

# Combining SELEX and the yeast three-hybrid system for in vivo selection and classification of RNA aptamers

JULIAN KÖNIG,<sup>1,3</sup> CHRISTIAN JULIUS,<sup>1,3,4</sup> SEBASTIAN BAUMANN,<sup>1</sup> MATTHIAS HOMANN,<sup>2</sup>  
H. ULRICH GÖRINGER,<sup>2</sup> and MICHAEL FELDBRÜGGE<sup>1</sup>

<sup>1</sup>Department of Organismic Interactions, Max-Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Strasse, 35043 Marburg, Germany

<sup>2</sup>Genetics, Darmstadt University of Technology, 64287 Darmstadt, Germany

## ABSTRACT

Aptamers are small nucleic acid ligands that bind to their targets with specificity and high affinity. They are generated by a combinatorial technology, known as SELEX. This in vitro approach uses iterative cycles of enrichment and amplification to select binders from nucleic acid libraries of high complexity. Here we combine SELEX with the yeast three-hybrid system in order to select for RNA aptamers with in vivo binding activity. As a target molecule, we chose the RNA recognition motif-containing RNA-binding protein Rrm4 from the corn pathogen *Ustilago maydis*. Rrm4 is an ELAV-like protein containing three N-terminal RNA recognition motifs (RRMs). It has been implicated in microtubule-dependent RNA transport during pathogenic development. After 11 SELEX cycles, four aptamer classes were identified. These sequences were further screened for their in vivo binding activity applying the yeast three-hybrid system. Of the initial aptamer classes only members of two classes were capable of binding in vivo. Testing representatives of both classes against Rrm4 variants mutated in one of the three RRM domains revealed that these aptamers interacted with the third RRM. Thus, the yeast three-hybrid system is a useful extension to the SELEX protocol for the identification and characterization of aptamers with in vivo binding activity.

**Keywords:** SELEX; three-hybrid system; RNA recognition motif; RNA/protein interaction; RNA aptamer

## INTRODUCTION

Aptamers are small DNA or RNA oligonucleotides that are selected to interact with a variety of different targets such as nucleic acids, polypeptides, sugar molecules, small organic compounds, as well as entire cells (<http://aptamer.icmb.utexas.edu/>; Homann and Göringer 1999; Göringer et al. 2003; Lee et al. 2004; Proske et al. 2005; Göringer et al. 2006; Ulrich 2006). Key usages are diagnostic and therapeutic applications where target proteins involved in human diseases are detected and inhibited by aptamers (Sullenger and Gilboa 2002; Nimjee et al. 2005). In addition, aptamers can be used to purify proteins or to investigate the sequence specificity of nucleic acid-binding

proteins. The latter is important for the investigation of RNA-binding proteins, e.g., for the identification of potential target sites within a given genome in order to characterize the corresponding protein/RNA linkage map (Murphy et al. 2003; Ulrich 2006).

RNA aptamers are identified by the SELEX technology (Ellington and Szostak 1990; Tuerk and Gold 1990). SELEX uses iterative cycles of target binding and sequence enrichment to select for high-affinity ligands from combinatorial nucleic acid libraries with complexities of up to  $10^{15}$  different molecules. Due to the iterative character of SELEX, even extremely rare sequences can be evolved from the initial library. However, for in vivo applications, the selected aptamers have to be evaluated on an individual basis.

The yeast three-hybrid system is a powerful technique to investigate interactions between RNA and proteins in vivo (SenGupta et al. 1996, 1999; Hook et al. 2005). It allows the identification of RNA and protein binding partners as well as the dissection of higher-order RNA/protein complexes (Bernstein et al. 2002). The system is based on the yeast two-hybrid system and consists of three components: the first hybrid contains the DNA-binding domain of LexA from *Escherichia coli* fused to the RNA-binding coat protein

<sup>3</sup>These authors contributed equally to this work.

<sup>4</sup>Present address: Institute of Neuropathology, University Hospital Zurich, Schmelzbergstrasse 12, 8091 Zurich, Switzerland.

Reprint requests to: Michael Feldbrügge, Department of Organismic Interactions, Max-Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Strasse, 35043 Marburg, Germany; e-mail: feldbrue@mpi-marburg.mpg.de; fax: +49-6421-178609.

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.334307>.

of phage MS2. The second hybrid is an RNA molecule that consists of the MS2-binding site and the RNA sequence of interest. The third hybrid is composed of the transcriptional activation domain of Gal4 from *Saccharomyces cerevisiae* and a given RNA-binding domain that interacts with the RNA of interest. The auxotrophic marker *HIS3* and *lacZ* (encoding imidazole-glycerol-phosphate dehydratase from *S. cerevisiae* and  $\beta$ -galactosidase from *E. coli*) are used as reporter genes. Their promoter regions contain cognate *lexA* operator sequences that are recognized by the first hybrid (Bernstein et al. 2002). The strength of this approach is the detection of RNA/protein interactions in vivo. In addition, it offers a rapid and simple experimental routine for detecting reporter gene expression. It has been used for in vivo selection of RNA-binding sites (Bernstein et al. 2005), but the number of transformants, typically  $10^7$ , limits the exploration of larger sequence space.

As a target molecule for the selection of aptamers we chose Rrm4 involved in pathogenic development of *Ustilago maydis*, the causative agent of corn smut disease (Becht et al. 2005). This RRM protein is essential for polar growth of the infectious hypha. Loss of Rrm4 abolishes formation of retraction septa at the distal pole and results in an increased number of bipolar growing cells (Becht et al. 2006). The protein is a component of particles that move along microtubules and is most likely involved in RNA transport from the nucleus to the cell poles. Rrm4 contains three N-terminal RRMs (RRM1–3) with characteristic spacing of embryonic lethal abnormal vision (ELAV)-type proteins (Soller and White 2004). A mutational analysis revealed that RRM1 is necessary for the determination of polarity during filamentation (Becht et al. 2006).

Here we followed the strategy to combine SELEX with the three-hybrid system (Cassiday and Maher 2001, 2003) to select RNA aptamers that interact with Rrm4 in vivo. Four aptamer classes were identified after 11 cycles of SELEX. Sequences from this cycle were tested for protein interaction in a yeast three-hybrid screen. Two out of four aptamer classes revealed in vivo binding to Rrm4 and members of both classes were shown to interact with the third RRM.

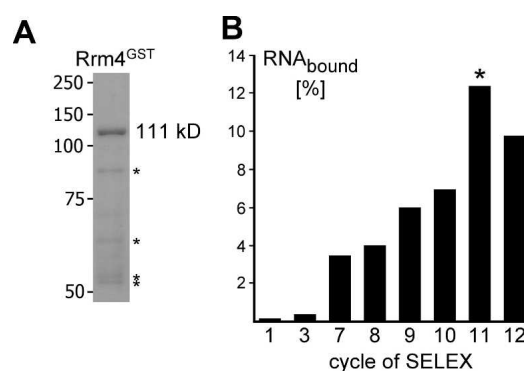
## RESULTS

### In vitro selection of Rrm4-binding RNA aptamers

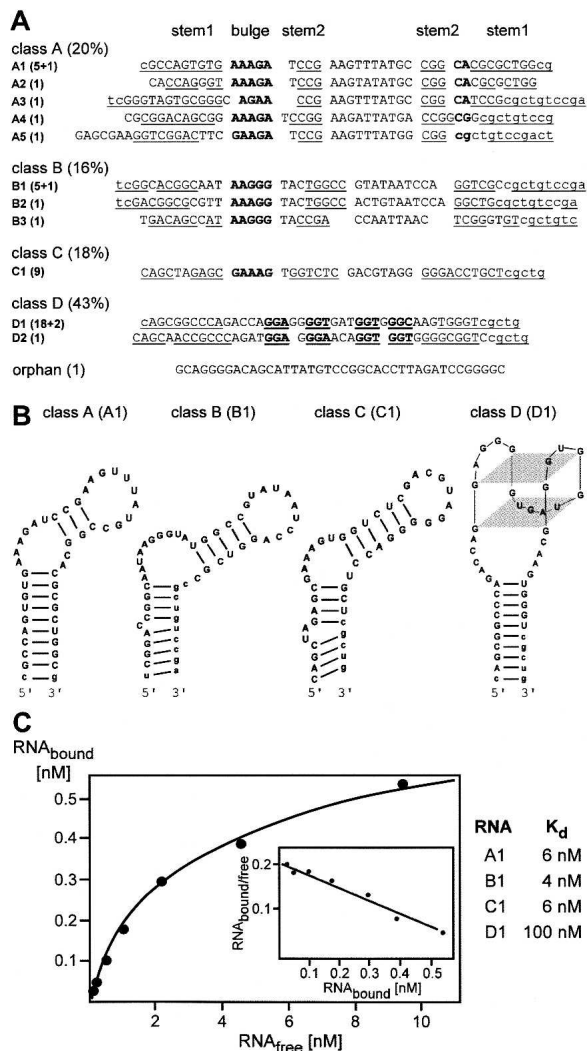
For in vitro selection we used a library of RNA molecules with an estimated complexity of  $3 \times 10^{14}$  unique sequences. The library contained a central domain of 40 randomized nucleotides (nt), flanked by primer binding sites of 21 and 24 nt. As a binding target we chose the Rrm4 protein from *U. maydis*. The polypeptide was expressed in *E. coli* as a C-terminal fusion protein with GST (111 kDa) and purified to near homogeneity by affinity chromatography (Rrm4<sup>GST</sup>; estimated grade of purity  $\sim 85\%$ )

(Fig. 1A). Twelve cycles of selection and amplification were performed. Bound RNA was separated from unbound material by filter adsorption. To deplete aptamers that bind to filter material or the GST domain we performed counterselections in cycles 3–4 and 7–12 (see Materials and Methods). The enrichment of Rrm4-specific binders was monitored by determining the percentage of bound RNA in each cycle. An increase was observed that reached a plateau in cycle 11 (Fig. 1B). The RNA pools from cycles 7–12 were analyzed for binding to filter material and Bruno<sup>GST</sup>. Bruno, an ELAV-type protein from *Drosophila melanogaster* (Kim-Ha et al. 1995), was expressed as GST fusion protein and purified under identical conditions. In both cases binding never exceeded 2%, indicating that the observed increase was due to the selection of aptamers specific for the Rrm4 part of Rrm4<sup>GST</sup>.

DNA templates from cycle 11 were cloned and 49 clones were sequenced. In total, 12 different sequences were identified (Fig. 2). These sequences were grouped into four classes and an orphan sequence on the basis of primary sequence and secondary structure predictions (Fig. 2A). Five members of class A (20% of the total sequences) exhibit high sequence identity to each other and are predicted to fold into a hairpin structure with a 8–14-nt stem, a highly conserved 10-nt AU-rich loop sequence (AAGUUUAUGC) and an asymmetric purine-rich internal bulge (AAAGA/CA) (Fig. 2B). Class B contains three members (16%), which can be folded into structures similar to class A members with loop sizes from 9 to 11 nt and different AU-rich loop sequences (Fig. 2B). Class C contains a single sequence that was found nine times (18%). It can be folded into a hairpin loop similar to classes A and B (Fig. 2B). However, the loop size is only 8 nt and its sequence lacks AU stretches. The two class D members (21 clones, 43%) are characterized by complementary



**FIGURE 1.** Course of the SELEX experiment using recombinant Rrm4<sup>GST</sup> as a target. (A) SDS-PAGE analysis of recombinant Rrm4<sup>GST</sup> preparations after staining with Coomassie brilliant blue. The full-length fusion protein (111 kDa) and putative degradation products (asterisks) are indicated. (B) Graphical representation of the percentages of bound RNA for different cycles of SELEX. DNA from the 11th cycle (asterisk) was cloned and 49 clones were sequenced.



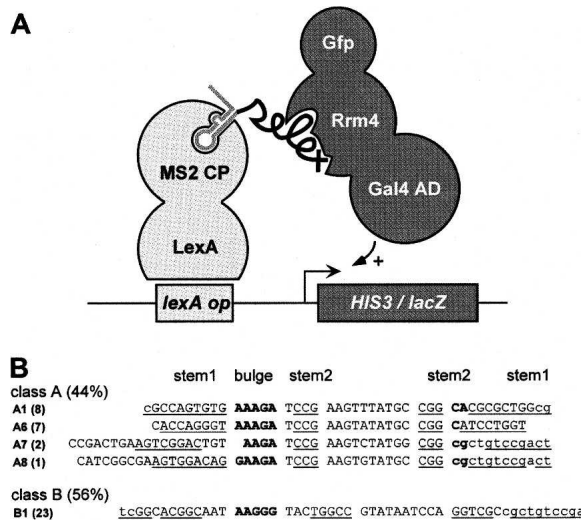
**FIGURE 2.** Classification of in vitro selected aptamer sequences. (A) Four aptamer classes were identified (A to D, relative abundances in parentheses). The designation of each aptamer is followed by its clone abundance in parentheses (+n, additional n sequences with single-point mutations) and the primary sequence of the randomized region (capital letters). Complementary sequences corresponding to stems 1 and 2 are underlined. Lower-case letters implicated in stem-loop formation indicate sequences derived from the constant primer binding sites. A purine-rich stretch in classes A–C, a conserved CA/G sequence in class A, and four GGN repeats in class D are printed in bold. (B) Secondary structure models for the most abundant member of classes A–D are shown. Models are based on structure predictions using Mfold and sequence similarities to known G-quartet structures (Zuker 2003; Homann et al. 2006). (C) Determination of equilibrium dissociation constants ( $K_d$ ). Representative binding isotherm for the binding of aptamer B1 to Rrm4<sup>GST</sup>. The inset shows a Scatchard plot of the binding data with the slope of the curve representing  $-1/K_d$ .

sequence stretches of 10–11 nt plus four repeats of a GGN triplet sequence. These sequence motifs are likely to fold into a G-quartet structure (Homann et al. 2006), located on top of a helical domain as shown in Figure 2B. Finally, a single orphan sequence (Fig. 2A) was identified, which showed no similarity to any of the other sequences.

Equilibrium dissociation constants ( $K_d$ ) for the interaction of the most abundant aptamers A1, B1, C1, and D1 with recombinant Rrm4<sup>GST</sup> were determined by filter adsorption. A representative binding isotherm for aptamer B1 with corresponding Scatchard analysis is shown in Figure 2C. Aptamers A1, B1, and C1 bind to Rrm4 with  $K_d$  values between 4 and 6 nM and thus represent high-affinity RNA/protein interactions. In contrast, the binding of aptamer D1 was about 20-fold weaker and showed a  $K_d$  in the range of 100 nM.

**In vivo selection of RNA aptamers interacting with Rrm4 using the yeast three-hybrid system**

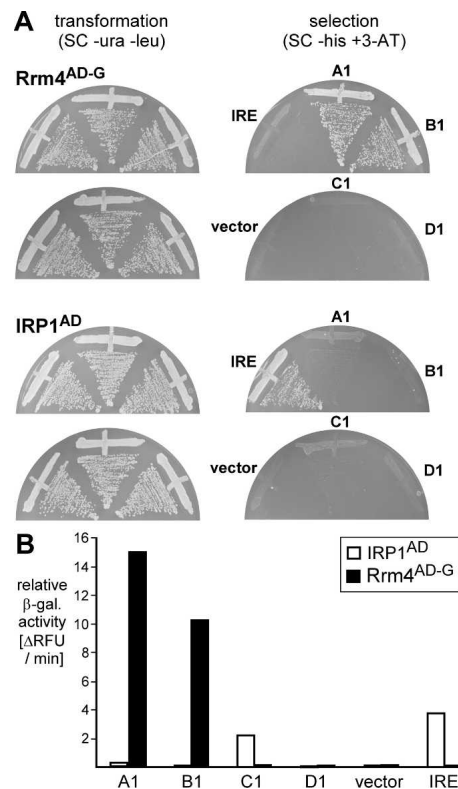
In order to identify aptamer sequences that interact with Rrm4 in vivo we applied the yeast three-hybrid system. Using a *LEU2* gene-containing yeast vector, Rrm4 was expressed as fusion protein with the Gal4 activation domain and the green fluorescence protein (eGfp, enhanced version, Clontech) at the N and C terminus (Rrm4<sup>AD-G</sup>), respectively, (Fig. 3A). The resulting plasmid was transformed into yeast strain L40-coat (SenGupta et al. 1996) containing the two reporter genes *lexA-op-lacZ* as well as *lexA-op-HIS3* and expressing the first hybrid LexA-MS2CP (Fig. 3A). A comparable fusion protein of Rrm4 and eGfp has been shown to be functional in *U. maydis* (Becht et al. 2006). Gfp fluorescence was used to verify, by microscopy, that Rrm4<sup>AD-G</sup> was expressed in the yeast nucleus (data not shown).



**FIGURE 3.** Yeast three-hybrid selection of RNA aptamers in vivo. (A) Schematic representation of the yeast three-hybrid system. For simplicity only one MS2 protein and binding site are indicated (the MS2 protein binds as a dimer and the RNA hybrid contains two MS2 binding sites). The label “selex” symbolizes the presence of sequences from the 11th cycle of SELEX. CP, AD, and op are abbreviations for coat protein, activation domain, and operator, respectively. (B) Forty-one clones were sequenced and the aptamers were classified as described in Fig. 2A.

To construct the hybrid RNA library we inserted DNA templates from the 11th SELEX cycle downstream from the MS2-binding sites of an *URA3* gene-containing vector (pIIIMS2–4). For quality control, the library DNA was transformed into *E. coli* cells and the inserts of 39 randomly selected clones were sequenced. All four SELEX classes were represented with a relative abundance of 18% (class A), 36% (class B), 31% (class D), and 15% (class C) (data not shown). Next, the library DNA was transformed into yeast strain L40-coat (SenGupta et al. 1996) expressing Rrm4<sup>AD-G</sup> (see Materials and Methods), and 10<sup>5</sup> transformants were pooled and tested for *HIS3* expression (growth on SC plates –his+5 mM 3-AT, see Materials and Methods). *HIS3* and *lacZ* reporter gene expression could be confirmed for all 104 randomly selected histidine prototrophs. DNA inserts of 41 randomly selected clones were sequenced after amplification by whole-cell PCR. Sequence comparison revealed that members of classes A and B were selected exclusively (44% and 56%, respectively) including three new variants of class A (Fig. 3B; sequences A6–A8). The latter aptamers share the characteristic primary and secondary structure elements of class A including the highly conserved AU-rich 10-nt loop region (Fig. 2B). Thus, class A and B sequences were detected with roughly equal frequencies during SELEX and three-hybrid selection, whereas class C and D aptamers were completely absent from this selection although they account for >60% of the in vitro selected sequences. This might indicate that, whereas aptamers of classes A and B bind in vivo, those of the remaining classes C and D were nonfunctional.

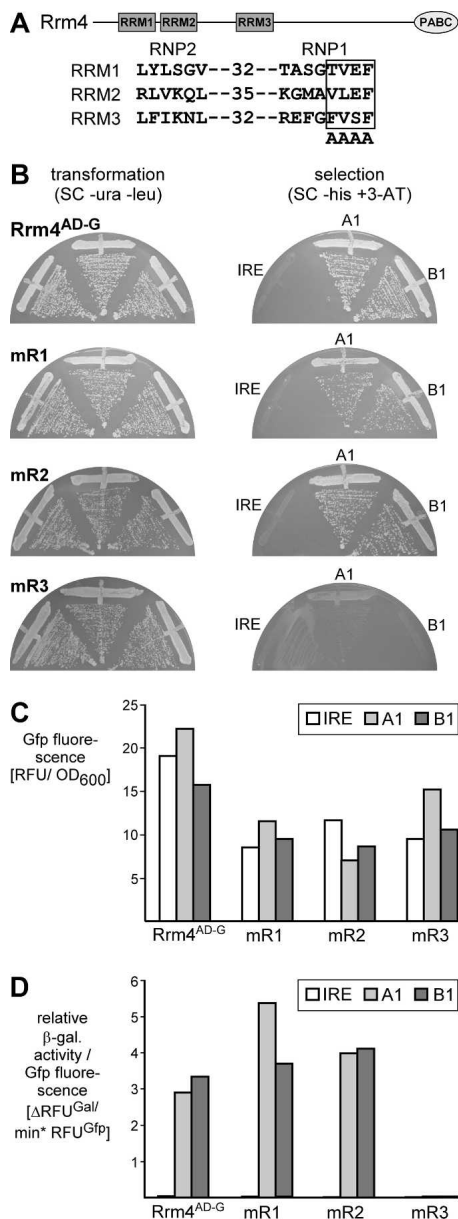
In order to verify this result, the most abundant member of each class was tested individually for its interaction with Rrm4<sup>AD-G</sup> in the three-hybrid assay. Plasmids encoding hybrid RNAs MS2-A1, MS2-B1, MS2-C1, and MS2-D1 were transformed into L40-coat expressing Rrm4<sup>AD-G</sup>. Only MS2-A1 and MS2-B1 were able to trigger *HIS3* reporter gene expression, whereas MS2-C1 and MS2-D1 were inactive (Fig. 4A, top). The specificity of the aptamer-mediated activation was tested in a control experiment using the iron response protein (IRP1) from mammals that has been shown to bind the iron responsive element (IRE) in vivo (SenGupta et al. 1996). As expected, only MS2-IRE was able to trigger high *HIS3* reporter gene expression (Fig. 4A, bottom). These results were further verified by quantifying  $\beta$ -galactosidase reporter gene activation using a fluorogenic substrate. In this assay, the change of relative fluorescence units (RFUs) over time was determined using whole-cell extracts of each transformant. Consistently, MS2-A1 and MS2-B1 were able to activate  $\beta$ -galactosidase expression in combination with Rrm4<sup>AD-G</sup> (Fig. 4B). In the case of IRP1<sup>AD</sup>, MS2-IRE, as well as MS2-C1, were able to elicit  $\beta$ -galactosidase expression, indicating a weak interaction of MS2-C1 with IRP1 (Fig. 4B). In summary, representatives from two aptamers classes, MS2-A1 and MS2-B1, interact specifically and efficiently with Rrm4 in vivo.



**FIGURE 4.** Three-hybrid assay with representative members from all classes. (A) *HIS3* expression was monitored as growth on selection plates (SC –his+5 mM 3-AT). Strain L40-coat expressed either Rrm4<sup>AD-G</sup> (top) or IRP1<sup>AD-G</sup> (bottom). The respective hybrid RNA is indicated next to the various plate sectors. (B)  $\beta$ -galactosidase expression was quantified using a fluorogenic substrate. The difference in relative fluorescence units (RFUs) per minute is given for strains expressing the different proteins (*inlay*) and hybrid RNAs (*bottom*) tested.

### Mapping the Rrm4 interaction domain of the RNA aptamers in vivo

In order to identify the domains of Rrm4 involved in the interaction with aptamers A1 and B1, we tested Rrm4 variants carrying point mutations in the RNP1 region of RRM1, RRM2, and RRM3 (Rrm4-mR1<sup>AD-G</sup>, Rrm4-mR2<sup>AD-G</sup>, and Rrm4-mR3<sup>AD-G</sup>) (Fig. 5A). The same point mutations in RRM1 and RRM3 have been analyzed before in *U. maydis* (Becht et al. 2006). The corresponding plasmids were cotransformed with hybrid RNA-encoding plasmids expressing MS2-A1, MS2-B1, and MS2-IRE hybrid RNAs into yeast strain L40-coat. *HIS3* expression (growth on SC plates –his+5 mM 3-AT) was detected in strains expressing hybrid RNAs MS2-A1 and MS2-B1 with Rrm4<sup>AD-G</sup>, Rrm4-mR1<sup>AD-G</sup>, and Rrm4-mR2<sup>AD-G</sup> but not with Rrm4-mR3<sup>AD-G</sup>, demonstrating that both aptamers interact with the third RRM domain (Fig. 5B). To control for the expression of Rrm4 mutants we determined Gfp fluorescence of the various Rrm4<sup>AD-G</sup> versions in combination with three different hybrid RNAs (RFUs per OD<sub>600</sub>).



**FIGURE 5.** Identifying the aptamer interaction domain of Rrm4 in vivo. (A) Schematic representation of the Rrm4 domain architecture (*top*; RRM and C-terminal PABC domains are given as rectangles and oval, respectively) and the introduced point mutations (*bottom*). The boxed amino acids in the RNA contact region RNP1 were mutated to alanine. The numbers indicate the distance between RNP1 and RNP2. (B) *HIS3* expression was monitored as growth on selection plates (SC -his+5 mM 3-AT). Strain L40-coat expressed different versions of Rrm4<sup>AD-G</sup> (Rrm4<sup>AD-G</sup>-mR1, Rrm4<sup>AD-G</sup>-mR2, and Rrm4<sup>AD-G</sup>-mR3 carry point mutations in RRM1, RRM2, and RRM3, respectively; labeled with mR1, mR2, and mR3 on the *left*). The respective hybrid RNA is indicated next to the various plate sectors (labeled on the *right*). (C) Expression of Rrm4<sup>AD-G</sup> versions (given at the *bottom*) was determined in strains expressing different hybrid RNAs (*inlay*) by quantifying Gfp fluorescence per OD<sub>600</sub>. (D) Relative β-galactosidase activity was determined using a fluorogenic substrate. To account for differences in protein expression, enzyme activity is given relative to Gfp fluorescence (ΔRFU<sup>Gal</sup> per minute per RFU<sup>Gfp</sup>). Different versions of Rrm4<sup>AD-G</sup> (*bottom*) were tested with various hybrid RNAs (*inlay*).

In all strains each variant was expressed, albeit Rrm4<sup>AD-G</sup> versions carrying point mutations were expressed with approximately twofold lower efficiency compared to wild type (Fig. 5C). To verify our results obtained during determining *HIS3* expression, we measured β-galactosidase reporter gene expression using a fluorogenic substrate as described above. Since reporter gene expression is dependent on the amount of Rrm4<sup>AD-G</sup> versions (Hook et al. 2005), we calculated the increase of fluorescence over time relative to the amount of Gfp fluorescence (Fig. 5D; ΔRFU<sup>Gal</sup> per min per RFU<sup>Gfp</sup>). Consistently, β-galactosidase activity was not detectable when Rrm4-mR3<sup>AD-G</sup> was tested with MS2-A1 or MS2-B1 hybrid RNAs (Fig. 5D). Thus, the yeast three-hybrid system allowed rapid identification of the aptamer interaction domain of Rrm4.

## DISCUSSION

SELEX selects and amplifies RNA aptamers from a large pool of combinatorial sequences. The yeast three-hybrid system detects RNA/protein interactions in vivo. Here we combine these techniques (Cassiday and Maher 2001, 2003) to select and characterize RNA aptamers that interact with the RRM protein Rrm4 in vivo. After 11 cycles of SELEX, a single round of selection applying the three-hybrid system was sufficient to reduce the complexity of the SELEX sequences to ~50% and identify a distinct subset of RNA aptamers that function in vivo. Secondary structure calculations of these selected aptamers revealed similarities consisting of two stem-loops, a purine-rich bulge and an AU-rich loop sequence (Fig. 2B). This related structure is in accordance with our observation that members of both classes interact with the same domain of the target Rrm4, i.e., the third RRM.

As outlined above RNA aptamers have the capability of binding to a variety of different targets. Therefore, the identification of RNA aptamers that interact with a known RNA-binding protein is no guarantee that the aptamers are directed against the RNA-binding domain of the target protein. Our approach to test different variants carrying point mutations in the RNA contact region of the RRM allowed the identification of RRM3 as the interaction domain. Thus, applying the yeast three-hybrid system is also a straightforward approach to map the target “epitope” of RNA aptamers.

Rrm4 has been implicated as functioning during polar growth of the infectious hyphae in *U. maydis*. Since it binds RNA in vivo, and as it moves bidirectionally along microtubules toward both ends of the tip cell, it has been hypothesized that Rrm4 is involved in mRNA transport from the nucleus to the poles (Becht et al. 2006). Currently, we are applying an in vivo UV cross-linking approach to identify target RNAs (Ule et al. 2005). The identified RNA aptamers will be helpful in guiding our search for the

identification of Rrm4 binding sites within candidate sequences.

Why does Rrm4 discriminate between the four SELEX classes in the yeast three-hybrid system? The most plausible explanation at least for the observation that Rrm4 interacts with members of classes A and B but not with members of class D is the difference in binding affinities. Members of classes A and B bind with high affinity ( $K_d$  values of  $\sim 6$  and  $4$  nM, respectively), whereas the member of class D exhibits  $\sim 20$ -fold lower affinity ( $K_d$  of  $\sim 100$  nM). This assumption is supported by recent work systemically investigating the relationship between reporter gene expression in the yeast three-hybrid system and in vitro affinity of an RNA/protein interaction (Hook et al. 2005). High-affinity interactions of FBF-1 from *Caenorhabditis elegans* with its wild-type binding sequence FBEa ( $K_d$  of  $\sim 10$  nM) can be detected on selection plates (SC  $-his$ ) even in the presence of  $100$  mM 3-AT. When testing low-affinity interactions such as the binding of FBF-1 to a different binding site PME ( $K_d$  of  $\sim 80$  nM), optimal growth on selection plates is detected only in the absence of 3-AT (Hook et al. 2005). This is consistent with our observation that members of class D exhibiting weak interaction with Rrm4 ( $K_d$  of  $\sim 100$  nM) do not grow on our selection plates containing  $5$  mM 3-AT. It remains to be solved why Rrm4 interaction with the class C aptamer is not detectable in the three-hybrid system although the binding affinity in vitro is high ( $K_d$  of  $\sim 6$  nM). Possible explanations are that this particular RNA aptamer sequence is expressed poorly as a hybrid RNA with the tandem MS2-binding sequences. Alternatively, the aptamer binding site in Rrm4 could be masked in the fusion protein, for example, by the activation domain. However, it is also conceivable that this particular aptamer is unable to fold correctly in vivo. The discrimination of the different aptamers for in vivo binding supports the dual approach combining SELEX and the three-hybrid system. It would have been impossible to predict the correct binding behavior of the various aptamers in vivo solely based on the in vitro binding data.

In previous studies the yeast three-hybrid system was used to select from a random pool in vivo binding sequences for human mRNA export factor Tap or the PUF protein FBF-1 from *Caenorhabditis elegans* (Kang et al. 1999; Bernstein et al. 2005). These studies were successful for the selection of new variants of known binding sequences. However, both studies were limited in exploring the sequence space by the number of yeast transformants, typically  $10^7$ . The combination of in vitro and in vivo techniques has the potential to overcome this obstacle. This has been demonstrated by combining SELEX with the three-hybrid system to identify an RNA aptamer in vivo that functions as an RNA decoy for the transcription factor NF- $\kappa$ B (Cassiday and Maher 2001, 2003). Aptamers with improved in vivo activity in the three-hybrid system were selected by testing mutagenized versions of the original

RNA aptamer or assaying binding candidates after a few cycles of SELEX under increased selection pressure for *HIS3* expression (using higher 3-AT concentrations) (Cassiday and Maher 2003). In the latter case an additional in vivo binder could be identified that might have been lost after excessive in vitro cycles. Here we demonstrate that the combination of SELEX with the three-hybrid system is a useful strategy not only for selection but also for classification of in vivo RNA aptamers. In addition the application of the three-hybrid system allows rapid mapping of the aptamer binding site of the interaction partner.

Future extensions of this approach could be the combination of rapid non-SELEX selection methods such as nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) (Berezovski et al. 2006) with the yeast three-hybrid system or the application of SELEX with the yeast one-hybrid system (Li and Herskowitz 1993) to identify DNA-binding sites of transcription factors.

Combination of SELEX and the three-hybrid system might become particularly attractive for drug discovery, RNA diagnostics, and therapeutics (Cassiday and Maher 2003). Since both techniques can be automated (Cox et al. 2002; Eulberg et al. 2005; Stelzl et al. 2005), high-throughput applications also are feasible.

## MATERIALS AND METHODS

### Strains, plasmids, and plasmid constructions

The yeast strain L40-coat with the following genotype was used: *MATa*, *ura3-52*, *leu2-3,112*, *his3 $\Delta$ 200*, *trp1 $\Delta$ 1*, *ade2*, *LexA-MS2 coat (TRP1)*, *LYS2::lexA-op-HIS3*, and *ura3::lexA-op-lacZ* (SenGupta et al. 1996). For cloning purposes and protein expression we used *E. coli* K-12 derivatives Top10 (Invitrogen) and BL21(DE3) $\Delta$ pLys (Stratagene). The plasmids used were pCR2.1-TOPO (Invitrogen) for cloning, pGEX-2T (GE healthcare) for protein expression, and pIIIMS2-1 as well as pACTII for the three-hybrid analysis (SenGupta et al. 1996). Plasmids were constructed as follows. In the case of pRrm4<sup>GST</sup>-pGEX (pUMa429), a 3.1-kb BamHI/EcoRI fragment from pRrm4<sup>AD-G</sup>-ACTIIA (pUMa427) was inserted into pGEX-2T. The resulting plasmid encodes the complete *rrm4* ORF (Becht et al. 2005) fused to the coding region of the enhanced fluorescence protein (eGfp, Clontech) via the short amino acid linker AANAAT. Plasmid pBruno<sup>GST</sup>-pGEX (pUMa435) contains a 2-kb BamHI/EcoRI fragment from pBruno-ACTII encoding the complete Bruno ORF in pGEX-2T. In the case of pACTIIA (pUMa399), a 2.1-kb BsrGI/EagI fragment from pIIIMS2-2 (Bernstein et al. 2002) encoding Ade2p was inserted into pACTII. In the case of pRrm4<sup>AD-G</sup>-ACTIIA (pUMa427), a 3.1-kb NcoI/AscI fragment from pRrm4G-NatR (Becht et al. 2006) was inserted into pACTIIA using oligonucleotide linkers. Plasmids pRrm4-mR1<sup>AD-G</sup>-ACTIIA (pUMa581), pRrm4-mR2<sup>AD-G</sup>-ACTIIA (pUMa594), and pRrm4-mR3<sup>AD-G</sup>-ACTIIA (pUMa526) are comparable to pRrm4<sup>AD-G</sup>-ACTIIA but contain the *rrm4* ORF with mutations in the RNP1 regions of RRM1, RRM2, and RRM3 (Fig. 5A). Mutations mR1, mR2, and mR3 were introduced using oligonucleotide primers

oMF750 5'-TATCTAGAGTCTGCGCCGCGGCCGCGCCGCTGG CTGTCTGGCCC-3', oMF764 5'-CGGCGCGCTGGGCGTCTTG TTCAGAGTAGGCCCGCCGCGGCCATACCCTTGAAGCC-3', and oMF755 5'-GAGAATTCGGGCGCGGCTGCCACCACTCC TGACGAAGCGC-3'. In the case of pIRP1<sup>AD</sup>-ACT1IA (pUMa412), a 3-kb MluI/BamHI fragment from pAD-IRP1 (Bernstein et al. 2002) was inserted into pACT1IA. The plasmid pIIIMS2-4 (pUMa486), which was used for the construction of the RNA hybrid library, was generated by annealing oligonucleotides oMF762 5'-CCGGAATCGATGCTTAATTAAGGCGCGC-3' with oMF763 5'-CCGGGCGCGCTTAATTAAGCATCGATT-3' and inserting the double-stranded product into the XmaI site of pIIIMS2-1 (Bernstein et al. 2002). Thereby, the multiple-cloning site contains the restriction endonuclease cleavage sites for ClaI, PacI, and AscI, allowing directional cloning of ClaI/AscI inserts. In the case of pIRE-IIIMS2-4 (pUMa492), oligonucleotides oMF636 5'-AAATCGATCCCAGGATCTGTATCTTGCTTCAACAGTGT TTGGACGGAACAGACCCGGATCGATAA-3' and oMF637 5'-TTATCGATCCGGGCTCTGTTCCG-TCCAAACACTGTTGAAGCA AGATACAGATCCTCGGGATCGATT-3' were annealed, cleaved with ClaI, and inserted into pIIIMS2-4. Plasmids pA1-IIIMS2-4 (pUMa550), pB1-IIIMS2-4 (pUMa546), pC1-IIIMS2-4 (pUMa866), and pD1-IIIMS2-4 (pUMa867) correspond to *E. coli* transformants that were generated during quality control of the library. All plasmid constructs were verified by DNA sequencing.

### Protein expression in *E. coli*

Recombinant Rrm4<sup>GST</sup>, Bruno<sup>GST</sup>, and GST were expressed as follows. BL21(DE3) $\Delta$ pLys cells harboring plasmids pRrm4<sup>GST</sup>-pGEX, pBruno<sup>GST</sup>-pGEX, or pGEX-2T were grown in dYT medium containing ampicillin (100  $\mu$ g/mL) and chloramphenicol (25  $\mu$ g/mL) at 37°C. At an OD<sub>600</sub> of 0.8, cells were induced with 0.2 mM IPTG for 16 h at 22°C; 600 mL cell suspension were harvested, washed, and resuspended in 20 mL buffer A (50 mM Tris/HCl pH 8.0, 150 mM NaCl, 5 mM DTT, protease inhibitor cocktail Complete, Roche). Cells were sonicated and insoluble cell debris was removed by centrifugation (4°C, 45 min, 50,000g). The supernatant was incubated with 3 mL of GSH-sepharose beads (GE healthcare) for 5 h at 4°C. Beads were washed on the column (five times) with buffer A. Elution of GST fusion proteins was performed using 20 mM reduced glutathione, which was subsequently removed by dialysis against buffer A for 20 h at 4°C. Finally, the protein concentration was determined (Bradford 1976).

### In vitro selection (SELEX)

The single-stranded DNA starting library (5'-GAATTCAGTCG GACAGCG(N)<sub>40</sub>GATGG-ACGAATATCGTCTCCC-3') was synthesized and amplified as described (Homann and Göringer 1999). It contained a central sequence stretch of 40 randomized nucleotides flanked by constant regions C1 and C2. In the initial cycle, 100  $\mu$ g of the library were transcribed in transcription buffer (20 mM Na<sub>2</sub>H<sub>2</sub>PO<sub>4</sub> pH 7.9, 8 mM MgCl<sub>2</sub>, 20 mM DTT, 4 mM spermidine) using 500 U T7 RNA polymerase (Stratagene), 160 U RNasin (Promega), and 1 mM NTPs for 1.5 h at 37°C in a volume of 1.2 mL. In subsequent cycles, 20–50  $\mu$ g PCR product were used as template for in vitro transcription. The reaction was terminated upon addition of MgCl<sub>2</sub> (final concentration of 50 mM) and DNase I (1 h 37°C; 1 U DNase I/ $\mu$ g DNA, Roche). RNAs were

extracted with phenol, precipitated with ethanol and redissolved in binding buffer (25 mM Tris/HCl pH 7.5, 5 mM MgCl<sub>2</sub>, 1 mM DTT, 100 mM NaCl, 4 mM KCl, 5 mM glucose). A final purification was achieved by size exclusion chromatography in the same buffer. Before each cycle of selection, all RNAs were pre-incubated at 37°C for 30 min. For the first cycle, 250  $\mu$ g RNA library were incubated with Rrm4<sup>GST</sup> (750 nM) in 1 mL binding buffer for 30 min at RT. In later cycles, the selection stringency was increased stepwise by decreasing the amount of Rrm4<sup>GST</sup> (from 750 nM in cycle 1 to 4nM in cycles 11–12), decreasing the amount of RNA (from 250  $\mu$ g in cycle 1 to 12.5  $\mu$ g in cycles 9–12), and increasing the reaction volume from 1 to 8 mL (cycles 11–12). Non-specific competitor RNA (total RNA from *U. maydis*) was added in cycle 4 (10  $\mu$ g/ $\mu$ g pool RNA) and in cycles 9–12 (100  $\mu$ g/ $\mu$ g pool RNA). RNA sequences with affinity for the filter material or the GST moiety of Rrm4<sup>GST</sup> were depleted from the pool by counterselection in cycles 3–4 and 7–12. This was achieved by passing the in vitro transcribed RNA through a filter prior to binding and by pre-incubation with GST using an RNA/GST ratio of 1:10.

Bound RNAs were partitioned from unbound material by filter adsorption to nitrocellulose/cellulose acetate mixed bed filters. Samples were filtered using a mild vacuum and filters were washed with binding buffer (at least fivefold excess to the initial binding volume). Selected RNAs were recovered by phenol extraction followed by ethanol precipitation. RNA was reverse-transcribed using 150 U MuLV-RT (USB) and 10  $\mu$ g primer oMF731 5'-GGTCGAGAATTCAGTCGGACAGCG-3' for 45 min at 42°C followed by 5 min at 48°C. The resulting DNA products were amplified by PCR with oligonucleotides oMF731 and oMF730 5'-GCCAAGCTTAATACGACTCACTATAGGGAGACGATATT CGTCCATC-3'. The resulting DNA was used as a template for the next cycle. DNA from cycle 11 was inserted into pCR2.1-TOPO and inserts were sequenced by dideoxy terminator sequencing.

### Classification of aptamer sequences and structure determination

The aptamer sequences were aligned using Clone Manager Professional Suite version 8 using the Multi-Way alignment type. The identification of sequence motifs and the final alignment were performed by visual inspection and with the help of secondary structure calculations using Mfold (Zuker 2003).

### Determination of equilibrium dissociation constants

The binding affinity of RNA aptamers to recombinant Rrm4<sup>GST</sup> was measured by filter adsorption. Uniformly (<sup>32</sup>P)-labeled RNA aptamers were incubated with Rrm4 (15 nM) for 20 min at RT in 50  $\mu$ L binding buffer (25 mM Tris/HCl pH 7.5, 5 mM MgCl<sub>2</sub>, 2 mM DTT, 100 mM NaCl, 5 mM KCl, 100 ng/ $\mu$ L BSA, 70 ng/ $\mu$ L tRNA<sub>yeast</sub>). Samples were filtered through nitrocellulose/cellulose acetate mixed bed filters and washed with 2 mL binding buffer. The percentages of bound RNA were determined by scintillation counting. Equilibrium dissociation constants (K<sub>d</sub>) were derived from experiments determining the bound percentage of labeled aptamer (0.1 nM) at increasing concentrations of unlabeled RNA (0.6–300 nM) in the presence of a constant Rrm4<sup>GST</sup> concentration (15 nM). K<sub>d</sub> values were calculated based on the Scatchard equation:  $r/[A] = n/K_d - r/K_d$ , with the slope of the plot

representing  $-1/K_d$  ( $[A] = [\text{free RNA}]$ ,  $r = [\text{bound RNA}]$ , and  $n = \text{number of binding sites}$ ).

### Yeast three-hybrid analysis

Construction of the yeast three-hybrid RNA library was done as follows. cDNA of the 11th SELEX cycle was PCR-amplified using primer oMF767 5'-AACATCGATGAGACGATATTCGTCCATC-3' and oMF768 5'-TATGGCGCGCCAGAATTCAGTCGGACAGCG-3' introducing ClaI and AscI restriction endonuclease sites 5' and 3' to the constant regions C1 and C2, respectively. PCR products were cleaved (ClaI/AscI) and inserted into the hybrid RNA encoding vector pIIIMS2-4. About  $4 \times 10^4$  transformants were pooled and plasmid DNA was isolated. For quality control, 1  $\mu\text{g}$  of the library DNA was transformed into *E. coli* and 39 randomly chosen clones were sequenced after whole-cell PCR using standard procedures.

Subsequently, 10  $\mu\text{g}$  of library DNA were transformed into strain L40-coat (Gietz and Woods 2002) that contained plasmid pRrm4<sup>AD-G</sup>-ACTIIA, and transformants were incubated on synthetic complete plates lacking leucine and uracil (SC –leu–ura; 2 d at 28°C) to select for the presence of both plasmids. About  $10^5$  transformants were pooled and incubated on SC plates lacking histidine and containing 5 mM 3-aminotriazole (SC –his+5 mM 3-AT; an inhibitor of His3p). One-hundred-four randomly chosen clones were verified to express both reporter genes according to growth on plates (SC –his+5 mM 3-AT) and  $\beta$ -galactosidase activity, which was determined using standard filter assays with X-gal as the chromogenic substrate (Bernstein et al. 2002). Forty-one randomly chosen clones were subjected to whole-cell PCR analysis with oligonucleotide primers oMF713 5'-AAATTGGAAGTTGGA TATGG-3' and oMF725 5'-CATGGTACGCTGTGGTGC-3' using standard procedures. The resulting PCR products were sequenced.

To test representative aptamers of each class individually, RNA (pA1-IIIMS2-4, pB1-IIIMS2-4, pC1-IIIMS2-4, pD1-IIIMS2-4, and pIRE-MS2-4) and protein-encoding plasmids (pRrm4<sup>AD-G</sup>-ACTIIA and pIRP1<sup>AD-G</sup>-ACTIIA) were cotransformed in strain L40-coat and incubated on SC –ura–leu plates (2 d at 28°C). At least three transformants were assayed on plates (SC –his+5 mM 3-AT; 3 d at 28°C).

To test different mutant versions of Rrm4 for RNA binding in vivo, RNA (pA1-IIIMS2-4, pB1-IIIMS2-4, and pIRE-MS2-4) and protein-encoding plasmids (pRrm4<sup>AD-G</sup>-ACTIIA, pRrm4-mR1<sup>AD-G</sup>-ACTIIA, pRrm4-mR2<sup>AD-G</sup>-ACTIIA, and pRrm4-mR3<sup>AD-G</sup>-ACTIIA) were cotransformed in strain L40-coat and assayed for reporter gene expression as mentioned.

### Fluorimetric measurements of $\beta$ -galactosidase activity

A densely grown culture of cells (SC –ura–leu, 28°C, shaking with 200 rpm) was diluted (optical density, OD<sub>600</sub> of ~0.2). After ~4 h of incubation (OD<sub>600</sub> of ~0.6–0.8)  $\sim 2 \times 10^8$  cells ( $10^7$  cells/mL represent OD<sub>600</sub> = 1) were harvested by centrifugation (room temperature, RT, 1000g, 5 min). The pellet was washed with 1 mL chilled reaction buffer (0.1 M Na<sub>2</sub>H<sub>2</sub>PO<sub>4</sub> pH 7.3; 1 mM MgCl<sub>2</sub>) and resuspended in 150  $\mu\text{L}$  chilled reaction buffer containing 45 mM  $\beta$ -mercaptoethanol. Upon adding of 0.3 g acid-washed glassbeads, cells were rigorously shaken (4°C, 1400 rpm, 5 min). After two rounds of centrifugation (4°C, 16,000g, 5 and 20 min

each), the protein concentration of the supernatant was determined (Bradford 1976). Fifty nanograms of total protein extract were assayed in the presence of 100  $\mu\text{L}$  CUG working solution (3-carboxyumbelliferyl  $\beta$ -D-galactopyranoside, FluoReporter quantification kit; Invitrogen) using a monochromator fluorescence reader (Tecan). Relative fluorescence units were determined at 25°C with excitation and emission wavelengths of 385 and 445 nm, respectively (bandwidth of 12.5 nm each). Measurements were performed in 1-min intervals (in total 30 min). Relative  $\beta$ -galactosidase activity was expressed as RFUs per minute. Negative values due to very low enzyme activity were considered to be zero. Two independent experiments were performed and representative results are shown.

To control expression of mutant Rrm4<sup>AD-G</sup> variants,  $\sim 3 \times 10^5$  cells were harvested by centrifugation (RT, 1000g, 5 min) and resuspended in 200  $\mu\text{L}$  H<sub>2</sub>O. Gfp fluorescence was measured with excitation and emission wavelengths of 485 nm and 520 nm, respectively (bandwidth of 7.5 nm each). The resulting RFUs were normalized to OD<sub>600</sub>. To account for expression differences of Rrm4<sup>AD-G</sup> variants, we normalized the relative  $\beta$ -galactosidase activity to the Gfp fluorescence. Two independent experiments were performed and representative results are shown.

### ACKNOWLEDGMENTS

We thank Dr. A. Brachmann and the laboratory members for valuable discussions and critically reading of the manuscript. We thank Janine Koepke for initial work on Rrm4 mutants in yeast and Julia Hohenner for excellent technical assistance. We are grateful to Drs. M. Wickens and P. MacDonald for the generous gift of the three-hybrid system and a Bruno expression plasmid, respectively. This work was supported by funds from Bayer CropScience and DFG Fe448/3 to M.F. as well as by an International Research Scholar grant from the Howard Hughes Medical Institute (HHMI) and DFG SFB579 to H.U.G.

Received October 6, 2006; accepted December 20, 2006.

### REFERENCES

- Becht, P., Vollmeister, E., and Feldbrügge, M. 2005. A role for RNA-binding proteins implicated in pathogenic development of *Ustilago maydis*. *Eukaryot. Cell* **4**: 121–133.
- Becht, P., König, J., and Feldbrügge, M. 2006. The RNA-binding protein Rrm4 is essential for polarity in *Ustilago maydis* and shuttles along microtubules. *J. Cell Sci.* **119**: 4964–4973.
- Berezovski, M., Musheev, M., Drabovich, A., and Krylov, S.N. 2006. Non-SELEX selection of aptamers. *J. Am. Chem. Soc.* **128**: 1410–1411.
- Bernstein, D.S., Buter, N., Stumpf, C., and Wickens, M. 2002. Analyzing mRNA-protein complexes using a yeast three-hybrid system. *Methods* **26**: 123–141.
- Bernstein, D., Hook, B., Hajarnavis, A., Opperman, L., and Wickens, M. 2005. Binding specificity and mRNA targets of a *C. elegans* PUF protein, FBF-1. *RNA* **11**: 447–458.
- Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**: 248–254.
- Cassiday, L.A. and Maher, L.J. 2001. In vivo recognition of an RNA aptamer by its transcription factor target. *Biochemistry* **40**: 2433–2438.



- Cassiday, L.A. and Maher, L.J. 2003. Yeast genetic selections to optimize RNA decoys for transcription factor NF- $\kappa$ B. *Proc. Natl. Acad. Sci.* **100**: 3930–3935.
- Cox, J.C., Hayhurst, A., Hesselberth, J., Bayer, T.S., Georgiou, G., and Ellington, A.D. 2002. Automated selection of aptamers against protein targets translated in vitro: From gene to aptamer. *Nucleic Acids Res.* **30**: e108.
- Ellington, A.D. and Szostak, J.W. 1990. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**: 818–822.
- Eulberg, D., Buchner, K., Maasch, C., and Klussmann, S. 2005. Development of an automated in vitro selection protocol to obtain RNA-based aptamers: Identification of a biostable substance P antagonist. *Nucleic Acids Res.* **33**: e45.
- Gietz, R.D. and Woods, R.A. 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol.* **350**: 87–96.
- Göringer, H.U., Homann, M., and Lörger, M. 2003. In vitro selection of high-affinity nucleic acid ligands to parasite target molecules. *Int. J. Parasitol.* **33**: 1309–1317.
- Göringer, H.U., Homann, M., Zacharias, M., and Adler, A. 2006. RNA aptamers as potential pharmaceuticals against infections with African trypanosomes. *Handb. Exp. Pharmacol.* **173**: 375–393.
- Homann, M. and Göringer, H.U. 1999. Combinatorial selection of high affinity RNA ligands to live African trypanosomes. *Nucleic Acids Res.* **27**: 2006–2014.
- Homann, M., Lörger, M., Engstler, M., Zacharias, M., and Göringer, H.U. 2006. Serum-stable RNA aptamers to an invariant surface domain of live African trypanosomes. *Comb. Chem. High Throughput Screen* **9**: 491–499.
- Hook, B., Bernstein, D., Zhang, B., and Wickens, M. 2005. RNA–protein interactions in the yeast three-hybrid system: Affinity, sensitivity, and enhanced library screening. *RNA* **11**: 227–233.
- Kang, Y., Bogerd, H.P., Yang, J., and Cullen, B.R. 1999. Analysis of the RNA binding specificity of the human Tap protein, a constitutive transport element-specific nuclear RNA export factor. *Virology* **262**: 200–209.
- Kim-Ha, J., Kerr, K., and MacDonald, P.M. 1995. Translational regulation of *oskar* mRNA by bruno, an ovarian RNA-binding protein, is essential. *Cell* **81**: 403–412.
- Lee, J.F., Hesselberth, J.R., Meyers, L.A., and Ellington, A.D. 2004. Aptamer database. *Nucleic Acids Res.* **32**: D95–D100.
- Li, J.J. and Herskowitz, I. 1993. Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. *Science* **262**: 1870–1874.
- Murphy, M.B., Fuller, S.T., Richardson, P.M., and Doyle, S.A. 2003. An improved method for the in vitro evolution of aptamers and applications in protein detection and purification. *Nucleic Acids Res.* **31**: e110.
- Nimjee, S.M., Rusconi, C.P., and Sullenger, B.A. 2005. Aptamers: An emerging class of therapeutics. *Annu. Rev. Med.* **56**: 555–583.
- Proske, D., Blank, M., Buhmann, R., and Resch, A. 2005. Aptamers: Basic research, drug development, and clinical applications. *Appl. Microbiol. Biotechnol.* **69**: 367–374.
- SenGupta, D.J., 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.* **93**: 8496–8501.
- SenGupta, D.J., Wickens, M., and Fields, S. 1999. Identification of RNAs that bind to a specific protein using the yeast three-hybrid system. *RNA* **5**: 596–601.
- Soller, M. and White, K. 2004. ELAV. *Curr. Biol.* **14**: R53.
- Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F.H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., et al. 2005. A human protein–protein interaction network: A resource for annotating the proteome. *Cell* **122**: 957–968.
- Sullenger, B.A. and Gilboa, E. 2002. Emerging clinical applications of RNA. *Nature* **418**: 252–258.
- Tuerk, C. and Gold, L. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**: 505–510.
- Ule, J., Jensen, K., Mele, A., and Darnell, R.B. 2005. CLIP: A method for identifying protein–RNA interaction sites in living cells. *Methods* **37**: 376–386.
- Ulrich, H. 2006. RNA aptamers: From basic science toward therapy. *Handb. Exp. Pharmacol.* **173**: 305–326.
- Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**: 3406–3415.