tRNA mimicking structures to control and monitor biological processes

Paul, Avishek

DOI:
10.33612/diss.166342562

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter 3

Photochemical control of bacterial gene expression based on trans encoded genetic switches

This chapter has been published online:

3.1. INTRODUCTION

Synthetic biology aims to develop genetic circuits to reprogram cell behaviour\(^1\),\(^2\). These complex genetic parts that allow perturbing and interpreting biological processes usually function through reacting to exogenous chemical inducers\(^3\),\(^4\). However, these chemical inducers generally result in a constitutive effect, i.e., in a continuous “ON” state, which limits their application to study inherently dynamic behaviour of cells\(^5\). In contrast, light can be considered as an excellent tool that matches those dynamics due to its high spatiotemporal resolution, complete bio-orthogonality and fine tunability in regard to its wavelength and intensity\(^6\),\(^7\). To date, two approaches have been widely used to control gene expression using light. One is the application of optogenetic tools, which generally control protein expression by regulating the interaction between a photo controllable transcription factor and its promoter through light irradiation\(^8\),\(^9\). This approach necessitates the expression of additional photo-responsive proteins which might cause burden on metabolic and cell signaling pathways\(^10\). The second photochemical approach\(^11\) relies on the installation of light responsive small molecules onto the bio-macromolecular scaffolds including nucleic acids\(^12\),\(^13\) and proteins\(^14\)-\(^16\) thereby providing an extra layer of control over their biological functions by light. This approach generally requires large synthetic efforts for covalently modifying the bio-macromolecules with photo-caging groups\(^17\) or photo-switches\(^18\),\(^19\). Moreover, the utilization of photocages only permits a single-way gene regulation event due to the deprotective removal of the caging groups under light illumination\(^20\).

**Figure 1.** a) Schematic of the Light-Off optogenetic tool. Upon light exposure, the light-switchable repressor LEVI (light-switchable repressor LexA-VVD) is
induced to form homodimers, which then bind to the operator sequence and repress the activity of promoter by blocking the attachment of RNA polymerase to the promoter. Removal of blue light results in gradual dissociation of the dimers and transcription activation. b) A caged monomeric building block, thymidine phosphoramidite, is incorporated into a deoxyoligonucleotide through standard DNA synthesis, rendering the oligomer inactive. Upon a brief irradiation with UV light, the caging group is removed, restoring the natural thymidine residue and thus the biological function of the DNA (e.g., the ability to undergo duplex formation, or catalytic activity). The 6-nitropiperonyloxymethylene (NPOM) caging group is shown in blue. c) The caged 10-23 DNAzyme D2 binds to its complementary RNA but has no catalytic activity due to incorporation of the caged thymidine at the crucial position T12 in the catalytic core. Irradiation at 365 nm removes the caging group, activates the DNAzyme, and induces RNA cleavage in presence of Mg$^{2+}$. d) Placement of a caging group on the E3 ligase ligand (blue triangle) prevents ternary complex formation between the E3 ubiquitin ligase (E3) and the protein of interest (POI). Treatment with light removes the caging group, enabling ternary complex formation, resulting in ubiquitin (Ub, orange circle) transfer and subsequent proteasomal degradation$^9$,$^{12,14}$.

To overcome these limitations, we present a novel photochemical approach to control bacterial gene expression based on the trans-encoded TMS (tRNA mimicking structure) genetic switch$^{21}$. As it was mentioned in the previous chapter, the TMS switch binds both flanking sites of the ribosome binding site (RBS) of the target mRNA, without disturbing the RBS and the gene of interest, to block ribosome entry and subsequent translation. Moreover, since RBS flanking sites do not contain any specific nucleotides, the same switch can target any gene without any alteration of the gene sequence. To reverse repression, an anti-repressor RNA was employed that binds the tRNA mimicking structure (TMS) and as a consequence pulls the TMS off from the mRNA, liberating the RBS to induce gene expression. Moreover, the modular TMS was reengineered to be responsive to other inputs than RNA. Bacterial tRNA structure harbors a loop containing the dihydrouridine base. This loop is called D-loop. With the replacement of the D-loop of TMS switch by a Neomycin B aptamer sequence$^{22,23}$, an RNA switch was obtained, which responds in a concentration-dependent manner to a Neomycin B derivative as input. Moreover, the ligand mediated gene switching was highly depending on the affinity of the ligand toward the aptamer component of the switch. Inspired by this observation, we
aimed to engineer the gene switch to be photo-responsive by fabricating a ligand bearing a photo-switchable unit that can undergo a light-induced structural change to exhibit two photo-isomers. We reasoned that these isomers would have different binding affinities and/or binding modes towards the sensor domain of the TMS switch and thus can induce photo-control of protein expression (Figure 2). In contrast to the photocaged systems, our design aims to control gene expression by regulating the noncovalent interactions between ligand and an aptamer domain on the riboswitch by light, and thus have the added benefits of being synthetically easily accessible and capable of regulating translation in both directions – from OFF to ON and from ON to OFF.

Figure 2. Concept to control gene expression by light employing the TMS switch and photo switchable ligands. The two photo-isomers exhibit different affinities to the sensor element of the TMS switch and hence allow photo-control of gene expression. The two isomers can be converted into each other by light and this structural change is enabled by an azobenzene moiety indicated by a white box.
connecting two ligands (shown in orange) that bind to the TMS switch in the magenta region.

3.2. RESULTS AND DISCUSSION

3.2.1. Design of the photo-switchable ligand binding to the TMS switch

In order to design the photo switchable ligand binding to the TMS switch, we decided to take paromomycin as the parent compound for further modifications. The reasons for this choice are twofold. First, paromomycin exhibits high binding affinity against the Neomycin B aptamer (Kd = 0.2 μM, Supplementary Figure 3.1) as determined by isothermal titration calorimetry (ITC)\(^{24}\). This binding is even stronger than that between the same sequence and NeomycinB-azide (Kd = 28 μM), which was used as input to control gene expression\(^{21}\) as it was described in the previous chapter. Second, given that paromomycin has a unique aminomethyl group with little steric hindrance, it can be readily functionalized by activated acylating reagents with high regioselectivities\(^{25}\). With these two considerations in mind, a type of photo-switchable dimeric paromomycin (Figure 3a) containing o-fluoroazobenzene moiety (F-dimer) as the photo-responsive linker was designed. The introduction of fluorine atoms in the ortho-positions not only allows the use of light in the visible range (\(\lambda > 500\) nm) instead of UV light to trigger the \(E\rightarrow Z\) isomerization, but also dramatically enhances the stability of the thermodynamically unfavourable \(Z\)-isomers\(^{26,27}\). We hypothesized that the \(Z\)-isomers would bind less efficiently to the TMS switch since the amino sugar rings in the \(Z\)-isomers are arranged in a more crowded configuration than in the \(E\) isomer, thus introducing steric hindrance and eventually blocking the binding to the aptamer domain of the TMS switch.

3.2.2. Synthesis and characterization of F-dimer

Following our design, the photo-switchable F-dimer was synthesized from 2,2’,6,6’-tetrafluoroazobenzene-4,4’-dicarboxylic acid by well-established amide bond formation procedures\(^{27}\) (Scheme 3.2). This compound was activated to form the corresponding N-hydroxysuccinimide (NHS) ester, which was reacted with two aminoglycoside units. In this way, the aminomethyl group in position 6’’’ of paromomycin was functionalized by activated acylating reagent with high regioselectivities\(^{25}\) forming the photo-switchable F-dimer. The crude product was purified by high performance liquid chromatography (HPLC) and characterized by 1D- and 2D-NMR (heteronuclear single quantum coherence (HSQC)) spectroscopy (Supplementary Figure 3.2 and Supplementary Figure
As shown in Figure 3c, the HSQC spectrum of the F-dimer shows a remarkable shift of the \(J(C6^\text{-}H)\) coupling of ring IV in comparison to the 2D-spectrum of pristine paromomycin (Figure 3b) proving the regioselective acylation reaction in the C6 position of ring IV.

The photo-physical properties of F-dimer in water were studied using UV-vis spectroscopy. The UV–vis spectrum of the \(E\)-isomer of the F-dimer exhibits a characteristic strong \(\pi \rightarrow \pi^*\) absorption band at short wavelength (\(\lambda = 317\) nm) and a weaker \(n \rightarrow \pi^*\) band centred around 446 nm (Figure 3d). Fluorine substituents in the ortho-positions, as \(\sigma\)-electron withdrawing groups, can lower the n-orbital energy of the \(Z\)-isomer, resulting in a blue-shift of the \(n \rightarrow \pi^*\) absorption band for the \(Z\)-isomer. As a result, the two isomers of F-dimer exhibit a separation of \(n \rightarrow \pi^*\) bands of 26 nm, which allows to selectively trigger the \(E \rightarrow Z\) isomerization with visible light at wavelengths longer than 500 nm. Thus, the solution was irradiated with a 530 nm LED coupled to an optical fiber. During the \(E \rightarrow Z\) photo-isomerization, hypochromism and bathochromic shifts were concomitantly observed in the region of \(\pi \rightarrow \pi^*\) absorption, alongside with an increase in absorbance coupled with shift towards shorter wavelengths in the region of \(n \rightarrow \pi^*\) absorption. During the \(Z \rightarrow E\) photo-isomerization upon irradiation at 455 nm, the \(\pi \rightarrow \pi^*\) absorption peak shifted from 295 nm to 317 nm while the \(n \rightarrow \pi^*\) absorption peak shifted from 420 nm to 446 nm until reaching the thermodynamic equilibrium state, also known as photo-stationary state. Both photo-isomerization processes exhibit three isosbestic points at 276 nm, 378 nm and 442 nm.
Figure 3. Structure and characterization of F-dimers by linking two paromomycin units with a difunctional azobenzene core. a) Chemical structure of F-dimer. b) HSQC spectrum (500 MHz, D$_2$O) of paromomycin. c) HSQC spectrum (500 MHz, D$_2$O) of F-dimer (see Supplementary Figure 3.3 for full spectrum). Red arrows indicate the shift of specific signals caused by the regioselective transformation of the amino group in position 6''' of ring IV: J(C6'''−H$_a$) and J(C6'''−H$_b$) coupling. d) Changes in the absorption spectra of a solution of F-dimer in MQ water (10 μM) upon irradiation at 530 nm (E→Z) and at 455 nm (Z→E). Blue and red dashed lines indicate the separation of π→π* bands between the two isomers of the F-dimer. e) Repetitive switching cycles of F-dimer upon alternating irradiation with visible light was used to isomerize F-dimer in both directions, producing PSSs containing 78% of Z-isomer with green light (530 nm) and 75% of E-isomer with blue light (λ = 455 nm). After eight cycles, no fatigue indicated by a reduction of absorbance was observed. The absorbance was measured at the maximum of the π-π* transition (317 nm) of the E-isomer.

Using 1D-NMR the ratio of the two photo-isomers of F-dimer at the photo-stationary state (PSS) were determined. Irradiation with visible light was used to isomerize F-dimer in both directions, producing PSSs containing 78% of Z-isomer with green light (530 nm) and 75% of E-isomer with blue light (455 nm) (Supplementary Figure 3.2a and Supplementary Figure 3.2b). Multiple E/Z photo-isomerization cycles did not result in any noticeable degradation, highlighting the robustness of the F-dimer switch (Figure 3e). As a consequence of the stabilization of the n-orbital of the Z-isomer through the introduction of ortho-fluoro substituents, thermal Z→E isomerization of F-dimer shows a half-life of more than 8 days at 37°C in the dark (Supplementary Figure 3.6). These photo-physical characteristics, i.e., irradiation at longer wavelength and high thermal stability, will be beneficial for controlling gene expression using light in living systems as shown below.

3.2.3. Studies on the interactions between photo-isomers and the aptamer domain of the TMS switch

With this photo-switch at hand, we firstly determined the binding affinity of the E and Z isomers against the sensor domain of the TMS switch. For this purpose, the isomers of the resulting F-dimer were isolated by HPLC with >95% purities (Supplementary Figure 3.2c and Supplementary Figure 3.2d) thanks to their extreme thermal stability. The purified isomers and a 23mer aptamer$^{22,23}$ that will be later incorporated into the TMS switch structure were then used for isothermal titration calorimetry (ITC) studies. It was observed that the two isomers can bind
to the sensor domain with different affinities – the E-isomer shows significantly higher binding affinity ($K_d = 0.85 \ \mu M$, Figure 4a) than that of the Z-isomer ($K_d = 73 \ \mu M$, Figure 4b) against the sensor domain. This result supports our hypothesis that the more crowded configuration of Z-isomer prevents its effective binding to the aptamer.

To obtain a better insight into the nature of interactions between the photo-isomers and the aptamer, we carried out atomistic molecular dynamics (MD) simulations of the binding between E/Z-F-dimer and the 23mer aptamer sequence by NAMD$^{28}$ and the CHARMM force fields$^{29,30}$. The structure of the aptamer was taken from the PDB that was previously determined by NMR$^{31}$. The particle mesh Ewald (PME)$^{32}$ method was used for the evaluation of long-range Columbic interactions. Figure 4c shows the initial and final states of each simulation (for the trajectories of F-dimers interacting with aptamer, see Supplementary Figure 3.7). Initially, only one paromomycin of the F-dimers was inserted in the binding pocket on the aptamer (Figure 4c left column). The Z-isomer causes severe conformational changes of the aptamer, while the E-isomer induces minor disturbance in the aptamer structure (Figure 4c right column). In the case of the Z-isomer, the first paromomycin, which initially binds to the aptamer groove, gets distorted and shifts away from the initial binding pose, while the second paromomycin opens up the RNA groove due to competitive binding to the RNA backbone caused by the bent structure preserved by the Z-azobenzene. In contrast, for the E-isomer, the first paromomycin stays more anchored in the RNA groove, while the second paromomycin transiently binds to the RNA or stays solvated in the solution due to the trans-configuration, which perturbs the aptamer structure less.
Figure 4. Study of the binding between the F-dimer isomers and the aptamer domain of the TMS switch. ITC titration curves of a) E isomer (70 μM in 20 mM phosphate buffer pH 7.5) into aptamer solution (7 μM in 20 mM phosphate buffer pH 7.5) and b) Z isomer (700 μM in 20 mM phosphate buffer pH 7.5) into aptamer solution (7 μM in 20 mM phosphate buffer pH 7.5); c) Initial and final snapshots for the E/Z-isomers with RNA aptamer (the nucleotides are numbered from 5’ to 3’ as 1 to 23); d) Contacting times of nucleotides with E-isomer during simulation. e) Contacting times of nucleotides with Z-isomer during simulation. ADE = adenine, CYT = cytosine, GUA = guanine, URA = uracil; f) RMSD for the nucleotides forming the initial binding pocket in the presence of E- or Z-isomer.

Next, we analysed the contacting nucleotides with E/Z-F-dimer. The nucleotide is considered to interact with F-dimers if it is within 3 Å of the E/Z-isomer. The number of contacting nucleotides in the Z case plummets at 1.75 μs, when the RNA groove opens significantly, while the number for E-isomer remains similar during the whole simulation time (Supplementary Figure 3.8). We further analysed the contact times of each nucleotide over the last 1.8 μs. The nucleotides which interact less than half of the total time are considered to be less involved in binding. The initial nucleotides forming the binding pocket were encircled red in Figure 4d and 4e. Nucleotides out of the initial binding pocket also show stable binding to the E/Z-isomers due to re-adjustment of binding during simulation. In
the Z case, one of the initial nucleotides (#10) marked in green shows reduced binding times, while all the initial nucleotides in the E case participate in the binding strongly. The total number of interacting nucleotides in the E case (18) is less than that of the Z-isomer (23). Besides, the contacting time shows bipolarity among the nucleotides in the E case, meaning that the nucleotides are binding either for long or very short times during the whole simulation process, while it is in a more random order in the Z case (Figure 4d and 4e).

The binding energy between F-dimers and aptamer was calculated by VMD plugin\textsuperscript{33} (Supplementary Figure 3.9). The total binding energy is more waved in the Z- than in the E configuration. Moreover, the total binding energy in the Z case fluctuates up at 1.75 µs due to opening up of the RNA aptamer. The average binding energy (total) for the Z isomer after 1.75 µs is -1602 kcal/mol, while it is -1701 kcal/mol for the E isomer, which indicates a stronger binding between the E isomer and the aptamer. Figure 4f shows the root-mean-square deviation (RMSD) for the initial nucleotides encircled in Figure 4d and 4e. The Z-isomer breaks the initial pocket indicated by the big fluctuation of the RMSD value; in contrast, the E-isomer only slightly disturbs the structure of the initial pocket.

In summary, during binding of the E isomer the “canonical” strong neomycin B – aptamer binding mode is maintained, whereas the Z isomer does not resemble the initial binding mode. The different binding modes of the two different isomers originate from the structural changes induced by E/Z-azobenzene moieties. The Z-isomer deforms the groove of the RNA aptamer and binds in a non-specific and weak manner, while upon binding of the E-isomer the specific binding mode is preserved due to the larger distance between the two paromomycin units. Based on these results, we anticipate that the Z-isomer would bind less than the E-isomer to the aptamer domain of the TMS switch and thus is unable to turn on protein production. Accordingly, control over gene expression by light can be achieved by switching the conformation of the photo-isomers and their interactions with the TMS switch.

3.2.4. Photochemical control over gene expression with F-dimer

After verifying the difference in binding between the photo-isomers against the sensor domain of the TMS switch, we seek to exploit the F-dimers to photochemically control gene expression in bacteria. Since the parent compound paromomycin is an antibiotic drug, we first examined the toxicity of the F-dimers in bacteria through a MIC test\textsuperscript{34}. It was found that both photo-switchable aminoglycoside isomers show no difference in antibacterial activity. Both exhibit
a similar MIC value of > 64 μM (Supplementary Figure 3.10). Noteworthy, modifying paromomycin with a photo-switchable unit reduces its original antibacterial activity, which is a prerequisite for using it as a non-toxic input signal to control gene expression.

To prove the feasibility of our concept, HPLC purified E- and Z-F-dimers were employed to regulate GFP expression in bacteria. The TMS switch and the GFP gene were expressed from two separate plasmids in Escherichia coli BL21(DE3) cells. The TMS switch was produced from an IPTG inducible T7 promoter and the GFP mRNA was transcribed from an arabinose inducible promoter. The cells were transformed with the plasmids encoding the switch and GFP and were grown in LB medium in presence of ampicillin and chloramphenicol. In the meantime, LB agar plates were prepared supplemented with these two antibiotics corresponding to the plasmids used in this study and inducers to express the switch and the GFP mRNA. When the transformed cells in the LB medium reached the early log phase (O.D. = 0.4), inducers to express the switch and the GFP were added into the medium. Later, when the O.D. of the cells reached 0.8 they were spread on two LB agar plates. Then 32 μM purified E- and Z-isomer of the F-dimer were added onto these plates, respectively. Simultaneously, two other plates were prepared as a positive control (plate containing inducer to express GFP only without adding any F-dimer, Figure 5a) and the negative control (plate containing neither inducers nor any F-dimer, Figure 5b). All the plates were kept in the dark at 37°C overnight. The results demonstrated that the E-isomer could successfully initiate GFP expression to the level similar to the positive control (Figure 5c), while the Z-isomer only induced very weak gene expression (Figure 5d). This observation bolsters our previous hypothesis – E-isomer could effectively bind to the sensor domain of the TMS switch, distort the switch structure resulting in detachment of the TMS switch from the target GFP mRNA and consequently initiation of GFP translation. Due to the crowded configuration of the two paromomycin units and the lower affinity, Z-isomer was unable to trigger the TMS switch as input. To quantify the GFP expression on each plate, we dispersed the bacterial colonies in PBS buffer for flow cytometry analysis. These experiments showed that the positive control as well as the sample containing the E-isomer showed two populations of cells, one with a very low level of GFP expression and another with a high GFP expression level (Figure 5e). We calculated the $\text{GFP}_E/\text{GFP}_Z$ ratio, which turned out to be as high as 110-fold. $\text{GFP}_E$ signifies the GFP expression of the cells occurring in presence of the
switch and the E-isomer while GFP\textsubscript{Z} refers to the GFP expression in presence of the switch and the Z-isomer.

Next, to determine the effect of the concentration of the E-isomer on the GFP expression pattern, we incubated the cells with the pure E-isomer employing different concentrations ranging from 64 \( \mu \text{M} \) to 4 \( \mu \text{M} \). These experiments indicated that the E-isomer can initiate GFP expression at 32 \( \mu \text{M} \) but not 16 \( \mu \text{M} \) (Supplementary Figure 3.11). To more precisely determine the turn on concentration of GFP expression three more concentrations of the E-isomer in between 16 and 32 \( \mu \text{M} \) were investigated, i.e., 20, 24 and 28 \( \mu \text{M} \). GFP expression was not visible at 20 \( \mu \text{M} \), but when 24 \( \mu \text{M} \) and 28 \( \mu \text{M} \) of E-isomer were applied (Supplementary Figure 3.12). This observation indicates that the TMS switch is highly sensitive to the concentration of the small molecule inputs.

**Figure 5.** Photochemical control over GFP expression with F-dimer isomers. a) Fluorescent images of the plates after overnight incubation at 37\( ^\circ \text{C} \): positive control (containing inducer to express GFP only without having any F-dimer); b) negative control (containing neither inducers nor any F-dimer); c) in the presence of 32 \( \mu \text{M} \) purified E-isomer; d) in the presence of 32 \( \mu \text{M} \) purified Z-isomer. e) Flow cytometry histograms of the GFP expression occurring on the plates from
(a) to (d). The histograms show bimodal populations of GFP expression for the positive control and cells containing E-isomer. f) Fluorescent image recording of in-situ activation of GFP expression where a photomask was placed on the right half of the plate.

Encouraged by these results, we intended to carry out the photo-isomerization of F-dimer in situ to control the bacterial gene expression. First, Z to E isomerization was studied. Following the procedures described above, 32 μM pure Z-isomer was added onto six plates with bacteria. Five of the plates were irradiated with light of a wavelength of 455 nm for 5 min, 10 min, 30 min, 1 h and 1.5 h, respectively, while the last plate was kept in the dark. Afterwards the plates were incubated at 37°C overnight. As shown in Supplementary Figure 3.13, the minimum exposure time required to partially turn on gene expression was 30 minutes and the plate that was treated with light for 1.5 hours showed maximum GFP expression. The ON/OFF ratio of this in situ switching was quantified to be around 112-fold. The ON/OFF ratio was calculated from the fluorescence of the GFP positive cells treated with light for 1.5 hours and fluorescence of the cells that were kept in dark. In contrast, the plate that was not illuminated with light and the plates that were irradiated for short periods of time showed no GFP expression. To confirm that the GFP expression was triggered only due to the Z to E isomerization but not because of any other effect of light irradiation, the same plate without Z-F-dimer was irradiated with light of a wavelength of 455 nm for 1.5 hours. It turned out that this plate did not show any GFP expression (Supplementary Figure 3.14). These observations proved that the Z-isomer was converted to the E-isomer and subsequently triggered GFP expression. Moreover, the non-irradiated plate containing the Z-isomer did not show any GFP fluorescence as the Z-isomer itself is unable to trigger GFP expression through structural alteration of the TMS switch and liberation of the RBS on the mRNA.

Next, E to Z isomerization induced in situ to switch-off GFP expression was investigated. Two agar plates with transformed cells were prepared. Initially 32 μM pure E-isomers were added to these plates. One plate was irradiated with light of a wavelength of 530 nm for 1.5 hours and then both plates were kept at 37°C for overnight incubation. It turned out that both plates showed similar level of GFP expression, suggesting that E to Z photo-isomerization was not as efficient as Z to E isomerization under the experimental conditions. We reasoned that the observed difference might originate from the presence of aptamer inside cells. To prove this, we performed an in vitro study where the photoisomerization of F-
dimers was measured in the presence of NeomycinB aptamer. In this study, one equivalent of the photo-isomers was incubated with two equivalents of aptamer for one hour. Then, their photoisomerization processes were analyzed using UV-vis spectroscopy (Supplementary Figure 3.15). It was observed that upon irradiation at 530 nm, the $E \rightarrow Z$ photoisomerization was significantly slowed down by the presence of RNA aptamer (Supplementary Figure 3.15a and Supplementary Figure 3.15b). In contrast, under irradiation at 455 nm ($Z \rightarrow E$), the presence of RNA aptamer barely influences the photoisomerization kinetics of F-dimers (Supplementary Figure 3.15c and Supplementary Figure 3.15d). These results can be explained by the spatial confinement effect which plays a pivotal role in photoisomerization of azobenzenes. Our photo-switchable compounds are composed of two bulky paromomycin molecules linked by an azobenzene group and the compounds require a high degree of conformational freedom to switch between the two photo-isomeric forms. However, when the $E$ isomer binds to the aptamer domain of the switch, the compound encounters a sterically hindered environment that confines space and hinders free rotation to convert the $E$ to $Z$ isomer.

Since the TMS switch is highly sensitive to the concentration of the $E$-isomer, we reasoned that to switch-off gene expression in-situ might be achievable at a concentration that is close to the observed ‘threshold’ concentration where obvious GFP production started. To test this, the applied concentration of the $E$-isomer was reduced to 22 $\mu$M for the same experiments as described above. As indicated from the flow cytometry results (Supplementary Figure 3.16), there was partial GFP fluorescence observed with 22 $\mu$M $E$-isomer but no GFP fluorescence was noticed when the plate was treated with 530 nm wavelength light, suggesting that a slight amount of transformation to the $Z$-form reduced the $E$-isomer concentration to cross the ‘threshold’ and thus effectively down-regulate the expression of GFP. Further improvement on the sensitivity of our system will be subject of future studies by either more elegant molecular design or increasing the intensity of the applied light.

As light as a stimulus holds great promise for locally activating a certain biological function, we next studied if the amalgamation of photo-responsive F-dimers and the TMS switch could spatially resolved address a certain population of bacteria on an agar plate. For this purpose, the transformed cells were spread on an agar plate and the $Z$-isomer was added on the plate. Then a photomask was placed to cover half of the plate. After 1-hour light irradiation (455 nm), the
photomask was removed, and the plate was incubated at 37°C overnight. Figure 5f shows that GFP was only expressed on the exposed part of the plate, while no GFP expression was observed on the other half of the plate where the photomask was originally placed. This experiment proves the capability of the F-dimer in combination with the TMS switch to realize spatial control over gene expression.

**Figure 6.** Photochemical control over φX174 E lysis gene expression with F-dimers. Images of the plates after overnight incubation at 37°C: a) positive control (containing inducer to express φX174 E lysis gene only without having any F-dimer); b) negative control (containing neither inducers nor any F-dimer); c) in the presence of 32 μM purified E isomer; d) in the presence of 32 μM purified Z isomer. e) In-situ activation of φX174 E lysis gene expression by Z to E photoisomerization.
3.2.5. Photochemical control over φX174 E lysis gene expression with F-dimers

Finally, to prove the general applicability of our approach, we replaced the target GFP gene with a bacteriophage lysis gene (φX174 E) to photo-chemically control bacterial cell growth in the presence of the TMS switch and photo-switchable F-dimer. φX174 E lysis gene refers to a membrane protein of 91 amino acids, which hinders peptidoglycan synthesis in the bacterial cell wall and thus causes lysis of the host cell during growth\textsuperscript{37}. Following the same experimental design for controlling GFP expression, the TMS switch and the lysis gene were expressed from two separate plasmids in \textit{Escherichia coli} BL21(DE3) cells. The cultured cells in which the expression of the bacteriophage lysis gene was initially inhibited by the TMS switch were spread on two agar plates. To these plates, 32 μM \textit{E}- and \textit{Z}-isomers were added, respectively. At the same time, two other plates were prepared as positive control (plate containing inducer to produce the lysis gene only, Figure 6a) and negative control (plate containing both inducers but without any form of F-dimer, Figure 6b). After overnight incubation at 37°C, the plate treated with \textit{E}-isomer did not show any bacterial growth, similar as the positive control plate (Figure 6c). In stark contrast, the plate treated with \textit{Z}-isomer was fully covered with bacterial colonies (Figure 6d). Since we already demonstrated that the isomers of F-dimer showed no difference in antibacterial activity from the MIC test, the drastic difference of bacterial growth must be a result of the different level of lysis gene production induced by the two isomers. To further prove the photo control of this system, 32 μM \textit{Z}-isomer was added to the agar plate with spread cells. Subsequently, the plate was illuminated with light of a wavelength of 455 nm for 1 hour and was kept overnight at 37°C. The plate did not show any sign of bacterial growth (Figure 6e). To rule out the effect of light, one plate was prepared containing transformed cells and \textit{Z}-isomer but without addition of the inducer to express the lysis gene. Another plate with the same cells contained both inducers to express the TMS switch and the lysis gene but no \textit{Z}-isomer. These plates were then illuminated with light of a wavelength of 450 nm for 1.5 h. After overnight incubation, both plates did not show any inhibition of bacterial growth (Supplementary Figure 3.17), indicating the \textit{in situ} switch-on of lysis gene expression originating from light induced conformational changes of the F-dimer.
3.3. CONCLUSION

In summary, we have demonstrated a powerful photo-responsive gene expression system that relies on a TMS switch and the corresponding photo-switchable F-dimer ligand. The use of fluorinated azobenzene as a photo-responsive unit enabled our system to operate with visible light, which represents an important feature for future development in living systems. By switching the conformations of F-dimer using light, we were able to regulate their interactions with the aptamer domain of the TMS switch and accordingly up- or down-regulate the level of the target protein production. Importantly, this working mechanism was further validated with the help of ITC measurements and computational studies. Spatially resolved activation of gene expression was realized by applying light at the desired position. In addition, since our tRNA based TMS switch is trans encoded, it can easily be accommodated to control any other gene with the same photo-switchable ligand and the same TMS switch. This modularity and the general applicability of our system was demonstrated by expanding the control of GFP production to bacteriophage lysis gene expression. In this regard, it is worth to note that recent synthetic biology approaches are involving engineered bacteria to deliver drugs in eukaryotic cells\textsuperscript{38}. This opens the avenue of designing future photo-responsive gene expression systems where the on-site production of functional and therapeutic proteins can be controlled with high spatiotemporal resolution by light. Moreover, the TMS switch can be developed for photo-chemical systems in eukaryotic cells by replacing the bacterial tRNA scaffold with a eukaryotic tRNA structure. We believe that this implementation will raise broad interest in the fields of chemical and synthetic biology.
3.4. EXPERIMENTAL SECTION

3.4.1. Synthesis of F-dimer

Scheme 3.1. Synthetic route of F₄-diacid (7).

2,6-difluoro-4-bromoaniline (2).

NBS (35.6 g, 200 mmol) was added to a solution of 2,6-difluoroaniline (1) (25.8 g, 200 mmol) in acetonitrile (300 mL) and stirred overnight at room temperature. The mixture was diluted with water and hexane. After separation of the layers, the organic phase was dried over MgSO₄, filtrated and the solvent was evaporated in vacuo. The resulting mixture was purified by column chromatography (DCM/hexanes : 1/1) to give 2 as a purple solid (33.5 g, 80%).

\(^1\)H NMR (300 MHz, CDCl₃) δ ppm 7.00 (dd, \(J = 6.1, 1.3\) Hz, 2 H), 3.75 (br s, 2 H). \(^1^3\)C NMR (75 MHz, CDCl₃) δ ppm 153.36, 150.12, 123.50, 114.72, 107.09.
4-amino-3,5-difluorobenzonitrile (3).

Compound 2 (20 g, 96 mmol) and CuCN (25.6 g, 288 mmol) were dissolved in DMF (100 mL). The mixture was refluxed overnight. After the solution was cooled down, the mixture was poured into a NH₃ 12% aqueous solution and extracted with ethyl acetate. The organic phase was dried over MgSO₄, filtered, and the solvent was evaporated in vacuo. The product was purified by column chromatography (DCM/hexanes : 2/1) to give 3 as white solid (9.2 g, 62%). ¹H NMR (300 MHz, CDCl₃) δ ppm 7.16 (dd, J = 6.1, 2.2 Hz, 2 H), 4.29 (br s, 2 H).

4-amino-3,5-difluorobenzoic acid (4).

Compound 3 (6.9 g, 45 mmol) was suspended in NaOH 1M (150 mL) and refluxed overnight. After the solution was cooled down, 1M HCl was added until hydrochloric salt 4 precipitated. The precipitate was then dissolved in EtOAc, dried over MgSO₄, filtered, and the solvent was evaporated in vacuo. The product was obtained as white solid and used without purification in the next step. ¹H NMR (300 MHz, DMSO-d₆) δ ppm 7.40 (dd, J = 4.2, 1.4 Hz, 2 H), 3.81 (br s). ¹³C NMR (75 MHz, DMSO-d₆) δ ppm 165.96, 150.65, 148.74, 130.65, 115.59, 112.22. HRMS-ESI: m/z = 172.0278 (calcd for [M - H]⁻, 172.0204).

Ethyl 4-amino-3,5-difluorobenzoate (5).

To a solution of acid 4 (5.92 g, 28 mmol) in EtOH (100 mL) was added H₂SO₄ (2 mL) and refluxed for 5 h. Saturated NaHCO₃ was added until the solution was neutralized (pH 7). The mixture was extracted with DCM and the organic phase was dried over MgSO₄ and filtered. The solvent was evaporated under reduced pressure to give 5 as a pale brown solid (4.2 g, 69% over two steps). ¹H NMR (300 MHz, CDCl₃) δ ppm 7.50 (dd, J = 7.2, 2.2 Hz, 2 H), 4.30 (q, J = 7.1 Hz, 2 H), 4.07 (br s, 2 H), 1.34 (t, J = 7.1 Hz, 3H). ¹³C NMR (75 MHz, CDCl₃) δ ppm 165.19, 152.21, 149.03, 128.78, 118.38, 112.55, 61.04, 14.25. HRMS-ESI: m/z = 202.0535 (calcd for [M + H]⁺, 202.0680).
**Diethyl-4,4’-(2,2’,6,6’-tetrafluoro)azobenzene dicarboxylate (6).**

Compound 5 (600 mg, 3 mmol) was dissolved in DCM (100 mL) and a freshly grinded mixture of KMnO₄ (5.1 g) and FeSO₄·7H₂O (5.1 g) was suspended in this solution. The mixture was refluxed overnight, filtered through celite, dried over MgSO₄, filtered, and the solvent was evaporated *in vacuo*. The product was purified by column chromatography (DCM/hexanes : 1/1) to give 6 as an orange/red solid (136 mg, 23%). ¹H NMR (300 MHz, CDCl₃) (E-isomer) δ ppm 7.76 (m, 4 H), 4.45 (q, J = 7.0 Hz, 4 H), 1.45 (t, J = 7.0 Hz, 6 H). ¹³C NMR (75 MHz, CDCl₃)(E-isomer) δ ppm 163.63, 156.75, 153.24, 133.80, 113.90, 62.18, 14.20. HRMS-ESI: m/z = 399.0999 (calcd for [M + H]⁺, 399.0968).

**F4-diacid (7).**

Compound **F4-diester 6** (136 mg, 0.34 mmol) and KOH (67 mg, 1.19 mmol) were dissolved in H₂O/THF 2:1 (4 mL) and the solution was heated at 70 °C for 2 h. The solution was diluted with water and ethyl acetate, the two layers were separated, and the aqueous layer was washed twice with ethyl acetate. The aqueous layer was acidified with HCl (1M) until the orange precipitate fell out from solution. The orange precipitate was filtered to yield 7 as an orange solid (113 mg, 97%). HRMS-ESI: m/z = 342.0264. (calcd for [M - H]⁻, 342.2056).

Into a solution of paramomycin sulfate salt (0.3 mmol) in water/methanol (1:1 10 mL), ion-exchange resin was added and the solution was stirred for 2 h at room temperature. Ion-exchange resin was removed by filtration and the aqueous solution was concentrated under reduced pressure to remove methanol and freeze dried.
Scheme 3.2. Synthetic route of F-dimer.

In a 5 ml round bottom flask, F4-diacid 7 (20.4 mg, 0.06 mmol) and N-hydroxysuccinimide (13.8 mg, 0.12 mmol) were dissolved in 20 mL dry DMF. After the addition of N,N'-dicyclohexylcarbodiimide (24.8 mg, 0.12 mmol), the reaction was continued for 12 h under inert atmosphere at room temperature. Precipitated dicyclohexylurea (DCU) was removed by filtration. F4-diacid-NHS was used for coupling without further purification.

Paromomycin (0.3 mmol) was dissolved in 20 mM sodium phosphate buffer (20 mL, pH 7.0), a solution of F4-diacid-NHS in DMF (2.5 mL) was added drop by drop in 3 h. The reaction was carried out overnight under inert atmosphere at room temperature. The solvent was evaporated. The obtained crude mixture was purified by column chromatography.
by using homogeneous mobile phase consisting of dichloromethane/methanol/aq. 25% ammonia (from 2:2:1 to 2:3:2 v/v/v) mixture. After evaporation of the solvent, the residue was re-dissolved in water (3 mL) and traces of silica were removed by filtration through a 0.45 mm-syringe filter. Lyophilization yielded the **F-dimer** as orange solid. Yield: 51.2 mg (0.03 mmol, 56%). 1H NMR (500 MHz, D2O) (E-isomer) δ ppm 7.69 (m, 4 H, azo), 5.73 (d, J = 3.5 Hz, 2H, 1-H'), 5.40 (d, J = 1.5 Hz, 2H, 1-H''), 5.27 (s, 2H, 1-H'''), 4.44 (t, J = 5.5 Hz, 2H, 3-H'''), 4.34 (m, 2H, 2-H'''), 4.31 (m, 2H, 3-H'''''), 4.26 (t, J = 5.5 Hz 2H, 5-H'''''), 4.22 (m, 2H, 4-H'''), 4.01 (t, J= 7.8 Hz, 2H, 4-H), 3.92-3.82 (m, 10H, 6a-H''', 5a-H'', 6a-H', 3-H', 5-H). 3.81-3.68 (m, 8H, 5b-H'', 5-H', 6b-H', 4-H'''), 3.72 (t, J = 9.5 Hz, 2H, 6-H), 3.64 (m, 2H, 6b-H'''), 3.60(m, 2H, 2-H'''), 3.58 (t, J = 9.75 Hz, 2H, 3-H), 3.44 (t, J = 7.75 Hz, 2H, 4-H'), 3.40 (dd, J = 11.0 Hz, J = 4.0 Hz, m, 2H, 2- H'), 3.37(m, 2H, 1-H), 3.25 (dt, J = 13.0 Hz; J = 4.0 Hz, 2H, 2-He), 1.86 (dd, J = 12.0 Hz, 2H, 2-Ha). 13C-signals based on HSQC (D2O, 500 MHz) δ ppm 112.23 (azo), 109.65 (C-1'''), 92.13 (C-1''), 96.024 (C-1'''''), 84.12 (C-5), 81.98 (C- 4'''), 77.60 (C-4), 76.32 (C-3'''), 73.91 (C-5''), 73.87 (C-2'''), 72.48 (C- 5'''), 72.23 (C-6), 69.26 (C-4''), 68.82 (C-3'), 67.68(C-3'''), 66.34 (C-4'''), 60.37 (2C, C-5'', C-6''), 53.76 (C-2''), 51.09 (C-2'''''), 49.61 (C-1), 48.82 (C-3), 40.29 (C-6'''''), 28.08 (C-2). HRMS-ESI: m/z = 1537.62; 769.48 (calcd for [M+H]+, 1537.60; [M+2H]^{2+}/2=769.30).

**3.4.2. Isothermal Titration Calorimetry (ITC) measurement**

ITC experiments were executed using the ultrasensitive ITC 200 calorimeter (MicroCal) at 25°C. First, we degassed all solutions for 15 min using a vacuum pump in order to prevent the formation of bubbles in the sample cell during the experiment. We filled the reference cell with degassed distilled water and rinsed the sample cell with the buffer two times. We filled the sample cell with 7 μM aptamer solution (prepared in phosphate buffer solution with pH 6.8) and calorimeter syringe with 70 μM target ligand solution (prepared in phosphate buffer solution with pH 6.8). A purge-refill cycle was performed during filling up the syringe to avert the formation of air bubbles inside the syringe. To determine the binding constant, the ligand solution was added to the cell containing aptamer solution in a stepwise manner. The instruction of the instrument was set up in such a way that 20 injections of 2 μl volume from the syringe were added into the aptamer solution with intervals of 60 or 120 seconds between each injection. Control experiments were done by titrating the phosphate buffer without ligand into the aptamer solution. Data were analysed by using the nonlinear curve-fitting functions for one binding site provided by the ORIGIN software of MicoCal.
3.4.3. Molecular Dynamic Simulations

The structure of RNA aptamer (23 nucleotides) was taken from the PDB and is based on a published NMR structure\(^1\). The E/Z-isomers and RNA aptamer were simulated by NAMD\(^2\) and CHARMM\(^3\) force fields. The particle mesh Ewald (PME)\(^5\) method was used for the evaluation of long-range Coulombic interactions. The time step was set to 2 fs. The simulations were performed in the NpT ensemble (\(p = 1\) bar and \(T = 310\) K), using Langevin dynamics with a damping constant of 1 ps\(^{-1}\). After 2000 steps of minimization, ions and water molecules were equilibrated for 2 ns around RNA and E/Z-isomers, which were restrained using harmonic forces with a spring constant of 2 kcal/ (mol Å\(^2\)). The last frames of restrained equilibration were used to start simulations of free molecules and RNA aptamer. The simulations lasted for 1.9–2 μs. The calculations of binding energy and RMSD were performed by VMD plugin\(^6\). The binding energy is composed of vdW and electrostatic energy, where the dielectric constant is set to 1 during the calculation.

3.4.4. Determination of MIC values of F-dimers

The F-dimers were serial diluted from 512μM to 0.25 μM in 15 mL tubes. Overnight cultures of \(E.\ coli\) ATCC 25922 were diluted to an optical density at 600 nm of 0.1 and 100 μl of this cell suspension was added to total volume of 500 μl LB medium containing the F-dimers at the given concentration. After overnight growth at 37°C and shaking at 220 rpm, the optical density at 600 nm was measured. Graphs were background-corrected by subtracting the OD of the LB medium without any bacteria.

3.4.5. Growth of the \(E.\ coli\) cells and expression of the TMS switch as well as the protein

The conditions of growth of the \(E.\ coli\) BL21(DE3) cells and the expression of the proteins along with the switch was maintained in the same way as it was described in the previous chapter. To grow the bacteria on the plates, culture from the logarithmic growth phase was spread on the agar plates. The plates contain appropriate antibiotics (ampicillin 50 μg/ml and chloramphenicol 25 μg/ml) and the inducers to express the switch and GFP (1 mM IPTG to express the switch and 0.1 % arabinose (w/v) to express the protein). After spreading the cells, depending on the experimental conditions, the plates were either kept overnight at 37°C or were exposed to light and then kept overnight at 37°C.
3.4.6. Flow cytometry measurements

Flow cytometry measurements were performed on a BD FACS Canto flow cytometer. The BD FACS Canto was calibrated with CST beads from BD Biosciences (Cat No: 655051). The bacterial plates were treated with 1X PBS to disperse the bacterial colonies into the PBS buffer and were diluted by 100-fold into fresh 1X PBS buffer prior measurement. For GFP fluorescence measurement, a 488 nm excitation filter (optical power 20mW) and 515-545 nm emission filter were used. For each sample, 50,000 events were recorded. All the samples were measured with low sample flow rate (approximately 12 µl/min). The flow cytometry analysis was performed using FlowJo software (version 10). Note: ON/OFF ratios were calculated based on the positive population only.

Authors contribution

Andreas Herrmann, Lifei Zheng, Avishek Paul, Jingyi Huang and Yanxiao Han conceived the research. Jingyi Huang carried out most of the synthesis and characterization of the photo-isomers. Yanxiao Han performed the molecular dynamic simulation study. Avishek Paul performed all the work dealing with bacteria including the in situ study with the photo-dimers and analysed the flowcytometry data. Moreover, he performed the ITC measurements. Xintong Yang synthesized additional F-dimer. Andreas Herrmann, Lifei Zheng, Avishek Paul, Lela Vuković and Petr Král wrote the manuscript.

References


**Appendix**

**Supplementary Figure 3.1.** ITC titration curve of paromomycin (70 μM in 20 mM phosphate buffer pH 7.5) into aptamer solution (7 μM in 20 mM phosphate buffer pH 7.5); Kd = 0.2 μM.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>0.717 ± 0.0045 Sites</td>
</tr>
<tr>
<td>K</td>
<td>4.87E6 ± 1.44E5 M⁻¹</td>
</tr>
<tr>
<td>ΔH</td>
<td>-1.791E4 ± 158.7 cal/mol</td>
</tr>
<tr>
<td>ΔS</td>
<td>-31.5 cal/mol/deg</td>
</tr>
</tbody>
</table>
Supplementary Figure 3.2. $^1$H-NMR (500MHz, D$_2$O) spectrum of F-dimer. a) PSS after green light (>500 nm) irradiation; b) PSS after blue light (455 nm) irradiation; c) pure Z-F-dimer after HPLC purification; d) pure E-F-dimer after HPLC purification. (The peak at 4.79 ppm is the solvent peak of D$_2$O.)
Supplementary Figure 3.3. HSQC spectrum (500 MHz, D$_2$O) of F-dimer.

Supplementary Figure 3.4. HRMS-ESI analysis of F-dimer.
**Supplementary Figure 3.5.** UPLC traces (recorded at the corresponding isosbestic points) of purified $E$- and $Z$-isomer.

**Supplementary Figure 3.6.** Determination of thermal half-life for F-dimer at 37 °C in water in the dark. First, PSS was reached upon >500 nm irradiation, after which the absorption was measured at $\lambda_{\text{max}} = 324$ nm.
Supplementary Figure 3.7. Trajectories of F-dimers interacting with RNA aptamer.

Supplementary Figure 3.8. Number of contacting nucleotides with F-dimers.
Supplementary Figure 3.9. Binding energy between Z/E-isomers and RNA aptamer.

Supplementary Figure 3.10. MIC test of the $E$ and $Z$ F-dimers to determine the working concentration for gene switching experiments. Pristine paromomycin was added as reference.
Supplementary Figure 3.11. The effect of the concentration of the E-isomer on the GFP expression. (a) GFP expression with 64 μM E-isomer; (b) GFP expression with 32 μM E-isomer; (c) GFP expression with 16 μM E-isomer; (d) GFP expression with 8 μM E-isomer; (e) GFP expression with 4 μM E-isomer.

Supplementary Figure 3.12. The effect of the concentration of the E-isomer on the GFP expression. (a) GFP expression with 28 μM E-isomer; (b) GFP expression with 24 μM E-isomer; (c) GFP expression with 20 μM E-isomer.

Supplementary Figure 3.13. In situ Z to E photo-isomerization induced switch-on of GFP expression under the irradiation at 455 nm for different times.
Supplementary Figure 3.14. The effect of light (455 nm) on GFP expression. (a) Bright field image showing complete bacterial growth on the plate; (b) Image of the plates under UV lamp. It was observed that no GFP expression occurred with the treatment of light only.

Supplementary Figure 3.15. Changes in the absorption spectra of a solution of F-dimer (10 μM): upon irradiation at 530 nm (E→Z) in the absence (A) and in the presence of RNA aptamer (B); upon irradiation at 455 nm (Z→E) in the absence (C) and in the presence of RNA aptamer (D).
**Supplementary Figure 3.16.** In situ $E$ to $Z$ photo-isomerization induced switch-off of GFP expression under the irradiation at 530 nm for 1.5h.

**Supplementary Figure 3.17.** (a) Bright field image of the plate containing $Z$ isomer only but without adding the inducer to express the lysis gene; (b) Bright field image of the plate containing the inducers to express the TMS switch and the lysis gene but without adding $Z$ isomer.
Supplementary References


