

The PUMILIO–RNA Interaction: A Single RNA-Binding Domain Monomer Recognizes a Bipartite Target Sequence[†]

Phillip D. Zamore,[‡] David P. Bartel,[‡] Ruth Lehmann,[§] and James R. Williamson^{*||}

The Whitehead Institute for Biomedical Research and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, Massachusetts 02129, Howard Hughes Medical Institute, The Skirball Institute, New York University Medical Center, New York, New York 10016, and The Skaggs Institute for Chemical Biology, Scripps Research Institute, La Jolla, California 92037

Received September 21, 1998; Revised Manuscript Received October 28, 1998

ABSTRACT: Translational repression of *hunchback* (*hb*) mRNA in the posterior of the *Drosophila* embryo requires two copies of a bipartite sequence, the Nanos Response Element (NRE), located in the 3′ untranslated region of the mRNA. The PUMILIO (PUM) protein is thought to bind the NREs and thereby repress *hb* translation. The RNA-binding domain of PUM defines an evolutionarily conserved family of RNA-binding proteins, the PUM-Homology Domain (PUM-HD) proteins, which have been identified in yeast, plants, and animals. The PUM RNA-binding domain, the *Drosophila* PUM-HD (DmPUM-HD), has been shown previously to recognize nucleotides in both the 5′ and 3′ halves of the NRE, suggesting that a dimer of PUM might recognize one NRE. Here, we analyze the RNA-binding affinity and stoichiometry of the DmPUM-HD and find that one DmPUM-HD monomer binds independently and with equal affinity to each NRE ($K_D \sim 0.5$ nM). We detect no cooperative interactions between DmPUM-HD monomers bound at adjacent sites. Our results imply that a single DmPUM-HD protein recognizes nucleotides in both the 5′ and 3′ NRE half-sites. Based on our estimate of the intraembryonic concentration of PUM (>40 nM), we propose that in vivo nearly all NREs are occupied by a PUM monomer.

Regulation of mRNA translation is a common strategy for controlling protein expression during development (1, 2). In the fruit fly *Drosophila melanogaster*, mRNAs encoding regulatory proteins are synthesized during oogenesis and are prepositioned within the egg prior to fertilization. Subsequently, these regulatory proteins repress or activate translation of maternally supplied mRNAs in specific regions within the early embryo. For example, maternal mRNA encoding the transcription factor *hunchback* (*hb*)¹ is uniformly distributed throughout the embryo, but the HUNCHBACK (HB) protein is synthesized in a gradient with its highest concentration at the anterior pole (3, 4). HB protein acts in the anterior half of the embryo to promote the development of structures such as head and thorax (5, 6). Ectopic HB in the posterior blocks the normal development of abdominal segments (7–9).

The concentration gradient of HB reflects the action of the NANOS (NOS) protein (10–13). Maternally supplied *nos* mRNA is itself regulated so as to create a gradient of NOS protein with its highest concentration at the posterior pole (12, 14–18). High concentrations of NOS produce high

levels of HB repression. As a consequence, the HB and NOS protein gradients are complementary. NOS is a metalloprotein with nonspecific affinity for RNA (19), but its biochemical mechanism is not understood.

Translational regulation of *hb* mRNA also requires the protein PUMILIO (PUM), which is distributed uniformly along the anteroposterior axis (20–22). Like *nos*, *pum* mRNA is synthesized during oogenesis and then transported to the developing egg. PUM binds two copies of a bipartite regulatory sequence in the *hb* 3′ untranslated region (UTR), the Nanos Response Element (NRE; 10, 23). Each NRE consists of a 5′ conserved pentanucleotide box A and a 3′ hexanucleotide box B sequence motif (24). Both NOS and PUM are required for deadenylation of NRE-containing mRNAs injected into embryos, suggesting that translational repression of *hb* is a consequence of deadenylation (25). Alternatively, NRE-directed poly(A) shortening and translational repression may be distinct processes that each require both NOS and PUM.

PUM is a 158 kDa, sequence-specific RNA-binding protein. The 39.6 kDa PUM RNA-binding domain, termed the *Drosophila melanogaster* PUM-Homology Domain (DmPUM-HD), contains 8 imperfect, 36-amino acid repeats that define a large class of evolutionarily conserved proteins found in animals, plants, and fungi. This class is referred to as the PUM-HD or the PUF domain family (26–29). No regions of PUM beyond the PUM-HD are required for specific recognition of the NRE in vitro. Furthermore, the isolated DmPUM-HD retains some function in translational repression in vivo (29).

[†] P.D.Z. is supported by a Charles P. Hood postdoctoral fellowship. R.L. is an investigator of the Howard Hughes Medical Institute. This work was supported by a grant from the NIH (GM-53320) to J.R.W.

* Corresponding author.

[‡] Massachusetts Institute of Technology.

[§] The Skirball Institute.

^{||} The Skaggs Institute for Chemical Biology.

¹ Abbreviations: aa, amino acid; Dm, *Drosophila melanogaster*; GST, glutathione S-transferase; *hb*, *hunchback*; HD, homology domain; NOS, NANOS; NRE, Nanos Response Element; PUM, PUMILIO; UTR, untranslated region.

To date, at least 22 PUM-HD proteins have been identified (A. Kuspa, personal communication; 27, 28). As with *Drosophila* PUM, human and *C. elegans* PUM-HD proteins have been shown to be sequence-specific RNA-binding proteins (27, 28). Like PUM, the *C. elegans* PUM-HD protein, FBF, regulates the expression of an mRNA (*fem-3*) by binding a 3' UTR sequence element (28), suggesting that other PUM-HD proteins also regulate mRNA stability or translation.

Little is known about how the PUM-HD functions in RNA binding. Detailed analysis of the RNA-binding properties of PUM-HD proteins has been difficult because recombinant PUM-HDs from several species are partially or completely insoluble unless fused to a protein such as glutathione *S*-transferase or maltose-binding protein, and these fusion protein partners significantly alter the properties of the PUM-HD proteins (unpublished observations). Furthermore, even when fused to another protein, the majority of soluble recombinant PUM-HD protein from humans, flies, and worms has been inactive (unpublished observations), preventing accurate determination of the affinity of the DmPUM-HD for the RNA. In the absence of such data, it has been difficult to assess the prevailing model that PUM protein bound stably to the NRE recruits other regulatory factors, such as NOS, to the RNA (23).

The DmPUM-HD recognizes sequences in both the 5' and 3' halves of each NRE. In particular, UGU triplets in both the 5' and 3' half-sites of NRE1 are critical for binding (27). Others have recently proposed that recognition of both 5' and 3' half-sites by PUM may reflect binding of a PUM dimer to a single NRE (29). Alternatively, given the repeated structure of DmPUM-HD, a single PUM protein could bind the entire NRE or even multiple NREs.

To determine the RNA-binding affinity and stoichiometry of a representative of the PUM-HD family, we have devised a method to express and purify active, soluble DmPUM-HD. Using this recombinant protein, it has been possible to quantitatively characterize the RNA-binding properties of a PUM-HD protein. We have determined the affinity of DmPUM-HD for the NRE and analyzed the multimerization state of the DmPUM-HD free in solution and bound to RNA. We find that the DmPUM-HD is a monomer in solution and that a single monomer binds to each NRE with substantially higher affinity than previously appreciated. Our *in vitro* results together with our estimation of the intraembryonic concentration of PUM suggest that *in vivo* most NREs will be occupied by PUM monomers.

MATERIALS AND METHODS

Protein Expression. DNA encoding the RNA-binding domain of PUM (amino acids 1091–1433) was inserted into the vector pTYB3 (New England Biolabs) to create an N-terminal fusion with the intein/chitin-binding domain protein. To facilitate intein-mediated cleavage of the fusion protein, a glycine codon was added between PUM amino acid 1433 and the catalytic cysteine residue at the N-terminus of the intein. The resulting plasmid, pTYB3-DmPUM-HD, was expressed in *E. coli* strain ER2566 (New England Biolabs) in LB broth containing 100 μ g/mL ampicillin. Twenty-five milliliters of a saturated overnight culture was added per 1 L of broth and grown at 37 °C until the OD

600 nm was 1.3. The culture was then cooled to room temperature and grown another 2 h in the presence of 0.5 mM isopropylthiogalactopyranoside. The bacteria were harvested by centrifugation, and the pellet was frozen overnight at –70 °C.

Protein Purification. The frozen bacterial pellet from 1 L of original culture was resuspended in 50 mL of ice-cold lysis buffer [50 mM tris(hydroxymethyl)aminomethane (Tris) hydrochloride, pH 8.0, 500 mM NaCl, 1 mM ethylenediaminetetraacetate (EDTA), 1 mM ethylene glycol-*O,O'*-bis-(2-aminoethyl)-*N,N,N',N'*-tetraacetate (EGTA), 5 mM dithiothreitol (DTT)] containing 1 mg/mL 4-(2-aminoethyl)benzenesulfonfyl fluoride (AEBSF) and 0.3 mM phenylmethanesulfonyl fluoride and lysed by sonication at 200 W with a standard tip (30 s on, 2 min off) for a total process time of 10 min. The lysate was cooled during sonication by a salt/ice bath. Insoluble material was removed by centrifugation (14000g, 20 min), and the supernatant was incubated for 1 h in batch with 12.5 mL of chitin–agarose (New England Biolabs) per liter of original culture. The chitin–agarose was then packed into a 2.5 cm diameter column. The column was washed extensively with lysis buffer, and cleavage of the bound fusion protein was initiated by equilibrating the chitin–agarose with lysis buffer containing 100 mM DTT. The air above the column was then purged with argon, and the column was capped and incubated >3 days at room temperature to ensure nearly complete cleavage. The column was then eluted with 3 column volumes of lysis buffer containing 6 M guanidine hydrochloride, 50 mM DTT, and 20% (v/v) glycerol. To refold the protein, the eluate was dialyzed 2 times, 18 h each, against 13 volumes of 2 \times binding buffer [20 mM 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES)–KOH, pH 7.4, 100 mM KCl, 1 mM EDTA, 0.02% (w/v) Tween-20] containing 1 mM EGTA, 20% glycerol, and 5 mM DTT. Insoluble material was removed by centrifugation, and the supernatant was incubated in batch for 1 h at 4 °C with 12.5 mL of fresh chitin–agarose (equilibrated in dialysis buffer) per liter of original culture. The chitin–agarose was removed by filtration through Whatman #1 filter paper prewetted in dialysis buffer, and the filtrate was concentrated ~100-fold using a Centriprep-30 concentrator (Amicon). The concentrated protein was applied to a Sephacryl S-200 HR, HiPrep 16/60 gel filtration column (Pharmacia) equilibrated in column buffer (20 mM HEPES–KOH, pH 7.4, 150 mM KCl, 1 mM EDTA, 5 mM DTT), and the protein was eluted at 0.8 mL/min. Fractions from a single major peak eluting at ~58 mL were pooled, concentrated to ~70 μ M in a Centricon-30 unit (Amicon), diluted with 3 volumes of column buffer containing 75% (v/v) glycerol, and stored at –20 °C. Approximately 0.5 mg of purified DmPUM-HD was isolated per liter of bacterial culture. The protein was stable with no loss of specific activity for at least 1 month.

To isolate the DmPUM-HD-intein–chitin-binding domain fusion protein used in Figure 5, the chitin–agarose column was eluted prior to the DTT-induced cleavage step. The protein was then refolded by dialysis and concentrated as described. The uncleaved fusion protein was stored at 4 °C in dialysis buffer without purification by gel filtration. The molar specific activity of the refolded fusion protein was approximately 10-fold lower than that of the purified DmPUM-HD.

Purity and Concentration Determination. DmPUM-HD purity was greater than 90% both by SDS–polyacrylamide gel electrophoresis and by gel filtration (monitored by both UV absorbance at 280 nm and refractive index). Furthermore, because the purified DmPUM-HD but not the intact fusion protein contains no tryptophan, the absence of tryptophan peaks in the spectrum indicated successful purification. Protein concentration was estimated in 6 M guanidine hydrochloride at 280 nm using a calculated extinction coefficient of $20\,860\text{ M}^{-1}\text{ cm}^{-1}$ which was confirmed by quantitative amino acid analysis of the protein sample prior to the addition of glycerol.

Gel Filtration. Purified DmPUM-HD (200 μL) was applied to a Superdex 200 HR 10/30 (Pharmacia) column equilibrated in $2\times$ binding buffer containing 2 mM DTT and was eluted at 0.4 mL/min. Size standards (BioRad) were chromatographed immediately prior to DmPUM-HD.

Amino Acid Analysis and HPLC Size-Exclusion Chromatography/Laser Light Scattering. Amino acid and light-scattering analyses were performed by the Keck Foundation Biotechnology Resource Laboratory at Yale University (<http://info.med.yale.edu/wmkeck/prochem.htm>; 30–32). Samples for light-scattering analysis were exchanged into 20 mM HEPES–KOH, pH 7.4, 100 mM KCl, 1 mM EDTA, 1 mM DTT by rapid desalting through Biospin P6 (BioRad) immediately before injection onto a Superdex 200 HR gel filtration column.

Native Gel Analysis. Transcription of 181 nt RNAs has been described previously (1, 27). Transcription of the 60 nt RNA (5'-GGG CAG AUC UUC ACU AUU AUC AUA UAA UCG UUG UCC AGA AUU GUA UAU AUG GAU CCC UUU-3') from a polymerase chain reaction template was as described (27). Short RNAs (37, 30, 26, and 22 nts) were transcribed from single-stranded oligonucleotide templates that included a double-stranded T7 RNA polymerase promoter (33), and were subsequently radiolabeled with poly-(A) polymerase and [α - ^{32}P]cordycepin (34). Binding reactions and native gels were as described previously (27). In equilibrium binding experiments, the ^{32}P -labeled RNA concentration was approximately 37 pM (2500 cpm). Native gels were dried and exposed to Fujix phosphor screens, scanned with a Fujix Bas2000, and analyzed with MacBas 2.5 (Fuji) software. Binding data were fit to the Langmuir isotherm as described in (35). Hill coefficients (n) were estimated by nonlinear least-squares fitting of the data to the equation:

$$\text{fraction of sites occupied} = \frac{[\text{protein}]^n}{K_D + [\text{protein}]^n}$$

using IgorPro 3.03 (Wavemetrics) software. Relative K_D values were calculated as described in (27). For stoichiometric binding experiments, the linearly increasing portion of the titration was fit to

$$\text{fraction bound} = m \left(\frac{[\text{protein}]}{[\text{RNA}]} \right) + b$$

using Kaleidagraph 3.05 (Abelbeck) software. Data points near the breakpoint in the titration were excluded to ensure $r^2 > 0.98$. Data beyond the break point were fit to a line with $m = 0$.

Preparative Gel Analysis. Preparative gel electrophoresis was conducted using 1.5 mm, 6% gels with 1.5 cm wide wells. BSA was omitted from the binding reactions (40 μL). Reactions contained approximately 3.8 μM RNA and 1 μM protein. For amino acid analysis, the complexes were detected by UV shadowing, and the appropriate regions of the gel excised.

Unimolecular Dissociation Rate. Off-rate time courses were initiated by diluting a 50 μL binding reaction 2-fold in binding buffer containing a 1000-fold excess of unlabeled over radiolabeled RNA (a 75-fold excess over protein) and adding Ficoll-loading buffer. Measurements were taken at protein concentrations 3–5-fold greater than the K_D . At each time point, a sample of the quenched binding reaction was loaded directly onto a native gel running at 2 W. After the last sample was loaded, the gel was run at 5 W for 1 h. k_{off} was calculated by fitting the data using IgorPro 3.03 software to the equation:

$$\text{fraction bound} = a + be^{-k_{\text{off}}t}$$

k_{off} was determined for the dissociation of a DmPUM-HD-NRE complex from an RNA containing a mutant NRE1 and a wild-type NRE2 and for the dissociation of two DmPUM-HD-NRE complexes to one DmPUM-HD-NRE complex from an RNA bearing two wild-type NREs. The latter rates were divided by 2 to obtain the rate for DmPUM-HD complex dissociating from a single NRE.

Estimation of the Intraembryonic PUM Concentration. Oregon R embryos were collected on yeasted apple juice agar plates at 25 °C for 105 min, washed with 0.7% (w/v) NaCl, dechorionated with 50% (v/v) bleach (Clorox), washed well with water, and dried with bibulous paper. Embryos were lysed in SDS-sample buffer containing 100 mM DTT and 0.1 mg/mL AEBSF. The total time from harvesting the embryos to lysis was about 15 min. Embryo lysates and 2-fold dilutions of purified GST-PUM I protein (27) were analyzed by standard Western blotting procedures. The nitrocellulose blot was blocked 30 min at room temperature, incubated overnight at 4 °C with rabbit anti-GST-PUM-I, diluted 1:5000, washed 5 times (10 min each), incubated 1 h at room temperature with ^{125}I -protein A (0.2 $\mu\text{Ci}/\text{mL}$), washed again, and exposed for 2 days to a phosphor screen. All incubations and washes were in phosphate-buffered saline containing 5% (w/v) nonfat dry milk and 0.1% (w/v) Tween-20.

RESULTS

Our previous studies of PUM identified the PUM-HD as the minimal RNA-binding domain and estimated its K_D to be 25–100 nM (27). However, the specific activity of the fusion protein used in those studies was very low, apparently due to aggregation. To obtain active, soluble protein, we have developed a new strategy for expressing and purifying recombinant DmPUM-HD. DmPUM-HD was expressed as a fusion protein using an intein construct that was cleaved with DTT under argon. The protein was refolded at low concentration, and monomeric protein was separated from aggregates by gel filtration. Storage in 50% glycerol stabilized the purified protein against reaggregation. With this improved protein preparation, it has now been possible to

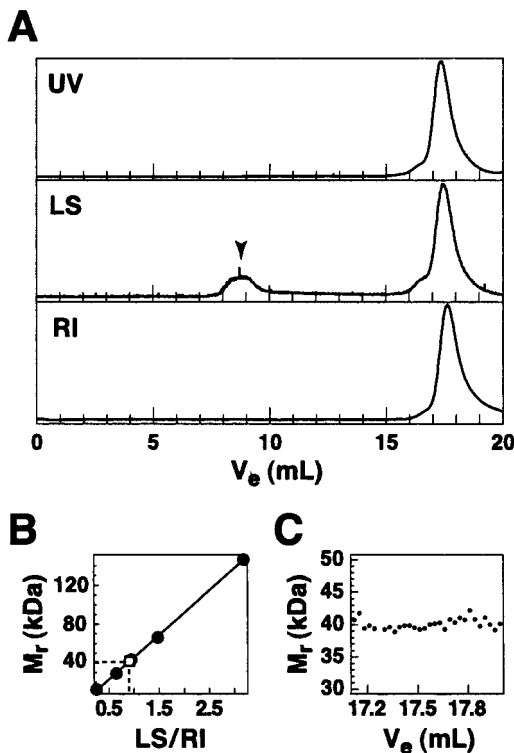


FIGURE 1: Native molecular mass determination by HPLC-size exclusion/laser light scattering. (A) Purified DmPUM-HD was fractionated by gel filtration. The effluent was monitored by UV absorbance (280 nm), light scattering (LS), and refractive index (RI) as a function of elution volume (V_e). The arrowhead above the light-scattering trace marks the position of the void volume of the gel filtration column. (B) The native molecular mass of the single DmPUM-HD peak (\square) observed in the UV and refractive index traces was calculated from its LS/RI ratio relative to a linear plot of LS/RI versus native molecular mass for known protein standards (\bullet). (C) The native molecular mass calculated from the LS data is shown for every 30 μ L slice across the DmPUM-HD peak.

measure the RNA binding affinity, cooperativity, and stoichiometry of a PUM-HD protein.

DmPUM-HD Is a Monomer in Solution. To determine the multimerization state of the free DmPUM-HD, the apparent size of the 39.6 kDa protein was analyzed by gel filtration at two different concentrations, ~ 60 and ~ 1 μ M. A single peak was detected with an apparent molecular mass of 32 kDa relative to protein size standards, and no peaks were observed in the region of the chromatogram appropriate for dimers or higher multimers.

Gel filtration analysis can result in an underestimate of molecular mass for proteins that interact with the gel filtration column matrix or that have elongated shapes. Therefore, the molecular mass was also estimated by HPLC-size exclusion/laser light scattering (30). In this technique, the protein is first resolved by gel filtration chromatography, and the light scattering (LS) and refractive index (RI) of the eluate are monitored (Figure 1A). The apparent molecular mass of the protein within a particular peak is determined from its LS/RI ratio relative to a linear plot of LS/RI versus molecular mass for known protein standards (Figure 1B). The LS/RI ratio for the peak of DmPUM-HD from gel filtration corresponds to a molecular mass of ~ 41.6 kDa, nearly identical to the calculated molecular mass of a monomer (39.6 kDa). The estimation of apparent molecular mass by

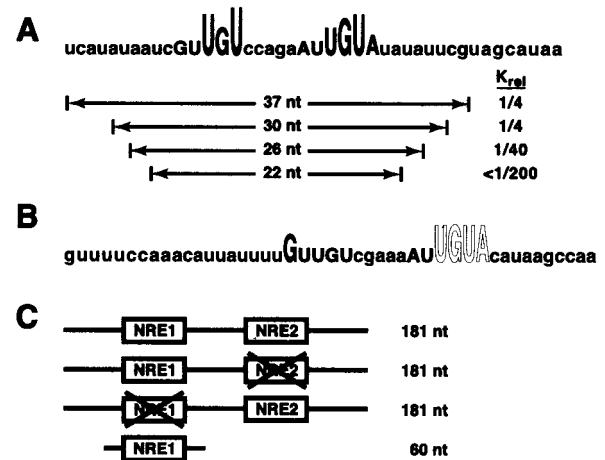


FIGURE 2: (A) NRE1 and flanking sequences. Capital letters indicate nucleotides conserved between NRE1 and NRE2, and between *Drosophila melanogaster hb* and *hb* from other insect species. The relative importance of specific nucleotides for DmPUM-HD binding is indicated by the height of the letters (log scale). Arrows denote the sequences of the 37, 30, 28, and 26 nt RNAs used in binding studies. The 37 nt RNA includes three and the 30, 26, and 22 nt RNAs include two additional 5'-guanosines to facilitate in vitro transcription. The binding affinity of DmPUM-HD, relative to 60 and 181 nt RNAs containing a single functional NRE, is indicated. (B) NRE2 and flanking sequences. The sequence begins with the *hb* 3' UTR nucleotide immediately following the last nucleotide in (A). Outlined text indicates nucleotides that were tested only in pairwise combination (UG \rightarrow AC and UA \rightarrow GC; 29). Effect of mutations on binding is indicated by the height of the letters (log scale). (C) RNAs used for binding affinity and stoichiometry measurements.

this method is independent of the elution position of the protein from the gel filtration column (30). Additionally, the molecular mass can be calculated from the LS and RI data without reference to protein standards (31, 32). The molecular mass determined directly, 40 kDa, agrees well with that estimated from the LS/RI ratio. The molecular mass of DmPUM-HD is constant across the gel filtration peak, indicating that it contains a single, monomeric species (Figure 1C). At the concentrations tested, no multimer peaks were detected. Based on the gel filtration and light-scattering data, we conclude that at concentrations as high as 60 μ M, the DmPUM-HD is a monomer.

The DmPUM-HD Binds Independently and Tightly to Each NRE. In vivo, identical mutations in each NRE are not equivalent: mutations in NRE1 more strongly disrupt abdominal development than do mutations in NRE2 (19). Thus, it is of interest to know if this difference reflects a difference in the affinity of PUM for each NRE. Consistent with PUM binding the two NREs with different affinities, the nucleotides important for PUM binding NRE1 and NRE2 have been reported to differ. Seven nucleotides (5', UGU; and 3', UGUA) were shown to be essential for high-affinity binding to NRE1 (Figure 2A; 27), whereas only a single 5' nucleotide (the initial guanosine) and not more than four 3' nucleotides (UGUA) were important for binding NRE2 (Figure 2B; 29). These data suggest that the DmPUM-HD might discriminate between NRE1 and NRE2.

To determine if the DmPUM-HD binds with different affinities to the two NREs, the equilibrium dissociation constant (K_D) of the DmPUM-HD was measured for each NRE, individually and together, using mobility shift experi-

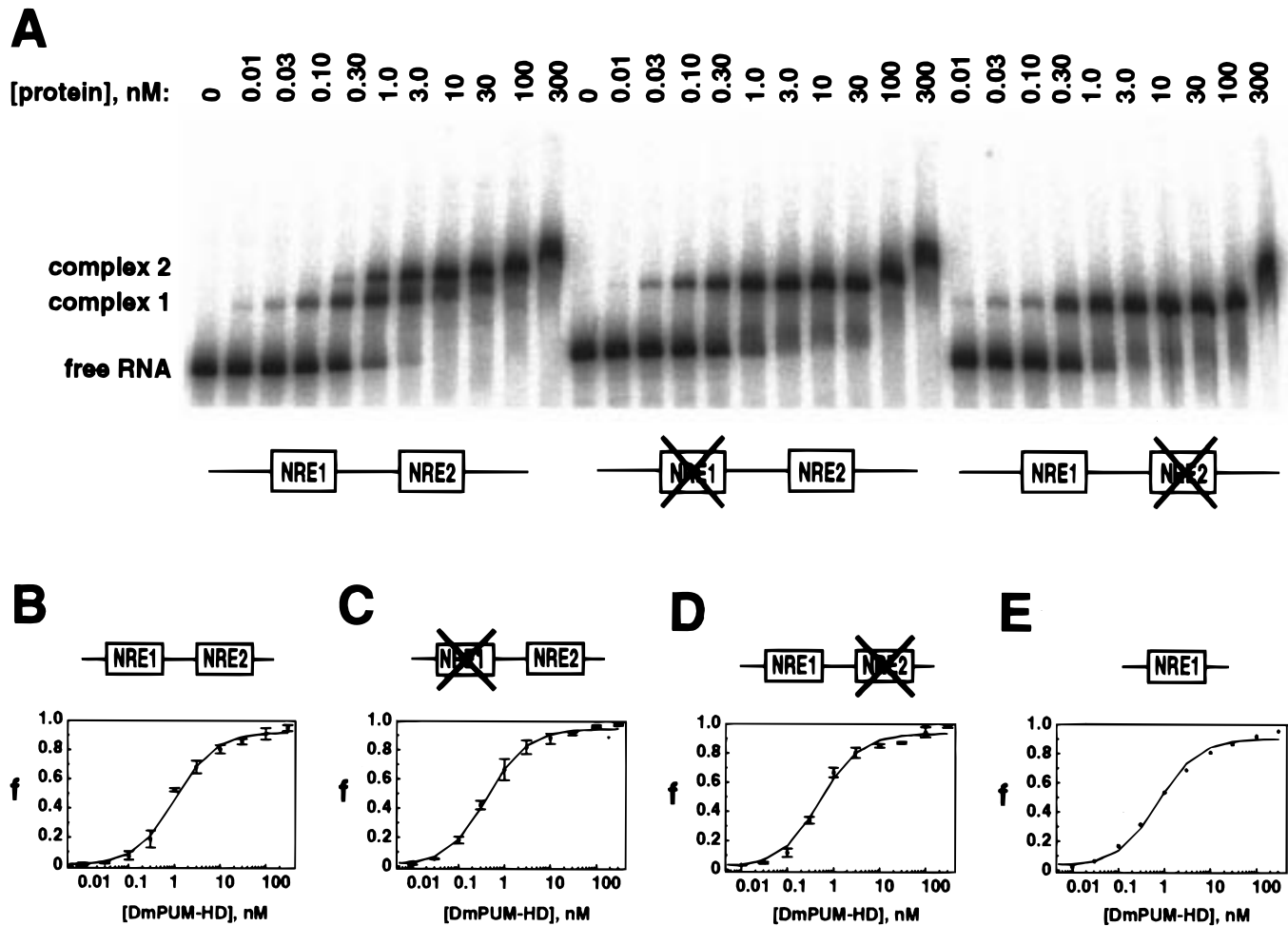


FIGURE 3: Analysis of DmPUM-HD binding to NRE1 or NRE2, and to both NREs together. (A) A representative native gel analysis of the binding of DmPUM-HD to the RNAs diagrammed below each panel. (B–D) Analysis of the equilibrium data according to the model that a single DmPUM-HD protein binds a single NRE. For each RNA (indicated above each panel), the average value of the fraction of sites occupied (f) for three trials is plotted for each protein concentration. Error bars indicate \pm standard deviation. The curves are the binding isotherms that correspond to the average K_D of the three trials. (E) Protein concentration versus the fraction of sites occupied (f) for a 60 nt RNA containing NRE1. The curve represents the best-fit binding isotherm to the data.

Table 1: Summary of Hill Coefficients and Binding Constants^a

NRE1	NRE2	length (nt)	Hill coeff	K_D (nM)	trials
wild type	wild type	181	1.00 ± 0.16	1.04 ± 0.29	3
mutant	wild type	181	0.97 ± 0.22	0.42 ± 0.07	3
wild type	mutant	181	1.16 ± 0.23	0.51 ± 0.08	3
wild type	none	60		0.73	1

^a For each 181 nt RNA, the Hill coefficient and dissociation constant (K_D) are reported as the average value \pm standard deviation. The K_D values were calculated using a Hill coefficient of 1.

ments. The RNAs used in these studies are diagrammed in Figure 2C. The RNAs contained the NRE in the context of the normal flanking sequences of the *hb* 3' UTR. In the mutant NRE, three conserved guanosines were replaced by uracil. This mutation disrupts NRE function *in vivo* (19), and eliminates DmPUM-HD binding *in vitro* (27). The 60 nt RNA contained only NRE1 and flanking sequences. Figure 3A shows a representative mobility shift assay, Figure 3B–E shows the analyses of the binding data, and Table 1 summarizes the results of our experiments. No difference is observed between the binding of DmPUM-HD to NRE1 (Figure 3D; $K_D = 0.51 \pm 0.08$ nM) and NRE2 (Figure 3C; $K_D = 0.42 \pm 0.07$ nM). An RNA containing two NREs

showed similar binding affinity (Figure 3B; $K_D = 1.04 \pm 0.29$ nM for each site), consistent with the idea that binding to each NRE is identical and independent. A single NRE in the context of only 60 nts bound with essentially the same affinity as the 181 nt RNAs (Figure 3E; $K_D \sim 0.73$ nM). However, shorter RNAs either reduced (37, 30, and 26 nt RNAs) or abolished (22 nt RNA) RNA binding, implying that sequences immediately flanking the NRE contribute to binding energy or RNA structure. Figure 2A indicates the sequences of these four shorter RNAs and summarizes the decrease in DmPUM-HD binding observed relative to the 181 nt RNAs containing one wild-type and one mutant NRE.

A Single DmPUM-HD Monomer Binds to Each NRE. Although the DmPUM-HD is a monomer free in solution, it was possible that a stable dimer forms upon RNA binding. The number of DmPUM-HD proteins bound was determined by direct titration of the 181 nt RNAs containing one or two NREs (Figure 4). The binding stoichiometry of DmPUM-HD for each RNA was determined from the amount of protein required to saturate binding to a fixed amount of RNA (Figure 4A,C), and also from the amount of RNA required to saturate binding to a fixed amount of protein (Figure 4B,D). The concentration of the fixed component was 100–

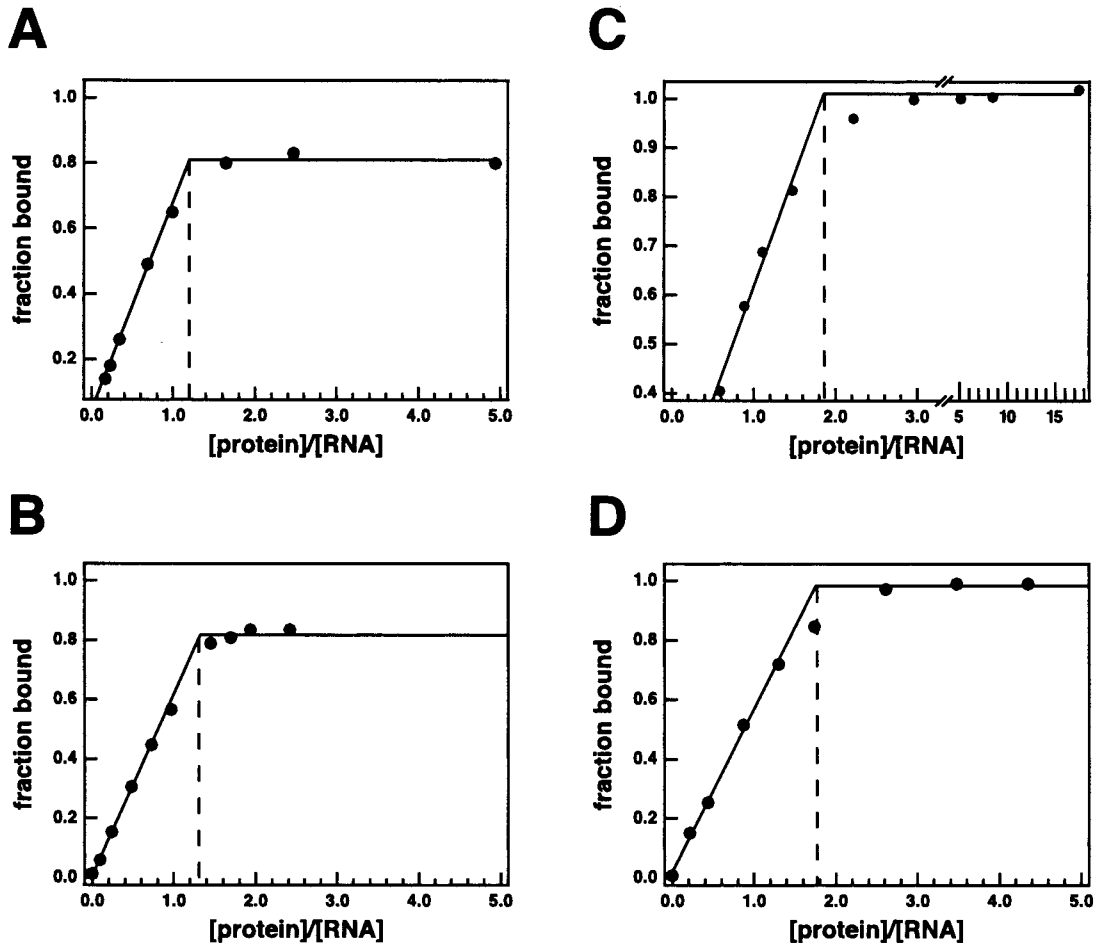


FIGURE 4: Stoichiometric binding of DmPUM-HD to RNA containing wild-type NRE1 and mutant NRE2 (left panels) or two wild-type NREs (right). In (A) and (C), a constant amount of protein (510 nM) was titrated with increasing amounts of RNA. In (B) and (D), a fixed amount of RNA (102 and 115 nM) was titrated with increasing amounts of protein.

1000-fold greater than the K_D to permit direct stoichiometric titration. Using the two complementary methods, we find that ~ 1.2 proteins bind the RNA containing a single wild-type NRE and that ~ 1.8 proteins bind the RNA containing two NREs. These data agree closely with a binding stoichiometry of 1 protein per NRE.

Because interpretation of the stoichiometry experiments assumes that all of the recombinant protein is active, mixing experiments were performed as an independent measure of stoichiometry that is unaffected by inactive protein. If a single DmPUM-HD monomer binds each NRE, then upon incubation of a mixture of a large protein (DmPUM-HD/CBD fusion) and a small protein (DmPUM-HD) with RNA containing only one functional NRE, complexes will be observed that contain either the large protein or the small protein, but not both. As predicted, when the large and small proteins were incubated with the single NRE RNA (Figure 5, lanes 10–18), only the RNA–protein complexes that form when the proteins are incubated individually were observed. When the proteins were incubated with an RNA containing two NREs (lanes 1–9), five protein–RNA complexes were observed corresponding to one small (complex 1S) or one large protein bound (complex 1L), two small (complex 2S) or two large proteins bound (complex 2L), and an additional complex with intermediate mobility consistent with one small and one large protein bound (arrowheads, lanes 9 and 14). If the DmPUM-HD bound as a dimer, a heterodimer of large

and small proteins bound to the single NRE RNA would be expected. In principle the heterodimer complex might not be observed because the preexisting large and small homodimers are exceptionally stable. This is unlikely, since the DmPUM-HD is a monomer free in solution at concentrations more than 1000-fold greater than those used in Figure 5.

We favor the model that a single protein binds a single NRE, rather than a dimer of DmPUM-HD binds two NREs, because in experiments in which the DmPUM-HD was incubated with two NRE-containing RNAs of different lengths (Figure 6), only the complexes that formed when each RNA was incubated with the protein separately were observed. These RNA mixing experiments show that the DmPUM-HD can only make stable contacts with one NRE at a time.

As a further test of the model that a single DmPUM-HD binds each NRE, the 1S and 2S protein–RNA complexes were isolated from a preparative native gel. The amount of protein in the gel slice was determined by quantitative amino acid analysis. The amount of RNA in each gel slice was calculated by determining the radioactivity in the protein–RNA complex, using the specific activity of the RNA. Quantitation of the protein and RNA in each complex yields a stoichiometry of 0.9 ± 0.1 protein per RNA (average of two trials \pm standard deviation) for the faster mobility

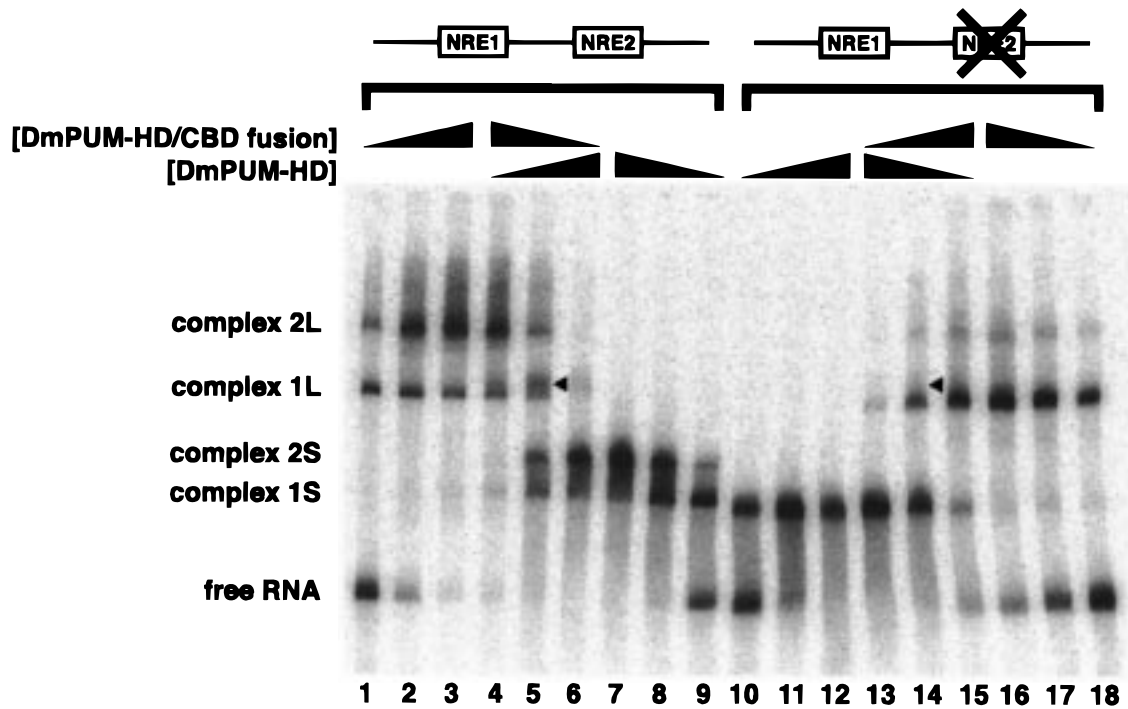


FIGURE 5: Native gel analysis of NRE binding by a 97.2 kDa DmPUM-HD/chitin-binding domain fusion protein, by the 39.6 kDa DmPUM-HD protein, and by mixtures of the two. In lanes 1–3 and 16–18, only fusion protein was added; in lanes 7–14, only DmPUM-HD. Lanes 4–6 and 13–15 contain mixtures of both proteins. For lanes 1–9, the RNA contains two NREs, whereas for lanes 10–18, the RNA contains a wild-type NRE1 and a mutant NRE2. The arrowheads indicate the position of the complex that forms on the two-NRE-containing RNA only when both the 97.2 and 39.6 kDa proteins are present. For each protein, the concentration increased ~ 6 -fold across the bases of the triangles. The small fraction of complex 2L formed with the fusion protein in lanes 14–18 likely results from nonspecific interactions between chitin-binding domains.

complex (1S) and 1.6 ± 0.2 proteins per RNA for the slower (2S).

In sum, three independent methods to determine the binding stoichiometry of the DmPUM-HD for the NRE support the proposal that a single monomer of the DmPUM-HD binds each NRE.

The DmPUM-HD Binds Noncooperatively to the NREs. The affinity of the DmPUM-HD for an RNA containing two NREs is no greater than its affinity for an RNA bearing a single wild-type NRE (see above). These data suggest the DmPUM-HD binds each NRE independently: i.e., a DmPUM-HD protein bound to one NRE does not increase the affinity of the protein for an adjacent NRE. As a further test of the hypothesis that the DmPUM-HD binds the two NREs in the *hb* 3'UTR noncooperatively, we examined our binding data by Hill analysis, using the model that a single DmPUM-HD protein binds each NRE. Binding data for both the single- and double-NRE RNAs give Hill coefficients close to 1.0 (Table), indicating that no cooperative interactions occur between DmPUM-HD proteins bound at adjacent sites.

In Vivo Concentration of PUM. To apply our *in vitro* measurements of PUM binding to the embryo requires knowledge of the intraembryonic concentration of PUM. For up to 2 h after fertilization, the *Drosophila* embryo is a syncytium of nuclei in a common cytoplasm. Antibody staining of whole embryos indicates that PUM protein is uniformly distributed in the cortical cytoplasm at this time (D. Barker and R. Lehmann, unpublished results; 22). It is therefore reasonable to estimate the *in vivo* PUM concentration as the total number of moles of PUM divided by the total volume of water in one embryo. We have used

quantitative Western blotting to measure the number of moles of PUM in an embryo at this time in development (Figure 7). The amount of full-length PUM in total lysates of a known mass of 15 min to 2 h old dechorionated embryos was compared to known amounts of the recombinant PUM derivative used to generate our anti-PUM antibody. By this assay, each embryo contains 320 ± 60 amol of PUM (average \pm standard deviation). Assuming the effective intracellular volume of an embryo (7.3 nL) is equal to the volume of water in an embryo (36), the intracellular concentration of PUM is 44 ± 8 nM.

Does PUM binding reach equilibrium *in vivo* early in development? To assess this idea, we estimated how rapidly equilibrium is reached in our *in vitro* experiments. Using the mobility shift assay, the unimolecular dissociation rate of the DmPUM-HD–NRE complex is $\sim 0.29 \text{ min}^{-1}$ ($t_{1/2} = 2.4 \text{ min}$; Figure 8). Using our K_D measurements to calculate k_{on} ($=k_{off}/K_D$), the time constant for the reaction reaching equilibrium, $\tau = 1/(k_{on}[PUM] + k_{off})$, is about 3 s for an intracellular PUM concentration of 44 nM. This analysis assumes that *in vivo* the *hb* RNA concentration is less than that of PUM, which is reasonable given that a 2-fold reduction in wild-type PUM protein does not perturb normal development. It is therefore reasonable to model the binding of PUM to *hb* mRNA in the cytoplasm as at or near equilibrium for most of the first 2 h of development. Since the K_D for PUM binding the NRE is < 1 nM, and the intracellular concentration of PUM is > 44 nM, the inherent binding affinity of PUM would be sufficient for nearly all NREs to be occupied by a PUM protein during early development.

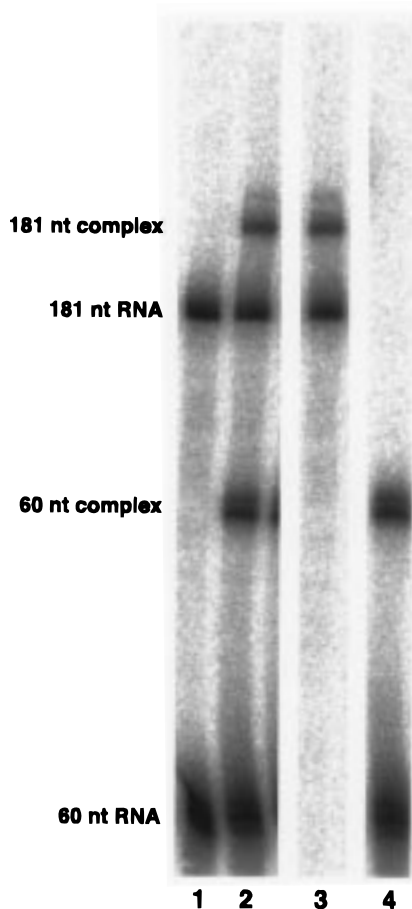


FIGURE 6: Native gel analysis of 0.5 nM DmPUM-HD binding NRE1-containing RNAs of different lengths. The positions of free RNA and RNA-protein complexes are indicated. Lane 1, no protein, 60 nt RNA only; lane 2, protein plus both RNAs; lane 3, protein plus 181 nt RNA only; lane 4, protein plus 60 nt RNA only. The three panels are all from the same image of the same gel.

DISCUSSION

We have determined the *in vitro* binding affinity and stoichiometry of the DmPUM-HD for RNAs containing either NRE1 or NRE2 or both NREs together. *In vitro*, a single DmPUM-HD binds each NRE independently. We detect no cooperative interactions between DmPUM-HD proteins bound to adjacent sites, but whether cooperativity is observed for binding of the intact 158 kDa protein remains to be tested. Nonetheless, we note that cooperativity is not a feature of the isolated RNA-binding domain, which by itself retains some function in translational regulation *in vivo* (29).

The finding that a single DmPUM-HD monomer binds an NRE implies that the protein makes contacts with nucleotides in both the 5' and 3' half-sites of the NRE. It is intriguing that the DmPUM-HD, which itself contains 8 direct repeats of a 36-amino acid motif, binds an RNA site that contains a repeated sequence element (UGU). One explanation would be that the DmPUM-HD is itself a pseudodimer, comprising two covalently linked RNA-binding subdomains. However, previous studies have found no PUM-HD subfragments that retain RNA-binding activity (27–29). It will be of interest to test if different repeats within the PUM-HD contact 5' and 3' half-site nucleotides.

Perhaps other PUM-HD proteins bind RNA as a monomer. Supporting this speculation, the RNA-binding domain of the

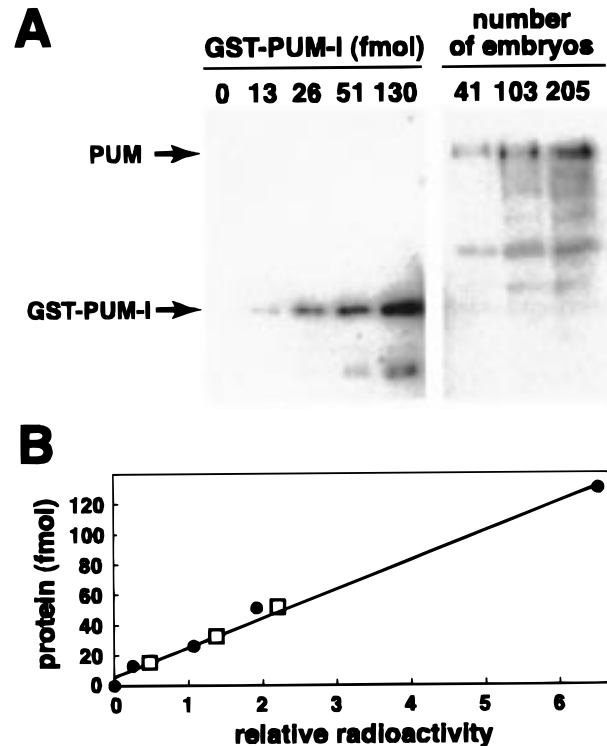


FIGURE 7: Quantitative Western blotting. (A) Western blot of known amounts of GST-PUM-I protein and of full-length PUM in 41, 103, and 205 embryos. (B) The amount of full-length PUM in 41, 103, and 205 embryos (□) was estimated from a standard curve of the band intensities of full-length GST-PUM-I (●). The concentration of full-length GST-PUM-I was determined in a parallel experiment by quantitative amino acid analysis of a gel slice corresponding to the full-length protein. The intraembryonic concentration of PUM was calculated assuming 9.5 $\mu\text{g}/\text{embryo}$ and 7.3 μg of water/embryo (36).

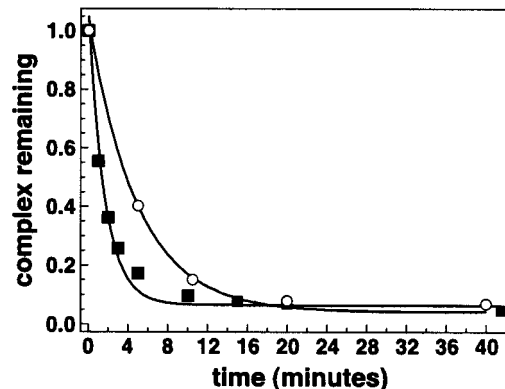


FIGURE 8: Representative data used to determine the unimolecular dissociation rate. Open circles (○) indicate the amount of 1S complex remaining for a 181 nt, wild-type NRE1-containing RNA, and closed squares (■) indicate the amount of 2S complex remaining for a 181 nt RNA containing both wild-type NRE1 and NRE2. The curves correspond to the best fit to the data.

human protein KIAA0099, whose function is unknown, has a mobility in native gels similar to that of DmPUM-HD (unpublished results) and forms similar fast- and slow-migrating complexes of RNAs with two NREs (27).

Surprisingly, our data suggest that the functional difference between the two NREs observed *in vivo* does not arise from a difference in affinity of PUM for NRE1 and NRE2; the DmPUM-HD binds each NRE with equal affinity. Instead, it must reflect functional differences between the position

of the two NREs or differences in the binding of a second factor to the NREs. Our finding that the DmPUM-HD binds both NREs with similar affinity seems at odds with published data that PUM recognizes seven nucleotides in NRE1 (27) but no more than five in NRE2 (29), and suggests that the precise nucleotide requirements for PUM binding to each NRE should be further examined.

Implications for Binding of PUM to NREs in Vivo. What do our *in vitro* binding studies say about the state of PUM *in vivo*? By quantitative Western blotting analysis, the intracellular concentration of PUM is >40 nM, 40–80-fold greater than the K_D of the DmPUM-HD for the NRE. This estimate likely represents a lower bound for the intracellular PUM concentration because losses in protein transfer from the denaturing gel to the nitrocellulose membrane may be greater for PUM (~158 kDa) than the concentration standard (~78 kDa). These data suggest that nearly all NREs are bound by PUM *in vivo*. There is therefore no need to invoke additional factors to achieve a high occupancy of the NREs by PUM.

If NREs throughout the embryo are occupied by PUM, how then is *hb* translation restricted to the anterior? One possibility is that the activity of PUM is directly regulated, for example, by posttranslational modification. We would expect NOS to play a role, directly or indirectly, in this process since its intracellular location determines where *hb* is repressed. The RNA-binding activity of PUM is unlikely to be regulated by NOS because PUM from embryos laid by NOS-mutant females binds RNA as well as that from embryos with wild-type mothers (unpublished observations; 23), but modifications may influence the function of RNA-bound PUM in repressing translation. Alternatively, additional protein factors might associate with NRE-bound PUM to repress translation. The distribution or activity of such factors would be expected to depend on NOS. Recently, Wharton and co-workers (29) mapped a single amino acid mutation in the DmPUM-HD that blocks PUM function *in vivo* but appears not to disrupt RNA binding *in vitro*. As these authors suggest, this mutation may identify a region of PUM that interacts with specific translation factors or recruits additional regulatory proteins to the NRE. Although it is tempting to propose that NOS might be one of these PUM-associated factors, to date no stable interactions between PUM and NOS have been detected (unpublished observations; 29). How and which proteins regulate translational repression by PUM bound to the NREs remain unknown. A critical step toward the identification of all the factors required for regulation of *hb* translation is likely to be the development of a cell-free system that recapitulates NRE-directed translational repression *in vitro*.

ACKNOWLEDGMENT

We thank Christopher Cilley for help constructing pTYB3-DmPUM-HD; Dan Treiber, Jeff Orr, Danny Chagnovich, and members of the Bartel, Lehmann, and Williamson laboratories for many helpful discussions and for encouragement in the course of this work; Michael Feldbrugge and Pamela Green for providing the 60 nt NRE1 transcription template; Adam Kuspa for sharing unpublished findings; and Terry Orr-Weaver for generously providing fly resources. We also thank Ewa Folta-Stogniew and Myron Crawford of the Keck

Foundation Biotechnology Resource Laboratory at Yale University for expertly performing light-scattering and quantitative amino acid analyses.

REFERENCES

- Curtis, D., Lehmann, R., and Zamore, P. (1995) *Cell* 81, 171–178.
- Wickens, M., Kimble, J., and Strickland, S. (1996) in *Translational Control* (Hershey, J. W. B., Mathews, M. B., & Sonenberg, N., Eds.) pp 411–450, Cold Spring Harbor Laboratory Press, Plainview, NY.
- Tautz, D., Lehmann, R., Schnurch, H., Schuh, R., Seifert, E., Kienlin, A., Jones, K., and Jäckle, H. (1987) *Nature* 327, 383–389.
- Tautz, D., and Pfeifle, C. (1989) *Chromosoma* 98, 81–85.
- Lehmann, R., and Nüsslein-Volhard, C. (1987) *Dev. Biol.* 119, 402–417.
- Hülkamp, M., Pfeifle, C., and Tautz, D. (1990) *Nature* 346, 577–580.
- Irish, V., Lehmann, R., and Akam, M. (1989) *Nature* 338, 646–648.
- Struhl, G. (1989) *Nature* 338, 741–744.
- Hülkamp, M., Schröder, C., Pfeifle, C., Jäckle, H., and Tautz, D. (1989) *Nature* 338, 629–632.
- Wharton, R. P., and Struhl, G. (1991) *Cell* 67, 955–967.
- Lehmann, R., and Nüsslein-Volhard, C. (1991) *Development* 112, 679–691.
- Wang, C., and Lehmann, R. (1991) *Cell* 66, 637–648.
- Gavis, E. R., and Lehmann, R. (1992) *Cell* 71, 301–313.
- Wang, C., Dickinson, L. K., and Lehmann, R. (1994) *Dev. Dyn.* 199, 103–115.
- Gavis, E. R., and Lehmann, R. (1994) *Nature* 369, 315–318.
- Dahanukar, A., and Wharton, R. P. (1996) *Genes Dev.* 10, 2610–2620.
- Smibert, C. A., Wilson, J. E., Kerr, K., and Macdonald, P. M. (1996) *Genes Dev.* 10, 2600–2609.
- Gavis, E., Lunsford, L., Bergsten, S., and Lehmann, R. (1996) *Development* 122, 2791–2800.
- Curtis, D., Treiber, D. K., Tao, F., Zamore, P. D., Williamson, J. R., and Lehmann, R. (1997) *EMBO J.* 16, 834–843.
- Lehmann, R., and Nüsslein-Volhard, C. (1987) *Nature* 329, 167–170.
- Barker, D. D., Wang, C., Moore, J., Dickinson, L. K., and Lehmann, R. (1992) *Genes Dev.* 6, 2312–2326.
- Macdonald, P. M. (1992) *Development* 114, 221–232.
- Murata, Y., and Wharton, R. P. (1995) *Cell* 80, 747–756.
- Kuwabara, P. E. (1998) *Curr. Biol.* 8, R278–R281.
- Wreden, C., Verrotti, A. C., Schisa, J. A., Lieberfarb, M. E., and Strickland, S. (1997) *Development* 124, 3015–3023.
- Kennedy, B., Gotta, M., Sinclair, D., Mills, K., McNabb, D., Murthy, M., Pak, S., Laroche, T., Gasser, S., and Guarente, L. (1997) *Cell* 89, 381–391.
- Zamore, P. D., Williamson, J. R., and Lehmann, R. (1997) *RNA* 3, 1421–1433.
- Zhang, B., Gallegos, M., Puoti, A., Durkin, E., Fields, S., Kimble, J., and Wickens, M. P. (1997) *Nature* 390, 477–484.
- Wharton, R. P., Sonoda, J., Lee, T., Patterson, M., and Murata, Y. (1998) *Mol. Cell* 1, 863–872.
- Wen, J., Arakawa, T., and Philo, J. (1996) *Anal. Biochem.* 240, 155–166.
- Wyatt, P. (1993) *Anal. Chim. Acta* 272, 1–40.
- Takagi, T. (1990) *J. Chromatogr.* 506, 409–416.
- Milligan, J., Groebe, D., Witherell, G., and Uhlenbeck, O. (1987) *Nucleic Acids Res.* 15, 8783–8798.
- Lingner, J., and Keller, W. (1993) *Nucleic Acids Res.* 21, 2917–2920.
- Batey, R. T., and Williamson, J. R. (1996) *J. Mol. Biol.* 262, 536–549.
- Mazur, P., Schneider, U., Jacobsen, K., and Mahowald, A. (1988) *Cryobiology* 25, 543–544.