RNA

Separable roles in vivo for the two RNA binding domains of *Drosophila* A1-hnRNP homolog

K. Zu, M. L. Sikes and A. L. Beyer

RNA 1998 4: 1585-1598

References

Article cited in:

<http://www.rnajournal.org/cgi/content/abstract/4/12/1585#otherarticles>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

Notes

To subscribe to *RNA* go to:
<http://www.rnajournal.org/subscriptions/>

Separable roles in vivo for the two RNA binding domains of a *Drosophila* A1-hnRNP homolog

KAI ZU,¹ MARTHA L. SIKES, and ANN L. BEYER

Department of Microbiology, University of Virginia, Charlottesville, Virginia 22908, USA

ABSTRACT

We analyzed the roles of the three domains of a *Drosophila* hnRNP A1 homolog by expression of wild-type and mutant versions of HRB87F/hrp36 in *Drosophila melanogaster*. HRB87F/hrp36 is one of two *Drosophila* proteins that is most similar to mammalian A1 hnRNP, and like A1, consists of two copies of the RNA-binding domain (RBD) motif followed by a glycine-rich domain (GRD). The role of the domains in nuclear localization and RNA binding to polytene chromosomal sites was determined. RBD-1 and the GRD were largely responsible for both the cellular location of the protein and for the typical chromosomal distribution pattern of the protein at sites of *Po* II transcription. RBD-1 also provided a role in the exon-skipping activity of the protein that was not provided by RBD-2. On the other hand, RBD-2 and the GRD were responsible for the very limited chromosomal distribution pattern seen upon heat shock, when HRB87F/hrp36 is sequestered at heat-shock puff 93D, which encodes a long nucleus-restricted RNA. Thus, these studies indicate that the two RBDs function independently of each other but in concert with the GRD. In addition, the self-association property of the GRD was strikingly evident in these overexpressed proteins.

Keywords: alternative splicing; glycine-rich domain; heat shock; RRM

INTRODUCTION

HnRNP proteins are a diverse family of ~20 different nuclear proteins that are isolated in association with pre-mRNA (reviewed in Dreyfuss et al., 1993). The best studied hnRNP proteins are the A/B-type, typified by the human A1 hnRNP protein. These proteins of 30–40 kDa are among the smallest, most basic in charge (pIs of 9–10), and most abundant of the hnRNP family, and have a 2×RBD:gly-rich domain structure, consisting of two copies of the ~80 amino acid RNA-binding domain (RBD or RRM) and a C-terminal glycine-rich domain (GRD). The function of these proteins has been the subject of much debate, with proposed roles in RNA packaging, mRNA export, and alternative RNA splicing. In terms of RNA packaging, it has been argued, based on their nuclear abundance and similar affinity for a wide variety of RNA sequences, that they will bind to all available pre-mRNAs (Abdul-Manan & Williams, 1996). This is supported by in vivo observations that proteins of this family associate with essentially all *Po*

II transcripts as they are being transcribed (Wu et al., 1991; Amero et al., 1992; Matunis et al., 1993), perhaps via nucleation at preferred binding sites (Burd & Dreyfuss, 1994). As discussed (Herschlag, 1995), A1 hnRNP is the best-known example of a nonspecific RNA chaperone, or protein that serves to prevent and resolve RNA misfolding. One of its roles in the nucleus presumably involves prevention of most secondary-structure formation in pre-mRNA, which would be counterproductive to rapid pre-mRNA processing. In addition to this packaging function for A/B hnRNPs, a role in mRNA export was proposed when it was found that A1 hnRNP shuttles continuously between the nucleus and cytoplasm and can be crosslinked to poly(A)⁺ RNA in both compartments (Piñol-Roma & Dreyfuss, 1992). Recent studies in *Xenopus* oocytes support the involvement of A1 in the mRNA export process (Izaurrealde et al., 1997). Lastly, a possible role for A/B hnRNPs in alternative splicing regulation was proposed from results with in vitro splicing systems (Mayeda & Krainer, 1992), and has been both supported (Cáceres et al., 1994; Yang et al., 1994) and challenged (Zu et al., 1996) by in vivo studies.

Properties of protein sequence that distinguish the A/B hnRNP family from other 2×RBD-Gly hnRNP proteins are characteristic amino acid constellations in the two RBDs, a GRD that is indeed very rich in glycine

Reprint requests to: Ann L. Beyer, Department of Microbiology, University of Virginia, HSC 441, 1300 Jefferson Park Ave., Jordan 7-59, Charlottesville, Virginia 22908, USA; e-mail: alb4h@virginia.edu.

¹ Present address: Department of Medicine, School of Medicine, State University of New York at Stony Brook, Stony Brook, New York 11794, USA.

(>40% glycine) and a characteristic pI of 9–10 (Haynes et al., 1991; Birney et al., 1993). There are two proteins in *Drosophila*, HRB98DE/hrp38 and HRB87F/hrp36, that fall into this rigorous A/B class, along with ten other proteins identified to date (Mayeda et al., 1998). Other insect 2×RBD:Gly proteins studied are acidic and are not as closely related to A1 (Siebel et al., 1992; Matunis et al., 1994; Visa et al., 1996). HRB87F/hrp36 protein is the subject of this report. [These proteins were termed HRB98DE and HRB87F to denote the position of their genes on polytene chromosomes (Haynes et al., 1990, 1991) but were renamed hrp38 and hrp36 according to their size when a larger group of fly hnRNP proteins was characterized (Matunis et al., 1992; Matunis MJ et al., 1992).] As discussed previously (Raychaudhuri et al., 1992; Zu et al., 1996), these two fly proteins share all properties tested with vertebrate A/B-type hnRNPs, including general association with *Pol* II transcripts, biochemical recovery in monomer hnRNP particles with bulk poly(A)⁺ pre-mRNA, and ability to promote exon skipping but not alternative 3' splice site selection when overexpressed. Nuclear localization and shuttling of A1 are mediated via a short sequence in the GRD termed M9 (Siomi & Dreyfuss, 1995; Weighardt et al., 1995; Michael et al., 1995). We have identified a similar sequence with the same function in HRB87F/hrp36 (Zu et al., in prep.). The *Hrb87F* gene is not essential (Zu et al., 1996) due to functional redundancy between HRB87F/hrp36 and HRB98DE/hrp38, one of which is required for viability (S. Haynes, pers. comm.). HRB87F/hrp36 is ubiquitously expressed in *Drosophila* (Haynes et al., 1991), accounts for at least half of the total protein in the strict A/B class of hnRNPs (Raychaudhuri et al., 1992), and associates promiscuously with sites of *Pol* II transcription on polytene chromosomes under normal growth conditions (Amereo et al., 1992). It undergoes a dramatic redistribution upon heat shock, however, as it is one of several hnRNP proteins that show a strong preference for heat-shock puff 93D, with the great majority of the protein accumulating at that locus (Hovemann et al., 1991).

The roles of the HRB87F/hrp36 protein domains in protein location, exon skipping, protein self-association, and RNA binding are the subject of this report. The two RBDs of HRB87F/hrp36 are members of a large RBD domain family (Birney et al., 1993). Structural studies of RBDs reveal a four-stranded β -sheet platform that contacts RNA (Nagai et al., 1990; Oubridge et al., 1994), with two highly conserved sequences within the RBD known as RNP-1 and RNP-2 juxtaposed in the central β strands of the RNA-binding platform. The conserved aromatic residues in these elements are directly involved in RNA binding via base-stacking interactions (Merrill et al., 1988; Oubridge et al., 1994) and can be mutagenized to diminish activity or RNA binding of an RBD (Cáceres & Krainer, 1993; Mayeda et al., 1994), as we have done in this study. Structural studies of the

two RBDs of A1 hnRNP reveal a configuration in which the two RNA-binding surfaces are held in close contact in an antiparallel arrangement with the RNA-binding surfaces facing the same side (Shamoo et al., 1997; Xu et al., 1997). We find that the two RBDs of the *Drosophila* A1 homolog HRB87F/hrp36 function independently and under different growth conditions to confer distinct binding specificities to the protein, with RBD-1 providing the nonspecific binding pattern to multiple sites of transcription under normal growth conditions and RBD-2 plus the GRD reproducing the very specific localization to a single locus upon heat shock.

The large GRD at the C terminus contributes many of the unusual properties of A/B hnRNP binding. It is ~120–200 residues long in proteins of this class and has a low sequence conservation, though a high conservation of glycine richness (40–50% gly), and a semi-regular spacing of flexible, aromatic, and positively charged or polar amino acids (Cobianchi et al., 1986) resulting in binding determinants along its entire length (Pontius, 1993). The GRD of HRB87F/hrp36 contains two RGG boxes (Kiledjian & Dreyfuss, 1992) near its N terminus and an M9-like nuclear shuttling signal near its C terminus. In addition to binding RNA directly (Kumar et al., 1990; Casas-Finet et al., 1993), the GRD of A1 hnRNP is responsible for protein–protein interactions (independent of RNA) between proteins with a GRD (Casas-Finet et al., 1993; Cartegni et al., 1996) and for the cooperative binding of A1 to RNA (Cobianchi et al., 1988; Nadler et al., 1991). We find that the GRD plays an important role in vivo, not only in nuclear localization of the protein, but also in specifying the quantitative deposition properties of the protein at sites of transcription and in protein sequestration at the 93D locus during heat shock.

RESULTS

Stable overexpression of wild-type and mutant HRB87F/hrp36 proteins in fruit flies

Figure 1A shows the structure of the wild-type and mutant HRB87F/hrp36 proteins used in this study. Transformed homozygous fly stocks were established in a wild-type background, each of which expressed a different mutant *Hrb87F* transgene under the control of the hsp70 promoter. All of the proteins were epitope tagged at their N terminus with the Flag epitope so that they could be distinguished immunologically from the endogenous HRB87F protein after transformation into flies. To abrogate the RNA binding properties of one or the other RBD, two highly conserved phenylalanines that occur in the RNP-1 box of the RBDs and participate directly in RNA binding (Merrill et al., 1988; Oubridge et al., 1994) were changed to aspartic acid residues, as was done previously for similar studies with human A1 hnRNP (Mayeda et al., 1994). Although

FIGURE 1. HRB87F/hrp36 protein mutants and their stable expression in flies. **A:** Schematic of the wild-type and mutant HRB87F/hrp36 proteins used in this study. The three-domain structure consists of two RNA-binding domains (RBDs) and one glycine-rich domain (GRD). Six versions of the gene were transformed into flies, all under the control of the hsp70 promoter and all tagged at their N terminus with the Flag epitope. The "X" within an RBD denotes two amino acid changes (phe → asp) within the conserved RNP-1 octamer motif. Specifically, for RBD-1, the sequence RGFGFITY was changed to RGDGDITY. For RBD-2, the sequence RGFAFIEF was changed to RGDADIEF. In addition to these RBD sequence changes, other mutant proteins had the GRD deleted or the GRD substituted by the RS domain from the *Drosophila* B52 protein (see Materials and Methods). The nomenclature used for these proteins is shown at the left. **B:** Western blot analysis of HRB87F/hrp36 proteins. Total larval proteins from the six homozygous transformed fly strains were isolated either before or 8 h after heat-shock induction of the transgene, and were probed with anti-Flag Ab after electrophoresis. The position of molecular weight standards is shown on the left. Flag-tagged proteins of the expected size (indicated by asterisks on the right side) were induced in each case and were stable up to 8 h after induction. The last two lanes (R1-R2-RS) were done at a different time but under exactly the same conditions. This strain consistently showed a certain level of basal expression of flag-tagged proteins, presumably because of genomic site of insertion of the P element. All experiments were done after hs induction when a protein of the expected size is the major product.

we have not done in vitro RNA-binding assays on these mutated proteins, there is ample evidence from previous studies that these changes to RNP-1 will severely compromise RNA-binding ability of the RBD as mediated via its RNA-binding β -sheet platform (e.g., Mayeda et al., 1994; Bouvet et al., 1997; Deardorff & Sachs, 1997). This disruption of RNA-binding ability by mutation of conserved aromatics is also predicted by biophysical (Merrill et al., 1988) and X-ray crystallographic analyses (Oubridge et al., 1994). The phenylalanines were substituted by aspartic-acid residues, which are expected to remain solvent exposed without disrupting the structure of the domain, but unable to participate in the base-stacking interactions important to interaction of RNA with the RNA-binding platform of the RBD. The wild-type protein is referred to as R1-R2-G in the figures, reflecting the RBD1-RBD2-GRD domain structure. The altered proteins include mutation of the first RBD by the two phe-to-asp changes (termed X-R2-G), mutation of the second RBD in the same fashion (termed R1-X-G), mutation of both RBDs (termed X-X-G), deletion of the GRD (termed R1-R2), and substitution of the GRD by the RS domain from the *Drosophila* SR protein B52 (termed R1-R2-RS) (Champlin et al., 1991).

The expression and stability of the mutant proteins were assayed by immunoblot analysis of total larval protein from the six transgenic stocks before and 8 h after heat-shock induction (Fig. 1B). In each case, an epitope-tagged protein of the expected size was made in response to heat shock and persisted at least 8 h after induction. This stability is similar to that of the wild-type HRB87F/hrp36 protein, which, when overexpressed in flies, peaked in amount at 8–12 h after heat-shock induction and persisted for up to 24 h (Zu et al., 1996), thus allowing effects of excess hnRNP proteins to be monitored independently of transient heat-shock effects. At their peak, the induced hnRNP proteins were present in amounts 5–10-fold above normal levels (Zu et al., 1996 and data not shown). These homozygous transgenic fly stocks have been maintained for several years. Their development is somewhat slower than wild-type flies, but induction of the transgenes does not induce any obvious developmental defects or lethality.

All three domains of the protein are required for exon-skipping activity in vivo

We have shown previously that overexpression of either HRB87F/hrp36 or HRB98DE/hrp38 in flies is able to induce an aberrant exon-skipping pattern in the endogenous pre-mRNA encoding *dopa-decarboxylase* (*Ddc*), in which both of two short internal exons are skipped (Shen et al., 1995; Zu et al., 1996) (see Fig. 2A). Although this particular alternative splicing activity is not a biologically relevant function of HRB87F/hrp36 (Zu et al., 1996), the endogenous *Ddc* splicing

FIGURE 2. Induction of exon skipping in *Ddc* pre-mRNA by excess HRB87F/hrp36. **A:** Schematic of the two naturally occurring alternative splicing patterns in *Ddc* pre-mRNA, as well as the aberrant exon-skipping pattern induced in the endogenous message by excess HRB protein (Shen et al., 1995; Zu et al., 1996). Open boxes, labeled A–D, are exons. Arrows indicate the locations and the 5'–3' orientation of the primers used for RT-PCR analysis. **B:** RT-PCR analysis of *Ddc* splicing in the presence of excess HRB87F/hrp36 mutant proteins showed that only the wild-type protein (lane 2) induced exon skipping. RNA was isolated from the hypoderm of third instar larvae from the different homozygous transgenic fly strains either without heat shock, or 8 h after a 1-h heat shock. Larval RNA was reverse transcribed and PCR amplified with *Ddc* exons A- and D-specific primers. The exon composition of the spliced products is shown on the right. **C:** RT-PCR analysis as in **B**, except that all flies had been heat shocked 12 h earlier. The flies used were either homozygous Ore-R nontransgenic flies (+/+ in lane 1), homozygous transgenic flies overexpressing wild-type HRB87F/hrp36 (lane 3), or various heterozygotes expressing the combinations listed above the lanes. Note that wild-type flies make only the ACD form as expected for larval hypoderm (lane 1), and that the homozygous wild-type overexpressors show a complete shift to the AD form (lane 3) as expected for a 12 h post-heat-shock time point (Zu et al., 1996). Two of the mutant proteins, R1-X-G (lane 5) and R1-R2-RS (lane 7) shifted the ratio to the AD form when co-overexpressed with the wild-type R1-R2-G protein. The other two mutant proteins (X-R2-G in lane 4 and X-X-G in lane 6) had no effect when co-overexpressed with R1-R2-G. In different experiments, the difference ratios for the two spliced forms were more reproducible than the total amount of PCR product. We suspect that the difference in the amount of total product between lanes is not as significant as it appears here.

assay remains convenient for determining hnRNP protein domains involved in protein–protein and/or protein–RNA interactions. The RT-PCR analysis in Figure 2B was done in flies homozygous for the various wild-type and mutant inducible transgenes. Larval hypodermal tissue was used, and thus the ACD splice form of *Ddc* mRNA is expected, in which the B exon is skipped (Fig. 2A). The analyses were done 8 h after induction of these transgenes to avoid possible heat-shock effects on splicing. When overexpressed, the wild-type hnRNP protein induced exon skipping as seen previously (ACD to AD form in Fig. 2B, lane 2) (Zu et al., 1996), whereas none of the mutant proteins had any effect (Fig. 2B, lanes 3–12). Thus all three of the domains are involved in this activity, as was shown previously for human hnRNP A1 in vitro (Mayeda et al., 1994).

In a second experiment, heterozygous flies were made that were capable of overexpressing two different transgenes at the same time. The wild-type HRB87F/hrp36 protein was overexpressed along with one of the mutant proteins to test for possible transdominant suppression effects on the exon-skipping activity (Fig. 2C). Surprisingly, none of the mutant proteins suppressed the effect, but two of the mutant proteins (R1-X-G and R1-R2-RS) enhanced the exon-skipping activity provided by one copy of the wild-type transgene (Fig. 2C, lanes 5 and 7) similar to a second copy of the wild-type transgene (lane 3). Thus, although a double-overexpressed dose of any of the mutant proteins alone had no effect on exon skipping (Fig. 2B), for two of the mutant proteins, a single overexpressed dose was able to influence exon skipping if present together with a single overexpressed dose of the wild-type protein (Fig. 2C). The properties in common to the two mutant proteins that enhanced exon-skipping activity of the wild-type protein were the presence of an intact RBD-1 plus the presence of a nuclear-localizing domain (see below). Thus, for exon-skipping activity, it is sufficient for one of the overexpressed proteins to have all three domains if the other one has RBD-1, suggesting a co-operative assembly of the proteins onto RNA.

RBD-1 and the GRD are required for the typical nuclear localization of the protein

Each of the proteins was localized within cells. Figure 3 shows whole-mount salivary glands from third instar larvae immunostained with antibody to the flag epitope 8 h after induction of the transgenic HRB proteins. Note that the wild-type protein (R1-R2-G in Fig. 3A) was strictly nuclear and that the GRD deletion (R1-R2) was excluded from the nucleus (Fig. 3E), consistent with the presence of an M9-type NLS in the HRB87F/hrp36 GRD as has been found for A1 hnRNP (Siomi & Dreyfuss, 1995; Weighardt et al., 1995). In further studies we have shown that an M9-like sequence in the GRD of HRB87F/hrp36 functions in both nuclear location and

FIGURE 3. Subcellular location of HRB87F/hrp36 wild-type and mutant proteins. Salivary glands from third instar larvae of transgenic flies in which the transgene had been induced by heat shock 8 h earlier were incubated with anti-Flag MAb and fluorescein-conjugated secondary Ab. The nomenclature for the mutant proteins is shown in Figure 1. The GRD is necessary but not sufficient for efficient nuclear localization.

nuclear shuttling (Zu et al., in prep.). The other protein with intact RBDs (R1-R2-RS) also localized primarily to the nucleus (Fig. 3F) as expected based on the nuclear-localizing properties of the RS domain (Li & Bingham, 1991). The remaining panels demonstrate that the RBDs influenced the localization properties of the GRD. The R1-X-G protein localized efficiently to the nucleus (Fig. 3B), indicating that the first RBD and the GRD can provide the signals required for wild-type protein location. The X-R2-G protein, however, had significant cytoplasmic and weaker nuclear staining (Fig. 3C). Specifically, RBD-2 acted as a positive factor for cytoplasmic location that was overridden by RBD-1 when both domains were intact. When both RBDs were crippled (X-X-G), the protein localized to both the nucleus and cytoplasm, though it was consistently more nuclear than the X-R2-G protein (Fig. 3C,D).

RBD-2 and the GRD reproduce the distribution pattern of the wild-type protein under heat-shock conditions

We analyzed the pattern of binding of wild-type and mutant HRB87F/hrp36 proteins to sites of transcription on polytene chromosomes. Previous studies have established that the protein has a wide distribution pattern during normal growth conditions (Amereo et al., 1992; Matunis et al., 1993), but a single strong site of accumulation (on heat-shock puff 93D) immediately after heat shock (Dangli et al., 1983; Hovemann et al., 1991; Zu et al., 1996). These two chromosomal staining patterns are shown for our epitope-tagged wild-type HRB87F/hrp36 protein in Figures 4A (heat shock) and 5A (normal growth). We asked which domains of the protein are responsible for these two very different native RNA binding patterns.

In considering the heat-shock pattern (Fig. 4), the inset in Figure 4A shows the distribution of wild-type HRB87F/hrp36 in an intact nucleus immediately after heat shock, showing that a single chromosomal locus accommodates the great majority of the protein; the chromosomal squash confirms that this is the 93D heat-shock puff (Fig. 4A, long arrow). For the mutant proteins, it can be seen from the chromosome squashes and from the whole-cell insets that all of the proteins with a GRD localized to a significant extent to 93D (indicated by the long arrows in Fig. 4A–F), even in X-X-G, in which both RBDs were crippled (Fig. 4D). In the absence of a GRD, the R1-R2 protein showed very little nuclear and chromosomal staining (Fig. 4E). The R1-R2-RS protein was nuclear and chromosomal and was distributed to multiple sites with only a slight preference for heat-shock puffs, including 93D (long arrow in Fig. 4F). Thus, the GRD was the primary determinant for preferential binding to the 93D locus. Interestingly, however, the RBDs modified the 93D specificity of the GRD. The X-R2-G protein bound as exclusively

FIGURE 4. Location of wild-type and mutant HRB87F/hrp36 proteins immediately after heat shock: X-R2-G is able to reproduce the R1-R2-G pattern. For each protein, a polytene chromosome squash is shown as well as an intact salivary gland cell (insets). Proteins were immunolocalized using anti-Flag MAb and fluorescein-conjugated secondary Ab. Long thin arrows in each panel indicate the heat-shock puff at 93D on the right arm of the third chromosome. Short, fat arrows in each panel indicate nucleoli. The large foci in the whole cell insets in **B** and **D** correspond to nucleoli and the single small foci in the whole cell insets shown in **A** and **C** correspond to a chromosomal locus (presumably 93D), as judged from comparison to the phase micrographs of these cells (not shown).

to 93D as did the wild-type protein (Fig. 4A,C), but both R1-X-G and X-X-G (Fig. 4B,D) showed significant and reproducible binding to nucleoli (marked with short arrows in all of the panels of Fig. 4 and visible as the large foci in intact nuclei in insets in Fig. 4B,D). This nucleolar localization is never seen for wild-type hnRNP proteins, even when overexpressed 20-fold, and presumably represents a low-affinity and non-physiological binding site during heat shock for GRD-containing proteins that lack a functional RBD-2. The R1-X-G protein also localized weakly to multiple chromosomal sites and to the cytoplasm immediately after heat shock (Fig. 4B, inset). Thus, the distribution patterns of the proteins reflect a combination of the properties of the remaining functional domains. No single domain was able to reproduce the wild-type pattern, but a combination of RBD-2 and the GRD in X-R2-G was able to localize exclusively to the 93D locus during heat-shock conditions.

RBD-1 and the GRD reproduce the distribution pattern of the wild-type protein under normal growth conditions

Localization of the same proteins on chromosomes 5–6 h after recovery from heat-shock induction (i.e.,

under normal growth conditions; Fig. 5A–F), also revealed distinct yet interacting roles of the three protein domains in RNA binding. As expected, the wild-type protein bound all over the chromosomes (Fig. 5A), and the GRD deletion (R1-R2), which does not enter the nucleus, showed no significant chromosomal binding (Fig. 5E). The RS domain substitution (R1-R2-RS; Fig. 5F) bound to multiple chromosomal sites like the wild-type protein, but did not show the signal amplification at highly active sites that is characteristic of hnRNP proteins, consistent with the absence of highly cooperative binding or protein aggregation mediated via the GRD. The RBD mutations showed interesting patterns. First, X-R2-G (Fig. 5C), which bound so efficiently to 93D immediately after heat shock (Fig. 4C), remained associated with 93D even at this late time point after heat shock (5 h) at which time heat-shock effects are dissipated. Two chromosome sets are shown for this mutant. On both, the arrows indicate the major signal at the 93D puff, which has not regressed as expected, perhaps indicating stabilization of the RNA at the site. (See arrow in wild-type pattern, Fig. 5A, showing a regressed 93D puff.) Thus, a combination of RBD-2 and the GRD, in the absence of the dominant effects of RBD-1, resulted in a protein that

FIGURE 5. Location of wild-type and mutant HRB87F/hrp36 proteins under normal growth conditions: R1-X-G is able to reproduce the R1-R2-G pattern. Shown are polytene chromosome squashes and intact nuclei (insets). Proteins were immunolocalized 5–6 h after heat induction of HRB protein synthesis using anti-Flag MAb and fluorescein-conjugated secondary Ab. Arrows in each panel indicate position 93D on the right arm of the third chromosome.

preferred the 93D locus over the other RNA transcripts being synthesized throughout the chromosomes.

The other single RBD mutation, R1-X-G, which showed a less normal binding pattern than X-R2-G immediately after heat shock (Fig. 4B), showed a more normal binding pattern than X-R2-G after recovery from heat shock in that it was distributed over many chromosomal sites (Fig. 5B). The pattern was somewhat different in a quantitative sense, but was similar qualitatively, as will also be shown in Figure 6. When both RBDs were crippled (X-X-G; Fig. 5D), the protein bound in large amounts to a few chromosomal sites that correspond to major sites of deposition of the wild-type protein, as will be shown in Figure 6. One of the major sites bound by X-X-G was 93D (arrow in Fig. 5D), but instead of the usual compact puff configuration, the protein occurred in a globular, stretchable form that frequently was smeared out from the site by the squashing procedure. On different chromosome sets the protein smear had different sizes and shapes but was very reproducibly present at 93D and a few other locations, presumably corresponding to an aggregation via GRD–GRD interactions of itself and wild-type HRB proteins, which are present in these flies. The insets in Figure 5 show the distribution of the proteins in intact cells. As shown previously (Zu et al., 1996), the overexpressed

R1-R2-G protein accumulated in an apparently nonsaturable manner on chromosomal sites (as did R1-X-G) and was not found in the nucleoplasm or cytoplasm (insets, Fig. 5A,B), whereas substitution of the GRD with the RS domain resulted in a more diffuse nuclear distribution (inset, Fig. 5F). The X-X-G protein occurred in large nuclear masses that had the same smeary appearance as seen in the squash preparation (inset, Fig. 5D). In summary, the combined results of Figures 4 and 5 suggest that RBD-1 and the GRD are mainly responsible for the chromosomal binding properties of the wild-type protein under normal growth conditions while RBD-2 and the GRD play that role during the heat-shock response. In addition, these results provide a strong visual demonstration of the self-aggregation properties of the GRD.

The specific chromosomal binding patterns of the wild-type and mutant HRB proteins were examined by mapping the bound sites on a specific (i.e., the X) chromosome (Fig. 6). For comparison to the overexpressed and mutant proteins, the X chromosome binding pattern of the endogenous HRB87F/hrp36 protein in a wild-type fly is shown (Fig. 6, left). This high resolution comparison confirmed the impression from Figure 5 that if RBD-1 was intact, (as in HRB87F, R1-R2-G, R1-X-G, and R1-R2-RS), the protein bound to a large num-

FIGURE 6. High-resolution comparison of the location of wild-type and mutant HRB87F/hrp36 proteins on the X polytene chromosome. Conditions as in Figure 5. In addition to overexpressed wild-type and mutant proteins (indicated by the R1-R2-G series nomenclature), also shown is the endogenous wild-type protein (labeled HRB87F) detected by anti-HRB Ab (Raychaudhuri et al., 1992). Mapping of specific chromosomal positions was done using the DAPI-stained version of each chromosome. One of these is shown on the far right. Different chromosomes were stretched in different ways during the squashing procedure—thus there is not necessarily a linearity between length and chromosome position. This is particularly evident in comparing the HRB87F and R1-R2-G chromosomes in the region between 9E and 15C. The unusual break in the X-X-G chromosome represents a chromosome that was originally L-shaped but was cut and pasted to have a more linear form. Late third instar larvae were used for this experiment, but they were not more precisely staged. Although care was taken to choose representative chromosomes, some of the differences in intensity of stain at specific sites may be because of developmental variations.

ber of sites. The overexpressed R1-R2-G pattern is qualitatively similar to the wild-type HRB87F pattern, though quantitatively amplified at some of the larger puffed sites. On the other hand, the R1-R2-RS protein, though qualitatively similar to the wild-type protein, showed much less signal amplification at puff sites. Although the RS domain has protein–protein interaction properties (Kohtz et al., 1994), its ability to confer self-association properties appears to be less than that of the GRD, at least in this context. The R1-X-G pattern showed multiple chromosomal sites with strong signals from the major sites. The signal from the weaker sites was dampened relative to the signal from the wild-type protein however. For a protein with self-association properties, this may represent a protein that binds RNA with lower affinity though with similar nonspecificity to the wild-type protein. The proteins without an intact RBD-1 showed fewer significant binding sites. For X-X-G, the major sites corresponded to the major sites of binding of the wild-type R1-R2-G protein, consistent with protein–protein interactions. (Since all of these transgenic flies were made in a wild-type background, the endogenous HRB87F/hrp36 and HRB98DE/hrp38 proteins were present in all cases.) The X-R2-G protein had the lowest occurrence of all the proteins on the X

chromosome (Fig. 6); its major site of deposition, 93D, is on the third chromosome (Fig. 5C). Thus, RBD-2 confers distinct RNA binding properties which are manifested upon heat shock but are masked under non-heat-shock conditions if RBD-1 is present. The first RBD is largely able to reproduce the promiscuous RNA binding properties of the wild-type protein in a qualitative sense; the GRD has an important contribution to the wild-type quantitative pattern.

DISCUSSION

Role of the RBDs in HRB87F/hrp36 protein function

The polytene chromosome-binding assay used in this study is a powerful *in vivo* RNA-binding assay. The chromosomes provide a well-characterized global array of all nascent RNA, which is the predominant form of nuclear pre-mRNA in these cells because the RNA exits the nucleus very rapidly once released from its site of synthesis (Zachar et al., 1993). Although the wild-type HRB87F/hrp36 protein is present in these flies and is bound to the same RNAs and thus may influence the RNA binding properties of the transgenic pro-

teins, we feel that this is a valid representation of the *in vivo* situation for the following reasons. First, it represents a typical complex environment in which protein–RNA interactions occur in a nucleus. Even if the experiments had been done in a HRB87F null situation, there would still have been abundant levels of its functionally redundant sister protein, HRB98DE/hrp38. Second, the distinct distribution patterns that we see with the different mutant proteins tells us that their unique binding properties are not being overridden by protein–protein interactions, though the latter do clearly play a role, as they undoubtedly do for any GRD-containing protein. Third, the same chromosomal distribution patterns are seen over a wide range of protein overexpression (which can be modulated by varying the length and temperature of the heat-shock induction (not shown)), and for the overexpressed wild-type protein (R1-R2-G), the distribution is the same as the endogenous wild-type protein (Fig. 6). Thus, saturation of binding sites does not appear to influence the localization of the proteins.

Using this chromosome binding assay, we found clear evidence for distinct roles in RNA binding for the two RBDs of HRB87F/hrp36. These two RBDs in A/B-type hnRNP proteins have evolved independently, suggesting different functions for the two domains (Haynes et al., 1991). Our results show that RBD-1 assumes the dominant role in conferring RNA-binding properties to the protein under normal growth conditions, conferring a polytene chromosome-binding pattern that is qualitatively similar to the wild-type protein pattern (Figs. 5, 6). Thus, RBD-1 appears mainly responsible for promiscuous binding to the general population of pre-mRNA, which is characteristic of A/B hnRNPs. In addition, RBD-1 provides a function in exon skipping that is not provided by RBD-2 (Fig. 2), and in concert with the GRD, RBD-1 confers normal nuclear location to the protein (Fig. 3). These latter two properties may be an indirect effect of the RNA-binding properties of RBD-1, but in any case, they are not provided by RBD-2. The finding that X-R2-G and X-X-G are not localized to the nucleus as efficiently as the wild-type protein in spite of the presence of an intact GRD and M9-like element indicates that RNA-binding properties of the protein influence cellular location, as suggested previously for A1 hnRNP (Weighardt et al., 1995).

On the other hand, RBD-2 appears to assume the dominant role under heat-shock conditions, a time at which HRB87F/hrp36 is known to show a very strong preference for binding to a single polytene chromosome locus, that is, the heat-shock puff at 93D (Hovemann et al., 1991). The X-R2-G protein shows normal (exclusive) binding to 93D upon heat shock, whereas none of the other mutant proteins do. However, non-heat-shock properties of X-R2-G are not characteristic of wild-type HRB87F/hrp36. Much of the X-R2-G protein accumulates in the cytoplasm in spite of an intact M9-like sequence within its GRD and the protein that

remains in the nucleus shows a preference for the 93D locus, even under non-heat-shock conditions when transcription is proceeding normally at other chromosomal sites (as indicated by expected developmental puffing patterns and by the health of these flies).

The 93D heat-shock locus is somewhat of an enigma—it is required for viability (Mohler & Pardue, 1982; Lakhota & Sharma, 1996), yet transcripts of 93D are not protein encoding. This locus (and similar loci in other *Drosophila* species, collectively known as *hsr- ω*) are similar in that they all make a transcript of ≥ 7 kb that is restricted to the nucleus (Pardue et al., 1992). Although there has been rapid sequence divergence at the loci, in each species examined the *hsr- ω* RNA contains 5–16 kb of tandem direct repeats with the single conserved feature being the nonamer AUAGGUAGG repeated at ~ 100 bp intervals (Pardue et al., 1992; Hogan et al., 1995). This sequence bears some similarity to splice junctions (and to spliced junctions). Interestingly, it also resembles a central octamer sequence in the SELEX winner sequence for the RBD-2:GRD fragment of A1 hnRNP, which is UAGGUCAG (Burd & Dreyfuss, 1994). Furthermore, a recent report on hnRNP A1 that found a higher intrinsic RNA-binding activity for RBD-2 as compared to RBD-1 was done using a SELEX-selected RNA that includes a sequence very similar to the 93D repeat: UAGGUUAGG (Mayeda et al., 1998). Thus, the specificity of wild-type HRB87F/hrp36 binding to 93D upon heat shock that is conferred by a combination of RBD-2 and the GRD may be because of the preference of RBD-2 for the nonamer repeat sequence. The R1-X-G protein has lost this specificity. Although a significant amount still binds 93D (as do all GRD-containing HRB87F/hrp36 proteins), there is high background binding to other chromosomal sites and especially to nucleoli. Other pre-mRNA binding proteins with GRDs are known to bind almost exclusively to 93D upon heat shock, such as HRB98DE/hrp38 (unpubl.), HRB57A (Buchenau et al., 1997), and Sxl (Samuels et al., 1994), and it is reasonable to hypothesize that the GRDs of these other proteins are involved in this unusual distribution pattern under stress conditions. Given the fact that the 93D locus is required for viability and that there seems to be a limit on the minimum number of sequence repeats at the locus (Hogan et al., 1995), it is tempting to speculate that protein sequestration at the locus may be an important survival strategy and a previously unrecognized function of some GRDs. It is not clear why these proteins are sequestered at 93D at heat shock, but it is perhaps best to have them out of the way during the specialized transcription and export of heat-shock mRNAs (Saavedra et al., 1996).

Our results confirm and extend earlier studies with A1 hnRNP. Although both A1 RBDs have RNA-binding properties (Shamoo et al., 1994; Mayeda et al., 1994; Burd & Dreyfuss, 1994) and there is an increase in

occluded site size when both RBDs are bound (14 ± 2 nt) versus one RBD (6 ± 1 nt), the tenfold increase in affinity seen with $2 \times$ RBD versus $1 \times$ RBD is not the 1,000-fold expected from the product of the affinities of the isolated RBDs (Shamoo et al., 1994, 1995). Both UV crosslinking studies (Merrill et al., 1988; Merrill & Williams, 1990) and physical studies on the A1 hnRNP protein (Casas-Finet et al., 1991, 1993) suggest that only one RBD at a time is bound to nucleic acid via specific interaction with an RNA-binding platform. Furthermore, the conserved phenylalanine residues in RBD-1 of A1 hnRNP account for 75% of total cross-linking, whereas the phenylalanines of RBD-2 account for 25%, leading Merrill et al. to conclude that RBD-1 is a better nucleic acid-binding polypeptide (Merrill et al., 1988). Our results are less easy to reconcile with a recent study on hnRNP A1 that concludes that the two RBDs are functionally nonequivalent, but that RBD-2 plays a more important role than RBD-1 in RNA binding and in alternative splicing (Mayeda et al., 1998). It is possible that the different results are because of different RNA substrates used in the two studies or to organism-specific differences in hnRNP protein function.

Recent structural analyses of the two RBDs of the A1 hnRNP protein (Shamoo et al., 1997; Xu et al., 1997) represent the first such studies on two linked RBDs and reveal that the two domains are held in close contact in an antiparallel arrangement with the RNA binding surfaces facing the same side, forming what appears to be an extended (doubled) RNA-binding surface. The amino acid pairs responsible for the close apposition and orientation of the two RBDs in A1 hnRNP are conserved in HRB87F/hrp36 and in all 12 members of the strict A/B hnRNP family (Mayeda et al., 1998). The structural data accommodate independent binding by the RBDs. In fact, given the antiparallel arrangement of the two RBDs, a single RNA strand would have to loop around to be bound by both RBD platforms (Shamoo et al., 1997; Xu et al., 1997). Our results indicate that the *specificity* of binding of HRB87F/hrp36 to polytene chromosome sites is conferred primarily by one or the other RBD at any given time.

Previous studies of proteins with more than one RBD have revealed situations in which one RBD has been shown to play the dominant role in RNA binding and function (Lutz-Freyermuth et al., 1990; Scherly et al., 1990; Flickinger & Salz, 1994; Dember et al., 1996), or situations in which the RBDs have been found to work in synergy to provide the RNA binding properties of the native protein (Zamore et al., 1992; Cáceres & Krainer, 1993; Zuo & Manley, 1993; Kanaar et al., 1995; Bouvet et al., 1997). A combination of these two modes of RNA binding seems to be operating in the poly(A) binding protein, which has four RBDs (Deardorff & Sachs, 1997). HRB87F/hrp36 is unusual in having apparently distinct and separable RNA-binding roles for its two RBDs, although *in vitro* binding studies suggest that this may

also be the case for the mHuC protein (Abe et al., 1996). Distinct roles have also been found for the two RBDs of yeast U1A snRNP protein using a genetic approach, though it is not clear that these both involve specific RNA binding (Tang & Rosbash, 1996). The significance of our results with HRB87F/hrp36 remains to be tested genetically, but such studies are now feasible.

Role of the glycine-rich domain in A/B hnRNP protein function

The GRD (or the dM9 sequence within the GRD) plays an important role in protein import into the nucleus, characteristic RNA binding under normal conditions, and exclusive binding at 93D under heat-shock conditions. The cytological assays reveal that the GRD confers apparently nonsaturable self-aggregation properties, whether these sites are all over the chromosomes (as in the wild-type protein), or at limited sites in large nuclear masses (as when both RBDs are mutated). In the latter situation, we believe that the localization patterns represent GRD interactions rather than specific RNA binding, because the major sites bound by X-X-G correspond to the major sites of deposition of the wild-type protein (which is present in these flies) under both normal (Figs. 5, 6) and heat-shock (Fig. 4) conditions. We also suspect that the nucleolar binding seen upon heat shock in those proteins with a GRD but without a functional RBD-2 (R1-X-G and X-X-G in Fig. 4) is because of hnRNP protein binding to one of several GRD-containing nucleolar proteins, such as fibrillarin or nucleolin.

In addition to direct binding of the GRD to RNA (Kumar et al., 1990; Casas-Finet et al., 1993), previous studies have also documented the self-association properties of the GRD of A1 hnRNP (Cobianchi et al., 1988) and of GRDs in general (Cartegni et al., 1996), whether or not RNA is present (Casas-Finet et al., 1993; Cartegni et al., 1996). It is probably this self-association characteristic that underlies additional properties of the domain, such as its ability to confer cooperativity of binding (Cobianchi et al., 1988; Nadler et al., 1991), to contribute to strand-annealing properties (Portman & Dreyfuss, 1994), and to contribute half of the free energy of binding of A1 hnRNP to RNA (Shamoo et al., 1994). Various models have been proposed to explain the properties of these quasirepetitive domains (Steinert et al., 1991; Pontius, 1993; Cartegni et al., 1996) which have a semiregular spacing of aromatics (on average every 6 amino acids in A1 and every 9 amino acids in HRB87F/hrp36) and basic or polar amino acids embedded in glycines. Although the models vary in particulars, they all predict an unusually flexible domain that can interact nonspecifically with RNA and with other GRDs via multiple, weak interactions that are easily disrupted and rearranged. Pontius (1993) discusses how the unstructured repetitive nature of

these domains is essential to their function in promoting association because multiple, transient interactions (both hydrophobic and electrostatic) can be formed regardless of the encounter orientation. There is a positive correlation between GRD length and affinity to RNA (Buvoli et al., 1990; Mayeda et al., 1994) as well as affinity to other GRDs (Cartegni et al., 1996). The critical importance of the aromatic residues in the interaction (Cartegni et al., 1996) is in agreement with the suggestion that stacking of aromatics (with each other or with RNA bases) is involved. Most of these previous studies have been done on the GRD of the A1 protein. The GRD of HRB87F/hrp36 is ~40% longer than that of A1. It has two RGG repeats near its N terminus and its middle region consists essentially of aromatics at intervals of 4–12 amino acids embedded in polar amides (20%) and glycine (60%).

HRB87F/hrp36 and alternative splicing

In the *Ddc* mRNA splicing assay (Fig. 2), no dominant negative effects on exon-skipping induction were seen when mutant HRB proteins were expressed. Instead, two of the mutant proteins enhanced the exon-skipping activity. These two proteins (R1-X-G and R1-R2-RS) are the only ones that bound reasonably well to multiple sites on polytene chromosomes as does the wild-type protein (Fig. 5). The fact that enhancement, rather than suppression, was seen (even in the presence of an RS domain) is consistent with a nonspecific heteromultimerization that sterically blocks splice site selection. Such a multimerization is also consistent with the self-association properties of the protein as revealed in the cytological assays, and with the ability of GRDs to associate with RS domains as well as with other GRD domains (Cartegni et al., 1996). It is not clear whether the apparent involvement of RBD-1 in the enhancement phenomenon is because of a direct requirement for RBD-1 to bind this region of *Ddc* RNA or rather is an indirect effect of protein localization at chromosomal sites of transcription. An alternative explanation for the results, that is, that excess RBD-1 is titrating a factor involved in splice-site selection for exon C of *Ddc* pre-mRNA, seems unlikely because excess R1-X-G alone did not induce exon skipping.

The R1-R2-RS mutant was originally made to test aspects of the model that proposes antagonistic effects between A/B hnRNPs and SR proteins in alternative splicing (Mayeda & Krainer, 1992); for example, would the RS domain function as predicted as a splicing activator if targeted to the RNA via other RBDs? Although it was somewhat surprising that R1-R2-RS acted in concert with the hnRNP protein to enhance (rather than suppress) exon skipping, there have been previous reports of RS domains functioning to inhibit splicing (Wang & Manley, 1995). Even more surprising, however, was the finding that overexpression of R1-R2-RS resulted

in no observable phenotypes or ill effects on the flies. When this same B52 RS domain was overexpressed in flies as part of B52 (which is homologous to human ASF/SF2), there were serious adverse effects on the development of the organism (Kraus & Lis, 1994). The simplest interpretation of a model that proposes RS domains as effector domains in alternative splicing (Graveley & Maniatis, 1998, and references therein) would have predicted at least some observable phenotypes given the observation that the R1-R2-RS protein appears to be deposited at many if not all sites of transcription on polytene chromosomes (Figs. 5, 6). One could argue that our resolution in this binding assay is low and that R1-R2-RS may not be in direct contact with the appropriate sites on the pre-mRNA, but it does seem to be at the site of action on *Ddc* pre-mRNA as shown by its ability to enhance exon skipping (Fig. 2B, lane 7). Further experiments will be required to resolve this issue, but one can conclude at this point that the ill effects of B52 overexpression require more than putting the RS domain onto RNA at inappropriate locations. The argument formulated below concerning the relatively weak binding of these RBDs to RNA is probably also relevant.

We have previously reported that either complete deletion or organism-wide overexpression of wild-type HRB87F/hrp36 results in few obvious negative effects on fruit flies, indicating that pre-mRNA processing proceeds normally with a range in A/B hnRNP protein level from about half the normal level to a 10–15-fold excess (Zu et al., 1996). As discussed, this is difficult to reconcile with the proposal that the concentration of these proteins relative to SR proteins influences alternative splicing patterns in many pre-mRNAs. We now report that overexpression of mutant versions of the HRB87F/hrp36 protein also has no obvious dominant negative effects, lending further support to our argument that these proteins do not as part of their normal function exist in specific multiprotein complexes that regulate alternative splicing of many pre-mRNAs. Overexpression or removal of other proteins involved in RNA processing has been shown to have significant deleterious effects, as have inappropriate expression and dominant negative mutant expression (e.g., Ring & Lis, 1994; Kraus & Lis, 1994; Romac & Keene, 1995; Wang & Manley, 1995). We suspect that the unusual behavior of these hnRNP proteins (with respect to not being deleterious when overexpressed in wild-type or mutant forms) is because of a combination of the low binding affinities and the low sequence specificity of the RBDs (Abdul-Manan & Williams, 1996) and the apparent nonsaturable self-association properties of the GRD, which provides a sink for excess protein. The slowed development seen when these proteins are overexpressed may be because of a slowed rate of mRNA export from the nucleus, which we are now testing.

MATERIALS AND METHODS

Plasmid constructions for P-element transformation of *Drosophila*

The addition of a FLAG epitope (Kodak) to the N terminus of the *Hrb87F* cDNA has been described (Zu et al., 1996). pBS-HRB87F was generated by inserting the 2.2 kb *EcoR*I fragment containing FLAG-Hrb87F cDNA into pBluescript II KS(+) (Stratagene) which had been modified to remove the *Hinc* II site in the polylinker sequence. pBS-HRB87F was used to make the following mutant constructs. The sequence of each construct was confirmed by DNA sequencing, and in those cases where new restriction sites were introduced, by restriction digestion.

pBS-X-R2-G, which contains Asp substitutions of both Phe 67 and Phe 69 in RNP1 of RBD-1, was constructed by the mega-primer PCR technique (Sarkar & Sommer, 1990). Mega-primer #1 was made using primer #1 (5' GCCTGAAG GCTCACTTCGAG) and primer #2 (5' GACTGGGAGTACGT GATATCACCGTCGCCGCGAGAGCGCTTCG), which converted two TTC codons for Phe to codons GAT and GAC for Asp and also created an *EcoRV* site. A second PCR reaction using mega-primer #1 and primer #4 (5' TTGTTCTTGATG GAGTGGG) resulted in a fragment that was used to substitute the 354-bp *Hinc* II fragment of the wild type cDNA in pBS-HRB87F.

pBS-R1-X-G, which contains Asp substitutions of both Phe 158 and Phe 160 in RNP1 of RBD-2, was generated in a similar approach. Mega-primer #2 was made using primer #3 (5' CCGCAAGAAGCGCGCGCGACGCCGACATTGAG TTGCATGACTAC) and primer #4 and the second PCR reaction used mega-primer #2 and primer #1.

pBS-X-X-G, which contains Asp substitutions of all 4 conserved Phe residues in both RBD-1 and RBD-2, was generated by substituting the 2.3-kb *Bsa*I fragment of pBS-X-R2-G, which spans RBD-2, with the same *Bsa*I fragment from pBS-R1-X-G.

pBS-R1-R2, the carboxyl terminal glycine-rich domain deletion mutant (amino acids 1–195), was constructed by deleting the terminal 1.5-kb *Bsa*I/*Spe*I fragment of *Hrb87F* cDNA. The junction sequence contained an in-frame stop codon and a new *EcoR* V site.

pBS-R1-R2-RS, which encodes a chimeric protein of HRB87F and B52, was made by joining the two RBDs of HRB87F (amino acids 1–196) to the carboxyl terminal RS domain (amino acids 186–376) of *Drosophila* B52 (dSRp55; Champlin et al., 1991). The RS domain of B52 was obtained by isolating the 0.9-kb *Bam*H I/*Spe*I fragment from pBS-B52 (kindly provided by J. Lis, Cornell). After filling in the 5' overhang of the *Bam*H I site, the fragment was used to replace the 1.5-kb *Bsa*I/*Spe*I from pBS-HRB87F, which contains the sequence encoding the GRD.

pBHS/K(–) (Park et al., 1994; kindly provided by P. Adler, University of Virginia) was used to generate intermediate plasmids for each mutant cDNA. The 2.2-kb *EcoR*I fragments of pBS-HRB87F, pBS-R1-X-G, pBS-X-R2-G, and pBS-X-X-G containing the wild-type or mutant cDNAs were inserted into the *EcoR*I site between the *Drosophila* *Hsp70* promoter and the SV40 polyadenylation signal of pBHS/K(–). The 0.7-kb *Xho*I/*Xba*I fragment of pBS-R1-R2 was inserted into the *EcoR*I site of pBHS/K(–) after filling the 5' overhangs of

*Xho*I, *Xba*I, and *EcoR*I sites. The 1.6-kb *EcoR*I fragment of pBS-R1-R2-RS was inserted into the *EcoR*I site of pBHS/K(–). The *Not*I/*Kpn*I fragments from the pBHS/K(–) plasmid series containing epitope-tagged Hrb87F wild-type and mutant cDNAs under the control of the *Hsp70* promoter and ending with the SV40 polyA signal were subcloned into the P-element transformation vector pW8 (Klemenz et al., 1987).

P-element mediated transformation was done by standard techniques (Roberts, 1986; Zu et al., 1996). Genetic crosses indicated that each stable homozygous line had an insert on a single chromosome. For induction of the transgenes, larvae were heat-shocked at 37° for 30–60 min.

Analysis of splicing products

RNA preparation and RT-PCR were performed as described previously (Shen et al., 1995).

Immunodetection methods

Western blots were done as described previously (Zu et al., 1996), using 50 µg of total larval protein in each lane.

Polytene chromosome squashes were done using the method of Zink & Paro (1989). Slides were incubated in a 1:600 dilution of anti-FLAG IgG in TBST (20 mM Tris-HCl, pH 7.7, 1.7% NaCl, 0.1% Tween-20) plus 5% dry milk overnight at 4°C. Slides were washed 3 times for 5–10 min in TBST, and then incubated in a 1:300 dilution of fluorescein-conjugated anti-mouse IgG in TBST, 5% dry milk, followed by 3 washes for 5–10 min in TBST. Slides were then stained for 1 min in freshly diluted 4,6 diamidino-2-phenylindole (DAPI) at 0.12 µg/mL in 180 mM Tris-HCl, pH 7.5, washed as above, and mounted in Vectashield.

For immunostaining of whole salivary glands, dissected glands were fixed in 4% formaldehyde, 50% acetic acid, and half-strength PBS for about 30 min. For immunostaining of dissected fat body, the tissue was fixed in 4% formaldehyde in PBS for about 30 min. The fixed tissues were washed 4 times for 5 min in PBS and 2 times for 5 min in PBT (PBS, 0.1% BSA, 0.3% Triton-X 100). Primary antibody incubation was with 1:300 dilution of anti-FLAG IgG in PBT with gentle mixing overnight at room temperature. Tissues were washed 5 times for 5 min in PBT, incubated with fluorescein-conjugated anti-mouse IgG (1:100) in PBT for 6 h at room temperature, washed with 4 changes in PBT for 6–24 h, and mounted in Vectashield for viewing.

For immunostaining of partially squashed salivary gland cells, glands were dissected in 45% acetic acid and 3.7% formaldehyde and were squashed between a siliconized coverslip and a slide by gently tapping with the back end of a pair of dissecting forceps until the desired degree of dispersal was obtained. These preparations were then treated as described in Zu et al. (1996) except that permeabilization was for 15 min and DAPI staining was for 1 min.

ACKNOWLEDGMENTS

We thank Carlos Madrid for early observations on localization of mutant proteins during heat shock. We thank Sarah French for generous help with molecular biology techniques

and for critical reading of the manuscript. We also thank Yvonne Osheim, Sue Haynes, Paul Adler, and anonymous reviewers for critical reading of the manuscript. This work was supported by NIH GM-39271 to A.L.B.

Received June 15, 1998; returned for revision July 21, 1998; revised manuscript received September 1, 1998

REFERENCES

- Abdul-Manan N, Williams KR. 1996. HnRNP A1 binds promiscuously to oligoribonucleotides: Utilization of random and homo-oligonucleotides to discriminate sequence from base-specific binding. *Nucleic Acids Res* 24:4063–4070.
- Abe R, Sakashita E, Yamamoto K, Sakamoto H. 1996. Two different RNA binding activities for the AU-rich element and the poly(A) sequence of the mouse neuronal protein mHuC. *Nucleic Acids Res* 24:4895–4901.
- Amero SA, Raychaudhuri G, Cass C, van Venrooij W, Habets WJ, Krainer AR, Beyer AL. 1992. Independent deposition of heterogeneous nuclear ribonucleoproteins and small nuclear ribonucleoprotein particles at sites of transcription. *Proc Natl Acad Sci USA* 89:8409–8413.
- Birney E, Kumar S, Krainer AR. 1993. Analysis of the RNA-recognition motif and RS and RGG domains: Conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res* 21:5803–5816.
- Bouvet P, Jain C, Belasco JG, Amalric F, Erard M. 1997. RNA recognition by the joint action of two nucleolin RNA-binding domains: Genetic analysis and structural modeling. *EMBO J* 16:5235–5246.
- Buchenau P, Saumweber H, Arndt-Jovin DJ. 1997. The dynamic nuclear redistribution of an hnRNP K-homologous protein during *Drosophila* embryo development and heat shock. Flexibility of transcription sites in vivo. *J Cell Biol* 137:291–303.
- Burd C, Dreyfuss G. 1994. RNA binding specificity of hnRNP A1: Significance of hnRNP A1 high-affinity binding sites in pre-mRNA. *EMBO J* 13:1197–1204.
- Buvoli M, Cobiainchi F, Bestagno MG, Mangiarotti A, Bassi MT, Biamonti G, Riva S. 1990. Alternative splicing in the human gene for the core protein A1 generates another hnRNP protein. *EMBO J* 9:1229–1235.
- Cáceres JF, Krainer AR. 1993. Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. *EMBO J* 12:4715–4726.
- Cáceres JF, Stamm S, Helfman DM, Krainer AR. 1994. Regulation of alternative splicing in vivo by overexpression of antagonistic pre-mRNA splicing factors. *Science* 265:1706–1709.
- Cartegni L, Maconi M, Morandi E, Cobiainchi F, Riva S, Biamonti G. 1996. HnRNP A1 selectively interacts through its gly-rich domain with different RNA-binding proteins. *J Mol Biol* 259:337–348.
- Casas-Finet JR, Karpel R, Maki AH, Kumar A, Wilson SH. 1991. Physical studies of tyrosine and tryptophan residues in mammalian A1 heterogeneous nuclear ribonucleoprotein: Support for a segmented structure. *J Mol Biol* 221:693–709.
- Casas-Finet JR, Smith JD, Kumar A, Kim JG, Wilson SH, Karpel RL. 1993. Mammalian hnRNP A1 and its constituent domains. Nucleic acid interaction, structural stability and self-association. *J Mol Biol* 229:873–889.
- Champlin DT, Frasch M, Saumweber H, Lis J. 1991. Characterization of a *Drosophila* protein associated with boundaries of transcriptionally active chromatin. *Genes & Dev* 5:1611–1621.
- Cobiainchi F, SenGupta DN, Zmudzka BZ, Wilson SH. 1986. Structure of rodent helix-destabilizing protein revealed by cDNA cloning. *J Biol Chem* 261:3536–3543.
- Cobiainchi F, Karpel RL, Williams KR, Notario V, Wilson SH. 1988. Mammalian hnRNP complex protein A1. Large scale overproduction in *E. coli* and cooperative binding to single-stranded nucleic acids. *J Biol Chem* 263:1063–1071.
- Dangli A, Grond C, Kloetzel P, Bautz EKF. 1983. Heat shock puff 93D from *Drosophila melanogaster*: Accumulation of a RNP-specific antigen associated with giant particles of possible storage function. *EMBO J* 2:1747–1751.
- Deardorff JA, Sachs AB. 1997. Differential effects of aromatic and charged residue substitutions in the RBDs of the yeast poly(A)-binding protein. *J Mol Biol* 269:67–81.
- Dember LM, Kim ND, Liu KQ, Anderson P. 1996. Individual RNA recognition motifs of TIA-1 and TIAR have different RNA binding specificities. *J Biol Chem* 271:2783–2788.
- Dreyfuss G, Matunis MJ, Burd CJ, Piñol-Roma S. 1993. HnRNP proteins and the biogenesis of pre-mRNA. *Ann Rev Biochem* 62:289–321.
- Flickinger TW, Salz HK. 1994. The *Drosophila* sex determination gene *snf* encodes a nuclear protein with sequence and functional similarity to the mammalian U1A snRNP protein. *Genes & Dev* 8:914–925.
- Graveley BR, Maniatis T. 1998. Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. *Mol Cell* 1:765–771.
- Haynes SR, Raychaudhuri G, Beyer AL. 1990. The *Drosophila* Hrb98DE locus encodes four protein isoforms homologous to the A1 protein of mammalian heterogeneous nuclear ribonucleoprotein complexes. *Mol Cell Biol* 10:316–323.
- Haynes SR, Johnson D, Raychaudhuri G, Beyer AL. 1991. The *Drosophila* Hrb87F gene encodes a new member of the A and B hnRNP protein group. *Nucleic Acids Res* 19:25–31.
- Herschlag D. 1995. RNA chaperones and the RNA folding problem. *J Biol Chem* 270:20871–20874.
- Hogan NC, Slot L, Traverse KL, Garbe JC, Bendena WG, Pardue ML. 1995. Stability of tandem repeats in the *Drosophila* melanogaster *HSR-omega* nuclear RNA. *Genetics* 139:1611–1621.
- Hovemann BT, Dessen E, Mechler H, Mack E. 1991. *Drosophila* snRNP associated protein P11 which specifically binds to heat shock puff 93D reveals strong homology with hnRNP core protein A1. *Nucleic Acids Res* 19:4909–4914.
- Izaurrealde E, Jarmolowski A, Beisel C, Mattaj IW, Dreyfuss G, Fischer U. 1997. A role for the M9 transport signal of hnRNP A1 in mRNA nuclear export. *J Cell Biol* 137:27–35.
- Kanaar R, Lee AL, Rudner DZ, Wemmer DE, Rio RC. 1995. Interaction of the Sex-lethal RNA binding domains with RNA. *EMBO J* 14:4530–4539.
- Kiledjian M, Dreyfuss G. 1992. Primary structure and binding activity of the hnRNP U protein: Binding RNA through RGG box. *EMBO J* 11:2655–2664.
- Klemenz R, Weber U, Gehring WJ. 1987. The white gene as a marker in a new P-element vector for gene transfer in *Drosophila*. *Nucleic Acids Res* 15:3947–3959.
- Kohtz JD, Jamison SF, Will CL, Zuo P, Lüthmann R, Garcia-Blanco MA, Manley JL. 1994. Protein-protein interactions and 5' splice-site recognition in mammalian mRNA precursors. *Nature* 368:119–124.
- Kraus ME, Lis JT. 1994. The concentration of B52, an essential splicing factor and regulator of splice site choice in vitro is critical for *Drosophila* development. *Mol Cell Biol* 14:5360–5370.
- Kumar A, Casas-Finet JR, Luneau CJ, Karpel RL, Merrill BM, Williams KR, Wilson SH. 1990. Mammalian heterogeneous nuclear ribonucleoprotein A1. Nucleic acid binding properties of the COOH-terminal domain. *J Biol Chem* 265:17094–17100.
- Lakhotia SC, Sharma A. 1996. The 93D (*hsr-omega*) locus of *Drosophila*: Non-coding gene with house-keeping functions. *Genetica* 97:339–348.
- Li H, Bingham PM. 1991. Arginine/serine-rich domains of the su(wa) and tra RNA processing regulators target proteins to a sub-nuclear compartment implicated in splicing. *Cell* 67:335–342.
- Lutz-Freyermuth C, Query CC, Keene JD. 1990. Quantitative determination that one of two potential RNA-binding domains of the A protein of the U1 snRNP complex binds with high affinity to stem-loop II of U1 RNA. *Proc Natl Acad Sci USA* 87:6393–6397.
- Matunis EL, Matunis MJ, Dreyfuss G. 1992. Characterization of the major hnRNP proteins from *Drosophila melanogaster*. *J Cell Biol* 116:257–269.
- Matunis MJ, Matunis EL, Dreyfuss G. 1992. Isolation of hnRNP complexes from *Drosophila melanogaster*. *J Cell Biol* 116:245–255.
- Matunis EL, Matunis MJ, Dreyfuss G. 1993. Association of individual hnRNP proteins and snRNPs with nascent transcripts. *J Cell Biol* 121:219–228.
- Matunis EL, Kelley R, Dreyfuss G. 1994. Essential role for a heterogeneous nuclear ribonucleoprotein in oogenesis: hrp40 is absent

- from the germ line in the dorsoventral mutant squid. *Proc Natl Acad Sci USA* 91:2781–2784.
- Mayeda A, Krainer AR. 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell* 68:365–375.
- Mayeda A, Munroe SH, Caceres JF, Krainer AR. 1994. Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. *EMBO J* 13:5483–5495.
- Mayeda A, Munroe SH, Xu R-M, Krainer AR. 1998. Distinct functions of the closely related tandem RNA recognition motifs in hnRNP A1. *RNA* 4:1111–1123.
- Merrill BM, Stone KL, Cobianchi F, Wilson SH, Williams KR. 1988. Phenylalanines that are conserved among several RNA binding proteins form part of a nucleic acid binding pocket in the A1 hnRNP. *J Biol Chem* 263:3307–3313.
- Merrill B, Williams KR. 1990. Structure/function relationships in hnRNP proteins. In: Strauss PR, Wilson SH, eds. *The eukaryotic nucleus, molecular biochemistry and macromolecular assemblies*, vol. 2. Caldwell New Jersey: Telford Press. pp 579–604.
- Michael WM, Choi M, Dreyfuss G. 1995. A nuclear export signal in hnRNP A1: A signal-mediated, temperature dependent nuclear protein export pathway. *Cell* 83:415–422.
- Mohler J, Pardue ML. 1982. Deficiency mapping of the 93D heat shock locus in *Drosophila*. *Chromosoma* 86:457–467.
- Nadler SG, Merrill BM, Roberts WJ, Keating KM, Lisbin MJ, Barnett SF, Wilson SH, Williams KR. 1991. Interactions of the A1 hnRNP and its proteolytic derivative, UP1, with RNA and DNA: Evidence for multiple RNA binding domains and salt-dependent binding mode transitions. *Biochemistry* 30:2968–2976.
- Nagai K, Oubridge C, Jessen TH, Li J, Evans PR. 1990. Crystal structures of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature* 348:515–520.
- Oubridge C, Ito N, Evans PR, Teo CH, Nagai K. 1994. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* 372:432–438.
- Pardue ML, Ballinger DG, Hogan NC. 1992. The heat shock response: Cells coping with transient stress. *Ann NY Acad Sci* 663:125–137.
- Park WJ, Liu J, Adler PN. 1994. The *frizzled* gene of *Drosophila* encodes a membrane protein with an odd number of transmembrane domains. *Mech Dev* 45:127–137.
- Piñol-Roma S, Dreyfuss G. 1992. Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature* 355:730–732.
- Pontius BW. 1993. Close encounters: Why unstructured, polymeric domains can increase rates of specific macromolecular association. *Trends Biochem Sci* 18:181–186.
- Portman DS, Dreyfuss G. 1994. RNA annealing activities in HeLa nuclei. *EMBO J* 13:213–221.
- Raychaudhuri G, Haynes SR, Beyer AL. 1992. Heterogeneous nuclear ribonucleoprotein complexes and proteins in *Drosophila melanogaster*. *Mol Cell Biol* 12:847–855.
- Ring HZ, Lis JT. 1994. The SR protein B52/SRp55 is essential for *Drosophila* development. *Mol Cell Biol* 14:7499–7506.
- Roberts DB. 1986. P element-mediated transformation. In: Roberts DB, ed. *Drosophila: A Practical Approach*. Oxford, England: IRL Press.
- Romac JM, Keene JD. 1995. Overexpression of the arginine-rich carboxy-terminal region of U1 snRNP 70K inhibits both splicing and nucleocytoplasmic transport of mRNA. *Genes & Dev* 9:1400–1410.
- Saavedra C, Tung KS, Amberg DC, Hopper AK, Cole SN. 1996. Regulation of mRNA export in response to stress in *Saccharomyces cerevisiae*. *Genes & Dev* 10:1608–1620.
- Samuels ME, Bopp D, Colvin RA, Roscigno RF, Garcia-Blanco MA, Schedl P. 1994. RNA binding by Sxl proteins in vitro and in vivo. *Mol Cell Biol* 14:4975–4990.
- Sarkar G, Sommer SS. 1990. The “megaprimer” method of site-directed mutagenesis. *Biotechniques* 8:404–407.
- Scherly D, Boelens W, Dathan NA, van Venrooij WJ, Mattaj IW. 1990. Major determinants of the specificity of interaction between small nuclear ribonucleoproteins U1A and U2B’ and their cognate RNAs. *Nature* 345:502–506.
- Shamoo Y, Abdul-Manan N, Patten AM, Crawford JC, Pellegrini MC, Williams KR. 1994. Both RNA binding domains in hnRNP A1 contribute toward single-stranded RNA binding. *Biochemistry* 33:8272–8281.
- Shamoo Y, Abdul-Manan N, Williams KR. 1995. Multiple RNA binding domains (RBDs) just don't add up. *Nucleic Acids Res* 23:725–728.
- Shamoo Y, Krueger U, Rice LM, Williams KR, Steitz TA. 1997. Crystal structure of the two RNA binding domains of human hnRNP A1 at 1.75 Å resolution. *Nature Struct Biol* 4:215–222.
- Shen J, Zu K, Cass C, Beyer AL, Hirsh J. 1995. Exon skipping by overexpression of a *Drosophila* hnRNP in vivo. *Proc Natl Acad Sci USA* 92:1822–1825.
- Siebel CW, Frescoe LD, Rio RC. 1992. The mechanism of somatic inhibition of *Drosophila* P-element pre-mRNA splicing: Multiprotein complexes at an exon pseudo-5' splice site control U1 snRNP binding. *Genes & Dev* 6:1386–1401.
- Siomi H, Dreyfuss G. 1995. A nuclear localization domain in the hnRNP A1 protein. *J Cell Biol* 129:551–560.
- Steinert PM, Mack JW, Korge BP, Gan SQ, Haynes SR, Stevens AC. 1991. Glycine loops in proteins: Their occurrence in certain intermediate filament chains, loricroins, and single-stranded RNA binding proteins. *Int J Biol Macromol* 13:130–139.
- Tang J, Rosbash M. 1996. Characterization of yeast U1 snRNP A protein: Identification of the N-terminal RBD binding site and evidence that the C-terminal RBD functions in splicing. *RNA* 2:1058–1070.
- Visa N, Alzhanova-Ericsson AT, Sun X, Kiseleva E, Bjorkroth B, Wurtz T, Daneholt B. 1996. A pre-mRNA-binding protein accompanies the RNA from the gene through the nuclear pores and into polysomes. *Cell* 84:253–264.
- Wang J, Manley JL. 1995. Overexpression of the SR proteins ASF/SF2 and SC35 influences alternative splicing in vivo in diverse ways. *RNA* 1:335–346.
- Weighardt F, Biamonti G, Riva S. 1995. Nucleo-cytoplasmic distribution of human hnRNP proteins: A search for the targeting domains in hnRNP A1. *J Cell Sci* 108:545–555.
- Wu Z, Murphy C, Callan HG, Gall JG. 1991. Small nuclear ribonucleoproteins and heterogeneous nuclear ribonucleoproteins in the amphibian germinal vesicle: Loops, spheres, and snurposomes. *J Cell Biol* 113:465–483.
- Xu R, Jokhan L, Cheng X, Mayeda A, Krainer AR. 1997. Crystal structure of human UP1, the domain of hnRNP A1 that contains two RNA-recognition motifs. *Structure* 5:559–570.
- Yang X, Bani MR, Lu SJ, Rowan S, Ben-David Y, Chabot B. 1994. The A1 and A1^B proteins of heterogeneous nuclear ribonucleoproteins modulate 5' splice site selection in vivo. *Proc Natl Acad Sci USA* 91:6924–6928.
- Zachar Z, Kramer J, Mims IP, Bingham PM. 1993. Evidence for channeled diffusion of pre-mRNAs during nuclear RNA transport in metazoans. *J Cell Biol* 121:729–742.
- Zamore PD, Patton JG, Green MR. 1992. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* 355:609–614.
- Zink B, Paro R. 1989. In vivo binding pattern of a trans-regulator of homeotic genes in *Drosophila melanogaster*. *Nature* 337:468–471.
- Zu K, Sikes ML, Haynes SR, Beyer AL. 1996. Altered levels of the *Drosophila* HRB87F/hrp36 hnRNP protein have limited effects on alternative splicing in vivo. *Mol Biol Cell* 7:1059–1073.
- Zuo P, Manley J. 1993. Functional domains of the human splicing factor ASF/SF2. *EMBO J* 12:4727–4737.