, TMSD unwinds the toehold hairpin, and the anti-Mango sequence is released. This leads to a refolding process in which the anti-Mango sequence blocks critical nucleotides of a Mango G-quadruplex, resulting in the formation of anti-Mango hairpins and repression of fluorogen binding.

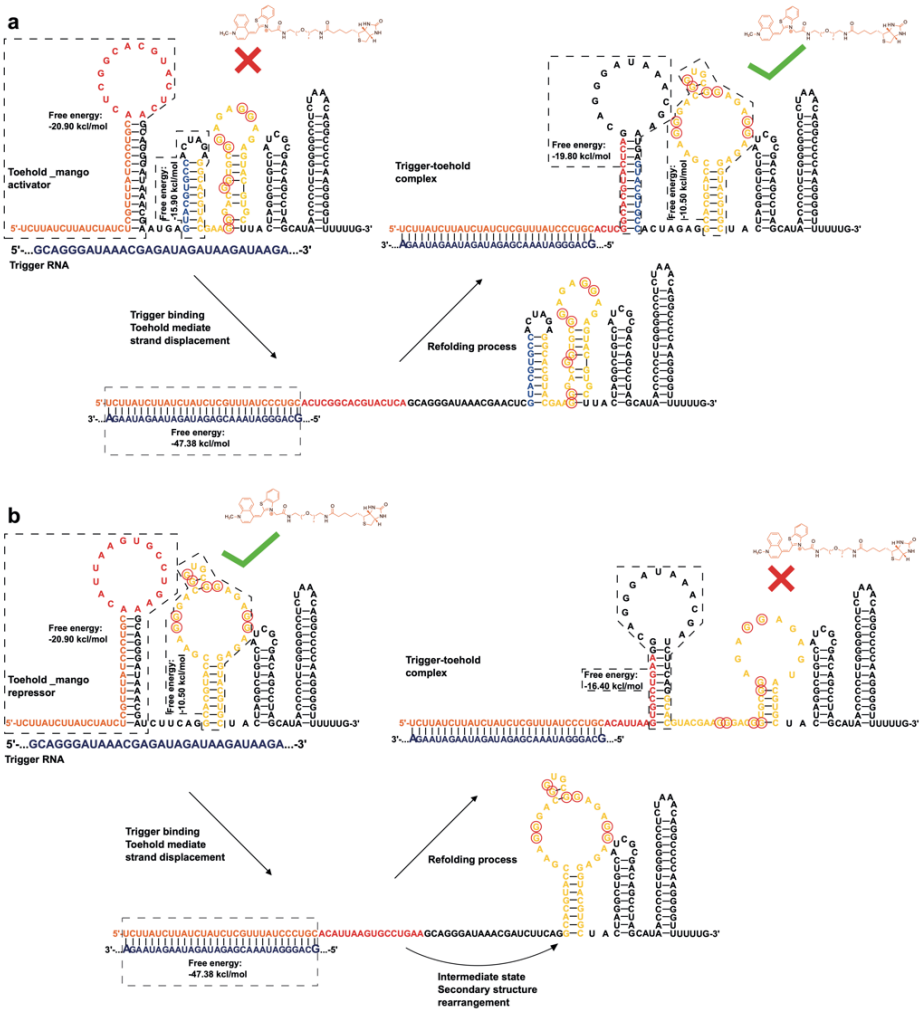


Figure S44. Schemes of toehold Mango switches with RNA sequences and switching mechanisms. a) Proposed mechanism of a toehold-Mango activator. The toehold hairpin includes a free toehold region (orange), an incumbent sequence, and an unpaired anti-anti Mango sequence (red). The anti-Mango hairpin includes an anti-Mango (blue) and Mango aptamer (yellow). The transcription terminator at the 3' end is shown in black. Critical nucleotides for G-quadruplex formation are marked with red circles. The trigger RNA (dark blue) binds to the toehold Mango switch, forming an intermediate complex which promotes refolding of the structure into an active Mango conformation. b) Proposed mechanism of a toehold-Mango repressor. The toehold hairpin has a free toehold region (orange), an incumbent sequence, and an unpaired anti-Mango sequence (red). The Mango aptamer (yellow) is initially folded and accessible for fluorogen binding. Trigger RNA (dark blue) binds to the toehold Mango repressor, which induces refolding of the aptamer into a fluorescent OFF state. The trigger RNAs are protected from degradation via a 5' hairpin and the 3' terminator hairpin (not shown). The free energy values were calculated using NUPACK.

## *In* vitro expression of toehold-Mango switches

We designed different AM sequences for the toehold-Mango switches and corresponding trigger RNAs, which were characterized *in vitro* as described in the main text for the Spinach aptamers. Overall, the toehold-Mango activator showed lower ON/OFF ratios compared to the toehold-Spinach activators, which is due to a higher fluorescence in the OFF state. The toehold-Mango repressor performed similarly, except for sequence AM_5 which had an ON/OFF ratio less than 2 times.

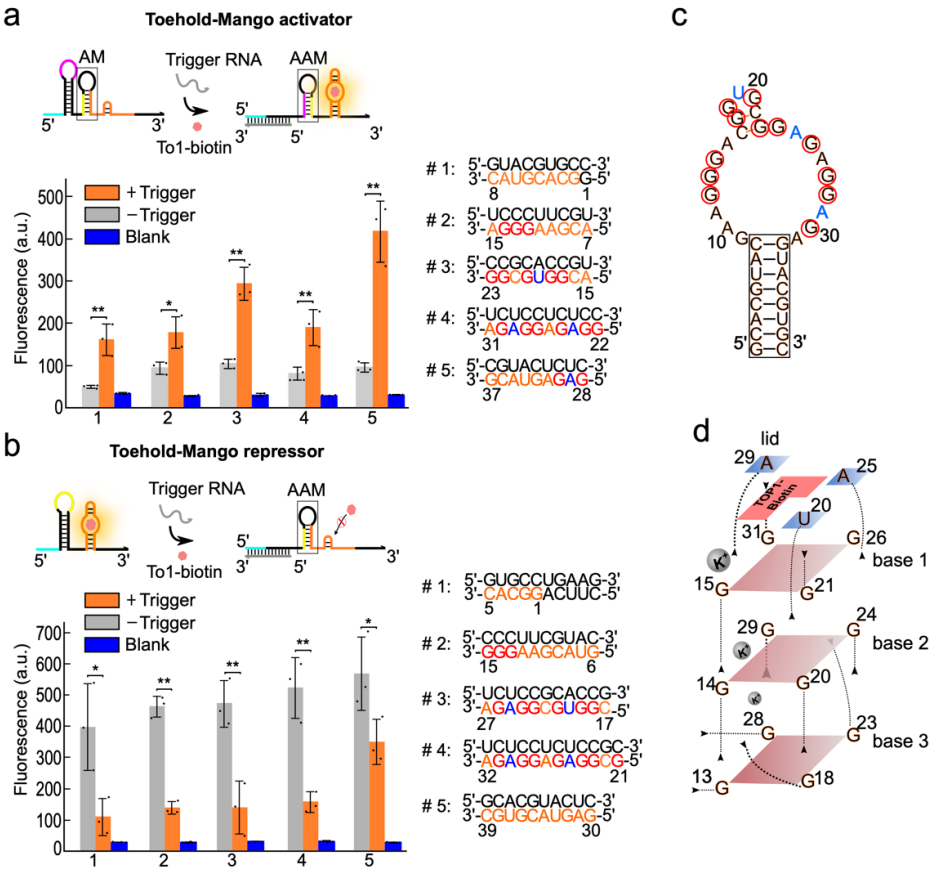


Figure S45. Design and characterization of different toehold-Mango switch variants. a), Schematic of switching process of toehold-Mango activator (cf. Fig. S44) and fluorescence values obtained for different versions in the presence and absence of trigger RNA. Sequences of the corresponding anti-Mango hairpin stems are shown on the right. b), Schematic of switching process of toehold-Mango repressor and fluorescence values in the presence and absence of trigger RNA. Sequences of the anti-Mango hairpin stems for the different version are shown on the right. c), NUPACK prediction of the secondary structure of the Mango aptamer without G quadruplexes. Critical guanine nucleotides are marked with red circles and triplex lid nucleotides are marked in blue. The lower stem of the Mango is highlighted with a box. d), Scheme of the G-quadruplex stack involved in binding the TO1-biotin fluorophore. Numbers indicate the nucleotide numbers corresponding to the structure in c). Fluorescence values in a) and b) are given as mean values of three independent replicates and error bars represent their standard deviation (s.d.). For the fluorescence intensities, Welch's t-tests were performed for each construct; *P < 0.05, **P<0.01 indicate conditions where the fluorescence intensity for the trigger RNA+ condition is statistically significantly different from that of the trigger RNA- condition.

## *In vitro* expression of adenine-Spinach switches

As for the guanine-Spinach switches, we varied the stability of the P1 stem of the adenine aptamer to optimize the switching behavior of adenine-dependent Spinach switches. P1 stem #1 and #2 have 6 base-pairs with the AS sequence and we utilized a wobble base-pair (G-U) to adjust their free energy slightly. As shown in Fig. S7 A, these minor adjustments of the stem lead to dramatic changes in performance. P1 stem #2, which differs from #1 by an A-U instead of a G-U bp shows a much better switching behavior than #1 and a low OFF fluorescence close to the Blank reference. P1 stems #3 and #4 have an additional base-pair to reduce the free energy of the adenine aptamer structure (to -12.98 kacl/mol and -13.28 kcal/mol in the absence of adenine)

in order to increase the fluorescence signal in the ON state. Surprisingly, we observed strongly increased leak signals from these versions in the OFF state and P1 stem #3 does not seem to switch at all.

In addition, we also varied the P1 stem of the adenine-Spinach repressors by adding additional base-pairs and mismatches (Fig. S46 b). P1 stem #4, which has the lowest free energy of -12.38 kcal/mol only exhibits an ON/OFF ratio of ≈ 3. Based on the sequence of P1 stem #4, we first inserted a mismatch (C-C) on the stem to reduce its stability (version #1), which indeed resulted in an increased ON/OFF ratio. Further adjustments (#2 and #3) led to reduced leak, but also reduced maximum fluorescence intensities.

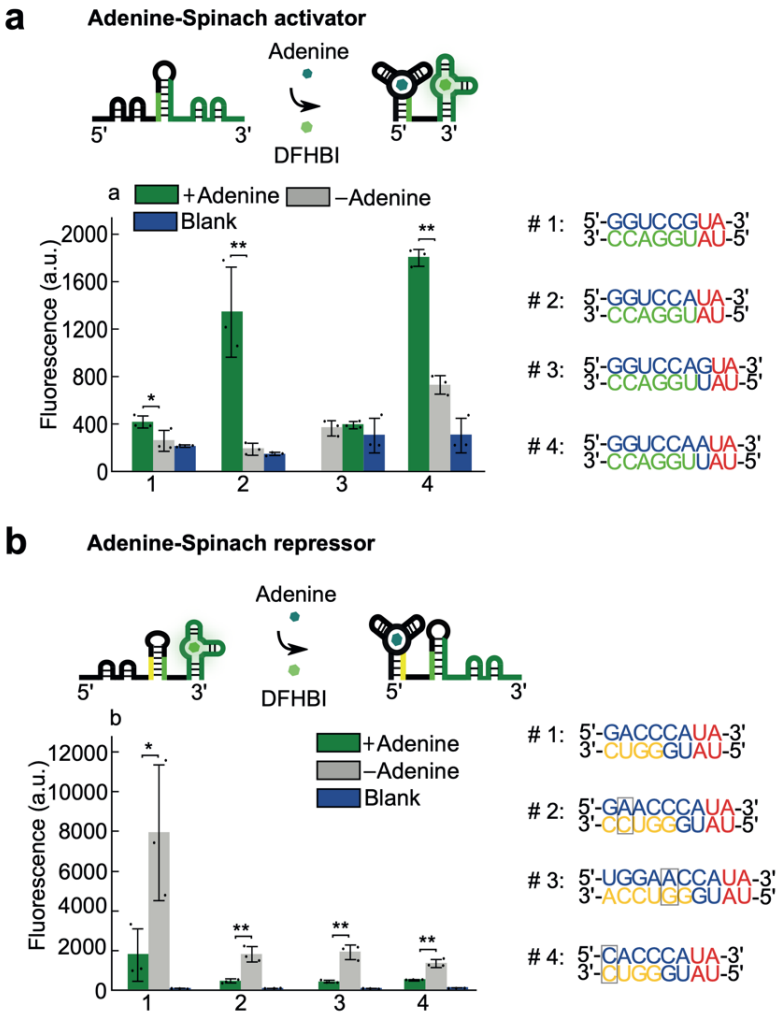


Figure S46. Fluorescence intensities obtained for different versions of adenine-Spinach activators a), and repressors b), utilize different P1 stem sequences in the presence of 0 µM or 50 µM guanine, respectively. Shown are the mean values of for three independent measurements, error bars represent their s.d. (statistical significance of the differences was determined via Welch's t-test, **p-value < 0.01). The corresponding P1 sequences are shown on the right. Conserved nucleotides are colored in red, variable nucleotides are shown in blue. Sequestered nucleotides of the Spinach aptamer are shown in green, while anti-Spinach nucleotides are colored yellow. Sequence mismatches on the P1 stem are highlighted with black boxes. The blue bars indicate the background level of a blank measurement containing TX mix and DFHBI.

## Orthogonality test for eight different toehold-stem sequences

Our riboswitch-inspired design for toehold-Spinach switches eliminates any sequence restriction for the trigger RNAs, allowing it to detect different RNA inputs by simply modifying the toehold and stem sequences on the toehold hairpin. To demonstrate the high orthogonality of our design, we created eight different toehold-Spinach activators and their corresponding trigger RNAs. We co-transcribed each activator with different triggers under the same conditions and measured the relative fluorescence intensities using a microplate reader. The results, shown in a heat-map (Fig. S47), indicate that our design exhibits high orthogonality when used with these different triggers.

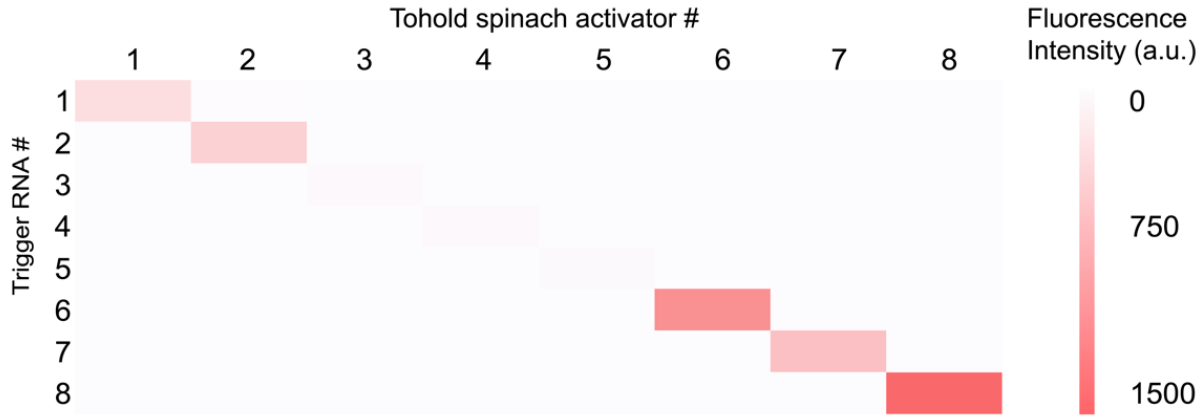


Figure S47. The heat map illustrates the toehold-Spinach activator crosstalk, with the color scale indicating the relative fluorescence intensities measured using a microplate reader. These measurements were taken 4 hours after incubation at 37°C for 64 combinations of trigger and switch. The scale ranges from the minimum to the maximum values of the fluorescence signal emitted by the toehold-Spinach activator.

*In vivo* switching of a toehold-Broccoli switch in *E. coli*

In addition to the Spinach and Mango switches, we also developed a toehold-Broccoli activator for potential *in vivo* sensor applications in *E. coli* BL21 DE3 cells. The broccoli aptamer is smaller than the previously developed Spinach aptamer[203] and has a higher level of fluorescence in *E. coli* even without a stabilizing tRNA modification[203,204]. To further enhance the *in vivo* stability of the switch, complementary clamp sequences were added to the 3' end of the switch RNA and the 5' end of the trigger RNA (Fig. S48 a). The clamp sequence on the trigger RNA can thus base-pair with the complementary sequence on the switch RNA, which forms a closed conformation with an internal loop and increases thermodynamic stability. We transformed a plasmid carrying the toehold-broccoli activator and its corresponding trigger RNA into *E. coli* BL21 DE3 cells and characterized the activator's switching behavior using a microplate reader and fluorescence microscopy (Olympus IX81) (Fig. S48 b-d). The *E. coli* cells were initially cultured in LB medium and M9 medium, followed by induction of T7 RNAP with 1 mM IPTG. Measurement was started after adding 100 µM DFHBI-1T. The results show that the toehold-broccoli activator can indeed be activated by the trigger RNA and generate GFP-like fluorescence inside the *E. coli* cells, resulting in an ON/OFF ratio of ≈ 4 after 2 hours of incubation (Fig. S48 c). Compared to other work using FLAPs for *in vivo* imaging, we only obtained modest fluorescence levels, which is likely caused by the low *in vivo* stability and efficiency of hybridization between trigger and switch.

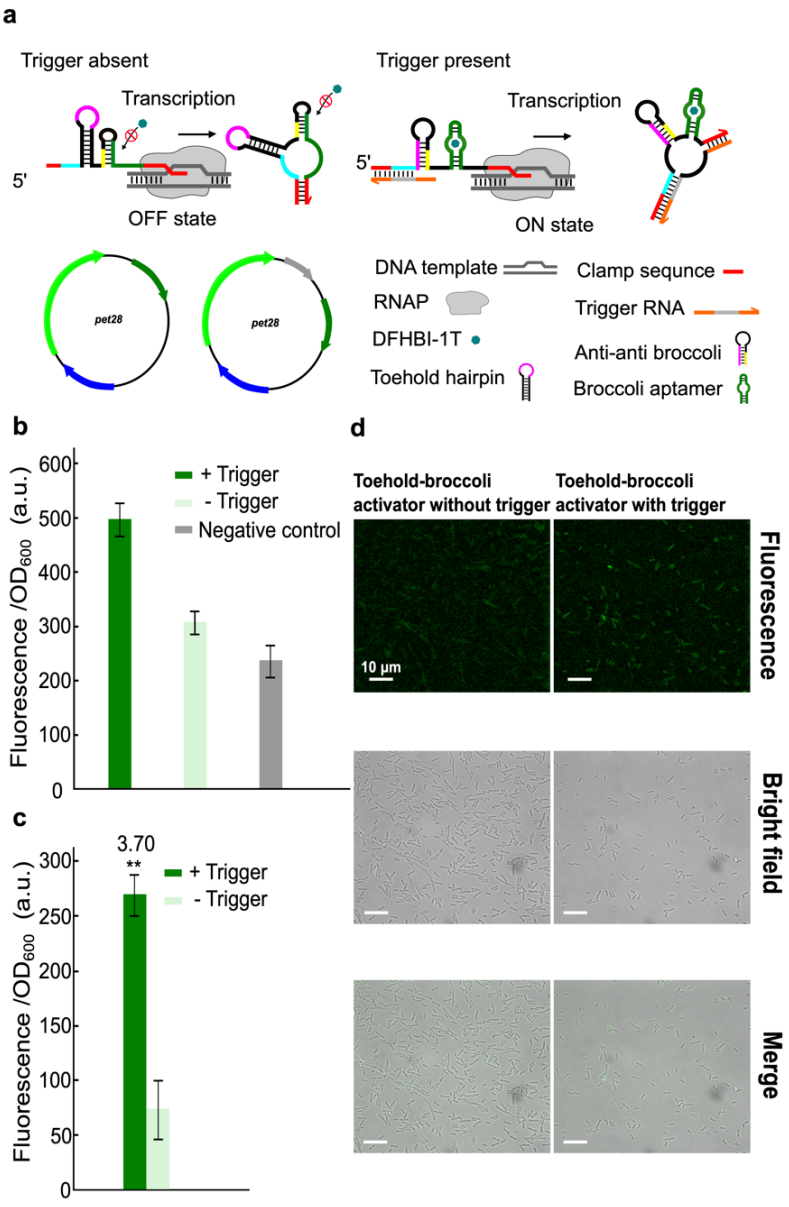


Figure S48. Design and Characterization of the Toehold-Broccoli Switch *in vivo*: a) Scheme of the switching process for the Toehold-Broccoli activator. b) Fluorescence intensities normalized by absorbance ($OD_{600}$), obtained from the Toehold-Broccoli switch in the presence and absence of trigger RNA in *E. coli* BL21 DE3 cells, measured using a plate reader in comparison with a negative control strain at 37°C. c) Same data as in b), but with signal level corrected by subtraction of the fluorescence/OD600 of the negative control. Each data point is the average of three independent experiments, with error bars indicating the standard deviation. Welch's t-tests were performed for each construct, and **P<0.01 indicates conditions where the fluorescence intensity for the + trigger condition is statistically significantly different from that of the - trigger. d) Fluorescence microscopy images of *E. coli* BL21 DE3 cells expressing the Toehold-broccoli switch with trigger RNA (right column) compared to cells expressing the Toehold-broccoli switch alone (left panel). Also, bright field and overlay images are shown. Scale bars are 10 µm.

## Switching toehold-spinach sensor with viral RNA sequence

To explore the potential application of toehold-spinach switches in the context of *in vitro* RNA sensing, we designed and constructed five Toehold-Spinach sensors that respond to subsequences of the SARS-CoV-2 genome. Specifically, the sensors were designed by adjusting their toehold and stem regions to contain sequences complementary to subsequences of the *orf1ab* polyprotein, *orf7*, spike protein, membrane glycoprotein and nucleocapsid phosphoprotein coding sequence (Fig. S49 a,)[205].

In order to assess their performance, SARS-CoV-2 Spinach sensors were *in vitro* transcribed and then mixed with a corresponding PAGE-purified target RNA that contains a 30 nt subsequence of the genome. As shown for the example of the membrane glycoprotein sequence, the Toehold-Spinach sensor allowed detection of target RNAs in the range 10 – 100 nM (Fig. S49 b). With a size of 30 kilobases, the SARS-CoV-2 genome contains many potential target sites for our sensors. Targeting several of these simultaneously can be used to increase the sensitivity of our approach. In order to explore this idea further, we constructed and purified a multi-target RNA that contains all the above mentioned five coding sequences (Fig. S49 c). As before, we *in vitro* transcribed single and mixed COVID-spinach sensors using 10 nM DNA templates for each and added purified multi-target RNA (500 nM) to the IVT mix. As shown in Figure 49 d, the single *orf7* sensor rapidly responds to the multi-target RNA and reaches its final level after ≈ 20 min of incubation. As desired, the multi-site binding assay provides an up to five-fold enhanced fluorescence signal, albeit with slowed kinetics.

Several studies have reported the viral loads of SARS-CoV-2 in nasal swab are in a range of 7.9 x $10^3$ to 3.7 x $10^8$ copies/mL (0.01 - 614 fm)[206], which is far below the minimum detection concentration of our sensors. Due to the detection limitation of the microplate reader, our Toehold-Spinach sensors alone were incapable of detecting such low RNA concentrations. To demonstrate the applicability of our toehold-spinach sensors for diagnostics of clinically relevant RNA concentrations of SARS-CoV-2, in principle, we combined our *orf7* sensor with the isothermal RNA amplification process known as NASBA (Nucleic acid sequence-based amplification)[207]. To this end, we redesigned and extended the trigger sequence (see Sequence and Primers list) to improve the efficiency of both reverse transcription and TMSD. Purified trigger RNAs were added to the NASBA master mix at a concentration of 10 pM, 1 pM, 100 fM respectively and incubated at 41°C for 90 minutes. The resulting amplified RNA products (3) were then added to the transcription mix containing the *orf7* sensor and incubated at 37°C. Within 1h, we observed a strong increase in fluorescence, whereas a negative control without RNA inputs did not show any signal (Fig. S49 f).

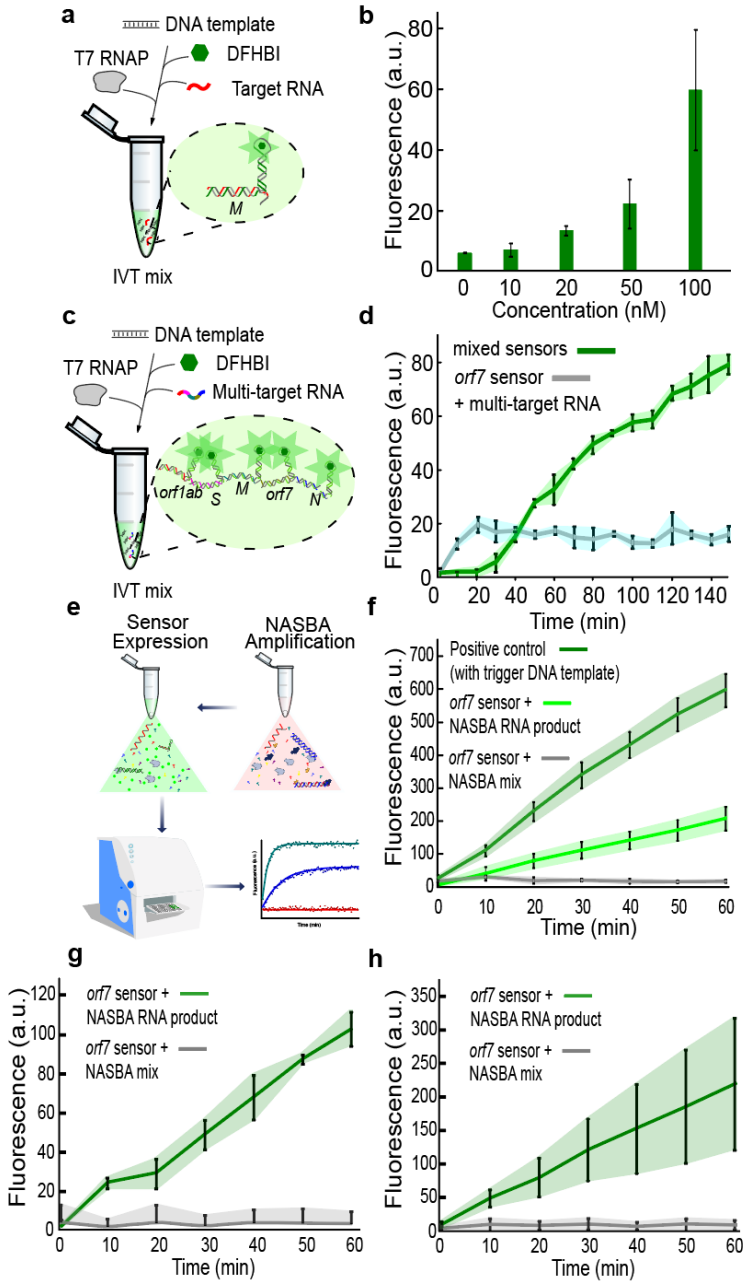


Figure S49. Design and characterization of toehold-spinach based RNA sensors. a) Detection of a single target RNA derived from the M gene of the SARS-CoV-2 genome. The toehold-spinach switch is transcribed from its DNA template in an *in vitro* transcription (IVT) reaction in the presence of DFHBI, after which the purified target RNA is added. b) Fluorescence signal recorded from a FLAP in the presence of different concentrations of target RNA after 1h of incubation at 37 °C. c) Multiple sensors can be utilized in parallel to interact with a long target RNA with several binding sites. d) Comparison of fluorescence intensity recorded with mixed SARS-CoV-2-spinach sensors compared to a single sensor as a function of time in the presence of 500 nM multi-target RNA at 37 °C. e) Simple workflow schematic for *in vitro* detection. RNA samples are isothermal amplified via NASBA and added to the transcription mix of sensor. The detection of amplified trigger RNA is indicated by fluorescence under microplate reader. f) Detection of *orf7* trigger RNA with input concentration of 10 pM after NASBA amplification in 1h of incubation at 37 °C. g), Detection of *orf7* trigger RNA with input concentration of 1 pM after NASBA amplification in 1h of incubation at 37 °C h), Detection of orf7 trigger RNA with input concentration of 100 fM after NASBA amplification incubation at 37 °C. The fluorescence outputs of the sensors are shown as the mean values of background subtracted fluorescence levels for three independent measurements, error bars represent their s.d.

Expanding the scope of our design approach, we formulated a switchable Cas12a gRNA construct, wherein the incorporation of a toehold hairpin (TH) at its 5' end was undertaken, comprising a 14-nt single-stranded toehold along with a 16-bp stem configuration (Fig. S50). We separated Cas12a handle sequence 4nt downstream away from the toehold hairpin. Within the loop region of this hairpin, an anti-handle sequence was strategically inserted. The interaction with Cas12a prompts the cleavage of both the designated target DNA as well as the FQ-labeled single-stranded DNA (ssDNA) probe (100nM), thereby inducing an increasing of fluorescence signal. In the presence of trigger, toehold (light blue)-mediated invasion of the hairpin stem releases the anti-handle, which leads to formation of anti-handle hairpin (AHH) and sequestration of the Cas12a handle. Consequently, the nuclease activity is repressed in this conformation.

According to the results of *in vitro* cleavage assay, it is evident that the toehold-sgRNA repressor exhibits strong repression of Cas12a nuclease activity when the trigger strand is present, leading to a notable ON/OFF ratios of $\approx$ 3 (Fig. S50 b). Additionally, the ssDNase assay illustrates a pronounced decrease in fluorescence signal from cleaved ssDNA probe within a 30 min following the addition of the trigger strand in stark contrast to the signal emanating from the non-repressed toehold-sgRNA complex.

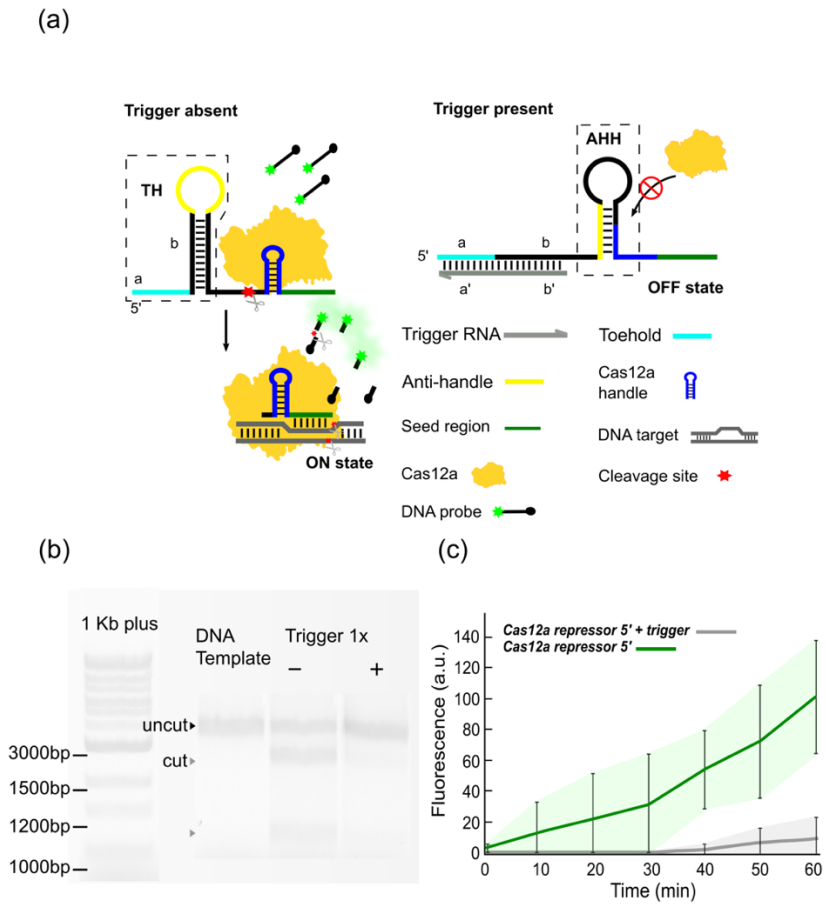


Figure S50. Design and Characterization of the Toehold-sgRNA repressor *in vivo*. a), schematic of Switchable sgRNA. In the absence of a trigger RNA (grey), the toehold hairpin (TH) confines an anti-handle sequence (yellow) within its loop region. The Cas12a handle is accessible by Cas12a, enabling it to bind to the target DNA and catalyzing the cleavage of the FQ-labeled (fluorophore (FAM)-quencher (Black Hole Quencher 1)-labeled) ssDNA

probe. This cleavage process removes quencher moieties and generates detectable fluorescent signals. When a trigger is present, the toehold (light blue)-mediated invasion of the hairpin stem leads to the release of the anti-handle. This release causes the formation of an anti-handle hairpin (AHH) and the sequestration of the handle, effectively repressing the formation of the complex. b), DNA cleavage assay. Target plasmid DNA was PCR amplified and used as substrates for digestion with the toehold-Cas12a repressors complex for 30 min incubation at 37°C. Experimental controls were conducted with the full LbCas12a/sgRNA duplex (+) or with only one component of the duplex (-). Gel electrophoresis separated the resulting products, and depending on the presence of trigger RNA, the cleavage of the linear template resulted in two cuts: fragments _A and B (grey arrow). c), detection of switchable sgRNA using Cas12a-FQ-reporter assay. Toehold-sgRNA repressor RNA was transcribed with and without trigger strand *in vitro* and reacts in Cas12a-FQ-reporter mix for 1h incubation at 37°C. The fluorescence outputs of the sensors are shown as the mean values of background subtracted fluorescence levels for three independent measurements, error bars represent their s.d

# Acknowledgments

I am profoundly grateful to all those who have provided unwavering support throughout my six-year journey towards completing my Ph.D. This remarkable adventure has been a truly memorable experience.

First and foremost, my heartfelt thanks go to Prof. Simmel for granting me the invaluable opportunity to embark on this thesis journey in Munich. Your guidance, kindness, and unwavering patience have not only shaped my research but also my personal growth. Your support has been my pillar of strength, helping me overcome countless challenges and persevere in my research. Our daily interactions during lunch and our discussions have been invaluable, offering me not only scientific knowledge but also profound insights into becoming a better scientist.

I extend my sincere gratitude to Ms. Susanne Kinzel, Ms. Daniela Härtwig, and Ms. Helene Budjarek for their constant assistance. Their contributions have made administrative and lab tasks appear far simpler than they are.

I would like to express my appreciation to Dr. Andrea Mückl, Dr. Günter Kugler, Dr. Eva Pensa, and Dr. Sandra Sagredo, Dr. Hari Raj Singh, Dr. Tobias Pirzer for their patient mentorship during my Ph.D. journey, especially Dr. Andrea Mückl who always give me many helps and cares during these years.

A heartfelt thank you goes out to GRK2062, the exceptional graduate school that exceeded my expectations. Beyond the generous research and travel support, I appreciate the help and guidance from Dr. Aurore Dupin and Dr. Mattheus Schwarz-Schilling during my early days with the community. Special thanks to Beate for her tireless efforts in organizing seminars, conferences, and workshops, which enriched our academic experience.

To my fellow students at GRK, thank you for the enjoyable meetings and enriching conversations. These interactions have been both fun and intellectually stimulating.

I am deeply grateful to the members of the E14 lab, a vibrant community of passionate individuals. I extend my heartfelt thanks to Louis, Lukas O, Thomas. F, Henning, Katze, Markus, and others

for their insightful discussions and invaluable advice. Your unwavering belief in my abilities has been a driving force. Thank you for your down-to-earth guidance when I needed it. I also thank my students, Patrick, Marta, Maria, Adli, and everyone else for the shared learning experiences during our projects.

I extend my appreciation to the E22 groups, particularly Prof. Rief and Andreas Walbrun, for their outstanding collaboration. During challenging times in my project, Andreas actively discussed solutions and encouraged me to overcome obstacles. His support in data analysis and single-molecule techniques has been tremendously beneficial. I am also thankful to Prof. Sulc and Dr. Matthies from the University of Arizona for their guidance in oxDNA simulations, which has expanded my skill set in simulations and modeling.

Lastly, heartfelt gratitude goes to my families, my parents, granny and other relatives and girlfriend. Their unwavering support, especially during the challenging times of the COVID-19 pandemic when I couldn't return home, has been my source of strength. Without their support, I could not have successfully completed my thesis.

Thank you all for being an integral part of my Ph.D. journey, contributing to its success and making it a truly memorable experience.

# Reference

1 Rodríguez, I. *et al.* A specific subdomain in φ29 DNA polymerase confers both processivity and strand-displacement capacity. *Proceedings of the National Academy of Sciences* **102**, 6407-6412 (2005).

2 Zhou, W. *et al.* A CRISPR–Cas9-triggered strand displacement amplification method for ultrasensitive DNA detection. *Nature communications* **9**, 5012 (2018).

3 Wang, T., Hellmer, H. & Simmel, F. C. Genetic switches based on nucleic acid strand displacement. *Current Opinion in Biotechnology* **79**, 102867 (2023).

4 Yurke, B., Turberfield, A. J., Mills, A. P., Simmel, F. C. & Neumann, J. L. A DNA-fuelled molecular machine made of DNA. *Nature* **406**, 605-608 (2000).

5 Fu, T. *et al.* DNA-based dynamic reaction networks. *Trends in biochemical sciences* **43**, 547-560 (2018).

6 Qian, L. & Winfree, E. in *International Workshop on DNA-Based Computers.* 114-131 (Springer).

7 Siu, K.-H. & Chen, W. Riboregulated toehold-gated gRNA for programmable CRISPR–Cas9 function. *Nature chemical biology* **15**, 217-220 (2019).

8 Hanewich-Hollatz, M. H., Chen, Z., Hochrein, L. M., Huang, J. & Pierce, N. A. Conditional guide RNAs: programmable conditional regulation of CRISPR/Cas function in bacterial and mammalian cells via dynamic RNA nanotechnology. *ACS central science* **5**, 1241-1249 (2019).

9 Collins, S. P., Rostain, W., Liao, C. & Beisel, C. L. Sequence-independent RNA sensing and DNA targeting by a split domain CRISPR–Cas12a gRNA switch. *Nucleic Acids Research* **49**, 2985-2999 (2021).

10 Oesinghaus, L. & Simmel, F. C. Switching the activity of Cas12a using guide RNA strand displacement circuits. *Nature communications* **10**, 1-11 (2019).

11 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).

12 Pardee, K. *et al.* Rapid, low-cost detection of Zika virus using programmable biomolecular components. *Cell* **165**, 1255-1266 (2016).
13 Kim, J. *et al.* De novo-designed translation-repressing riboregulators for multi-input cellular logic. *Nature chemical biology* **15**, 1173-1182 (2019).
14 Wang, T. & Simmel, F. C. Riboswitch-inspired toehold riboregulators for gene regulation in Escherichia coli. *Nucleic acids research* **50**, 4784-4798 (2022).
15 Broker, T. R. & Lehman, I. Branched DNA molecules: intermediates in T4 recombination. *Journal of molecular biology* **60**, 131-149 (1971).
16 Radding, C. M., Beattie, K. L., Holloman, W. K. & Wiegand, R. C. Uptake of homologous single-stranded fragments by superhelical dna: Iv. branch migration. *Journal of molecular biology* **116**, 825-839 (1977).
17 Green, L. S. *et al.* Inhibitory DNA ligands to platelet-derived growth factor B-chain. *Biochemistry* **35**, 14413-14424 (1996).
18 Panyutin, I. G. & Hsieh, P. The kinetics of spontaneous DNA branch migration. *Proceedings of the National Academy of Sciences* **91**, 2021-2025 (1994).
19 Hellyer, T. J. & Nadeau, J. G. Strand displacement amplification: a versatile tool for molecular diagnostics. *Expert review of molecular diagnostics* **4**, 251-261 (2004).
20 Zhang, D. Y. & Winfree, E. Control of DNA strand displacement kinetics using toehold exchange. *Journal of the American Chemical Society* **131**, 17303-17314 (2009).
21 Jiang, Y. S., Bhadra, S., Li, B. & Ellington, A. D. Mismatches improve the performance of strand-displacement nucleic acid circuits. *Angewandte Chemie* **126**, 1876-1879 (2014).
22 Lakin, M. R., Youssef, S., Polo, F., Emmott, S. & Phillips, A. Visual DSD: a design and analysis tool for DNA strand displacement systems. *Bioinformatics* **27**, 3211-3213 (2011).
23 Broker, T. R. An electron microscopic analysis of pathways for bacteriophage T4 DNA recombination. *Journal of molecular biology* **81**, 1-16 (1973).
24 Kaiser, C., Goldman, D. & Bustamante, C. Single Molecule Analysis of Protein Folding on the Ribosome. *Biophysical Journal* **100**, 28a (2011).
25 Qu, X. *et al.* The ribosome uses two active mechanisms to unwind messenger RNA during translation. *Nature* **475**, 118-121 (2011).
26 Collin, D. *et al.* Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature* **437**, 231-234 (2005).
27 Woodside, M. T., García-García, C. & Block, S. M. Folding and unfolding single RNA molecules under tension. *Current opinion in chemical biology* **12**, 640-646 (2008).
28 Hyeon, C. & Thirumalai, D. Mechanical unfolding of RNA hairpins. *Proceedings of the National Academy of Sciences* **102**, 6789-6794 (2005).
29 Andreasson, J. O. *et al.* Crowdsourced RNA design discovers diverse, reversible, efficient, self-contained molecular switches. *Proceedings of the National Academy of Sciences* **119**, e2112979119 (2022).
30 Szczelkun, M. D. *et al.* Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proceedings of the National Academy of Sciences* **111**, 9798-9803 (2014).
31 Rutkauskas, M. *et al.* Directional R-loop formation by the CRISPR-Cas surveillance complex cascade provides efficient off-target site rejection. *Cell reports* **10**, 1534-1543 (2015).
32 Manosas, M. *et al.* Mechanism of strand displacement synthesis by DNA replicative polymerases. *Nucleic acids research* **40**, 6174-6186 (2012).
33 Green, A. A. *et al.* Complex cellular logic computation using ribocomputing devices. *Nature* **548**, 117-121 (2017).
34 Wang, S., Emery, N. J. & Liu, A. P. A novel synthetic toehold switch for microRNA detection in mammalian cells. *ACS synthetic biology* **8**, 1079-1088 (2019).
35 Angenent-Mari, N. M., Garruss, A. S., Soenksen, L. R., Church, G. & Collins, J. J. A deep learning approach to programmable RNA switches. *Nature communications* **11**, 1-12 (2020).
36 Poppleton, E., Romero, R., Mallya, A., Rovigatti, L. & Šulc, P. OxDNA. org: a public webserver for coarse-grained simulations of DNA and RNA nanostructures. *Nucleic acids research* **49**, W491-W498 (2021).
37 Bouchiat, C. *et al.* Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophysical journal* **76**, 409-413 (1999).
38 Stigler, J. & Rief, M. Hidden Markov Analysis of Trajectories in Single-Molecule Experiments and the Effects of Missed Events. *ChemPhysChem* **13**, 1079-1086 (2012).
39 Žoldák, G., Stigler, J., Pelz, B., Li, H. & Rief, M. Ultrafast folding kinetics and cooperativity of villin headpiece in single-molecule force spectroscopy. *Proceedings of the National Academy of Sciences* **110**, 18156-18161 (2013).

40 Machinek, R. R., Ouldridge, T. E., Haley, N. E., Bath, J. & Turberfield, A. J. Programmable energy landscapes for kinetic control of DNA strand displacement. *Nature communications* **5**, 5324 (2014).
41 Haley, N. E. *et al.* Design of hidden thermodynamic driving for non-equilibrium systems via mismatch elimination during DNA strand displacement. *Nature communications* **11**, 2562 (2020).
42 Liu, H. *et al.* Kinetics of RNA and RNA: DNA hybrid strand displacement. *ACS Synthetic Biology* **10**, 3066-3073 (2021).
43 Snodin, B. E. *et al.* Introducing improved structural properties and salt dependence into a coarse-grained model of DNA. *The Journal of chemical physics* **142** (2015).
44 Šulc, P. *et al.* Sequence-dependent thermodynamics of a coarse-grained DNA model. *The Journal of chemical physics* **137** (2012).
45 Srinivas, N. *et al.* On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic acids research* **41**, 10641-10658 (2013).
46 Irmisch, P., Ouldridge, T. E. & Seidel, R. Modeling DNA-strand displacement reactions in the presence of base-pair mismatches. *Journal of the American Chemical Society* **142**, 11451-11463 (2020).
47 Romano, F., Chakraborty, D., Doye, J. P., Ouldridge, T. E. & Louis, A. A. Coarse-grained simulations of DNA overstretching. *The Journal of chemical physics* **138** (2013).
48 Engel, M. C. *et al.* Force-induced unravelling of DNA origami. *ACS nano* **12**, 6734-6747 (2018).
49 Šulc, P., Ouldridge, T. E., Romano, F., Doye, J. P. & Louis, A. A. Simulating a burnt-bridges DNA motor with a coarse-grained DNA model. *Natural Computing* **13**, 535-547 (2014).
50 Belmont, P., Constant, J.-F. & Demeunynck, M. Nucleic acid conformation diversity: from structure to function and regulation. *Chemical Society Reviews* **30**, 70-81 (2001).
51 Sugimoto, N. *et al.* Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry* **34**, 11211-11216 (1995).
52 Ouldridge, T. E., Louis, A. A. & Doye, J. P. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *The Journal of chemical physics* **134** (2011).
53 Serra, M. J. & Turner, D. H. in *Methods in enzymology* Vol. 259 242-261 (Elsevier, 1995).
54 Woodside, M. T. *et al.* Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proceedings of the National Academy of Sciences* **103**, 6190-6195 (2006).
55 Green, C. & Tibbetts, C. Reassociation rate limited displacement of DNA strands by branch migration. *Nucleic Acids Research* **9**, 1905-1918 (1981).
56 Broadwater, D. B. & Kim, H. D. The effect of basepair mismatch on DNA strand displacement. *Biophysical journal* **110**, 1476-1484 (2016).
57 Dunn, K. E., Trefzer, M. A., Johnson, S. & Tyrrell, A. M. Investigating the dynamics of surface-immobilized DNA nanomachines. *Scientific reports* **6**, 29581 (2016).
58 Broadwater, D. B., Cook, A. W. & Kim, H. D. First passage time study of DNA strand displacement. *Biophysical Journal* **120**, 2400-2412 (2021).
59 Bustamante, C., Marko, J. F., Siggia, E. D. & Smith, S. Entropic elasticity of λ-phage DNA. *Science* **265**, 1599-1600 (1994).
60 Pei, Y. *et al.* Single-molecule resettable DNA computing via magnetic tweezers. *Nano Letters* **22**, 3003-3010 (2022).
61 Morgens, D. W. *et al.* Genome-scale measurement of off-target activity using Cas9 toxicity in high-throughput screens. *Nature communications* **8**, 15178 (2017).
62 Klein, M., Eslami-Mossallam, B., Arroyo, D. G. & Depken, M. Hybridization kinetics explains CRISPR-Cas off-targeting rules. *Cell reports* **22**, 1413-1423 (2018).
63 Zhang, J., Zong, W., Hong, W., Zhang, Z.-T. & Wang, Y. Exploiting endogenous CRISPR-Cas system for multiplex genome editing in Clostridium tyrobutyricum and engineer the strain for high-level butanol production. *Metabolic engineering* **47**, 49-59 (2018).
64 Khodakov, D., Wang, C. & Zhang, D. Y. Diagnostics based on nucleic acid sequence variant profiling: PCR, hybridization, and NGS approaches. *Advanced drug delivery reviews* **105**, 3-19 (2016).
65 Valeri, J. A. *et al.* Sequence-to-function deep learning frameworks for engineered riboregulators. *Nature communications* **11**, 1-14 (2020).
66 Wang, M. D. *et al.* Force and velocity measured for single molecules of RNA polymerase. *Science* **282**, 902-907 (1998).
67 Lang, M. J., Fordyce, P. M., Engh, A. M., Neuman, K. C. & Block, S. M. Simultaneous, coincident optical trapping and single-molecule fluorescence. *Nature methods* **1**, 133-139 (2004).

68 Bustamante, C., Bryant, Z. & Smith, S. B. Ten years of tension: single-molecule DNA mechanics. *Nature* **421**, 423-427 (2003).
69 Scharf, S. Cloning with PCR. *PCR Protocols: A guide to methods and applications*, 84-91 (1990).
70 Ouldridge, T. E., Šulc, P., Romano, F., Doye, J. P. & Louis, A. A. DNA hybridization kinetics: zippering, internal displacement and sequence dependence. *Nucleic acids research* **41**, 8886-8895 (2013).
71 Wang, M. D., Yin, H., Landick, R., Gelles, J. & Block, S. M. Stretching DNA with optical tweezers. *Biophysical journal* **72**, 1335-1346 (1997).
72 Suren, T. *et al.* Single-molecule force spectroscopy reveals folding steps associated with hormone binding and activation of the glucocorticoid receptor. *Proceedings of the National Academy of Sciences* **115**, 11688-11693 (2018).
73 Gebhardt, J. C. M., Bornschlögl, T. & Rief, M. Full distance-resolved folding energy landscape of one single protein molecule. *Proceedings of the National Academy of Sciences* **107**, 2013-2018 (2010).
74 Gebhardt, J. C. M., Clemen, A. E.-M., Jaud, J. & Rief, M. Myosin-V is a mechanical ratchet. *Proceedings of the National Academy of Sciences* **103**, 8680-8685 (2006).
75 Schlierf, M., Berkemeier, F. & Rief, M. Direct observation of active protein folding using lock-in force spectroscopy. *Biophysical journal* **93**, 3989-3998 (2007).
76 Gardiner, C. (Springer-Verlag, Berlin, 2009).
77 Schink, S. *et al.* Quantitative analysis of the nanopore translocation dynamics of simple structured polynucleotides. *Biophysical journal* **102**, 85-95 (2012).
78 Waters, L. S. & Storz, G. Regulatory RNAs in bacteria. *Cell* **136**, 615-628 (2009).
79 Prasanth, K. V. & Spector, D. L. Eukaryotic regulatory RNAs: an answer to the 'genome complexity'conundrum. *Genes & development* **21**, 11-42 (2007).
80 Burgyán, J. & Havelda, Z. Viral suppressors of RNA silencing. *Trends in plant science* **16**, 265-272 (2011).
81 Zamore, P. D., Tuschl, T., Sharp, P. A. & Bartel, D. P. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *cell* **101**, 25-33 (2000).
82 Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-1712 (2007).
83 Nudler, E. & Mironov, A. S. The riboswitch control of bacterial metabolism. *Trends in biochemical sciences* **29**, 11-17 (2004).
84 Mandal, M., Boese, B., Barrick, J. E., Winkler, W. C. & Breaker, R. R. Riboswitches control fundamental biochemical pathways in Bacillus subtilis and other bacteria. *cell* **113**, 577-586 (2003).
85 Mandal, M. & Breaker, R. R. Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nature structural & molecular biology* **11**, 29-35 (2004).
86 Isono, K. & Kitakawa, M. Cluster of ribosomal protein genes in Escherichia coli containing genes for proteins S6, S18, and L9. *Proceedings of the National Academy of Sciences* **75**, 6163-6167 (1978).
87 Kalnins, A., Otto, K., Rüther, U. & Müller-Hill, B. Sequence of the lacZ gene of Escherichia coli. *The EMBO journal* **2**, 593-597 (1983).
88 Wong, P., Gladney, S. & Keasling, J. D. Mathematical model of the lac operon: inducer exclusion, catabolite repression, and diauxic growth on glucose and lactose. *Biotechnology progress* **13**, 132-143 (1997).
89 Maki, Y., Yoshida, H. & Wada, A. Two proteins, YfiA and YhbH, associated with resting ribosomes in stationary phase Escherichia coli. *Genes to cells* **5**, 965-974 (2000).
90 Yamanaka, K., Zheng, W., Crooke, E., Wang, Y. H. & Inouye, M. CspD, a novel DNA replication inhibitor induced during the stationary phase in Escherichia coli. *Molecular microbiology* **39**, 1572-1584 (2001).
91 Mangan, M. W. *et al.* The integration host factor (IHF) integrates stationary-phase and virulence gene expression in Salmonella enterica serovar Typhimurium. *Molecular microbiology* **59**, 1831-1847 (2006).
92 Freundlich, M., Ramani, N., Mathew, E., Sirko, A. & Tsui, P. The role of integration host factor in gene expression in Escherichia coli. *Molecular microbiology* **6**, 2557-2563 (1992).
93 Goosen, N. & van de Putte, P. The regulation of transcription initiation by integration host factor. *Molecular microbiology* **16**, 1-7 (1995).
94 Ditto, M. D., Roberts, D. & Weisberg, R. A. Growth phase variation of integration host factor level in Escherichia coli. *Journal of bacteriology* **176**, 3738-3748 (1994).

95 Chang, D. E., Smalley, D. J. & Conway, T. Gene expression profiling of Escherichia coli growth transitions: an expanded stringent response model. *Molecular microbiology* **45**, 289-306 (2002).
96 Jones, C. & Holland, I. B. Role of the SulB (FtsZ) protein in division inhibition during the SOS response in Escherichia coli: FtsZ stabilizes the inhibitor SulA in maxicells. *Proceedings of the National Academy of Sciences* **82**, 6045-6049 (1985).
97 Zadeh, J. N. *et al.* NUPACK: Analysis and design of nucleic acid systems. *Journal of computational chemistry* **32**, 170-173 (2011).
98 Fernández-Coll, L. & Cashel, M. Contributions of SpoT hydrolase, SpoT synthetase, and RelA synthetase to carbon source diauxic growth transitions in Escherichia coli. *Frontiers in microbiology* **9**, 1802 (2018).
99 Xia, C., Fan, J., Emanuel, G., Hao, J. & Zhuang, X. Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. *Proceedings of the National Academy of Sciences* **116**, 19490-19499 (2019).
100 Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C. & Teichmann, S. A. The technology and biology of single-cell RNA sequencing. *Molecular cell* **58**, 610-620 (2015).
101 Qian, Y. *et al.* Programmable RNA sensing for cell monitoring and manipulation. *Nature* **610**, 713-721 (2022).
102 Jiao, C. *et al.* RNA recording in single bacterial cells using reprogrammed tracrRNAs. *Nature Biotechnology*, 1-10 (2023).
103 Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099-1102 (2010).
104 Mandal, M. & Breaker, R. R. Gene regulation by riboswitches. *Nature reviews Molecular cell biology* **5**, 451-463 (2004).
105 Mulhbacher, J. m. & Lafontaine, D. A. Ligand recognition determinants of guanine riboswitches. *Nucleic acids research* **35**, 5568-5580 (2007).
106 Kirchner, M. & Schneider, S. Gene expression control by Bacillus anthracis purine riboswitches. *Rna* **23**, 762-769 (2017).
107 Soukup, J. K. & Soukup, G. A. Riboswitches exert genetic control through metabolite-induced conformational change. *Current opinion in structural biology* **14**, 344-349 (2004).
108 Stevens, J. T. & Carothers, J. M. Designing RNA-based genetic control systems for efficient production from engineered metabolic pathways. *ACS synthetic biology* **4**, 107-115 (2015).
109 Davidson, E. A. & Ellington, A. D. Synthetic RNA circuits. *Nature chemical biology* **3**, 23-28 (2007).
110 Baeumner, A. J., Cohen, R. N., Miksic, V. & Min, J. RNA biosensor for the rapid detection of viable Escherichia coli in drinking water. *Biosensors and Bioelectronics* **18**, 405-413 (2003).
111 Mutalik, V. K., Qi, L., Guimaraes, J. C., Lucks, J. B. & Arkin, A. P. Rationally designed families of orthogonal RNA regulators of translation. *Nature chemical biology* **8**, 447-454 (2012).
112 Groher, F. *et al.* Riboswitching with ciprofloxacin—development and characterization of a novel RNA regulator. *Nucleic acids research* **46**, 2121-2132 (2018).
113 Groher, A. *et al.* (2019).
114 Felletti, M., Stifel, J., Wurmthaler, L. A., Geiger, S. & Hartig, J. S. Twister ribozymes as highly versatile expression platforms for artificial riboswitches. *Nature communications* **7**, 12834 (2016).
115 Wurmthaler, L. A., Sack, M., Gense, K., Hartig, J. S. & Gamerdinger, M. A tetracycline-dependent ribozyme switch allows conditional induction of gene expression in Caenorhabditis elegans. *Nature communications* **10**, 491 (2019).
116 Weenink, T., van der Hilst, J., McKiernan, R. M. & Ellis, T. Design of RNA hairpin modules that predictably tune translation in yeast. *Synthetic Biology* **3**, ysy019 (2018).
117 Laing, C. & Schlick, T. Computational approaches to RNA structure prediction, analysis, and design. *Current opinion in structural biology* **21**, 306-318 (2011).
118 Aguirre-Hernández, R., Hoos, H. H. & Condon, A. Computational RNA secondary structure design: empirical complexity and improved methods. *BMC bioinformatics* **8**, 1-16 (2007).
119 Dirks, R. M., Lin, M., Winfree, E. & Pierce, N. A. Paradigms for computational nucleic acid design. *Nucleic acids research* **32**, 1392-1403 (2004).
120 Shabalina, S. A., Spiridonov, A. N. & Ogurtsov, A. Y. Computational models with thermodynamic and composition features improve siRNA design. *BMC bioinformatics* **7**, 1-16 (2006).
121 Ceroni, F. *et al.* Burden-driven feedback control of gene expression. *Nature methods* **15**, 387-393 (2018).
122 Brophy, J. A. & Voigt, C. A. Principles of genetic circuit design. *Nature methods* **11**, 508-520 (2014).

123 Isaacs, F. J. *et al.* Engineered riboregulators enable post-transcriptional control of gene expression. *Nature biotechnology* **22**, 841-847 (2004).
124 Chappell, J., Takahashi, M. K. & Lucks, J. B. Creating small transcription activating RNAs. *Nature chemical biology* **11**, 214-220 (2015).
125 Lehr, F.-X. *et al.* Cell-free prototyping of AND-logic gates based on heterogeneous RNA activators. *ACS synthetic biology* **8**, 2163-2173 (2019).
126 Pardee, K. *et al.* based synthetic gene networks. *Cell* **159**, 940-954 (2014).
127 Hong, F. *et al.* Precise and programmable detection of mutations using ultraspecific riboregulators. *Cell* **180**, 1018-1032. e1016 (2020).
128 Sadat Mousavi, P. *et al.* A multiplexed, electrochemical interface for gene-circuit-based sensors. *Nature chemistry* **12**, 48-55 (2020).
129 Perdrizet, G. A., Artsimovitch, I., Furman, R., Sosnick, T. R. & Pan, T. Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. *Proceedings of the National Academy of Sciences* **109**, 3323-3328 (2012).
130 Chauvier, A. *et al.* Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation. *Nature communications* **8**, 13892 (2017).
131 Rentmeister, A., Mayer, G., Kuhn, N. & Famulok, M. Conformational changes in the expression domain of the Escherichia coli thiM riboswitch. *Nucleic acids research* **35**, 3713-3722 (2007).
132 Hollands, K. *et al.* Riboswitch control of Rho-dependent transcription termination. *Proceedings of the National Academy of Sciences* **109**, 5376-5381 (2012).
133 Proshkin, S., Mironov, A. & Nudler, E. Riboswitches in regulation of Rho-dependent transcription termination. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1839**, 974-977 (2014).
134 Bastet, L. *et al.* Translational control and Rho-dependent transcription termination are intimately linked in riboswitch regulation. *Nucleic acids research* **45**, 7474-7486 (2017).
135 Brennan, C. A., Dombroski, A. J. & Platt, T. Transcription termination factor rho is an RNA-DNA helicase. *Cell* **48**, 945-952 (1987).
136 Konan, K. V. & Yanofsky, C. Rho-dependent transcription termination in the tna operon of Escherichia coli: roles of the boxA sequence and the rut site. *Journal of Bacteriology* **182**, 3981-3988 (2000).
137 Iost, I., Guillerez, J. & Dreyfus, M. Bacteriophage T7 RNA polymerase travels far ahead of ribosomes in vivo. *Journal of bacteriology* **174**, 619-622 (1992).
138 Dong, H., Nilsson, L. & Kurland, C. G. Gratuitous overexpression of genes in Escherichia coli leads to growth inhibition and ribosome destruction. *Journal of bacteriology* **177**, 1497-1504 (1995).
139 Segall-Shapiro, T. H., Meyer, A. J., Ellington, A. D., Sontag, E. D. & Voigt, C. A. A 'resource allocator'for transcription based on a highly fragmented T7 RNA polymerase. *Molecular systems biology* **10**, 742 (2014).
140 Salis, H. M. in *Methods in enzymology* Vol. 498 19-42 (Elsevier, 2011).
141 Wiegand, D. J., Lee, H. H., Ostrov, N. & Church, G. M. Establishing a cell-free Vibrio natriegens expression system. *ACS synthetic biology* **7**, 2475-2479 (2018).
142 Gregory, G. J. & Boyd, E. F. Stressed out: Bacterial response to high salinity using compatible solute biosynthesis and uptake systems, lessons from Vibrionaceae. *Computational and structural biotechnology journal* **19**, 1014-1027 (2021).
143 Helmling, C. *et al.* Life times of metastable states guide regulatory signaling in transcriptional riboswitches. *Nature Communications* **9**, 944 (2018).
144 McDowell, J. C., Roberts, J. W., Jin, D. J. & Gross, C. Determination of intrinsic transcription termination efficiency by RNA polymerase elongation rate. *Science* **266**, 822-825 (1994).
145 Yarnell, W. S. & Roberts, J. W. Mechanism of intrinsic transcription termination and antitermination. *Science* **284**, 611-615 (1999).
146 Neidhardt, F. C., Ingraham, J. L. & Schaechter, M. Physiology of the bacterial cell: a molecular approach. *(No Title)* (1990).
147 So, L.-h. *et al.* General properties of transcriptional time series in Escherichia coli. *Nature genetics* **43**, 554-560 (2011).
148 Levine, E., Zhang, Z., Kuhlman, T. & Hwa, T. Quantitative characteristics of gene regulation by small RNA. *PLoS biology* **5**, e229 (2007).
149 Arkin, A., Ross, J. & McAdams, H. H. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected Escherichia coli cells. *Genetics* **149**, 1633-1648 (1998).
150 Mihailovic, M. K. *et al.* High-throughput in vivo mapping of RNA accessible interfaces to identify functional sRNA binding sites. *Nature communications* **9**, 4084 (2018).

151 Vogel, J. & Luisi, B. F. Hfq and its constellation of RNA. *Nature Reviews Microbiology* **9**, 578-589 (2011).
152 Lorenz, R. *et al.* ViennaRNA Package 2.0. *Algorithms for molecular biology* **6**, 1-14 (2011).
153 Hoffman, R. M. Application of GFP imaging in cancer. *Laboratory investigation* **95**, 432-452 (2015).
154 Meyer, A., Pellaux, R. & Panke, S. Bioengineering novel in vitro metabolic pathways using synthetic biology. *Current opinion in microbiology* **10**, 246-253 (2007).
155 Bach, E. A. *et al.* GFP reporters detect the activation of the Drosophila JAK/STAT pathway in vivo. *Gene Expression Patterns* **7**, 323-331 (2007).
156 Steiner, T. *et al.* Synthetic biology of proteins: tuning GFPs folding and stability with fluoroproline. *PLoS One* **3**, e1680 (2008).
157 Paige, J. S., Wu, K. Y. & Jaffrey, S. R. RNA mimics of green fluorescent protein. *Science* **333**, 642-646 (2011).
158 Dolgosheina, E. V. *et al.* RNA mango aptamer-fluorophore: a bright, high-affinity complex for RNA labeling and tracking. *ACS chemical biology* **9**, 2412-2420 (2014).
159 Warner, K. D. *et al.* A homodimer interface without base pairs in an RNA mimic of red fluorescent protein. *Nature chemical biology* **13**, 1195-1201 (2017).
160 Sunbul, M. *et al.* Super-resolution RNA imaging using a rhodamine-binding aptamer with fast exchange kinetics. *Nature Biotechnology* **39**, 686-690 (2021).
161 Neubacher, S. & Hennig, S. RNA structure and cellular applications of fluorescent light-up aptamers. *Angewandte Chemie International Edition* **58**, 1266-1279 (2019).
162 Huang, H. *et al.* A G-quadruplex–containing RNA activates fluorescence in a GFP-like fluorophore. *Nature chemical biology* **10**, 686-691 (2014).
163 DasGupta, S., Shelke, S. A., Li, N.-s. & Piccirilli, J. A. Spinach RNA aptamer detects lead (II) with high selectivity. *Chemical communications* **51**, 9034-9037 (2015).
164 You, M., Litke, J. L. & Jaffrey, S. R. Imaging metabolite dynamics in living cells using a Spinach-based riboswitch. *Proceedings of the National Academy of Sciences* **112**, E2756-E2765 (2015).
165 Dey, S. K. *et al.* Repurposing an adenine riboswitch into a fluorogenic imaging and sensing tag. *Nature Chemical Biology* **18**, 180-190 (2022).
166 Huang, K. *et al.* FASTmiR: an RNA-based sensor for in vitro quantification and live-cell localization of small RNAs. *Nucleic acids research* **45**, e130-e130 (2017).
167 Bertucci, A., Porchetta, A. & Ricci, F. Antibody-templated assembly of an RNA mimic of green fluorescent protein. *Analytical chemistry* **90**, 1049-1053 (2018).
168 Ren, K. *et al.* In situ genetically cascaded amplification for imaging RNA subcellular locations. *Journal of the American Chemical Society* **142**, 2968-2974 (2020).
169 Soni, R., Sharma, D., Krishna, A. M., Sathiri, J. & Sharma, A. A highly efficient Baby Spinach-based minimal modified sensor (BSMS) for nucleic acid analysis. *Organic & Biomolecular Chemistry* **17**, 7222-7227 (2019).
170 Roth, A. & Breaker, R. R. The structural and functional diversity of metabolite-binding riboswitches. *Annual review of biochemistry* **78**, 305-334 (2009).
171 Zhang, D. Y. & Seelig, G. Dynamic DNA nanotechnology using strand-displacement reactions. *Nature chemistry* **3**, 103-113 (2011).
172 Simmel, F. C., Yurke, B. & Singh, H. R. Principles and applications of nucleic acid strand displacement reactions. *Chemical reviews* **119**, 6326-6369 (2019).
173 Gilbert, S. D., Mediatore, S. J. & Batey, R. T. Modified pyrimidines specifically bind the purine riboswitch. *Journal of the American Chemical Society* **128**, 14214-14215 (2006).
174 Regulski, E. E. & Breaker, R. R. in *Post-transcriptional gene regulation* 53-67 (Springer, 2008).
175 Ottink, O. M. *et al.* Ligand-induced folding of the guanine-sensing riboswitch is controlled by a combined predetermined–induced fit mechanism. *Rna* **13**, 2202-2212 (2007).
176 Kim, J. N. & Breaker, R. R. Purine sensing by riboswitches. *Biology of the Cell* **100**, 1-11 (2008).
177 Lemay, J.-F., Penedo, J. C., Tremblay, R., Lilley, D. M. & Lafontaine, D. A. Folding of the adenine riboswitch. *Chemistry & biology* **13**, 857-868 (2006).
178 Calin, G. & Croce, C. MicroRNAs and chromosomal abnormalities in cancer cells. *Oncogene* **25**, 6202-6210 (2006).
179 Zhang, C. *et al.* Cancer diagnosis with DNA molecular computation. *Nature nanotechnology* **15**, 709-715 (2020).
180 Bhadra, S. & Ellington, A. D. in *Methods in enzymology* Vol. 550 215-249 (Elsevier, 2015).
181 Bhadra, S. & Ellington, A. D. A Spinach molecular beacon triggered by strand displacement. *Rna* **20**, 1183-1194 (2014).

182 Espah Borujeni, A., Mishler, D. M., Wang, J., Huso, W. & Salis, H. M. Automated physics-based design of synthetic riboswitches from diverse RNA aptamers. *Nucleic acids research* **44**, 1-13 (2016).
183 Zhao, B. & Zhang, Q. Characterizing excited conformational states of RNA by NMR spectroscopy. *Current opinion in structural biology* **30**, 134-146 (2015).
184 Carothers, J. M., Goler, J. A., Kapoor, Y., Lara, L. & Keasling, J. D. Selecting RNA aptamers for synthetic biology: investigating magnesium dependence and predicting binding affinity. *Nucleic acids research* **38**, 2736-2747 (2010).
185 Compton, J. Nucleic acid sequence-based amplification. *Nature* **350**, 91-92 (1991).
186 Boussebayle, A. *et al.* Next-level riboswitch development—implementation of Capture-SELEX facilitates identification of a new synthetic riboswitch. *Nucleic acids research* **47**, 4883-4895 (2019).
187 Wayment-Steele, H. K. *et al.* RNA secondary structure packages evaluated and improved by high-throughput experiments. *Nature Methods* **19**, 1234-1242 (2022).
188 Vezeau, G. E., Gadila, L. R. & Salis, H. M. Automated design of protein-binding riboswitches for sensing human biomarkers in a cell-free expression system. *Nature communications* **14**, 2416 (2023).
189 Dirks, R. M. & Pierce, N. A. A partition function algorithm for nucleic acid secondary structure including pseudoknots. *Journal of computational chemistry* **24**, 1664-1677 (2003).
190 Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research* **31**, 3406-3415 (2003).
191 Do, C. B., Woods, D. A. & Batzoglou, S. CONTRAfold: RNA secondary structure prediction without physics-based models. *Bioinformatics* **22**, e90-e98 (2006).
192 Schaeffer, J. M., Thachuk, C. & Winfree, E. in *International Workshop on DNA-Based Computers.* 194-211 (Springer).
193 Doye, J. P. *et al.* Coarse-graining DNA for simulations of DNA nanotechnology. *Physical Chemistry Chemical Physics* **15**, 20395-20414 (2013).
194 Kosuri, S. & Church, G. M. Large-scale de novo DNA synthesis: technologies and applications. *Nature methods* **11**, 499-507 (2014).
195 Anishchenko, I. *et al.* De novo protein design by deep network hallucination. *Nature* **600**, 547-552 (2021).
196 Cheng Tan, Y. Z., Zhangyang Gao, Hanqun Cao, Stan Z. Li. Hierarchical Data-efficient Representation Learning for Tertiary Structure-based RNA Design. *arXiv preprint arXiv:1511.08458* **arXiv:2301.10774** (2023).
197 Pelz, B. *Enzyme mechanics studied by single molecule force spectroscopy*, Technische Universität München, (2014).
198 Tolić-Nørrelykke, S. F. *et al.* Calibration of optical tweezers with positional detection in the back focal plane. *Review of scientific instruments* **77** (2006).
199 Feng, C. *et al.* Accurate de novo prediction of RNA 3D structure with transformer network. *bioRxiv*, 2022.2010. 2024.513506 (2022).
200 Paige, J. S., Nguyen-Duc, T., Song, W. & Jaffrey, S. R. Fluorescence imaging of cellular metabolites with RNA. *Science* **335**, 1194-1194 (2012).
201 Schärfen, L. & Neugebauer, K. M. Transcription regulation through nascent RNA folding. *Journal of molecular biology* **433**, 166975 (2021).
202 Watters, K. E., Strobel, E. J., Yu, A. M., Lis, J. T. & Lucks, J. B. Cotranscriptional folding of a riboswitch at nucleotide resolution. *Nature structural & molecular biology* **23**, 1124-1131 (2016).
203 Filonov, G. S., Moon, J. D., Svensen, N. & Jaffrey, S. R. Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *Journal of the American Chemical Society* **136**, 16299-16308 (2014).
204 Masuda, I. *et al.* A genetically encoded fluorescent tRNA is active in live-cell protein synthesis. *Nucleic acids research* **45**, 4081-4093 (2017).
205 Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265-269 (2020).
206 Yu, F. *et al.* Quantitative detection and viral load analysis of SARS-CoV-2 in infected patients. *Clinical Infectious Diseases* **71**, 793-798 (2020).
207 Cordray, M. S. & Richards-Kortum, R. R. Emerging nucleic acid–based tests for point-of-care detection of malaria. *The American journal of tropical medicine and hygiene* **87**, 223 (2012).