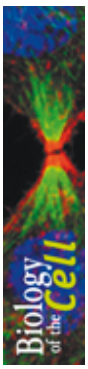
Tethering of proteins to RNAs by bacteriophage proteins.

Cécile Keryer-Bibens, Carine Barreau, Howard Beverley Osborne

Primer Review for Biology of the Cell



Tethering of proteins to RNAs by bacteriophage proteins.

Cecile Keryer-Bibens, Carine Barreau[#] and H.Beverley Osborne*.

UMR 6061 Génétique et Développement,

CNRS/ Université de Rennes 1,

IFR 140 Génétique Fonctionnelle, Agronomie et Santé,

Faculté de Médecine, CS 34317,

2 avenue du Pr. Léon Bernard,

35043 Rennes cedex, France

***Corresponding author H.B.Osborne,**

Téléphone 33 (0) 223 23 45 23; Fax 33 (0) 223 23 44 78; e-mail beverley.osborne@univ-rennes1.fr .

 # present address: Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

Running title: Tethering Proteins to RNAs.

Keywords: MS2 coat protein, Lambda N protein, RNA-binding proteins

Abbreviations

EMSA, Electrophoretic mobility shift assays; UTR, untranslated region.

Abstract.

Many steps in the control of gene expression are dependent on RNA-binding proteins most of which are bi-functional, in as much as they both bind to RNA and interact with other protein partners in a functional complex. A powerful approach to study the functional properties of these proteins *in vivo*, independently of their RNA-binding ability, is to attach or tether them to specifically engineered reporter mRNAs whose fate can be easily followed. Two tethering systems that have been mainly used in eukaryotic cells, namely the MS2/MS2-coat protein system and the lambda N/B box system. In this Primer Review we first describe several studies in which these tethering systems have been used and an overview of most applications is given in tabloid form. We next describe the major features of these two systems and finally we highlight a number of points that should be considered when designing experiments using this approach.

Introduction.

In eukaryotic cells gene expression is regulated both the nucleus and the cytoplasm. Apart from the initial and final steps, transcription and post-translation modifications or protein degradation respectively, the nuclear and cytoplasmic regulatory processes largely involve RNA-binding proteins. These proteins bind either to specific RNA targets that possess defined sequence elements recognised by the RNA-binding protein or to RNA in general. For instance the Hu family of RNA-binding proteins are associated with mRNAs containing AU-rich elements (for a review see Barreau et al. 2005). In contrast, the polyA binding protein will bind to any polyadenylated RNA, that is, virtually all cellular mRNAs except those encoding histones. Due to the importance of RNA-binding proteins in cellular responses one of the major tasks in molecular cell biology is to determine the biological role of these proteins and the mechanisms by which they achieve their various functions. Indeed, RNA-binding proteins are involved in a wide spectrum of both regulated and constitutive functions. These include, in the nucleus, pre-mRNA splicing, mRNA export to the cytoplasm and in the cytoplasm, mRNA localisation, translation and stability. Much ground work can and has been done by *ex-vivo* biochemical studies, but the real proof of biological function requires *in vivo* analyses.

One of the classical ways to study the functional role of a RNA-binding protein when bound to a specific mRNA is to use reporter genes, with or without the cognate sequence

element , encoding gene products that are more easily quantified than those translated from most natural target mRNAs. However, it is not always possible to ensure that, *in vivo*, preventing the binding of one RNA-binding protein does not facilitate the association of another. In addition, this approach requires the careful and precise mapping of the binding site within the RNA and the verification that, in the context of the cell, only the protein under study can associate with the delimited sequence element. A different approach to studying the function of RNA-binding proteins *in vivo* is to tether the protein to a reporter mRNA. This powerful technique is based on the use of two recombinant molecules. Firstly, a hybrid mRNA composed of the coding region of the protein under study fused in-frame to the RNA-binding domain of a well characterised RNA-binding protein and secondly a recombinant reporter mRNA encoding an easily measurable protein and the sequence element (tether) recognised with high affinity by the exogenous RNA-binding domain of the fusion protein. When co-expressed in a cell the fusion protein should be bound or tethered to the reporter mRNA via the exogenous RNA-binding domain thereby imposing on the reporter mRNA the functional consequences of its association with a mRNA. In Figure 1 are schematically depicted a number of situations that can occur. The first case, line A, is the simplest situation where the association of a RNA-binding protein is sufficient to provoke the regulated function. The tethered protein should then be fully functional. In case B, the RNA-binding protein requires a partner and if the tethered protein can form a similar complex (line B-1) there will be a functional read-out. In some cases the partner may be the function provoking protein, then tethering it alone will be sufficient to cause the regulatory events (line B-2). The tethering approach also allows the function of a protein to be studied independently of its binding to cognate sites in the RNA. By including in the fusion protein only specific regions or domains of the studied RNA-binding protein, its domain structure can be determined as functional units and the requirement for other proteins as partners or co-factors can be assessed (line B-3). However, it must be remembered that the tethering approach is based on an artificial situation that may not recapitulate the normal *in vivo* scenario. For instance, structural changes caused by the binding of a protein to its cognate RNA site will not be reproduced (Case C in Figure 1). Furthermore, the geometry of the complex tethered to the reporter mRNA may not be compatible with all or part of the functions that can be expressed by the tethered protein (Case B-4 in Figure 1 and see section on Practical Considerations). Hence, negative results in a RNA-tethering assay are particularly uninformative and, in most cases, can not be taken as evidence that a particular function is not associated with the tethered protein.

The tethering approach has been used with success by a number of laboratories and principally two tethering proteins of bacteriophage origin have been used; the coat protein from the MS2/R17 bacteriophage (Bardwell and Wickens 1990) and the 1-22 peptide of lambda N protein (De Gregorio et al. 1999). Although other RNA-binding proteins of eukaryotic origin can and have been used as a tether (see De Gregorio et al. 1999) there are in principle draw-backs. For instance, the Iron Responsive Protein will bind with high affinity to the Iron Responsive element, however such elements exist in the 5' or 3' untranslated regions (UTR) of a number of mRNAs (for a recent review see Pantopoulos 2004) and the expression of the hybrid protein could affect the fate of these mRNAs in addition to that of the reporter.

The first use of MS2 coat protein as a tether to catch an RNA was described by Bardwell and Wickens (1990) in a protocol to purify RNA-protein complexes. The basic idea in this protocol was built on by SenGupta et al. (1996) when they developed the three hybrid system in yeast cells which can be used to identify the RNA targets for a RNA-binding protein. The first tethering assays *per se* were reported by Coller et al. (1998) and (Graveley and Maniatis 1998). Coller et al. (1998) fused the poly(A) binding protein 1 to MS2 coat protein and showed that, in yeast cells, poly(A) binding protein 1 tethered to the 3'UTR can stabilise the targeted reporter independently of binding to the poly(A) tail but dependent on translation of the mRNA. By using various deletion mutants of the poly(A) binding protein 1, they genetically separated the domains required for mRNA stabilisation and poly(A) binding. Graveley et al. used this same approach to study SR-protein function in splicing reactions. First they showed that the RS-domain of these proteins can act as a splicing activator independently of the RNA-binding domain function, indicating that the general and enhancer-dependent functions of SR proteins can be uncoupled (Graveley and Maniatis 1998). In a further study, by changing the position and the number of MS2 coat protein binding sites, they analysed the effect on splicing efficiency of the distance between the tethered SR-protein and the intron and the number of bound SR-proteins (Graveley et al. 1998). These first studies illustrate the variety of information that can be obtained using a tethering assay. Concomitant with the work described above, Bertrand et al. (1998) used tethering via the MS2 coat protein to follow mRNA localisation. By expressing a recombinant ASH1 mRNA containing 6 MS2 sites in the 3'UTR and Green Fluorescent Protein fused in frame to the MS2 coat protein they were able to follow the localisation of ASH1 mRNA to the bud tip in *Saccharomyces cerevisiae*. The first report using the lambda N peptide to tether a protein to a mRNA containing box B sequences addressed the question of the role of eIF4G in translation

initiation (De Gregorio et al. 1999). In this case the protein was tethered to the mRNA in front of the AUG initiation codon by a single box B element.

Tethering via the lambda N peptide or the MS2 coat protein appears to produce similar results. Indeed, functional studies of the factors required for Nonsense-Mediated Decay of mRNA have used both systems (Lykke-Andersen et al. 2000, 2001; Gehring et al. 2003; Gehring et al. 2005; Kunz et al. 2006) and, as non-conflicting results were obtained, these studies secondarily indicate that the functionality of the complex formed around the tethered protein is independent of the tethering protein. (Bardwell and Wickens 1990; De Gregorio et al. 1999; Pantopoulos 2004)

However, the tethering systems based on these two bacteriophage proteins have fundamental differences that may not be immediately obvious to the experimenter. For instance, the lambda N peptide binds to cognate RNA sequence as a monomer whereas the functional MS2 coat protein is an obligate dimer. At concentrations above 1μM, which can be achieved *in vivo*, the wild type MS2 coat protein aggregates so that specific mutants devoid of this capacity must be used (see below). Both proteins bind to stem-loop structures but again the constraints are different. For MS2 coat protein only the loop sequence is important whereas lambda N protein binds to nucleotides in both the loop and the stem.

Although the experimental approach described above can be used as a simple tool box, we think that it would benefit from a better conceptual understanding of the mechanisms involved. Hence we now describe the salient features of these two proteins and their cognate RNA binding sites that are important for their use as tethering proteins and target sites. We then consider some of the practical aspects of tethering assays with the MS2 coat protein and the lambda N peptide. This methodological approach has been previously described in several reviews (Coller and Wickens 2002; Baron-Benhamou et al. 2004), that by Baron-Benhamou et al. more particularly addresses protocols for using the lambda N/ B box system.

MS2/R17 coat protein based tethering.

The RNA bacteriophages of *Escherichia coli* are divided into four different groups based on seriological cross-reactivity, replicase specificity and physical properties (Fiers 1979). The MS2 and closely related R17 bacteriophages are in group I. During the late stages of *Escherichia coli* infection by these bacteriophages, the translation of the replicase gene is repressed due to binding of the coat protein to specific sequence elements in the replicase

mRNA. The binding sites of the MS2 and R17 coat proteins and the coat proteins themselves are highly similar; they will be considered here as interchangeable.

Coat protein binding sites.

The translational repression of the replicase mRNA is achieved by the specific binding of the bacteriophage coat protein to a segment of the bacteriophage RNA that contains the initiation codon (AUG) of the replicase reading frame (Bernardi and Spahr 1972 and references therein; reviewed in Johansson et al. 1997). Further work has shown that within this segment of RNA the AUG initiation codon is positioned within a stem-loop (Jansone et al. 1979). The minimum binding site for the R17 coat protein is a 21 nucleotide stem-loop that has a particular secondary structure (see Figure 2A) (Carey et al. 1983b; Romaniuk et al. 1987). The identity of the nucleotides forming the stem does not appear to influence the affinity of the coat protein for the RNA. The sequence of the loop which contains 4 nucleotides (AUUA) is more important and mutation of the adenosines reduces the affinity of the R17 coat protein for the RNA by at least 100 fold (Carey et al. 1983b; Romaniuk et al. 1987). Several studies showed that although changing the loop sequence to ACUA did not affect R17 coat protein binding, the affinity was increased when the loop sequence was changed to AUCA (K_d for wt loop 1-3 10^{-9} M; K_d for AUCA loop 2-6 10^{-10} M) (Carey et al. 1983b; Lowary and Uhlenbeck 1987; Romaniuk et al. 1987; Wu et al. 1988) (see Figure 2A). The increased affinity of the C-loop for R17 coat protein is attributed to a slower dissociation rate of the complex (Lowary and Uhlenbeck 1987). The C-loop sequence is now generally used as a high affinity binding site for the R17 or MS2 coat proteins in binding or tethering assays (see Table 1).

The second feature of the RNA stem-loop required for coat protein binding is an unpaired adenosine two nucleotides up-stream of the loop. Deletion of this bulged-out base (Carey et al. 1983b) or substitution for a cytosine (Romaniuk et al. 1987) drastically decreases the affinity of the RNA for the R17 coat protein. Truncation from the 5' or the 3' side of the binding site showed that the bottom of the stem (below the bulged out A) should contain at least three paired nucleotide to maintain wild type affinity (Carey et al. 1983b). The identity of the unpaired nucleotides 5' and 3' to the stem does not appear to be important (Romaniuk et al. 1987; Wu et al. 1988) and they may even be dispensable (Dertinger et al. 2001).

Coat protein.

The MS2 and R17 coat proteins are bifunctional, acting both as translational repressors of the replicase mRNA and as structural proteins of the bacteriophage particle. They have a molecular weight of 13.7 kDa and contain 129 amino acids. In solution these proteins exist stably as dimers. At concentrations above 1μM they aggregate to form structures that closely resemble bacteriophage capsids that can be devoid of RNA (Johansson et al. 1997). The capsids are icosahedral structures composed of 90 dimer elements bound to the bacteriophage RNA which is thereby protected inside the capsid. The X-ray structure of the MS2 capsids has been determined (see Ni et al. 1995; Valegard et al. 1997 and references therein). This crystallographic data shows that in the dimer the two MS2 coat protein molecules are orientated head to tail or anti-parallel; the N terminal region of one molecule is in close proximity to the C terminal regions of the other molecule (see Figure 2B). One RNA molecule is bound to the MS2-coat protein dimer. The geometry of the RNA binding site and the amino acids involved in contacts with the RNA have been defined both genetically (Peabody and Lim 1996) and by X-ray crystallographic analysis (Valegard et al. 1997). The RNA binding site spans across the surface composed of 10 β-sheets (2x5) formed by the conjunction of the two coat protein monomers (see Figure 2C).

When the coat proteins aggregate into capsids they are no longer available to bind to a mRNA undergoing translation (Carey et al. 1983a) and this leads to an apparent decrease in affinity. A number of mutations have been introduced into the coat protein that, by preventing capsid formation, increase the apparent affinity of the protein for the stem-loop structure in the target mRNA (Peabody and Ely 1992; LeCuyer et al. 1995). The structural studies of MS2 coat protein (Ni et al. 1995; Valegard et al. 1997) have shown that the 15 amino acid loop between the F and G β-strands (see Figure 1B) are important for interaction between the separate dimers in the bacteriophage particle. Accordingly, either deletion of residues 67-79 from this loop (Peabody and Ely 1992), denoted as "dlFG", or mutations within the loop (LeCuyer et al. 1995) can prevent aggregation with a concomitant increase in the apparent affinity. The most efficient FG loop mutant was the V75E;A81G mutation that increased the concentration required for *in vitro* capsid assembly at least 1000 fold and prevented capsid formation in *E.coli* cells without significantly affecting either binding affinity or specificity. The W82R mutation within the FG loop also prevents capsid assembly. Although the dlFG and W82R mutations prevent bulk capsid formation, multimers higher than dimer do form. These multimers do not appear to be in equilibrium with the dimers proficient to bind RNA as

none are observed when the dimers are isolated and re-analysed by gel filtration (Peabody and Ely 1992). A similar analysis for the other point mutations within the FG loop has not been reported. When the crystallographic structure of these mutant protein dimers was compared to that of the coat proteins associated in capsids no significant difference in the overall structure was observed, although the conformation of the FG loop is different (Ni et al. 1995). Mutations have also been isolated that increase the affinity of the coat protein for the RNA stem-loop structure (Lim and Peabody 1994). However no increase in the specificity of the mutant protein for the wild-type or C-loop sequence was observed (Lim and Peabody 1994).

The MS2 coat protein dimer contains two head-to-tail molecules. This suggested that a functional dimeric protein could be synthesised from a single mRNA by a genetic fusion of the two reading frames. When the two complete reading frames were fused with a single amino acid linker, a complex pattern of oligomerisation was observed. Deleting two or three amino acids around the initiation (ATG) codon of the second reading frame produced a dimer with self association properties and an affinity to the MS2 stem loop similar to that of the wild type protein (Peabody and Lim 1996). Another property of this genetically fused protein is an increased stability. This was observed when the octapeptide Flag sequence was inserted into the reading frame at different positions. The protein synthesised from the genetic fusion was more tolerant to these insertions (Peabody 1997).

Lambda N protein /B box system

The lambda bacteriophage is the prototype of a number of bacterial DNA viruses collectively called "lambdoid phages" (for a review see Arber 1983). In many DNA bacteriophages, the transcription of the phage encoded genes is temporally regulated by sequential promoter usage and the synthesis of antitermination proteins. After infection, transcription is initiated from two promoters (pR and pL) in the bacteriophage genome allowing the synthesis, respectively, of the transcriptional repressor *cro* and the antiterminator N. In the absence of the antiterminator N, transcription terminates after these two genes. The antiterminator function of N requires two genetically defined regions, nut_R and nut_L, situated within the *cro* and *N* genes, and it allows the read-though of the termination sites 3' to these genes. In fact, the regulatory protein N is a RNA-binding protein that binds to a specific sequence element in the transcribed RNA. Once bound to the RNA, N participates in a complex containing the host cell factors NusA and the RNA-polymerase. When present in this

complex, the polymerase acquires the capacity to read-through the two termination sites so that genes down-stream of *cro* and N are transcribed (Barik et al. 1987; Horwitz et al. 1987).

N-binding sites.

The nutR and nutL sites contain two conserved sequence elements, a 8-12 bp element called box A and about 5 to 10 bp downstream a 15 bp element called box B that has the characteristics of an interrupted palindrome with the ability to fold into a hairpin structure (for a review of early work see Friedman and Gottesman 1983). The lambda N protein binds to the RNA box B element (Lazinski et al. 1989). It should be noted that nutR and nutL box B elements differ by one base in the loop (Figure 3A). In the analysis of sequence requirements for lambda N binding to the box B element the majority of the studies have used one of two approaches, either the quantification of antitermination efficiency as a measure of functional sequence elements (Doelling and Franklin 1989; Lazinski et al. 1989; Franklin 1993), or quantification of protein binding by Electrophoretic Mobility Shift Assays (EMSA) (Chattopadhyay et al. 1995; Tan and Frankel 1995; Su et al. 1997a). In general the results from these two types of analysis concord but there are some differences. For instance, Doelling and Frankin (1989) observed that changing any of the bases in the loop sequence constrained antitermination of transcription by lambda N protein, whereas Chattopadhyay et al. (1995) observed using EMSA that G_6 and A_8 of the loop sequence are the most critical for N protein binding. We will concentrate here on the data relative to direct protein binding as this is a major concern in a tethering assay.

The structural requirements for N-protein binding to the box B elements were extensively studied by Chattopadhyay et al. (1995). These authors observed that in the B box/protein N complex only the 5' arm of the box B stem was protected from RNase attack and they suggested that binding of N to the box B element was asymmetrical. Structural studies of an oligomer containing the box B sequence either in the presence or absence of N-protein (see below) have confirmed the asymmetric nature of the RNA-protein association in this complex (Legault et al. 1998). The optimal binding of N protein to the box B element also requires a specific sequence of the stem (Chattopadhyay et al. 1995). Using several biochemical methods the box B element was shown to bind one N-protein with an affinity (Kd) of $1.3 \pm 0.4\ 10^{-9}$ M (Van Gilst et al. 1997). This value for the Kd is about 10 fold higher than that measured by EMSA ($10\text{-}20\ 10^{-9}$ M) (Chattopadhyay et al. 1995) but both are

significantly higher than the Kd for non-specific binding to RNA, DNA or mutant box B sequences ($1 – 2\ 10^{-6}$ M) (Van Gilst et al. 1997).

Lambda N protein

The lambda N protein is a small (12.2 kDa, 107 amino acids) basic protein. The homologous proteins encoded by other bacteriophages such as P22, although they conserve the same function but with different RNA sequence specificity, do not show any significant sequence similarity apart from a 18 amino acid region near the N-terminus (Franklin 1985). In fact this region alone is sufficient to confer sequence specific binding which was demonstrated using hybrid proteins between different regions of several lambdoid N-proteins (lambda, *Φ*21 and P22). The binding specificity was found to be defined by the origin of the N-terminus and the arginine amino acids in this region were shown to be particularly important (Lazinski et al. 1989; Franklin 1993). Subsequently, (Tan and Frankel 1995) showed that a synthetic peptide composed of the 22 N-terminal amino acids of lambda N-protein (Figure 3B) could bind to the box B RNA element with an affinity and specificity close to that of the full length protein. The N-terminal methionine appears to be dispensable for binding (Cilley and Williamson 1997; Su et al. 1997a) but removal of the three C-terminal amino acids from this peptide reduces the Kd about 5 fold (Cilley and Williamson 1997). This is compatible with the scanning mutagenesis data of Su et al (Su et al. 1997a) using the 2-22 peptide derived from lambda N. These authors observed that the mutation of amino acids 21 (Ala) or 22 (Asn) to glycine and alanine respectively, reduced binding by about 50%. In this same study, only three amino acids could be mutated without affecting binding: threonine 5; glutamine 13 and lysine 19. In addition, none of the arginine amino acids could be changed to lysine without affecting the binding, although some changes were more deleterious than others. In contrast, changing lysine 14 to an arginine was without effect (Su et al. 1997a). When the binding of the 2-22 lambda N peptide to various mutant box B elements was studied, the relative importance of the different nucleotides in the loop and at the top of the stem was similar to that observed for the full length protein, only the mutation of the fourth loop nucleotide from A to C was without effect (Cilley and Williamson 1997).

Using a fluorescence based binding assay Austin et al. (Austin et al. 2002) observed that the binding affinities of the P22 N 1-21 peptide for the P22 box B was greater than that of the lambda N 1-22 peptide for the lambda box B sequences. This led these authors to make

various mutant lambda N 1-22 peptides, based on the comparison with the P22 peptide sequence, which had an enhanced affinity (about 100 fold) for the lambda box B. To our knowledge these mutant peptides have not been incorporated into hybrid proteins for use in tethering assays.

Structural changes incurred by binding.

Arginine rich domains that bind to RNA have been identified in a number of RNA-binding proteins (for a review see Weiss and Narayana 1998). Like the peptide of HIV-Rev that recapitulates binding of the protein to the cognate RNA hairpin, the 1-22 lambda N peptide has the potential to take on an α-helical conformation. Disruption of this conformation by mutation of Glu 9 to proline completely abolished binding to the box B element whereas changing this amino acid to alanine, that conserved the helical structure, only reduced binding (Tan and Frankel 1995). This is in accordance with the scanning mutagenesis data of Su et al. (Su et al. 1997a). Non-specific RNA binding (K_d > 5 μM) was not affected by the helical content of the peptide (Tan and Frankel 1995). Most interestingly, concomitant studies demonstrated that binding of either the complete lambda N protein (Van Gilst et al. 1997) or the 2-22 peptide (Su et al. 1997b) to the box B element induced a random coil to α-helical transition. For the peptide the α-helical content increased from only about 10% in the absence of RNA to 80% in the protein-RNA complex. This random coil to α-helix transition is specific to binding to the box B element (Van Gilst et al. 1997).

The RNA moiety of the lambda N / box B complex also undergoes structural changes on binding (Su et al. 1997b). In the absence of the peptide, the box B RNA consists of a stem and a flexible loop. When the RNA is bound by the peptide the U_5A_{11} base pair is stabilised and pairing of G_6 and A_{10} occurs as a sheared GA pair (reversed Hoogsteen), a structure first observed in the GNRA tetraloop (for a review see Conn and Draper 1998). This GA base pairing is coherent with previous observations that mutation of either, or both, of these bases abrogated functional binding of lambda N protein or the peptide (Doelling and Franklin 1989; Chattopadhyay et al. 1995; Tan and Frankel 1995; Cilley and Williamson 1997; Van Gilst et al. 1997) and the induced α-helical transition is also absent with a $A_{10}C$ mutant box B RNA (Su et al. 1997b).

The NMR structure of the lambda N peptide / box B RNA (nutL sequence) complex has been determined and confirmed the above observation (Legault et al. 1998) (Figure 3C).

The N_{1-22} peptide formed two α-helical segments (amino acids 4-10 and 12-21) and a local perturbation at Arg_{11} caused a 120° bend between these two segments. The peptide bound to the major groove of the RNA stem but this interface only involved C_4 to A_{10} and the 5' phosphate of G_{11}. The RNA stem had the conformation of a regular A-form helix with the G_6A_{10} base pair giving a particular twist so that A_7, A_8 and A_{10} were stacked in the loop and G_9 was excluded. The orientation of G_9 permitted stacking of the guanosine base with the ribose of A_8. In the nutR box B element this ninth base is an adenosine which should similarly stack with the ribose of the preceding base. Deleting this base only reduces N-binding about 2 fold, but binding of the host cell factor NusA was abrogated (Legault et al. 1998), showing that this base is required to form a functional antitermination complex. The structure of the loop induced by N peptide binding largely explains the sequences requirements of the box B element described above.

Practical considerations.

In tables 1 and 2 are summarised some of the studies in which the tethering approach has been used. To date it would appear that the MS2 coat protein (Table 1) has been used in these studies to a greater extent than the lambda N peptide (Table 2). The reason for this is not evident and may simply be historical. As shown in Table 1, both the dlFG and the V75E;A81G mutations of the MS2 coat protein have been used to tether functional proteins to a RNA. The genetically fused MS2 coat protein dimer has also been used.

In the bacteriophage, translational repression by the MS2 coat protein and antitermination by the lambda N protein are achieved with one of the cognate stem-loop structures. However, at least for the binding of MS2 coat protein to non-bacteriophage RNA non-convergent data have been reported as to the number of stem-loops required. Bardwell and Wickens (1990) observed that two sites were required to achieve R17 coat protein binding to a hybrid RNA containing the binding sites within a non-R17 sequence background. Coller et al. (1998) used two MS2 sites on the basis that the R17 coat protein exhibited cooperative binding when two sites were present (apparent Kd about $2x10^{-9}$ M) (Witherell et al. 1990). However, it should be noted that the above studies were done using the wild type protein and no cooperative binding was observed for some of the non-aggregating mutants of MS2 coat protein (LeCuyer et al. 1995). Using a non-aggregating mutant of MS2 coat protein, Ko and Gunderson (2002) obtained a tethering of splicing factors to RNAs containing a single

stem-loop. As mentioned above the R17 and MS2 coat proteins bind as a dimer to the stem loop and in the experiments of Ko and Gunderson (2002) the coat protein (non-aggregating mutant form) was translated as two tandem copies so that a single polypeptide chain forms a binding proficient protein.

The number of binding sites used for the lambda peptide has also varied (Table 2). In their initial study De Gregorio et al. (1999) used only one box B sequence, however in a subsequent study of the Y14 and hUpf3b factors of Nonsense-Mediated Decay, at least 3 box B elements were required to elicit a measurable effect of the tethered proteins (Gehring et al. 2003).

A number of factors can affect the results obtained in a tethering assay:

i) The sequence into which the MS2 coat protein or lambda N peptide binding sites are embedded may be of importance and some of the variability as to the number of sites required may be due to the presence of extraneous sequences. For instance, Witherell et al. (1990) showed that binding of MS2 coat protein can be prevented by occlusion of the binding site(s) by alternative structures.

ii) The position within the reporter mRNA may also affect the observed effect. In most cases the tethering site(s) have been placed within the 3' UTR of the reporter mRNA. Often this has been the region of the endogenous target mRNAs to which the studied protein bound but furthermore, tethering a protein near to the 5' end of the mRNA can inhibit translation independently of its normal function(s) (Stripecke et al. 1994; Grskovic et al. 2003).

iii) Depending on the function expressed by the tethered protein the number of binding sites may be of importance. For instance, increasing the number of MS2 binding site did not enhance the translational stimulation conferred by the histone hairpin-binding protein (Gorgoni et al. 2005). In contrast, the translation of the reporter mRNA increased with the number of MS2 tether site for the RNA-binding protein Dazl (Collier et al. 2005). Furthermore, the local concentration of binding sites on the RNA (defined by the number of binding sites) may be important if the intracellular concentration of the fusion protein is limiting. This consideration suggests that increasing the number of MS2 coat protein or lambda N peptide binding sites can only be beneficial. In at least two studies reporter RNA containing increasing numbers of binding sites have been used. In their studies of Nonsense Mediated Decay factors, Lykke-Anderson et al. (2000) used 2, 4, 6 and 8 MS2 sites inserted into the 3' UTR of the β-globin gene and Gehring et al. (2003) used 2, 3 or 5 box B elements

within the same reporter. In the absence of tethered Nonsense Mediated Decay factors, no effect of the number of binding sites on the steady state level of the reporter mRNA was noted in either of these studies. However, using a different reporter system (Firefly and Renilla luciferases driven from a bidirectional promoter (see Barreau et al. 2006) we observed that increasing numbers of the MS2 binding site (2, 4, 6, 8) cause a clear decrease in the steady state level of the reporter mRNA (Figure 4). The reason for these apparent differences is not clear but may be related to the intrinsic stability of the different reporter mRNAs used.

iv) The stoichiometry of the tethered protein should also be considered. Many RNA-binding proteins function as dimers. When a protein is fused in frame to MS2 coat protein, which binds as a dimer to the cognate site, then two copies of the protein being studied will also be present. This however will not be the case if the genetic fusion of two MS2 coat proteins is used. The lambda N peptide also binds as a monomer, so in these two cases two binding sites will be required to allow dimerisation of the fused protein. The work of Ko and Gunderson (2002) is illustrative of this point. By using various numbers of MS2 binding sites they showed that poly(A) polymerase activity could be inhibited by one copy of U2AF65 or U170K whereas two copies of SRP75 or U1A were required. The inhibitory activity of these proteins was correlated with the number of Poly(A) polymerase regulatory domains present in these proteins.

v) Stoichometry was also evoked to explain the lack of inhibition produced by tethering the translational inhibitor drosophila sex-lethal to a reporter mRNA (Grskovic et al. 2003). The authors showed that the RNA-binding domain and the 7 amino acids C-terminal to this domain were sufficient to confer translational inhibition. Structural studies have shown that the two RNA-recognition motifs in this domain only become orientated when bound to the cognate RNA binding site (Handa et al. 1999). Therefore, an additional possibility is that the structural change induced by RNA-binding, which is not reproduced when sex-lethal is tethered to the RNA, is required to form a motif recognised by cofactors necessary for translational inhibition.

vi) The orientation of the fused proteins will also depend on the tethering protein. The two monomers in the MS2 coat protein dimer that binds to RNA are in anti-parallel orientation (see above) and hence the fused proteins will be positioned on either side of the complex. In general, fusion proteins are made with a hinge or linker region between the two moieties which may have sufficient flexibility for the two monomers of the fused protein to associate. However, the anti-parallel orientation of the MS2 coat protein moiety will probably influence

the orientation of the fused protein. This would be particularly important if the monomers of the fused protein under study associate in a parallel orientation. In such a case we would predict that a longer hinge/linker would be required. When the tandem copy form of MS2 coat protein or the lambda N peptide are used then the two monomers of the fused protein will have a parallel orientation. Accordingly, the inverse situation arises where, if the monomers of the fused protein associate in an anti-parallel orientation, a longer hinge/linker will be required.

vii) A last consideration is the geometry or the topology of the fusion protein. Should the tethering moiety be placed N or C terminal to the protein of interest? In the MS2 coat protein dimer the N and C termini of each monomer are on opposite sides of the binding complex (see Figure 2B) but spatially relatively close to the complementary termini of the other monomer. Hence, it is unlikely that the geometry of the fusion protein affects the binding capacity of the MS2 coat protein moiety. However, the N or C terminal fusion geometry may affect the functionality of the fused protein under study. The lambda N peptide in binding as a monomer is orientated relative to the RNA stem loop. In the full length N-protein the RNA-binding peptide is situated at the N-terminus. Hence, placing the fused protein after the peptide (C-terminal fusion) will reproduce the "natural" situation.

Conclusions.

We have described here the specific characteristics of the two tethering system that have been the most used experimentally. It is clear that each of the tethering systems has certain advantages in particular situations. Therefore, in designing experiments using the tethering assay, we would recommend that these characteristics are fully taken into account within the constraints of the structural requirements for the protein to be fused with the tethering moiety. Furthermore, as shown by Ko and Gunderston (2002) by comparing the responses obtained with different binding stoichiometries it should be possible to deduce information about the oligomerisation of the studied protein.

Acknowledgements. The authors thank P. Osborne for the art work of figure 1. Work in the author's laboratory is supported by the Centre National de la Recherche Scientifique and the Ministère de la Recherche (ACI BCMS 369). C. K-B was supported by technology transfer grants from the Region Bretagne (PRIR 560-412) and the Rennes Metropole/CNRS.

References (* highlighted references).

Arber, W. (1983) A Beginner's Guide to Lambda Biology. In: *Lambda II* (eds R. W. Hendrix, J. W. Roberts, F. W. Stahl and R. A. Weisberg), pp. 381-394. Cold Spring Harbour: Cold Spring Harbour Laboratory.

Austin, R.J., Xia, T., Ren, J., Takahashi, T.T. and Roberts, R.W. (2002) Designed arginine-rich RNA-binding peptides with picomolar affinity. J. Am. Chem. Soc. **124**, 10966-10967.

Bardwell, V.J. and Wickens, M. (1990) Purification of RNA and RNA-protein complexes by an R17 coat protein affinity method. Nucleic Acids Res. **18**, 6587-6594.

Barik, S., Ghosh, B., Whalen, W., Lazinski, D. and Das, A. (1987) An antitermination protein engages the elongating transcription apparatus at a promoter-proximal recognition site. Cell **50**, 885-899.

*Baron-Benhamou, J., Gehring, N.H., Kulozik, A.E. and Hentze, M.W. (2004) Using the lambdaN peptide to tether proteins to RNAs. Methods Mol. Biol. **257**, 135-154.

Barreau, C., Paillard, L. and Osborne, H.B. (2005) AU-rich elements and associated factors: are there unifying principles? Nucleic Acids Res. **33**, 7138-7150.

Barreau, C., Watrin, T., Beverley Osborne, H. and Paillard, L. (2006) Protein expression is increased by a class III AU-rich element and tethered CUG-BP1. Biochem. Biophys. Res. Commun. **347**, 723-730.

Bernardi, A. and Spahr, P.F. (1972) Nucleotide sequence at the binding site for coat protein on RNA of bacteriophage R17. Proc. Natl. Acad. Sci. U S A **69**, 3033-3037.

*Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H. and Long, R.M. (1998) Localization of ASH1 mRNA particles in living yeast. Mol. Cell **2**, 437-445.

Carey, J., Cameron, V., de Haseth, P.L. and Uhlenbeck, O.C. (1983a) Sequence-specific interaction of R17 coat protein with its ribonucleic acid binding site. Biochemistry **22**, 2601-2610.

Carey, J., Lowary, P.T. and Uhlenbeck, O.C. (1983b) Interaction of R17 coat protein with synthetic variants of its ribonucleic acid binding site. Biochemistry **22**, 4723-4730.

Chattopadhyay, S., Garcia-Mena, J., DeVito, J., Wolska, K. and Das, A. (1995) Bipartite function of a small RNA hairpin in transcription antitermination in bacteriophage lambda. Proc. Natl. Acad. Sci. U S A **92**, 4061-4065.

Chou, C.F., Mulky, A., Maitra, S., Lin, W.J., Gherzi, R., Kappes, J. and Chen, C.Y. (2006) Tethering KSRP, a decay-promoting AU-rich element-binding protein, to mRNAs elicits mRNA decay. Mol. Cell. Biol. **26**, 3695-3706.

Cilley, C.D. and Williamson, J.R. (1997) Analysis of bacteriophage N protein and peptide binding to boxB RNA using polyacrylamide gel coelectrophoresis (PACE). RNA **3**, 57-67.

*Coller, J. and Wickens, M. (2002) Tethered function assays using 3' untranslated regions. Methods **26**, 142-150.

*Coller, J.M., Gray, N.K. and Wickens, M.P. (1998) mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation. Genes Dev. **12**, 3226-3235.

Collier, B., Gorgoni, B., Loveridge, C., Cooke, H.J. and Gray, N.K. (2005) The DAZL family proteins are PABP-binding proteins that regulate translation in germ cells. EMBO J. **24**, 2656-2666.

Conn, G.L. and Draper, D.E. (1998) RNA structure. Curr; Opin. Struct. Biol. **8**, 278-285.

*De Gregorio, E., Preiss, T. and Hentze, M.W. (1999) Translation driven by an eIF4G core domain in vivo. EMBO J. **18**, 4865-4874.

Dertinger, D., Dale, T. and Uhlenbeck, O.C. (2001) Modifying the specificity of an RNA backbone contact. J. Mol. Biol. **314**, 649-654.

Dickson, K.S., Thompson, S.R., Gray, N.K. and Wickens, M. (2001) Poly(A) polymerase and the regulation of cytoplasmic polyadenylation. J. Biol. Chem. **276**, 41810-41816.

Doelling, J.H. and Franklin, N.C. (1989) Effects of all single base substitutions in the loop of boxB on antitermination of transcription by bacteriophage lambda's N protein. Nucleic Acids Res. **17**, 5565-5577.

Expert-Bezancon, A., Sureau, A., Durosay, P., Salesse, R., Groeneveld, H., Lecaer, J.P. and Marie, J. (2004) hnRNP A1 and the SR proteins ASF/SF2 and SC35 have antagonistic functions in splicing of beta-tropomyosin exon 6B. J. Biol. Chem. **279**, 38249-38259.

Fiers, W. (1979) Structure and function of RNA bacteriophages. In: *Comprehensive Virology* (eds H. Fraenkel-Conrat and R. R. Wagner), pp. 69-204. New York: Plenum Publishing Corporation.

Franklin, N.C. (1985) "N" transcription antitermination proteins of bacteriophages lambda, phi 21 and P22. J. Mol. Biol. **181**, 85-91.

Franklin, N.C. (1993) Clustered arginine residues of bacteriophage lambda N protein are essential to antitermination of transcription, but their locale cannot compensate for boxB loop defects. J. Mol. Biol. **231**, 343-360.

Franks, T.M. and Lykke-Andersen, J. (2007) TTP and BRF proteins nucleate processing body formation to silence mRNAs with AU-rich elements. Genes Dev. **21**, 719-735.

Friedman, D. and Gottesman, M. (1983) Lytic Mode of Lambda Development. In: *Lambda II* (eds R. W. Hendrix, J. W. Roberts, F. W. Stahl and R. A. Weisberg). Cold Spring Harbour: Cold Spring Harbour Laboratory.

Gehring, N.H., Kunz, J.B., Neu-Yilik, G., Breit, S., Viegas, M.H., Hentze, M.W. and Kulozik, A.E. (2005) Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. Mol. Cell **20**, 65-75.

Gehring, N.H., Neu-Yilik, G., Schell, T., Hentze, M.W. and Kulozik, A.E. (2003) Y14 and hUpf3b form an NMD-activating complex. Mol. Cell **11**, 939-949.

Gorgoni, B., Andrews, S., Schaller, A., Schumperli, D., Gray, N.K. and Muller, B. (2005) The stem-loop binding protein stimulates histone translation at an early step in the initiation pathway. RNA **11**, 1030-1042.

Graveley, B.R., Hertel, K.J. and Maniatis, T. (1998) A systematic analysis of the factors that determine the strength of pre-mRNA splicing enhancers. EMBO J. **17**, 6747-6756.

Graveley, B.R. and Maniatis, T. (1998) Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. Mol Cell. **1**, 765-771.

Gray, N.K., Coller, J.M., Dickson, K.S. and Wickens, M. (2000) Multiple portions of poly(A)-binding protein stimulate translation in vivo. EMBO J. **19**, 4723-4733.

Grskovic, M., Hentze, M.W. and Gebauer, F. (2003) A co-repressor assembly nucleated by Sex-lethal in the 3'UTR mediates translational control of Drosophila msl-2 mRNA. EMBO J. **22**, 5571-5581.

Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y. and Yokoyama, S. (1999) Structural basis for recognition of the tra mRNA precursor by the Sex-lethal protein. Nature **398**, 579-585.

Horwitz, R.J., Li, J. and Greenblatt, J. (1987) An elongation control particle containing the N gene transcriptional antitermination protein of bacteriophage lambda. Cell **51**, 631-641.

Jansone, I., Berzin, V., Gribanov, V. and Gren, E.J. (1979) The regulatory region of MS2 phage RNA replicase cistron. III. Characterization of fragments resulting from S1 nuclease digestion. Nucleic Acids Res. **6**, 1747-1760.

Johansson, H.E., Liljas, L. and Uhlenbeck, O.C. (1997) RNA Recognition by the MS2 Phage Coat Protein. Seminars in Virology **8**.

Kertesz, S., Kerenyi, Z., Merai, Z., Bartos, I., Palfy, T., Barta, E. and Silhavy, D. (2006) Both introns and long 3'-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. Nucleic Acids Res. **34**, 6147-6157.

*Ko, B. and Gunderson, S.I. (2002) Identification of new poly(A) polymerase-inhibitory proteins capable of regulating pre-mRNA polyadenylation. J. Mol. Biol. **318**, 1189-1206.

Kunz, J.B., Neu-Yilik, G., Hentze, M.W., Kulozik, A.E. and Gehring, N.H. (2006) Functions of hUpf3a and hUpf3b in nonsense-mediated mRNA decay and translation. RNA **12**, 1015-1022.

Larralde, O., Smith, R.W., Wilkie, G.S., Malik, P., Gray, N.K. and Clements, J.B. (2006) Direct stimulation of translation by the multifunctional herpesvirus ICP27 protein. J. Virol. **80**, 1588-1591.

Lazinski, D., Grzadzielska, E. and Das, A. (1989) Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. Cell **59**, 207-218.

LeCuyer, K.A., Behlen, L.S. and Uhlenbeck, O.C. (1995) Mutants of the bacteriophage MS2 coat protein that alter its cooperative binding to RNA. Biochemistry **34**, 10600-10606.

Legault, P., Li, J., Mogridge, J., Kay, L.E. and Greenblatt, J. (1998) NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. Cell **93**, 289-299.

Lim, F. and Peabody, D.S. (1994) Mutations that increase the affinity of a translational repressor for RNA. Nucleic Acids Res. **22**, 3748-3752.

Liu, J., Rivas, F.V., Wohlschlegel, J., Yates, J.R., 3rd, Parker, R. and Hannon, G.J. (2005) A role for the P-body component GW182 in microRNA function. Nat. Cell Biol. **7**, 1261-1266.

Long, R.M., Gu, W., Lorimer, E., Singer, R.H. and Chartrand, P. (2000) She2p is a novel RNA-binding protein that recruits the Myo4p-She3p complex to ASH1 mRNA. EMBO J. **19**, 6592-6601.

Lowary, P.T. and Uhlenbeck, O.C. (1987) An RNA mutation that increases the affinity of an RNA-protein interaction. Nucleic Acids Res. **15**, 10483-10493.

Lykke-Andersen, J., Shu, M.D. and Steitz, J.A. (2000) Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. Cell **103**, 1121-1131.

Lykke-Andersen, J., Shu, M.D. and Steitz, J.A. (2001) Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. Science **293**, 1836-1839.

McDonald, D., Hope, T.J. and Parslow, T.G. (1992) Posttranscriptional regulation by the human immunodeficiency virus type 1 Rev and human T-cell leukemia virus type I Rex proteins through a heterologous RNA binding site. J. Virol. **66**, 7232-7238.

Minshall, N. and Standart, N. (2004) The active form of Xp54 RNA helicase in translational repression is an RNA-mediated oligomer. Nucleic Acids Res. **32**, 1325-1334.

Minshall, N., Thom, G. and Standart, N. (2001) A conserved role of a DEAD box helicase in mRNA masking. RNA **7**, 1728-1742.

Ni, C.Z., Syed, R., Kodandapani, R., Wickersham, J., Peabody, D.S. and Ely, K.R. (1995) Crystal structure of the MS2 coat protein dimer: implications for RNA binding and virus assembly. Structure **3**, 255-263.

Pantopoulos, K. (2004) Iron metabolism and the IRE/IRP regulatory system: an update. Ann. N. Y. Acad. Sci. **1012**, 1-13.

Peabody, D.S. (1997) Subunit fusion confers tolerance to peptide insertions in a virus coat protein. Arch. Biochem. Biophys. **347**, 85-92.

Peabody, D.S. and Ely, K.R. (1992) Control of translational repression by protein-protein interactions. Nucleic Acids Res. **20**, 1649-1655.

Peabody, D.S. and Lim, F. (1996) Complementation of RNA binding site mutations in MS2 coat protein heterodimers. Nucleic Acids Res. **24**, 2352-2359.

Piganeau, N., Schauer, U.E. and Schroeder, R. (2006) A yeast RNA-hybrid system for the detection of RNA-RNA interactions in vivo. RNA **12**, 177-184.

Pillai, R.S., Artus, C.G. and Filipowicz, W. (2004) Tethering of human Ago proteins to mRNA mimics the miRNA-mediated repression of protein synthesis. RNA **10**, 1518-1525.

Romaniuk, P.J., Lowary, P., Wu, H.N., Stormo, G. and Uhlenbeck, O.C. (1987) RNA binding site of R17 coat protein. Biochemistry **26**, 1563-1568.

Ruiz-Echevarria, M.J. and Peltz, S.W. (2000) The RNA binding protein Pub1 modulates the stability of transcripts containing upstream open reading frames. Cell **101**, 741-751.

Sanford, J.R., Gray, N.K., Beckmann, K. and Caceres, J.F. (2004) A novel role for shuttling SR proteins in mRNA translation. Genes Dev. **18**, 755-768.

SenGupta, D., Zhang, B., Kraemer, B., Pochart, P., Fields, S. and Wickens, M. (1996) A three-hybrid system to detect RNA-protein interactions in vivo. Proc. Natl. Acad. Sci. U. S. A. **93**, 8496-8501.

Stripecke, R., Oliveira, C.C., McCarthy, J.E. and Hentze, M.W. (1994) Proteins binding to 5' untranslated region sites: a general mechanism for translational regulation of mRNAs in human and yeast cells. Mol. Cell. Biol. **14**, 5898-5909.

Su, L., Radek, J.T., Hallenga, K., Hermanto, P., Chan, G., Labeots, L.A. and Weiss, M.A. (1997a) RNA recognition by a bent alpha-helix regulates transcriptional antitermination in phage lambda. Biochemistry **36**, 12722-12732.

Su, L., Radek, J.T., Labeots, L.A., Hallenga, K., Hermanto, P., Chen, H., Nakagawa, S., Zhao, M., Kates, S. and Weiss, M.A. (1997b) An RNA enhancer in a phage transcriptional antitermination complex functions as a structural switch. Genes Dev. **11**, 2214-2226.

Tan, R. and Frankel, A.D. (1995) Structural variety of arginine-rich RNA-binding peptides. Proc. Natl. Acad. Sci. U. S. A. **92**, 5282-5286.

Tanaka, K.J., Ogawa, K., Takagi, M., Imamoto, N., Matsumoto, K. and Tsujimoto, M. (2006) RAP55, a cytoplasmic mRNP component, represses translation in Xenopus oocytes. J. Biol. Chem. **281**, 40096-40106.

Valegard, K., Murray, J.B., Stonehouse, N.J., van den Worm, S., Stockley, P.G. and Liljas, L. (1997) The three-dimensional structures of two complexes between recombinant MS2 capsids

and RNA operator fragments reveal sequence-specific protein-RNA interactions. J. Mol. Biol. **270**, 724-738.

Van Gilst, M.R., Rees, W.A., Das, A. and von Hippel, P.H. (1997) Complexes of N antitermination protein of phage lambda with specific and nonspecific RNA target sites on the nascent transcript. Biochemistry **36**, 1514-1524.

Vasudevan, S. and Steitz, J.A. (2007) AU-rich-element-mediated upregulation of translation by FXR1 and Argonaute 2. Cell **128**, 1105-1118.

Weiss, M.A. and Narayana, N. (1998) RNA recognition by arginine-rich peptide motifs. Biopolymers **48**, 167-180.

Wilkie, G.S., Gautier, P., Lawson, D. and Gray, N.K. (2005) Embryonic poly(A)-binding protein stimulates translation in germ cells. Mol. Cell. Biol. **25**, 2060-2071.

Witherell, G.W., Wu, H.N. and Uhlenbeck, O.C. (1990) Cooperative binding of R17 coat protein to RNA. Biochemistry **29**, 11051-11057.

Wu, H.N., Kastelic, K.A. and Uhlenbeck, O.C. (1988) A comparison of two phage coat protein-RNA interactions. Nucleic Acids Res. **16**, 5055-5066.

Key issue	*Effector Protein*	*Tethering protein*		*Tether sequence*		*Nb of sites*	*Organism*	*References*
		form	*position*	*form*	*position*			
Analysis of protein function	Rev, Rex	Full length	**C**	wt t	Equivalent intronic sequence	2,4	CV1 cells, COS7 cells	(McDonald et al. 1992)
	SR proteins	ns	N	C-loop	downstream splicing site	1,2	Hela cells nuclear extract	(Graveley et al. 1998)
	Pub1p	ns	N	ns	Open reading frame	2	Yeast	(Ruiz-Echevarria and Peltz 2000)
	Upf proteins	V75E,A81G	C or N	ns	3' UTR	2, 4, 6	Hela cells	(Lykke-Andersen et al. 2000)
	PABP	ns	N	ns	3' UTR	2	Xenopus oocytes, yeast	(Gray et al. 2000)
	RNPS1	V75E,A81G	C	ns	3' UTR	6	HEK293	(Lykke-Andersen et al. 2001)
	PAP	ns	N	ns	3' UTR	2	Xenopus oocytes	(Dickson et al. 2001)
	Xp54	ns	N	ns	3'UTR	2	Xenopus oocytes and eggs	(Minshall and Standart 2004)
	hnRNPA1	ns	C	C-loop	Intron	2	Hela cells nuclear extract	(Expert-Bezancon et al. 2004)
	SF2/ASF	ns	N	ns	3' UTR	2	Xenopus oocytes	(Sanford et al. 2004)
	ePABP	ns	N	ns	3' UTR	2	Xenopus oocytes	(Wilkie et al. 2005)
	DAZL	ns	N	ns	3'UTR	1, 3, 9	Xenopus oocyte	(Collier et al. 2005)
	CUG-BP1	V75E,A81G	C	C-loop	3' UTR	2, 4	NIH 3T3, HeLa cells	(Barreau et al. 2006)
	TTP, BRF	V75E,A81G	N	ns	3'UTR	6	Hela cells	(Franks and Lykke-Andersen 2007)
Dissection of multiple functions	PAB	single or tandem	N	C-loop or low affinity mutant	3' UTR	2	Yeast	(Coller et al. 1998)
	Xp54	ns	N	ns	3' UTR	2	Xenopus oocytes	(Minshall et al. 2001)
	SR proteins PRD domains	Tandem	N	C loop	3' UTR	1,2	Hela cells	(Ko and Gunderson 2002)
	RAP55	ns	Ns	Ns	3'UTR	Ns	Xenopus oocytes	(Tanaka et al. 2006)
	ICP27	ns	N	ns	3' UTR	2	Xenopus oocytes	(Larralde et al. 2006)
	KSRP	V75E,A81G	N	ns	3' UTR	6	Hela cells	(Chou et al. 2006)
Dissection of complexes	UIA	ns	ns	C-loop	5' end	2	Xenopus oocytes	(Bardwell and Wickens 1990)
	She3p	ns	ns	C-loop	3' UTR just after stop codon	6	Saccharomyces cerevisiae	(Long et al. 2000)
Protein-RNA interactions	LexA	ns	C	C-loop	3' end	2	Yeast	(SenGupta et al. 1996)
	SLBP	ns	N	ns	3'UTR	1, 3	Xenopus oocyte	(Gorgoni et al. 2005)
RNA-RNA interactions	LexA	ns	C	C-loop	ns	2	Yeast	(Piganeau et al. 2006)
mRNA localisation	GFP	deltaFG	N	C-loop	3' UTR just after stop codon	6	Yeast	(Bertrand et al. 1998)

Table 1: Overview of some applications of the MS2 coat protein/MS2 tethering system:
The V75E, A81G and dlFG forms of the MS2 coat protein are non-aggregating mutants. In the tandem form two monomers are translated from the same mRNA. The C-loop form of the MS2 tether sequence is a high affinity form (see Figure 2 and text). ns: non specified.

Key issue	*Effector Protein*	*Tether position*	*Nb of sites*	*Organism*	*References*
	dSXL	3'UTR and/or 5'UTR	1, 2	Drosophila embryos extracts	(Grskovic et al. 2003)
Analysis of protein function	hAGO2	3'UTR	1, 2, 3, 5	HeLa	(Pillai et al. 2004)
	Ago1, Ago2	3'UTR	5	HeLa cells	(Liu et al. 2005)
	UPF1	3' UTR or 5' UTR	5	Nicotinia tabacum or benthamiana	(Kertesz et al. 2006)
	FRX1, AGO2	3'UTR	5	HEK 293	(Vasudevan and Steitz 2007)
Dissection of multiple functions	eIF4G central domain	inter-cistronic region	1	HeLa cells	(De Gregorio et al. 1999)
	hUpf3a,b	3' UTR or just before stop codon	4, 5	HeLa cells	(Kunz et al. 2006)
Dissection of complexes	hUpf3b, Y14	3' UTR	2, 3, 5	HeLa cells	(Gehring et al. 2003)

Table 2: Overview of some applications of the LambdaN/B box tethering system:
The tethering protein is always the Lambda N peptide 1-22 positioned at the N-terminus of the fusion protein. The tethering sequence is the lambda nutL form of the B box (see Figure 3).

Figure Legends.

Figure 1: Schematic illustration of different situations that may be observed with tethered RNA-binding proteins.

The RNA molecule is shown as a linear structure, the orange box corresponds to the open reading frame. The 5' extremity is to the left. The natural binding site (black cross) and the tethering site (green vertical bar) are placed arbitrarily in the 3' UTR. The RNA-binding protein is represented by a flower connected by a stem to the roots that contain the RNA-binding domain. The yellow star represents the tethering protein.

Case A, the simplest situation where the association of the RNA-binding protein alone is sufficient to provoke the regulatory function.

Case B, a second protein, shown here as a bee, is required to make the functional complex. If this second protein alone can provoke the regulatory function then its direct tethering will be sufficient (case B-2). Also removing the stem and roots required for binding to the normal RNA element will still allow a functional complex to form (case B-3). However, no functional complex will form if the geometry of the bound protein is incorrect (case B-4).

Case C, the association of the RNA-binding protein with the RNA causes a structural change that is depicted here as a change in colour of the flower's centre. The tethering assay, by circumventing the association between the protein and the RNA, does not reproduce the structural change and hence does not produce a functional complex although all the elements are present.

Figure 2: MS2 RNA stem-loop and MS2-coat protein.

A. The wild type (left) and high affinity C-loop (right) forms of the MS2 stem loop are shown. The numbering is from the first base of the AUG which is the initiation codon of the replicase mRNA. B. Diagram of the association between two MS2 coat protein monomers (red and grey). The β-sheets are lettered A to G and the N and C termini of each monomer is indicated (derived from (Ni et al. 1995).

C. The sequence of the MS2 coat protein, the numbering is relative to the processed protein from which the initial methionine is removed. The amino acids structured in β-sheets are boxed and the corresponding sheet (A to G) indicated in bold below. The helical segments are double underlined and labelled H1 and H2. The amino acids involved in RNA-binding in either the upper or lower sub-units of the dimer are indicated by * or + respectively.

Figure 3: The B box RNA and lambda N-peptide.

A. The lambda nut L and nut R sequences.

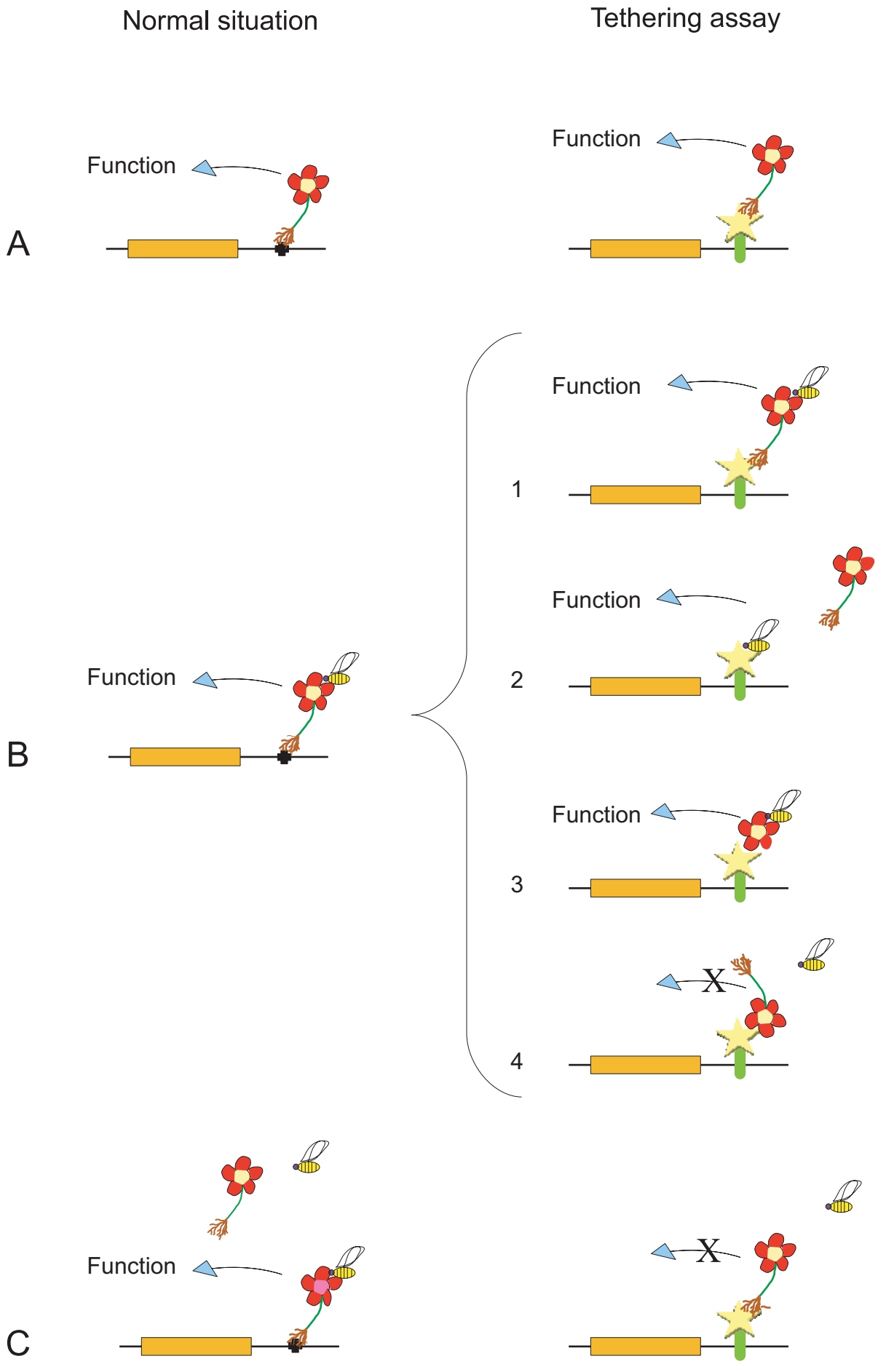
B. Sequence of the 1-22 peptide of the lambda N protein.

C. Representation of the association between the nut L stem loop and the lambda N (1-22) peptide. The RNA is shown as a dotted surface, the phosphate backbone is represented by spheres for C3', C5', O3', O5', O1P, O2P and P atoms. The C^{α} peptide backbone is shown as a smoothed ribbon (reproduced from (Legault et al. 1998).

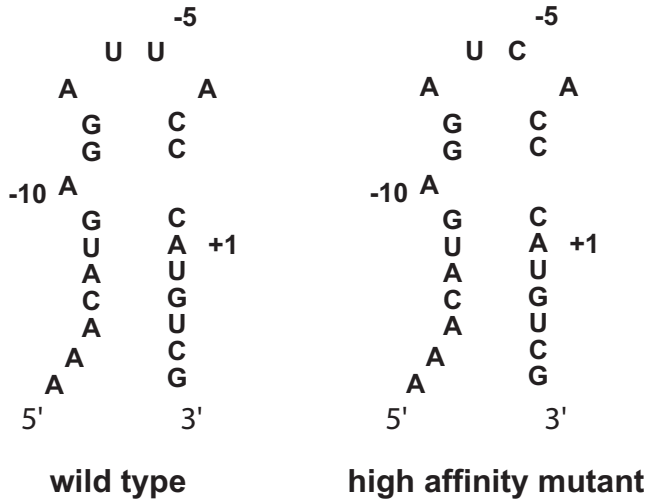
Figure 4. Effect of the number of MS2 motifs on luciferase mRNA stability in HeLa cells.

A. Two sets of HeLa cells (8×10^4 for each transfection) were cultured overnight and then independently co-transfected using the calcium phosphate procedure with 500 ng of plasmid encoding the MS2 coat protein and 500 ng of plasmid encoding the Renilla and Firefly luciferases under the control of a bi-directional CytoMegalo Virus promoter (see (Barreau et al. 2006). The 3'UTR of the Renilla mRNA contained the indicated number of MS2 binding sites. 28 hours later total RNA was extracted and analysed by Northern blot using a Renilla specific probe (top panel) and then a Firefly specific probe (bottom panel). After washing the blots were revealed using a Storm 860 Phophoimageur. NTF, not transfected.

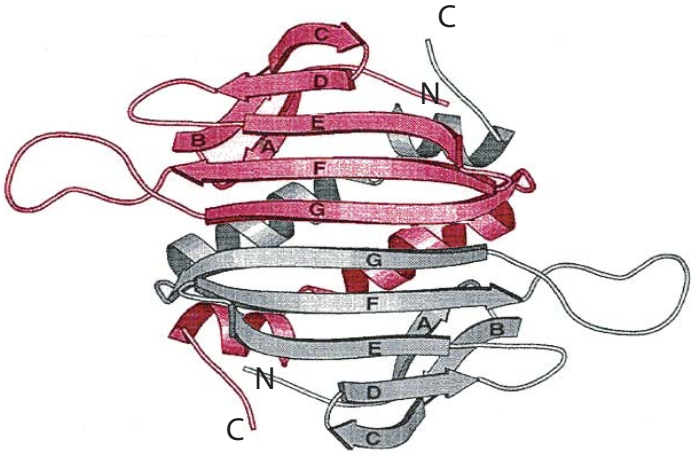
B. Ratio of the Renilla luciferase to Firefly luciferase mRNAs (average ± standard deviation) calculated from the data shown in panel A using the ImageQuant software. The ratio measured when the Renilla luciferase mRNA had no MS2 element (0 MS2) was set to 100%.


Normal situation
Tethering assay
A
Function
Function
B
Function
1
Function
2
Function
3
X
4
C
Function
X

A MS2 stem loop sequence



B Diagram of MS2 coat protein dimer

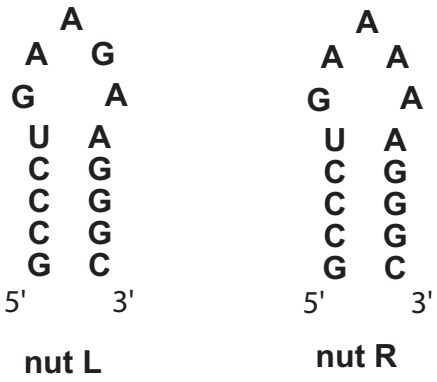


C MS2 coat protein sequence

```
                                                        *
1    (m)asnf tqfvlvdn ggtg dvtvapsnfan gvaewiss
                                                        +
                    A              B     C          D

                  *  *  *  *            *  *  *  *
36   nsrs qaykvtcsvrqss a qnrkytikvevp
                 + + + + +    + + + +
                 E              F

                            * *
66   kvatqtvggvelpvaaw rsylnmeltip ifat
                                 G

98   nsdcelivkamqgll kdgn pipsaiaa nsgiy
           H1                  H2
```

A

box B sequences



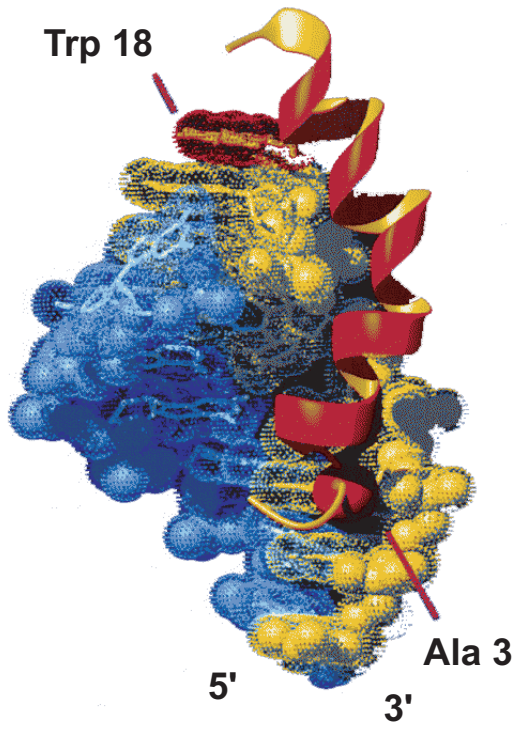
B

Lambda N-(1-22) peptide

MDAQTRRRERRAEKQAQWKAAN

1 22

C



Keryer-Bibens et al. **Figure 4**