

# Purification of RNA and RNA-protein complexes by an R17 coat protein affinity method

Vivian J. Bardwell<sup>1</sup> and Marvin Wickens<sup>1,2\*</sup>

<sup>1</sup>Cell and Molecular Biology Program and <sup>2</sup>Department of Biochemistry, University of Wisconsin – Madison, Madison, WI 53706, USA

Received July 23, 1990; Accepted September 28, 1990

## ABSTRACT

We describe an affinity chromatography method to isolate specific RNAs and RNA-protein complexes formed *in vivo* or *in vitro*. It exploits the highly selective binding of the coat protein of bacteriophage R17 to a short hairpin in its genomic RNA. RNA containing that hairpin binds to coat protein that has been covalently bound to a solid support. Bound RNA-protein complexes can be eluted with excess R17 recognition sites. Using purified RNA, we demonstrate that binding to immobilized coat protein is highly specific and enables one to separate an RNA of interest from a large excess of other RNAs in a single step. Surprisingly, binding of an RNA containing non-R17 sequences to the support requires two recognition sites in tandem; a single site is insufficient. We determine optimal conditions for purification of specific RNAs by comparing specific binding (retention of RNAs with recognition sites) to non-specific binding (retention of RNAs without recognition sites) over a range of experimental conditions. These results suggest that binding of immobilized coat protein to RNAs containing two sites is cooperative. We illustrate the potential utility of the approach in purifying RNA-protein complexes by demonstrating that a U1 snRNP formed *in vivo* on an RNA containing tandem recognition sites is selectively retained by the coat protein support.

## INTRODUCTION

Complexes of RNA and proteins are critical for diverse aspects of RNA metabolism, including splicing, translational regulation and ribosome assembly. Isolation of specific RNA-protein complexes is an important initial step toward understanding their function.

Affinity purification methods exist for the isolation of RNA-protein complexes formed *in vitro* (1–4). In one such method, RNAs are transcribed *in vitro* in the presence of a low concentration of biotinylated UTP. These RNAs can then be purified by their affinity for streptavidin attached to agarose beads (2,3). In another method, a short RNA (4), or 2'-O-methyl RNA (1), containing a covalently attached biotin, is hybridized to the RNA of interest. The hybrid, together with any attached proteins or other factors, can then be purified using streptavidin beads (1,4).

In this report we describe an alternative affinity purification method that exploits the high affinity of the coat protein of the *Escherichia coli* bacteriophage R17 for a short (21 nucleotide) hairpin in its genomic RNA (5–7). In this method, R17 coat protein is covalently attached to beads. The RNA to be isolated contains the protein's cognate recognition sequence and is therefore selectively retained by the beads. This general method may prove useful in the analysis of a variety of RNA-protein interactions, whether the RNA has been transcribed *in vitro* or *in vivo*.

## MATERIALS AND METHODS

### Preparation of R17 Phage

Preparation of R17 coliphage was performed essentially as described, with minor modifications (8,9). *Escherichia coli* strain Q13 (*rna-19, his-95, tyrA6, relA1, pnp-13, spoT1, metB1, E. coli* Genetic Stock Center strain # 4947) was grown in 2 to 6 liters of LB supplemented with calcium and glucose (1% tryptone-0.5% yeast extract-0.5% NaCl-0.4% glucose-2mM CaCl<sub>2</sub>-pH 7.5) to an absorbance (600nm) of 0.4 to 0.5 (0.8 to 1.0 × 10<sup>8</sup> cells per ml). CaCl<sub>2</sub> and glucose were added to the media just before use. The cells were then infected with R17 phage at a multiplicity of infection of 3 to 10. Phage, in 1/50 volume of LB, were added to the cells, swirled slowly for 10 min, and incubated an additional 10 min before resuming normal shaking. Cell lysis was detectable 3 hours post-infection and was complete after 5 hours. After lysis, enough chloroform was added to achieve a final concentration of 1%. Shaking was continued for 15 min. Cellular debris was removed by centrifugation in a Sorvall GSA rotor at 8,000 rpm, at 5°C for 10 min. The concentration of the phage in the supernatant typically was 10<sup>12</sup> plaque forming units per ml.

The following protocol indicates the amounts needed to prepare phage from 1 liter of supernatant. 58g NaCl and 100g polyethylene glycol (PEG 8000) were added to one liter of phage solution. After stirring at 4°C to dissolve the NaCl and PEG, the solution was stored at 4°C for at least 90 min. Precipitated phage were collected by centrifugation at 8,000 rpm at 5°C for 15 min in a GSA rotor (Sorvall). The phage pellet was resuspended in 10 ml of 10 mM Tris-HCl, pH 8.0. Phage were then reprecipitated with 0.58g NaCl and 1g PEG and transferred

\* To whom correspondence should be addressed

to 30 ml Corex tubes. To recover phage, the solution was centrifuged at 8,000 rpm at 5°C for 15 min in a SS-34 rotor (Sorvall). The phage pellet was resuspended in 11.0 ml 10 mM Tris-HCl, pH 8.0, plus 6.5 g optical grade CsCl (Bethesda Research Laboratories). Phage were separated from residual cellular debris by buoyant density equilibrium centrifugation at 25,500 rpm at 5°C for 36 to 60 hours, in an SW40Ti rotor (Beckmann). Each SW40Ti tube can contain material from a four liter culture. The opalescent phage band, identified by visual inspection, was collected with a syringe inserted 0.3 cm below the band and dialyzed against 1 L of 10 mM Tris-HCl (pH 8.0) per ml of phage, at 4°C for 3 hours with one buffer change. Phage were then stored at 4°C. The extinction coefficient of a 1 mg per ml solution of purified phage is 8.03 absorbance units at 260 nm (8). As judged by absorbance, 8 to 13 mg of phage (generally in approximately 1 ml) are recovered from 1 liter of cleared lysate. This value is often considerably higher than that calculated from the number of plaque forming units per ml, presumably due to the presence of defective phage particles.

For large scale preparation of phage (e.g. 30 liters in a fermenter), phage were concentrated from supernatant by PEG precipitation as above. Alternatively, if facilities are available, phage in the supernatant can be concentrated by filtration through a Pellicon Cassette (100,000 molecular weight exclusion filter; Millipore). Phage were then further purified as described above.

#### Preparation of R17 Coat Protein from Purified Phage

Coat protein was prepared essentially as described (9). To disrupt the phage particle and precipitate the RNA, ten ml glacial acetic acid was mixed with 5 mls of phage at approximately 10 mg per ml (as judged by  $A_{260}$ ). The mixture was kept on ice for one hour, with vortexing for approximately 30 secs every 10 min. RNA was removed by centrifugation for 15 min at 4°C at 15,600×g. The supernatant, containing the phage coat protein, was dialyzed against 10 liters of distilled water (1 liter of water per 0.5 ml phage) at 4°C for 3 hours resulting in a final concentration of approximately 20 mM acetic acid. After dialysis, the coat protein solution was centrifuged briefly at 4°C to remove particulate matter. Protein concentration was determined by absorbance ( $E_{280} = 1.54 \times 10^4 M^{-1}$  (10)) and a molecular weight of 13,700 g/mole (11). Typically, from 50 mg (5 ml) of CsCl-banded phage, 30 ml of coat protein at 1 to 2 mg per ml was recovered. R17 coat protein was stored at 4°, at this acidic pH, because it begins to aggregate above pH 5 with concentrations greater than  $10^{-6}M$  (5).

#### Preparation of R17 Coat Protein Beads

Coat protein was coupled to the activated support, Affigel-10 (Biorad), via the N-hydroxysuccinimide ester of the resin. For every mg of coat protein, 100 mg of Affigel-10 beads were used. The beads were washed with 20 volumes of ice-cold deionized water by vacuum filtration, avoiding complete dryness at any point. Washed beads were weighed, added to the protein and incubated on a rotary mixer at 4°C for 24 hours. Coupling is more efficient at a pH close to the pI of the protein, 8.65 (11), but because coat protein aggregates under those conditions, we chose to use the lower pH (approximately 3.5) of the coat protein-acetic acid solution. After coupling, the beads were allowed to settle. The supernatant, containing unreacted coat protein, was removed. The beads were extensively washed with 3 mM acetic acid as follows: twice quickly with 5 volumes, overnight with 10 volumes on a rotary mixer, at 4°C, and twice more with 5

volumes. A small amount of coat protein precipitated out of solution during the coupling. Therefore the beads were allowed to settle out of solution the first two times to avoid collecting the aggregated protein with the beads. After that the beads can be recovered by centrifugation (2000×g). The coupled coat protein beads were stored as a 50% slurry in 3 mM acetic acid at 4°C and retained binding activity for at least 4 months. Approximately 1 mg of protein was coupled per ml of beads. The unreacted coat protein can be stored at 4 °C and re-used in a second coupling.

#### RNA Structures, Sequences and Nomenclature

**Structures and nomenclature.** The general structures of the RNAs used in this study are diagrammed in Figure 1A. To simplify nomenclature, RNAs are named by the sequences they contain, in a 5' to 3' order. The R17 site is designated R and the SV40 late polyadenylation region is designated S. Thus R/S RNA contains a single R17 recognition site followed by the SV40 sequence, while 2R RNA contains two R17 recognition sites and no adjacent sequence.

**Sequence of R17 recognition sites.** The R17 recognition sites used in this report contain a transition mutation in the loop of the hairpin (sequence in Fig. 1B). This mutation (a U to C change) increases the affinity of coat protein for its site, and greatly stabilizes the RNA-coat protein complex (7). In most experiments, two tandem R17 recognition sites are used. The sequence of the two tandem sites is presented in Fig. 1B, drawn in the form of two RNA stem-loop structures.

#### Preparation of Plasmids and RNAs

**A single R17 recognition site.** Two complementary DNA oligonucleotides, carrying the R17 recognition site, flanked by *Xba* I and *Pst* I recognition sites were annealed and inserted into mp18. This generated mp18R1.

**Single R17 recognition site preceded by the SV40 late polyadenylation region (S/R RNA).** The *Hind*III site at position +70 of pSPSV-141/+70 (12) was changed to a *Xba*I site by insertion of a *Xba*I linker. This allowed the -141/+70 region of the SV40 late polyadenylation site to be transferred into mp18R1, using a *Bam*HI-*Xba*I fragment of -141/+70 (*Bam*HI at -141, *Xba*I at +70) and the *Bam*HI and *Xba*I sites of mp18R1. This generated mp18SVR1. mp18SVR1 was cut with *Bam*HI and *Pst*I and this SV40-R17 fragment was transferred into pSP65 (an SP6 transcription vector) to create pSVR1. S RNA (-141/+55) was prepared by transcription of *Dra*I cut pSVR1 (Fig. 1). S/R RNA can be prepared by transcription of *Pst*I cut pSVR1.

**Single R17 recognition site followed by the SV40 late polyadenylation region (R RNA and R/S RNA).** pSVR1 was cut with *Eco*RI and *Xba*I to remove the SV40 sequences. The ends of the remaining DNA were filled in and the plasmid religated to generate pVB531. SV40 sequences were reinserted after the R17 recognition site by using a *Pst*I-*Hind*III fragment (-141/+70) from pSPSV-141/+70 and the *Pst*I-*Hind*III sites of pVB531. This generated pVB532a. R RNA and R/S RNA were prepared by transcription of *Pst*I and *Dra*I cut pVB532a respectively (Fig. 1).

**Tandem R17 recognition sites (2R RNA).** The R17 recognition site-containing *Xba* I fragment from pVB531 was isolated and ligated to itself. The ligated products were separated on a native

polyacrylamide gel, isolated, ligated into mp19 and sequenced. pVB535 is a direct repeat two R17 recognition sites in mp19.

**Tandem R17 recognition sites followed by the SV40 late polyadenylation site (2R/S RNA)** To generate the appropriate clone, designated pVB536, a three way ligation was performed, involving a *Hind*III-*Bam*HI fragment (two R17 recognition sites) from pVB535, a *Bam*HI-*Eco*RI fragment (-141/+70 SV40 polyadenylation sequences) from pSPSV-141/+70, and *Hind*III-*Eco*RI cut pSP64. 2R RNA and 2R/S RNA were prepared by transcription of *Bam*HI and *Dra*I cut pVB536 respectively (Fig. 1).

**Tandem R17 recognition sites at position +27 of the human U1 snRNA (2R/U1 RNA).** A tandem R17 recognition site flanked by *Bam*HI sites was generated by insertion of a *Bam*HI linker at the *Hind*III site 5' of the R17 recognition sites in pVB536. This tandem R17 recognition site *Bam*HI fragment was transferred into *Bcl*I-cut ND101, generating pVB544a. ND101 contains the human U1 gene (13,14). The *Bcl*I site into which the tandem R17 sequences were inserted, is at +27 relative to the transcription start at +1 (13,14) Upon injection into frog oocytes, pVB544a directs the synthesis of a U1 snRNA containing two R17 sites inserted at position +27.

#### Transcription *in vitro*

RNAs were prepared by run-off transcription using SP6 polymerase (15) in the presence of 1 mM diguanosine triphosphate and 0.025 to 3 mCi per ml [<sup>32</sup>P]UTP or 50 nCi per ml [<sup>3</sup>H]UTP.

#### Chemical Synthesis of a Single R17 Site

A 22 nucleotide RNA was containing the high affinity variant of the R17 recognition site (7) was synthesized chemically, and was a generous gift of Angus Lamond and Brian Sproat (EMBL, Heidelberg). Full length RNA was purified by gel electrophoresis. This RNA was used as the eluting agent in Figure 6.

#### Injection of Oocytes and Preparation of Oocyte Extract

Individual oocyte nuclei were injected with approximately 5ng of DNA mixed with 1 μCi [<sup>32</sup>P]GTP, and incubated at 20°C for 19 hours. To prepare deproteinized RNA from oocytes, oocytes were homogenized in Homogenization Buffer (50 mM Tris-HCl pH 7.9-5 mM EDTA-2% SDS-0.3M NaCl) in a 1 ml ground glass homogenizer, using at least 25 μl per oocyte. The homogenate was extracted with phenol/chloroform (1/1) and precipitated with ethanol. RNA was redissolved in 2 μl of water per oocyte. To prepare RNA-protein complexes from oocytes (Fig. 10B), oocytes were homogenized in Buffer D (20% [vol/vol] glycerol-20 mM *N*-2-hydroxyethylpiperazine-*N*-2-ethane-sulfonic acid [HEPES]-KOH [pH 7.9]-0.125mM EDTA-100mM KCl-0.5 mM DTT) containing 20 mM vanadyl-ribonuclease complexes (Bethesda Research Labs) plus 5mM DTT. 25 to 40 μl of homogenization buffer was used per oocyte. Generally, 10 to 20 oocytes were homogenized together. Yolk was removed by centrifugation at 15,600×g for 10 min 5°C. The supernatant was used without further purification.

#### Fractionation of snRNPs by Cesium Chloride Density Gradient Centrifugation

The procedure used for CsCl density gradient centrifugation was a modification of a previously described method (16,17). Oocyte extract (0.5 ml), prepared in Buffer D as described above, was

adjusted to 15 mM MgCl<sub>2</sub> by the addition of 1 M MgCl<sub>2</sub>. Solid CsCl (Optical Grade, Bethesda Research Labs) was then added to a final density of 1.6 g/ml. 0.5 ml Buffer D containing 0.5 mM DTT, 15mM MgCl<sub>2</sub> and CsCl to a density of 1.3 g/ml was underlayered with 0.5 ml of this extract and centrifuged in a TLA-100.2 rotor (Beckmann) at 100,000rpm for 4 hours at 4°C. The gradient was fractionated, from the top, into 100μl fractions. The density of each fraction was determined by weighing 50 μl samples. Fractions were dialyzed into Buffer D containing 15 mM MgCl<sub>2</sub>. 20 μl samples were extracted with phenol/chloroform (1/1) and analyzed by electrophoresis through a 6% polyacrylamide gel. 10 μl samples were used for affinity purification.

#### Binding of RNA to Coat Protein Beads

Binding, washing and quantitation were all performed in microfuge tubes. In a standard binding reaction, 10 μl of beads (20 μl of slurry) were incubated with 50 μl of 10mg/ml heparin in TMK Buffer (100 mM Tris-HCl [pH 7.8]-80 mM KCl-10 mM MgAcetate (5)) for 5 min at room temperature in a silanized 1.5 ml microfuge tube. 1.5 ml of TMK was added, the solution vortexed briefly and spun at 15,600×g for 15 seconds to pellet the beads. The buffer was removed to the level of the beads with a drawn out glass micropipette. 30 μl of TMK containing 25 μg of heparin and 5 to 20 fmol of RNA was added to the beads. The mixture was vortexed on an automatic vortexer (VWR Vortexer 2, maximum speed) for 25 min at room temperature. The beads were washed 3 or 4 times with 1.5 ml TMK spinning 15 seconds each time to pellet the beads. To determine what fraction of the RNA had bound to the beads, the radioactivity associated with the beads in the microfuge tube was quantitated using Cerenkov radiation.

For some experiments (Fig. 3, 4 and 7), binding, washing and elution was done in mini-columns formed in yellow pipette tips plugged with silanized glass wool. Binding in microfuge tubes and columns was comparable, although nonspecific binding to yellow pipette tips varied, and sometimes was high. To isolate RNA from a large volume we found it preferable to vortex beads with the solution rather than to pass it over a column of beads multiple times.

#### Elution of RNA from Coat Protein Beads

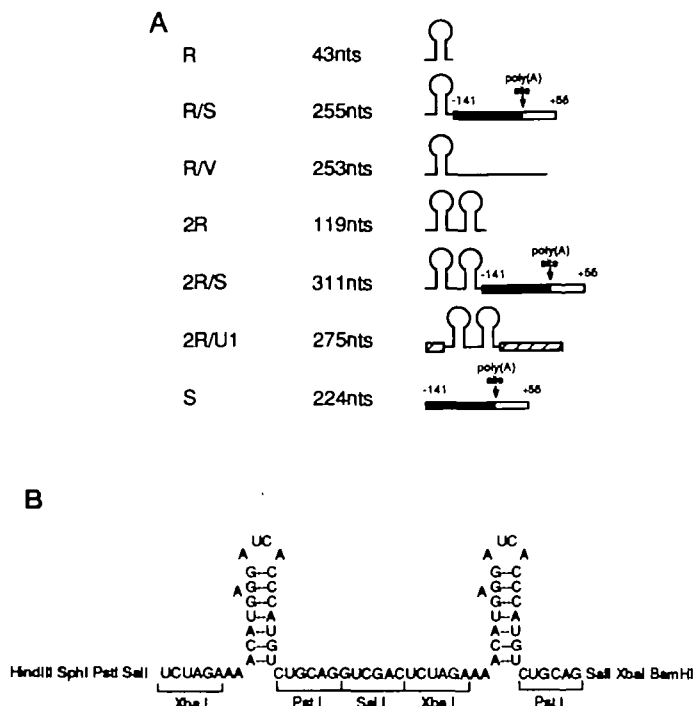
To test elution conditions, 100μl of the elution solution to be tested was added to the beads. This mixture was then incubated for 30 min. The beads were collected by centrifugation then washed once with the same solution. To determine the fraction of the radioactive RNA that had eluted, the radioactivity remaining associated with the beads in the microfuge tube was quantified using Cerenkov radiation. To test elution using excess recognition sites (Fig. 6A), a chemically synthesized single recognition oligonucleotide (see above) in 15 μl TMK, was added to the beads. Incubation was continued for 20 min. Beads were then washed with 500 μl TMK.

To elute RNA from minicolumns, the elution solution was added to the column and incubated for 5 to 30 min. The column was then washed with 100 to 2000 μl elution solution.

## RESULTS

### General Strategy

We describe here an affinity method that permits the isolation of specific RNAs and RNA-protein complexes. The general approach, diagrammed in Figure 2, hinges on the specific



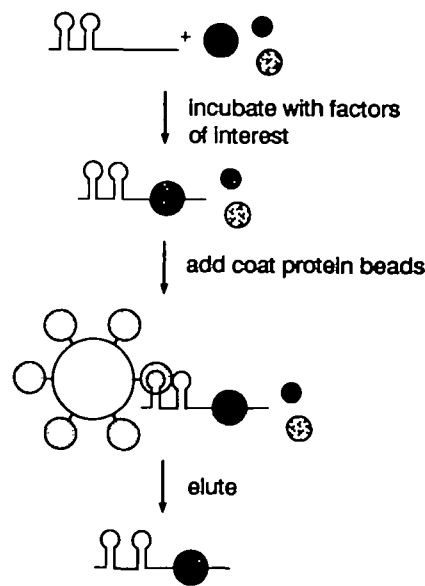
**Figure 1.** DNA double strand scission induced by murine topoisomerase II in SV40 DNA. Upper panel: cleavage in the presence of 0.5  $\mu$ M doxorubicin; lower panel: cleavage in the absence of drug (The scale ordinate is amplified 6-fold relative to the upper panel). Densitometer scanning curves of DNA cleavage patterns analyzed by means of agarose gels are corrected to give a linear measure of cleavage frequency per base pair. Stars indicate the labeled 5'-end termini and arrows the sequenced fragments (see Materials and Methods); ORI, replication origin; MAR, major nuclear-matrix associated region (7, 12).

interaction between the R17 coat protein and a high affinity variant ( $K_a = 3.5 \times 10^{10}$ ; ref. 7) of its cognate recognition site (Fig. 1B; ref. 7). A chimeric RNA containing two R17 recognition sites and the RNA sequence of interest is prepared either *in vitro*, by transcription with a phage polymerase, or *in vivo*, by cellular transcription of a transfected or injected DNA template. The chimeric RNA binds to appropriate factors in the cell or extract. The resulting RNA-factor complexes then can be selectively retained on a support to which R17 coat protein has been covalently coupled. RNAs that lack recognition sites are not retained. To recover the specific RNA molecules and any associated factors, the beads are treated with either an excess quantity of R17 recognition sites, with SDS, or with high concentrations of salt.

### Specific Retention: Two Sites are Better than One

To test the efficacy of the method, we analyzed the binding of six different RNAs to coat protein beads. These RNAs contained either 0, 1 or 2 recognition sites, and, in some cases, additional non-R17 sequences. Their structures are diagrammed in Figs. 1 and 3. Approximately equal amounts of radioactivity of each of the six labeled RNAs were mixed together and incubated with coat protein beads. RNAs retained by the beads were eluted with SDS. To determine which RNAs had bound to the beads and which had not, we analyzed 'bound' and 'flow through' RNAs by gel electrophoresis. The results are presented in Fig. 3.

Whereas RNA containing a single R17 recognition site and no additional sequences binds efficiently, RNA lacking



**Figure 2.** Outline of the method. Purification of a hypothetical specific RNA-protein complex from a mixture of proteins. See text for details.

recognition sites does not. Similarly, a short RNA containing only two tandem sites, and no additional sequences, binds to the beads. We conclude that coat protein, after attachment to beads, still binds specifically to its cognate site, and that, with RNAs that do not contain any non-R17 sequences, a single site and two sites in tandem are retained with comparable efficiency.

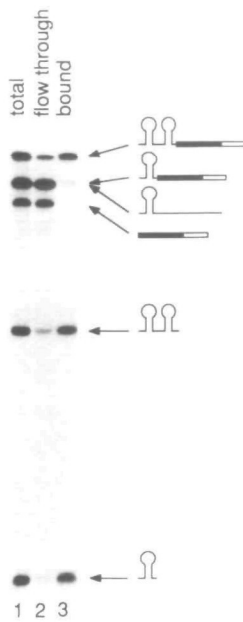
In striking contrast, two sites are necessary for efficient retention of RNAs that contain non-R17 sequences. This is demonstrated by comparing the RNAs that contain the sequence spanning the polyadenylation site of SV40 late mRNAs. SV40 RNA with just a single R17 recognition site binds very poorly, but an identical RNA with two recognition sites binds efficiently.

These data (Fig. 3) lead to two conclusions. First, the presence of non-R17 sequences interferes with binding to a single recognition site. The inhibitory effect of adjacent RNA sequences is not specific for the SV40 sequence, as it is observed with sequences derived from a procaryotic vector (Fig.3), and with sequences derived from U1 snRNA (data not shown). The second conclusion is that the inhibitory effect of additional sequences is relieved by inserting a second R17 recognition site next to the first. An RNA of this type, bearing two tandem sites, is retained with an efficiency nearly that of a single site in isolation.

Based on recent studies of coat protein binding in solution (18), we suspect that the stimulation of binding by the presence of a second site reflects cooperative binding of coat protein (see Discussion). Whatever the explanation, the data in Fig. 3 establish an important technical point: For retention of 'long' RNAs by immobilized coat protein, two recognition sites are better than one. For this reason, in all subsequent experiments, we used RNAs containing two adjacent sites.

### RNA Transcribed *In Vivo* Binds Coat Protein

To determine whether RNA transcribed *in vivo* would bind to coat protein beads, even in the presence of a vast excess of cellular RNA, we performed the following experiment. We constructed a DNA containing two R17 sites inserted into the human U1

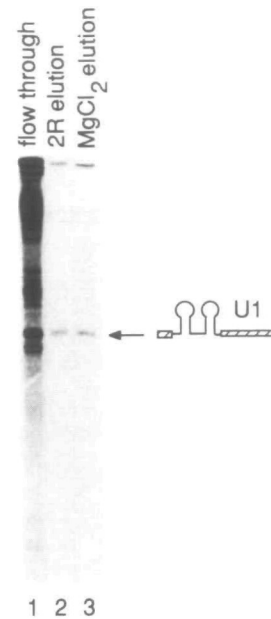


**Figure 3.** Specific retention: two sites are better than one. Approximately equal amounts of radioactivity of each of six different RNAs plus 50  $\mu\text{g}$  of heparin were mixed in 10  $\mu\text{l}$  TMK buffer, applied to a 10  $\mu\text{l}$  coat protein bead column and incubated for 25 min. The first 200  $\mu\text{l}$  of TMK wash was collected as the 'flow through', the column further washed with 800  $\mu\text{l}$  TMK and bound material was eluted with 200  $\mu\text{l}$  1% SDS. RNA was precipitated and analyzed by electrophoresis on a 6% denaturing polyacrylamide gel. Lane 1, a mixture of the six RNAs not applied to the column; lane 2, RNAs which flowed through the column; lane 3, RNAs which bound to the coat protein beads and then were eluted. Structures of the RNAs are indicated on the right. R/S and R/V RNAs co-migrate on this gel.

snRNA gene, at position +27 relative to the transcription start at +1. This insertion should not disrupt the U1 promoter (14). The chimeric U1/R17 gene was injected into nuclei of *Xenopus* oocytes, together with alpha  $^{32}\text{P}$  labeled GTP. After 19 hours, RNA was prepared by phenol/chloroform extraction. The deproteinized RNA was incubated with coat protein beads and the 'bound' and 'flow through' fractions were analyzed by denaturing polyacrylamide gel electrophoresis. Only the U1 RNA containing two R17 recognition sites was retained (Fig. 4). All other labeled RNAs were not retained even though they were much more abundant than the U1/R17 species. Furthermore, the oocyte RNA preparation contains a 1000-fold excess of unlabeled RNAs relative to the labeled species, yet specific retention is observed. We conclude that binding to the coat protein beads is highly specific even in the presence of a large quantity of non-target RNA.

### Optimal Binding Conditions

Conditions for optimal binding in solution of coat protein to a single recognition site have been reported previously (5,6). In the following series of experiments (Fig. 5), we optimized conditions for the binding of immobilized coat protein to an RNA containing two recognition sites. To do so, we compared the binding of two RNAs to the coat protein matrix. One RNA contains two recognition sites followed by the SV40 late polyadenylation sequences (2R/S). The retention of this RNA reflects specific binding. The other RNA contains only SV40 sequences (S), and reflects non-specific binding. The results demonstrate that, over a broad range of conditions, binding is highly specific and selective. The breadth of the optima contrasts



**Figure 4.** RNA transcribed *in vivo* binds coat protein. Deproteinized RNA from one oocyte, in 10  $\mu\text{l}$  TMK, was applied to a 10  $\mu\text{l}$  coat protein column and incubated for 25 min. The first 200  $\mu\text{l}$  of buffer was collected as the 'flow through'. After further washing RNA was specifically eluted for 30 min with 10 pmol of 2R RNA in 10  $\mu\text{l}$  TMK. RNA that still remained bound was eluted with 0.4M  $\text{MgCl}_2$ . RNA was prepared and analyzed as in Fig. 3. Lane 1, RNAs which flowed through the column; lane 2, RNAs eluted with 5 pmol 2R RNA for 30 min; lane 3, remaining RNAs eluted with 0.4 M  $\text{MgCl}_2$ . The position of 2R/U1 RNA is indicated on the right. In this experiment the column was pre-washed with heparin but heparin was not included in the binding reaction.

dramatically with the restricted conditions in which coat protein binds to a single site in solution (ref. 5; see Discussion).

### Time and temperature of incubation (Fig. 5A)

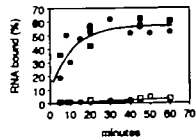
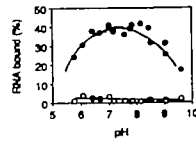
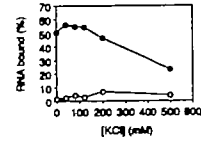
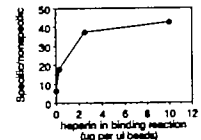
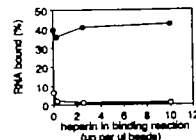
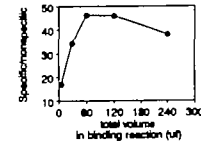
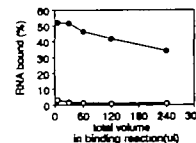
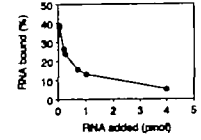
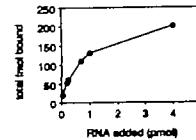
Specific binding increased between zero and 30 min. At 30 min, approximately 55% of the RNA containing recognition sites had bound (Fig. 5A, black symbols). At that same time, less than 2% of the RNA lacking the sites was bound (Fig. 5A, open symbols). These data corroborate those in Fig. 3 and 4. The extent of specific binding at room temperature (circles) and at 4°C (squares) was comparable.

### pH and salt concentration (Figs. 5B and 5C)

The binding reaction exhibited a broad pH optimum of 6.5 to 8.5 at room temperature (Fig. 5B) and was not significantly affected by salt concentrations between 0 and 200 mM KCl (Fig. 5C).

### Nonspecific competitor (Fig. 5D)

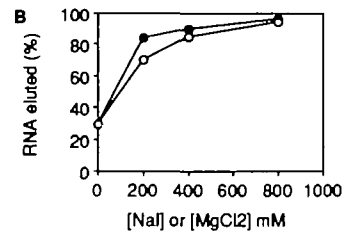
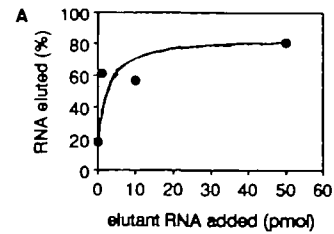
Agarose beads, Eppendorf tubes and coat protein have a small but finite capacity to bind RNA nonspecifically. In order to minimize the background due to this nonspecific binding, beads and tubes were treated with the polyanion, heparin, before incubation with RNA. Heparin was also included in the binding reaction. To determine the optimal heparin concentration, increasing amounts of heparin were added to a fixed amount of beads during the binding reaction. Addition of heparin at 2.5  $\mu\text{g}$  per  $\mu\text{l}$  of beads was sufficient to decrease non-specific binding from 7% (with no heparin added) to less than 1% (Fig. 5D; left panel). This increased the ratio of specific to nonspecific binding from 6 to 40 (Fig. 5D; right panel).

**A. Time and temperature****B. pH****C. KCl concentration****D. Concentration of non-specific competitor****E. Volume of incubation****F. Capacity of beads**

**Figure 5. Optimal binding conditions.** (A) Time and temperature. The percentage of 2R/S RNA (black symbols) and S RNA (open symbols) that bound to coat protein beads was determined after different times of incubation. Circles represent incubation at 25°C and squares represent incubation at 4°C. The results of three time courses are combined in this figure. (B) pH. The percentage of 2R/S RNA (filled circles) and S RNA (open circles) that bound to coat protein beads at the pH indicated was determined. 100 mM K-MES was used below pH 7 and 100 mM Tris-HCl was used at pH 7 and above. (C) Potassium chloride concentration. The percentage of 2R/S RNA (filled circles) and S RNA (open circles) that bound to coat protein at various [KCl] was determined. The concentration of magnesium acetate was kept constant, at 10 mM. (D) Concentration of non-specific competitor (heparin). Left panel: The percentage of 2R/S RNA (filled circles) and S RNA (open circles) that bound to coat protein beads in the presence of the indicated amounts of heparin (x-axis) was determined. Right panel: The data are re-plotted to indicate, on the y-axis, the ratio of specific (2R/S) to non-specific (S) binding. (E) Volume of incubation. Left panel: The percentage of 2R/S RNA (filled circles) and S RNA (open circles) that bound to coat protein beads was determined at various volumes of incubation. The amounts of RNA (10 fmol) and of beads (10 µl) were kept constant. Right panel: The data are re-plotted to indicate, on the y-axis, the ratio of specific (2R/S) to non-specific (S) binding. (F) Capacity of beads. Left panel: The amount of 2R/S RNA (filled circles) that bound to coat protein was determined using various concentrations of RNA. Right panel: The data are re-plotted to indicate, on the y-axis, the percentage of the input RNA that bound to the beads.

**Volume of binding (Fig. 5E)**

An advantage of affinity purification is the ability to isolate a specific molecule from even a relatively large volume using a tractably small quantity of beads. We therefore tested the extent of specific binding in increasing volumes of incubation, keeping



**Figure 6. Elution of RNA that has bound to the beads.** (A) Elution with excess recognition sites. 2R/S RNA was bound to coat protein beads under standard conditions. RNA that had bound was eluted by incubating the beads in the presence of various concentrations of unlabeled R RNA for 20 mins on ice. The percentage of bound RNA that was eluted is indicated on the y-axis. (B) Elution with high salt concentrations. 2R/S RNA was bound to coat protein beads under standard conditions. RNA that had bound was eluted by incubating the beads in the presence of various concentrations of MgCl<sub>2</sub> (closed circles) or NaI (open circles). The percentage of the bound RNA that was eluted is indicated on the y-axis.

the amounts of RNA and beads constant (Fig. 5E). Increasing the volume to 10 times that of the beads (10 µl of beads in 100 µl total volume) had little effect on specific binding, but reduced nonspecific binding significantly (Fig. 5E, left panel). As a result, the ratio of specific to non-specific binding was increased by increasing the volume (Fig. 5E, right panel).

**Capacity of beads (Fig. 5F)**

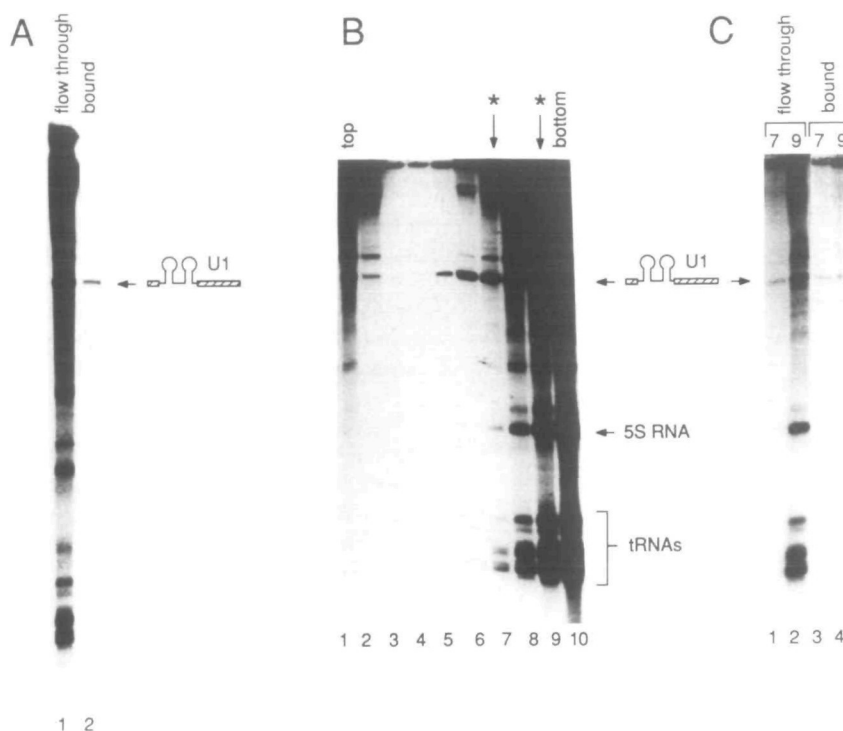
To test the capacity of the coat protein beads for RNA, a fixed quantity of beads (10 µl) was incubated with increasing amounts of RNA containing recognition sites. Ten µl of beads bound up to 200 fmol of RNA without apparently reaching saturation (Fig. 5F, left panel). However, with increasing amounts of added RNA, a progressively smaller fraction of the RNA was retained (Fig. 5F, right panel).

**Elution Strategies**

To recover the bound RNA and associated proteins, we reasoned that an excess of recognition sites should effectively 'displace' RNAs already bound to the immobilized coat protein. Figure 6A demonstrates that this is the case. Labeled 2R/S RNA was bound to immobilized coat protein under standard conditions. To elute, the beads were incubated for 20 minutes with various amounts of a single R17 recognition site RNA, synthesized chemically (a generous gift of A. Lamond and B. Sproat, EMBL, Heidelberg). With 1 pmol of elutant RNA, which represents approximately a 100-fold excess over the amount of RNA specifically bound to the matrix, 60% of the labeled RNA was eluted (Fig. 6A; see also Fig. 4, lane 2). Increasing the quantity of elutant RNA 50-fold only increased the extent of RNA released to 80%. We conclude that elution by excess recognition sites is effective. It may prove particularly useful in cases in which the biological activity of specific RNA-protein complexes is to be assayed after elution.

Three other, non-specific agents can also be used to elute RNA from the coat protein matrix. NaI or MgCl<sub>2</sub> each releases bound





**Figure 7.** Isolation of U1 snRNP containing an snRNA with two recognition sites from *Xenopus* oocytes. (A) *Xenopus* oocytes were injected with the plasmid carrying a 2R/S gene and [ $^{32}$ P] GTP. After 19 hours, a crude homogenate was prepared. 10  $\mu$ l of this crude homogenate, without any deproteinization, was incubated with coat protein beads (10  $\mu$ l) in a column for 30 min. The first 100  $\mu$ l of TMK wash was collected as the 'flow through'. After further washing, material that had remained bound was eluted with 1% SDS. Samples were extracted with phenol/chloroform and precipitated with ethanol. Lane 1, RNAs present in the fraction that flowed through the column; lane 2, RNAs present in the fraction that bound to the column. The position of 2R/U1 RNA is indicated on the right. (B) The distribution of labeled RNAs in a CsCl/MgCl<sub>2</sub> density gradient. Oocytes were injected and incubated as in (A). The crude *Xenopus* extract, without phenol extraction or ethanol precipitation, was centrifuged in a CsCl/MgCl<sub>2</sub> gradient as described in Materials and Methods. From each fraction of the gradient,  $^{32}$ P-labeled RNA was prepared by extraction with phenol/chloroform and precipitation with ethanol. Labeled RNAs were analyzed by electrophoresis through a 6% polyacrylamide gel containing 7M urea. The position of the 2R/U1 RNA is indicated. The density of the fractions ranged from 1.29 g/ml in fraction 1 to 1.52 g/ml in fraction 9. Fractions 7 and 9 (indicated with asterisks) were used in panel (C). (C) Isolation of U1 snRNP. 10  $\mu$ l of fractions 7 and 9 from the gradient shown in (B) were applied to 10  $\mu$ l coat protein columns and incubated for 30 min. The first 100  $\mu$ l of TMK buffer wash was collected as the 'flow through'. After further washing the snRNPs which remained bound were eluted with 0.4 M MgCl<sub>2</sub>. RNA was prepared as in (A). Lane 1, RNA extracted from fraction 7 material that flowed through the column; lane 2, RNA extracted from fraction 9 material that flowed through the column; lane 3, RNA extracted from fraction 7 material that bound to the coat protein beads; lane 4, RNA extracted from fraction 9 material that bound to the coat protein beads.

RNA (Fig. 6B), as does 0.1% SDS (not shown). Obviously these eluting agents are likely to dissociate any RNA-protein complexes and so disrupt biological activity. Nonetheless, they may be useful for the detection of a specific molecule for which a probe is already available.

#### Isolation of U1snRNP *Xenopus* Oocytes

To determine whether RNA-protein complexes formed *in vivo* could be isolated using the affinity technique, oocytes were injected with a plasmid carrying the 2R/U1 snRNA gene and alpha  $^{32}$ P GTP. In the oocyte, this gene directs the synthesis of human U1 snRNA containing two R17 recognition sites near its 5' end (see Fig. 4), which is assembled into a snRNP. 19 hours after injection, oocytes were homogenized. The crude homogenate, without further purification, was incubated with coat protein beads. RNA was prepared from the material that bound to the matrix, and from the material that had flowed through the column. These RNAs were then analyzed by electrophoresis (Fig. 7A). Approximately 25% of the total 2R/U1 present in the homogenate was retained by the matrix. All other labeled RNAs were not retained, even though they were present in considerable excess.

To distinguish whether the RNA that had bound to the beads did so as a snRNP or as 'naked' RNA, free of any associated proteins, we examined the binding of partially purified snRNP particles. In the presence of Mg<sup>++</sup>, snRNPs are stable at high ionic strength and so can be separated from free RNA by CsCl/MgCl<sub>2</sub> equilibrium centrifugation (16,17). A crude homogenate of oocytes that had been injected with the 2R/U1 gene and [ $^{32}$ P] GTP was fractionated in this manner. A portion of each fraction of the CsCl/MgCl<sub>2</sub> gradient was deproteinized and the RNA analyzed by electrophoresis. Figure 7B shows the profile of oocyte RNAs present in the CsCl density gradient. RNA not associated with protein pellets to the bottom of the centrifuge tube while snRNPs equilibrate at characteristic densities (17). 2R/U1 snRNP is the major labeled snRNP found in these extracts. After dialysis to remove the CsCl, fractions 7 and 9 (Fig. 7A) were incubated with coat protein beads. Fraction 7 contains snRNP particles (buoyant density 1.44) while fraction 9 contains RNAs with few proteins bound (buoyant density 1.52). In both fractions, the 2R/U1 snRNA species was retained while other RNAs were not (Fig. 7C). We conclude therefore that snRNPs formed on RNAs containing R17 recognition sites are retained on the coat protein beads.

## DISCUSSION

In this paper we describe an affinity purification method for the isolation of specific RNAs and RNA-protein complexes. It is based on the affinity of the RNA bacteriophage R17 coat protein for a short RNA sequence (5,7). We demonstrate that RNAs containing two recognition sequences are specifically retained by coat protein beads. RNA-protein complexes formed *in vivo* can also be isolated on the support, as illustrated by the purification of U1 snRNPs from a crude *Xenopus* oocyte extract. Both with naked RNA and with U1 snRNP, the specificity of retention is dramatic.

To retain RNAs that carry foreign sequences, two R17 recognition sites are necessary: Six different RNAs with a single recognition site, either at the 5' end, the 3' end or in the middle of additional RNA sequences, all bind coat protein very poorly either in solution or on beads (Fig. 3, and data not shown). Similar results have recently been obtained in solution (18). Yet only a single site is required for binding of coat protein to R17 genomic RNA (19), and isolated single sites, without any flanking sequences, bind efficiently (e.g., ref. 5). The inhibition by foreign flanking sequences may be due to the formation of alternative secondary structures that disrupt the recognition site, or to direct interference with binding by the foreign RNA.

The inhibitory effect of foreign flanking sequences is relieved by the insertion of a second site. Cooperative interactions between coat proteins bound to adjacent sites, as have recently been described in solution (18), could account for this effect. In solution, cooperative binding to adjacent sites occurs over a much broader range of conditions than does binding to a single site, presumably due to the different optima of the protein-protein and RNA-protein interactions (18). In our experiments, retention of RNA by immobilized coat protein occurs over a wide range of pH, monovalent cation concentration, and temperature, suggesting that cooperative binding is involved.

Two sites are necessary for binding to the coat protein matrix, but are not always sufficient. An RNA that contains two sites that are three nucleotides closer than in the RNA we have described here, and surrounded by different sequences, fails to bind (data not shown). Similarly, it has recently been reported that sequences between two sites can interfere with binding in solution, by forcing the RNA into an alternative secondary structure (18). Although we have observed only one case in which an RNA containing two sites failed to bind to coat protein beads, this exception emphasizes the importance of testing the binding of the RNA of interest as 'naked' RNA before embarking on the isolation of RNA-protein complexes.

The affinity method we have described is simple and the needed reagents are easy to prepare. Binding of RNA to the beads is rapid, efficient and highly selective. It may be possible to retain the biological activity in many factors of interest after purification by eluting the column with excess R17 recognition sites.

For any particular RNA-protein complex, the utility of the method we have described will be a function of that complex's abundance, and its stability in the presence of heparin and after dilution during washing of the support. Retention requires that assembly of the complex of interest not be prevented by the presence of two recognition sites, and that the two sites remain accessible after the complex has formed. Clearly, further studies will be required to establish whether the method will have general utility. In principle, however, the ability to isolate RNA-protein complexes formed *in vivo* is attractive. The studies presented here

invite exploration of the method as a means of purifying RNA-protein complexes from diverse biological systems.

## ACKNOWLEDGMENTS

We are especially grateful to Jim Dahlberg for suggesting the idea of using R17 coat protein for affinity purification. We thank Ray Gesteland and Paul Hines for advice on R17 infections, phage preparation and coat protein preparation, Olke Uhlenbeck for advice on R17 coat protein binding, and Phil Zamore and Barbara Ruskin for advice on snRNP purification. Pauline Stephenson and Steve Ogg helped with plasmid construction. We thank Mary Whitmer and Phil Johnson for help with the growth of R17 phage, Henry Neuman de Vegvar for the plasmid ND101, and Angus Lamond and Brian Sproat for chemical synthesis of a single recognition site. We thank all members of the Wickens laboratory past and present for helpful discussions and encouragement and Jim Becker and Laura Vanderploeg for technical assistance.

DNA oligonucleotide synthesis was performed by the University of Wisconsin Protein Sequence-DNA Synthesis Facility. V.B. was supported by a Natural Sciences and Engineering Research Council of Canada postgraduate fellowship. This work is supported by the Public Health Service research grant GM31892 and Research Career Development Award GM00521 from the National Institutes of Health to M.W.

## REFERENCES

1. Blencowe, B.J., Sproat, B.S., Ryder, U., Barabino, S. and Lamond, A. (1989) *Cell* **59**, 531–539.
2. Grabowski, P.J. and Sharp, P.A. (1986) *Science* **233**, 1294–1299.
3. Rouault, T.A., Hentze, M.W., Haile, D.J., Harford, J.B. and Klausner, R.D. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5768–5772.
4. Ruby, S.W. and Abelson, J. (1988) *Science* **242**, 1028–1035.
5. Carey, J., Cameron, V., de Haset, P.L. and Uhlenbeck, O.C. (1983) *Biochemistry* **22**, 2601–2610.
6. Carey, J. and Uhlenbeck, O.C. (1983) *Biochemistry* **22**, 2610–2615.
7. Lowary, P.T. and Uhlenbeck, O.C. (1987) *Nucleic Acids Res.* **15**, 10483–10493.
8. Kolakofsky, D. (1971) *Methods Mol. Biol.* **1**, 267–277.
9. Sugiyama, T., Herbert, R.R. and Hartman, K.A. (1967) *J. Mol. Biol.* **25**, 455–463.
10. Weber, K. and Konigsberg, W. (1975) In Zinder, N. (ed.), *RNA phages*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
11. Weber, K. (1967) *Biochemistry* **6**, 3144–3153.
12. Zarkower, D. and Wickens, M. (1987) *EMBO J.* **6**, 177–186.
13. Neuman de Vegvar, H.E. (1989) Ph.D. thesis. University of Wisconsin-Madison.
14. Neuman de Vegvar, H.E., Lund, E. and Dahlberg, J.E. (1986) *Cell* **47**, 259–266.
15. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) *Nucleic Acids Res.* **12**, 7035–7056.
16. Ruskin, B., Zamore, P.D. and Green, M.R. (1988) *Cell* **52**, 207–219.
17. Leyay-Taha, M., Reveillaud, I., Sri-Widada, J., Brunnel, C. and Jeanteur, P. (1986) *J. Mol. Biol.* **189**, 519–532.
18. Witherell, G.W., Wu, H.N. and Uhlenbeck, O.C. (in press)
19. Lodish, H.F. and Zinder, N.D. (1986) *J. Mol. Biol.* **19**, 333–348.