Sex-lethal (Sxl) is a RNA-binding protein, containing two conserved RNA binding domains (RBDs) and a glycine-rich region, which functions as a regulator of alternative splicing in Drosophila sex determination. Previous work demonstrated that Sxl monomers interact cooperatively upon binding to target RNAs and that the cooperativity depends on the glycine-rich N terminus. Here we use band shift experiments to show that RNA binding patterns are altered when Sxl is combined with other proteins having similar glycine-rich domains, including mammalian heterogeneous nuclear (hn) RNP L and Drosophila Hrb87F (an hnRNP A/B homolog). Direct involvement of the Sxl glycine-rich region in protein interactions was verified by Far-Western analysis. Two interaction domains, the Sxl N terminus and the Sxl first RNA binding domain, were suggested by the yeast two-hybrid assay. In a systematic examination of the RNA binding properties of Sxl domains, it was found that the Sxl N termini as well as the RBDs influence RNA binding specificity. Finally, selection of the Sxl optimal binding site (SELEX) confirms the importance of U-runs in the Sxl binding site and suggests a second type of non-U-run target that may be associated with RNA secondary structure.

How specific and general splicing factors find their target pre-mRNA sequences and determine the correct splicing pattern remains a major question in understanding alternative splicing. One emerging picture is that individual factors, identified by either biochemical or genetic methods, work in larger complexes by interacting with other factors. Understanding how the splicing factors function in these systems relies on studying their interactions with target RNAs and with other splicing proteins.

Various structural studies on a number of known RNA-binding proteins and splicing factors have identified several functional domains responsible for RNA or protein interactions. One highly conserved domain (RBD, RRM, or RNP-CS) has been intensively studied as an independent functional motif for specific RNA binding (1–4). Another conserved domain, the RS domain, has also been intensively studied as a mediator of multiple interactions with other splicing proteins (5, 6). However, the involvement of the overall protein structure in RNA binding and protein interactions has been relatively neglected. Also, potential interaction or effector domains besides the well known RS domain have not been clarified.

We have been using the sex determination pathway in Drosophila to study splicing regulation by specific factors. In this system, we are provided with a well defined hierarchy of genes regulated by alternative splicing. The Sex-lethal (Sxl) protein is a sex-specific RNA-binding protein that regulates the alternative splicing of a number of pre-mRNAs of downstream genes, as well as its own pre-mRNA (7, 8). Sxl interacts directly with multiple cis-elements on its target pre-mRNAs (9–13). We have previously studied how Sxl binds to Sxl pre-mRNA and found that Sxl monomers cooperate through their N termini when binding to pre-mRNA regions containing two adjacent binding sites, each consisting of a short U-run (12). Because Sxl binds RNA at some distance from the Sxl splice sites it regulates, we postulated that in addition to the interactions between Sxl molecules, Sxl might also interact with other specific or general splicing factors (12). This seems especially likely during autoregulation, which involves a number of additional components identified by genetic studies (14–16).

The Sxl N terminus, which we had demonstrated to be required for cooperative RNA binding (12), is very rich in glycine, serine, asparagine, and proline. A similar structure is found in a number of known RNA-binding proteins, such as mammalian hnRNP proteins A1, B2, L (18–20) and Drosophila A/B-like proteins (21–23). Among these proteins, A1 has been found to counteract SR protein SF2/ASF in alternative splicing (24); L was shown to enable processing of intronless pre-mRNA (25). Glycine-rich regions are found in a variety of other known splicing factors, including the U1 snRNP component 70K protein (26), U2AF35 (27), SF2/ASF (28, 29), and Drosophila P element splicing regulator PSI (30). Glycine-rich regions in all these proteins might function as common protein interaction domains.

In this report, we show that the glycine-rich Sxl N terminus influences interactions with hnRNP proteins containing RNA binding and glycine-rich domains. When we systematically analyzed the putative domains of Sxl for direct interaction in other assays, we confirmed the importance of the N terminus. Moreover, the first RBD may be sufficient to mediate interactions in vivo. We also show that the N and C termini, lying outside the RBDs, play a role in targeting Sxl to different RNAs. Finally, selection of the Sxl optimal binding site (SELEX) confirms the importance of U-runs in Sxl binding and suggests a second type of non-U-run target.

EXPERIMENTAL PROCEDURES

Production of glutathione S-transferase (GST) fusion proteins from bacteria, 32P-labeled RNAs by in vitro transcription, and conditions for...
in vitro RNA binding using band shift or filter binding methods were described in detail previously (12). As described by Wang and Bell (12) for Sxl and variant proteins, binding conditions were performed using a high protein/RNA ratio, and protein concentrations were on the order of 1 μM. UV cross-linking was performed by exposing an RNA binding assay (such as the “prey” vector containing the U1 RNA light-stratagene) for 15 min. The plate was kept on ice during the exposure.

DNA Constructs—The plasmids for making GSTK, GSTPTB/hnRNPin (University of Pennsylvania), M. Garcia-Blanco (Duke University Medical Center), and M. Green (University of Massachusetts Medical Center) were similarly used. Additional GST fusion plasmids were created by cloning BamHI-EcoRI fragments into pGEX-2T as follows. The cDNA clone of HrbSp7 (a gift from S. Haynes) (National Institutes of Health) was PCR-amplified with primers that added BamHI immediately upstream of the ATG(CGGGATCCATG) and an EcoRI site downstream of the open reading frame. The snf cDNA, pBC-D25 (provided by J. Romac and J. Keene, Duke University Medical Center), was transferred into pGEX-2T using the flanking BamHI and EcoRI sites in the original PET3a vector (31). The cDNA of the snf/mutant was isolated from snf/mutant homologous flies by reverse transcriptase-PCR that placed a BamHI site immediately upstream of the ATG and an EcoRI site after the open reading frame.

To make the GSTs/snfl construct, snf cDNA was reamplified by PCR using 5′-CATCTCGATGAGCGAGTCCTACCCCAAC3′ (upstream) and 5′-GGACACGCGGCTCTTGGCGGAAGCGTATT3′ (downstream) to add BamHI and AluI sites then ligated into the same sites of GST-2T-Sxlf1 (12) to replace the two Sxl RBDS with those of snf. The Sxl early cDNA fusion was created by PCR amplification that added a BamHI site immediately before the ATG and continued to a unique BamHI site. The PCR product was used to replace the BamHI to BamHI region of the plasmid pGEX-2T-Sxlf1. For the shorter version of the alternative Sxl exon 3 splicing, the BamHI to SfI region of a unique Sxl cDNA MS11 (32) was used to replace the equivalent region of Sxl cDNA CF1 (7).

The Sxlf1 deletion recombinant mutants used unique sites BspMI, BspHI, AflII, and AluI, SnfXIII9 deletions 38 amino acids. The N terminus is 117 amino acids, RBD-1 is 80 aa, RBD-2 is 86 aa, and the C terminus is 71 aa. The 5′ end of all the C-terminal deletion cDNAs has a 14 bp XhoI linker, CTCAGTCTTACTAG (all linkers from New England Bio-labs), which contains a TAG stop codon in all three reading frames. To maintain the correct reading frame of mutants N1, N3, and B2, an 8 bp XhoI linker, CCTCGAAGG, was similarly used for making N4. All constructs without the open reading frame.

The Drosophila U1 snRNA cDNA clone was a gift from Dr. S. Mount (University of Maryland). It was PCR-amplified with primers 5′-CATCTCGATGAGCGAGTCCTACCCCAAC3′ (upstream) and 5′-GGACACGCGGCTCTTGGCGGAAGCGTATT3′ (downstream) to add BamHI and AluI sites then ligated into the same sites of GST-2T-Sxlf1 (12) to replace the two Sxl RBDS with those of snf. The Sxl early cDNA fusion was created by PCR amplification that added a BamHI site immediately before the ATG and continued to a unique BamHI site. The PCR product was used to replace the BamHI to BamHI region of the plasmid pGEX-2T-Sxlf1. For the shorter version of the alternative Sxl exon 3 splicing, the BamHI to SfI region of a unique Sxl cDNA MS11 (32) was used to replace the equivalent region of Sxl cDNA CF1 (7).

The Sxlf1 deletion recombinant mutants used unique sites BspMI, BspHI, AflII, and AluI, SnfXIII9 deletions 38 amino acids. The N terminus is 117 amino acids, RBD-1 is 80 aa, RBD-2 is 86 aa, and the C terminus is 71 aa. The 5′ end of all the C-terminal deletion cDNAs has a 14 bp XhoI linker, CTCAGTCTTACTAG (all linkers from New England Bio-labs), which contains a TAG stop codon in all three reading frames. To maintain the correct reading frame of mutants N1, N3, and B2, an 8 bp XhoI linker, CCTCGAAGG, was similarly used for making N4. All constructs without the open reading frame.

Interactions between Sxl and hnRNP L—Previous work using band shift analyses demonstrated several properties of Sxl binding (12). First, Sxl binds a U-run site; second, Sxl binds cooperatively to RNA containing an activation domain were performed. Ten μl of fusion protein at about 1 mg/ml was added to 5 μl of 10 × HMK buffer (200 mM Tris, pH 7.5, 1 mM NaCl, 120 mM MgCl2), 5 μl of [γ-32P]ATP and 50 units HMK (Sigma). The reaction was incubated at 37 °C for 30 min. To the reaction, 20 μl of glutathione beads (1:1 in MTPBS (150 μM NaCl, 16 mM Na2HPO4, 4 mM NaH2PO4, pH 7.3)) was added and rotated for 5 min at room temperature. The beads were washed with 100 μl of MTPBS three times, and the bound proteins were then released with 30 μl of 20 mM Tris, pH 8, 1.5 M NaCl, and 5 mM reduced glutathione three times. All the fractions were pooled and stored at −80 °C.

To make in vitro translated Sxl probe for Western, the cDNA portion of pGEX-2TK-Sxlf1 from BamHI to EcoRI was blunt-ended and ligated to EcoRV of pcDNA3 (Invitrogen). [35S]Met-labeled Sxl protein was produced in the TNT reticulocyte lysate (Promega) according to the manufacturer protocol. The Sxlf1 Sxl band was present after translation and absent from the control of pcDNA alone (data not shown).

For Western blotting, various proteins separated by 12% SDS-PAGE were transferred to nitrocellulose. Incubation and washing followed the procedures of Kaelin et al. (33). In vitro translated protein at 105 to 106 cpm/ml was added. To check for similar protein levels, the filter was also incubated with anti-GST antibody recognizing all the fusion proteins (Pharmacia). After exposure to x-ray film to visualize the labeled Sxl binding, a Western blot was developed (data not shown).
cleotide region of the Sxl pre-mRNA containing the sequence U₉AU₈ and modifications in which the U-runs have been systematically disrupted to (UC)-runs. Consistent with previous results, the RNAs containing zero, one, or two U-runs are bound by Sxl either not at all, as a monomer, or as a dimer, respectively (Fig. 1A, lanes 1, 4, and 7; Ref. 12). In contrast, protein L (in the form of a GST-L fusion) could bind, probably as a dimer judging by its slow mobility, to all three of these RNAs (Fig. 1A, lanes 3, 6, and 9).

When Sxl and L were both added to the two-U-run RNA, a strong, broad band was observed, intermediate in size, between the Sxl and L bands (Fig. 1A, lane 8). The predominance of the intermediate complex suggests that the binding of Sxl and L influence one another, possibly by a protein interaction of the same type previously observed between Sxl and GSTSxl, which together produced a similar intermediate band (12). Like the observation on two-U-run RNA, with Sxl and L together on one-U-run RNA, an intermediate band was again observed, but it was of lower mass due to a single Sxl binding site (Fig. 1A, compare lanes 5 and 8). Strikingly, even when there was no U-run on the RNA, and so no Sxl binding site, a weak but distinct intermediate band was still observed (Fig. 1A, lane 2). The presence of an intermediate band in the absence of a Sxl binding site suggests two possibilities. A Sxl-L protein interaction might occur without any binding of Sxl to RNA; alternatively, an interaction with L may allow Sxl to bind RNA that lacks the normal U-run site of Sxl.

To show that the interaction between Sxl and L is due to the Sxl N terminus, we performed the same binding assay with SxlN1 (lacking the first 38 N-terminal amino acids) and SxlN2 (lacking the entire N terminus of 116 amino acids). It has been shown previously that SxlN1 no longer demonstrates cooperativity when binding RNA; thus, at an appropriate protein concentration, SxlN1 forms two discrete bands on two-U-run RNA in which one or two binding sites are filled (12). Here, each N-terminal deletion protein forms two bands on the two-U-run RNA although the bands are not well separated on this particular gel (Fig. 1A, lanes 10 and 13). On the same two-U-run RNA, protein L produced less of the intermediate band with SxlN1 than with Sxl (Fig. 1A, compare lanes 8 and 11).

Most clearly, with SxlN2, there was no intermediate band, and instead, there were discrete L and SxlN2 bands (Fig. 1A, compare lanes 8 and 14). The loss of the intermediate band, especially with SxlN2, strongly suggests that the Sxl N terminus is important for the Sxl-L interaction, just as previously demonstrated for the Sxl-Sxl interactions (12). These experiments were performed several times at various protein concentrations (data not shown).

In a reciprocal experiment, portions of protein L were removed to create a C-terminal deletion (GST LΔC) lacking 267 amino acids, and an N-terminal deletion (GST LΔN) lacking 36 amino acids of the glycine-rich domain. The C-terminal deletion LΔC binds RNA less well than the entire protein but forms a strong intermediate band with Sxl (Fig. 1B, lanes 5 and 6). Thus, LΔC behaves similarly to the entire L protein. In contrast, LΔN forms a band with an unexpectedly fast migration rate (Fig. 1B, lane 8). Although LΔN contains a deletion of only 27 amino acids, it migrates much faster than the entire L protein (Fig. 1B, lane 4) and even migrates faster than the much smaller protein LAC, which lacks 267 amino acids (Fig. 1B, lane 6). This faster migration rate suggests that the N-terminal deletion protein may bind as a monomer (or small multimer) while the normal protein binds as a dimer (or larger multimer). This might be analogous to the loss of cooperativity in RNA binding that was previously observed for SxlN1 (12).

Furthermore, Sxl appears to enhance the binding of LΔN; however, unlike the other situations examined, two complexes are observed that might be Sxl-Sxl and Sxl-LΔN (Fig. 1B, lane 7). One interpretation is that LΔN protein no longer interacts with itself, but still has enough of the glycine-rich region left for interaction with Sxl. Finally, replacement of the missing glycine-rich sequences in LΔN with the N-terminal 21 amino acids of Sxl (GST L/S) leads to restoration of the intermediate complex (Fig. 1B, lane 9).

In summary, Sxl and hnRNP L show different patterns of RNA binding when they are in combination compared with
when each is separate, suggesting the likelihood of a protein interaction. Deletions analysis suggests that this interaction is likely to be mediated by the glycine-rich domains of the two proteins.

**Sxl Interacts with Other Sxl Isoforms and with Another hnRNP Protein**—There are two natural variants of the Sxl N terminus. The N-terminal 26 amino acids of Sxl (isoform cF1, called Sxl herein (7)), which is the product of the late Sxl transcripts, are replaced by a different sequence of 24 amino acids in the early Sxl protein (SxlcE1, from cDNA clone cE1 (36)). In addition, alternative splicing of the late transcripts generates another isoform, SxlcF1S, in which the N terminus is missing eight amino acids just after the SxIN1 deletion (32). Given that SxlcE1 is known to have functions similar to Sxl (36), we predicted that all isoforms would interact with Sxl. As shown in Fig. 2A, lanes 1–4, that appears to be the case. Intermediate bands are evident between Sxl and GSTSxlcE1 or GSTSxlcF1S. Since two-U-run RNA was used as the substrate, both proteins are assumed to interact as they bind RNA. These interactions are virtually identical to those between Sxl and GSTSxlcF1S, which we studied in detail previously (12). However, it appears that GSTSxlcE1 interacts with Sxl less well compared with the other two proteins (Fig. 2A, lanes 1 and 2).

To see whether known hnRNP proteins in addition to protein L could interact with Sxl, we performed assays with *Drosophila* Hrb87F (hrp36), which is an hnRNP A/B type protein with a glycine-rich domain at its C terminus (21–23). Interestingly, GSTHrb87F alone binds at a very low level to two-U-run RNA, in combination with Sxl it shows a strong shifted band above the band with Sxl alone (Fig. 2A, lanes 7 and 8). A UV cross-linking experiment shows that GSTHrb87F alone is capable of a low level of binding to the two-U-run RNA; a protein concentration at least 100-fold higher than Sxl is required for equivalent levels of binding (Fig. 2B). Apparently Sxl interacts with Hrb87F to stabilize or strengthen its binding to a weak RNA site.

For comparison, two proteins lacking a glycine-rich domain were tested in this assay. Splicing factor U2AF65 and polypyrimidine track binding protein, PTB/hnRNP I, both bind poly-
Sex-lethal Interactions with Protein and RNA

Fig. 3. Interactions between parts of Sxl. A. Sxl deletion constructs and their interactions with intact Sxl. Interactions measured using the yeast two-hybrid system are listed as β-galactosidase activities averaged from three to six repeats, arising from activation of a LexA operator-lacZ reporter plasmid, or as + or − according to the ability to grow on plates lacking leucine, arising from activation of a LexA binding site-LEU2 chromosomal gene. The bait was intact Sxl in the form of a LexA/Sxl fusion; the prey proteins were the Sxl mutants that were fused to an activation domain. B. yeast two-hybrid analysis of the interactions between different Sxl deletions. Interactions are indicated as in panel A. The asterisk indicates that when SxIC3 is used as bait it activates the reporter gene in the absence of a prey plasmid; interactions were, therefore, impossible to measure. C. Far-Western blot of various GST-Sxl fusion proteins with in vitro transcribed and translated Sxl as the probe. Cross-reacting bands that are larger than the correct fusion proteins and that lack GST are seen in lanes 10–12 (designated with an arrow). The weak upper band in lane 6 is also nonspecific. The SxIC3 protein (lane 9) appears as several small bands. Protein levels were equalized to show equivalent staining after SDS-PAGE; identity was verified by Western blot with anti-GST.

Singly or together bind little or not at all (Fig. 3C, lanes 10–12, SxIB1, B2, B12), and deletion of either RB has no effect on the interaction (Fig. 3C, lanes 13 and 14, SxID1 and D2). Surprisingly, the C-terminal region alone (SxIN4) interacts well with Sxl (Fig. 3C, lane 6). However, it is possible that the C terminus is normally prevented from intermolecular interactions by other parts of Sxl structure because not all proteins having a C terminus interact with Sxl.

We also used the yeast two-hybrid assay to analyze interactions between Sxl and a series of altered Sxl proteins (Fig. 3A). Activities were measured using lacZ or LEU2 reporter genes that respond to interactions between wild-type Sxl fused to the LexA DNA binding domain (the bait) and Sxl deletion mutants fused to an activation domain (the prey). Some of the differences in activity can be attributed to a level of protein expression in yeast cells that is lower for intact Sxl than for any of the Sxl deletions (data not shown). In this assay, the N-terminal region appears neither necessary nor sufficient for a significant interaction (Fig. 3A). Nevertheless, the decrease in activity from SxIN1 to SxIN2 does suggest that the N terminus might play some role in protein interaction in this system (Fig. 3A, SxIN1, SxIN2).

It is more obvious that RB-1 is important for interaction (Fig. 3A). For example, the N-terminal deletions (Fig. 3A, SxIN1–N4) do not completely lose activity until RB-1 is lost. In addition, the isolated RB-1 (Fig. 3A, SxIB1) is capable of interacting with Sxl. However, the presence of RB-1 is not always sufficient because the C-terminal deletions SxIC1 and SxIC2 contain RB-1 but do not interact with Sxl. This might be due to the influence of overall protein structure; for example, without the C terminus, the N terminus might block the ability of RB-1 to interact with intact Sxl. Alternatively, the C-terminal deletion proteins may interact so strongly with each other (see below) that this overwhelms the measured interaction with intact Sxl.

When individual domains were used as bait and prey, the ability of RB-1 to behave as an interacting domain became even more apparent (Fig. 3B). The two proteins containing RB-1, SxIB1 and SxIC2, each interacted strongly with itself and with the other protein but did not interact with other regions (Fig. 3B). This is in striking contrast to the failure of SxIC2 to interact with intact Sxl (Fig. 3A, SxIC2). Thus, it appears that for each of the C-terminal deletions, the deletion proteins might interact so strongly among themselves that this overwhelms the interaction that is measured with intact Sxl. Such an explanation may also be invoked to explain the lack of apparent interaction of the isolated N terminus. As an additional consideration, it is possible that interactions between two RNA binding domains are dependent upon the presence of cellular RNA. If RNA molecules are involved, they might bring together the bait and prey proteins with or without the involvement of direct protein-protein contacts.

The N terminus alone activates transcription of the lacZ reporter gene even in the absence of the prey plasmid (Fig. 3B, SxIC3). Apparently, the artificially exposed Sxl N terminus works as a transcriptional activation domain when fused to the LexA DNA binding domain. This suggests that the N terminus can interact with other proteins, albeit with the transcription apparatus. Nevertheless, the N terminus can be assayed as the prey plasmid because it is already fused to an activation domain, but even here it shows no obvious interaction with any other Sxl domain (Fig. 3B, SxIC3). Again, self-interaction may play a role, and competition between the RB-1 interaction and the N-terminal interaction may also be involved. Unfortu-
nately, the interaction between the two isolated N termini cannot be tested.

In considering a final difference between the two assays, it is unclear why the C terminus (SxlN4) interacts with Sxl on the Far-Western blot (Fig. 3C, lane 6) but shows no evidence of interaction in the yeast two-hybrid system (Fig. 3A and B, SxlN4). This may result from differences in fusion proteins used in each assay. In yeast, a LexA-Sxl fusion must interact with a SxlN4-activation domain fusion. In the Far-Western assay, the GST-SxlN4 fusion was probed with nonfusion Sxl.

In summary, Far-Western analysis clearly shows the importance of the glycine-rich Sxl N terminus in protein interactions. The yeast two-hybrid assay appears to measure two different interaction domains, the N terminus and RBD-1.

Influence of Sxl Protein Structure on Binding Properties—The Sxl deletions used above were also tested for RNA binding ability. The following three RNA substrates were used: (a) RNA S5A, which contains a U-run RNA from the Sxl regulated 3' splice site, contains a run of 8 Us (U-run RNA); S5A, from the Sxl downstream common 3' splice site, contains no U-runs (no-U RNA); and Drosophila U1 snRNA (U1 snRNA). Protein amounts were adjusted where possible to produce approximately equal binding to U-run S5A RNA. A, a summary of the band shift analysis. For each mutant protein, the specificity is indicated as normal (+), nonspecific (ns), or new. The RNA binding affinity was estimated from the gels shown in panel A after correcting for the varying amounts of protein used in each lane.

Determination of the Sxl Optimal Binding Site (SELEX)—To study further how Sxl interacts with RNA, we performed selection/amplification (SELEX) of binding sites from a random pool of 26-mers to identify the Sxl optimal binding sites. Listed in Fig. 5A are 15 RNA sequences that bound Sxl. A run of 8 or more undisrupted Us, and a disrupted U12 (Fig. 5B, lanes 8–11), correlates with stronger binding compared with those with a U-run of 7 or less, or a disrupted run of 8, 9, or 10 Us (Fig. 5B, lanes 1–7). The only RNA with 16 Us showed very strong binding and formation of a higher complex, apparently with two Sxl proteins that bound cooperatively (Fig. 5B, + and + lanes 11). This conclusion is in agreement with our previous
report that Sxl binds to the U-runs on the Sxl and tra pre-mRNA, as a monomer or dimer according to the run length (12). The low frequency of long U-runs obtained may be because such sequences are more prone to mutagenesis during PCR amplification. Two sequences lack a U-run (RNAs 14 and 15). It is interesting that two bands are present initially in these RNAs, presumably due to secondary structure, but Sxl appears to bind only the upper band (Fig. 5B, lanes 14 and 15). Sxl does not bind to the control RNA, C2, which also lacks U-runs.

We next used some of these newly isolated sequences as binding substrates for Sxl and the isolated Sxl RNA binding domains (SxIB1 and SxIB2). They were tested in a filter binding assay together with the same RNAs used in Fig. 4. U-run RNA S5A, U1 snRNA, and no-U-run RNA S8A. Reflecting the band-shift experiment in Fig. 5B, wild-type Sxl bound the SELEX sequences 15, 6, 10, and 8 with increasing affinity (Fig. 5C, Sxl). The first Sxl RBD (SxIB1) binds all substrates containing U-runs (RNAs 6, 8, 10) with somewhat reduced affinity compared with the wild-type protein (Fig. 5C, SxIB1 (hatched bars)) similar to the band shift results in Fig. 4 (compare SxIB1 and Sxl on U-run RNA). However, modifying the results of Fig. 4, in this assay, it is seen that RBD-1 (SxIB1) actually has a specificity somewhat different from wild-type Sxl because it bound to the SELEX RNA 15, which lacks U-runs, much more strongly than wild-type Sxl (Fig. 5C, compare Sxl (black bars) and SxIB1 (hatched bars) on RNA 15). The second Sxl RBD (SxIB2, gray bars), in agreement with the results shown in Fig. 4, binds U-run containing RNAs 6 and S5A weakly and binds U1 snRNA quite well (Fig. 5C). None of the proteins bound to the control RNA C1.

In summary, all the tested SELEX sequences having U-runs behaved similarly to U-run RNA S5A, binding Sxl better than RBD-1 or RBD-2. In contrast, the no-U-run RNA 15 and U1 snRNA were preferred by RBD-1 or RBD-2, respectively. Overall, the SELEX results argue for two types of binding: one containing U-runs influenced by their sequence context, the other somewhat weaker and lacking U-runs but associated with secondary structure.

DISCUSSION

The Glycine-rich Region as a Protein-Protein Interaction Motif—It is currently thought that for Sxl to regulate the splicing of its own pre-mRNA, it must directly contact other splicing proteins. In this way, Sxl would be able to bridge the considerable distance between its intronic binding sites and the splice sites it regulates. It has already been demonstrated that Sxl molecules interact: cooperative interactions occur between two Sxl monomers as they bind to adjacent U-run sites on the Sxl pre-mRNA (12). Because the cooperativity was shown to be dependent upon the glycine-rich N terminus, that region was thought to be a protein interaction domain. It was also proposed that the large glycine-rich region may simultaneously interact with Sxl and with additional proteins (12).

In this study, we used band-shift assays as before to demonstrate that Sxl can interact on a small region of Sxl pre-mRNA with other proteins having a glycine-rich region. In the case of human hnRNP L, when combined with Sxl, an RNA-protein complex is formed that is intermediate in size between that formed by either protein alone. It is especially noteworthy that the intermediate complex forms even on RNA that lacks any Sxl binding site, suggesting that either L protein helps Sxl bind to a site that Sxl does not bind alone, or Sxl binds the L protein without itself binding RNA. If the first possibility is correct, it would be similar to the interaction observed between Sxl and Drosophila hnRNP protein Hrb87F. Hrb87F binds extremely weakly to the U-run that Sxl normally binds; however, in combination with Sxl, Hrb87F binding is much improved (Fig. 2A and B).

These experiments also address the question of whether, when two cooperating Sxl monomers are bound to RNA, they might be capable of simultaneously interacting with additional proteins. When L and Sxl were combined upon RNA containing two adjacent Sxl binding sites, to which Sxl alone would bind cooperatively (12), an intermediate band was observed (Fig. 1A, lane 2). This result suggests that a simultaneous interaction between two Sxl molecules, L protein, and RNA may be possible.
Although these experiments were meant to test the in vitro interaction properties of the Sxl N terminus, it would not be surprising if the interaction between Sxl and certain Drosophila hnRNP proteins turns out to have relevance in vivo. A number of mammalian hnRNP proteins, including A1, F, and I/PtB, have been shown to directly influence splicing (40–43). In addition, it was recently shown that overexpression of Drosophila Hrb98DE, which is an hnRNP A/B protein, causes exon skipping in vivo (43).

Direct evidence for the involvement of the glycine-rich N terminus in protein interactions was provided by Far-Western analysis of the interactions between two Sxl molecules. The Sxl N-terminal region alone is sufficient for an interaction with Sxl (Fig. 3C). Consistent with these results, when the Sxl termini were added to snf, which is a protein that has two RBDs but no glycine-rich region, it acquired the ability to interact with Sxl (Fig. 2C).

The yeast two-hybrid assay gave results consistent with the idea that the N terminus is important for the interaction between Sxl molecules (Fig. 3A). However, the assay was complicated by the fact that the isolated Sxl N terminus is intrinsically able to activate transcription, presumably by inappropriate protein interaction with transcription factors (Fig. 3B). In addition, it appears likely that in yeast, the C-terminal deletions, including the isolated N terminus, might each form a tight interaction with itself that precludes interactions with other partners (Fig. 3, A and B). The N terminus might also interact tightly with unknown cellular proteins to similarly interfere with the measured interactions.

The behavior of the Sxl N terminus as a transcriptional activation domain is consistent with other observations of similar behaviors by glycine-rich regions of certain hnRNP proteins. For example, when the glycine-rich N terminus of hnRNP P2 (also called TLS/FUS) is fused by chromosomal translocation to CHOP, a C/EBP family DNA binding transcription factor, it acts as a transcriptional activator that produces dominant enhancement of liposarcomas (44–47). Similarly, the glycine-rich N terminus of RNA-binding protein EWS becomes oncogenic when fused to a series of DNA binding domains (for example, see Ref. 48). Finally, when the glycine-rich domain from Drosophila SARFH, a homolog of EWS and TLS/FUS, is fused to CHOP, the fusion protein leads to cell transformation (47, 49).

Unexpectedly, we found that in the yeast system the isolated RBD-1 behaved as an interaction domain, although in the Far-Western assay it did not (Fig. 3C). It may be that certain interactions are stabilized by yeast cellular RNA, possibly through transcripts containing poly(U) stretches. RBD-1 may be able to interact with other proteins only after binding to RNA; alternatively, activation may be achieved indirectly if the two proteins containing RBD-1, as bait and prey, both bind to a single cellular transcript. With regard to the importance of the RBD for protein interactions, we note that interactions between wild-type Sxl monomers were previously found to be dependent on RNA binding (12), and it is possible that normally the glycine-rich domain may coordinate with an RBD.

The glycine-rich domains found in various RNA-binding proteins do not have any clear structural features beyond the likelihood that stretches of amino acids such as glycine or proline result in flexible coils. These regions may be important for a variety of protein interactions in splicing, acting analogously to the RS domains, which connect a number of different proteins involved in general and regulated splicing (50). The interactions of Sxl with itself, with hnRNP L, and with Hrb87F provide preliminary evidence for the possibility of such a network.

Sxl Binding to RNA—We also examined the role of the Sxl RNA binding domains, and their functional interplay with the glycine-rich and C-terminal regions, in the recognition of RNA sequences. To date, most of the studied cases involving proteins with two or more RBDs have shown that an isolated RBD domain loses RNA binding specificity, or affinity, or both (1–5, 13, 51). For Sxl, we find that each isolated RBD has acquired an altered specificity, which is either mildly altered in the case of RBD-1 (SxlB1) or radically altered in the case of RBD-2 (SxlB2). In addition, the isolated pair of RBDs (SxlB12) has become nonspecific. Somewhat surprisingly, this lack of specificity can be rescued by addition of either the N or C terminus. This result suggests that an important factor in binding specificity is the establishment of an overall structural stability rather than any particular protein sequence. We would suggest that when two RBDs must act together, the overall structural context can be important; in the case of Sxl, this context can be provided by either the N or C terminus.

There are several differences between our results and those of a similar study by Kanaar et al. (13). Briefly, Kanaar et al. found that each RBD alone showed nonspecific binding and lower binding affinity than the intact protein, with very low affinity shown by RBD-2. The isolated pair of RBDs showed correct specificity and even stronger binding than the intact Sxl protein. Surprisingly, their intact protein was less specific in recognizing U-runs than the RBD pair. We note three points of difference. First, we found that Sxl RBD-1 shows normally normal specificity with regard to recognizing U-runs although with somewhat lower affinity. Nevertheless, we also find that RBD-1 specificity is not completely normal because it bound a SELEX sequence that lacks U-runs more strongly than full-length Sxl (Fig. 5C, compare S5 and S5A). The discrepancy may arise from differences in extent of RBD-1 where we excluded the linker region between the two RBDs while Kanaar et al. (13) included it. Alternatively, the testing of different RNAs may have led to different results. Second, we find that RBD-2 has not simply lost affinity but rather has changed its specificity so that it exhibits a surprisingly strong affinity toward Drosophila U1 snRNA. When the C terminus is added to RBD-2 (SxlN3), it regains some ability to bind U-runs and becomes generally nonspecific like the RBD-2 of Kanaar et al. (13). This result suggests that some of the differences may be due to the RBD length. Third, and importantly, we find no evidence for a negative role of the termini that decreases both binding affinity and specificity of the RBD pair. Instead, we find that the two RBDs together have lost all binding specificity, and it is not regained until either the N or C terminus is added (SxlC1, SxlN2). Moreover, we observe that the entire protein shows strong binding and accurate specificity capable of distinction between U-run and no-U-run polypyrimidine tracts. Finally, in contrast to Kanaar et al. (13), we and others have previously demonstrated that Sxl does not bind the U-rich β2 polypyrimidine tract under our chosen conditions (9, 12). These differences may result from different methods of cloning and isolating the proteins.

As part of our study on Sxl interactions with RNA, we analyzed the Sxl binding site. Our results of selection and amplification of the optimal binding site (SELEX) do not concur with consensus sequences previously reported by others in which Sxl binding sites included AU,NAGU (52) or U5(U/G)UU(U/G)U5 where Gs interrupting the Us are favored (43). These reported consensus sequences, emphasized because of their appearance within the Sxl binding site on tra pre-mRNA, are not found in most of the natural binding targets of Sxl on Sxl pre-mRNA (12, 53, 54) or male-specific lethal-2 RNA (55). Nevertheless, there is some resemblance to the consensus of Singh
et al. (43) where several sequences contain U-runs flanked by Gs, and one strong binding sequence has a U-run interrupted by Gs (RNA8). Aside from the interpretation, the results of Sakashita and Sakamoto (52) were very close to ours with regard to the U-run lengths, locations, and binding strength, as well as the identification of a few sequences lacking U-runs. Our observation of the sequences lacking U-runs suggests that a second type of somewhat weaker binding site associated with RNA secondary structure may exist.

Acknowledgments—We thank G. Dreyfuss, S. Jamison, M. Garcia-Blanco, M. Green, S. Haynes, J. Keene, and S. Mount for generously providing cDNA clones. We thank Gail Miyasato for excellent assistance with constructs and purifications.

REFERENCES

1. Burd, C. G., and Dreyfuss, G. (1994) EMBO J. 13, 1197–1204
2. Tacke, R., and Manley, J. L. (1995) EMBO J. 14, 3540–3551
3. Zuo, P., and Manley, J. L. (1993) EMBO J. 12, 4727–4737
4. Caceres, J. F., and Krainer, A. R. (1993) EMBO J. 12, 4715–4726
5. San Juan, A. B., Santamaria, P., and Sanchez, L. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 8769–8773
6. Wu, J. Y., and Maniatis, T. (1993) J. Biol. Chem. 268, 3645–3650
7. Cavalli, A., Mandahl, N., and Ron, D. (1993) J. Cell Biol. 122, 1874–1886
8. Kobayashi, M., Kuroda, M. I. (1995) Cell 81, 867–877
9. Mayeda, A., and Krainer, A. R. (1992) Genes & Dev. 6, 3584–3602
10. Kaelin, W. G., Jr., Krek, W., Sellers, W. R., DeCaprio, J. A., Ajchenbaum, F., Fuchs, C. S., Chattenden, T., Li, Y., Farnham, P. J., Bhanar, M. A., Livingston, D. M., and Flemington, E. K. (1992) Cell 70, 351–364
11. Dreyfus, G. (1994) EMBO J. 13, 2575–2587
12. Wang, J., and Bell, L. R. (1994) Genes & Dev. 8, 2072–2085
13. Matunis, M. J., Matunis, E. L., and Dreyfuss, G. (1992) Mol. Cell. Biol. 12, 1163–1176
14. Mayeda, A., and Krainer, A. R. (1992) Genes & Dev. 6, 3645–3650
15. Patton, J. G., Mayer, S. A., Tempst, P., and Nadal-Ginard, B. (1991) Genes & Dev. 5, 1224–1236
16. Gil, A., Sharp, P. A., Maniatis, T., and Garcia-Blanco, M. A. (1991) Genes & Dev. 5, 1237–1251
17. Horabin, J. I., and Schedl, P. (1993) Mol. Cell. Biol. 13, 2992–3000
18. Haynes, S. R., Johnson, D., Raychaudhuri, G., and Beyer, A. L. (1991) Nucleic Acids Res. 19, 25–31
19. Matunis, M. J., Notario, V., and Wilson, S. H. (1986) J. Biol. Chem. 261, 1063–1071
20. Cavalli, A., Mandahl, N., and Ron, D. (1993) J. Cell Biol. 122, 1874–1886