Minireview

Regulating eukaryotic gene expression with aptamers

Jean-Jacques Toulmé*, Carmelo Di Primo, David Boucard

European Institute of Chemistry and Biology, INSERM U386, Université Victor Segalen, Bordeaux, France

Received 19 March 2004; accepted 24 March 2004

Available online 22 April 2004

Edited by Horst Feldmann

Abstract Aptamers are RNA or DNA oligonucleotides identified within a randomly synthesized library, through an in vitro selection procedure. The selected candidates display a predetermined property of interest with respect to a given target. Successful selection has been carried out against targets ranging from small (amino acids, antibiotics) to macro-molecules (proteins, nucleic acids). They generally show an affinity in the nanomolar range and a high specificity of target recognition. Interestingly, aptamers selected against purified targets in the test tube retain their properties within cells. RNA aptamers can be generated in situ from an appropriate DNA construct or delivered as nuclease-resistant oligonucleotide analogues. For example, aptamers recognizing RNA structure through loop–loop interactions modulate the *trans*-activation of in vitro transcription mediated by the TAR RNA element of human immunodeficiency virus type 1. Consequently, they constitute both exquisite tools for functional genomics analysis and promising prototypes of therapeutic agents. Natural aptameric motifs have been identified within mRNA sequences, which upon binding to a metabolite control the expression of the encoded gene, which is generally involved in the biosynthesis of this particular metabolite.

Keywords: In vitro selection; RNA structure; Viral RNA; Surface plasmon resonance

1. Introduction

Over the last 25 years a number of strategies have been developed that make use of nucleic acids for artificially regulating gene expression. The regulatory oligomers are designed for generating specific complexes with a pre-determined target. Consequently, the effect is restricted to a single gene making such oligonucleotides selective gene inhibitors. To this end either nucleic acids or proteins can be targeted (Fig. 1). In the old antisense approach, a sequence complementary to part of an RNA (generally messenger or pre-messenger RNA) prevents the reading of the encoded information [1]. Numerous successful results have been reported both in vitro and in vivo with antisense oligomers as well as with ribozymes which cleave the target RNA at the hybridization site. Very recently RNA interference opened new avenues for the inhibition of translation in many organisms, through double-stranded RNA generating so-called siRNA [2,3]. Triple-stranded structures can be generated from a duplex DNA composed of one purine strand and, consequently, one pyrimidine strand. This antigene strategy has been demonstrated to efficiently prevent transcription of various genes [4,5]. The binding of a protein (let us say a transcription factor) to its target site can be competed out upon supply of a decoy, an oligomer whose sequence is identical to the protein target site; this consequently prevents the biological effect of the factor on the downstream gene [6]. The aforementioned strategies are rational ones: the regulatory oligomer is designed according to established rules, e.g., Watson–Crick base pairing for antisense and siRNA. A combinatorial approach termed Systematic Evolution of Ligands by EXponential enrichment (SELEX) described in 1990 offers an alternative for the design of oligonucleotide ligands of high affinity, termed aptamers [7,8].

This strategy can be used against a wide variety of target molecules leading to highly selective ligands of interest for both therapeutic and diagnostic purposes. Iterative cycles of selection and amplification allowed the identification of RNA and DNA aptamers against proteins or nucleic acids. As far as these targets mediate the control of the expression of a gene, the cognate aptamers constitute specific artificial modulators of biological processes, mimicking the behaviour of decoy or antisense sequences. In contrast to rational strategies there is no pre-requisite to the design of aptamers. Interestingly, gene regulation can also be achieved through the association between RNA aptamers and small molecules, as exemplified by riboswitches. In this review, examples will be given of the specific control of gene expression by aptamers targeted to different kind of targets. We will focus a large part of it on the targetting of RNA structures.

2. Kissing aptamers

Numerous RNA motifs play a key role in the transmission of the genetic information. For instance, the *trans*-activating responsive (TAR) element is a 59 nucleotide long imperfect RNA hairpin located at the very 5′-end of the human

* Corresponding author. Present address: IECB – INSERM U386, 2 rue Robert Escarpit, 33607 Pessac Cedex, France. Fax: +33-5-57-57-10-15.
E-mail address: jean-jacques.toulme@bordeaux.inserm.fr (J.-J. Toulmé).

Abbreviations: HIV-1, human immunodeficiency virus type 1; Tat, *trans*-activator protein; DIS, dimerisation initiation site; SPR, surface plasmon resonance; LNA, locked nucleic acid

0014-5793/$22.00
doi:10.1016/j.febslet.2004.03.111

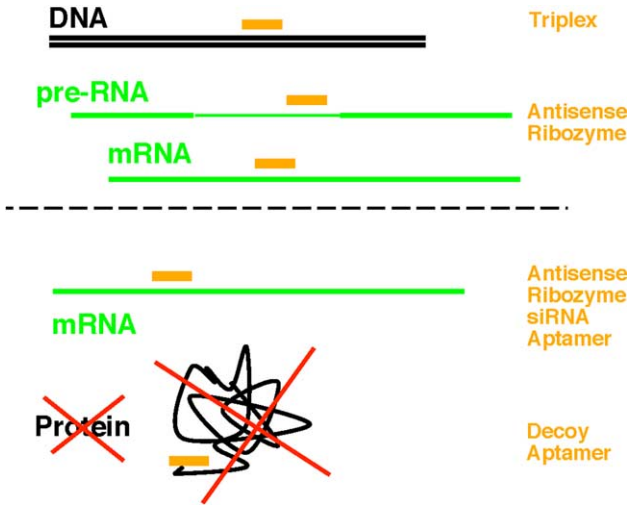

Fig. 1. Strategies (listed to the right) for artificially regulating gene expression by means of synthetic oligonucleotides. The specific binding of an oligomer to DNA, pre-mRNA or mRNA in the nucleus (top), to mRNA or to a protein (bottom) in the cytoplasm ultimately results in the selective inhibition of protein function.

immunodeficiency virus (HIV) mRNA. The viral protein *trans*-activator protein (Tat) binds specifically to a bulge in the upper part of the TAR element and recruits the cellular cyclin T1 and the associated kinase CDK9, at the apical loop of the hairpin. This multi-protein–RNA complex hyperphosphorylates the carboxy terminal domain of the RNA polymerase II, making this enzyme more processive [9,10]. Therefore, specific ligands of the TAR RNA that would inhibit the binding of Tat are expected to severely hinder the transcription of the HIV genome and the development of the retrovirus. The TAR RNA structure was used as a target for the in vitro selection of aptamers.

2.1. R06 aptamer

In vitro selection against TAR identified RNA aptamers recognizing the folded structure [11]. After ten selection rounds the sequenced candidates revealed a consensus octamer, 5′-GUCCCAGA-3′, the six central bases of which were complementary to the TAR RNA loop. Band shift assays were performed between the radiolabelled target and these candidates leading to a dissociation equilibrium constant of 20–50 nM under the selection conditions. Computer analysis used to predict the secondary structure of these aptamers revealed that they folded as imperfect hairpins. The 8-mer consensus sequence presented in the apical loop includes six bases complementary to the TAR loop. A minimal binding motif could be defined for the aptamer of highest affinity, R06, corresponding to the top part of the 98-nt parent candidate. The truncated aptamer, which retains similar TAR binding properties, is a perfect hairpin with a 8-bp stem and the octameric sequence displayed in its apical loop (Fig. 2A). Enzymatic footprints gave strong evidence for loop–loop interaction or kissing-complex formation between TAR and R06. Further analysis of the target-aptamer complex by band shift assay demonstrated that Watson–Crick base pairing between the loops was crucial for the stability of the bimolecular complex: a point mutation in the R06 loop was detrimental for binding to TAR. The affinity for the target could be restored by introducing a compensatory mutation in the TAR loop. Therefore, R06 can discriminate between two hairpins differing by a single base in the hexanucleotide loop.

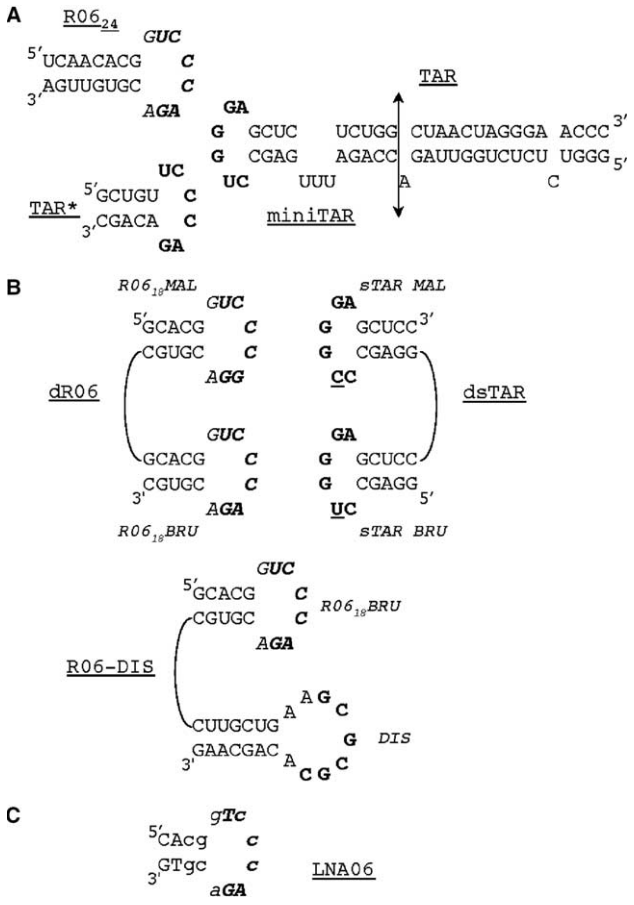

Fig. 2. Sequence and secondary structure of loop–loop interacting hairpins. (A) R06 is the RNA aptamer identified by in vitro selection against TAR (the selected consensus sequence is in italic). TAR* is a rationally designed TAR ligand. Bases susceptible to base-pair are in bold. MiniTAR, a truncated version of TAR, has been used for kissing complexes studies. (B) dsTAR is the double target model (see text), it consists in the TAR BRU hairpin linked to TAR MAL by tri-ethylene glycol phosphate units (thin line). BRU and MAL refer to two HIV-1 strains differing by one base in the apical loop (the mutated base is underlined). The double ligand, dR06, consists in two hairpin aptamers, R06MAL and R06BRU, complementary to their respective target. These hairpins were linked by the same linker as the double target. R06-DIS is the double ligand designed to recognize simultaneously the TAR and DIS hairpins within the 5′-LTR of HIV-1 (see text). (C) Mixmer LNA/DNA hairpin derivative of the R06 aptamer recognizing TAR BRU. Capital and lower case letters represent LNA and DNA bases, respectively.

However, complementarity between the loops did not fully account for the binding capability of R06, suggesting that non-canonical interactions between TAR and the aptamer were crucial. Compared to the aptamer, the selected octameric sequence was a very poor ligand of TAR with an association constant 2–3 orders of magnitude lower than the full-length aptamer. Thermal denaturation experiments monitored by UV spectroscopy revealed that the complex formed between TAR and the 8-mer consensus sequence was characterized by a melting temperature (T_m) lowered by 27 °C compared to that of the full-length aptamer–TAR complex [12]. Binding to immobilized TAR, monitored by surface plasmon resonance

(SPR), could not even be observed when the 8-mer sequence was injected over the sensorchip surface, while at a similar concentration the hairpin aptamer gave rise to clear association and dissociation phases [12]. Therefore, even if the loop–loop interaction was driven primarily by Watson–Crick base pairing, stability of the complex required the interacting region to be presented in a structured context.

The structure of the kissing complex formed by TAR and TAR* (Fig. 2A), a rationally designed hairpin with a six-nucleotide loop complementary to the TAR loop and a 5-bp stem, was shown by NMR spectroscopy to be bent towards the major groove of the loop–loop helix with an angle of 30° [13]. A quasi-continuous stacking of base pairs from one stem helix to the other one, through the loop–loop helix was observed. These structural features were also observed in other kissing complexes derived from NMR experiments [14–16]. In contrast an almost straight coaxial stacking of the helices was observed in the crystal structure of the 23S RNA kissing complex [17] and in the DIS–DIS one [18], a kissing complex that constitutes the initial step of the dimerization process of the human immunodeficiency virus type 1 (HIV-1) genome. Interestingly, molecular dynamics (MD) performed on the TAR–TAR* solution structure reconciled these observations [19]. The initial curvature of the structure disappeared during the first nanosecond of the simulation run and TAR–TAR* adopted an almost straight coaxial structure, indicating that this latter conformation was likely the most stable one. The three-dimensional structure of TAR–R06 is not known yet but one can reasonably expect that this complex adopts an overall conformation similar to that observed for other RNA–RNA loop–loop complexes. In particular, TAR–R06 as other kissing complexes was recognized by the structure specific Rop protein encoded by the ColE1 plasmid from *Escherichia coli* [20].

The 8 nt loop of the anti-TAR R06 aptamer corresponding to the selected consensus shows selected G and A residues flanking the six nucleotide sequence complementary to the TAR loop [11], suggesting that they might play a role for the stability of the TAR–aptamer complex. This was demonstrated by investigating the properties of different combinations of loop closing residues [11,12]. Purine,purine combinations gave rise to the less destabilized TAR–aptamer complexes, the A,G combination, for instance, being almost equivalent to the selected G,A one. In contrast, pyrimidine, pyrimidine combinations drastically decreased the stability of the complex. It is worth noting that complex stability was not directly related to the aptamer stem stability. The G,C combination, for instance, that provided the aptamer stem with an additional Watson–Crick base pair, generated an aptamer of increased thermal stability compared to the selected aptamer ($\Delta T_m = +11$ °C) but the resulting bimolecular complex with the viral target was less stable ($\Delta T_m = -16$ °C). The U,A combination that gave rise to a TAR*-like aptamer resulted in a −17 °C decreased thermal stability of the complex that exactly matched the stability obtained with TAR*, suggesting that the increased stability of the TAR–R06 complex over TAR–TAR* really originated in the G and A residues closing the aptamer loop. This result validated the usefulness of an in vitro combinatorial approach over a rational one to identify high affinity RNA ligands.

Kinetic analysis by SPR further emphasized the stabilizing role of the G and A residues [12]; the decreased stability of the complexes formed with the mutated aptamers originated in a faster dissociation reaction, while the association reaction remained unchanged. In the selection buffer which contained 3 mM Mg^{2+}, binding of TAR* to TAR was hardly detected. Structural studies have suggested that the stabilizing effect of this cation on loop–loop complexes would result from direct binding in two specific metal ion binding sites made by phosphate clusters flanking the major groove of the loop–loop helix [13,14,21].

MD was used recently to investigate the role of the residues closing the aptamer loop [19]. As the three-dimensional structure of TAR–R06 is not established yet, the TAR–TAR* solution structure was taken as the starting structure. The TAR* hairpin was converted into an R06 like aptamer by substituting the U closing the loop on the 5′-side by a G, thus generating the characteristic loop closing G,A combination. The choice of the initial conformation of these residues was obviously crucial. The MD results on TAR–TAR* showed that the UA base pair next to the TAR* loop adopted a conformation characterized by a large C1′–C1′ interglycosidic distance, which likely favoured stacking of the bases at the stem–loop junctions. Taking also into account the preliminary NMR experiments on TAR–R06, which suggested that the GA base pair did not adopt a sheared conformation, the G,A bases were positioned in a *cis* Watson–Crick/Watson–Crick conformation. The results obtained during the MD simulation runs suggested that the stabilizing role of G,A likely resulted from inter-backbone hydrogen bonds that were not observed in the TAR–TAR* structure and an optimized stacking of the bases at the stem–loop junctions.

Stacking interactions as key structural determinants for stable kissing complexes were also demonstrated in the case of a loop–loop interaction derived from a transient RNA–RNA kissing complex that regulates the replication of the *E. coli* plasmid ColE1, by introducing 2-aminopurine fluorescent probe, an analogue that forms with uracil a Watson–Crick base pair isosteric with AU at the stem–loop junctions [22]. Kinetic analysis of the interaction by fluorescence-detected stopped-flow experiments showed that loop–loop association followed a two-step mechanism: an initial encounter reaction was followed by a slower kinetic step that might reflect an isomerization reaction for optimizing the stacking interactions at the stem–loop junctions.

Other in vitro selection of RNA candidates against RNA hairpins further supported the idea that purine-purine base pairs might be preferred for closing hairpin loops involved in RNA loop–loop interactions. Scarabino et al. identified hairpin aptamers that bound to the anticodon loop of the yeast $tRNA^{phe}$ through kissing complex formation [23]. As reported for the R06 anti-TAR aptamer, the 7 nt loop of the aptamer complementary to the anticodon loop was flanked by G and A residues. In vitro selection against the DIS hairpin that regulates dimerization of the HIV-1 RNA genome also selected kissing aptamers showing mostly purine-purine combinations to close the loop [24]. However, the A,A combination was preferred over all others, including GA, confirming that the identified AA sheared pair that flanked the auto-complementary hexanucleotide sequence was crucial for dimerization [25].

In vitro selection of DNA candidates against TAR RNA led also to aptamers that formed loop–loop complexes with the viral target. However, footprinting studies [26] and NMR experiments [27] showed that the conformation of the RNA–DNA interacting loops differed from that of the TAR–RNA

aptamer complex. The loop–loop interaction involved only five base pairs instead of six for TAR–R06. In contrast to this latter complex that has no linking residue to connect the loop–loop helix to the TAR or aptamer stems, one RNA and two DNA residues constituted linkers in the TAR–DNA aptamer complex. Difference in geometry between RNA–RNA and DNA–RNA loop–loop complexes was further supported by the fact that the TAR–DNA aptamer was not recognized by the Rop protein. The RNA–DNA loop–loop helix was even different from that of linear RNA–DNA duplexes as it was not recognized by the *E. coli* RNase H protein. These results indicate that depending on the chemistry of the random pool, the selection process evolved not only to optimize the Watson–Crick interactions that primarily drive loop–loop interactions but also to favour non-canonical interactions that significantly contribute to the thermodynamic stability of kissing complexes.

2.2. Chemically modified aptamers

In vitro selected DNA or RNA aptamers are ligands that display high affinity and strong selectivity for their target and might be of interest for biological applications such as the regulation of gene expression. When used in a cellular context, efficiency of the selected candidates decreased inevitably as their life-time is drastically reduced by nuclease degradation. Several modifications were developed in the frame of the antisense strategy to circumvert this limitation and even to improve the affinity for the targeted sequences [28]. Two approaches can be used to generate nuclease-resistant aptamers. The first one takes advantage of some unnatural nucleotides that can be enzymatically incorporated by the polymerases during the selection process. Phosphorothioate linkage in place of phosphodiester is compatible with the SELEX enzymes [29]. 2′-Fluoro and 2′-amino pyrimidines are chemically modified nucleotides that were successfully used to generate nuclease resistant aptamers [30,31].

The second approach which consists in introducing post-selection, chemical modifications in the selected aptamer is risky. The chemistry of the random pool dictates the structure of the complex, reflecting the exquisite adaptation of the aptamer to its target. A DNA version of the R06 RNA aptamer is a poor TAR ligand and vice versa. Then, any modification that will alter the geometry of the nucleotides could affect affinity and specificity of the selected candidates for the target. In this context, modifications that will retain the conformation of the parent aptamers will generate a priori good mimics. 2′-*O*-methyl (2′-OMe), N3′ → P5′ phosphoramidate deoxynucleotide (NP-DNA) and "locked nucleic acid" (LNA) (Fig. 3) modifications confer resistance to nucleases and adopt the N-type (C3′-endo) conformation characteristic of the RNA [32].

Indeed, fully modified 2′-OMe and NP-DNA R06 derivatives formed complexes with TAR, characterized by similar or even slightly higher affinity constants than the parent RNA–RNA complex [33,34]. Moreover, these derivatives retained the key structural determinants identified by in vitro selection. In particular, the crucial G,A loop-closing combination of the aptamer still contributed to the thermodynamic stability of the complex indicating that they adopted an overall conformation close to that of the selected TAR–RNA complex. However, these chemical modifications introduced subtle changes in the geometry of the resulting kissing complex and the loop–loop helices differed from that observed with linear hybrids. In

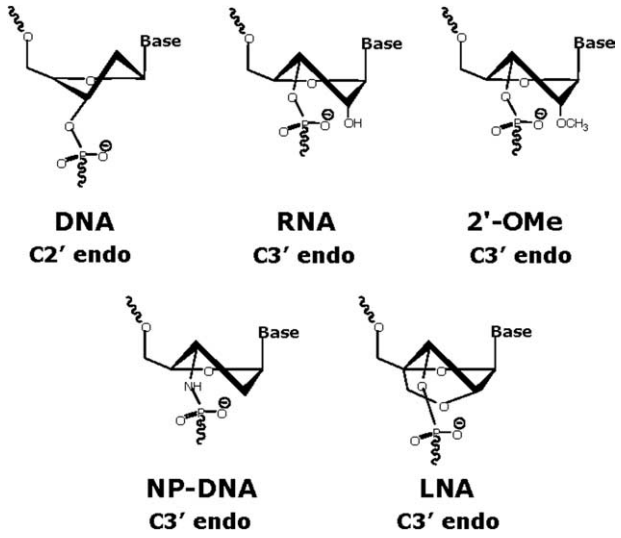

Fig. 3. Schematic representation of the sugar conformation adopted by DNA, RNA, 2′-OMe, NP-DNA and LNA monomers.

contrast to the RNA–RNA kissing complex, neither 2′-OMe- nor NP-DNA–TAR complexes were recognized by the Rop protein. Moreover, the increased stability per modified residue of 2′-OMe-RNA ($\Delta T_m = +0.5$ °C) [35] or NP-DNA–RNA ($\Delta T_m = +2.5$ °C) [36] versus RNA–RNA linear duplexes was not observed. This further underlined the non-canonical conformation of the loop–loop region.

The results obtained with LNA analogues illustrate that post-selection modification of aptamers is not trivial and that C3′ endo-conformation of the incorporated modified nucleotides will not always guarantee success. LNA is a recently introduced chemical modification that generates the most stable hybrids ever characterized with a ΔT_m of +3 °C and +10 °C per LNA residue upon binding to DNA and RNA, respectively [37]. In contrast to fully 2′-OMe and NP-DNA versions of R06, the LNA aptamer did not form a stable complex with TAR. Neither were good ligands hairpins that displayed an LNA stem with a DNA loop or vice versa (Darfeuille et al., unpublished).

Having in mind that flexibility of the ribose ring was restricted due to the 2′-*O*,4′-C-methylene linkage and that oligomers which alternated DNA and LNA nucleotides adopted an overall A-type conformation [37–39], a series of mixmer LNA/DNA hairpins were synthetized. No rules really dictated the positions at which DNA or LNA residues were introduced. One derivative, LNA06, (Fig. 2C) with the G and A residues closing the loop being DNA, led to a complex as stable as the one obtained with the parent RNA aptamer. As previously observed with 2′-OMe and NP-DNA modified aptamers, the LNA modification did not generate a complex of increased stability. Further analysis of the interaction between the mixmer DNA/LNA derivative and the viral target showed that stacking interactions at the stem–loop junctions were still crucial and related to the identity (G and A) and the chemistry (DNA) of the nucleotides closing the loop of the hairpin analogue. Surprisingly, the LNA/DNA antisense octamer (corresponding to the hairpin loop) displayed the same affinity for the target as the aptamer, even if it behaved kinetically differently, whereas its RNA version hardly bound to TAR RNA. However, kissing over antisense interactions provided

recognition with increased specificity, validating the usefulness of in vitro selection (Darfeuille et al., unpublished).

The biological effect of 2′-OMe and NP-DNA anti-TAR aptamer derivatives was evaluated. Both analogues were inhibitors of the Tat-mediated in vitro transcription with an IC_{50} of about 400 nM compared to >4 μM for the in vitro selected RNA aptamer [33,34]. This effect was specific as the R06 analogue with a loop-closing C,U combination, which did not bind to TAR, did not inhibit Tat-mediated transcription. As demonstrated with the NP-DNA hairpin and a Tat peptide, the inhibitory effect was likely in part related to a competition between the viral protein and the chemically modified aptamer for binding to TAR. Interestingly, as the binding sites of these ligands do not overlap this suggested that interaction of the aptamer analogue induced structural changes that prevent Tat binding or inhibited the conformation changes of the TAR RNA taking place upon Tat binding.

3. From mono- to bi-functional aptamers

Aptamers can be used to finely tune gene regulation by interacting with regulatory RNAs through loop–loop interactions (see above). Examples of natural RNA regulators working through similar types of complexes have been described in prokaryotes [40]. Most of these natural regulators consist in several structural motifs and engage several interactions simultaneously. One could take advantage of such an organization to rationally design RNA aptamers that could interact with their target at multiple binding sites. Compared to strategies that target one individual structural motif, this approach is expected to increase both the stability and the specificity of the resulting complexes.

3.1. Double loop–loop complexes

In order to validate multivalent structured RNA aptamers as ligands of increased affinity and specificity for highly structured RNA targets, a model of interaction consisting in a double kissing complex was first developed on the basis of the previously described TAR–R06 complex. Two variants of the TAR RNA element – namely BRU and MAL – are known which essentially differ by a single base in the loop. The double target consisted in truncated versions of BRU (TAR BRU) and MAL (TAR MAL) TAR RNA hairpins, connected to each other by three tri-ethylene glycol phosphate (C9PEG) units to generate dsTAR (Fig. 2B). The bivalent ligand was derived from a shortened version of the RNA aptamer R06 (R06BRU) that had been raised against TAR BRU (Fig. 2B). An aptamer selective for the TAR MAL hairpin (R06MAL) was merely generated from R06BRU by introducing an $A \rightarrow G$ compensatory mutation in the aptamer loop. This aptamer, R06MAL, was shown to bind its target through a loop–loop interaction, as well [33,34]. The stems of these two hairpins have been shortened down to five base pairs and linked to each other by the same C9PEG linker as the double target thus giving rise to dR06. Moreover, mutants of the double ligand have been produced by substituting the three Cs in the loop by three As given that this mutation drastically decreases the stability of the BRU and MAL complexes.

Under 1 mM Mg^{2+}, the R06BRU/dsTAR and the R06MAL/dsTAR complexes displayed K_d values equal to 6.2 and 1.0 nM, respectively, as measured by band-shift assays. The dissociation equilibrium constant for the dR06/dsTAR complex could not be determined since at the first concentration used, 0.4 nM, 90% of the radiolabelled ligand was shifted. The enthalpy of formation, ΔH, of each complex was deduced from the variation of the melting temperature monitored by UV absorption spectroscopy in function of the total RNA concentration. Under 0.1 mM Mg^{2+}, dR06, R06BRU and R06MAL, respectively, displayed ΔH values of −407.4, −63.3 and −66.1 kcal mol^{-1} (Boucard et al., unpublished). These results demonstrated that multiple structured interactions can generate complexes of increased stability compared to the parent individual complexes. In addition, the link between the two hairpins gave rise to a cooperative interaction since the enthalpy of the double complex was three times the sum of the two individual ones. This could be explained by a structural rearrangement once the two sites of the double aptamer are bound to their respective target. This hypothesis was supported by the fact that, on a native polyacrylamide gel, the dR06/dsTAR complex migrated much faster than the complex formed by the two single ligands mixed together and dsTAR. Hence, the increased stability of the double kissing complex over the single ones would stem not only from an increased number of base pairs but also from a peculiar conformation that would be thermodynamically favourable.

SPR experiments carried out under 1 mM Mg^{2+} have shown no significant differences between the association rates (Boucard et al., unpublished). The on-rates observed for mono- and bi-functional complexes were actually in the same range of values. In contrast, the dissociation rate of the double kissing complex ($k_{off} = 3.810^{-5}$ s^{-1}) was about one order of magnitude lower than the ones observed for the monovalent complexes. The study of this model has thus shed light upon the advantages of using multifaceted structural interactions to target highly organized RNA elements.

3.2. Simultaneous targetting of RNA hairpins in the HIV-1 genome

This bi-modal interaction strategy has then been applied to the design of bivalent RNA ligands capable of recognizing the 5′-untranslated region (5′-UTR) of the HIV-1 genome by multiple interactions. This region is actually highly structured and is composed of several hairpins that constitute regulatory elements for different steps of the viral cycle, making these structural motifs valuable targets for our strategy [41]. The Dimerization Initiation Site (DIS) has been shown to trigger the retroviral RNA dimerization through the formation of a kissing complex between two DIS elements which contain a palindromic sequence in the hairpin loop [42,43]. We generated a double hairpin aptamer by linking a DIS hairpin to R06BRU through a hexa-ethylene glycol phosphate unit. Band shift assays have been performed to determine the apparent dissociation constants, K_d, of the radiolabelled single or double ligands upon the 5′-UTR. The apparent binding constant for DIS/5′-UTR is higher than 200 nM as no complex was detected with the DIS hairpin alone in the concentration range used for these experiments. The K_d values obtained for the R06/5′-UTR complex suggested that this aptamer only interacted with the TAR hairpin with a K_d of 3.6 nM, in fair agreement with previous results. The double aptamer, R06-DIS, could also dimerize through its DIS element but this interaction did not prevent the binding of this ligand to the 5′-UTR. The bifunctional aptamer ligand displayed a

remarkable increased affinity for the 5′-UTR compared to R06 alone ($K_d < 0.4$ nM). This would suggest that both sites of the bivalent ligand simultaneously interacted with the 5′-UTR thus leading to a very stable complex (Boucard et al., unpublished).

The inhibitory effect of these ligands was tested in an in vitro reverse transcription assay. No significant effect was observed with either the DIS hairpin or with the R06 aptamer. In contrast, R06-DIS displayed a specific inhibitory effect with an IC_{50} of about 100 nM. These results clearly demonstrated the correlation between the affinity and the inhibitory effect of the aptamers. Several aptamers directed against the reverse transcriptase had been previously demonstrated to be good inhibitors of the HIV-1 reverse transcription in vitro or in cultured cells [44–46]. The "bifunctional aptamer" strategy therefore extends the repertoire of potential targets for the design of antiviral molecules.

This work proved that one can design aptamers that could interact with their targets by multiple binding sites by combining previously identified individual aptamers. However, one should take care of the correct folding of such aptamers in order that each domain within the complex behaves as the isolated element. This could be achieved by re-inforcing the stem regions in order to prevent intramolecules rearrangement as previously achieved, for instance, for aptamers targeted to the drosophila B52 protein [47].

3.3. Aptamers targeting proteins

Proteins are the most frequently used targets for the development of aptamers. RNA and DNA ligands have been identified for nuclear, cytoplasmic and membrane proteins [48,49]. As specific binders, aptamers are potential inhibitors of protein function. This might result either from an interaction at the catalytic site of an enzyme or from a competition with the natural ligand, let us say a DNA or an RNA sequence.

A number of examples are available in the field of viruses (HIV, HTLV, HCV). An RNA aptamer 37 nt long was identified against the protein Tat of HIV-1 [50]. The binding constant of this aptamer for Tat was two orders of magnitude higher than that of the natural TAR element making it a very efficient competitor for the authentic RNA site. Interestingly, this aptamer, composed of two inverted repeats of a TAR-like motif, is a selective inhibitor of the Tat-dependent *trans*-activation of transcription [51]. As it does not contain the apical loop of the retroviral TAR element, this aptamer does not trap cellular proteins. Consequently, it shows effects restricted to TAR mediated transcription and does not interfere with cellular processes involving the host protein cyclin T1 and CDK9. Other regulatory retroviral proteins have been successfully targeted such as the Rex protein of HTLV-1 [52] or the Rev protein of HIV-1 [53] that modulate the early phases of the viral life cycle through the transport of incompletely spliced RNA to the cytoplasm.

Applying the SELEX procedure to the HIV-1 reverse transcriptase led Tuerk et al. [54] to the identification of an RNA pseudo-knot which displayed both an extremely high binding constant ($K_d \approx 25$ pM) and a strong inhibitory effect. The interest of this aptamer was later on evaluated in cultured cells. The expression of this aptamer, within a chimeric gene, was driven by a Pol III promoter in order to ensure an appropriate cytoplasmic localization. In stably transfected T-lymphoid cells, the pseudo-knot construct specifically reduced the replication capability of the virus [46]. In another study it was demonstrated that this pseudo-knot aptamer was also a potent reverse transcriptase inhibitor in bacterial cells [44]. Therefore, aptamers selected in vitro against a purified protein retain their binding efficacy on endogenous target and may modulate their biological function. Furthermore, no toxic effect of intracellular aptamers was reported. The main limitations is their delivery. This is less problematic when the target is localized outside the cell. 2′-Fluoro-pyrimidine containing aptamers were isolated, against the HIV-1 surface protein gp120 which mediates the interaction with host-cell receptors. These high affinity ligands ($K_d = 5$–100 nM) neutralize HIV-1 infectivity in human peripheral blood mononuclear cells [55]. Moreover, these aptamers were also demonstrated to neutralize clinical isolates of HIV-1.

The use of aptamers is not restricted to the inhibition of viral proteins. Specific control of gene expression has also been achieved with aptamers targeted to transcription factors. RNA aptamers raised against NF-κB recognized its target in yeast. Interestingly, the in vitro selected aptamers were further optimized through yeast genetic selection, the combination of in vitro and in vivo processes providing ligands with decoy properties [56].

3.4. Regulation of gene expression through riboswitches

Aptamers have been converted into switches for turning on and off the expression of a downstream gene. This is illustrated by the work of Werstuck and Green who selected aptamers to small molecules (antibiotics or Hoechst dyes) and inserted these aptamers into the 5′-UTR of a reporter gene [57]. The binding of the drug stabilizes the aptamer structure and prevents the scanning of the 5′ leader region or the binding of the ribosome to the mRNA. Consequently, the translation of the reporter gene was repressed both in an in vitro assay and in cultured CHO cells upon the administration of the drug. The effect was shown to be specific and induced exclusively by the relevant drug: a related molecule did not show any inhibitory effect. Other examples of similar conditional expression have been described in yeast: the insertion of tetracycline-aptamer near the start codon reduces the expression of the downstream GFP gene [58].

Interestingly, in the last few years several examples of regulatory processes were described that make use of "natural aptamers". These switches are mediated by an RNA structure which constitutes a highly selective binding site for a metabolite. Upon association between the metabolite and the mRNA, the structure element undergoes a conformational change that leads to the alteration in the expression of this gene. As the regulated gene is directly related to the production of the metabolite, these sensors offer a direct link between the biochemical surrounding and the genetic information.

The regulation of thiamine genes, a co-factor of key enzymes in carbohydrate metabolism, is negatively controlled by thiamine and thiamine pyrophosphate. An evolutionary conserved element – the thi box – has been identified in the 5′-untranslated region of those *thi* genes. It was demonstrated that 5′ leaders of *E. coli thiC* and *thiM* mRNAs bind directly to thiamine pyrophosphate. The association of the metabolite induces a conformational change correlated to the control of the operon function. This response is highly specific as several thiamine analogues do not trigger the switch [59]. Other RNA elements have been identified which respond to other metab-

olites such as adenosylcobalamin or flavin mononucleotide [60]. The FMN sensor binds riboflavin about 3 orders of magnitude less tightly than the mononucleotide which differs by only one phosphate group [59]. Therefore, the RNA aptamer motif managed for producing stabilizing interactions between anionic species, likely through the involvement of metal ions.

Remarkably, these metabolite-binding elements are found in the genome of numerous prokaryotic species but also in eukaryotes. The thiamine pyrophosphate (TPP) domain was identified in fungi and in plants, and was shown to bind TPP, suggesting that riboswitches might also control gene expression in higher organisms [61]. However, the mechanism by which the sensor mediates the regulation can be different. In *E. coli* the TPP-induced alteration sequesters the Shine and Dalgarno sequence in the complex, thus preventing translation. In *Bacillus subtilis*, the TPP-dependent structural change generates a terminator hairpin thus inducing transcription termination. In *Arabidopsis thaliana* the TPP binding domain is located in the 3′-UTR, whereas in *Neurospora crassa* it lies in an intron suggesting that in these organisms gene regulation might occur at the processing stability and splicing levels, respectively [61]. Therefore, the signal (the riboswitch) is not part of the function (the gene regulation) a situation reminiscent to what was previously described for the iron responsive element (IRE). This RNA imperfect hairpin controls the expression of genes involved in the metabolism of iron through the specific association with IRE-binding proteins. Located either in the 5′- or in the 3′-untranslated region, the IRE element acts at the translational or at the RNA stability level of the ferritin and transferrin messages, respectively [62].

4. Conclusion

Aptamers have been shown to be ligands of both strong affinity and high specificity. K_{ds} in the low nanomolar range are frequently observed. Examples are available of RNA aptamers that discriminate between molecules differing by a single methyl or a phosphate group. This review summarizes examples showing the potential of aptamers as gene regulators. Both riboswitches and artificial aptamers are functional in vivo demonstrating the biological interest of RNA scaffolds as ''sensor'' elements. The potential interest of aptamers has also been demonstrated in the field of diagnostics. Phase III clinical trial is underway for an aptamer targeted to VEGF in the case of wet macular degeneration.

Acknowledgements: D. Boucard is the recipient of a FRM/Sidaction fellowship. This work was supported in part by the Agence Nationale de Recherches sur le Sida and by the Fondation pour la Recherche Médicale. We are grateful to Dr. Fabien Darfeuille for sharing unpublished results.

References

[1] Hélène, C. and Toulmé, J.J. (1990) Biochim. Biophys. Acta 1049, 99–125.
[2] Agrawal, N., Dasaradhi, P.V., Mohmmed, A., Malhotra, P., Bhatnagar, R.K. and Mukherjee, S.K. (2003) Microbiol. Mol. Biol. Rev. 67, 657–685.
[3] Bartel, D.P. (2004) Cell 116, 281–297.
[4] Casey, B.P. and Glazer, P.M. (2001) Prog. Nucleic Acid Res. Mol. Biol. 67, 163–192.
[5] Praseuth, D., Guieysse, A.L. and Hélène, C. (1999) Biochim. Biophys. Acta 1489, 181–206.
[6] Tomita, N., Morishita, R., Tomita, T. and Ogihara, T. (2002) Curr. Opin. Mol. Ther. 4, 166–170.
[7] Tuerk, C. and Gold, L. (1990) Science 249, 505–510.
[8] Ellington, A.D. and Szostak, J.W. (1990) Nature 346, 818–822.
[9] Karn, J. (1999) J. Mol. Biol. 293, 235–254.
[10] Rana, T.M. and Jeang, K.T. (1999) Arch. Biochem. Biophys. 365, 175–185.
[11] Ducongé, F. and Toulmé, J.J. (1999) RNA 5, 1605–1614.
[12] Ducongé, F., Di Primo, C. and Toulmé, J.J. (2000) J. Biol. Chem. 275, 21287–21294.
[13] Chang, K.Y. and Tinoco Jr., I. (1997) J. Mol. Biol. 269, 52–66.
[14] Lee, A.J. and Crothers, D.M. (1998) Structure 6, 993–1005.
[15] Kim, C.H. and Tinoco Jr., I. (2000) Proc Natl. Acad. Sci. USA 97, 9396–9401.
[16] Mujeeb, A., Clever, J.L., Billeci, T.M., James, T.L. and Parslow, T.G. (1998) Nat. Struct. Biol. 5, 432–436.
[17] Ban, N., Nissen, P., Hansen, J., Moore, P.B. and Steitz, T.A. (2000) Science 289, 905–920.
[18] Ennifar, E., Walter, P., Ehresmann, B., Ehresmann, C. and Dumas, P. (2001) Nat. Struct. Biol. 8, 1064–1068.
[19] Beaurain, F., Di Primo, C., Toulmé, J.J. and Laguerre, M. (2003) Nucleic Acids Res. 31, 4275–4284.
[20] Darfeuille, F., Cazenave, C., Gryaznov, S., Ducongé, F., Di Primo, C. and Toulmé, J.J. (2001) Nucleos. Nucleot. Nucleic Acids 20, 441–449.
[21] Jossinet, F., Paillart, J.C., Westhof, E., Hermann, T., Skripkin, E., Lodmell, J.S., Ehresmann, C., Ehresmann, B. and Marquet, R. (1999) RNA 5, 1222–1234.
[22] Rist, M. and Marino, J. (2001) Nucleic Acids Res. 29, 2401–2408.
[23] Scarabino, D., Crisari, A., Lorenzini, S., Williams, K. and Tocchini-Valentini, G.P. (1999) EMBO J. 18, 4571–4578.
[24] Lodmell, J.S., Ehresmann, C., Ehresmann, B. and Marquet, R. (2000) RNA 6, 1267–1276.
[25] Lodmell, J.S., Ehresmann, C., Ehresmann, B. and Marquet, R. (2001) J. Mol. Biol. 311, 475–490.
[26] Boiziau, C., Dausse, E., Yurchenko, L. and Toulmé, J.J. (1999) J. Biol. Chem. 274, 12730–12737.
[27] Collin, D., van Heijenoort, C., Boiziau, C., Toulmé, J.J. and Guittet, E. (2000) Nucleic Acids Res. 28, 3386–3391.
[28] Toulmé, J.J., Di Primo, C. and Moreau, S. (2001) Progr. Nucleic Acid Res. Mol. Biol. 69, 1–46.
[29] Andéola, M.L., Calmels, C., Michel, J., Toulmé, J.J. and Litvak, S. (2000) Eur. J. Biochem. 267, 5032–5040.
[30] Jellinek, D., Green, L.S., Bell, C., Lynott, C.K., Gill, N., Vargeese, C., Kirschenheuter, G., McGee, D.P., Abesinghe, P. and Pieken, W.A., et al. (1995) Biochemistry 34, 11363–11372.
[31] Lin, Y., Qiu, Q., Gill, S.C. and Jayasena, S.D. (1994) Nucleic Acids Res. 22, 5229–5234.
[32] Egli, M. and Gryaznov, S.M. (2000) Cell Mol. Life Sci. 57, 1440–1456.
[33] Darfeuille, F., Arzumanov, A., Gait, M.J., Di Primo, C. and Toulmé, J.J. (2002) Biochemistry 41, 12186–12192.
[34] Darfeuille, F., Arzumanov, A., Gryaznov, S., Gait, M.J., Di Primo, C. and Toulmé, J.J. (2002) Proc. Natl. Acad Sci. USA 99, 9709–9714.
[35] Monia, B.P., Lesnik, E.A., Gonzalez, C., Lima, W.F., McGee, D., Guinosso, C.J., Kawasaki, A.M., Cook, P.D. and Freier, S.M. (1993) J. Biol. Chem. 268, 14514–14522.
[36] Gryaznov, S.M., Lloyd, D.H., Chen, J.K., Schultz, R.G., DeDionisio, L.A., Ratmeyer, L. and Wilson, W.D. (1995) Proc. Natl. Acad. Sci. USA 92, 5798–5802.
[37] Bondensgaard, K., Petersen, M., Singh, S.K., Rajwanshi, V.K., Kumar, R., Wengel, J. and Jacobsen, J.P. (2000) Chemistry 6, 2687–2695.
[38] Petersen, M., Nielsen, C.B., Nielsen, K.E., Jensen, G.A., Bondensgaard, K., Singh, S.K., Rajwanshi, V.K., Koshkin, A.A., Dahl, B.M., Wengel, J. and Jacobsen, J.P. (2000) J. Mol. Recognit. 13, 44–53.

[39] Petersen, M., Bondensgaard, K., Wengel, J. and Jacobsen, J.P. (2002) J. Am. Chem. Soc. 124, 5974–5982.
[40] Brunel, C., Marquet, R., Romby, P. and Ehresmann, C. (2002) Biochimie 84, 925–944.
[41] Berkhout, B. (2000) Adv. Pharmacol. 48, 29–73.
[42] Muriaux, D., Fosse, P. and Paoletti, J. (1996) Biochemistry 35, 5075–5082.
[43] Paillart, J.C., Skripkin, E., Ehresmann, B., Ehresmann, C. and Marquet, R. (1996) Proc. Natl. Acad. Sci. USA 93, 5572–5577.
[44] Nickens, D.G., Patterson, J.T. and Burke, D.H. (2003) RNA 9, 1029–1033.
[45] Joshi, P. and Prasad, V.R. (2002) J. Virol. 76, 6545–6557.
[46] Chaloin, L., Lehmann, M.J., Sczakiel, G. and Restle, T. (2002) Nucleic Acids Res. 30, 4001–4008.
[47] Shi, H., Hoffman, B.E. and Lis, J.T. (1999) Proc. Natl. Acad. Sci. USA 96, 10033–10038.
[48] Toulmé, J.J. (2000) Curr. Opin. Mol. Ther. 2, 318–324.
[49] Brody, E.N. and Gold, L. (2000) Rev. Mol. Biotech. 74, 5–13.
[50] Yamamoto, R., Murakami, K., Taira, K. and Kumar, P.K.R. (1998) Gene Ther. Mol. Biol. 1, 451–465.
[51] Yamamoto, R., Katahira, M., Nishikawa, S., Baba, T., Taira, K. and Kumar, P.K.R. (2000) Genes Cells 5, 371–388.
[52] Baskerville, S., Zapp, M. and Ellington, A.D. (1999) J. Virol. 73, 4962–4971.
[53] Konopka, K., Lee, N.S., Rossi, J. and Duzgunes, N. (2000) Gene 255, 235–244.
[54] Tuerk, C., Macdougal, S. and Gold, L. (1992) Proc. Natl. Acad. Sci. USA 89, 6988–6992.
[55] Khati, M., Schuman, M., Ibrahim, J., Sattentau, Q., Gordon, S. and James, W. (2003) J. Virol. 77, 12692–12698.
[56] Cassiday, L.A. and Maher III, L.J. (2003) Proc. Natl. Acad. Sci. USA 100, 3930–3935, Epub 2003 Mar 13.
[57] Werstuck, G. and Green, M.R. (1998) Science 282, 296–298.
[58] Suess, B., Hanson, S., Berens, C., Fink, B., Schroeder, R. and Hillen, W. (2003) Nucleic Acids Res. 31, 1853–1858.
[59] Winkler, W.C., Cohen-Chalamish, S. and Breaker, R.R. (2002) Proc. Natl. Acad. Sci. USA 99, 15908–15913, Epub 2002 Nov 27.
[60] Nudler, E. and Mironov, A.S. (2004) Trends Biochem. Sci. 29, 11–17.
[61] Sudarsan, N., Barrick, J.E. and Breaker, R.R. (2003) RNA 9, 644–647.
[62] Hentze, M.W. and Kuhn, L.C. (1996) Proc. Natl. Acad Sci. USA 93, 8175–8182.