



HAL
open science

Recombinant RNA technology: the tRNA scaffold.

Luc Ponchon, Frédéric Dardel

► **To cite this version:**

Luc Ponchon, Frédéric Dardel. Recombinant RNA technology: the tRNA scaffold.. Nature Methods, 2007, 4 (7), pp.571-6. 10.1038/nmeth1058 . hal-00218342

HAL Id: hal-00218342

<https://hal.science/hal-00218342>

Submitted on 27 Jan 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Recombinant RNA technology : The tRNA scaffold

Luc Ponchon, Frédéric Dardel

Cristallographie & RMN Biologiques, Université Paris Descartes, CNRS, 4 avenue de l'Observatoire, 75006, Paris, France. Correspondance to F.D. (frederic.dardel@univ-paris5.fr)

During the past decades, RNA has emerged as a major player in most cellular processes. Understanding these processes at the molecular level requires homogeneous RNA samples for structural, biochemical and pharmacological studies. So far, this has been a bottleneck, as the only methods for producing such pure RNA were *in vitro* syntheses. Here, we describe a generic approach for expressing and purifying structured RNA in *E. coli*, using a series of tools which parallel those available for recombinant proteins. Our system is based on a camouflage strategy, the "tRNA scaffold", in which the recombinant RNA is disguised as a natural RNA and thus hijacks the host machinery, escaping cellular ribonucleases. This opens the way to large scale structural and molecular investigations of RNA function.

Introduction

Systematic studies of protein structure, function and interactions such as structural genomics and double-hybrid approaches have greatly benefited from the development of efficient recombinant expression systems. Simultaneously, new roles have been evidenced for structured RNA in translation, chromosome maintenance, viral replication, as riboswitches or ribozymes¹⁻³. As a result, these RNA are considered as promising drug targets, as for instance bacterial rRNA (antibiotics) or telomerase RNA⁴. Owing to their large structural repertoire, folded RNA are also interesting molecules for constructing self-assembling nano-objects, some of which have been used for delivering therapeutic siRNA^{5,6}. As opposed to proteins, systematic biochemical, structural and pharmacological studies of RNA have however been hampered by difficulties in obtaining homogeneous samples in large quantities. So far, RNA has mostly been produced *in vitro* using either T7 RNA polymerase transcription⁷ or chemical synthesis⁸, which are costly and cumbersome. Producing recombinant RNA *in vivo* could be an alternative, but thus far has been plagued by a number of obstacles, such as heterogeneity of the products, degradation by ribonucleases and difficulty in purifying the RNA from cell extracts. For example, RNA transcripts overexpressed *in vivo* using the T7 system⁹ are long and heterogeneous, as a consequence of terminator read-through, and have never been

purified. 5S ribosomal RNA¹⁰ and tRNA¹¹⁻¹⁴ are two exceptions which have been successfully expressed in *E. coli*. In the latter case, it works because tRNA are recognized by cellular enzymes that precisely process the primary transcript, incorporate modified nucleotides and repair their 3'-end¹⁵. This results in a single product of defined length, a key feature for structural studies. Furthermore, tRNA adopt a compact 3D structure which makes them resistant to both unfolding and nucleases.

We used this robustness and precise processing to express other structured RNA *in vivo* and designed a generic method using tRNA as a protective scaffold. We extracted the desired recombinant RNA with high yield and purified it to homogeneity using standard chromatography. We employ tools for tagging the RNA with affinity or fluorescent aptamers and present a scheme for efficiently cleaving off the tags. Overall, the set of methods described here offers *in vivo* tools for the RNA world which were up to now only available for recombinant proteins.

Results

The tRNA scaffold approach

We first tried to directly express structured RNA such as aptamers with vectors previously shown to work with tRNA. However, the resulting transcripts were heterogeneous and failed to accumulate at a significant level. Hence it is the tRNA structure, rather than the vector that allows stable expression. Based on this finding we constructed tRNA-RNA fusions, similar to GST fusions used for protein expression. In order to retain the 3D structure of the tRNA that is recognized by cellular factors, it was necessary to keep the acceptor, T Ψ C and D stems (**Fig. 1a**). This left the anticodon stem as a region for inserting the desired RNA. The resulting RNA chimeras are expected to be recognized by tRNA processing enzymes. Indeed, bacteria naturally produce such a tRNA chimera, transfer-messenger RNA (tmRNA), involved in trans-translation, a rescue mechanism for ribosomes trapped on defective messenger RNA¹⁶. tmRNA carries a tRNA-like domain¹⁷ with a large, complex structure inserted in the lower stem. Since tmRNA is known to be efficiently processed by tRNA

specific factors, it seemed reasonable to assume that an artificially engineered chimera using such a tRNA scaffold would be similarly handled by bacterial cells, in much the same way as intron-containing tRNA are processed in eukaryotes.

We therefore constructed various plasmids for expressing such recombinant RNA. The anticodon stem-loop was replaced by a sequence containing restriction sites, designed to maintain the 2D structure (**Fig. 1b**). Two tRNA scaffolds were used: human tRNA^{Lys₃} and *E. coli* tRNA^{Mct}. We selected these because NMR assignments of both tRNAs have been reported, and they have been shown to be efficiently expressed in a recombinant system^{13,14}. Also, the availability of two different scaffolds is advantageous in cases where RNA misfolding occurs because of spurious complementarity between the tRNA and the inserted sequence. **Figure 1b** shows the schematic organization of the corresponding expression vectors (DNA sequence in **Supplementary Fig. 1**). In principle, any structured RNA terminated by a stem could thus be inserted into the tRNA scaffold, including aptamers, ribozymes, viral RNA sequences.

Production and characterization of a tRNA chimera

As a test case, we focused on the epsilon sequence of human hepatitis B virus (HBV)^{18,19}. This RNA consists of two stems separated by a bulge on which reverse transcription initiates. We expressed chimeras with both tRNA^{Lys₃} (**Fig. 1c**) and tRNA^{Mct} scaffolds. The corresponding RNA (112 nucleotides) is strongly expressed by the recombinant strains, representing about 20% of total RNA and roughly 60% of all small, tRNA-size molecules. As described under Methods, the tRNA chimera can be easily visualized by performing RNA “minipreps” from small scale overnight cultures. Alternatively, this RNA can be purified on a larger scale by standard chromatography (**Fig. 1d**). Typical yields are 10-50 mg of purified RNA per liter. By growing cells in labeled medium, it is possible to enrich the recombinant RNA with stable isotopes, such as ¹⁵N and ¹³C and to confirm by NMR that the RNA is correctly folded: the ¹H-¹⁵N TROSY spectrum of the chimeric RNA (**Fig. 2**) shows a number of well resolved peaks, indicative of a defined, folded structure. We could unambiguously assign most of these peaks to either the tRNA^{Lys₃} or the HBV epsilon parts. In addition, the

spectra show that the tRNA chimera contains modified nucleotides: they show a number of characteristic shifted resonances (**Fig. 2**): ribothymidine (rT), pseudouridine (Ψ), 7-methylguanosine (m^7G), 4-thiouridine (s^4U). Hence, *in vivo*, this fused construct is recognized as a genuine tRNA by the cell. It is correctly processed to a single species of defined length and is modified at expected sites by structure-specific enzymes. Modified nucleotides are known to stabilize the tertiary fold of tRNA, in addition, some of them (*e.g.* s^4U at position 8) can provide convenient reactive groups for covalently attaching fluorescent labels or immobilizing RNA on solid support²⁰.

Applicability of the tRNA scaffold approach

We evaluated the general applicability of the method by constructing a series of chimeras and testing for their expression *in vivo*. We selected a number of diverse inserts, varying in size, topology of the secondary structure, biological origin and function (**Fig. 3**). Over twenty such constructs have been tested, corresponding to chimeras ranging from 74 to 376 nucleotides in length (**Fig. 3** and **Supplementary Fig. 2**). For inserts corresponding to structured RNA domains, all the chimera were efficiently expressed and processed to a product of defined length (**Fig. 4a**). This was independent of the origin of the RNA insert: artificial, viral, bacterial and human sequences were all efficiently produced and both branched and nested stem-loop structures were apparently well expressed by the system (yields are given in **Supplementary Table 1**). Some inserts nevertheless failed to be expressed, for instance, a construct containing the template region of the human telomerase RNA yielded no detectable product. This is possibly because it contained single-stranded regions which led to nuclease degradation *in vivo*. Longer inserts also gave smaller yields, suggesting that complex domains might occasionally misfold and get degraded. In order to minimize this problem, all tentative constructs were submitted to a computer prediction of their 2D structure.

In order to test the function of these recombinant molecules, we first analyzed a number of ligand-binding RNAs, such as the aptamer selected against sephadex²¹ (**Supplementary Fig. 2**). The corresponding RNA could be bound to sephadex G100, directly from whole cell extracts and further eluted with soluble dextran (**Fig. 4b**). We obtained similar results with

aptamers selected against streptavidin²² or malachite green²³ (**Supplementary Fig. 3**). We also tested the ability of several recombinant constructs to spontaneously self-assemble and form oligomers: the dimerization initiation site (DIS) of HIV-1 genomic RNA²⁴ induces dimerization of the corresponding chimera (**Supplementary Fig. 4**), as in the viral context. Similar results were obtained with the chimera containing the pre-head RNA of bacteriophage ϕ 29 (**Supplementary Fig. 2**), a structural RNA that self-assembles into a ring structure required for packaging the viral DNA^{25,26}. We performed all the above assays using the entire tRNA scaffold chimeras, indicating that within the tRNA context, the inserted RNA are correctly folded and functional.

Affinity tags and pull-down assay

In order to enhance the capabilities of the system and streamline purification, we constructed expression vectors allowing the expression of tripartite chimeras containing both an sephadex aptamer-based affinity tag and the RNA of interest inserted in the tRNA scaffold. We made combined constructs for duck and human HBV epsilon RNA (**Fig. 3c**), and for an *E. coli* 23S rRNA fragment (**Supplementary Fig. 2**) and used the sephadex affinity tag to purify the expressed RNA directly from cell lysates under native conditions, in the presence of proteins (**Fig. 4b**). Since the tRNA chimeras are folded correctly and functional within cells, they can be used for *in vivo* studies. By immobilizing RNA chimeras on the affinity matrix, it is possible to perform specific pull down assays, *e.g.* the specific binding of ribosomal protein L20 to its target site on the 23S rRNA²⁷ (**Fig. 4c**). We also derived similar constructs with a different topology from the streptavidin aptamer (**Supplementary Fig. 2**).

Release of the RNA insert from the tRNA scaffold

Although for some studies, the entire RNA chimera might be acceptable or even beneficial, in other cases it might be preferable to cleave the desired RNA off the scaffold. Several strategies were considered: using either a dedicated ribozyme or DNAzyme, or cleaving with RNase H using a pair of guide DNA oligonucleotides. Dual *cis*-acting ribozyme constructs have been used for *in vitro* transcribed RNA, but this approach is not compatible

with the tRNA scaffold topology²⁸. In our hands, the method that worked best is RNase H cleavage (**Fig. 4d**). It produces a defined product in a quantitative manner, furthermore, guide oligonucleotides for RNase H cleavage are targeted against the constant part of the scaffold and hence, the same pair can be used for all constructs. Previous reports of specific RNA cleavage using RNase H used modified DNA nucleotides²⁹ to improve specificity. In the present case, we used standard DNA oligonucleotides and obtained a single defined product. In order to test the yield and purity using this method, a large scale preparation was performed, using a tRNA chimera carrying the bacterial 16S ribosomal RNA A-site (**Supplementary Fig. 2**). Approximately 0.8 μ mole of doubly labeled RNA was obtained from 2 liters of ¹⁵N/¹³C labeled medium. The recombinant product was 33 nucleotides long, with cleavage sites located at the junction of the cloverleaf structure (**Fig. 5**). This could be a consequence of the GC-richness of the top of the anticodon-like stem which would resist strand displacement by the ends of the DNA oligonucleotides.

We analyzed this purified labeled RNA by gel electrophoresis (**Supplementary Fig. 5**) and by NMR. A typical ¹⁵N-¹H experiment is shown in **Figure 5**, revealing the NH imino groups involved in base pairing. The NMR spectrum shows a unique set of peaks, a direct evidence that a single product with a unique conformation is present in the sample. It also shows the presence of one A-U and ten to eleven G-C base pairs, in keeping with the predicted secondary structure. This RNase H cleavage strategy has now been successfully applied to six different constructs with the same guide DNA oligonucleotides. In all cases, quantitative cleavage was observed, suggesting that it is a broadly applicable method.

Discussion

In the present work, we demonstrate that it is possible to use this tRNA scaffold approach to express a large range of RNA chimera. The bacterial machinery recognizes and processes these constructs to single species which accumulate stably within the transformed cells. This demonstrates that, as postulated at start, the tRNA structure is robust and can accommodate and protect RNA inserts with diverse sizes and structures (**Fig. 3**). In all tested cases, the

recombinant RNA were folded and functional, even within the chimera. The tRNA scaffold approach is thus a broadly applicable strategy for producing large quantities of virtually any structured RNA fragment. The procedures are all straightforward, involving only standard and parallelizable techniques. The objectives were to make it possible to express any RNA or collection of RNA fragments in a matter of a few days. For instance, construction of the recombinant expression plasmids is extremely simple: for small RNA inserts, it suffices to ligate a pair of complementary DNA oligonucleotides in the selected vector. It should thus be possible to produce libraries of relevant RNA fragments, such as aptamers or non coding RNA, for systematic functional analyses.

Several RNA fragments produced and purified in this work are of major pharmacological interest : Domain III of hepatitis C virus internal ribosome entry site, the epsilon sequence of human and duck HBV viruses, HIV-1 genomic RNA fragments, domains CR4-CR5 and H/ACA of human telomerase RNA and antibiotic binding regions of bacterial ribosomal RNA (Constructs shown in **Supplementary Fig. 2**). Tens of milligram quantities of pure and stable RNA could thus be reproducibly obtained from liter-scale cultures. Such amounts are required to perform large scale screening of pharmacologically active compounds targeting RNA or for large scale production of “therapeutic” RNA constructs. Screening for ligands specific of these RNA targets using flow-injection NMR³⁰ is currently under way.

Generic purification procedures using affinity tags should open the way to a new range of applications: systematic structural approaches, both by NMR and by crystallography, high throughput parallel studies of RNA-protein interactions and combinatorial strategies. In addition, it should now be possible to perform novel *in vivo* approaches, such as co-expression of RNA-protein complexes. The recombinant RNA approach is also cost effective, which is important for studies requiring incorporation of expensive stable isotopes, and for applications requiring large amounts of material, such as medium and high-throughput screening of RNA-specific ligands. For the latter applications, however cleavage of the target RNA from the scaffold is likely to be required to eliminate tRNA binding-ligands, at least in secondary

screens. With the availability of fluorescent tags, and an efficient generic cleavage system, the complete spectrum of tools available for recombinant proteins is now also available for RNA.

Methods

Vector construction. We derived all vectors from pBSTNAV¹², by assembling synthetic overlapping oligonucleotides (purchased from Eurogentec). We constructed six such vectors, based on either human tRNA^{Lys3} or *E. coli* tRNA^{Met}, using either no tag, a streptavidin or a sephadex aptamer tag. We subsequently used these vectors for inserting the various DNA oligonucleotides coding for the RNA fragments using a similar strategy (**Fig. 1b**). The DNA sequence of all plasmid constructs was checked (Millegen, France). The cloning sites are *EagI* (5') and *SacII* (3') for the naked scaffold vectors and *SaI* (5') and *AatII* (3') for the aptamer containing vectors (**Supplementary Fig. 1**).

RNA expression and purification. We used freshly transformed *E. coli* cells (strain JM101Tr) in all expression experiments. This proved crucial to maintain a high expression level from the constitutive *lpp* promoter. For controlling expression levels, we performed RNA “minipreps” on 5 ml overnight cultures grown in LB medium. We pelleted and resuspended cells in 50 μ l of 10 mM Magnesium acetate, 1 mM Tris-HCl, pH 7.4 (Buffer L). We extracted the suspension with saturated phenol, ethanol precipitated the aqueous phase and analyzed the RNA contents analyzed by 8% polyacrylamide-urea gel electrophoresis. We visualized bands by UV shadowing. For large scale production (1 liter cultures), RNA accumulation was optimal after 12-18h of growth following inoculation, in the early stationary phase, and higher yields were obtained on richer media. We thus performed most cultures overnight in 2 x TY medium, starting from freshly transformed cells. As for minipreps, we pelleted and resuspended cells in 8 ml of buffer L. We added one volume of saturated phenol to the tube, which was gently agitated for one hour at room temperature, to ensure complete cell lysis. After centrifugation, we collected the aqueous phase and added 1/5 volume of 5 M NaCl to precipitate high molecular weight nucleic acids, which were then removed by centrifugation. We ethanol-precipitated the supernatant containing small RNA, allowed the pellet to dry and

then dissolved it in 20 mM sodium phosphate, pH 7.0 (Buffer A). We loaded the RNA extract on an ion-exchange column (Q-Hiload, Amersham) equilibrated in buffer A and eluted it with a 0 to 1 M NaCl gradient in the same buffer. We analyzed fractions by electrophoresis on 8% urea-acrylamide gels and pooled and ethanol-precipitated those containing the desired RNA. Yields were 10 to 50 mg of purified RNA per liter, with a purity of 80 to 90 %, depending on the construct, after this single chromatographic step. Whenever required, we could achieve further purification by gel filtration: we dissolved the RNA in 20 mM sodium phosphate, 100 mM NaCl, pH 7.0, loaded it on a Superdex-75 column (Amersham) equilibrated in the same buffer and eluted it at 0.5 ml/min. We pooled and ethanol-precipitated fractions containing the RNA. For structural studies, we dissolved the purified RNA in water, dialyzed it and lyophilized it.

NMR spectroscopy. We dissolved purified RNA samples in 300 μ l of 10% $^2\text{H}_2\text{O}$ /90% H_2O (final concentration 1 mM). We acquired spectra were at 15°C (288 K) on a 600 MHz Bruker Avance spectrometer. We recorded ^1H - ^{15}N TROSY and HMQC experiments as previously described³¹.

Affinity purification. We pelleted an overnight culture (10 ml) of the overproducing strain. We suspended cells in 400 μ L of 50 mM potassium phosphate, 5 mM MgCl_2 , 100 mM NaCl, pH 7.5 (Buffer A) and disrupted them by sonication. We removed debris by centrifugation and incubated the supernatant with either streptavidin-coated iron oxide magnetic beads (Dyna), or Sephadex G100 (Roche) suspended in buffer A. We washed the beads three times with buffer A and performed elution using 200 μ L of buffer A supplemented either with 5 mM *d*-biotin or with 200 mg/ml of soluble dextran (Mr 1500, Fluka), for the streptavidin and sephadex aptamer constructs, respectively.

Pull down assay. We extracted the tRNA chimera (tRNA/sephadex aptamer; tRNA/sephadex aptamer/23S rRNA) as described above (RNA "minipreps") and dissolved them in buffer A. We then incubated them with ribosomal protein L20C from *E. coli* (gift from

F. Allemand, IBPC, Paris) and sephadex G100 beads for 1 hour at 4 °C. We washed the beads three times with the same buffer. We eluted bound RNA and attached protein using 200 mg/ml of soluble dextran (Mr 1500, Fluka) in the same buffer, transferred them into SDS gel loading buffer and electrophoresed them on a 12 % SDS–polyacrylamide gel.

RNAse H cleavage. We purified recombinant, His-tagged *E. coli* RNAse H by Ni-NTA agarose and gel filtration chromatography. We first annealed the target RNA to the guide oligonucleotides in H₂O (40 μM each, 1:1 ratio with the target RNA) by heating at 95°C for 2 min. and cooled them down slowly to room temperature. We then added digestion buffer to a final concentration of 50 mM Tris-HCl, 100 mM NaCl, 10 mM MgCl₂, pH 7.5. We started the cleavage reaction by adding 60 pmoles of *E. coli* RNAse H per nmole of RNA, allowed it to proceed for 3h at 37°C and stopped it by ethanol precipitation. We purified preparative digests by ion-exchange chromatography (Mono-Q, Roche), under denaturing conditions (4M urea).

Acknowledgments

This research was supported by the 6th framework program of the European Union (grant FSG-V-RNA). The authors thank F. Allemand (Institut de Biologie Physico-Chimique, Paris) for the kind gift of purified *E. coli* L20 protein and Drs. C. Tisné and S. Nonin for assistance with the NMR experiments.

Authors contributions

F.D. was responsible for project planning. L.P. performed the experiments. F.D. and L.P. discussed the results, drafted the paper and revised the manuscript.

List of supplementary items

Supplementary Figure 1. Structure of the expression vector inserts.

Supplementary Figure 2. Structures of the various RNA chimera that were successfully expressed using the tRNA scaffold strategy.

Supplementary Figure 3. Malachite green aptamer function.

Supplementary Figure 4. Dimerization of the tRNA chimera containing HIV-1 genomic RNA dimerization initiation site (DIS).

Supplementary Figure 5. PAGE analysis of the cleavage product of a 16S rRNA A site chimera.

Supplementary Table. Overexpression yields of various tRNA chimera.

References

1. Breaker, R.R. Natural and engineered nucleic acids as tools to explore biology. *Nature* **432**, 838-45 (2004).
2. Holbrook, S.R. RNA structure: the long and the short of it. *Curr. Opin. Struct. Biol.* **15**, 302-8 (2005).
3. Puerta-Fernandez, E., Romero-Lopez, C., Barroso-delJesus, A. & Berzal-Herranz, A. Ribozymes: recent advances in the development of RNA tools. *FEMS Microbiol. Rev.* **27**, 75-97 (2003).
4. Hermann, T. & Westhof, E. RNA as a drug target: chemical, modelling, and evolutionary tools. *Curr. Opin. Biotechnol.* **9**, 66-73 (1998).
5. Jaeger, L. & Chworos, A. The architectonics of programmable RNA and DNA nanostructures. *Curr. Opin. Struct. Biol.* **16**, 531-43 (2006).
6. Guo, P. RNA nanotechnology: engineering, assembly and applications in detection, gene delivery and therapy. *J. Nanosci. Nanotechnol.* **5**, 1964-82 (2005).
7. Milligan, J.F., Groebe, D.R., Witherell, G.W. & Uhlenbeck, O.C. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **15**, 8783-8798 (1987).
8. Marshall, W.S. & Kaiser, R.J. Recent advances in the high-speed solid phase synthesis of RNA. *Curr. Opin. Chem. Biol.* **8**, 222-9 (2004).
9. Studier, F.W., Rosenberg, A.H., Dunn, J.J. & Dubendorff, J.W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60-89 (1990).
10. Moore, P.B. et al. Preparation of 5S RNA-related materials for nuclear magnetic resonance and crystallography studies. *Methods Enzymol.* **164**, 158-74 (1988).

11. Masson, J.-M. & Miller, J.H. Expression of synthetic tRNA genes under the control of a synthetic promoter. *Gene* **47**, 179-183 (1986).
12. Meinnel, T., Mechulam, Y. & Fayat, G. Fast purification of a functional elongator tRNA^{met} expressed from a synthetic gene in vivo. *Nucleic Acids Res.* **16**, 8095-6 (1988).
13. Tisné, C., Rigourd, M., Marquet, R., Ehresmann, C. & Dardel, F. NMR and biochemical characterization of recombinant human tRNA(Lys)³ expressed in *Escherichia coli*: identification of posttranscriptional nucleotide modifications required for efficient initiation of HIV-1 reverse transcription. *RNA* **6**, 1403-1412 (2000).
14. Wallis, N.G., Dardel, F. & Blanquet, S. Heteronuclear NMR studies of the interactions of ¹⁵N-labeled methionine-specific transfer RNAs with methionyl-tRNA transformylase. *Biochemistry* **34**, 7668-77 (1995).
15. Deutscher, M.P. Ribonucleases, tRNA nucleotidyltransferase, and the 3' processing of tRNA. *Prog. Nucleic Acid Res. Mol. Biol.* **39**, 209-240 (1990).
16. Keiler, K.C., Waller, P.R. & Sauer, R.T. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**, 990-3 (1996).
17. Komine, Y., Kitabatake, M., Yokogawa, T., Nishikawa, K. & Inokuchi, H. A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **91**, 9223-7 (1994).
18. Junker-Niepmann, M., Bartenschlager, R. & Schaller, H. A short cis-acting sequence is required for hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA. *EMBO J.* **9**, 3389-96 (1990).
19. Nassal, M. & Rieger, A. A bulged region of the hepatitis B virus RNA encapsidation signal contains the replication origin for discontinuous first-strand DNA synthesis. *J. Virol.* **70**, 2764-73 (1996).
20. Budker, V.G. et al. Photoaffinity reagents for modification of aminoacyl-tRNA synthetases. *FEBS Lett.* **49**, 159-62 (1974).
21. Srisawat, C., Goldstein, I.J. & Engelke, D.R. Sephadex-binding RNA ligands: rapid affinity purification of RNA from complex RNA mixtures. *Nucleic Acids Res.* **29**, E4 (2001).
22. Srisawat, C. & Engelke, D.R. Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins. *RNA* **7**, 632-41 (2001).
23. Grate, D. & Wilson, C. Laser-mediated, site-specific inactivation of RNA transcripts. *Proc. Natl. Acad. Sci. USA* **96**, 6131-6 (1999).

24. Paillart, J.C., Skripkin, E., Ehresmann, B., Ehresmann, C. & Marquet, R. A loop-loop "kissing" complex is the essential part of the dimer linkage of genomic HIV-1 RNA. *Proc. Natl. Acad. Sci. USA* **93**, 5572-7 (1996).
25. Guo, P., Zhang, C., Chen, C., Garver, K. & Trottier, M. Inter-RNA interaction of phage phi29 pRNA to form a hexameric complex for viral DNA transportation. *Mol. Cell* **2**, 149-55 (1998).
26. Zhang, F. et al. Function of hexameric RNA in packaging of bacteriophage phi 29 DNA in vitro. *Mol. Cell* **2**, 141-7 (1998).
27. Raibaud, S. et al. How bacterial ribosomal protein L20 assembles with 23 S ribosomal RNA and its own messenger RNA. *J. Biol. Chem.* **278**, 36522-30 (2003).
28. Price, S.R., Ito, N., Oubridge, C., Avis, J.M. & Nagai, K. Crystallization of RNA-protein complexes. I. Methods for the large-scale preparation of RNA suitable for crystallographic studies. *J. Mol. Biol.* **249**, 398-408 (1995).
29. Lapham, J. & Crothers, D.M. RNase H cleavage for processing of in vitro transcribed RNA for NMR studies and RNA ligation. *RNA* **2**, 289-96 (1996).
30. Tisné, C. & Dardel, F. Optimisation of a peptide library for screening specific RNA ligands by flow-injection NMR. *Comb. Chem. High Throughput Screen.* **5**, 523-9 (2002).
31. Tisné, C., Roques, B.P. & Dardel, F. The annealing mechanism of HIV-1 reverse transcription primer onto the viral genome. *J. Biol. Chem.* **279**, 3588-95 (2004).

Figure legends

Figure 1. The tRNA scaffold strategy. **(a)** Schematic structure of standard tRNA, showing some of the conserved tertiary interactions (thin lines). The scaffold used for inserting structured RNAs is boxed. **(b)** Cloning vectors used for expressing tRNA-RNA chimeras in *E. coli*. The anticodon stem is replaced by a polylinker composed of unique *EagI*, *EcoRV* and *SacII* restriction sites. This specific combination of restriction sites ensures that the corresponding RNA transcript will fold into an anticodon-like stem and loop structure. **(c)** tRNA-RNA chimera with the epsilon sequence from human HBV virus (boxed) inserted into the human tRNA^{Lys3} scaffold. The cloning sites are grayed. **(d)** Purification profile of the tRNA-HBV chimera shown in panel **c**. *E. coli* transformed with the corresponding recombinant plasmid were grown overnight at 37°C in rich medium. Total small RNA were obtained after whole-cell phenol-extraction. This RNA extract fractionated on a Q-Hiload column, as described in Methods. Fractions were analyzed by urea-PAGE and RNA bands were visualized by UV shadowing. The leftmost lane corresponds to the total RNA extract and successive fractions are shown from left to right. Various *E. coli* tRNA eluted first, followed by small amounts of 5S ribosomal RNA, whereas the tRNA-HBV chimera eluted last, at a NaCl concentration of about 500 mM. After this single purification step, the tRNA chimera is ~90% pure, and the final yield is 50 mg/liter.

Figure 2. Using tRNA chimeras in structural studies. 600 MHz NMR ¹H-¹⁵N TROSY spectrum of the ¹⁵N-labeled tRNA-HBV epsilon chimera (inset) revealing the NH imino groups engaged in base pair hydrogen bonds. Selected peaks corresponding to the previously assigned tRNA scaffold and HBV epsilon moieties are labeled in black and red, respectively. Primed numbers correspond to the HBV epsilon numbering. The peak corresponding to s⁴U8 (leftmost peak) is folded in the ¹⁵N dimension and its abnormal ¹⁵N shift (~179 ppm) is characteristic of the 4-thiolation of the base. Incorporation of pseudouridine Ψ55 is evidenced by the presence of a second imino group (N1), at a characteristic ¹⁵N chemical shift of 135.5 (top right). Similar shifted resonances were observed for m⁷G46, T54 and dihydrouridines (not shown)

Figure 3. Structures of the several RNA chimera that were successfully expressed using the tRNA scaffold strategy. tRNA^{Lys3} derived constructs: **(a)** scaffold; **(b)** human HBV epsilon; **(c)** sephadex aptamer/human HBV epsilon; **(d)** E. coli 23S rRNA, 5' region.

Figure 4. Expression and purification of recombinant RNA. **(a)** Overexpression of various constructs. 1: tRNA^{Lys}/HBV epsilon, 2: tRNA^{Lys}/Sephadex aptamer/Duck HBV epsilon, 3 : tRNA^{Lys}/Malachite green aptamer, 4: tRNA^{Mct}/Sephadex aptamer/HBV epsilon, 5: tRNA^{Lys}/Sephadex aptamer/HBV epsilon, 6: tRNA^{Mct}/streptavidin aptamer/HBV epsilon, control: vector with no insert. Crude RNA minipreps were electrophoresed on an 8% polyacrylamide-urea gel and revealed by UV shadowing. The white triangles indicate the overexpressed RNA band. tRNA^{Mct} based constructs (lanes 4 and 6) were consistently slightly less expressed than the tRNA^{Lys} based chimera (lanes 1-3 and 5). **(b)** Purification of a tRNA^{Mct}/sephadex aptamer chimera from a crude cell lysate under native conditions, using sephadex G100 beads, and soluble dextran for elution. **(c)** Pull-down assay. Two tRNA chimera were used, containing either the sephadex aptamer alone (control RNA), or the sephadex aptamer and helices 40 and 41 of *E. coli* 23S rRNA, which correspond to ribosomal protein L20 primary binding site (23S RNA). They were incubated with purified L20 and sephadex beads. Supernatant was removed, beads were washed with buffer and molecules bound were analyzed by SDS gel electrophoresis and proteins were visualized by coomassie staining. Both RNA chimera were bound (not shown), but only the 23S RNA chimera retained the L20 protein (white triangle). **(d)** Time course of RNase H cleavage of the HBV RNA from the tRNA^{Lys} scaffold (chimera shown in Fig. 3-A2), using "scissor" DNA oligonucleotides complementary to the tRNA scaffold. Control: untreated RNA chimera.

Figure 5. NMR analysis of RNA inserts cleaved with RNase H. **(a)** Structure of the tRNA-16S rRNA A-site chimera that was expressed by growing cells in doubly ¹⁵N-¹³C labeled medium. The two "scissor" DNA oligonucleotides were 19-mers complementary to the scaffold nucleotides sidlined in grey. The RNase H cleaving sites are indicated by blue triangles, they are located six nucleotides within the DNA oligonucleotide sequences. The

corresponding cleaved product is highlighted in yellow. **(b)** ^1H - ^{15}N HMQC correlation spectrum of the cleaved and purified 16S rRNA A-site fragment, centered on the imino proton window. Spectral regions corresponding to A-U and G-C pairs are indicated and individual peaks are labeled with blue triangles. One A-U and ten to eleven G-C pairs are seen, in keeping with the secondary structure shown in panel **(a)**. Additional resonances, corresponding to non-canonical pairs are indicated.

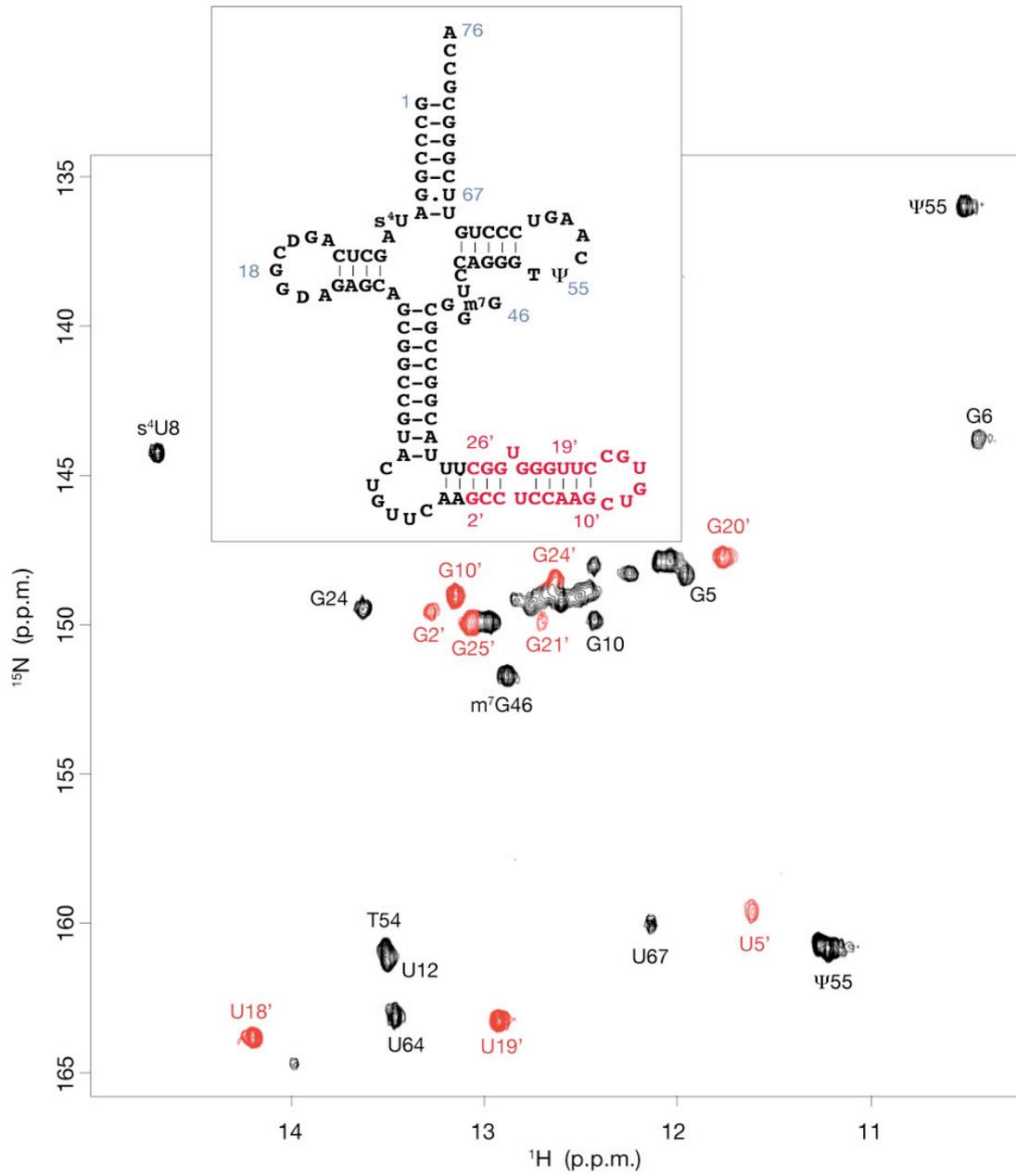


Figure 2

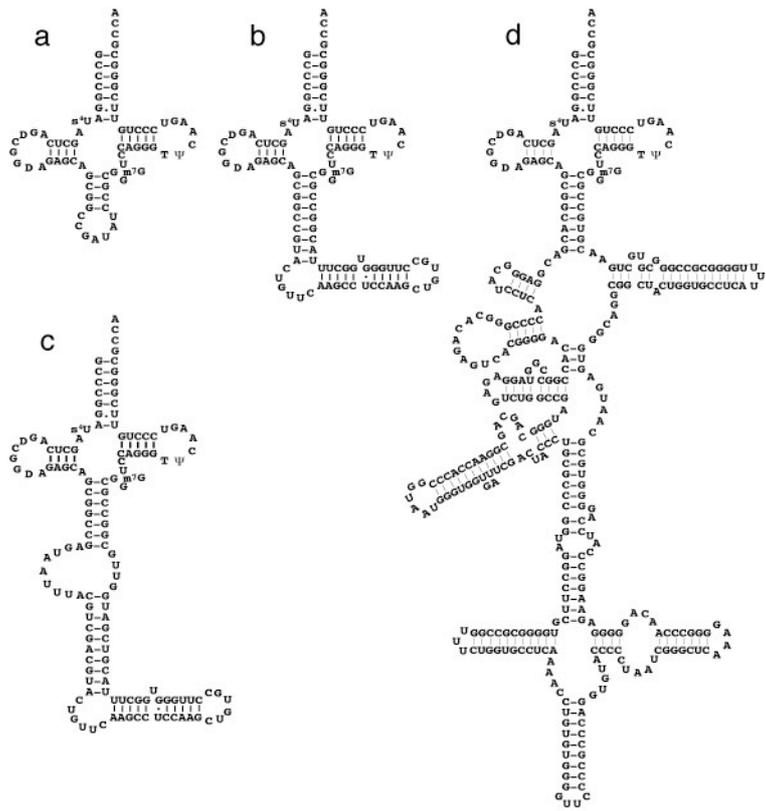


Figure 3

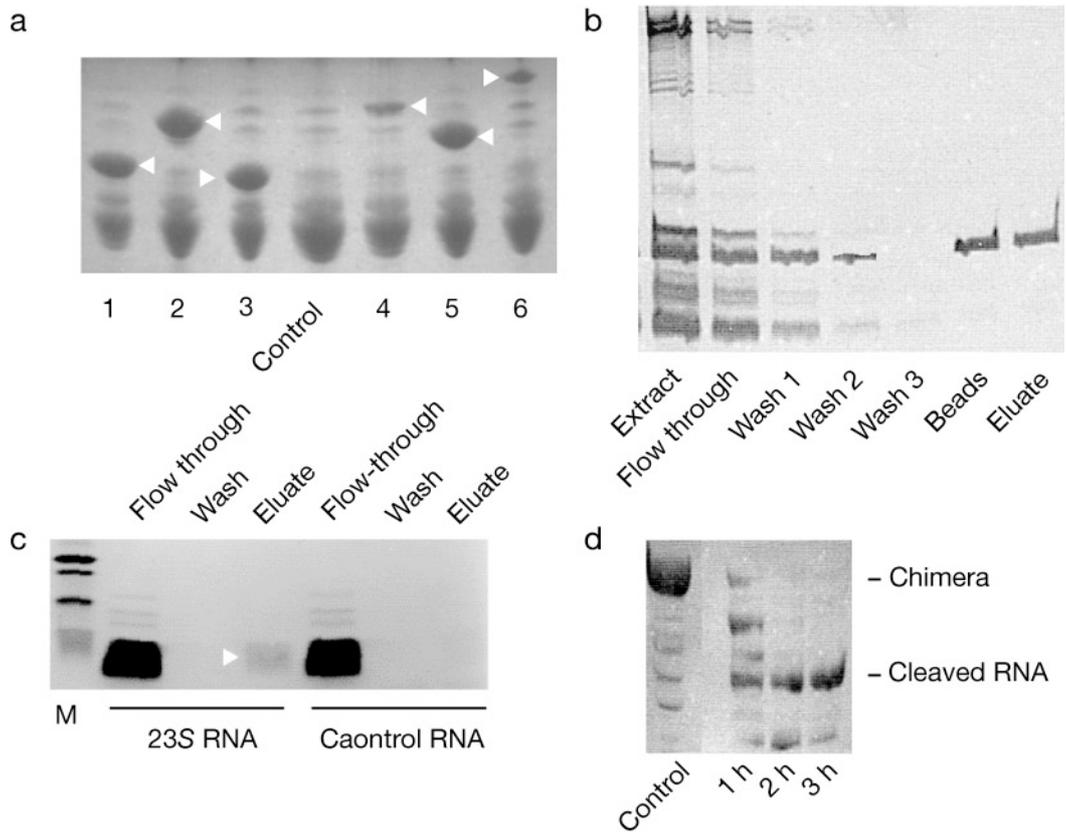


Figure 4

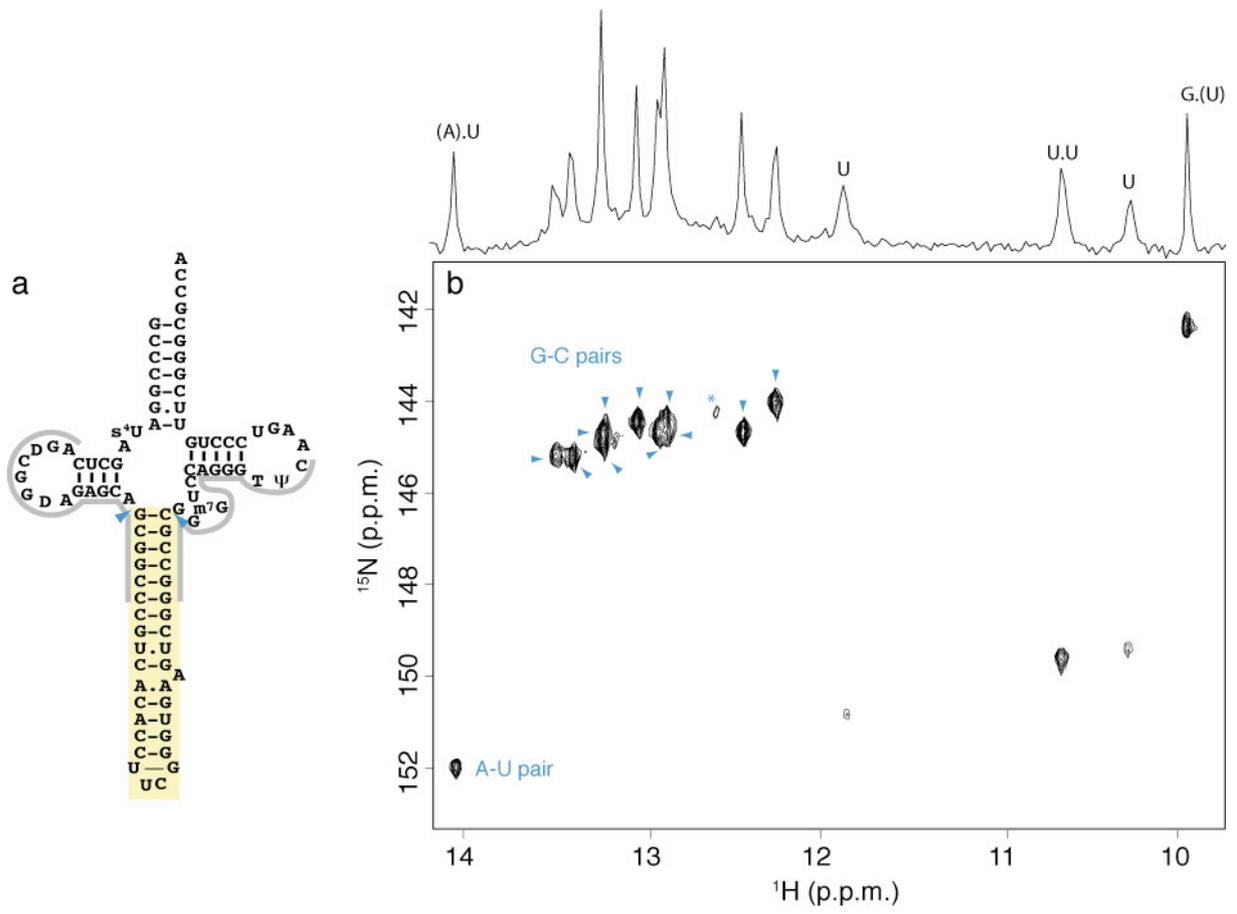
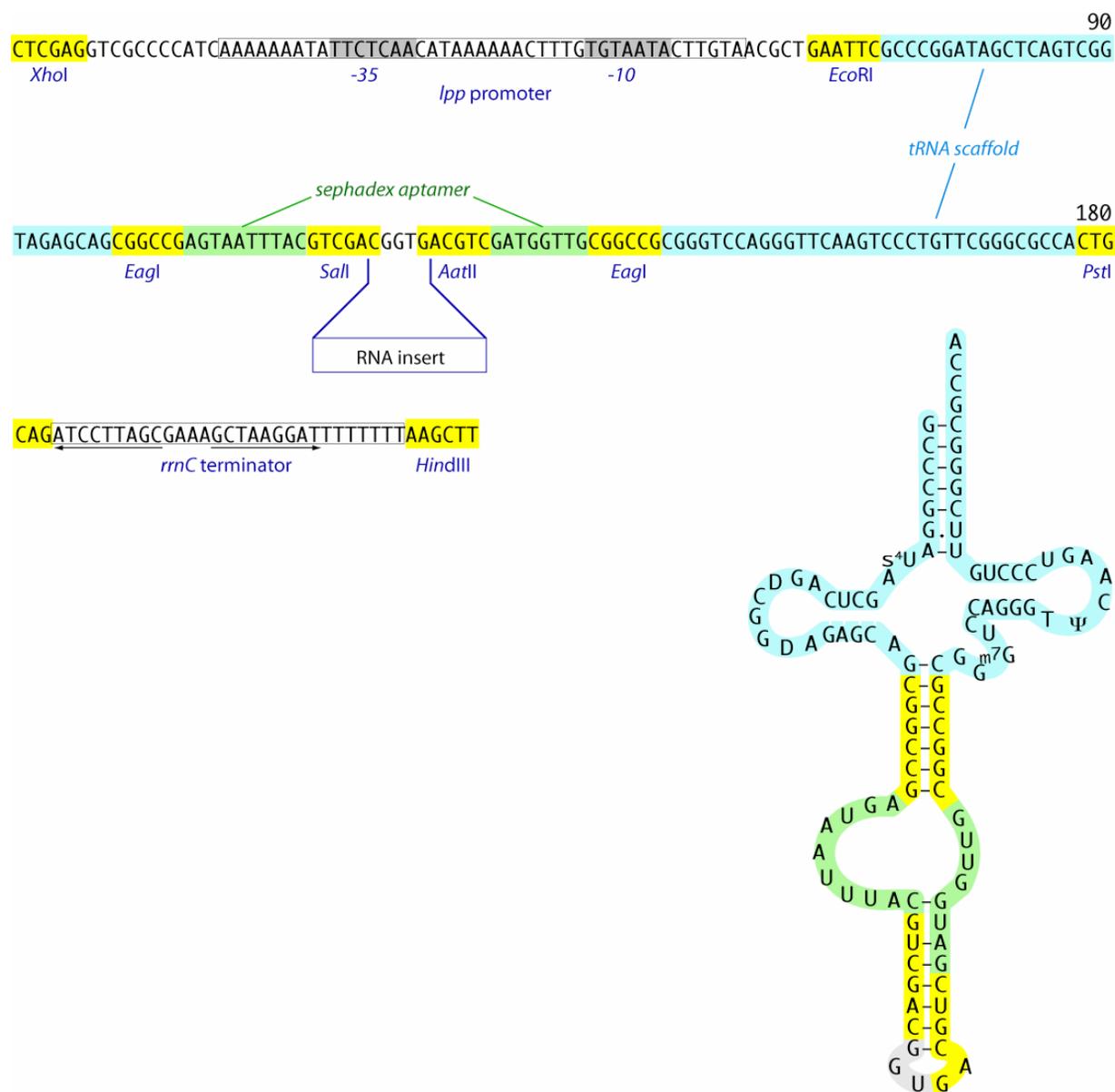


Figure 5

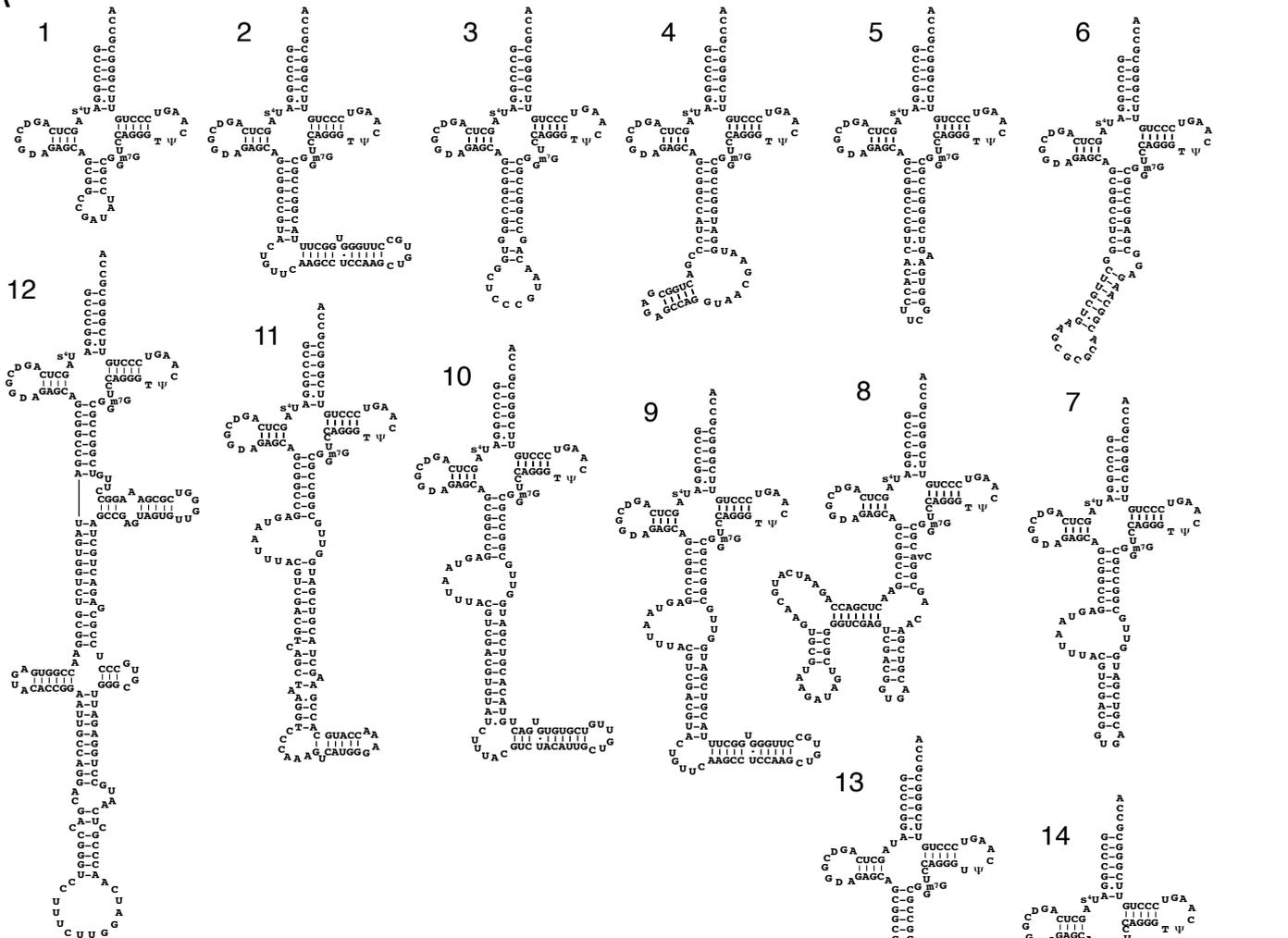
Supplementary Figure 1



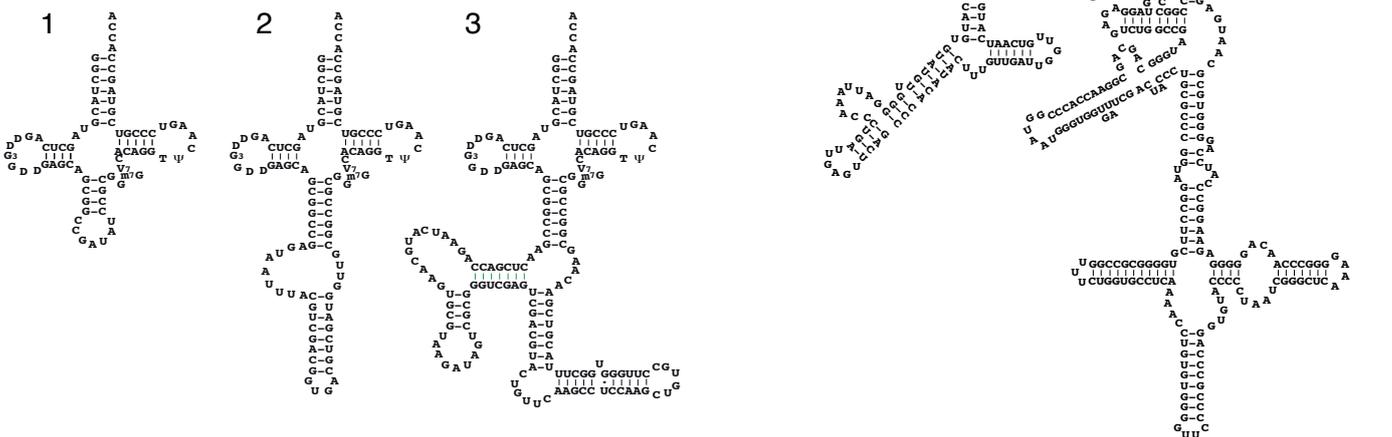
Structure of the expression vector inserts. Shown are the DNA sequence tRNA^{Lys}₃ scaffold with the sephadex aptamer and the corresponding RNA sequence of processed transcript using the same color coding. Relevant restriction sites are indicated in yellow. Within the RNA transcript region, they were selected to maintain the appropriate 2D structure.

Supplementary Figure 2

A

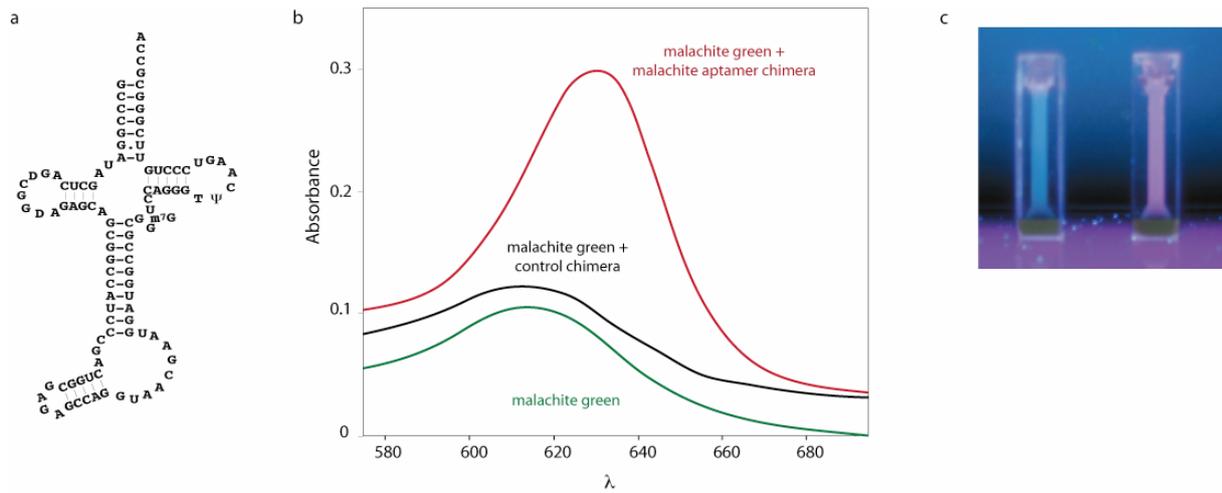


B



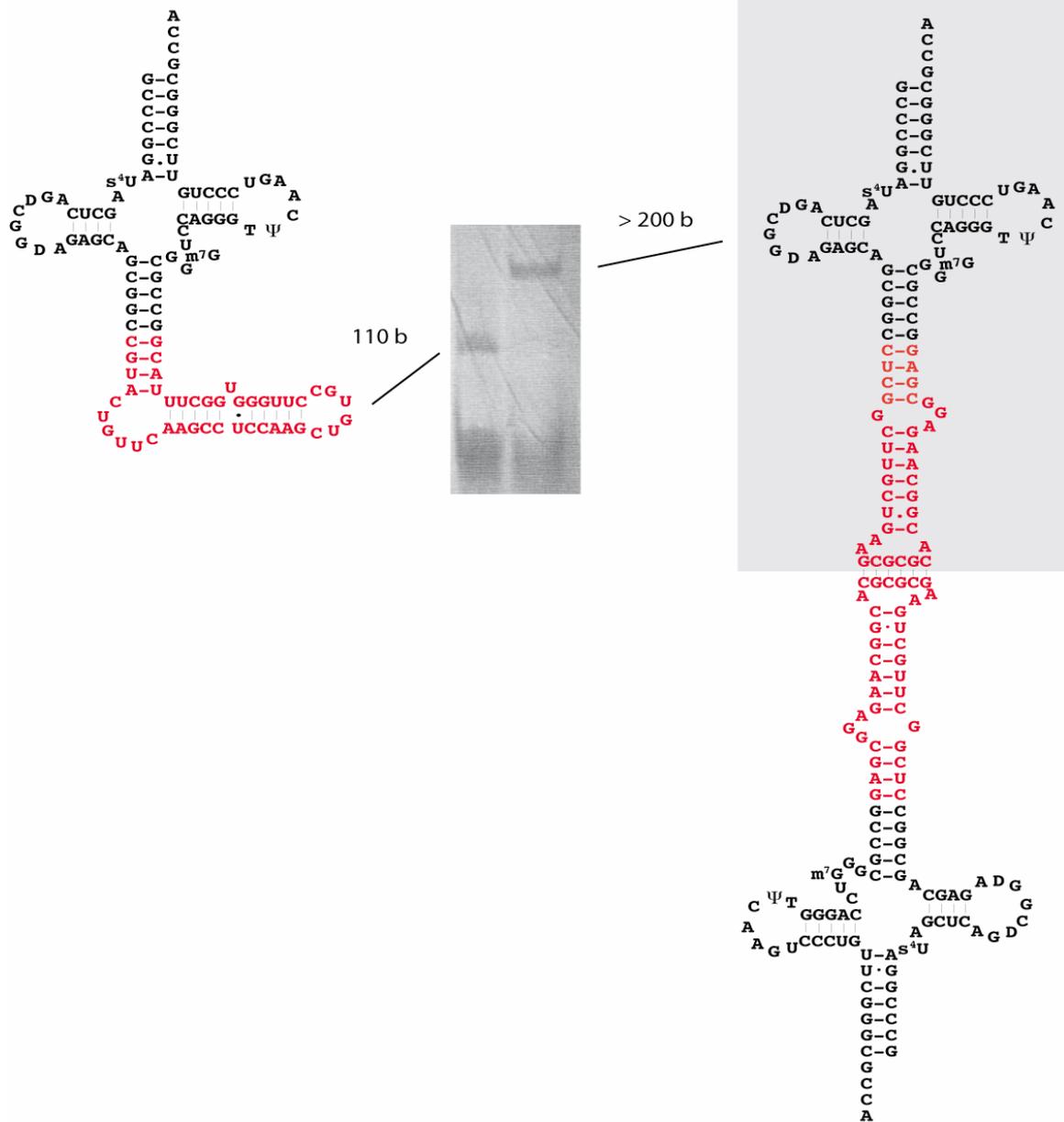
Structures of the various RNA chimera that were successfully expressed using the tRNA scaffold strategy. (A) tRNALys3 derived constructs: 1, scaffold; 2, human HBV epsilon; 3, *A. aeolicus* tmRNA domain; 4, malachite green aptamer; 5, *E. coli* 16S ribosomal RNA decoding site; 6, HIV-1 dimerization initiation site; 7 sephadex aptamer; 8, streptavidin aptamer; 9, sephadex aptamer/human HBV epsilon, short construct; 10, sephadex aptamer/human HBV epsilon, long construct; 11, sephadex aptamer/duck HBV epsilon; 12, HCV IRES domain III; 13 bacteriophage phi29 pRNA; 14, *E. coli* 23S rRNA, 5' region. (B) tRNAMet derived constructs: 1, scaffold; 2, sephadex aptamer; 3 streptavidin aptamer/human HBV epsilon.

Supplementary Figure 3



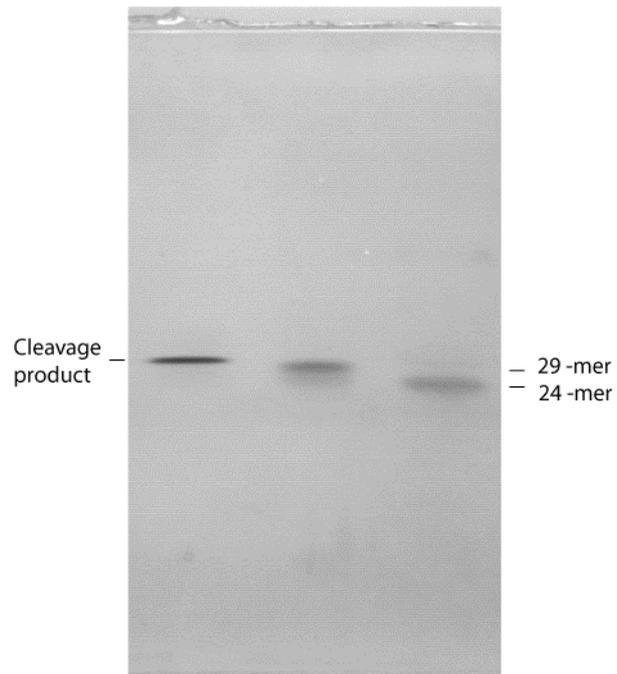
Malachite green aptamer function. **(a)** 2D structure of the chimera containing the malachite green aptamer. **(b)** Absorbance spectra of malachite green (green trace), malachite green with a non specific RNA chimera (tRNA^{Lys} / HBV epsilon, black trace) and malachite green complexed with the specific aptamer RNA chimera. Concentration of the dye is the same in all three experiments. **(c)** Redshift of the malachite green fluorescence in the presence of the specific RNA chimera: left cuvette, malachite green, right cuvette, malachite green complexed with the RNA chimera. The K_d of the dye for the RNA, measured by fluorescence titration, is 50 nM (not shown).

Supplementary Figure 4



Dimerization of the tRNA chimera containing HIV-1 genomic RNA dimerization initiation site (DIS). The apparent RNA size was monitored by native polyacrylamide gel electrophoresis, using RNA minipreps. tRNA^{Lys}/HBV epsilon chimera (left) was used as a control as it has the same size as the HIV-1 DIS chimera (110 bases vs 112 bases). The HIV-1 DIS chimera migrates as an RNA with an apparent size of ~200 bases, in keeping with the formation of a stable complex.

Supplementary Figure 5



PAGE analysis of the cleavage product of a 16S rRNA A site chimera. After purification, the cleaved recombinant RNA was run alongside two oligonucleotide markers and the bands were revealed by silver staining.

Supplementary Table

tRNA ^{Lys} derived constructs	Length (bases)	yield (mg per liter of culture)
human HBV epsilon, short construct (Suppl. Fig. 2A-2)	110	21-50
A. aeolicus tmRNA domain (Suppl. Fig. 2A-3)	91	23
malachite green aptamer (Suppl. Fig. 2A-4)	105	20-25
E. coli 16S ribosomal RNA decoding site (Suppl. Fig. 2A-5)	94	22
HIV-1 dimerization initiation site (Suppl. Fig. 2A-6)	104	16
phage phi29 prohead RNA (Suppl. Fig. 2A-13)	184	22
sephadex aptamer (Suppl. Fig. 2A-7)	115	20-25
HCV IRES domain III (Suppl. Fig. 2A-12)	218	8
sephadex aptamer/duck HBV epsilon (Suppl. Fig. 2A-11)	159	10-15

Overexpression yields of various tRNA chimera. Amounts were estimated from the UV absorbance (260 nm) of the purified recombinant products.