RNA (1998), *4*:471–478. Cambridge University Press. Printed in the USA.

Isoleucine:RNA sites with associated coding sequences

IRENE MAJERFELD and MICHAEL YARUS

Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado 80309-0347, USA

ABSTRACT

An RNA family that binds isoleucine with K_d = 200–500 μM was repetitively isolated from a ribonucleotide transcript pool containing 50 randomized positions. The RNA site is specific, discriminating against branched side chains of different size (valine—one methylene smaller than isoleucine) by at least 1.3 kcal/mol and against the shape (linear) of norleucine by 0.6 kcal/mol. The binding site was localized by sequence comparison, by synthesis of mutant and truncated active derivatives, and by chemical modification–interference. The binding site is small, no more than 10–12 nt, containing an asymmetric internal loop (6 over 2 nt) that includes the isoleucine codon AUU and a sequence of four G's, two of which are involved in G·U and G·C base pairs. Areas of U/G concentration like these may signal a hydrophobic RNA site.

Keywords: aliphatic; amino acid; genetic code; hydrophobic; selection–amplification

INTRODUCTION

Selection–amplification in vitro has been used to isolate RNAs with affinity for several amino acids, starting from sequence-randomized RNA pools. Characterized RNAs bind valine (Majerfeld & Yarus, 1994), citrulline (Famulok, 1994), and eight independent RNA sites were isolated with affinity for arginine (Connell et al., 1993; Connell & Yarus, 1994; Famulok, 1994; Tao & Frankel, 1995; Geiger et al., 1996). For a review of RNA amino acid binding sites, see Yarus (1998). To explore the capability of RNA to form a hydrophobic pocket, we have previously isolated an RNA that binds valine in preference to isoleucine (Majerfeld & Yarus, 1994). Affinity of this RNA for its ligand was moderate (K_d = 12 mM), but it showed stereospecificity, preferring L- over D-valine 15-fold ($\Delta\Delta G = 1.7$ kcal mol^{-1}). The valine-binding RNA also distinguished side chains of different size (e.g., alanine, α-amino-butyrate, or isoleucine had $\Delta\Delta G = 1.6$–1.8 kcal mol^{-1} versus valine), shape (linear norvaline had $\Delta\Delta G = 1.6$ kcal mol^{-1}), and polarity (valine versus isosteric, but polar threonine showed $\Delta\Delta G = 2.0$ kcal mol^{-1}). We now report the isolation of an RNA that conversely binds isoleucine in preference to valine. These new isoleucine sites show that RNA can display varied aliphatic specificities and also provide a new site in which amino acid binding is strongly dependent on nucleotides of a coding triplet within the site (Yarus, 1998).

Reprint requests to: Michael Yarus, Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado 80309-0347, USA; e-mail: yarus@stripe.colorado.edu.

RESULTS

Selection and analysis of the final RNA pool

Using selection–amplification, we isolated an RNA that binds isoleucine from an RNA transcript pool with 50 randomized nucleotides. RNA was adsorbed to an affinity matrix, ≈10 mM isoleucyl–agarose, and subsequently eluted with equimolar L-isoleucine. After 13 cycles of selection, pooled RNA sequences were no longer random, as visualized by partial T1 RNAse digestion (Ciesiolka et al., 1996). After two further selection–amplification cycles, RNAs were cloned and sequenced (Ciesiolka et al., 1996). The final pooled RNAs had increased median elution volume to 1,500 μL (versus ≈300 μL for randomized RNA), but a distinct isoleucine elution peak was not visible in the pool profile, despite the clear isoleucine elution of individuals purified from the pool (below). A pool selected for affinity to another hydrophobe, the amino acid valine, had behaved in the same way (Majerfeld & Yarus, 1994).

Fifty isolates were sequenced and grouped according to sequence homologies. Two repetitively isolated motifs could be identified. These two families were

termed the "AUA" and the "AUU" families. Figure 1 shows the sequences and partial secondary structures of representative clones.

The AUA family appeared to descend from two independent molecules in the initial randomized pool. Representative clones were tested for column affinity and for elution by free isoleucine and related molecules to determine binding specificities. IL11 bound to isoleucyl agarose with $K_d \cong 1.6$ mM, which is similar to the average of the final selected population. However, it

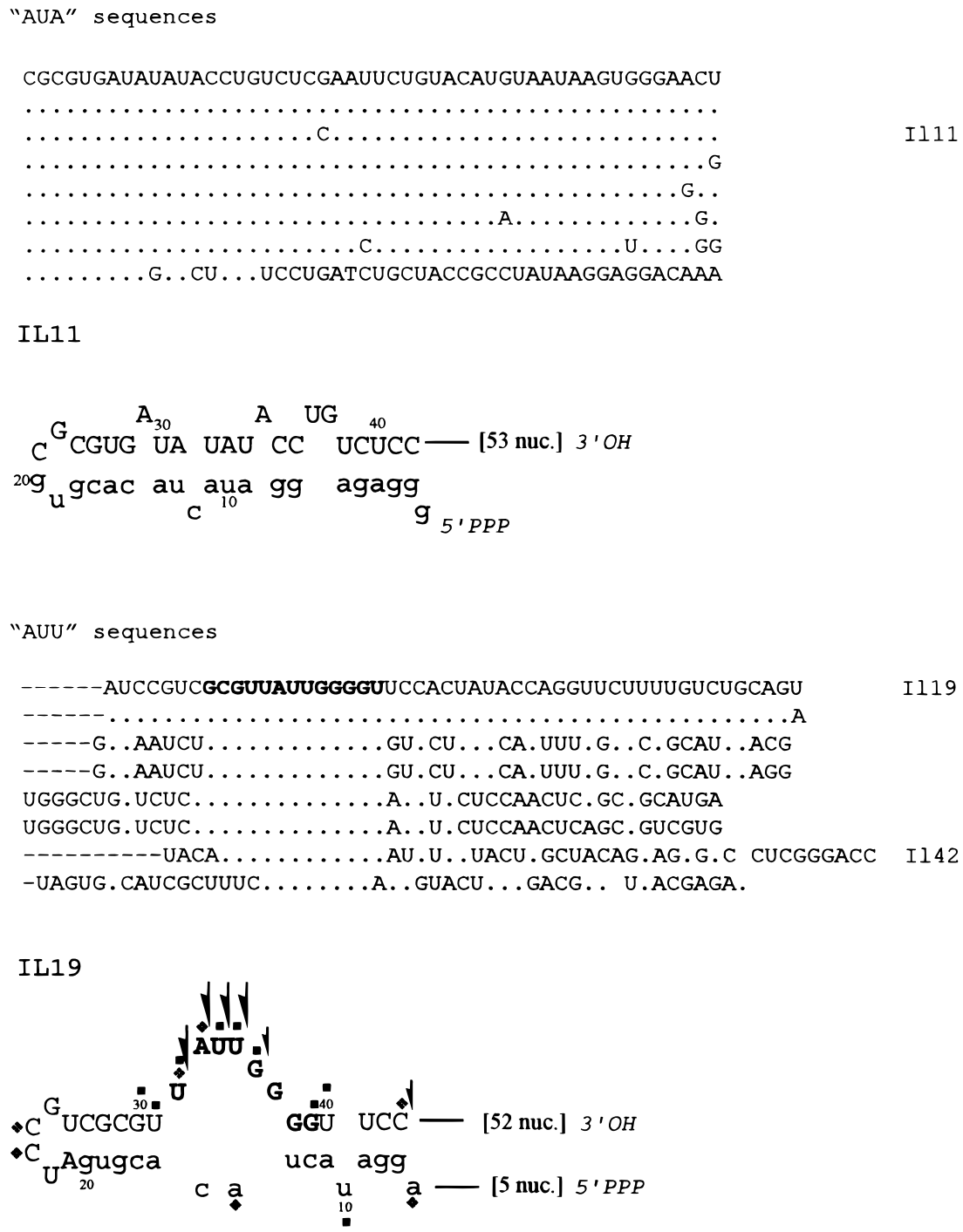

FIGURE 1. Alignment of the randomized regions of the AUA and AUU sequences, partial foldings of representative clones and summary of structural probing data for the AUU clone. Dots indicate nucleotides conserved with respect to the first sequence in each group, dashes indicate primer nucleotides. Conserved nucleotides in the AUU family motif are bold. Computed minimal free energy structures (Mfold; Jaeger et al., 1989) for clones IL11 and IL19 are shown with initially random nucleotides capitalized and fixed nucleotides in lower case. Pb^{2+} cleavages are indicated by half arrows; longer for more frequent cleavage. Susceptibilities to DMS and CMCT modifications are shown by diamonds and squares, respectively.

showed little specificity, being equally eluted by isoleucine, alanine, valine, and methyl-amine ($K_d \approx 10$ mM). This nonspecific family, which appeared to be indifferent to the amino acid side chain, was not characterized in detail.

The AUU family (eight clones) was specific. Sequence variation among isolates (Fig. 1) suggests derivation from ≥5 initial molecules. In selection buffer, RNAs IL19 and IL42 bound the column and free isoleucine with K_d = 200–500 μM, 5–10 times better to isoleucine than to free alanine or valine, and favored L- over D-isoleucine ninefold. The sequence shared among this entire family is the sequence UAUUGGGG; flanking sequences are partially conserved.

The calculated (Jaeger et al., 1989) most-stable structure places the sequence UAUUGG in an internal loop opposite two bulged fixed nucleotides (Fig. 1). Two conserved G's form base pairs, including one GU base pair. This secondary structure was confirmed for IL19 by limited Pb^{2+} hydrolysis, which is diagnostic for regions sufficiently flexible to align themselves for cleavage of the phosphodiester (Ciesiolka et al., 1992). Higher Pb^{2+} sensitivity is observed for loop nt U32–G36. A characteristic alternation of sensitivity within the loop seen in the full-sized molecule is preserved in all smaller derivatives, suggesting that the loop backbone confirmation is preserved in detail in all these RNAs (Fig. 2). The secondary structure is further confirmed by susceptibility to DMS and CMCT modifications of Watson–Crick sites on the bases. Most relevantly, the looped structure for the conserved UAUUGG was confirmed by all probes (Fig. 1).

Specificity of isoleucine binding

Isoleucine agarose binding by AUU RNAs required Zn^{2+}. IL19 and IL42 did not bind the column in selection buffer with 7.5 mM $MgCl_2$ as the only divalent. Restoration of 0.1 mM Zn^{2+} also restored binding; other selection divalents did not. Binding improved with Zn^{2+} up to at least 0.5 mM. Further experiments use only 7.5 mM Mg^{2+} and 0.4 mM Zn^{2+} as divalents. Addition of Zn^{2+} gave several protections and enhancements in and outside the conserved loop at concentrations as low as 40 μM (not shown), suggesting a Zn^{2+}-RNA structure. It is possible, in addition, that Zn^{2+} amino acid is an active ligand species.

We estimated dissociation constants by isocratic elution from isoleucyl agarose (Table 1; Ciesiolka et al., 1996). Full-sized RNAs and 5′ fragments thereof (see below, Fig. 3) bound free isoleucine with K_d = 200–500 μM. Discrimination against side chains of different configuration or size was evident. For example, the truncated transcript IL19-54 bound to isoleucine with a K_d = 400 μM and to norleucine (the same size, but linear in side-chain shape) with K_d = 1 mM ($\Delta\Delta G$ of 0.6 kcal/mol). Valine, one methylene shorter than isoleucine, bound with $K_d \geq 3.5$ mM to the truncates ($\Delta\Delta G \geq$ 1.3 kcal/mol). IL19-54 discriminates shape less strongly than a previously selected valine binding RNA ($\Delta\Delta G$ = 1.6 kcal/mol for norvaline; Majerfeld & Yarus, 1994), but is similar per methylene group ($\Delta\Delta G$ = 1.6 kcal/mol; Majerfeld & Yarus, 1994). Amino acid amides were better eluants, and may better resemble the column-bound affinity ligand used for selection. Nevertheless, their side chains were not distinguished as well as free amino acids (Table 1). The presence of the amide apparently obstructs specific side-chain contacts.

Thus, the specific hydrophobic RNA site for valine (Majerfeld & Yarus, 1994) is not unique; sites for other aliphatic hydrophobic shapes can be selected. Valine and isoleucine RNAs contain in common an asymmetric internal loop and a sequence of four G's, one of which (in isoleucine RNA) or two (in valine RNA) are involved in G·U base appositions. Other evidence also suggests that G's and G·U's may be pivotal in creation of a hydrophobic RNA site (see below).

Identification of the binding site

Our argument requires the location of this amino acid site. Extensive sequence variation among the AUU family (Fig. 1) suggested that the binding site could be small, centered on the conserved sequence UAUUG GGG. Terminal random truncation experiments using alkaline hydrolysis gave 3′ boundaries at nt 43 for IL19 and the comparable place in IL42. Thus, the 5′ domain of these RNAs was sufficient for binding and specific elution (not shown).

This was confirmed by specific truncation. Using an *Eco*R II site present in IL19 DNA (nt 50–54), we made a 54-nt truncated transcript, IL19-54. Probing of IL19-54 by Pb^{2+} hydrolysis and chemical modification showed that the secondary structure of this 5′ domain was unaffected (Figs. 2, 3). Zn^{2+} dependence and ligand specificity was also maintained. Isocratic elution of truncated IL19-54 RNA from an isoleucyl agarose column illustrates its specificity; isoleucinamide is a notably better eluant than valinamide (Fig. 4a). Therefore, the truncated RNA detects removal of one methylene group from the side chain.

Further sequence changes confine the site to the RNA's internal loop. Truncate IL42-32 is a shortened version of the 5′ domain of IL42 with an altered flanking hairpin. IL42-32b had a different nucleotide substituted at every position outside the conserved internal loop. In the latter, the fourth conserved G was changed to an A in an AU base pair. Pb^{2+} probing of these 32-mers confirmed the conserved family loop (Figs. 2, 3).

These 32-mers also required Zn^{2+} for column binding. Although overall affinity of IL42-32b RNA for the affinity matrix, isoleucine agarose, was decreased, specificity for free ligands determined by affinity chromatography (Fig. 4b) was similar to the parental RNA

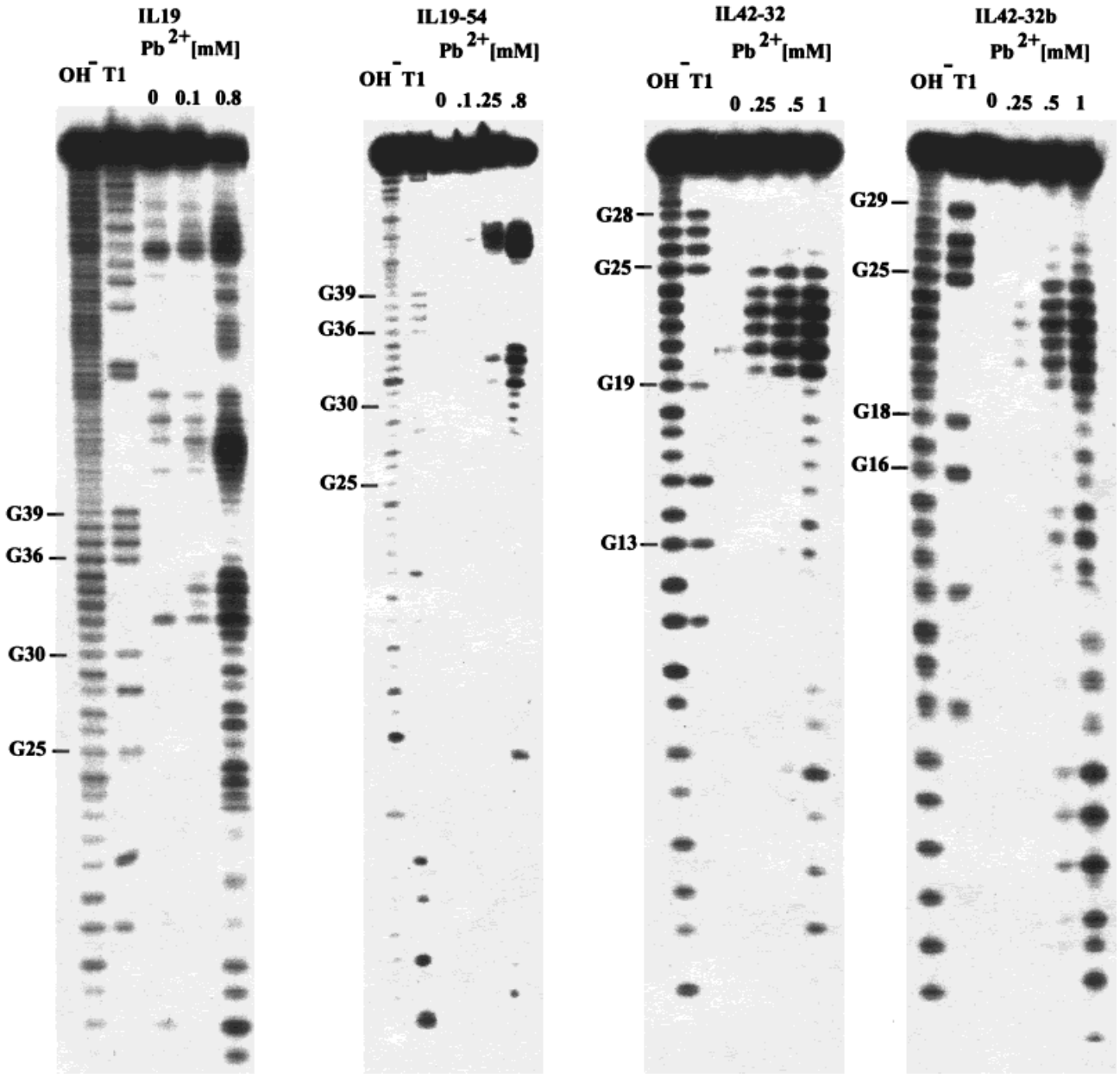

FIGURE 2. Structural probing of IL19 and truncated derivatives with Pb^{2+}. RNA was [^{32}P] labeled at the 5′ terminus and the products of partial Pb^{2+} hydrolysis were identified by separation in 12% polyacrylamide (IL19 and IL19-54) or 15% polyacrylamide (IL42-32 and IL42-32b). Each partial Pb^{2+} hydrolysate is run alongside an NaOH ladder and a partial RNase T1 digest as markers.

(Table 1). These results, taken together, show that the isoleucine binding site resides within the conserved internal loop, which is the only feature conserved in all molecules.

Involvement of bases in the conserved sequence was confirmed by CMCT modification–interference. CMCT modifies the N-1 of G and the N-3 of U, nucleotides frequent in the conserved loop. IL19-54 RNA was modified at 25 °C or at 65 °C, applied to isoleucyl agarose, then washed and specifically eluted. Molecules with modifications at positions that are required for binding or folding are enriched in the column flow through; modifications at crucial nucleotides are depleted in the bound and eluted fraction. Results are summarized in Figure 5.

With the exception of U35, modification of all G's and U's in the conserved sequence interfered with binding when denatured RNA was modified. Thus, the normal base pairing faces of all bases save U35 must be intact for folding of the amino acid site. Modification of nucleotides proximal to the loop (U13, G18, U19, and U40) interfered with binding, but to a lesser extent. CMCT reaction on both sides and in both strands at more distant sites is innocuous; thus, the site is sharply centered at the loop. Modifications performed at 25 °C were

TABLE 1. Specificity of isoleucine binding RNAs.[a]

	K_d (mM)									
		Amino acids					Amino acid amides			
	Ile-agarose	Ile	Norleu	Val	Ala	Gly	Ile	Leu	Val	Ala
Full-length RNA										
IL19	0.4 (2)	0.4 (2)		1.7 (2)	1.7 (2)					
IL42	0.5 (6)	0.5 (2)		1.7 (1)	1.7 (5)					
Truncated RNA										
IL19-54	0.3 (8)	0.4 (6)	1.0 (1)	>3.5 (1)	>3.5 (4)	3.0 (1)	0.1 (4)	0.7 (1)	0.3 (1)	0.3 (1)
IL42-32	1.0 (1)	0.3 (1)			>1.2 (1)					
IL42-32 b	8 (2)	0.2 (2)			1.2 (2)					

[a]Dissociation constants (K_d) for agarose-bound isoleucine and free ligands were determined by isocratic competitive affinity chromatography. Numbers in parenthesis are number of repeats. The error of means is estimated at ± 20%.

consistent but, as predicted by the secondary structure model, G38 and G39 were not accessible to modification in folded RNA. Modification at U41, which lies outside of the loop, results in some loss of binding; it appears that, at this position, the modification can be better accommodated upon folding of the RNA. In summary, the crucial nucleotides are thereby confined to conserved loop sequences and helix adjacent, in agreement with data above. All data are accounted for by the notion that all structure required for the amino acid site is supplied by specific loop nucleotides flanked by arbitrary helices.

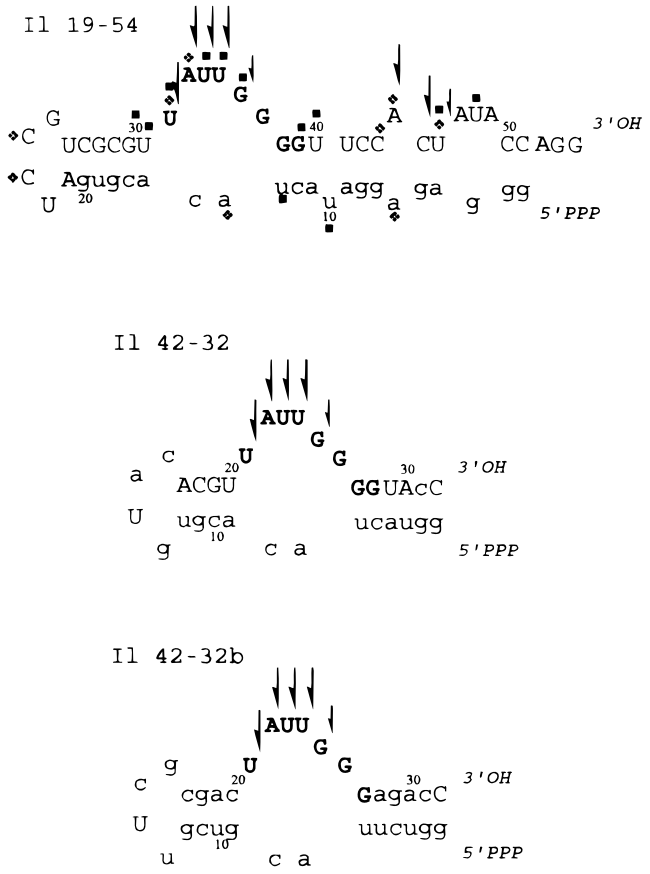

FIGURE 3. Computed minimal free energy structures and summary of structural probing data for truncates IL19-54, IL42-32, and IL42-32b. Upper case indicates nucleotides initially randomized and lower case indicates both PCR primer sequences and nucleotides whose identity was arbitrarily fixed. Symbols as in legend to Figure 1.

DISCUSSION

From an initial RNA pool of about 2×10^{13} 95-mers, we have isolated an RNA that binds isoleucine ($K_d = 200$–$500\ \mu$M) and does so specifically. The binding site is contained within a small asymmetric loop. An internal loop generally of not more than 10–12 nt, as exemplified by the isoleucine site, appears to be a common feature of most RNA amino acid binding sites (Yarus, 1998). The AUU family exemplifies this tendency, and also has a highly conserved sequence that includes an isoleucine triplet implicated in amino acid binding by modification studies.

Such functional coding triplets were not found in the selected valine site (Majerfeld & Yarus, 1994), but are frequent among in vitro-selected arginine-binding RNAs (Yarus, 1998) and have been found in a natural binding site, the group I intron (Yarus, 1988). Thus, such sequences exist in amino acid sites specific for arginine and isoleucine, at both extremes of side-chain polarity. This supports a stereochemical basis for the genetic code (Woese et al., 1967) in which direct RNA affinity for amino acids may have comprised the first coding. (For more complete discussion, see Yarus, 1998.)

Similarities between the valine and isoleucine sites are easily found. Both contain conserved strings of G's with apposed U's. Aliphatic amino acid side-chain interactions have long been known (la Cour et al., 1985) in which the aliphatic region is apposed to the face of G bases, suggesting that the conserved G's may be involved in a predictable interaction.

The same is true of the U base. Tat protein of bovine immunodeficiency virus (BIV) requires an isoleucine

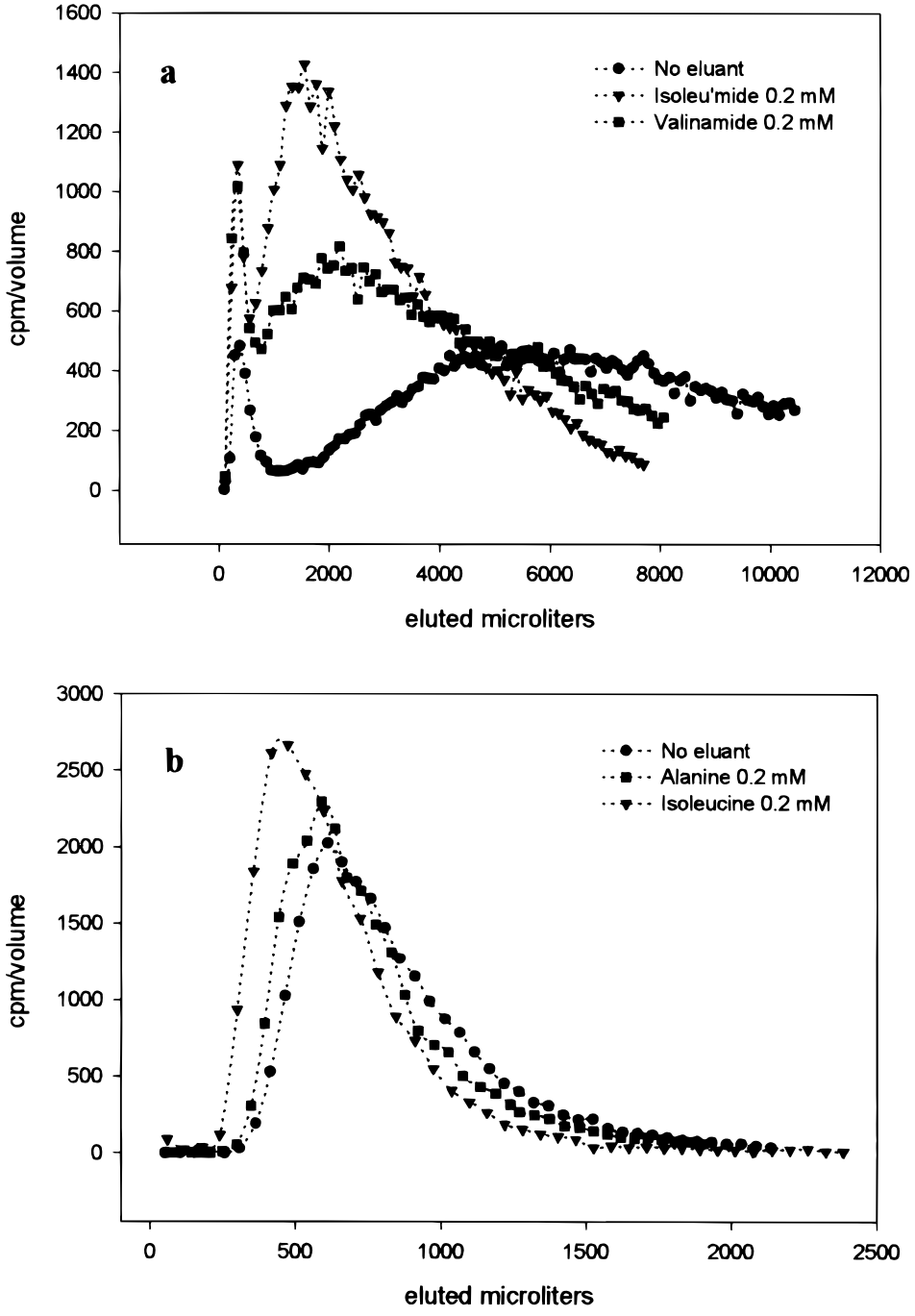

FIGURE 4. Isocratic affinity elution from isoleucine agarose. Approximately 10 pmol of internally labeled ^{32}P-RNA were denatured at 65 °C for 3 min, brought to column buffer composition (buffer B + 0.4 mM $ZnCl_2$ and 0.2 mM eluant when present) and allowed to renature for 10 min at room temperature. **a:** IL19-54 RNA. Column size 0.2 mL. **b:** IL42-32 b RNA. Column size 0.4 mL.

residue for binding to its target, TAR RNA (Chen & Frankel, 1995). Two recently published NMR structures of BIV TAR–Tat peptide complexes (Puglisi et al., 1995; Ye et al., 1995) direct this critical isoleucine toward the hydrophobic edge (C5-C6) of a specific U base, presented (Ye et al., 1995) in the TAR RNA major groove by its tertiary structure. Given prominent conserved U's within the isoleucine and valine (Majerfeld & Yarus, 1994) binding RNAs, BIV TAR-Tat peptide structure suggests that U edges may also be prominent among elements that can be assembled into specific hydrophobic RNA sites.

There is independent evidence for the role of G's and U's. These nucleotides are prominent in other hydro-

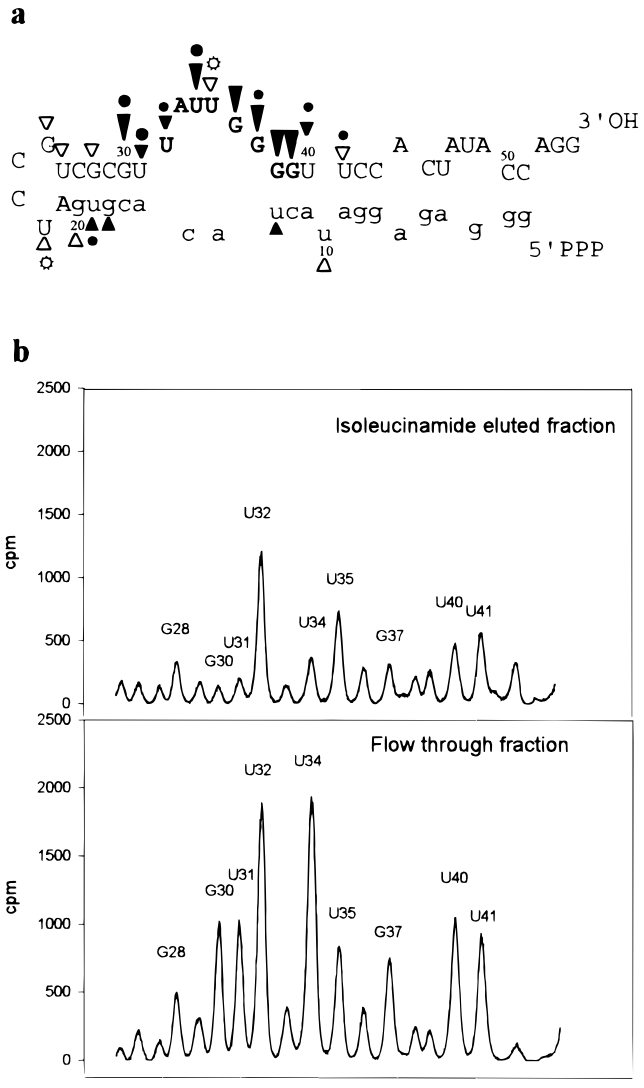

FIGURE 5. CMCT modification–interference on truncate RNA IL19-54. Chemical modification was performed at 25 °C and at 65 °C and the modified RNA was bound to and eluted from an isoleucyl agarose column. **a:** Summary of results. Triangles and circles indicate modified positions after treatment of unfolded and folded RNA, respectively. Filled symbols represent modifications that interfere with binding and elution. Larger symbols indicate greater interference. Open symbols mark modifications that do not interfere with binding and elution from the column. **b:** Phosphorimager profile of reverse transcription analysis for the segment G28–U41 of the flow through and specifically eluted fractions of RNA modified at 25 °C. U32 is a position either very sensitive to spontaneous cleavage or the site of a reverse transcription stop; interference was evaluated over this background.

phobic sites, and larger hydrophobic ligands seem to require larger G/U-rich structures. Recently isolated DNA (Li & Sen, 1996) and RNA (Conn et al., 1996) enzymes for the metalation of the hydrophobic ligand porphyrin, selected on the basis of affinity to a transition-state analogue, have sequences exceptionally rich in G and U residues. Such consistency suggests that the well-known association of codons centered on U (NUN′ codons) with hydrophobic amino acids may be explicable in terms of a primordial chemical affinity.

MATERIALS AND METHODS

Selection procedure

Isoleucine was linked through its carboxyl to EAH Sepharose 4B (Pharmacia) as described previously (Majerfeld & Yarus, 1994) to approximately 10 μmol/mL and the selection step was performed on a 0.2-mL column. Selection buffer was 10 mM HEPES, pH 7.0, 300 mM NaCl, and 7.5 mM $MgCl_2$ (buffer B) and 0.1 mM each of $ZnCl_2$, $CaCl_2$, and $MnCl_2$ and 0.01 mM $CuCl_2$. Elution buffer had, in addition, 10 mM isoleucine. Starting at cycle 7, 10 mM glycine was added to both buffers to discourage enrichment of species with exclusive affinity for the amino group of the ligand. The initial RNA pool consisted of 830 pmol of RNA (95 nt) transcribed from 2×10^{13} DNA templates; there are 50 randomized positions flanked by PCR primers that included a T7 RNA polymerase promoter and *Eco*R I and *Hin*d III restriction sites for cloning. The RNA pool was folded by heating in water to 65 °C for 3 min, adjusted to buffer B, and cooled to 25 °C over 10 min, then the other divalent ions were added. The RNA was applied to the column and washed with 5–20 column volumes of selection buffer (increasing stringency as the selection progressed) before specific elution. RNA eluted by isoleucine was precipitated and reverse transcribed. The cDNA was amplified by PCR and transcribed to produce a pool of RNA for the next selection. After cycle 2, selection was preceded by counterselection on an acetylated column to remove RNA with affinity for the resin. About 65% of the front running peak was collected for selection.

Lead hydrolysis

Lead cleavage was performed in buffer B with 0.4 mM $ZnCl_2$ as described (Ciesiolka et al., 1995). 5′-^{32}P-labeled RNA was 0.25 μM and $Pb(OAc)_2$ was added at concentrations ranging between 0 and 1.0 mM.

Chemical modification and modification interference

DMS and CMCT modifications were performed as described (Krol & Carbon, 1989) except that DMS treatment was done in 80 mM HEPES, pH 7.0, 50 mM NaCl, 7.5 mM $MgCl_2$, and 0.4 mM $ZnCl_2$. The CMCT reaction buffer was 50 mM sodium borate, pH 8.0, 42 mM KCl, 7.5 mM $MgCl_2$, and 0.4 mM $ZnCl_2$. Treatment was followed by two precipitations. Modifications were identified by primer extension with AMV reverse transcriptase and denaturing polyacrylamide electrophoresis. For IL19-54 modification–interference, the modification was either at 65 °C for 25 s (no $ZnCl_2$ in the reaction) or after renaturation at 25 °C for 15 min. RNA treated at 65 °C was precipitated twice and renatured in buffer B + 0.4 mM $ZnCl_2$. CMCT was removed from RNA treated at 25 °C by passage through a 1-mL P6 column in buffer B + 0.4 mM $ZnCl_2$. In both cases, modified RNA was applied to a 0.2-mL isoleucyl–agarose column, washed with eight column volumes of buffer B + 0.4 mM $ZnCl_2$, and eluted with buffer containing 0.2 mM isoleucinamide. Flow through, wash, and specifically eluted fractions were pooled.

Determination of K_d

RNA dissociation constants (K_d) for isoleucine agarose and free ligands were determined by isocratic competitive affinity chromatography (Connell et al., 1993; Ciesiolka et al., 1996). Approximately 10 pmol of internally labeled ^{32}P-RNA were heated at 65 °C for 3 min, brought to column buffer conditions (buffer B + 0.4 mM $ZnCl_2$ and 0.2 mM eluant when present), and allowed to renature for 10 min at room temperature before loading onto the column. Median elution volumes were used to calculate K_d with no correction for an unbound fraction (see Fig. 4). This leading peak consists of aggregated molecules, as judged from rerunning on native gel electrophoresis (not shown). Aggregated material declines in amount as the molecule is simplified (compare Fig. 4a and b), being relatively small or absent in truncated RNAs. Aggregates do not separate well from the main RNA peak in full-length molecules; uniformly neglecting this relatively small fraction increases K_d, but by an amount judged small with respect to other errors.

Received December 30, 1997; returned for revision January 26, 1998; revised manuscript received February 9, 1998

REFERENCES

Chen L, Frankel AD. 1995. A peptide interaction in the major groove of RNA resembles protein interactions in the minor groove of DNA. *Proc Natl Acad Sci USA 92*:5077–5081.

Ciesiolka J, Gorski J, Yarus M. 1995. Selection of an RNA domain that binds Zn^{2+}. *RNA 1*:538–550.

Ciesiolka J, Illangasekare M, Majerfeld I, Nickles T, Welch M, Yarus M, Zinnen S. 1996. Affinity selection–amplification from randomized ribonucleotide pools. *Methods Enzymol 267*:315–335.

Ciesiolka J, Lorenz S, Erdmann VA. 1992. Structural analysis of three prokaryotic 5S rRNA-ribosomal–protein complexes by means of Pb(II)-induced hydrolysis. *Eur J Biochem 204*:575–585.

Conn MM, Prudent JR, Schultz PG. 1996. Porphyrin metalation catalyzed by a small RNA molecule. *J Am Chem Soc 118*:7012–7013.

Connell GJ, Illangasekare M, Yarus M. 1993. Three small ribonucleotides with specific arginine sites. *Biochemistry 32*:5497–5502.

Connell GJ, Yarus M. 1994. RNAs with dual specificity and dual RNAs with similar specificity. *Science 264*:1137–1141.

Famulok M. 1994. Molecular recognition of amino acids by RNA-aptamers: An L-citrulline binding RNA motif and its evolution into an L-arginine binder. *J Am Chem Soc 116*:1698–1706.

Geiger A, Burgstaller P, von der Eltz H, Roeder A, Famulok M. 1996. RNA aptamers that bind L-arginine with sub-micromolar dissociation constants and high enantioselectivity. *Nucleic Acids Res 24*:1029–1036.

Jaeger JA, Turner DH, Zucker M. 1989. Improved predictions of secondary structure for RNA. *Proc Natl Acad Sci USA 86*:7706–7710.

Krol A, Carbon PA. 1989. A guide for probing native and small nuclear RNA and ribonucleoprotein structures. *Methods Enzymol 180*:212–227.

la Cour TF, Nyborg J, Thirup S, Clark BFC. 1985. Structural details of the binding of GDP to elongation factor Tu from *E. coli* as studied by X-ray crystallography. *EMBO J 4*:2385–2388.

Li Y, Sen D. 1996. A catalytic DNA for porphyrin metallation. *Nature Struct Biol 3*:743–747.

Majerfeld I, Yarus M. 1994. An RNA pocket for an aliphatic hydrophobe. *Nature Struct Biol 1*:287–292.

Puglisi JD, Chen L, Blanchard S, Frankel AD. 1995. Solution structure of a bovine immunodeficiency virus Tat–TAR peptide–RNA complex. *Science 270*:1200–1203.

Tao J, Frankel AD. 1995. Arginine-binding RNAs resembling TAR identified by in-vitro selection. *Biochemistry 35*:2229–2238.

Woese CR, Dugre DH, Dugre SA, Kondo M. Saxinger WC. 1967. On the fundamental nature and evolution of the genetic code. *Cold Spring Harbor Symp Quant Biol 31*:723–736.

Yarus M. 1988. A specific amino acid binding site composed of RNA. *Science 240*:1751–1758.

Yarus M. 1998. Amino acids as RNA ligands; a DIRT theory for the origin of the code. *J Mol Evol*. In press.

Ye X, Kumar RA, Patel DJ. 1995. Molecular recognition in the bovine immunodeficiency virus Tat peptide–TAR RNA complex. *Chem Biol 2*:827–840.