Published in final edited form as:

Biochim Biophys Acta. 2009 ; 1789(6-8): 469–476. doi:10.1016/j.bbagra.2009.05.003.

Compensatory Mutations in the L30e Kink-Turn RNA-Protein Complex

James J. Schweppe^{a,b}, Chaitanya Jain^c, and Susan A. White^a

^aDepartment of Chemistry, Bryn Mawr College, Bryn Mawr, PA 19010, USA

Abstract

The *S. cerevisiae* ribosomal protein L30e is an autoregulatory protein that binds to its own pre-mRNA and mature mRNA to inhibit splicing and translation, respectively. The L30e RNA binding element is a stem-asymmetric loop-stem that forms a kink-turn. A bacterial genetic system was designed to test the ability of protein variants to repress the expression of reporter mRNAs containing the L30e RNA-binding element. Initial screens revealed that changes in several RNA nucleotides had a measurable effect on repression of the reporter by the wild type protein. RNA mutants that reduce repression were screened against libraries of randomly mutagenized L30e proteins. These screens identified a glycine to serine mutation of L30e, which specifically restores activity to an RNA variant containing a U that replaces a helix-capping G. Similarly, an asparagine to alanine mutation was found to suppress a substitution at a position where the L30e RNA nucleotide extends out into the protein pocket. In addition, a compensatory RNA mutation within a defective RNA variant was found. The identification of these suppressors provides new insights into the architecture of a functional binding element and its recognition by an important RNA-binding protein.

Keywords

RNA-protein interactions; Kink-turn RNA; ribosomal protein; two-plasmid screen; suppressor mutations; L7Ae Superfamily

1. Introduction

The yeast (*Saccharomyces cerevisiae*) ribosomal protein L30e is an autoregulatory protein that binds to its spliced or unspliced transcript to inhibit translation or splicing, respectively [1-8]. L30e is an essential protein thought to reside at the ribosomal subunit interface, and autoregulation confers a survival advantage in yeast [9,10]. Biochemical and structural work have shown that the RNA binding site is comprised of two helical stems separated by a three-nucleotide bulge that enables the stems to come together at an acute angle [11-13]. Related motifs found in bacterial and archaeal ribosome structures revealed similar tight superimposable RNA bends and notable sequence similarities [14]. This bent RNA with its

Address correspondence to: Dr. Susan A. White, Department of Chemistry, Bryn Mawr College, Bryn Mawr, PA, 19010, USA. Tel. 610-526-5107; Fax 610-526-5086; E-mail: swwhite@brynmawr.edu.

^bDepartment of Chemistry and Biochemistry, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA

^cDepartment of Biochemistry and Molecular Biology, University of Miami Miller School of Medicine, 516 Gautier Medical Research Building, 1011 N.W. 15th Street, Miami, FL 33136, USA

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

characteristic sequence was termed the kink-turn, or the K-turn motif, and has also been found in archaeal RNAs that guide the covalent modification of ribosomal RNAs [15].

The hallmark features of K-turn RNAs are a canonical stem, three unpaired nucleotides, and a non-canonical stem having two G:A pairs adjacent to the bulged nucleotides (Fig. 1A)[14]. The importance of the G:A pairs in the L30e complex was underscored by an *in vitro* selection experiment in which the four purines were found to be nearly invariant (Fig 1B)[16]. In the L30e RNA-protein complex, this structure is further stabilized by purines that stack atop the two stems and an A-minor hydrogen bonding interaction between the two stems. RNA K-turns are most often complexed with proteins, and several of these proteins are homologous to L30e and members of the L7Ae protein class [14,17,18]. These proteins contain alternating regions of alpha helices and beta strands that fold into a central, four-stranded beta sheet flanked by two perpendicular pairs of alpha helices. It is primarily the three protein regions between the secondary structure elements on one face of the protein that contact the RNA. The yeast L30e primary structure is shown in Figure 1C, and the highly conserved RNA interface regions from a variety of organisms are shown in Figure 1D [19].

In order to characterize the interactions between the L30e RNA and protein, we developed a genetic screen to identify L30e protein suppressors that restore high affinity binding to mutant RNAs. Because of its autoregulatory nature, we feared overexpression of L30e would not be possible in yeast and opted instead to carry out the genetic screen in bacteria. Two plasmids, one bearing the gene for the yeast L30e protein and the other bearing the L30e K-turn sequence inserted just upstream of the *lacZ* reporter gene, were transformed into bacteria (Fig. 2) [20-26]. The L30e protein is expressed from the repressor plasmid and binds to its cognate RNA sequence in the reporter mRNA. In the *in vivo* screen, strong RNA-protein binding corresponds to greatly reduced expression of the reporter enzyme, β -galactosidase, and thus parallels one of the L30e autoregulatory activities in yeast.

Finding a protein mutation that compensates for an RNA mutation that weakens binding is an example of a gain-of-function mutation. An L30e protein variant bearing such a mutation would be termed a suppressor protein. The most straightforward explanation for such a compensatory change would be the existence of a direct contact between the mutated RNA nucleotide and the protein amino acid. Random and directed protein mutational strategies were employed to identify such L30e mutants. These screens resulted in the identification of two L30e suppressors, each containing a single amino-acid substitution that could recognize specific RNA mutants. A second-site RNA suppressor that was able to correct the defective binding of a mutant RNA was also identified. These findings provide new insights into the interaction between L30e and its RNA-recognition element. Throughout this work, protein residues will be referred to by their three letter abbreviations to distinguish them from RNA variants.

2. Materials and Methods

2.1 Bacterial strains and plasmids

All experiments were carried out in *E. coli* strains WM1 and WM1/F' that were made competent by the CCMB method [21,23,26]. The WM1/F' strain was used so that toxic L30e expression could be induced by addition of IPTG. Bacterial strains and parent plasmids for cloning, pACYC184, pREV1 and pLacZ-rep, were described by Jain and Belasco [21]. The DNA sequence encoding yeast L30e was amplified by PCR from pMalc-L32 provided by Dr. J. Warner [5]. The amplified fragment bearing *NdeI* and *Sall* restriction sites on the ends was inserted into the similarly prepared pREV1 plasmid. This plasmid, named pRPL30, is a medium copy number plasmid that confers chloramphenicol resistance. pLacZ-LBE was created by inserting the L30e RNA binding element (LBE) into pLacZ-rep using the unique *HindIII* and *KpnI* restriction sites immediately upstream of the Shine-Dalgarno (S/D) region of the *IS10*-

LacZ fusion gene. This reporter plasmid is a high copy number plasmid and confers resistance to ampicillin (Fig. 2).

2.2 Site-directed mutations in reporter plasmid

Nucleotides in the NC stem and the unpaired loop nucleotides were systematically mutated to find point mutations that moderately reduced repression. All mutations to the L30e RNA recognition sequence in the reporter plasmid were generated using Stratagene's Quik Change kit using DNA oligomers synthesized by Invitrogen. All plasmids were sequenced at the University of Pennsylvania DNA Sequencing Facility.

2.3 Random and targeted mutations in repressor plasmid

Error prone PCR was used to produce a library of random pRPL30 mutants [27,28]. In brief, the 315 base pair *RPL30* gene was amplified using 50 mM Tris HCl, pH 9.0, 50 mM NaCl, 7.5 mM MgCl₂, 0.2mg/mL BSA, 200 μ M dGTP and dATP, 1 mM dCTP and dTTP, 0.5 Units Taq polymerase, 0.33 μ M sense and anti-sense primers, and 50 ng (20 fmol) pRPL30 with 0.5 mM MnCl₂ added last. Following 22 amplification cycles, the amplified product was ligated into the pREV1 vector as before. Based on the sequencing of about 60 transformants in XL-1 cells (Stratagene) and potential suppressors, an average error rate of 3 percent was estimated. Targeted mutagenesis, based on the structure of the L30e RNA-protein complex, was performed for Lys28, Asn47, Pro49, Arg52, Asn74, Phe85, and Gly88. In each case, the corresponding codons were randomized.

2.4 Screening and measurement of Repression Ratios

Bacteria containing a reporter plasmid and a mutagenized protein expression library were grown on LB agar plates containing ampicillin, chloramphenicol, IPTG, and X-gal. Expression of the repressor plasmid is regulated by an IPTG-dependent promoter, and the amount of IPTG was adjusted to allow maximal L30e protein expression without being toxic to the host bacteria. Based on the results obtained with similar systems, strong binding of the L30e protein to the reporter transcript would result in less β -galactosidase activity, visualized as white colonies, whereas weaker protein binding would result in more β -galactosidase production and blue colonies [21-23]. For each RNA mutant approximately 5000 potential suppressor colonies were visually screened.

For quantitative measurements, colonies displaying the phenotype suggestive of suppressor function were grown to saturation in duplicate and used to inoculate fresh cultures that were then grown to mid-log phase. The cultures were harvested, cooled on ice, and assayed for β -galactosidase activity, as described by Miller [29]. The repression ratio is the average β -galactosidase activity of the reporter in the absence of an RNA-binding protein divided by the β -galactosidase activity of the reporter in the presence of L30e or its variants. Mutant L30e proteins exhibiting an increase in repression ratio over wild type L30e protein suggested suppression of the loss-of-function RNA mutation. As potential suppressors were identified, plasmid DNA from the remaining bacterial cells was purified and sequenced.

2.5 In vitro synthesis of RNAs and proteins

Partially double stranded DNA templates were annealed and used as templates for *in vitro* RNA transcription using Ambion's MegaShortscript Kit. *In vitro* transcribed RNAs were purified by denaturing gel electrophoresis. RNAs were visualized by UV shadowing, extracted from the gel, and eluted by soaking the gel slice overnight in 0.5 M NaOAc and 1 mM EDTA. Following centrifugation, the supernatant containing the RNAs was decanted and then concentrated by ethanol precipitation. RNA pellets were washed with 70 percent ethanol and then resuspended in TE buffer (pH 8). RNA concentrations were measured spectroscopically, and 25 pmoles

were dephosphorylated with calf intestinal alkaline phosphatase (NEBiolabs) and end-labeled with γ - ^{32}P ATP and T4 polynucleotide kinase. Radiolabeled RNAs were gel purified as described above except that transfer RNA was used as a carrier to aid ethanol precipitation and RNA bands were visualized by autoradiography.

L30e proteins were expressed from pMalc-L30e plasmids in JM109 *E. coli* strains as maltose-binding fusion proteins. Protein mutants were generated via site-directed mutagenesis of the pMalc-L30e plasmid as previously described. Protein expression was induced using 100 μM IPTG for several hours at 37 °C. Fusion proteins were purified from 250 mL cultures using 5 mL gravity flow amylose columns (NEBiolabs). Protein purity was estimated from Coomassie-stained SDS gels, and concentrations were measured using the Bradford assay with BSA as a protein concentration standard (BioRad).

2.6 In vitro measurement of binding affinities

Dissociation constants, K_d , were determined using purified maltose binding protein fusion proteins, MBP-L30e, in nitrocellulose filter binding assays and qualitatively verified in electrophoretic mobility shift assays (EMSA). These assays were performed in the following solution conditions: 75 mM KCl, 30 mM Tris, pH 8, 2 mM MgCl_2 , 1 mM DTT, 500 ng/ μL BSA, 40 ng/ μL tRNA, 0.05 unit/ μL RNase inhibitor (5, 11). RNAs were renatured in 350 mM KCl, 30 mM Tris (pH 8) and 10 mM DTT, by heating to 60°C and slowly cooling to room temperature. Fusion protein titration reactions were incubated at room temperature for 10-20 minutes and contained a constant, subnanomolar, amount of radiolabeled RNA. Filter binding reactions were loaded onto nitrocellulose filters soaked in the binding buffer minus the RNA or protein components. Filters were rinsed with 100-200 μL cold binding buffer, and the retained radioactivity was counted in Ecolume (ICN) scintillation fluid. EMSA reactions were electrophoresed at 2-8 °C on 10% (29:1 acrylamide: bisacrylamide) native gels in 0.5XTBE (50 mM Tris, 50 mM boric acid, 1 mM EDTA). 20 μL samples containing 10% glycerol and the reagents described above were run for 4-6 hours at 100 V. Filter binding data were plotted using Kaleidagraph (Synergy Software) and fit to a hyperbolic binding isotherm to determine the binding constant, K_d .

2.7 Structure and Graphics

The L30e RNA-protein complex was visualized using PyMol and the coordinate file 1T0K.pdb [13,30]. *In silico* single residue mutations were made and energy minimized locally. The program ENTANGLE was used to aid in cataloguing interactions between the RNA and protein [31].

3. Results

Preparation of the Two-Plasmid System

A bacterial two-plasmid system has been previously described for studying RNA-protein interactions [21]. In this system, a first plasmid contains a modified *lacZ* reporter gene with an RNA-binding element located close to the translation initiation region. A second plasmid expresses the RNA-binding protein that recognizes this RNA element. When the protein is bound to its cognate RNA located on the reporter mRNA, ribosomal assembly is obstructed, which results in the repression of reporter translation. One particularly valuable use of this system has been to identify protein variants that can bind to mutant RNA targets, hence providing detailed information about specific protein-RNA contacts [20,21].

This two-plasmid strategy was applied to the L30e protein and its target RNA. Although the initial two-plasmid constructs produced very low repression ratios, following optimization of

the reporter construct and conditions for protein expression, a 450-fold repression of the reporter construct by L30e was achieved (Fig. 2 and Table 1).

Previous work showed that RNA LBE point mutants could have repression ratios equal to one when paired with wild type L30e protein in the two-plasmid system [32]. Such mutants are expected to bind to the protein with extremely low affinity, if at all. Likewise, single amino acid changes are capable of drastically lowering the repression ratio. For example, binding assays indicate that the L30-Phe85Ala binds RNA very weakly and when tested in the two-plasmid system, this pair yields a repression ratio of 9.3 [33,34]. Thus, L30e RNA and protein sequence variants that were known to have high or low binding affinities behaved as expected in the two-plasmid assay.

Identification of candidate RNAs for suppression

In order to identify protein mutants that suppress deleterious RNA mutations, we first needed to identify L30e RNA mutations that decrease the wild type protein repression ratio. The goal was to disrupt individual RNA-protein contacts and not the RNA's ability to fold into a kink-turn. We thus decided that the ideal RNA mutation would be one that disrupted protein binding to a moderate, but not a catastrophic, extent. Systematic RNA mutations of the L30e RNA binding site, LBE, were made, which complemented RNAs studied in earlier SELEX experiments or *in vitro* binding experiments [16,32]. Of the four nucleotides that comprise the two G:A pairs of the non-canonical stem, only one, G11, may be changed without producing a repression ratio of one, which reflects a total loss of protein binding (Table 1). In addition, changing the polarity of A:G and G:A pairs, singly or in tandem, disrupted repression completely. Consistent with SELEX results, mutations in A12, G58, and A59 are not tolerated and have repression ratios equal to one, whereas G11A and G11U RNA mutants have repression ratios of 89 and 16, respectively. Since these repression ratios are between 1 and 450, these RNAs presumably have a moderate affinity for wild type L30e protein and were considered as candidates for a protein suppressor screen.

Mutations of the unpaired nucleotides of the kink-turn RNA produced more varied results in screening against wild type L30e protein (Table 1). Adenosine 55, the most variable position in the SELEX experiment, retained strong repression when mutated to U or C, but the A55G variant had a repression ratio of 59, indicative of moderate repression. Likewise, the G56U mutant, has a moderate repression ratio of 10, but G56A and G56C show very high and low repression ratios, respectively. Finally, A57U and A57C retain near wild type repression ratios, whereas A57G has a moderate repression ratio of 36. In the two groups of stem or loop RNA mutations studied, there are several with intermediate repression ratios, and thus these RNAs are presumed to be capable of forming a kink-turn that has compromised binding to wild type protein. These candidate RNAs that retain moderate repression ratios are G11A, G11U, A55G, G56U, and A57G and are written in bold in Table 1. Interestingly, the mutation of G56 to adenine negates the A55G mutation and restores the repression ratio binding to its wild type level. Thus, this is an RNA suppressor.

We were also interested in determining how well the repression ratios correlate with the binding of L30e to the mutant RNAs. For this purpose, several mutant RNAs were transcribed and their affinity for purified L30e was determined (Experimental Procedures). Examination of Table 1 shows that there are exceptions to the expected inverse relationship between repression ratios and binding affinities. While it is true that all RNAs having a repression ratio of one have weak binding affinities of at least 500 nM, the converse is not true. In most cases, higher repression ratios did indeed correspond to stronger binding affinities with two exceptions. A55G RNA has near wild type affinity for the L30e protein but a low repression ratio. Although, based on its low repression ratio, this RNA was selected as a candidate for suppressor screening, perhaps it is not surprising that no suppressor was found since filter binding indicated strong protein

binding. The G56A RNA had a large repression ratio, but relatively weak protein binding. Thus, the repression ratio depends not only on the L30e protein's affinity for the RNA leader sequence, but also on structural and geometric factors that allow the bound protein to interfere differently with reporter mRNA translation.

Screening protein mutants against candidate RNA

The L30e protein coding region was amplified by error-prone PCR and cloned into the pREV1 vector. The candidate RNAs containing the moderately defective mutants, G11A, G11U, A55G, G56U, or A57G, were then screened against the random library of L30e protein mutations. To test the efficacy of the screening procedure, preliminary studies using a wild-type reporter and a defective protein library spiked with a small proportion of wild-type L30e cloned indicated that such clones could indeed be recovered. Transformation conditions were optimized so that each candidate RNA in the two-plasmid screen produced 200-300 colonies per LB plate and generally more than 5000 colonies were examined for each two-plasmid screen. We expect that number of colonies screened would have included most of the amino-acid changes that are possible through single base-pair mutation at any codon position. Several transformant colonies that were lighter blue than control transformants expressing wild-type L30e were picked and subjected to duplicate Miller Assays. Such colonies may be expected to harbor L30e mutants that can bind to the RNA variants better than wild-type L30e.

Screening against G11A, G11U, A55G or the A57G RNAs failed to identify a randomly mutated protein suppressor that reproducibly increased the repression ratio. However, using the G56U variant, a Gly88Ser candidate suppressor was identified. In Miller assays the Gly88Ser mutation increased the repression ratio to the G56U RNA mutant by a factor of over four from 12.6 to 53 (Table 2). This mutant is specific for the G56U RNA mutation because enhanced binding was not observed when Gly88Ser was tested with wild type RNA. In additional controls, repressor plasmids bearing either the wild type or the Gly88Ser sequence were unable to repress reporter plasmids containing no LBE, or LBE with deletions of G56 or A57 (data not shown). Thus, the G56U mutation is suppressed specifically by the Gly88Ser mutation. Furthermore, replacing the serine with alanine, which removes serine's hydroxyl group, largely removes this suppression, though the Gly88Ala mutation was still slightly better than wild-type L30e at repression of wild type and G56U LBE RNAs. In order to find additional suppressors at position 88, the three nucleotides of this codon were completely randomized and this library was re-screened against the G56U reporter. Confirming the original findings, several isolates of Gly88Ser were found, along with an additional suppressor, Gly88Cys. In summary, replacement of the glycine at position 88 with a serine, alanine, or cysteine increases the ability of the L30e protein variant to repress expression β -galactosidase when the G56U mutation is present in the transcript leader sequence (Fig. 2).

Targeted protein mutants against candidate RNAs

Although the screening using a PCR mutagenized library did not yield suppressors to the G11A, G11U, A55G or the A57G RNAs, in order to not overlook potential suppressors, libraries of targeted protein mutations were constructed and screened against the appropriate RNA mutants. This can be important because PCR mutagenized libraries may not contain critical amino-acid changes that require two or three positions of a codon to be altered. The L30e-LBE co-crystal shows an interaction between the RNA position G11 and the protein residue lysine 28. Therefore, three-nucleotide codon that corresponds to lysine 28 was completely randomized (13). The targeted library was screened against G11A and G11U RNAs, but no suppressors were found. G56U RNA was also screened against all possible substitutions at Phe85, a residue that it interacts with in the crystal structure, and again no suppressors were identified.

Finally, specific L30e amino-acids that interact with Adenosine 57 were randomized. A57 is the only unstacked nucleotide, and it extends away from the RNA into a protein pocket (Fig. 1B). Earlier structural and biochemical work showed important Asn47 and Asn74 contacts to A57 [35]. Accordingly, these residues and others within the protein pocket were targeted for random mutagenesis. Separate randomized codon libraries for Asn 47, Pro49, Arg52, and Asn 74 were screened against wild type and A57 mutant RNAs. This led to the identification of Asn74Ala as a specific suppressor. The A57G RNA mutation decreases the RR for wild type protein from 500 to 36, but the Asn74Ala protein increased the RR for this RNA to 75, suggesting a suppressor effect (Table 2). The Asn74Ala mutant yielded near wild type repression ratios with the wild type LBE plasmid, indicating a specific effect on the mutant RNA. This result may be interpreted to suggest that Asn74 makes a specific RNA contact.

In vitro binding assays of candidate protein suppressors

In principle, the suppression of β -galactosidase expression should approximately parallel the binding affinity for the suppressor protein and RNA mutant. The G56U and A57G RNA variants were produced as 36-nt RNA transcripts, and the protein mutants were produced and purified as MBP-L30e fusion proteins. Radiolabeled RNA was subjected to qualitative electrophoretic mobility shift assays and quantitative nitrocellulose filter binding assays to measure *in vitro* RNA-protein binding affinities. A typical binding isotherm is shown in Figure 3. As suggested, based on the repression ratios, Asn74Ala binds more strongly to A57G RNA than does the wild type protein (Table 4). Furthermore, Gly88Ser and Gly88Ala bind more strongly to the G56U RNA. Both of these suppressors showed clear bandshifted complexes in EMSA experiments (data not shown). For reasons that are unclear, Gly88Cys does not suppress the G56U RNA mutation *in vitro*. Overall, the experiments confirm that the Gly88Ser and Asn74Ala proteins are specific suppressors both *in vivo* and *in vitro*.

4. Discussion

A bacterial two-plasmid screen was used to explore the interaction between the L30e protein and its RNA binding site. The nucleotides and amino acids located at the binding interface were systematically varied 1) in order to find RNAs that retain moderate protein binding ability and 2) to screen for protein mutants whose binding to variant RNAs is strengthened by the RNA and protein co-variations.

K-turn RNA sequence variations

In previous work, it was shown that of the seven nucleotides of the internal loop of the L30e binding site, G11, A12, G58, and A59 tolerated little change whereas A55, G56, and A57 could be changed without dramatically sacrificing protein binding-affinity [16,32]. Figure 4A shows the L30e interaction near the internal loop nucleotides. In the first screen using the two-plasmid system, G11A, G11U, A55G, G56U and A57G were found to have intermediate repression ratios when screened against wild type L30e protein. This is consistent with structural work that demonstrated that G11 takes part in a G:A pair that defines the NC-stem while A55, G56, and A57 remain unpaired (Fig 1A and B). That G11A and G11U RNAs retain moderate protein binding based on their intermediate repression ratios without forming the G:A pairs found in the consensus K-turn is somewhat surprising. However, both variants have been found in Kt-11 and Kt-23 in the ribosome [14,36]. Accordingly, either an A or a U may be modeled to replace G11 in the L30e RNA K-turn [34]. The identities of the three unpaired nucleotides are less essential for K-turn formation. Adenosine 55 and G56 stack on the canonical and non-canonical stems respectively and may be replaced, while A57 protrudes into the protein. Nucleotides essential for the architecture of the RNA kink-turn, the G58:A12 pair and the A59 that participates in the A-minor interaction, were not part of the screen because substitution by other nucleotides led to very low repression ratios indicative of very weak RNA-protein binding. If

protein binding requires the formation of a K-turn, then mutations that greatly destabilize the K-turn would be expected to bind protein weakly, and a single amino acid change would be unlikely to restore binding.

Structural context of suppressors

Screens against G11A, G11U, and A55G failed to find a protein suppressor of weakened RNA binding. Examination of the RNA-protein complex suggests that these residues make important RNA-RNA contacts. For example, the ribose of A55 interacts with the A12-N1 and, as discussed above, G11 forms a sheared pair with A59. Experimental work on a model kink-turn suggests that the 2' OH of the nucleotide equivalent to A55 forms an essential hydrogen bond [37]. For nucleotide substitutions G11A, G11U, and A55G, substitutions may weaken binding by altering RNA-RNA contacts. Presumably, protein binding is weakened by altering the RNA structure, and single amino acid changes are incapable of restoring the interactions required for optimal protein binding.

The logical place to look for an amino acid suppressor is the one nucleotide with multiple protein contacts and no interactions with RNA. Adenosine 57 in the L30e RNA is such a nucleotide as it extends into a loose protein pocket, and its base makes no RNA contacts (Fig. 4B). However, only one substitution at this position, a guanosine, reduces the repression ratio, whereas other nucleotides at position 57 yield wild type or higher repression ratios (Table 3). Thus, contrary to homologous K-turn RNA-protein systems, sequence requirements at this position are not strict [38]. Screening of the randomized library of L30e proteins failed to identify a suppressor, so targeted mutations were individually introduced at amino acid positions 47, 49, 52, and 74. Out of this screen, only Asn74Ala was identified as a partial but specific suppressor of the A57G L30e RNA. This nucleotide is situated in a loose pocket composed of Asn74, which forms a hydrogen bond with the adenine N6, Pro49 that makes a hydrophobic interaction with the central portion of the purine, and Arg52, which makes contacts with the ribophosphate. The exocyclic amino group of G may clash with the Asn74 amide group and replacement by the smaller alanine avoids this potential steric or electronic overlap. On the other hand, according to the crystal structure of the unbound L30e protein, there are two possible structures for the amino acid residues 74-81, and the temperature factors for asparagine 74 and its immediate neighbors are quite high [39]. Thus, this is a region of protein flexibility, and several individual amino acid or nucleotide substitutions may be accommodated within this flexible pocket. Although Asn47, Pro49, Arg52, and Asn74 are universally conserved in eukaryotes (Figure 1D), position 74 is generally a serine or threonine in archaea, but *Sulfolobus tokodaii* has an alanine substitution.

Another mutation, Gly88Ser specifically suppresses the RNA G56U mutation. This suppressor was found by screening an L30e protein library containing randomly introduced mutations over its entire length. A subsequent round of codon randomization at position 88 identified both serine and cysteine as suppressors, but only serine 88 suppresses the G56U mutation both *in vivo* and *in vitro*. However, the crystal structure shows little contact between Gly88 and the RNA and no specific contact with G56 (Fig 4C). Analysis of the structure of the complex shows that Gly88 makes contacts with the RNA's A57 phosphate that is adjacent to the uracil substitution. Thus, it is plausible that stacking is reduced between U56, the protein residue Phe85, and the A12-G58 pair atop the NC-stem, but that the amino acid mutations allow compensatory stabilizing contacts to be made. Therefore, we can speculate that the uracil substitution stacks less efficiently with Phe85, but local adjustments allow Ser88 to make favorable hydroxyl contacts. Modeling based on the L30e RNA-protein complex shows that minor, local structural rearrangements can improve the U56-Phe85 stacking interaction and remove the steric clash between Ser88 and Asn74 [34]. In several other K-turn RNA-protein crystals examined, the amino acid residues that interact with those equivalent to G56 and A57

form part of an extended loop between an α -helix and the fourth β -sheet strand [38]. In this short stretch, the RNA and protein backbones are roughly parallel and make several contacts (Fig. 4C). A BLAST alignment of diverse L30e proteins shows that close relatives of *S. cerevisiae* have a glycine at position 88. However, the *S. pombe* L30e protein has a serine, and many higher organisms, including insect, animals, and plants, have a cysteine at the equivalent position (Fig 1D) [19]. Interestingly, both the screening for compensatory mutations and the phylogenetic comparisons suggest that only glycine, serine, and cysteine are permissible. In fact, in the L7Ae bound to a K-turn based on the box H/ACA sRNA, a uridine is extruded, and its phosphate contacts L7Ae-Ser92, the equivalent of L30e-Gly88 [15].

Finally, an RNA compensatory mutation, G56A, was found to suppress the defect of A55G, suggests that A55G is allowable only with G56A. Examination of the sequences of L30e RNA along with other K-turns confirms that no K-turn RNA has guanosines at the helix capping positions 55 and 56, simultaneously, as would result from the A55G mutation. Figure 4D shows these two purines, and modeling demonstrates that if both are guanines their amino groups may clash [34]. By changing G56 to A56, such a clash can be avoided. Moreover, when nucleotides 55 and 56 are both adenosines, the repression ratio increases two-fold (Table 1).

In summary, the bacterial two-plasmid screen has been shown to be useful in studying a yeast RNA-protein interaction. Specific suppressors of deleterious RNA mutations were of several types: an RNA suppressor that potentially removes a steric clash, an amino acid substitution that may allow local favorable interactions to occur, and the gain-of-function associated with a direct RNA-protein contact. At the outset, it was anticipated that suppressors would be indicative a direct contact at the RNA-protein interface, but the screen was able to identify less direct contacts as well.

Acknowledgments

We are grateful to James R. Williamson (The Scripps Research Institute) for initial discussion of the two-plasmid system and to Cheryl Selah (Bryn Mawr College) for technical help. This work was supported by a grant from the National Institutes of Health to S. A. W. (GM62778-01).

Abbreviations

BLAST	Basic Local Search Alignment Tool
BSA	Bovine Serum Albumin
DTT	Dithiothreitol
IPTG	Isopropyl β -D- thiogalactopyranoside
K_d	Dissociation constant
LB	Luria-Bertani
LBE	L30e RNA binding element
MBP	

	Maltose Binding Protein
NC	Non-canonical
OD	Optical density
ONPG	o-nitrophenyl- β -galactoside
RR	Repression Ratio
S/D	Shine-Dalgarno
TE	Tris-Ethylenediaminetetraacetate
X-gal	5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside

References

1. Dabeva MD, Beittenmiller MA, Warner JR. Autogenous regulation of splicing of the transcript of a yeast ribosomal protein gene. *Proc Natl Acad Sci USA* 1986;83:5854–5857. [PubMed: 3526341]
2. Dabeva MD, Warner JR. The yeast ribosomal protein L32 and its gene. *J Biol Chem* 1987;262:16055–16059. [PubMed: 3316213]
3. Dabeva MD, Warner JR. Ribosomal protein L32 of *Saccharomyces cerevisiae* regulates both splicing and translation of its own transcript. *J Biol Chem* 1993;268:19669–19674. [PubMed: 8366109]
4. Eng FJ, Warner JR. Structural basis for the regulation of splicing of a yeast messenger. *Cell* 1991;65:797–804. [PubMed: 2040015]
5. Vilardell J, Warner JR. Regulation of splicing at an intermediate step in the formation of the spliceosome. *Genes Dev* 1994;8:211–220. [PubMed: 8299940]
6. Vilardell J, Warner JR. Ribosomal protein L32 of *Saccharomyces cerevisiae* influences both the splicing of its own transcript and the processing of rRNA. *Mol Cell Biol* 1997;17:1959–1965. [PubMed: 9121443]
7. Vilardell J, Chartrand P, Singer RH, Warner JR. The odyssey of a regulated transcript. *RNA* 2000;6:1773–1780. [PubMed: 11142377]
8. Vilardell J, Yu SJ, Warner JR. Multiple functions of an evolutionarily conserved RNA binding domain. *Mol Cell* 2000b;5:761–766. [PubMed: 10882112]
9. Li B, Vilardell J, Warner JR. An RNA structure involved in feedback regulation of splicing and translation is critical for biological fitness. *Proc Natl Acad Sci USA* 1996;93:1596–1600. [PubMed: 8643676]
10. Halic M, Becker T, Frank J, Spahn CM, Beckmann R. Localization and dynamic behavior of ribosomal protein L30e. *Nat Struct Mol Biol* 2005;12:467–468. [PubMed: 15864315]
11. Li H, Dalal S, Kohler J, Vilardell J, White SA. Characterization of the pre-mRNA binding site for yeast ribosomal protein L32: the importance of a purine-rich internal loop. *J Mol Biol* 1995;250:447–459. [PubMed: 7616567]
12. Mao H, White SA, Williamson JR. A novel loop-loop recognition motif in the yeast ribosomal protein L30 autoregulatory RNA complex. *Nat Struct Biol* 1999;6:1139–1147. [PubMed: 10581556]
13. Chao JA, Williamson JR. Joint X-ray and NMR refinement of the yeast L30e-mRNA complex. *Structure (Camb)* 2004;12:1165–1176. [PubMed: 15242593]

14. Klein DJ, Schmeing TM, Moore PB, Steitz TA. The kink-turn: a new RNA secondary structure motif. *EMBO J* 2001;20:4214–4221. [PubMed: 11483524]
15. Rozhdestvensky TS, Tang TH, Tchirkova IV, Brosius J, Bachellerie JP, Huttenhofer A. Binding of L7Ae protein to the K-turn of archaeal snoRNAs: a shared RNA binding motif for C/D and H/ACA box snoRNAs in archaea. *Nucleic Acids Res* 2003;31:869–877. [PubMed: 12560482]
16. Li H, White SA. RNA aptamers for yeast ribosomal protein L32 have a conserved purine-rich internal loop. *RNA* 1997;3:245–254. [PubMed: 9056762]
17. Kuhn JF, Tran EJ, Maxwell ES. Archaeal ribosomal protein L7 is a functional homolog of the eukaryotic 15.5kD/Snu13p snoRNP core protein. *Nucleic Acids Res* 2002;30:931–941. [PubMed: 11842104]
18. Koonin EV, Bork P, Sander C. A novel RNA-binding motif in omnipotent suppressors of translation termination, ribosomal proteins and a ribosome modification enzyme. *Nucleic Acids Res* 1994;22:2166–2167. [PubMed: 7518079]
19. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990;215:403–410. [PubMed: 2231712]
20. Bouvet P, Jain C, Belasco JG, Amalric F, Erard M. RNA recognition by the joint action of two nucleolin RNA-binding domains: genetic analysis and structural modeling. *EMBO J* 1997;16:5235–46. [PubMed: 9311984]
21. Jain C, Belasco JG. A structural model for the HIV-1 Rev-RRE complex deduced from altered-specificity Rev variants isolated by a rapid genetic strategy. *Cell* 1996;87:115–125. [PubMed: 8858154]
22. Jain, C.; Belasco, JG. mRNA Formation and Function. Academic Press; New York, NY: 1997. Ch 14: A rapid genetic method for the study of RNA binding proteins; p. 263-284.
23. Jain C, Belasco JG. Rapid genetic analysis of RNA-protein interactions by translation repression in *Escherichia coli*. *Methods Enzymol* 2000;318:309–332. [PubMed: 10889996]
24. Jain C, Kleckner N. IS10 mRNA stability and steady state levels in *Escherichia coli*: indirect effects of translation and role of RNA function. *Mol Microbiol* 1993;9:233–247. [PubMed: 7692216]
25. Chang AC, Cohen SN. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* 1978;134:1141–1156.
26. Hanahan D, Jessee J, Bloom FR. Plasmid transformation of *Escherichia coli* and other bacteria. *Methods Enzymol* 1991;204:63–113. [PubMed: 1943786]
27. Cadwell, RC.; Joyce, GF. PCR Methods and Appl. Vol. 2. Cold Spring Harbor Laboratory Press; 1992. p. 28-33.
28. Cadwell, RC.; Joyce, GF. PCR Methods Appl. Vol. 6. Cold Spring Harbor Laboratory Press; 1994. p. S136-40.
29. Miller, JH. Experiments in Molecular Genetics. Cold Spring Harbor Lab. Press; Cold Spring Harbor, N.Y: 1972.
30. DeLano, WL. The PyMOL Molecular Graphics System. DeLano Scientific; Palo Alto, CA, USA: 2002.
31. Allers J, Shamoo Y. Structure-based analysis of protein-RNA interactions using the program ENTANGLE. *J Mol Biol* 2001;311:75–86. [PubMed: 11469858]
32. White SA, Hoeger M, Schweppe JJ, Shillingford A, Shipilov V, Zarutskie J. Internal loop mutations in the ribosomal protein L30 binding site of the yeast L30 RNA transcript. *RNA* 2004;10:369–377. [PubMed: 14970382]
33. White SA, Shipilov V, Mao H, Williamson JR. Identifying critical residues at the yeast ribosomal protein L30 RNA interface. *Nucleic Acids Sym Ser* 1999;41:4–7.
34. Schweppe, JJ. Ph D Thesis. 2005. Characterization of the *Saccharomyces cerevisiae* ribosomal protein L30 using an in vivo two-plasmid genetic screen; p. 157-191.
35. Shipilov V, White SA. A conserved asparagine makes an essential contact to an RNA adenosine or cytidine. *J Biomol Struct Dyn* 1999;11:75–78.
36. Lescoute A, Leontis NB, Massire C, Westhof E. Recurrent structural RNA motifs, Isostericity Matrices and sequence alignments. *Nucleic Acids Res* 2005;33:2395–2409. [PubMed: 15860776]

37. Liu, J.; Lilley, DM. RNA. Vol. 13. RNA; 2007. The role of specific 2'-hydroxyl groups in the stabilization of the folded conformation of kink-turn; p. 200-210.
38. Nottrott S, Hartmuth K, Fabrizio P, Urlaub H, Vidovic I, Ficner R, Luhrmann R. Functional interaction of a novel 15.5kD [U4/U6.U5] tri-snRNP protein with the 5' stem-loop of U4 snRNA. EMBO J 1999;18:6119–33. [PubMed: 10545122]
39. Chao JA, Prosad GS, White SA, Stout CD, Williamson JR. Inherent protein structural flexibility at the RNA-binding interface of L30e. J Mol Biol 2003;326:999–1004. [PubMed: 12589748]

Figure 1. RNA and protein sequences

(A) Secondary structure of consensus K-turn [14]. Boxed nucleotides form the G:A motif, and circled nucleotides participate in an A-minor interaction. (B) Secondary structure and structural schematic of the *S. cerevisiae* L30e RNA transcript K-turn. Arrows indicate RNA substitutions that allowed protein binding in a SELEX experiment [16]. Closed and open circles represent canonical and non-canonical base pairing, and the star indicates the A-minor interaction. [14]; (C) L30e protein primary structure. The helix and sheet secondary structure is indicated as well as three regions that interact with the RNA, R-I, R-II, and R-III. Key protein residues are numbered. (D) Eukaryotic L30e protein phylogenies of selected regions aligned using BLAST [19]

Figure 2. Reporter and repressor plasmids

The reporter plasmid expresses β -galactosidase under control of the L30e RNA binding element (LBE) whose secondary structure is shown. The lower stem of the LBE has been truncated by one base pair compared to the wild type LBE. Expression of the L30e repressor is controlled by IPTG.

Figure 3. Binding Isotherms for Wild Type and Suppressor Fusion Proteins

Nitrocellulose filter-binding experiments were performed as described in the Experimental Section using radiolabeled wild type and G56U 36-nt RNAs. The lines indicate the best fit to the equation:

$$\% \text{Retainde} = \text{Background} + \frac{\text{Plateau} \times [\text{MBP} - \text{L30}] / K_d}{1 + [\text{MBP} - \text{L30}] / K_d}$$

and were calculated using Kaleidagraph (Synergy Software).

Figure 4. Structures of the L30e RNA-protein complex

Structures were drawn using PyMol using the structure file 1TOK.pdb [13,29]. The RNA tube chains are on the bottom, and the protein chain is on the top. (A) The L30e RNA-protein complex. L30e protein residue Gly88 is depicted as space-filling. L30e RNA positions G56 and A57 are drawn with sticks. (B) The A57 binding pocket of the L30e RNA-protein complex. A57 is shown as space-filling. The L30e peptide backbone is shown with important side chains depicted as sticks. (C) The G56-Gly88 region of the L30e RNA-protein complex. The protein and RNA chains with the G56 base are quasi-parallel. Phe85, Arg86, Val87, and Gly88 are shown as sticks as are RNA residues G56 and A57. (D) L30e RNA helix capping residues A55 and G56, shown as sticks, stack on top of the G13:C54 base pair, shown in thin lines, of the canonical stem. G56 stacks on the A12:G58 pair, shown in thin lines, of the non-canonical stem. If a G replaces A55, its amino group potentially clashes with the G56 amino group.

Table 1**Repression Ratios and Binding Affinities for RNA variants with wild-type L30e protein**

The repression ratios (RR) represent averages of several assays. All of the mutant RNA reporters were tested with the wild type pRPL30, as previously described. Dissociation constants (K_d) were reported in [31] or in [32] as indicated by *. Uncertainties are based on the standard deviations of at least three assays. Mutant RNAs that retain moderate protein affinities are in bold and were selected for screening with mutant proteins.

RNA	RR	K_d (nM)
<i>RNAs bearing single mutations</i>		
WT (LBE in Fig. 2)	450 ± 120	30 ± 10
G11A	89 ± 22	
G11C	1.4 ± 0.2	500 ± 150
G11U	16 ± 7.8	
A12C	1.2 ± 0.1	1050 ± 300
A12G	0.9 ± 0.2	
A12U	1.3 ± 0.2	
A55C	525 ± 60	
A55G	59 ± 13	15 ± 5
A55U	890 ± 220	
G56A	1000 ± 400	300 ± 100
G56C	0.9 ± 0.2	1200 ± 400
G56U	10 ± 3.8	
G56 deletion	1 ± 0.1	>1000 *
A57C	1010 ± 280	15 ± 5
A57G	36 ± 8	60 ± 20
A57U	490 ± 26	
A57 deletion	0.7 ± 0.1	>1000 *
G58A	0.8 ± 0.2	1500 ± 500
G58C	0.9 ± 0.2	
G58U	1.0 ± 0.2	
A59C	0.7 ± 0.1	
A59U	0.7 ± 0.1	
A59G	0.7 ± 0.2	450 ± 150
<i>RNAs bearing multiple mutations</i>		
G11A + A59G	1.3 ± 0.3	
A12G + G58A	1.0 ± 0.2	
G11A + A12G + G58A + A59G	0.9 ± 0.2	
A55G + G56A	490 ± 40	

Table 2**Specificity of Gly88Ser Suppression of G56U RNA**

Repression ratios with standard deviations are the average of several Miller Assays. Each data set included positive and negative controls and where appropriate, values have been normalized to the wild type repression ratio done in parallel. Suppressor data is shown in bold.

Repressor (Protein)	Reporter (RNA)	RR
WT-L30	G56U	12.6 ± 1.5
Gly88Ser	G56U	53 ± 20
Gly88Ala	G56U	16.5 ± 0.7
Gly88 Cys	G56U	106 ± 30
WT-L30	WT (LBE)	552 ± 172
Gly88Ser	WT (LBE)	448 ± 151
Gly88Ala	WT (LBE)	604 ± 214
Gly88Cys	WT (LBE)	290 ± 10

Table 3**Repression ratios of the Asn74Ala protein mutant with the A57G and wild-type reporters**

Each Miller Assay was completed several times in parallel with wild type positive controls. The Asn74Ala suppressor of A57G is shown in bold.

Repressor (Protein)	Reporter (RNA)	RR
WT-L30	A57G	36 ± 9
Asn74Ala	A57G	75 ± 20
WT-L30	WT (LBE)	500 ± 148
Asn74Ala	WT (LBE)	456 ± 112

Table 4***In vitro* dissociation constants (K_d)**

Dissociation constants and standard deviations for wild type and mutant RNAs and proteins were determined by averaging several nitrocellulose filter-binding assays. All proteins were fused to the C-terminus of the maltose binding protein. K_d s for *in vitro* suppressors are shown in bold.

	RNA		
	WT(LBE) (nM)	G56U (nM)	A57G (nM)
Protein			
WT-L30	15.2 ± 3.6	382 ± 69	38.7 ± 8.1
Gly88Ala	14.9 ± 4.1	223 ± 53	
Gly88Cys	20.3 ± 4.8	725 ± 142	
Gly88Ser	12.8 ± 3.7	193 ± 34	
Asn74Ala	12.8 ± 2.5		19.6 ± 2.8