Phys. Biol. **6** (2009) 025007 (10pp) doi:10.1088/1478-3975/6/2/025007

RNA synthetic biology inspired from bacteria: construction of transcription attenuators under antisense regulation

Alexandre Dawid[1], **Bastien Cayrol and Hervé Isambert**[2]

Institut Curie, Research Division, CNRS UMR168, 11 rue P. & M. Curie, 75005 Paris, France

E-mail: herve.isambert@curie.fr

Received 30 September 2008
Accepted for publication 12 February 2009
Published 1 July 2009
Online at stacks.iop.org/PhysBio/6/025007

Abstract

Among all biopolymers, ribonucleic acids or RNA have unique functional versatility, which led to the early suggestion that RNA *alone* (or a closely related biopolymer) might have once sustained a primitive form of life based on a single type of biopolymer. This has been supported by the demonstration of processive RNA-based replication and the discovery of 'riboswitches' or RNA switches, which directly sense their metabolic environment. In this paper, we further explore the plausibility of this 'RNA world' scenario and show, through synthetic molecular design guided by advanced RNA simulations, that RNA can also perform elementary regulation tasks on its own. We demonstrate that RNA synthetic regulatory modules directly inspired from bacterial transcription attenuators can efficiently *activate* or *repress* the expression of other RNA by merely controlling their folding paths 'on the fly' during transcription through simple RNA–RNA antisense interaction. Factors, such as NTP concentration and RNA synthesis rate, affecting the efficiency of this kinetic regulation mechanism are also studied and discussed in the light of evolutionary constraints. Overall, this suggests that direct coupling among synthesis, folding and regulation of RNAs may have enabled the early emergence of autonomous RNA-based regulation networks in absence of both DNA and protein partners.

M This article features online multimedia enhancements

1. Introduction

Life, as we know it, relies on a handful of different biopolymers: DNA, RNA, proteins, heteropolysaccharides (from an extracellular matrix or bacterial cell walls), etc. But the biological functions of these various biopolymers appear intricately intermingled in all extant life forms. This is best illustrated by the circular dependence of biopolymer synthesis: DNA is transcribed into RNA and then itself translated into proteins, which in turn control the expression and replication of the DNA blueprint. While this complex interdependence has presumably been in use for more than 2.5 BY, maybe even 3.5 BY corresponding to the oldest known prokaryote looking fossils, it is probable that evolution started with a simpler form of primitive life based on a single biopolymer.

While there is still no definitive proof of this scenario, it has long been suggested that RNA (or a closely related biopolymer) might have once sustained such a primitive form of life, commonly referred to as the 'RNA world' [1, 2]. Indeed, RNA stands out from the other biopolymers for its pivotal role in several fundamental cellular functions, such as the 4- to 20-letter genetic translation by tRNAs or the peptidyl transferase activity of rRNAs for protein synthesis. In addition, RNA also exhibits striking functional versatility from genetic blueprint (mRNA, RNA virus genomes, etc) to enzymes (ribozymes, rRNA, etc) and regulators of gene expression (ncRNA, miRNA, siRNA, etc), while the functional scope of DNA, proteins and other biopolymers is typically more specialized.

[1] Present address: AMOLF–Kruislaan 407, 1098 SJ Amsterdam, The Netherlands.
[2] Author to whom any correspondence should be addressed.

The plausibility of this 'RNA world' hypothesis which has long been associated with the demonstration of processive RNA-based replication [3–9] recently gained further support with the discovery of 'riboswitches' [10–14] or RNA switches that directly probe the concentration of small metabolites to control the expression of specific genes. Yet, beyond the ability to accurately replicate and adapt to varying metabolic environments, 'primitive' RNA, if they ever existed, must have also regulated their own expressions in the absence of both DNA and proteins. Indeed, elementary regulatory functions such as 'gene-to-gene' activation or repression are a likely prerequisite for the emergence and evolutionary adaptation of any 'multigene' living system.

In this area, recent advances to unravel the design principles of genetic circuits have come with the development of synthetic biological networks inspired by bottom-up design approaches of engineers and physicists. In particular, a variety of simple synthetic biological networks have been successfully designed using regulatory proteins, i.e. transcription factors [15–19], as well as 'riboregulators', i.e. regulatory RNA controlling the translation of proteins [20–23]. Synthetic biological systems including cell-free translation have even been constructed [24]. More recently, Penchovsky and Breaker [25] have designed synthetic ribozymes activated via antisense interactions, and Kim *et al* [26] have designed a bistable translation-free circuit relying on the control of transcription initiation by phage T7 RNA polymerase through hybrid RNA–DNA interactions using split T7 DNA promoter regions [27].

In this paper, we show through synthetic biology design that RNA 'alone' can efficiently activate or repress the expressions of other RNA by controlling their folding paths 'on the fly' during transcription through simple RNA–RNA antisense interactions. Although regulatory and target RNA are still produced by transcription from an RNA polymerase and a DNA template, actual regulatory decisions solely rely on RNA–RNA antisense interactions controlling the formation of alternative folded structures of the nascent RNA chains.

The synthetic RNA regulatory modules presented here are in fact directly inspired from a transcription antitermination mechanism that is well known in bacteria. Yet, in bacteria, RNA terminator/antiterminator switches do not typically operate under antisense RNA control but instead under protein or metabolite control [28]. There are, however, a few interesting exceptions, such as the control of a plasmid copy number through antisense RNA-mediated transcription attenuation [29, 30] or tRNA-mediated transcription antitermination in *B. subtilis* and other Gram-positive bacteria [31, 32]. Yet, the regulation efficacy in these natural examples is typically around 50–60% (i.e. about two-fold variations) even in the presence of additional proteins or other co-factors *in vitro*. By contrast, the synthetic RNA regulatory modules we have designed in this paper exhibit >80–90% activation or repression efficacy *in vitro* (i.e. five- to ten-fold variations) without any additional co-factors.

2. Results

2.1. Principles of transcription attenuators

The synthetic RNA regulatory modules we have designed are directly inspired from transcription attenuator switches that are commonly found in bacteria [28]. These regulatory switches rely on alternative co-transcriptional folding paths of mRNA or operon 5′ leader regions (figure 1(*A*)) that form either (i) a terminator hairpin (figure 1(*B*) (right)) promoting premature transcription termination [33] or (ii) an antiterminator helix (figure 1(*B*) (left)) preventing the formation of the terminator, thereby enabling transcription to go to completion.

Terminators of transcription are simple hairpin helices of typically 10–15 GC rich base pairs directly followed by a 6- to 9-nucleotide stretch of U rich bases [33], while antiterminators are any local base pairing arrangements able to prevent the formation of a specific terminator during RNA transcription [34].

Hence, by contrast with the control of transcription initiation by transcription factors, which typically bind to DNA promoter regions under equilibrium conditions, transcription attenuator switches must operate out of equilibrium, while the nascent RNA chain is still being transcribed from its 5′ to 3′ ends. In practice, the antiterminator has to form within a short time window (e.g. 50–300 ms depending on the transcription rate) *before* the 3′ strand region of the abortive terminator is added along the nascent RNA chain. Moreover, the early formed antiterminator must also remain in place once the terminator is available to compete for alternative folds.

In practice, the antiterminator efficacy at preventing terminator formation can simply rely on a significant overlap with the hairpin loop of the terminator (figure 1(*A*)). For example, a 4–5 base pair overlap readily prevents the nucleation of the terminator from the *top* of its hairpin, which could otherwise enable a rapid exchange between the two helices through a simple branch migration process, as demonstrated in [35]. In addition, the alternative nucleation of the terminator from its *base* (figure 1(*B*) (center)) leads to a topologically blocked transition intermediate caused by the local entanglement of the antiterminator helix *inside* the loop of the partially formed terminator hairpin (figure 1(*B*) (center)). Such a *knotted* intermediate [36] is a simple yet efficient way to prevent the formation of the complete terminator with an antiterminator of comparable length and free energy.

This ability of the antiterminator to block terminator formation, through sequence overlaps with terminator hairpin loops and topologically knotted intermediates, has been directly used in the RNA synthetic regulatory modules we have designed and tested below.

Similar antiterminator *versus* terminator arrangements are also typically observed in 5′ leader regions of bacteria that include natural transcription attenuators, as illustrated in figure 1(*C*).

2.2. Design of synthetic RNA regulatory modules

While natural transcription attenuators are usually controlled upon metabolite or protein binding to the nascent RNA leader

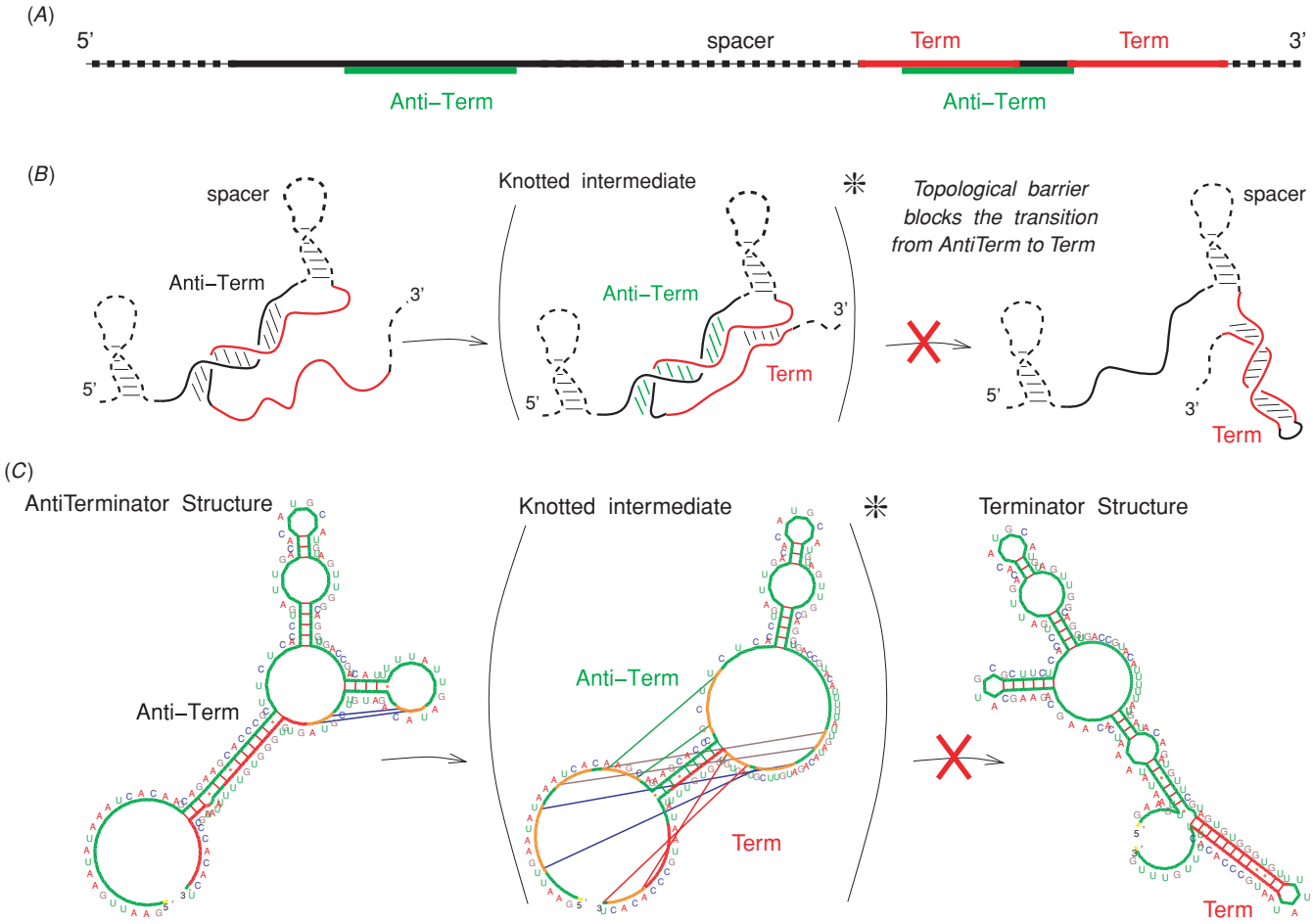

Figure 1. Principles of transcription attenuators under topological control. (*A*) Relative positions of antiterminator and terminator helices in a transcription attenuator switch (see text). A spacer region enables intermolecular antisense interactions to occur before the antiterminator may form during transcription. (*B*) Formation of the antiterminator prevents subsequent terminator formation through the depicted topologically blocked intermediate (see text). (*C*) Example of a natural transcription attenuator in the 5′ untranslated region of *infC* gene from *Bacillus subtilis* [31].

region [28], we have instead designed synthetic regulatory modules under RNA–RNA antisense control, in order to explore the ability of RNA to regulate, on its own, the expression of other RNA. Interestingly, a few examples of natural transcription attenuators are also known to be regulated through antisense interactions [31, 32]. Yet, their regulation efficiency is typically around 50–60% (i.e. about two-fold variations) even in the presence of additional proteins or other co-factors *in vitro*.

In order to design efficient synthetic regulatory modules based on transcription attenuator control, we have encoded specific features on their molecular sequence and folding path.

First, sequences of the regions targeted by antisense regulatory RNAs have been chosen to be essentially symmetric, in order to limit both their intramolecular folding and possible intermolecular self-interaction. These features favor rapid antisense recognition and interaction with their specific regulatory RNA in solution.

Second, sequence spacers have been added downstream of the antisense target regions to provide enough time for the intermolecular antisense interactions to occur during transcription, that is, before intramolecular base pairs eventually form with the antisense target region (e.g. to form an antiterminator as in figures 1(*A*) and (*B*)).

Third, folding paths and antisense control of potential RNA regulatory modules were thoroughly tested using a kinefold online server [36] (see section 5). The predictions of this server, which simulates the stochastic folding dynamics of RNA or DNA sequences including knotted and pseudoknotted structures, have already been successfully benchmarked on a variety of small pseudoknots [37], as well as on longer sequences such as 394-nt *Tetrahymena* group I intron [38]. Kinefold predictions of folding paths have also been compared to folding experiments probing natural RNA [37], designed RNA [35] or DNA [39] sequences. In addition, the molecular basis of these stochastic folding simulations was directly tested with detailed analysis of single molecule micromechanical unfolding/refolding experiments on small RNA motifs and the 16S ribosomal RNA of *E. coli* [40].

Applying these design principles and advanced RNA simulations, we have successfully constructed two synthetic RNA regulatory modules with opposite functions, that is, one synthetic repressor module S1/R1 and one synthetic activator module S2/A2. Their sequences are given in section 5.

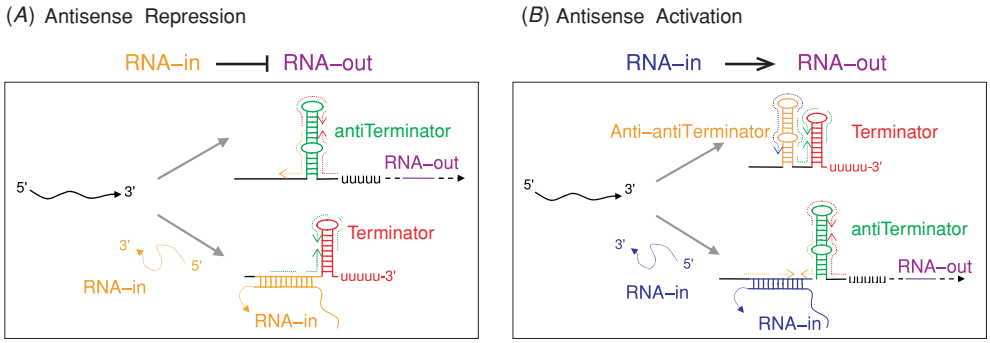

Figure 2. Design of synthetic transcription attenuators under antisense regulation. (*A*) Synthetic antisense repressor module. (*B*) Synthetic antisense activator module. See the main text.

Design of a synthetic RNA repressor. The synthetic repressor of transcription elongation is directly based on the transcription attenuator principle depicted in figures 1(*A*) and (*B*). Its regulation under antisense control is illustrated in figure 2(*A*). In the absence of an antisense regulatory RNA, RNA-in, an antiterminator forms enabling transcription to go to completion to synthethize a full RNA-out transcript. Conversely, in the presence of RNA-in, an antisense interaction sequesters the first strand of the antiterminator to prevent its formation, thereby inducing a terminator to form and promote premature termination of transcription. Advanced simulations using the kinefold online server [36–38] have enabled us to optimize two sequences, R1=RNA-in and S1=RNA-out, to best fit the proposed design principles and constraints numerically before testing the actual construct experimentally (see below and section 5). The structure snapshots in figures 3(*A*) and (*B*) summarize the simulated folding paths for this specific repressor module and its regulation under antisense interaction, following the design principles outlined in figure 2(*A*). The R1 sequence is simply attached upstream of S1 via a non-pairing linker for the simulation studies. Note that R1 remains eventually bound by antisense interaction to S1 after having induced its premature termination (figure 3(*B*)). The animations of the corresponding folding paths, figures 3(*A*) and (*B*), are available at stacks.iop.org/PhysBio/6/025007.

Design of a synthetic RNA activator. The second regulatory module is a synthetic activator of transcription elongation. Its regulation under antisense control, illustrated in figure 2(*B*), is adapted from the previous repressor design with the addition of an upstream region that promotes the early formation of an anti-antiterminator helix preventing, as its name suggests, the formation of the antiterminator. In the absence of the antisense regulatory RNA, RNA-in, the anti-antiterminator is formed enabling the terminator to fold and induce premature transcription termination of the RNA-out transcript. Conversely, in the presence of the regulatory sequence, RNA-in, the antisense interaction sequesters the first strand of the anti-antiterminator to prevent its eventual formation, thereby enabling the antiterminator to form and thus preventing the formation of the terminator. Hence, RNA-in promotes transcription of the full RNA-out transcript. Again, simulations using kinefold have enabled us to optimize two sequences, A2=RNA-in and S2=RNA-out, to best fit the proposed design principles and constraints before testing the actual construct experimentally (see below and section 5). The structure snapshots in figures 3(*C*) and (*D*) summarize the simulated folding paths for this specific activator module and its regulation under antisense interaction, as depicted in the design principles outlined in figure 2(*B*). The A2 sequence is simply attached upstream of the S2 sequence via a non-pairing linker for the simulation studies. Note that A2 only transiently interacts with the S2 nascent transcript to prevent its premature termination (figure 3(*D*)). The antisense interaction between A2 and S2 is then displaced through a branch migration exchange with a stronger intramolecular interaction encoded downstream on the S2 transcript [35]. This enables continuous recycling of A2 regulatory RNA with multiple turnover of S2 activation. The animations of the corresponding folding paths, figures 3(*C*) and (*D*), are available at stacks.iop.org/PhysBio/6/025007.

2.3. Experimental results

Each synthetic regulatory module was then experimentally tested following sequence optimization and simulation studies (see above). The sequences of the repressor and activator modules designed above were first constructed and cloned as described in section 5. PCR fragments with different promoters for either *E. coli* or phage T7 RNA polymerases were then synthesized, purified and used as DNA templates for *in vitro* transcription assays.

Efficient synthetic RNA regulatory modules. The most efficient regulatory modules in terms of expression variations were obtained with the *E. coli* RNA polymerase and are presented in figure 4(*A*) for the repressor module and in figure 4(*B*) for the activator module. The analysis of their efficiency was done by separating RNA transcripts on denaturing gels stained with ethidium bromide.

The repressor module S1/R1 (figure 4(*A*)) exhibits an excellent regulatory efficiency starting from a negligible basal repression in the absence of R1 (i.e. $<1\%$ of the 'S1 stop' transcript without R1) to almost complete repression under micromolar concentration of R1 (i.e. $>90\%$ of the 'S1 stop' transcript for [R1]>1 μM) (figure 4(*A*)). This contrasts with

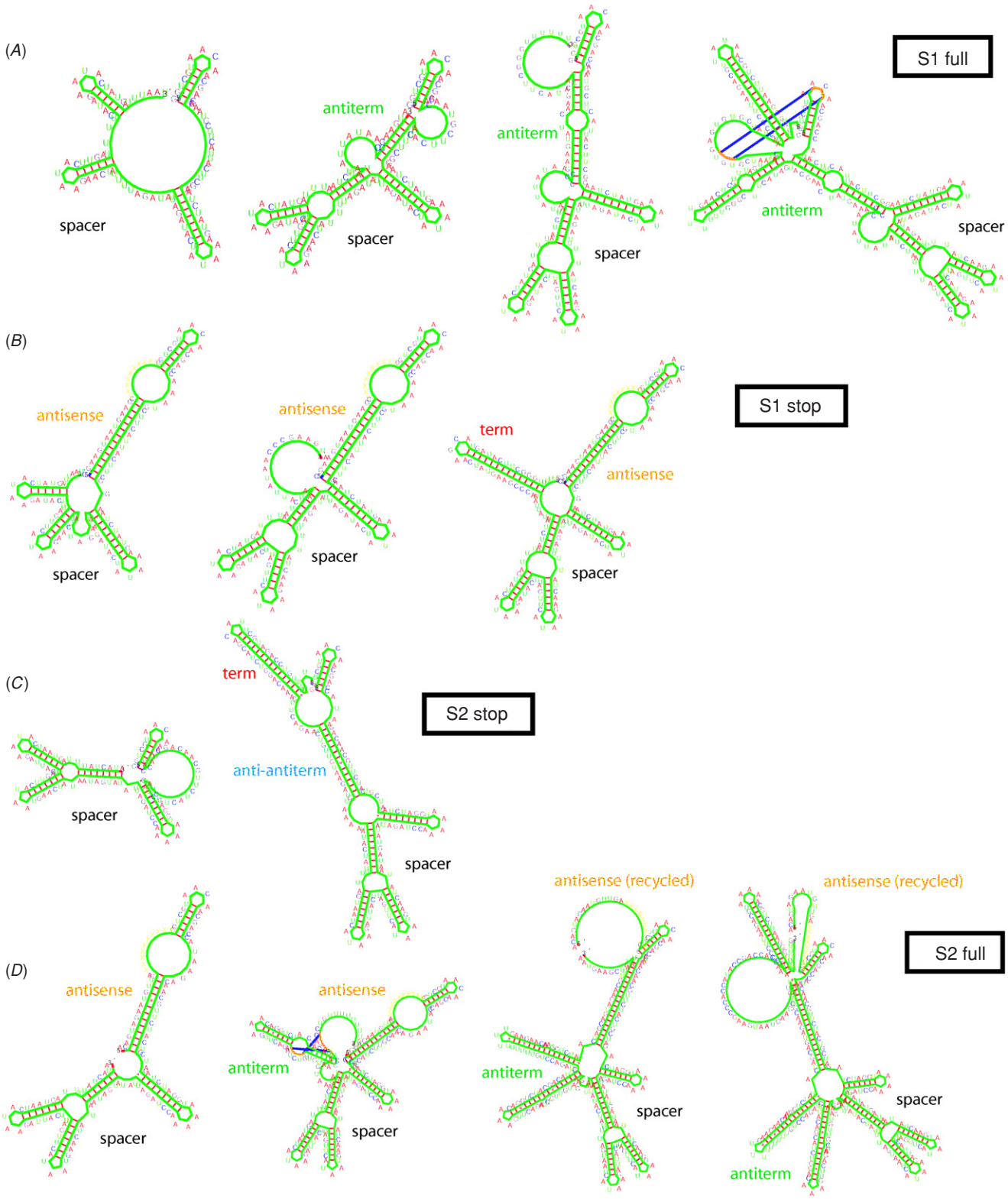

Figure 3. Snapshots of co-transcriptional folding intermediates along the folding paths. (*A*) S1. (*B*) S1 under R1 repression. (*C*) S2. (*D*) S2 under A2 activation. See the main text for details and stacks.iop.org/PhysBio/6/025007 for complete mpeg animations.

the relatively low 'regulatable fraction' of around 50% usually exhibited in natural systems [41].

The activator module S2/A2 (figure 4(*B*)) also exhibits a good regulatory efficiency although with a narrower range of expression variation than the repressor module S1/R1. Starting from weak but non-negligible basal activation in the absence of A2 (i.e. about 17% of the 'S2 full' transcript without A2), we obtained full activation under large A2 concentration

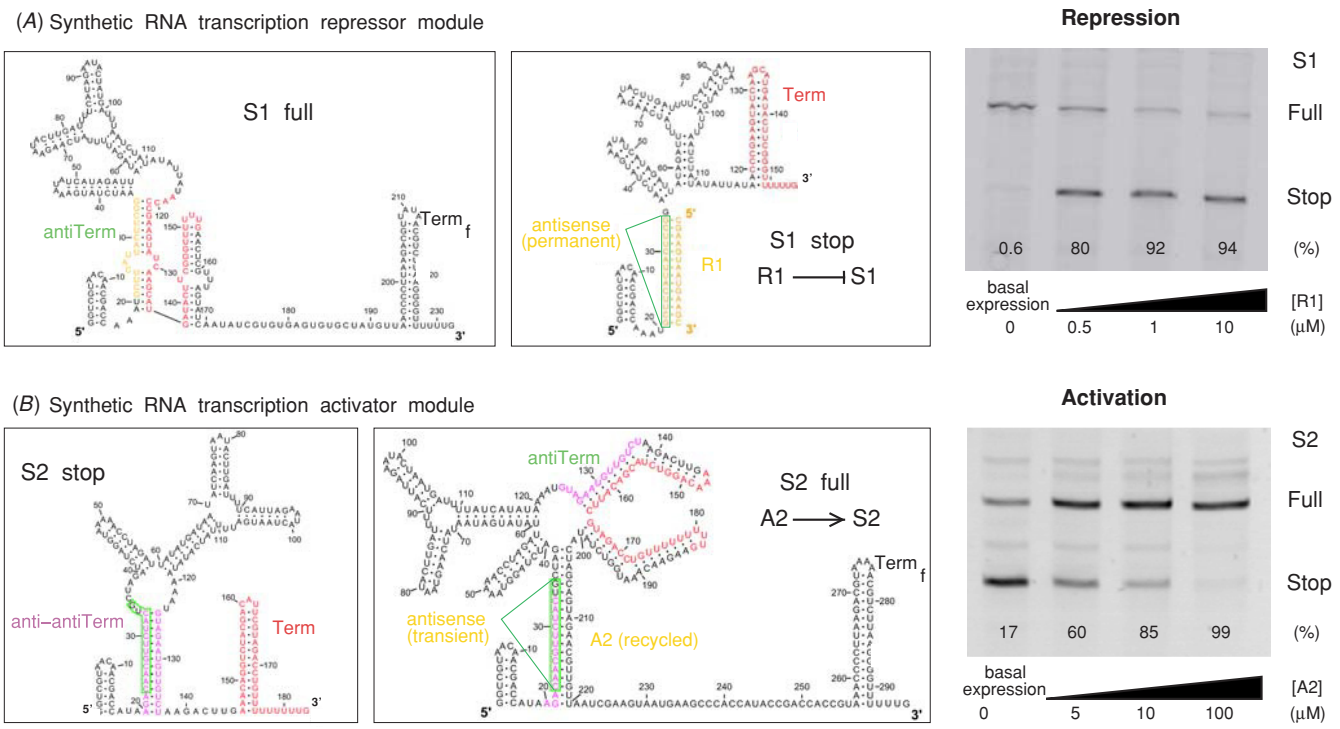

Figure 4. Efficacy of synthetic RNA regulatory modules. (*A*) Synthetic antisense repressor module S1/R1. (*B*) Synthetic antisense activator module S2/A2. See the main text.

(i.e. >85% of the 'A2 full' transcript for [A2] > 10 μM) (figure 4(*B*)).

A constitutive feature of both repressor and activator modules is that they have been designed to operate through RNA–RNA antisense interactions which must take place as the nascent RNA chain is still being transcribed. This kinetic regulatory mechanism, unlike thermodynamic controls of transcription initiation, implies that factors affecting transcription rates and/or the efficacy of transcription termination are expected to have direct effects on the regulatory efficiency of these synthetic RNA repressor and activator modules. This is because the transcription rate should directly control the time window available for RNA–RNA recognition and interaction, while the efficacy of transcription termination stems from the actual readout of the terminator signal by the RNA polymerase.

To investigate these points and explore the limitations of these designed regulatory modules, we have studied two factors that are known to affect both the rate of transcription and the efficacy of transcription termination. These are (i) the effect of varying NTP concentrations and (ii) the use of different RNA polymerases.

Effect of varying NTP concentrations. Decreasing the concentration of ribonucleotides, ATP, UTP, GTP and CTP, in solution is known to increase delays at pause sites and decrease the overall rate of transcription [42, 43]. In accordance with these results, figure 5 demonstrates that decreasing NTP concentrations threefold from [NTP] = 300 μM (figure 5(*A*)) to [NTP] = 100 μM (figure 5(*B*)) reduces both the critical activation concentration by A2 by about threefold and the spurious basal activation of S2 (from

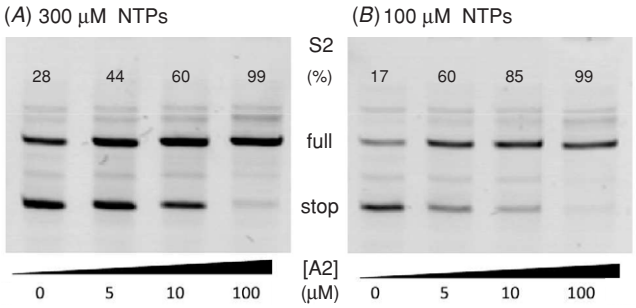

Figure 5. Effect of the NTP concentration. Synthetic antisense activator module S2/A2 with (*A*) 300 μM and (*B*) 100 μM NTPs.

28% to 17%). This is indeed in agreement with the increase of the time window enabling antisense activation by A2 and the role of transcription pauses in transcription termination at terminator sites [42, 43].

Effect of different RNA polymerases. Similarly, using the faster RNA polymerase of phage T7 (figure 6(*A*)) instead of the *E. coli* RNA polymerase, as above, on an S1/R1 repressor module (figures 4(*A*) and 6(*B*)) we found (i) that the critical concentration of repressor R1 increases by more than tenfold ($C_{T7} > 10C_{EC}$) and (ii) that spurious activation of S1 also increases at high R1 concentration (65% spurious activation with T7 RNApol instead of 6% with *E. coli* RNApol (figures 6(*A*) and (*B*)) at [R1] = 10 μM).

The increase of critical repression concentration is in good agreement with the reported six-fold increase in transcription velocity [44, 45] between *E. coli* RNApol ($v_{EC} \simeq 50$ nt s^{-1}) and T7 RNApol ($v_{T7} \simeq 300$ nt s^{-1}), assuming a

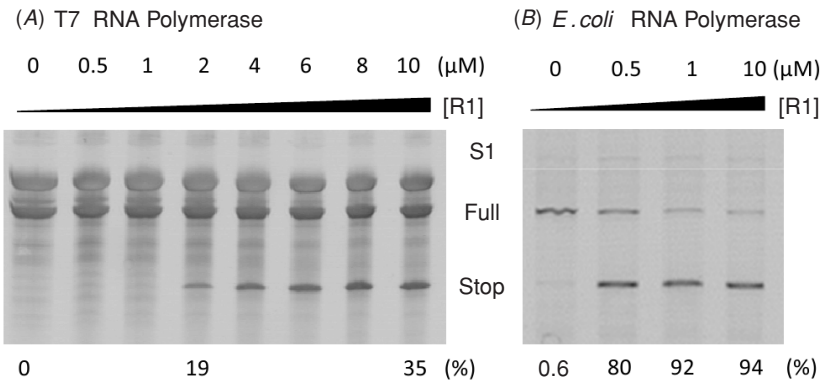

Figure 6. Effect of the RNA polymerase. Synthetic repressor module S1/R1 transcribed with (*A*) T7 RNApol and (*B*) *E. coli* RNApol.

diffusion-limited process for R1 antisense interaction. Indeed, the available time window Δt for antisense interaction should scale as $\Delta t \sim C^{-2/3}$ under a diffusion-limited process, where C is the antisense concentration in solution. This leads to the following scaling relation, $C \sim v^{3/2}$, between the antisense concentration and the transcription rate $v = L/\Delta t$, where L corresponds to the fixed length of the sequence linker delaying the antiterminator formation. Hence, we expect $C_{T7} \simeq 6^{3/2} C_{EC} > 10 C_{EC}$, in broad agreement with the above observation (figures 6(*A*) and (*B*)).

On the other hand, the low efficacy of the phage T7 RNA polymerase at interpreting terminator signals (figure 6(*A*)) is likely related to its natural function in the context of bacterial infection. Indeed, the only known natural terminator on phage T7 genome has about 50% functional efficacy to allow for significant readthrough transcription of further downstream genes [46]. In any event, this poor efficacy at interpreting terminator signals makes the T7 RNA polymerase largely inadequate in the context of synthetic regulatory modules based on termination/antitermination mechanisms. In addition, we also found that T7 RNApol tends to terminate transcription unspecifically at many long helices irrespective of their actual location in the co-transcriptionally formed structure of the nascent RNA transcript. This further limits the possibilities of using T7 RNApol with synthetic RNA regulatory modules.

3. Discussion

3.1. Simulations versus experiments comparison

By and large, the design of synthetic antisense regulatory modules based on the principle of transcription attenuation (figure 1) and kinefold simulations [36–38] proved to be reliable when compared to the observed experimental properties of the corresponding constructs. This is in line with previously reported agreements between this type of RNA folding/unfolding simulations and the corresponding experiments [36–38, 40, 47]. In fact, most of the discrepancies we have found in designing the regulatory modules presented here did not appear to be due to erroneous folding path predictions, but could generally be attributed to other causes, such as the unanticipated formation of antisense dimers (which prompted us to enforce sequence symmetry to prevent any type of self-interaction; see above) or even the formation of triple hybrid helices DNA/DNA/RNA on the dsDNA template (not shown).

But the main problem we encountered, with the experimental construction of RNA regulatory modules, concerned the defective readout of the expected folded structures by the RNA polymerase, especially with T7 RNApol transcription. This initially turned out to be a serious problem, as our first experimental regulatory modules were tested with this polymerase (figure 6(*A*)) which transcribes faster and more abundantly than bacterial RNA polymerases, owing to its simple single subunit structure. However, we found that termination signals were poorly interpreted by T7 RNApol, which rarely exceeds 60% termination efficacy, presumably due to its own natural requirement for readthrough transcription [46]. In addition, we also found that T7 RNApol tends to terminate transcription unspecifically at many long helices. All in all, this limits the possibilities of using T7 RNApol with synthetic RNA regulatory modules.

3.2. Constraints and evolution of the regulatory modules

In this study, we demonstrate through synthetic biology design that RNA transcripts can efficiently *activate* or *repress* the expression of other RNAs by controlling their folding paths 'on the fly' during transcription through simple RNA–RNA antisense interactions. Although the practical efficacy of these regulatory modules hinges on the imperfect readout of terminator signals by an RNA polymerase as discussed above, the actual regulatory decisions solely rely on RNA–RNA antisense interactions which control the formation of alternative folded structures of the nascent RNA chains.

In addition, beyond this particular case of transcription attenuators under antisense regulation, it is clear that relatively limited information is necessary to encode efficient RNA regulatory modules: it essentially amounts to encoding the relative lengths and positions of helices forming successively during transcription (figures 2 and 3) as discussed in more detail in [35]. In addition, the use of knotted intermediate structures (figure 1) preventing a branch migration exchange between competing helices of comparable length and strength also provides a simple but reliable mean to encode efficient folding paths and their regulation under antisense control.

Yet, it remains that the intrinsically kinetic control of the RNA regulatory modules we have designed relies on RNA–RNA antisense interactions, which must form within a short time window. Such kinetic control of RNA regulation clearly brings specific operative constraints (short time scale, micromolar antisense concentration, etc) as compared to more elaborate regulatory systems based on thermodynamic control of the initiation of transcription. This may explain why more complex regulatory systems, with flexible thermodynamic control and possibly other additional partners than solely RNA [48] eventually appeared in the course of evolution, although simple RNA-based regulatory modules similar to those designed here are still found in extant bacteria [31, 32].

4. Conclusion

All in all, the synthetic RNA repressor and activator we have designed here suggest that efficient RNA-based regulation modules might have easily emerged and continuously adapted the same way the functional native structures have evolved through mutation drift and adaptative selection. Such antisense RNA regulatory modules could have provided simple 'all-RNA' mechanisms to control the expression of other RNA in the absence of any elaborate control at the level of transcription initiation.

From an ancestral 'RNA World' perspective, this direct coupling among RNA synthesis, folding and regulation under simple RNA–RNA antisense interaction may have enabled the early emergence of autonomous RNA-based networks relying solely on intra- and intermolecular base pair interactions. Interestingly, Lincoln and Joyce [49] have recently reported the design of two cross-replicating RNA enzymes catalyzing each other's ligation assembly from half-enzyme sequence pieces. Hence, while there is still some way to go to demonstrate that a purely RNA-based living system with elementary regulation capacity can be designed by combining a processive RNA-based RNA polymerase with an RNA-based regulatory system, decisive steps continue to be made in that direction [49].

5. Material and methods

5.1. Design of the RNA regulatory modules

The design of each RNA switch is accomplished in two main steps. First, the global architectures of the two alternative folds (terminator fold or antiterminator fold) are drawn, defining the position, size and order of formation of the helices that form during the folding process toward each final fold (see figure 1). Only standard Watson–Crick base pairing applying to RNA is considered (canonical A:U, and G:C pairs, as well as G:U pairs). At this step, although the precise sequence is not yet defined, the anticipated constraints of complementarity between bases in alternative folds are taken into account. Second, the sequence is designed by an iterative process according to the following method: first, sequences of the target regions and the corresponding antisense RNAs are designed such that internal folding or self pairings are prevented; this is mainly done by avoiding as much as possible the use of complementary bases for the same strand and by imposing a symmetry center into the sequence. Then, an initial candidate sequence satisfying the complementarity constraints is specifically assigned for the rest of the sequence, and computational simulations are performed to test its co-transcriptional folding using the online kinefold server [36] at http://kinefold.curie.fr. This program performs a stochastic simulation of RNA folding during elongation of the transcript and returns a time-lapse movie of the co-transcriptional folding process as well as a helix tracing graph. The dynamics of the process is not assessed at the level of individual base pairs but rather at the level of helix nucleation and subsequent exchange during transcription. Each simulation corresponds to a single stochastic realization, so that three to five simulations are performed for any condition (see below) to ensure a reasonable statistic on the folding path of a specific sequence. Results are inspected to control that the sequence performs the proper desired folding path during elongation. Otherwise, the sequence is modified in order to correct for the improper folding and simulations are repeated. To test for the proper functioning of the switch, simulations are performed with and without the presence of the corresponding antisense RNA. Moreover, as a specificity test, simulations with non-complementary antisense RNA are also performed to ensure that they do not interfere with the native folding path of the RNA switch (e.g. the antisense A2 for the activator is used as a specificity test for the repressor S1, and more generally parts of the sequence designed to stay single stranded in the final structure and that may therefore accumulate in the environment are used as non-complementary antisense RNA). Furthermore simulations are also performed in the presence of combinations of multiple complementary and/or multiple non-complementary antisense-RNA in combination to assess for possible combinatorial effects. Sequences that pass through all the controls are used for *in vitro* experiments.

5.2. Oligonucleotides, enzymes, plasmids, strains and cloning

DNA and RNA oligonucleotides were purchased from IBA (Göttingen Germany). The *E. coli* RNA polymerase was purchased from USB Corporation. T7 RNA polymerases, nucleotides and other materials were purchased from Promega and New England Biolabs. DNA substrates coding for the RNA regulatory modules have been obtained as follows. First the DNA sequences coding for the RNA regulatory modules as well as for the core sequence of the T7 10 late promoter are constructed from 6 oligonucleotides phosphorylated, hybridized and ligated, before being inserted into the multiple-cloning site of a pUC19 plasmid and cloned in a DH5alpha *E. coli* strain.

The repressor sequence has been obtained from the following three oligonucleotide pairs and cloned between the BamHI and EcoRI sites of pUC19:

*S*1_1w : GATCCTAATACGACTCACTATAGGTCGTAACAACGA-
CCAAATGCTTCATTACTTCGGAATCTATGAAATATCA-
TAGATTATAGATTTTATCAAGAAT

*S*1_1c : TCAAGTATTCTTGATAAAATCTATAATCTATGATAT-

TTCATAGATTCCGAAGTAATGAAGCATTTGGTCGTT-
GTTACGACCTATAGTGAGTCGTATTAG

*S*1_2w : ACTTGATTTTCATAGAATACTATGATTTAATCTATA-
TATTATAACCCGAAGTATCAAGCATGATACTTCGGG-
TTTTTTGAACTCGTTTAGTATC

*S*1_2c : CGATATTGATACTAAACGAGTTCAAAAAACCCGAAG-
TATCATGCTTGATACTTCGGGTTATAATATATAGAT-
TAAATCATAGTATTCTATGAAAA

*S*1_3w : AATATCGTGTGAGTGTGCTATGTTAACCCCTTAAGA-
CGTTATAACGTCTTAAGGGGTTTTTTG

*S*1_3c : AATTCAAAAAACCCCTTAAGACGTTATAACGTCTTA-
AGGGGTTAACATAGCACACTCACA

The activator sequence has been obtained from the following oligonucleotide pairs and cloned between the SalI and BamHI sites of pUC19:

*S*2_1w : TCGACTAATACGACTCACTATAGGTCGTAACAACGAC-
CATAAGACAACGTTCTACTGCTAGATCTAGGTAAAAA-
CCTAGATTATATGATAATATCAAGTAATACTTGATT

*S*2_1c : AATGAAAATCAAGTATTACTTGATATTATCATATAAT-
CTAGGTTTTTACCTAGATCTAGCAGTAGAACGTTGTC-
TTATGGTCGTTGTTACGACCTATAGTGAGTCGTATTAG

*S*2_2w : TTCATTAGAATACTAATGATTTTATCATATAAAATGT-
AGAATGTTGTCTAAGACTTGAAACAGGTCTACGACAT-
TCGTAGACCTGTTTTTTTTTGAAGAACAAATGGTCTA-
TAC

*S*2_2c : CTGCTAGTATAGACCATTTGTTCTTCAAAAAAAAACA-
GGTCTACGAATGTCGTAGACCTGTTTCAAGTCTTAGA-
CAACATTCTACATTTTATATGATAAAATCATTAGTAT-
TCT

*S*2_3w : TAGCAGTAGAACGTTGTTAATCGAAGTAATGAAGCCC-
ACCATACCGACCACCGTAAACCCCTTAAGACGTAAAA-
ACGTCTTAAGGGGTTTTTTTTTG

*S*2_3c : GATCCAAAAAAAAACCCCTTAAGACGTTTTTACGTCT-
TAAGGGGTTTACGGTGGTCGGTATGGTGGGCTTCATT-
ACTTCGATTAACAACGTTCTA

PCR fragments of these sequences were synthesized, purified and used as a DNA substrate for the *in vitro* transcription assay. For transcription by the T7 RNA polymerase, primers M13RSP [AGCGGATAACAATTTCACACAGGA] and M13-FSP [GTTTTCCCAGTCACGACGTTGTA] were used. For transcription by the *E. coli* RNA polymerase, primer PT7A1 [ATATAAAGCTTAAAAGATTAATTTAAAATTTATCAAAAAGAGTAT—TGACTTAAAGTCTAACCTATAGGATACTTACAGCCGGTCGTAACA—ACGACCA] along with M13-FSP was used to substitute the 10 late promoter of T7 by the strong major promoter T7A1 for the *E. coli* RNA-polymerase. Antisense RNA R1 [CGAAGUAAUGAAGC] and A2 [ACAGUAGAAUGUUGA] were purchased from IBA (Göttingen).

5.3. Transcription assays and PAGE analysis

In vitro transcriptions by the *E. coli* RNA polymerase were performed for 30 min up to several hours at 37 °C (figure 4(*A*)) or 27 °C (figures 4(*B*), 5(*A*) and 5(*B*)) in 20 μL of a transcription buffer (20 mM Tris-HCl pH8, 20 mM KCl, 10 mM MgCl2, 10 mM DTT, 300 μM NTPs and 50 μg mL^{-1} BSA) with 1 pmol of the DNA template and one unit of the RNA polymerase. *In vitro* transcriptions by the *E. coli* T7 polymerase were performed according to New Englands Biolabs protocols. RNA transcripts were purified using the RNeasy Minikit (Qiagen) and analyzed using denaturing PAGE separation in 15% acrylamide gel containing 8 M urea. The sample was pre-heated at 80 °C in a denaturing gel-loading buffer containing 60% formamide, 12 mM EDTA, 0.03% bromophenol blue and 0.03% xylene cyanol. Electrophoresis was performed at 30 V cm^{-1} for 1 h at ∼50 °C. RNA migration was analyzed after staining by ethidium bromide and quantified using a trans-illuminator.

Acknowledgments

We thank H Putzer, C Lemoine, V Viasnoff and L Hirschbein for discussion. This work was supported by AC grant DRAB04/117, ANR grant PNANO06-015-01 and HFSP grant RGP36/2005.

References

[1] Woese C 1968 *The Genetic Code* (New York: Harper and Row)
[2] Gilbert W 1986 The RNA world *Nature* **319** 618
[3] Been M D and Cech T R 1988 RNA as an RNA polymerase: net elongation of an RNA primer catalyzed by the Tetrahymena ribozyme *Science* **239** 1412–6
[4] Doudna J A and Szostak J W 1989 RNA-catalysed synthesis of complementary-strand RNA *Nature* **339** 519–22
[5] Johnston W K, Unrau P J, Lawrence M S, Glasner M E and Bartel D P 2001 RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension *Science* **292** 1319–25
[6] Bartel D P, Doudna J A, Usman N and Szostak J W 1991 Template-directed primer extension catalyzed by the Tetrahymena ribozyme *Mol. Cell Biol.* **11** 3390–4
[7] McGinness K E and Joyce G F 2003 In search of an RNA replicase ribozyme *Chem. Biol.* **10** 5–14
[8] Lawrence M S and Bartel D P 2005 New ligase-derived RNA polymerase ribozymes *RNA* **11** 1173–80
[9] Zaher H S and Unrau P J 2007 Selection of an improved RNA polymerase ribozyme with superior extension and fidelity *RNA* **13** 1017–26
[10] Nahvi A, Sudarsan N, Ebert M S, Zhou X, Brown K L and Breaker R R 2002 Genetic control by metabolite binding mRNA *Chem. Biol.* **9** 1043–9
[11] Winkler W C, Cohen-Chalamish S and Breaker R R 2002 An mRNA structure that controls gene expression by binding FMN *Proc. Natl Acad. Sci.* **99** 15908–13
[12] Winkler W C and Breaker R R 2003 Genetic control by metabolite-binding riboswitches *Chem. Biol. Chem.* **4** 1024–32
[13] Winkler W C and Breaker R R 2005 Regulation of bacterial gene expression by riboswitches *Ann. Rev. Microbiol.* **59** 487–517
[14] Breaker R R 2008 Complex riboswitches *Science* **319** 1795–7

[15] Elowitz M B and Leibler S 2000 A synthetic oscillatory network of transcriptional regulators *Nature* **403** 335–8
[16] Gardner T S, Cantor C R and Collins J J 2000 Construction of a genetic toggle switch in *Escherichia coli Nature* **403** 339–42
[17] Guet C C, Elowitz M B, Hsing W and Leibler S 2002 Combinatorial synthesis of genetic networks *Science* **296** 1466–70
[18] Hasty J, McMillen D and Collins J J 2002 Engineered gene circuits *Nature* **420** 224–30
[19] Atkinson M R, Savageau M A, Myers J T and Ninfa A J 2003 Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli Cell* **113** 597–607
[20] Isaacs F J, Dwyer D J, Ding C, Pervouchine D D, Cantor C R and Collins J J 2004 Engineered riboregulators enable post-transcriptional control of gene expression *Nature Biotechnol.* **22** 841–7
[21] Bayer T S and Smolke C D 2005 Programmable ligand-controlled riboregulators of eukaryotic gene expression *Nature Biotechnol.* **23** 337–43
[22] Isaacs F J, Dwyer D J and Collins J J 2006 RNA synthetic biology *Nature. Biotechnol.* **24** 545–54
[23] Davidson E A and Ellington A D 2007 Synthetic RNA circuits *Nature Chem. Biol.* **3** 23–8
[24] Noireaux V, Bar-Ziv R and Libchaber A 2003 Principles of cell-free genetic circuit assembly *Proc. Natl Acad. Sci.* **100** 12672–7
[25] Penchovsky R and Breaker R R 2005 Computational design and experimental validation of oligonucleotide-sensing allosteric ribozymes *Nat. Biotechnol.* **23** 1424–33
[26] Kim J, White K S and Winfree E 2006 Construction of an *in vitro* bistable circuit from synthetic transcriptional switches *Mol. Syst. Biol.* **2** 68
[27] Jiang M, Rong M, Martin C and McAllister W T 2001 Interrupting the template strand of the t7 promoter facilitates translocation of the DNA during initiation, reducing transcript slippage and the release of abortive products *J. Mol. Biol.* **310** 509–32
[28] Grundy F J and Henkin T M 2004 Regulation of gene expression by effectors that bind to RNA *Curr. Opin. Microbiol.* **7** 126–31
[29] Novick R P 1989 Staphylococcal plasmids and their replication *Ann. Rev. Microbiol.* **43** 537–65
[30] Brantl S 2002 Antisense-rNA regulation and RNA interference *Biochim. Biophys. Acta.* **1575** 15–25
[31] Putzer H, Condon C, Brechemier-Baey D, Brito R and Grunberg-Manago M 2002 Transfer RNA-mediated antitermination *in vitro Nucl. Acids Res.* **30** 3026–33
[32] Grundy F J, Winkler W C and Henkin T M 2002 tRNA-mediated transcription antitermination *in vitro*: codon–anticodon pairing independent of the ribosome *Proc. Natl Acad. Sci.* **99** 11121–6
[33] Wilson K S and vol Hippel P H 1995 Transcription termination at intrinsic terminators: the role of the RNA hairpin *Proc. Natl Acad. Sci.* **92** 8793–7
[34] Yarnell W S and Roberts J W 1999 Mechanism of intrinsic transcription termination and antitermination *Science* **284** 611–5
[35] Xayaphoummine A, Viasnoff V, Harlepp S and Isambert H 2007 Encoding folding paths of RNA switches *Nucl. Acids Res.* **35** 614–22
[36] Xayaphoummine A, Bucher T and Isambert H 2005 Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots *Nucleic Acids Res.* **33** W605–10
[37] Isambert H and Siggia E D 2000 Modeling RNA folding paths with pseudoknots: application to hepatitis delta virus ribozyme *Proc. Natl Acad. Sci.* **97** 6515–20
[38] Xayaphoummine A, Bucher T, Thalmann F and Isambert H 2003 Prediction and statistics of pseudoknots in RNA structures using exactly clustered stochastic simulations *Proc. Natl Acad. Sci.* **100** 15310–5
[39] Viasnoff V, Meller A and Isambert H 2006 DNA nanomechanical switches under folding kinetics control *Nano Lett.* **6** 101–4
[40] Harlepp S, Marchal T, Robert J, Leger J-F, Xayaphoummine A, Isambert H and Chatenay D 2003 Probing complex RNA structures by mechanical force *Eur. Phys. J.* E **12** 605–15
[41] Heidrich N and Brantl S 2007 Antisense RNA-mediated transcriptional attenuation in plasmid pIP501: the simultaneous interaction between two complementary loop pairs is required for efficient inhibition by the antisense RNA *Microbiology* **153** 420–7
[42] Nudler E and Gusarov I 2003 Analysis of the intrinsic transcription termination mechanism and its control *Methods Enzymol.* **371** 369–82
[43] Borukhov S and Nudler E 2008 RNA polymerase: the vehicle of transcription *Trends Microbiol.* **16** 126–34
[44] Chamberlin M and Ring J 1973 Characterization of T7-specific ribonucleic acid polymerase: I. General properties of the enzymatic reaction and the template specificity of the enzyme *J. Biol. Chem.* **248** 2235–44
[45] Chamberlin M and Ring J 1973 Characterization of T7-specific ribonucleic acid polymerase: II. Inhibitors of the enzyme and their application to the study of the enzymatic reaction *J. Biol. Chem.* **248** 2245–50
[46] Jeng S T, Gardner J F and Gumport R I 1990 Transcription termination by bacteriophage T7 RNA polymerase at rho-independent terminators *J. Biol. Chem.* **265** 3823–30
[47] Pan T and Sosnick T 2006 RNA folding during transcription *Ann. Rev. Biophys. Biomol. Struct.* **35** 161–75
[48] Schroeder R, Barta A and Semrad K 2004 Strategies for RNA folding and assembly *Nature Rev. Mol. Cell. Biol.* **5** 908–19
[49] Lincoln T A and Joyce G F 2009 Self-sustained replication of an RNA e *Science* 1167856