4838–4840 *Nucleic Acids Research, 1996, Vol. 24, No. 23*

A tri-hybrid system for the analysis and detection of RNA–protein interactions

Ulrich Putz, Paul Skehel[1] and Dietmar Kuhl*

Center for Molecular Neurobiology (ZMNH), University of Hamburg, Martinistraße 52, D-20246 Hamburg, Germany and [1]Division of Neurophysiology, National Institute for Medical Research, London NW7 1AA, UK

Received August 19, 1996; Revised and Accepted October 8, 1996

ABSTRACT

A modification of the two-hybrid system is described for the *in vivo* reconstruction of specific RNA–protein interactions. In this tri-hybrid system, the DNA binding and transcription activation domains of the yeast transcriptional activator GAL4 are brought together via the interaction of recombinant fusion proteins with a recombinant RNA. The method provides a system for studying RNA–protein interactions with the genetic advantages of the two-hybrid system. It may be used to detect specific RNA-binding proteins or target RNAs from a library of cDNAs, or to analyse the structural specificity of identified RNA–protein interactions.

Interactions between RNAs and proteins are fundamental to many cellular processes. Transcription, posttranscriptional modifications, assembly of supramolecular structures such as ribosomes, RNA stabilisation, packaging of viral RNAs, transport of RNA and RNA localisation are all processes that rely on RNA–protein interactions.

RNA–protein interactions have been studied traditionally using extensive biochemical assays such as RNA band shifts, footprinting and RNA–protein cross-linking. A disadvantage of these techniques is that interacting proteins often exist in low abundance and are consequently difficult to detect. In addition, these techniques do not easily allow the identification of target RNAs recognised by a known RNA-binding protein. A further major disadvantage is that these techniques do not allow direct identification of the cognate genes encoding the proteins or RNAs of interest.

The yeast two-hybrid system is a widely employed genetic screen for detecting protein–protein interactions (1–3). We developed a modification of this system to detect RNA–protein interactions in order to take advantage of the genetic basis of the system. The two-hybrid system utilises the modular structure of particular transcription activators, allowing the DNA binding domains and transcription activation domains to function independently. The DNA binding and the transcription activation domains are expressed as two separate polypeptides fused to heterologous sequences. Any interaction between the heterologous polypeptides of the hybrid proteins brings together the two components of the transcription activator inducing the expression of genes under the control of appropriate upstream activating sequences (UAS). Selectable markers such as β-galactosidase or amino acid auxotrophy can then be used to report recombinant protein–protein interactions.

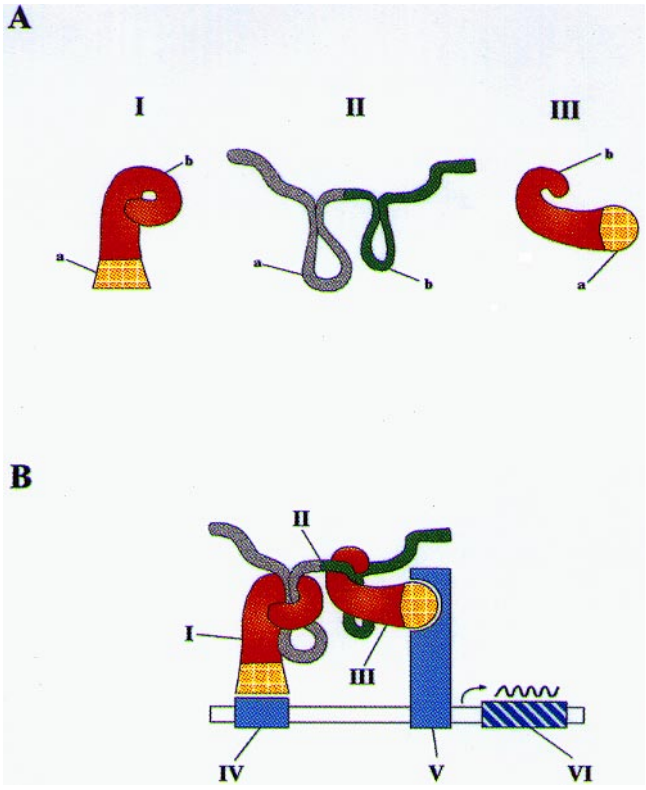

Figure 1. The basic strategy of the tri-hybrid method. (**A**) Schematically shows the components. The first hybrid-protein (I) contains the DNA-binding domain of GAL4 (Ia) fused to the RRE-RNA-binding protein RevM10 (Ib). A hybrid-RNA (II) containing the RRE sequence (IIa) and a target RNA sequence X (IIb). The second hybrid-protein (III) contains the activation domain of GAL4 (IIIa) fused to a protein Y (IIIb) capable of recognising the target RNA X on the RNA-hybrid. (**B**) Upon productive interaction of the three hybrids a reconstituted GAL4 transcription factor (I+II+III) bound to a GAL4 responsive promoter (IV) stimulates the basal transcriptional machinery (V) of the lacZ gene and the nutritional reporter gene HIS3 (VI).

To modify this system for the detection of RNA–protein interactions, it is necessary to construct a system where the association of the DNA-binding and transcription activation domains is dependent on an RNA–protein interaction.

The RevM10 mutation of the HIV-1 Rev protein binds specifically to the Rev responsive element (RRE) sequence in the *env* gene (4,5), but unlike the wild-type protein, is unable to bind to certain cellular proteins and does not promote the export of

*To whom correspondence should be addressed. Tel: +49 40 4717 4739; Fax: +49 40 4717 6595; Email: kuhl@plexus.uke.uni-hamburg.de

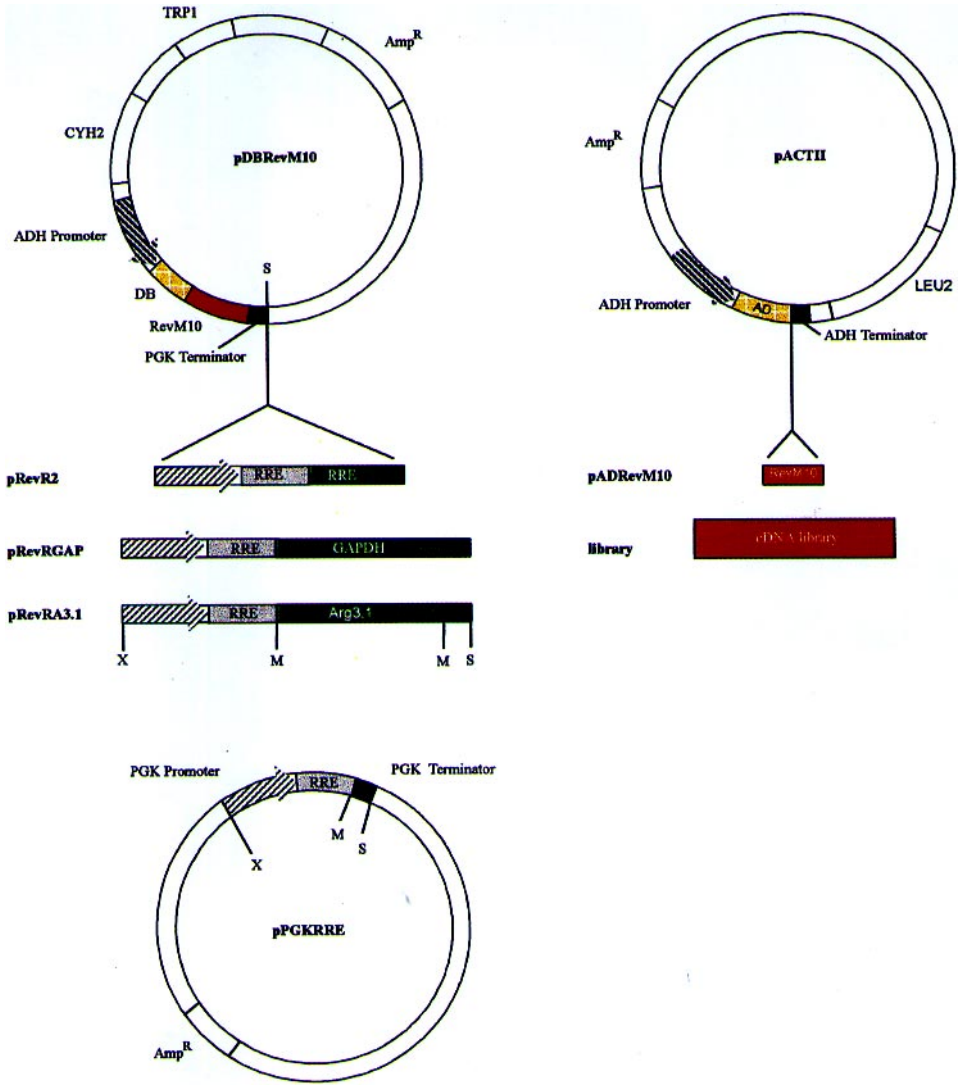

Figure 2. The plasmids used in the tri-hybrid system. The upper plasmids are *E.coli*–yeast shuttle vectors and were used in the transfection and mating experiments of Tables 1 and 2. pPGKRRE was used as an intermediate to clone hybrid RNAs as a transcription unit cassette into pDBRevM10. The transcription units generating the hybrid proteins and the hybrid RNAs are indicated. The respective promoters and terminators are depicted. *E.coli* and yeast replication origins are not shown. *E.coli* selectable marker for ampicillin resistance, Amp^R, and the yeast selectable markers TRP1 and LEU2 are also shown. Cyloheximide sensitivity in yeast is conferred by CYH2. Relevant restriction sites are abbreviated as follows: M, *Mlu*I; S, *Sal*I; X, *Xho*I. Methods. The pDBRevM10 plasmid was constructed in three steps. The RevM10 coding sequence was generated in a PCR reaction using pCRev+M10 (11,12) as a template, with primer sequences: 5′-CTCGAGAAGCTTACCGCCAC-CATGGCAGGAAGAAGCGGAGAC-3′, and 5′-AAGCTTATAGATCTTCTTTAGCTCCTGACTCCA-3′. The PCR product was cloned into pCR^{TM}II (Invitrogen). The RevM10 sequence was then removed as an *Nco*I–*Eco*RV fragment, and cloned into the *Nco*I/*Sma*I sites of a pAS2 (13) derivative, in which a 300 bp *Bam*HI–*Sal*I fragment from pPGK (14) containing the PGK transcription terminator had been cloned. pPGKRRE was constructed by inserting an RRE containing *Eco*RI fragment from pGEMRREter [a derivative of pGEMRRE containing nucleotides 13–224 of the RRE (15) and has an additional *Mlu*I site 3′ adjacent to the RRE sequence] into the *Eco*RI site of pPGK (14). pRevR2 which contains a duplication of the RRE was constructed by removing the PGK terminator with *Mlu*I and *Sal*I from pPGKRRE. The *Mlu*I site was blunted and a blunted *Sac*I–*Sal*I fragment from pPGKRRE containing an RRE followed by the PGK terminator was inserted. From the resulting plasmid, pRRE2, the RRE-RRE transcription unit cassette was released by restriction with *Xho*I and *Sal*I and cloned into the *Sal*I site of pDBRevM10. pRevRGAP was constructed by using a 1272 bp filled in *Xba*I–*Hin*cII fragment containing the entire coding sequence of the GAPDH mRNA derived from pGPDN5 (16). This fragment was cloned into the filled *Mlu*I site of pPGKRRE. From the resulting plasmid, pRGAP, the RRE-GAPDH transcription unit cassette was released by cleaving with *Xho*I and *Sal*I and cloned into the *Sal*I site of pDBRevM10. pRevRA3.1 was constructed by using the *Mlu*I fragment from pSPORT1 Arg3.1 (8) which contains the entire cDNA of Arg3.1 (3018 bp). This fragment was inserted into the *Mlu*I site of pPGKRRE. From the resulting plasmid pRA3.1, the RRE-Arg3.1 transcription unit cassette was released by cleaving with *Xho*I and *Sal*I and cloned into the *Sal*I site of pDBRevM10. pADRevM10 was constructed by inserting the *Nco*I–*Eco*RV fragment containing RevM10 (see above) between the *Nco*I and *Sma*I site of pACTII (17,18). The screened library was constructed by releasing $>2\times10^6$ independent cDNAs with an average size of 500–2000 bp from a seizure induced hippocampal pSPORT1 library (8) with *Bam*HI and *Sma*I. These fragments were cloned between the *Bam*HI and *Sma*I site of pACTII.

RNA from the nucleus to the cytoplasm (6,7). The method described takes advantage of this specific RNA binding activity to produce a genetic system for detecting and analysing RNA–protein interactions. The basic strategy of the method is illustrated in Figure 1 and specific plasmids used in the following examples are described in Figure 2.

As an example we expressed both the DNA binding domain and transcription activation domain of GAL4 as fusion proteins

with RevM10. No GAL4 activity was detected in yeast expressing these fusion proteins alone or in combination (Table 1). When, however, a recombinant RNA carrying two copies of the RRE was co-expressed with the two RevM10 fusion proteins, functional GAL4 activity was reconstituted (Table 1). A trimeric ribonuclear protein complex, therefore, may be formed in the nucleus, bringing together the GAL4 DNA binding and transcriptional activation domains and recreating functional GAL4 activity. These results demonstrate the ability of a recombinant RNA to interact with two fusion proteins in a tri-hybrid system. The methodology can be extended to detect specific RNA-binding proteins from a library of cDNAs.

Table 1. Tri-hybrid interaction of the RevM10 protein with the RRE RNA sequence

Expressed hybrid-protein	β-galactosidase activity	
	–RRE	+RRE
DB – RevM10 + AD – RevM10	–	+++
DB – RevM10	–	–
AD – RevM10	–	–

The yeast strain CG-1945 was transfected with plasmids expressing hybrid proteins containing the DNA binding domain of GAL4 fused to RevM10 (pDBRevM10 or pRevR2), the transcription activation domain of GAL4 fused to RevM10 (pADRevM10), and a recombinant RNA containing two copies of the RRE (pRevR2). The maps and details of construction of the expression plasmids used are given in Figure 2. Transformants were analysed for histidine independent growth and β-gal expression. Functional GAL4 activity was detected only when all three recombinant molecules were co-expressed. Relevant expressed proteins are indicated. '+RRE' indicates the presence of an RNA containing two copies of the RRE; '–RRE' indicates the absence of the hybrid RNA. Clones that developed visible blue colour within 4 h were registered as +++. If no colour developed after 16 h, clones were recorded as –. Transformants were grown on media selective for the given plasmids. When growth was a measure of GAL 4 activity the medium lacked histidine and contained 30 mM 3-amino-1,2,4-triazole.

Arg3.1 is an immediate early gene induced by synaptic stimulation in the mammalian brain (8,9). The induced transcripts are localised to the soma and dendrites of neurons. Dendritic localisation is believed to be achieved by specific RNA–protein interactions, although the identities of such proteins have not yet been reported. We have used the full length Arg3.1 transcript fused to an RRE, to screen a library of cDNAs expressed as transcription activation domain fusion proteins. The Arg3.1 mRNA was tethered to the DNA binding domain of GAL4 via the interaction between the RevM10 fusion and the RRE present on the recombinant RNA. Any polypeptide capable of interacting with the Arg3.1 message will recruit the transcription activation domain into the ribonuclear protein complex, and reconstitute GAL4 activity.

Several clones were isolated in this way. As an example the data from one clone are shown in Table 2. The identified polypeptide interacts specifically with the Arg3.1 transcript; no interaction is seen with recombinant RNAs containing two RREs or with an RRE fused to the glyceraldehyde-3-phosphate-dehydrogenase (GAPDH) mRNA. These clones and the details of their interaction with the Arg3.1 transcript are the subject of current research.

This tri-hybrid method, therefore, provides a system for studying RNA–protein interactions that has the genetic advantages that the two-hybrid system provides for protein–protein interactions. It can be used to reconstruct *in vivo* a known RNA–protein interaction for further analysis, and as shown, to identify specific RNA-binding proteins expressed from a library of cDNAs. The method could potentially also be used to identify target RNAs of specific RNA-binding proteins. While this manuscript was in preparation, a similar approach by SenGupta *et al.* using a different experimental design was reported (10).

Table 2. Detection of a polypeptide specifically interacting with the dendritically localised transcript Arg3.1

Expressed hybrid-protein	β-galactosidase activity		
	ARG	RRE	GAPDH
DB – RevM10 + AD – protein 1	+++	–	–

The yeast strain CG-1945 was co-transfected with pRevRA3.1 expressing both the hybrid protein DB-RevM10 and a recombinant RNA containing an RRE fused to the Arg3.1 transcript, and with a library of pACTII derived plasmids harbouring brain derived cDNAs expressed as fusions with the GAL4 transcription activation domain. From ~10^5 clones analysed 110 were initially able to grow on medium lacking histidine and containing 30 mM 3-amino-1,2,4-triazole. Of these His^+ clones 89 scored positive in a subsequent filter assay for β-galactosidase activity. These clones were cured of pRevRA3.1 expressing DB-RevM10 and the recombinant RNA, and re-examined for β-galactosidase activity. Twelve library clones were able to activate transcription independently and did not depend on the interaction with the other two hybrids. For 29 clones β-galactosidase activity could be reconfirmed by mating to a Y187 strain carrying the pRevRA3.1 plasmid. The specificity of the Arg3.1 interaction was tested against recombinant RRE RNAs containing an additional RRE (RRE) (pRevR2), or the GAPDH transcript (pRevR-GAP). In these tests 21 clones showed interaction with the Arg3.1 transcript but also with control transcripts and therefore presumably represent general RNA-binding proteins. For the eight remaining clones β-galactosidase expression was specifically dependent on the presence of the Arg3.1 transcript. The data for one clone are shown as an example. Expressed hybrid-proteins are indicated. 'ARG', 'RRE' and 'GAPDH' indicate the presence of the corresponding hybrid transcript in the yeast. Maps of the plasmids used are depicted in Figure 2.

ACKNOWLEDGEMENTS

We thank Drs Stephen J. Elledge, Oktavian Schatz and Brehon Laurent for the generous gift of plasmids and yeast strains; Drs Schatz and Laurent for helpful discussions, Dr Chica Schaller for helpful comments on the manuscript, and Julia Adlivankina-Kuhl for preparation of the figures.

REFERENCES

1 Allen,J.B. *et al.* (1995) *Trends Biochem. Sci.* **20**, 511–516.
2 Phizicky,E.M. and Fields,S. (1995) *Microbiol. Rev.* **59**, 94–123.
3 Fields,S. and Song,O. (1989) *Nature*, **340**, 245–246.
4 Zapp,M.L. and Green,M.R. (1989) *Nature*, **342**, 714–716.
5 Daly,T.J. *et al.* (1989) *Nature*, **342**, 816–819.
6 Fritz,C.C.*et al.* (1995) *Nature*, **376**, 530–533.
7 Stutz,F. and Rosbash,M. (1994) *EMBO J.* **13**, 4096–4104.
8 Link,W. *et al.* (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 5734–5738.
9 Lyford,G.L. *et al.* (1995) *Neuron*, **14**, 433–445.
10 SenGupta,D.J. *et al.* (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 8496–8501.
11 Malim,M.H. *et al.* (1988) *Nature*, **335**, 181–183.
12 Malim,M.H. *et al.* (1989) *Cell*, **58**, 205–214.
13 Harper,J.W. *et al.* (1993) *Cell*, **75**, 805–816.
14 Kang,Y.S.. *et al.* (1990) *Mol. Cell. Biol.* **10**, 2582–2590.
15 Malim,M.H. *et al.* (1990) *Cell*, **60**, 675–683.
16 Fort,P. *et al.* (1985) *Nucleic Acids Res.*, **13**, 1431–1442.
17 Elledge,S.J. *et al.* (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1731–1735.
18 Li,L. *et al.* (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 5012–5026.