Divergent RNA binding specificity of yeast Puf2p

YAHAV YOSEFZON,1 YVONNE Y. KOH,2 JACQUELINE J. CHRITTON,2 AVIGAIL LANDE,1 LIMOR LEIBOVICH,3 LAVI BARZIV,1 CHRISTINE PETZOLD,2 ZOHAR YAKHINI,3,4 YAEL MANDEL-GUTFREUND,1 MARVIN WICKENS,2,5 and YOAV ARAVA1,5

1Department of Biology, Technion–Israel Institute of Technology, Haifa 32000, Israel
2Department of Biochemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706, USA
3Department of Computer Science, Technion–Israel Institute of Technology, Haifa 32000, Israel
4Agilent Laboratories Israel, Petach-Tikva 49527, Israel

ABSTRACT

PUF proteins bind mRNAs and regulate their translation, stability, and localization. Each PUF protein binds a selective group of mRNAs, enabling their coordinate control. We focus here on the specificity of Puf2p and Puf1p of Saccharomyces cerevisiae, which copurify with overlapping groups of mRNAs. We applied an RNA-adapted version of the DRIM algorithm to identify putative binding sequences for both proteins. We first identified a novel motif in the 3′ UTRs of mRNAs previously shown to associate with Puf2p. This motif consisted of two UAAU tetranucleotides separated by a 3-nt linker sequence, which we refer to as the dual UAAU motif. The dual UAAU motif was necessary for binding to Puf2p, as judged by gel shift, yeast three-hybrid, and coimmunoprecipitation from yeast lysates. The UAAU tetranucleotides are required for optimal binding, while the identity and length of the linker sequences are less critical. Puf1p also binds the dual UAAU sequence, consistent with the prior observation that it associates with similar populations of mRNAs. In contrast, three other canonical yeast PUF proteins fail to bind the Puf2p recognition site. The dual UAAU motif is distinct from previously known PUF protein binding sites, which invariably possess a UGU trinucleotide. This study expands the repertoire of cis elements bound by PUF proteins and suggests new modes by which PUF proteins recognize their mRNA targets.

Keywords: Puf proteins; RNA binding proteins; PUM domain

INTRODUCTION

RNA binding proteins (RBPs) control the amount and nature of proteins produced from a gene. In the nucleus, they regulate RNA processing and export in the cytoplasm, they control mRNA stability, translation, and localization (Scherrer et al. 2010; Tsvetanova et al. 2010). A single RBP can bind and control tens to hundreds of functionally related mRNAs (Keene and Tenenbaum 2002; Hieronymus and Silver 2003; Gerber et al. 2004; Hogan et al. 2008; Kanitz and Gerber 2010; Kershner and Kimble 2010), enabling one protein to coordinate many mRNAs underlying a biological event.

PUF proteins, named after the two founding members, Drosophila Pumilio and Caenorhabditis elegans fem-3 binding factor (FBF) (Zamore et al. 1997; Zhang et al. 1997), exemplify this coordinate control. Members of the PUF family contain a PUF repeat region (or PUM-homology domain, PUM-HD), which usually comprises eight tandem PUF repeats, each containing about 40 amino acids. This domain is sufficient for RNA binding and interacts with protein coregulators (e.g., Kraemer et al. 1999; Sonoda and Wharton 1999). PUF proteins regulate mRNA expression by binding to short sequences in the 3′ UTR of their target mRNAs and recruiting other proteins that promote mRNA translational repression, degradation, and localization. The binding sites of several PUF proteins have been identified biochemically and through informatics (Gerber et al. 2004; Bernstein et al. 2005; Opperman et al. 2005; Galgano et al. 2008; Hogan et al. 2008; Morris et al. 2008; Stumpf et al. 2008a; Koh et al. 2009). To date, all binding sites that have been confirmed experimentally comprise 8–11 nt and include a UGUU tetranucleotide core motif. Crystal structures of PUM domains reveal common PUM scaffolds and diverse strategies for achieving RNA specificity (Edwards et al. 2001; Wang et al. 2001; Miller et al. 2008; Wang et al. 2009; Zhu et al. 2009). The overall structures of human Pumilio, C. elegans FBF, and Saccharomyces cerevisiae Pu3p and Pu4p bound to RNA are very similar. Invariably, each PUF repeat folds into a three-helix domain; these domains stack on one another to form a crescent. The...
RESULTS

Prediction of Puf2p binding site

A genome-wide affinity purification study identified mRNAs that were associated with Puf2p in yeast extracts (Gerber et al. 2004). While originally no consensus motif was found among the associated RNAs (Gerber et al. 2004), Hogan et al. (2008) applied a MEME motif-search algorithm (Bailey et al. 2009) and detected a new motif for Puf2p, UAAUAUUW.

We sought to characterize the Puf2p binding site using the DRIM algorithm (discovery of rank imbalanced motifs; see Materials and Methods) (Eden et al. 2007). An advantage of DRIM in the analysis of microarray data is that it does not require prior selection of the affected targets and therefore can be applied to an entire data set without defining an arbitrary threshold of enrichment.

We applied the algorithm to the original data set, seeking 3’ UTR motifs that were most enriched in the immunopurification with Puf2p (Gerber et al. 2004). The most significant motif (P-value < 10^{-48}) found among RNAs associated with Puf2p included two UAAU tetranucleotides separated by a 3-nt linker (Fig. 1A). This motif was also enriched in the 5’ UTRs and the coding regions, but to a lesser extent (P-values of 2 \times 10^{-6} and 3 \times 10^{-10}, respectively). To examine
whether the length of the linker between the two UAAU elements was conserved among the 3′ UTRs, we calculated the statistical enrichment of different linker lengths. All variants contained the two core sites separated by a linker of variable length, from a range of 0 to 10 nt (Fig. 1B). A 3-nt linker was most prevalent among Puf2p targets.

Figure 1C provides a visual rendering of the DRIM output: 4997 genes passed the quality selection criteria in the published microarray experiment (Gerber et al. 2004). The 3′ UTRs of these mRNAs are depicted as gray lines (proportional to their lengths) and are ordered along the y-axis according to their enrichment among Puf2p-associated mRNAs. Occurrences of UAAUNNUAAU motifs within these 3′ UTRs are indicated in red, positioned by distance from the termination codon along the x-axis. The entire list of genes and the position of the motif (also for the 5′ UTR and the coding regions) appear in Supplemental Table 1. The putative Puf2p binding motif is highly enriched among those mRNAs that are ranked high in the list. Optimum enrichment of this motif was attained among the top 336 genes: 69 contained the motif in their 3′ UTRs. Note that in several of the sequences that appear in the top of the list, the consensus motif occurs multiple times in a single 3′ UTR. Among the 69 targets that contained the motif, eight of the targets contained two motifs, and a single target contained four. The statistical signiﬁcance of the enrichment is visible in Figure 1D, where the motif accumulated occurrences (pink) are plotted versus the expected accumulated occurrences in a random data set (black). While in a random data set the occurrences are distributed uniformly (hence a diagonal line), in the actual data set the UAAU motif is enriched in the 3′ UTRs at the top of the list. We refer to the UAAUNNUAAU sequence as the “dual UAAU” motif.

In vitro validation of Puf2 predicted motif

To test binding of Puf2p to the predicted motif experimentally, we performed electrophoretic mobility shift assays (EMSA). In this assay, we combined increasing concentrations of puriﬁed GST-Puf2p PUM domain (Fig. 2A) with a 40-nt RNA derived from the PMP2 3′ UTR that contains the dual UAAU motif (Fig. 2B,E). Puf2p-PUM domain bound the PMP2 RNA with an apparent \( K_d \) of 53 ± 6 nM (Table 1), in the range of other PUF proteins (Bernstein et al. 2005; Koh et al. 2009). This binding is not due to the GST moiety because the GST-Puf5p fusion protein did not bind the dual UAAU motif (see Fig. 6 below). Moreover, partial puriﬁcation of Puf2p PUM domain (after cleavage of its GST moiety) did not change its apparent \( K_d \) (62 nM) (data not shown).

To simplify identiﬁcation of the minimal binding element in the PMP2 RNA, we created a second RNA oligo that disrupts an adjacent UAAU sequence and shortens the RNA to 30 nt. This RNA is referred to as PMP2*. Puf2p bound the PMP2* RNA and its mutated form. EMSA analysis of GST-Puf2p binding as in B. The sequence for the PMP2* RNA oligo and the mutated UAAU motif is shown above the panel. (E,F) Apparent \( K_d \)’s were estimated by ﬁtting a binding curve to the percentage of bound RNA from the EMSA experiments performed in B–D. The curves shown represent Puf2p binding to PMP2 (E), PMP2* (circles), or PMP2* Mut (triangles) (F).
sequence still bound Puf2p in vitro, albeit with a twofold reduced apparent $K_d$ (111 nM). The affinity was further reduced fivefold by mutations that disrupted both UAAU elements (Table 1), indicating that binding to the PMP1 sequence was specific.

### Analysis of association with TAP-Puf2

To test whether the dual UAAU motif was important for binding of Puf2p to endogenous mRNAs, we generated two yeast strains: one that expresses the wild-type (WT) PMP2 transcript and one that expresses a mutant version of the dual UAAU motif (Fig. 3A). Both strains also expressed TAP-tagged Puf2p. Association of the PMP2 transcript with the TAP-tagged Puf2p was assayed by affinity purification and Northern blot analysis (Fig. 3B). As a negative control, the same blots were hybridized with a probe to an mRNA that is not associated with Puf2 (ACT1) (Gerber et al. 2004). No signal was detected in the purified samples even after longer exposures of the blot, excluding nonspecific binding. Furthermore, to normalize for differences between the strains, the Northern blots were also probed for a second Puf2p-associated mRNA, PMP3 (Gerber et al. 2004). A reproducible $\sim$40% decrease in association of the mutated transcript with Puf2p compared with the WT PMP2 construct was observed (Fig. 3B). This confirms that the dual UAAU motif contributes to Puf2p association in vivo.

### Puf2p interactions in vivo

To further define the Puf2p binding site, we used the yeast three-hybrid system (Stumpf et al. 2008b). In this analysis, the interaction between the Puf2p–Gal4 activation domain fusion protein and an RNA sequence results in activation of a $\beta$-galactosidase reporter (Fig. 4A). $\beta$-galactosidase activity in this assay correlates with RNA–protein affinity as measured in vitro over a 10-fold to 100-fold range (Hook et al. 2005).

We first tested a 27-nt RNA sequence derived from the PMP2 3’ UTR containing the dual UAAU motif. The combination of the PMP2 RNA and Puf2p resulted in activation of the $\beta$-galactosidase reporter, while the vector control RNA did not (Fig. 4B). We again eliminated the complication of the additional UAAU by introducing two point mutations, converting that UAAU sequence to ACAU (Fig. 4B). This RNA is referred to as PMP2*.

The resulting RNA bound well and provided a template for all future mutagenesis studies. Transversion mutations in PMP2* in the upstream UAAU (mut1), the downstream UAAU (mut2), or both (mut1-2) eliminated binding (Fig. 4B). We conclude that both UAAU elements are important for the interaction with Puf2p.

To identify the minimal segment of RNA capable of binding well to Puf2p, we prepared RNAs in which all the nucleotides flanking the two UAAU sequences were converted to alternating CA repeats (Fig. 4C). A segment of RNA containing only UAAUAAUUAAU derived from PMP2, flanked by repeating CA sequences, bound as well as the WT sequence. In that background, mutation of a single U from either the 5’ or 3’ end reduced binding modestly; removal of a nucleotide from both ends simultaneously reduced binding to undetectable levels. We conclude that the minimal sequence required for high-affinity binding is the 11-nt sequence UAAUAAUUAAU.

To analyze the effect of the sequence between the two UAAU elements—the “linker”—we tested RNAs in which the spacing was varied from 0 to 6 nt (Fig. 4D). All of the RNAs bound with either WT or slightly reduced affinity. We also varied the identity of the nucleotides in the linker. While, in most cases, these changes did not affect binding, an $\sim$20-fold difference was observed between the WT, WT mutant, and 12 nt. The results demonstrate that the linker is critical for high-affinity binding, and that a specific sequence is required for optimal interaction.

### Table 1: Summary of in vitro binding assay results

<table>
<thead>
<tr>
<th>RNA</th>
<th>Sequence of Oligo</th>
<th>$K_d$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMP2</td>
<td>UUAAUUGCUGAUAUAUAUUAUGUG</td>
<td>53 ± 6</td>
</tr>
<tr>
<td>PMP2*</td>
<td>ATTGACGCAUAUAUAUUAUGUG</td>
<td>54 ± 8</td>
</tr>
<tr>
<td>PMP2* Mut.</td>
<td>ATTGACGCAUAUAUAUUAUGUG</td>
<td>354 ± 51</td>
</tr>
<tr>
<td>PMP2* Linker</td>
<td>ATTGACGCAUAUAUAUUAUGUG</td>
<td>170 ± 30</td>
</tr>
<tr>
<td>PMP2* Linker</td>
<td>ATTGACGCAUAUAUAUUAUGUG</td>
<td>57 ± 16</td>
</tr>
<tr>
<td>PMP1</td>
<td>AAGAAAUGGAAUAUAUAUUAUGUG</td>
<td>111 ± 11</td>
</tr>
<tr>
<td>PMP1 Mut.</td>
<td>AAGAAAUGGAAUAUAUAUUAUGUG</td>
<td>515 ± 56</td>
</tr>
</tbody>
</table>

The indicated oligos were subjected to EMSA with increasing concentrations of purified Puf2p PUM domain, and $K_d$ values were calculated as in Figure 2. The dual UAAU motif is shaded in gray, and mutated bases are in lowercase.
AAU linker to the modified CCA linker (Fig. 4D). The higher affinity of the motif with an AAU linker may be due to the fact that it also contains a dual UAAU motif without a linker; a motif that Pu2 can bind efficiently. These three-hybrid results indicate that Pu2 can efficiently bind motifs with varying lengths of linker. To assess whether this also was true in vitro, we performed EMSAs with PMP2 RNAs of variable linker lengths. We observed a threefold drop in affinity with a 1-nt linker, while a 6-nt linker had no effect on binding (Table 1). We conclude that the presence of the linker is relatively unimportant for binding and that a range of linker lengths and sequences support the interaction with Pu2p.

To test whether binding of Pu2p to the dual UAAU motif requires the same amino acids used in other PUF–RNA interactions, we analyzed mutations in several PUF repeats. Specifically, two conserved residues that contact the RNA in other PUF proteins were mutated to alanine in either two or three adjacent PUF repeats (R1/R2 and R1/R2/R3) (Fig. 5A). The mutant proteins were expressed at similar levels as WT Pu2p yet did not interact with any RNA tested (Fig. 5A,B). We conclude that the same residues that contact RNA in other PUF proteins likely do so in Pu2p.

In addition to the PUM domain, Pu2p contains a predicted RNA recognition motif (RRM) between amino acids 315 and 397. We tested whether this RRM contributes to the specific binding to the dual UAAU motif, by analyzing binding of an in-frame deletion of the RRM. The RRM deletion protein exhibited a similar binding profile as WT Pu2p (cf. Figs. 4B and 5C). We conclude that the RRM does not dramatically affect the RNA-binding specificity of Pu2p to the dual UAAU motif.

Pu1p and Pu2p specificities are distinct from those of other PUF proteins

To compare the RNA-binding specificity of Pu2p to those of other PUF proteins, we tested whether Pu2p bound the sites of other PUF proteins, and conversely, whether the dual UAAU motif bound other PUF proteins. Pu2p did not bind the Puf3p site (COX17 site B) or the Puf4p and Puf5p sites in the HO 3’ UTR in the three-hybrid assay (Fig. 6A). Conversely, Pu3p, Pu4p, and Pu5p failed to bind the dual UAAU motif, but did bind their own cognate sites. We further tested selectivity biochemically using EMSA with yeast Pu5p. Purified Pu5p (the PUM domain) failed to bind detectably to the PMP2 RNA, even at protein concentrations as high as 1000 nM (Fig. 6B). In contrast,
Puf5p efficiently binds its own site (in HO) already at 80 nM (Fig. 6B).

Interestingly, analysis of the binding specificity of the five canonical PUF proteins revealed that Puf1p and Puf2p behave similarly; Puf1p efficiently bound the dual UAAU motif but not the Puf3p and Puf4p/Puf5p motifs (Fig. 6A). Their shared binding specificity is consistent with the fact that yeast Puf1p and Puf2p share a high degree of sequence relatedness (35% identity and 50% similarity), and both interact preferentially with mRNAs encoding membrane-associated proteins (Gerber et al. 2004). DRIM analysis of Puf1p-bound RNAs revealed that they too were enriched in the sequence UAAUNNNUAAU (Supplemental Fig. 1A). A linker length of 3 nt was preferred (Supplemental Fig. 1B), as observed with Puf2p-associated RNAs.

To determine whether the yeast Puf1p binding profile was similar to that of Puf2p, we assayed Puf1p binding to select RNAs in the three-hybrid system. Puf1p bound PMP2 and PMP2* with high affinity (Fig. 7). As with Puf2p, binding of Puf1p to the PMP2* RNA was eliminated when either one or both UAAU motifs were mutated. Taken together, we conclude that Puf1p and Puf2p have a similar binding specificity that is distinct from that of other PUF proteins.

**DISCUSSION**

The DRIM algorithm, which searches for common motifs in a ranked list of genes, predicted a novel sequence motif that is significantly enriched in the 3’ UTRs of mRNAs bound by Puf2p (Fig. 1). This motif, designated the dual UAAU motif, is comprised of two tetranucleotides (UAAU) separated by a 3-nt linker. In vitro and in vivo assays revealed that the PUM domain of Puf2p binds the dual UAAU motif directly and that mutating either UAAU motif leads to a significantly lower interaction, while the distance and the type of bases between them lead to a weaker effect. These empirical measurements yield the consensus UAAUN0–6UAAU, which is different from the canonical PUF binding site (UGURN1–4UA).

Binding of the PMP1 RNA, which contains a 1-nt addition compared with the consensus motif, is reduced twofold. However, this RNA appears to associate with
Puf2p in vivo (Gerber et al. 2004). We infer that variants of the UAAU motif may be bound in vivo, broadening the spectrum of RNAs controlled by Puf2p. A related motif (UAAUAAAUWW) was predicted for Puf2p by Hogan et al. (2008) using a different computational method. While the enrichment of the dual UAAU motif is more significant (mHG $P$-value of $10^{-21}$), we cannot exclude the possibility that the UAAUAAAUW motif is also functional in vivo, thereby further increasing the complexity of Puf2–RNA interactions. Moreover, Puf2 (and Puf1) have a predicted RRM in addition to the PUM domain. Our results indicate that this domain is not important for binding of Puf2 to the dual UAAU motif. However, considering the broad and significant involvement of RRM's in RNA metabolism (Clery et al. 2008), this motif is likely to further expand the targets of Puf2 by binding to yet-to-determined sequences.

Puf2p diverges from the simplest condition, exemplified by human Pumilio, in which one PUF repeat interacts with one RNA base (Wang et al. 2001). Our results with Puf2p emphasize the diversity of PUF recognition in two ways. First, the binding site does not possess the UGUR sequence seen in all known targets of PUF proteins. Second, Puf2p recognizes 11 nt, while the PUM domain of Puf2p contains only six recognizable PUF repeats (Gerber et al. 2004). Other PUF proteins with eight canonical repeats recognize longer binding sites, either due to base flipping or additional nucleotide binding pockets (Bernstein et al. 2005; Opperman et al. 2005; Zhu et al. 2009). Since the Puf2p binding site possesses two copies of the UAAU sequence, the protein may bind as a dimer. Similarly, C. elegans PUF-5 binds UGURYYYYUGU, with two copies of the UGU sequence at each end of the core recognition site, and is analogous in that respect (Stumpf et al. 2008a). In neither instance do we detect any evidence of sequential binding of two monomers in gel shift studies, but the proteins could exist as homodimers in solution.

A linker length of 3 nt between the two UAAU elements is preferentially enriched among the mRNAs immunopurified with Puf2p ($P$-value of $10^{-48}$ vs. $5 \times 10^{-18}$ for a gap of 2 nt and $7 \times 10^{-14}$ for the empty gap; two and zero are the most enriched gap lengths after three) (Fig. 1B). Similarly, the linker sequence AAU was enriched among those motifs containing a 3-nt linker. However, the presence of the linker element had only a marginal effect on binding of Puf2p, as did variations in its length or the identity of the linker nucleotides. The enrichment of a 3-nt linker and the AAU sequence among mRNAs associated with Puf2p and Puf1p in vivo likely reflect additional biological functions beyond binding of these proteins. We suggest that the linker spaces between two sets of contacts with UAAU elements. This situation may be analogous to that seen in other PUF proteins. For example, several PUF proteins “flip” bases away from the protein (see Introduction). The identities of flipped bases can vary without affecting binding. We suggest two models of Puf2p binding. In the first, a single Puf2p molecule binds the dual UAAU element, and linker nucleotides in the complexes are flipped away from the protein. In the second, two molecules of Puf2p bind, each one interacting with a single UAAU.

We identified the Puf2p motif in 122 genes of the approximately 5000 analyzed S. cerevisiae genes (<3%), yet >50% of these (69 genes) appeared among the mRNAs that are highly associated with Puf2p (top 336 genes). GO term analysis for these 69 genes revealed enrichment for genes encoding membrane proteins ($P$-value = 0.001). This enrichment, however, is not necessarily due to the presence of a UAAU motif in their 3′ UTRs, because similar enrichment is observed also for Puf2p-associated mRNAs that do not have the UAAU motif. 3′ UTRs with multiple occurrences of the UAAU motif tend to be ranked higher (e.g., 11 of the 12 genes with more than one motif are ranked among the top 336 genes). The presence of more than one motif in the PMP2 3′ UTR may explain why mutations disrupting only one dual UAAU motif did not completely abolish Puf2p association (Fig. 4B). Multiple iterations of binding sites within 3′ UTRs is a well-established contributor to RBP–RNA interactions (Macdonald and Struhl 1988; Kiszalski et al. 1994; Chen and Shyu 1995; Deshler et al. 1997; Loya et al. 2008). Interestingly, DRIM analysis revealed enrichment of the dual UAAU motif also in 5′ UTR and ORF regions of the mRNA. Most of the mRNAs with a binding site within their 5′ UTR or ORF do not have a binding site within their 3′ UTR. While the $P$-values were lower than in the 3′ UTR and binding was not verified experimentally, these results further exemplify the complexity and diversity of Puf protein binding.

Puf1p appears to have a sequence specificity similar to that of Puf2p, based on both computational analysis of the RNAs with which it associates, as well as direct three-hybrid tests. In immunopurification/microarray studies (Gerber et al. 2004), 90% (36 of 40) of mRNAs bound to Puf1p are also bound by Puf2p, consistent with the overlap in their specificities. However, 75% of the mRNAs that bind Puf2p do not detectably associate with Puf1p. If the immunopurification results reflect the in vivo situation, it would imply that additional factors contribute to the specificity of Puf2p.
and enable it to bind those mRNAs. Those might include differences in RNA binding preferences that we have not detected, or collaboration with other protein factors.

In summary, Puflp and Puflp bind RNA elements that differ from other described PUF binding sites. The nature of its recognition by Puf2p and, specifically, which bases and Puf2p repeats are important for binding will require structural analysis. Regardless, the Puf2p–RNA interaction expands the repertoire of PUF proteins and diversifies their array of RNA specificities.

MATERIALS AND METHODS

Computational analysis of sequences (DRIM)

DRIM (discovery of rank imbalanced motifs) is a software application that identifies sequence motifs in lists of ranked DNA and RNA sequences (Eden et al. 2007; http://bioinfo.cs.technion.ac.il/drim/; the mHG statistics to discover motifs that are statistically enriched at the top of a ranked list of sequences). This statistical approach uses threshold optimization in this ranked list. Briefly, the flexible threshold process computes a hypergeometric tail (HGT) at each possible cutoff of the ranked list and then optimizes the threshold to yield the most significant enrichment of the motif under consideration. Exact P-values are obtained as described in Eden et al. (2007). Because DRIM was originally developed to identify enriched motifs in transcription factor target sequences (Eden et al. 2007), the algorithm first had to be adapted to search for enriched motifs in RNA sequences as well. One important feature of the new version of DRIM is the option to differentially weight multiplicities of motif occurrences. New features of the RNA version of DRIM are described at http://bioinfo.cs.technion.ac.il/drim/.

We applied DRIM to analyze a list consisting of 4997 S. cerevisiae 3′ UTR sequences, ranked according to Pufl binding, as reported by Gerber et al. (2004). The yeast 3′ UTR sequences were obtained from UCSC (S. cerevisiae genes), according to their experimentally defined lengths (Nagalakshmi et al. 2008). The motifs reported herein were thus found to be enriched in the UTRs ranked at the top of this experimentally derived list.

Protein purification

The PUM domain of Puflp (amino acids 361–832) was cloned into pGex6P1 as a GST fusion protein with a 6× His tag, and transformed into BL-21 Escherichia coli cells. Bacteria (0.5 L) were grown on LB to logarithmic phase (OD600 = 0.6–0.9) and induced with 0.1 mM IPTG and growing the cells overnight at 16°C. Cells were harvested by centrifugation and washed with 50 mM Tris–HCl at pH 8.0 and 10% sucrose. Cell pellets were frozen in liquid nitrogen and resuspended in 25 mL of Lysis Buffer (20 mM HEPES-KOH at pH 7.4, 0.5 M NaCl, 5 mM DTT, 0.02% Tween 20, 1 mM PMSF, 1 μg/mL leupeptin, 1 μg/mL Pepstatin). The cells were lysed with lysozyme (final concentration of 0.5 mg/mL). For 30 min on ice and were frozen in liquid nitrogen. After thawing the cells, 2 mM MgCl2 and 10 mg/mL DNase I were added and incubated for 30 min on ice. The lysate was frozen in liquid nitrogen, thawed, and clarified by centrifugation at 18,000 rpm for 40 min at 4°C. Because the Puf2-PUM domain was insoluble and found to accumulate in inclusion bodies, the pellet that contains Pufl-PUM was dissolved in Binding buffer (6 M urea, 15 mM imidazole, 0.5 M NaCl, 20 mM Tris–Cl at pH 8) and subjected to centrifugation at 18,000 rpm for 40 min at 4°C. The supernatant, which contains the dissolved proteins, was filtered through a 0.45-μm filter and loaded on an Ni-NTA Agarose bead column pre-equilibrated with Binding buffer. Proteins were eluted with 6 mL of Elution buffer and were dialyzed into decreasing concentration of urea in 1× PBS for the refolding process. Proper refolding was assessed by the ability of the GST moiety to bind glutathione Sepharose resin. Purity and concentration of the protein were estimated by SDS gel electrophoresis.

The PUM domain of Puflp (amino acids 126–626) was cloned into pGex6P1. GST-Puflp protein purification was performed by the same method described above except that after the centrifugation (18,000 rpm for 40 min at 4°C) the supernatant, which includes the protein, was passed through glutathione Sepharose resin (pre-equilibrated with Lysis buffer). The proteins were eluted from the glutathione Sepharose resin with Elution buffer, and protein purity and concentration were estimated by SDS gel electrophoresis.

EMSA

RNA oligonucleotides (40 pmol; Dharmacon) were end-labeled with [γ-32P]ATP by T4 polynucleotide kinase for 1 h at 37°C and purified with G-50 columns. For binding reaction, 20 fmol of labeled RNA was incubated for 1 h at room temperature with varying concentrations of GST-Pufl or GST-Pufl5 in EMSA binding buffer (Koh et al. 2009). Loading dye (6% glycerol, 0.06% Bromophenol Blue) was added to each reaction, and samples were loaded on a pre-run, nondenaturing 5% polyacrylamide gel (Bio-Rad). Gels were resolved in 1× TBE buffer at 100 V for 35 min at 4°C and were exposed to a PhosphorImager screen. Signals were quantified with the TINA program. Apparent Kd values were calculated using GraphPad Prism software (version 5.02), by fitting a binding curve to the percentage of bound RNA (which was calculated from densitometry readings of the shifted and unshifted RNA in each reaction) against the protein concentration. Apparent Kd values are represented as an average of two to four repetitions, with standard errors shown in Table 1.

Constructs for yeast three-hybrid assay

A cDNA encoding S. cerevisiae Puflp (GenBank accession number NM_001184139), comprising amino acids 1–1016, was cloned into an activation domain vector, pGADT7. The PUF repeat mutants used in Figure 5 were created by site-directed mutagenesis. Amino acid changes were as follows: Puflp R1/R2 mut, N591A, T592A, T627A, and W628A; Puflp R1/R2/R3 mut, N591A, T592A, T627A, W628A, N664A, and Y665A. For specificity testing, cDNAs encoding S. cerevisiae Puflp (NM_001181748) amino acids 1–1091, Pufl5 (NM_001181833) amino acids 511–879, Puflp (NM_001180879) amino acids 534–888, and Pufl5 (NM_001181043) amino acids 26–860 were each cloned into pGADT7.

A 27-nt region of the PMP2 3′ UTR was cloned into the Xmal and SpHl sites of the p3HR2 vector (Stumpf et al. 2008a). The 23rd nucleotide in that region was changed from U to A to avoid five consecutive uridines that will terminate RNA transcription by Pol III (Stumpf et al. 2008b). All other PMP2 RNA constructs were created the same way. The Puflp binding site B of COX17 3′ UTR and the first 50 nt of the HO 3′ UTR were similarly cloned into the p3HR2 vector.

1486 RNA, Vol. 17, No. 8
Yeast three-hybrid assay

Three-hybrid assays were performed in the YBZ-1 yeast strain as described previously (Hook et al. 2005) with the following modifications: 50 μL of a saturated culture was diluted in 1 mL of selective media and allowed to grow for 3 h to reach an OD_660 of 0.1 to 0.2. β-Galactosidase activity was normalized to an OD_660 of 0.1; values are represented as an average of three repetitions, with standard deviations shown in each figure.

Yeast protein extracts and Western blotting

The relative expression levels of Puf2p protein in the yeast three-hybrid assay were determined using Western blot analysis. Yeast were transformed with the pGADT7-Puf2p (amino acids 1–1016) plasmid as well as the RNA expression plasmid p3HR2. Cultures were grown to saturation in selective media at 30°C and then diluted into fresh selective media and grown for 4 h at 30°C. One milliliter of log-phase culture was pelleted, resuspended in 100 μL of ESB (2% SDS, 80 mM Tris at pH 6.8, 10% glycerol, 1.5% DTT), heated for 3 min to 100°C, and then lysed using glass beads and vortexing for 5 min. Extracts were then heated for 1 min to 100°C and stored at −20°C.

For Western blot analysis, 10 μL of yeast extract from each expression construct was loaded onto an SDS polyacrylamide gel. After electrophoresis, proteins were transferred onto a nitrocellulose membrane and then blocked with TBST + milk (1% milk [w/v]). To detect the HA-tagged Puf2p fusion proteins, we incubated the blot with HA-11 antibody (Covance) at a 1:100 ratio in TBST + milk for 1 h at room temperature. The blot was then washed with TBST (1× TBS, 0.1% Tween 20) three times and then probed with the secondary antibody, HRP-conjugated goat anti-mouse, for 1 h at room temperature in TBST + milk. After the blot was washed three times with TBST, HRP was detected using the Pierce ECL Western Blotting Substrate kit.

RNA affinity purification

Exponentially growing cells (1 L) were collected by centrifugation at 4000 rpm for 4 min at 4°C and resuspended and washed two times in 40 mL of TAP–Buffer A (20 mM Tris–Cl at pH 8, 140 mM KCl, 1.8 mM MgCl₂, 0.1% NP-40, and 0.01 mg/mL heparin). Cells were then resuspended in 5 mL of TAP–Buffer B (buffer A + 0.5 mM DTT, 1 mM PMSF, 0.5 μg/mL leupeptin, 0.8 μg/mL Pepstatin, 0.2 mg/mL heparin, 100 U/mL RNasin, and 20 U/mL DNase I), and aliquots of 1 mL were transferred to screw-capped microfuge tubes supplemented with 1 mL of chilled glass beads. Cells were lysed by two rounds of 90-sec pulses in a bead beater. The lysate was transferred to a clean microfuge tube and cleared from cell debris by centrifugation at 10,000 rpm for 10 min at 4°C. For unpurified control (“Input”), one-tenth of the lysate was precipitated and one-tenth was taken for Northern analysis. Affinity purification was performed in 15-mL tubes, by incubating 4.5 mL of lysate with 400 μL of IgG beads (pre-equilibrated with TAP–Buffer A) for 2 h at 4°C with constant rolling. Bound material was washed by four rounds of centrifugation at 3000 rpm at 4°C and addition of 4 mL of ice-cold TAP–Buffer C (20 mM Tris–Cl at pH 8, 140 mM KCl, 1.8 mM MgCl₂, 0.01% NP-40, 0.5 mM DTT, 10 U/mL RNasin). Proteins and RNA were eluted by TEV-protease cleavage for 2 h at 16°C in a TEV reaction buffer (provided by the manufacturer). The eluted RNA was extracted by phenol–chloroform, subjected to ethanol precipitation, and resuspended in 50 μL, and one-third was subjected to Northern analysis (Eldad et al. 2008).

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank Gal Shaul and Naama Hazan for their role in developing DRIM for RNA and for data analysis. We thank Israel Steinfeld for providing access to mHG implementation. We also are grateful for the assistance of Laura Vanderploeg with the figures. We appreciate help and discussions from the members of the Arava and Wickens laboratories. This work was supported by ISF grant number 1193/09 to Y.A., NIH grants GM32892 and GM50942 to M.W., the A*STAR Program (to Y.K.), an NIH Genetics Training Grant (to J.C.), and the Google Anita Borg Memorial Scholarship (to LL.).

Received March 1, 2011; accepted May 5, 2011.

REFERENCES


