# The Sex-lethal amino terminus mediates cooperative interactions in RNA binding and is essential for splicing regulation

**Jiwu Wang and Leslie R. Bell**

Molecular Biology Program, Department of Biological Sciences, University of Southern California, Los Angeles, California 90089-1340 USA

***Sex-lethal* (*Sxl*) acts as a binary switch that regulates *Drosophila* sexual differentiation and dosage compensation and also maintains a stable female state through autoregulation. As part of a cascade of genes that are regulated by sex-specific splicing, *Sxl* controls the sex-specific splicing of *transformer* (*tra*) RNA and also its own RNA. Sxl contains two RNP-CS (RNA-binding) domains and is known to bind *tra* pre-mRNA near the alternative 3′ splice site, thus blocking use of that site to give the female-specific splicing pattern. Here, we test how Sxl protein interacts with *Sxl* RNA during autoregulation. We show that Sxl not only binds *Sxl* pre-mRNA near the alternative 3′ splice site but also at distant, multiple sites surrounding the *Sxl* alternative exon. Moreover, Sxl binds cooperatively at these multiple sites. The Sxl amino terminus is essential for the cooperative interaction and is also required for regulatory activity in vivo. It appears that this region of Sxl protein, which resembles regions in some other RNA-binding proteins, is a domain that mediates protein–protein interactions during RNA binding and plays an important role in splicing regulation.**

[*Key Words*: Sex determination; RNA binding; Sex-lethal; alternative splicing]

Received June 3, 1994; revised version accepted July 19, 1994.

---

*Drosophila* has adapted the general mechanism of alternative RNA splicing to make a sensitive, stable, binary switch that regulates the choice between the male and female cell fates (Baker 1989; McKeown 1992a,b). This use of alternative splicing as a developmental switch is a minor variation on its nearly ubiquitous use for expressing protein isoforms in a tissue- or stage-specific manner (Smith et al. 1989). Although alternatively spliced genes are extremely common, only a few of the *trans*-acting factors controlling the many differential splicing patterns have been identified (Green 1991; McKeown 1992a). In the case of *Drosophila* sex determination, a part of the pathway consists of an ordered hierarchy of genes that are regulated at the level of pre-mRNA splicing; therefore, these genes define a set of splicing regulators and their respective RNA targets.

At the top of the splicing hierarchy is the gene *Sex-lethal* (*Sxl*) (Cline 1988; Nagoshi et al. 1988). *Sxl* has an autoregulatory activity that serves to direct the splicing of its own pre-mRNA (Bell et al. 1991), and it has been demonstrated to directly regulate the splicing of *transformer* (*tra*), the next gene in the somatic sexual differentiation pathway (McKeown et al. 1988; Nagoshi et al. 1988). tra protein in turn, together with tra-2 protein, regulates the splicing of *doublesex* (*dsx*) (Nagoshi et al. 1988). In addition to regulating *tra* in the somatic sexual differentiation pathway, *Sxl* also dictates germ-line sex determination (Schüpbach 1985; Steinmann-Zwicky et al. 1989; Oliver et al. 1993) and negatively regulates the dosage compensation system that hyperactivates the single-copy X chromosome in males (Cline 1979; Lucchesi and Skripsky 1981; Gorman et al. 1993; Bernstein and Cline 1994). Because of its role in dosage compensation, a mutation in the *Sxl* gene can lead to the death of the fly. It has been postulated that *Sxl* regulates itself as well as all of its downstream genes at the level of splicing (Bell et al. 1988).

Genetic analyses have established that *Sxl* acts as an on/off switch to set the sexual state: When *Sxl* is "on", development will be female, and when *Sxl* is "off", development will be male (Cline 1978, 1979). Reflecting the different activities in the two sexes, *Sxl* transcripts are differentially spliced in males and females (Bell et al. 1988). Male transcripts are spliced to include the male-specific exon, exon 3, which contains a nonsense codon that causes premature termination of Sxl protein translation (see Fig. 1, below); in contrast, the female *Sxl* transcripts skip exon 3 and encode full-length Sxl protein. The continuous maintenance of the female state depends on the regulation of *Sxl* RNA splicing by Sxl protein, such that exon 3 is excluded through autoregulation in females and exon 3 is included by default in males where *Sxl* is absent (Cline 1984; Bell et al. 1991). We note that *Sxl* is initially regulated at the transcriptional level in response to the X chromosome to autosome (X/A) ratio (Keyes et al. 1992). The initial female-

specific expression of Sxl protein starts the autoregulatory loop by controlling the alternative splicing of the previously mentioned sex-specific transcripts, which appear a short time later and persist throughout development.

It was originally suggested that Sxl regulates both itself and other genes by binding directly to substrate pre-mRNAs because it contains two copies of a highly conserved domain, termed the ribonucleoprotein (RNP) consensus sequence RNA-binding domain (RNP-CS-type RBD), RNA recognition motif (RRM), or RNP motif (Bell et al. 1988; Bandziulis et al. 1989; Kenan et al. 1991; Mattaj 1993). This motif is found in many RNA-binding proteins, including proteins of heterogenous nuclear RNP (hnRNP) and small nuclear RNP (snRNP) complexes. Like Sxl, many of these proteins have more than one RNP-CS domain.

The interaction between Sxl protein and *tra* RNA has been investigated extensively. Sxl directs the splicing pattern of the female-specific *tra* transcript, as compared with the default pattern of the *tra* non-sex-specific transcript (Boggs et al. 1987). As a consequence of using the female-specific site, an early translation stop codon is avoided and functional tra protein is expressed exclusively in females. In males, where Sxl is absent and the non-sex-specific 3′ splice site is used, tra protein is not expressed. In vivo experiments have demonstrated that Sxl acts to block the use of the *tra* non-sex-specific 3′ splice site; so the splicing machinery uses the downstream, female-specific 3′ splice site. Sosnowski et al. (1989) have termed this the "blockage model". Because the normal sex-specific regulation was destroyed by mutations at a stretch of 8 U bases just upstream of the *tra* non-sex-specific 3′ splice site, it was concluded that this $U_8$ stretch is the site of Sxl action (Sosnowski et al. 1989). Direct binding has been demonstrated in vitro, showing that Sxl protein binds to the $U_8$ stretch on *tra* RNA (Inoue et al. 1990; Valcárcel et al. 1993). A direct binding competition between Sxl and U2AF, a general splicing factor that binds to the polypyrimidine tract at the 3′ splice site to facilitate effective binding of U2 snRNP (Ruskin et al. 1988; Zamore et al. 1992), has been shown recently to be responsible for the regulation of *tra* splicing (Valcárcel et al. 1993). It was hypothesized by analogy with *tra* that a similar $U_8$ stretch just upstream of the *Sxl* male-specific exon (exon 3) could be the site at which Sxl protein exercises its autoregulatory function (Sosnowski et al. 1989).

Does Sxl protein interact with its own RNA in a manner similar to its interaction with *tra* RNA? Several lines of evidence suggest that it may not. *Sxl* has a more complex biological role than *tra* because it must regulate a number of genes, whereas *tra* apparently regulates only *doublesex*. *Sxl* has an autoregulatory activity responsible for maintaining the female determined state; thus, its activity state is critical for the pathway at all times. In contrast, *tra* is constantly under the control of *Sxl*. Furthermore, although the regulatory mechanisms of *Sxl* and *tra* splicing may appear superficially the same, the structures of the two genes are quite dissimilar. The entire *Sxl* gene is ~20 kb and the region of the regulated exon and surrounding introns is ~4 kb, whereas the entire *tra* gene is <1 kb and the alternative introns are ~70 and 250 bp (Boggs et al. 1987; Bell et al. 1988; Samuels et al. 1991). In addition, *Sxl* is regulated by exon skipping, whereas *tra* is regulated by a choice between alternative 3′ splice sites. Finally, recent in vivo studies have demonstrated that multiple *cis*-acting elements surrounding the male-specific exon are required in the autoregulation of *Sxl* splicing (Sakamoto et al. 1992; Horabin and Schedl 1993a). Additional studies point to the 5′ splice site rather than the 3′ splice site as the region of importance (Horabin and Schedl 1993a; Flickinger and Salz 1994). These results indicate that different molecular mechanisms may be used in the regulation of *Sxl* and *tra* splicing.

The purpose of this study was to determine how Sxl binds its own pre-mRNA. In addition to observing RNA binding at the 3′ splice site of the male exon (exon 3), we also identified additional binding sites surrounding the male exon. Most significantly, we found cooperative binding properties of the Sxl protein. The portion of Sxl protein that is responsible for the cooperativity shares amino acid similarity with other RNA-binding proteins and is essential for Sxl's function as a splicing regulator in vivo. These results suggest that, as has been well documented for transcription, the use of multiple binding sites and cooperative interactions can be important for the control of alternative splicing.

## Results

### *Sxl protein binds specifically near the regulated 3′ splice site of* Sxl

To determine whether Sxl protein can bind specifically to the 3′ splice site of the *Sxl* male-specific exon (exon 3), as predicted by the blockage model, we performed in vitro RNA-binding assays using bacterially produced Sxl protein (isoform cF1, referred to as Sxl in this paper; Bell et al. 1988). This Sxl protein has been shown to induce the female splicing pattern of *Sxl* and *tra* transcripts when ectopically expressed in male transgenic flies (Bell et al. 1991). We compared Sxl binding to a transcript (S5A) encompassing the 3′ splice site of the *Sxl* male-specific exon and a transcript (S8A) that includes the 3′ splice site of the unregulated, common exon 4. In a band-shift experiment, S5A entered a low-mobility band in a protein concentration-dependent manner, whereas S8A did not (Fig. 1A). S8A did not change mobility until a smear, probably from nonspecific binding, was observed when the protein concentration reached >10 times that required to shift half of the other transcript. In other controls, Sxl at these concentrations did not bind to RNAs encoded by the polylinker regions of the plasmids that were used to produce the in vitro transcripts (data not shown). These data indicate that Sxl protein binds to the region of the *Sxl* male-specific 3′ splice site but not the downstream common 3′ splice site.

Similar binding experiments were performed with

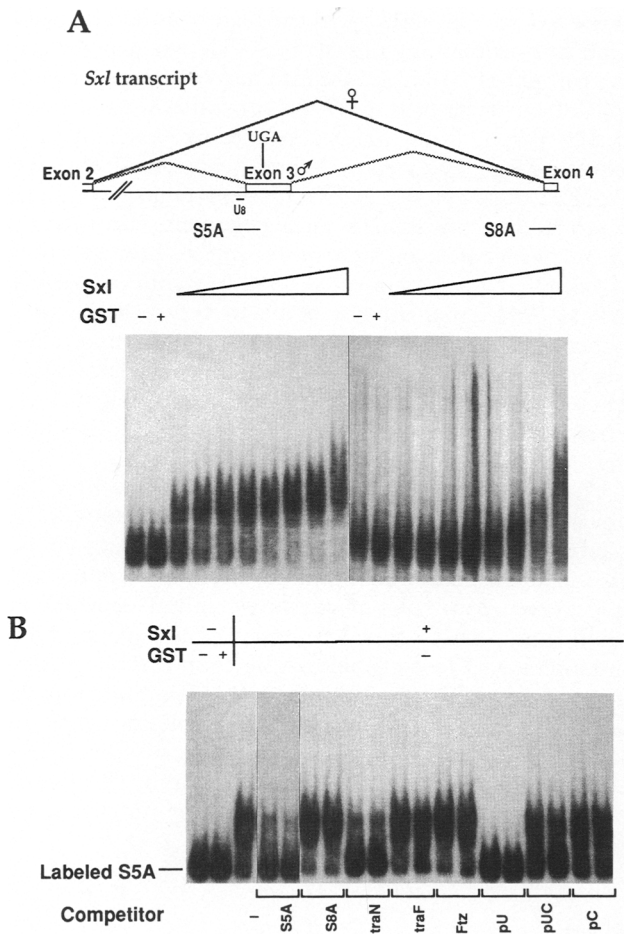


**Figure 1.** Sxl protein binds specifically to the 3′ splice site region of the *Sxl* male-specific exon. (*A*) The alternatively spliced male-specific and female-specific transcripts of *Sxl* are shown schematically for the region from exons 2 to 4. The diagram is roughly to scale except for the deleted part of one intron. Exon 3 is the male-specific exon. The $U_8$ sequence of the male-specific 3′ splice site and the translation termination codon are also shown. The transcripts used as binding substrates, S5A and S8A, are indicated above the corresponding gel panels. Sxl protein cF1 (Sxl) produced in *E. coli* was used at increasing concentrations, from 0.1 to 1 μM, in the mobility shift assay shown. The portion of GST that is encoded by the expression vector pGEX-2T was included as a control at a comparable concentration. (*B*) Competition binding experiment. Labeled *Sxl* S5A transcript was mixed with 0.6 μM Sxl. The band of free probe and lanes with different competitors are indicated. Competitors S5A and S8A from the Sxl transcript are diagramed in *A*; traN contains the region of the *tra* non-sex-specific 3′ splice site; traF contains the *tra* female-specific 3′ splice site; Ftz encompasses the *fushi tarazu* intron as a negative control; (pU, pUC, and pC) Poly(U), poly(UC), and poly(C). Each pair of competition reactions includes either a 20-fold excess (first lane) or a 50-fold excess (second lane) of competitor RNA relative to the labeled S5A. Competitor was measured in molar excess (S5A through Ftz) or by weight excess (pU, pUC, pC).

transcripts encoded by *tra*. Not surprisingly, the RNA containing the regulated, non-sex-specific 3′ splice site-formed protein–RNA complexes with Sxl protein (data not shown). Sxl also did not bind to either the *tra* female-specific 3′ splice site or the 3′ splice site of the *Drosophila fushi tarazu* transcript. These results indicate that our in vitro-binding system agrees with the in vivo and in vitro results of other investigators (Sosnowski et al. 1989; Inoue et al. 1990; Valcárcel et al. 1993).

The RNA-binding affinity of Sxl that we observed is rather low compared with previous studies, as is the ratio of specific to nonspecific binding affinity. The previously reported apparent binding affinity of Sxl for *tra* RNA, from the protein concentration that shifted half of the RNA substrate, was $K_d \sim 10^{-9}$ M (Valcárcel et al. 1993). In comparison, U2AF binds to various polypyrimidine tracts with dissociation constants ($K_d$) from $\sim 10^{-5}$ to $10^{-8}$ M (Zamore et al. 1992), which covers the approximate range we observed for Sxl. Our detailed calculations are described in a subsequent section. Several experimental differences between the two binding systems, such as the salt concentration (about three times higher in our binding reactions), the differences between the RNA substrates, and the determination of active protein concentrations, could affect the quantitative outcomes.

Various unlabeled RNAs were tested as competitors for the binding of Sxl to the *Sxl*-regulated 3′ splice site (transcript S5A). Both S5A and the region of the *tra* non-sex-specific 3′ splice site (traN), which had been observed to form complexes with Sxl protein, acted as competitors in a dosage-dependent way (Fig. 1B, S5A, traN). Those RNAs that had not been able to form complexes with Sxl protein did not compete even in 50-fold molar excess (Fig. 1B, S8A, traF, Ftz). Consistent with the known importance of the $U_8$ region of *tra*, poly(U) competed away the binding nearly completely with 20-fold excess, whereas poly(UC) and poly(C) showed very weak, if any, competition. We note that like all our binding assays, the competition was performed with ~100-fold molar excess of Sxl protein and $10^5$-fold by weight excess of yeast tRNA. These competition experiments further verified the specificity of Sxl binding and also suggested that Sxl protein recognizes the regulated 3′ splice site regions of both *Sxl* and *tra* by a mechanistically similar means, which involve the recognition of U bases.

### *Sxl protein can bind to multiple sites on the* Sxl *pre-mRNA*

We wanted to know whether, in addition to binding the regulated 3′ splice site, Sxl protein could also bind to other regions of the *Sxl* transcript surrounding the regulated male exon. This region contains several U-runs, which could be potential Sxl-binding sites. Figure 2 shows all of the U-runs of at least 6 U bases between *Sxl* exon 2 and exon 4, which includes the alternatively spliced, male-specific, exon 3. We divided this region into transcripts of several hundred bases and assayed binding with Sxl protein as before. As shown in Figure 2, in addition to the region of the male-specific 3′ splice site (S5), there are several other regions that bind Sxl

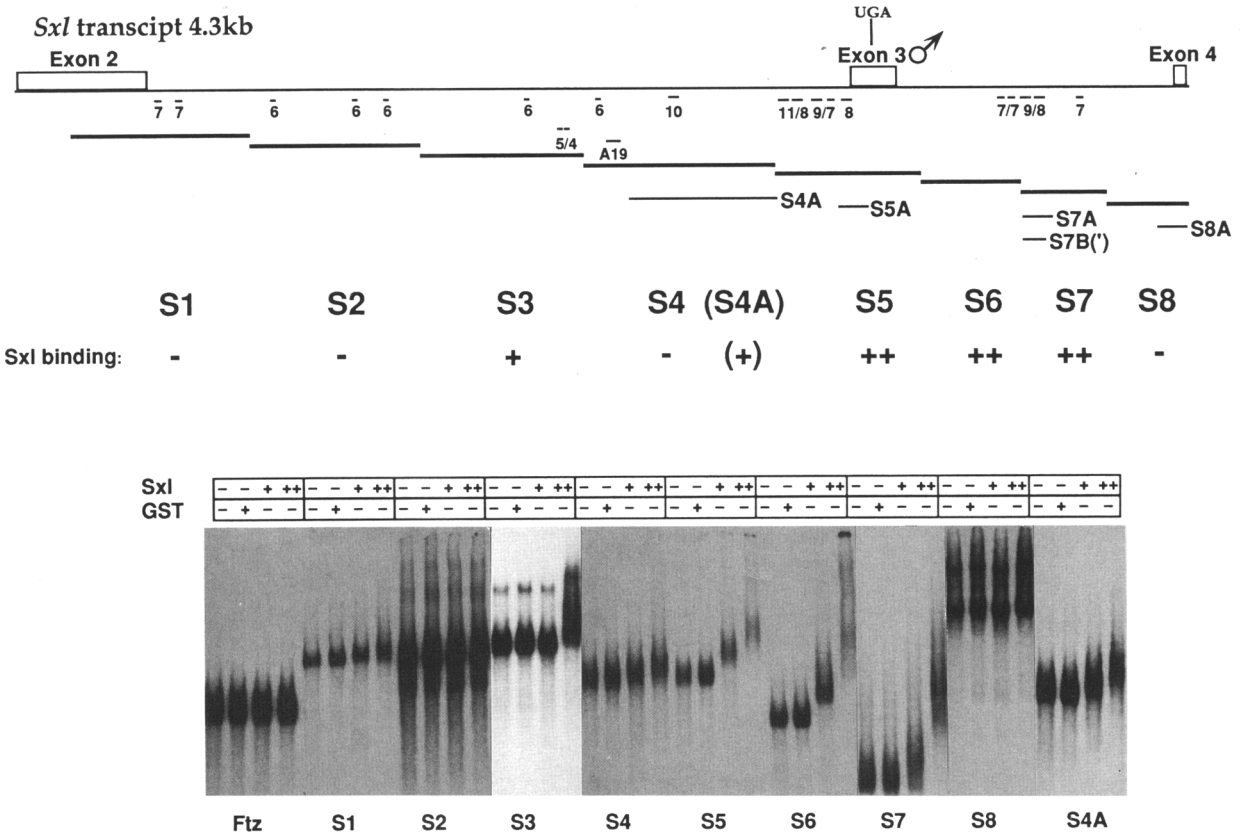


**Figure 2.** Sxl binds to multiple sites surrounding the *Sxl* male-specific exon. A portion of the *Sxl* transcript, 4.3 kb, from exon 2 to exon 4 is shown at the *top*. The RNA-binding substrates in this region are displayed together with all U-runs of at least 6 bases; the 5/4 of S3 is included as an exception discussed in the text. The numbers indicate the length of the U-runs. (A19) The $A_{19}$ sequence on S4A. The complete sequences of adjacent pairs of U-runs are as follows: 5/4 is $U_5CU_4CU$, 11/8 is $U_{11}CU_8$, 9/7 is $U_9GCAUAU_7$, 7/7 is $U_7GAU_7$, 9/8 is $U_9AU_8$. The bandshift assays are shown at the *bottom*. GST was used as a control protein as in Fig. 1. (+) A low level (0.9 μM); (++) a higher level (1.2 μM) of *Sxl* was added. The results are summarized beneath the diagram: (+) weak binding; (++) stronger binding. To give a general picture of the distribution and length of all *Sxl*-binding substrates, transcripts encoded by this region that were used in other sections are also shown.

protein specifically (S3, S4A, S6, S7). Closer inspection revealed a general correlation between positive binding and runs of eight or more U's, but context is clearly also important and eight contiguous U's are not always necessary. S1, S2, and S8 have U-runs of seven or less, and none were bound by Sxl protein. S4A, S5, and S7 have runs of at least 8 U's and were positive in binding. However, the S4 RNA, which is slightly larger than S4A and includes an $A_{19}$ stretch, was unable to form complex with Sxl. Two transcripts, S3 and S6, lack $U_8$ runs but bound Sxl. The first, S3, has only a $U_6$ as the longest run, but subdivision of S3 showed that a $U_5CU_4CU$ sequence, not the $U_6$, was actually responsible for the binding (data not shown). The second, S6, was clearly shifted even though its longest U-runs are two $U_7$'s, but these short U-runs are very close to each other, as part of the sequence $U_2AU_7GAU_7$. The binding of S3 and S6 suggests that U-rich stretches need not contain 8 U's in a row to be sufficient for Sxl binding if they are present in a favorable, still undefined, context.

We observed that when the RNA has multiple U-runs, complexes of different, distinct, migration rates were detected and these correlated with the protein level (Fig. 2, binding with S5, S6, and S7). Because these changes of protein levels were rather small, we speculated that the higher complexity might result from multimerization of Sxl proteins on RNAs with multiple binding sites.

### *Binding of Sxl protein to single and adjacent double U-runs: Evidence for cooperative interaction*

The above results led us to postulate that each run of ~8 U's could be a single Sxl-binding site, and when there are two adjacent U-runs, they might bind two proteins. To achieve the higher resolution necessary, we used S7A, a short transcript that contains a double-U sequence ($U_9AU_8$) from the first 82 bases of S7 (Fig. 3A), and another short transcript that contains a single $U_8$ sequence, S5A, which had been used in Figure 1. As shown in Figure 3A (S7A and S5A), a more slowly migrating protein–RNA complex was formed with double-U RNA than with single-U RNA. Furthermore, increasing Sxl protein

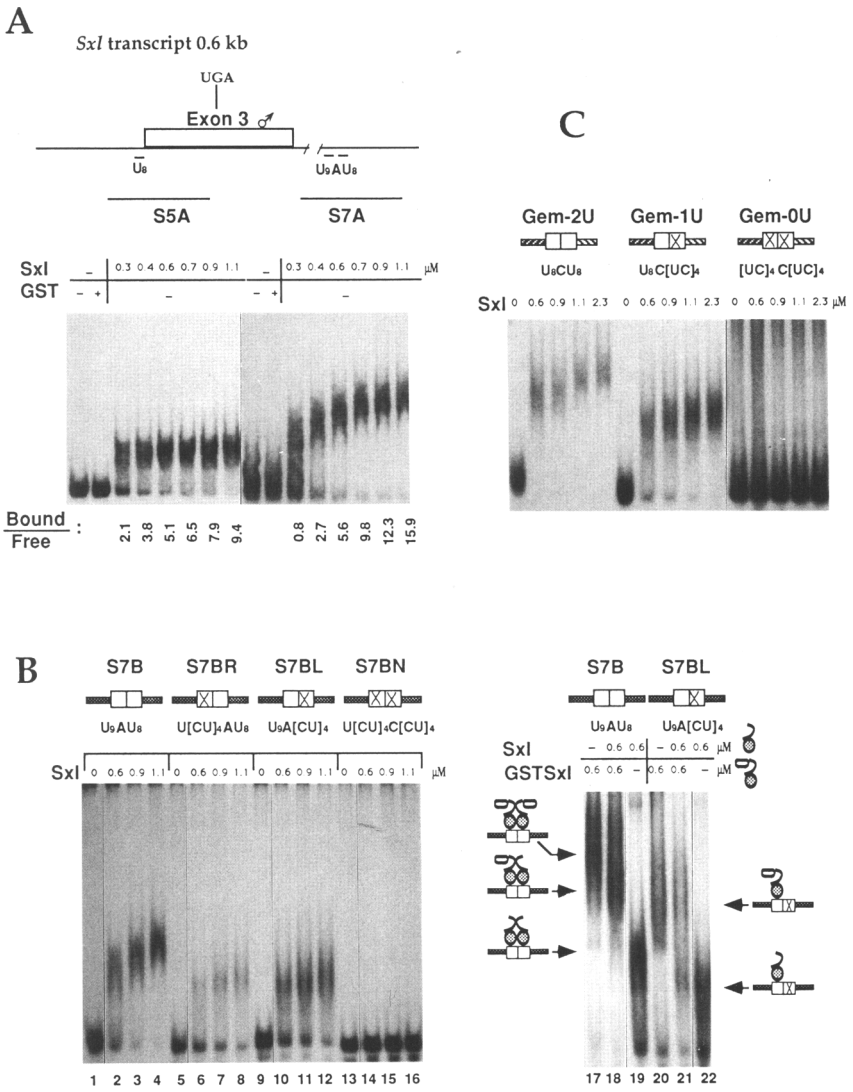


**Figure 3.** Double adjacent U-runs show multiple Sxl binding with cooperativity. (*A*) Comparison of Sxl binding to *Sxl* transcripts with one or two U-runs. The substrate RNAs S5A and S7A, also shown in Fig. 2, are schematically described at the *top* of the corresponding gel panels. The concentrations of Sxl are indicated above the lanes. The fractions of bound vs. free RNA are shown below the lanes. (*B*) The binding with S7B (see Fig. 2) and its derivatives. The U-run related sequences of each substrate are shown above the corresponding lanes. An open box indicates an intact U-run; a boxed X means that the U-run was disrupted by C's. (*Left*) An exemplary band-shift assay with these RNAs. The constants mentioned in the text were the results of binding of Sxl at several additonal concentration points, repeated approximately three times at each concentration. (*Right*) The combination of Sxl and GST–Sxl. The designated different RNA–protein complexes are explained in the text. The concentrations given for Sxl are active concentrations, whereas those of GST–Sxl are only approximate. The aspects of Sxl protein structure diagramed are GST (open oval), Sxl amino terminus (curved line), core of Sxl protein (stippled circle). (*C*) The binding with artificial double-U or single-U substrates, shown as for *B*.

concentration caused a more drastic increase in the bound fraction of double-U RNA (S7A) than single-U RNA (S5A) (Fig. 3A, bound/free ratio). It seems likely that a complex composed of one RNA and multiple proteins ($P_nR$) was formed when two runs of U's were available. The rather smeary appearance and the intermediate migration rate of these complexes with lower amounts of Sxl protein (Fig. 3A, first two lanes of S7A) could represent the disassociation–association state of weak binding during electrophoresis (Lane et al. 1992) and/or the changes in the structures of the RNA molecules or protein–RNA complexes at different binding levels.

Although the binding patterns were clearly different with single-U (S5A) and double-U (S7A) transcripts, the relative contributions of the U-runs versus the flanking non-U sequences could not be determined. At lower levels of Sxl, the binding of single-U RNA S5A is actually greater than that of the double-U RNA S7A (e.g., at 0.3 μM Sxl the bound/free ratio is 2.1 for S5A but 0.8 for S7A; Fig. 3A). Only at the higher levels of Sxl does the binding of double-U S7A equal or surpass that of single-U S5A.

To demonstrate in more controlled assays that the formation of different complexes depends on the number of U-runs, the next experiments used a double-U transcript and synthesized variants that differed only in the number of U-runs. We compared S7B, which includes the first 56 bases of S7A with the two adjacent U-runs, and S7B derivatives in which one or both U-runs were disrupted by C's. As shown in Figure 3B, RNAs containing a single U-run (S7BL or S7BR, lanes 5–12) formed a single shifted band, and it migrated more rapidly than bands formed with the double-U RNA (S7B, lanes 1–4). When both U-runs were destroyed (S7BN, lanes 13–16), no shifted band was observed.

The difference in binding patterns was also evident when we made completely artificial substrates in which two $U_8$, one $U_8$ with one $(UC)_4$, or two $(UC)_4$ sequences, were placed into the context of the pGEM4 polylinker. We again observed the formation of higher order com-

plex with two U-runs compared with the complex with one U-run (Fig. 3C). There was no binding to the two $(UC)_4$ runs. The conclusion could also be made from these results that a stretch of 8 U's is sufficient to be recognized as a binding site for Sxl protein.

We then took advantage of the fact that the Sxl fusion with glutathione *S*-transferase (GST–Sxl) displayed the same binding specificity and similar affinity as unfused Sxl protein, to demonstrate that the higher order complexes result from multimerization of Sxl proteins on RNA. Owing to its larger size, GST–Sxl alone shifted double-U RNA (S7B) to a band that migrated much more slowly than the band with Sxl alone (Fig. 3B, lanes 17,19). When the two proteins were combined, a new band migrating between the GST–Sxl–S7B and Sxl–S7B bands could be observed (Fig. 3B, lane 18). The novel band appears to be of a discrete complex, which is probably composed of one fused and one unfused Sxl protein on the double-U RNA. No such intermediate band was found when the same experiment was performed with single-U RNA (S7BL, Fig. 3B, lanes 20–22; S7BR, not shown), indicating that the formation of the ternary complex depends on the availability of two binding sites on the RNA. This also suggests that Sxl protein exists as monomers in solution and that the multimerization is dependent on RNA binding.

Two independent tests for the proposal that multimerization is dependent on RNA binding were conducted. First, an affinity–coprecipitation experiment was performed in which the GST–Sxl and Sxl proteins, at concentrations higher than those used in regular RNA-binding assays, were mixed in RNA-binding buffers prior to precipitation by glutathione–agarose beads. The GST–Sxl fusion protein was found attached to the beads as expected. If Sxl could multimerize with GST–Sxl in solution, as it does upon binding to double sites on RNA, a coprecipitation of Sxl with GST–Sxl should be observed. However, after precipitation, no unfused Sxl protein was detected by SDS-PAGE (data not shown). As a second test, similar to the method used to determine the dimerization of λ repressors as a prerequisite for DNA binding (Pirrotta et al. 1970), an order of addition experiment in which the order of diluting Sxl and adding RNA substrate was switched showed kinetically that the Sxl multimerization is not required before RNA binding (data not shown; see Materials and methods). We conclude that the majority of Sxl molecules exist as monomers in solution and can multimerize when binding to specific transcripts containing two adjacent binding sites.

Thus far, we have shown that two Sxl monomers bind to two adjacent U-run binding sites. Next, we argue that these two Sxl molecules interact on the RNA in a cooperative fashion. The initial suggestion of cooperativity came from the observation that when Sxl protein bound to RNA with double sites, only the higher order complex, presumably $P_2R$ (two proteins, one RNA), was observed. The smeary appearance of the presumed $P_2R$ complex at low Sxl concentrations was noted above. Lack of a discrete intermediate complex often suggests cooperative binding that results in almost simultaneous filling of all sites. Accordingly, quantitative analysis showed that RNAs with double binding sites showed a steeper response to the increase in protein concentration (Fig. 4). This became more obvious when we calculated the dissociation constants from the bandshift assays exemplified in Figure 3B. The dissociation constants are $K_{1(S7BR)} = 7.7 \pm 1.8 \times 10^{-7}$ M and $K_{1(S7BL)} = 11.1 \pm 4.9 \times 10^{-7}$ M for the binding with the single-U transcripts S7BR and S7BL, respectively; the overall dissociation constant for S7B (double-U) binding, $K_{2(S7B)} = 5.6 \pm 1.2 \times 10^{-14}$ $M^2$, was deduced given $n = 2$ in the equation $K_n = [P]^n[R]/[P_nR]$. The sample standard deviation of $K_{n(S7B)}$ was much higher when $n$ was assigned to any other numbers, in agreement with the previous suggestion that $P_2R$ is probably the form of the $(Sxl)_n$–S7B complex. Because $K_{2(S7B)}$ is ~15-fold less than $K_{1(S7BR)} \cdot K_{1(S7BL)}$, a positive cooperativity of ~15-fold is needed to explain this equilibrium state.

Being aware that the lack of intermediate complex could result from "disproportionation" artifacts (Kleinschmidt et al. 1991), in which there is preferential loss of $PR$ complex compared with $P_2R$ during electrophoresis, and that this could also affect the quantitation of the binding affinity, we performed filter binding experiments with various RNAs containing single or double sites. The results generally correlated very well with those from bandshift assays (Fig. 4). The apparent sig-

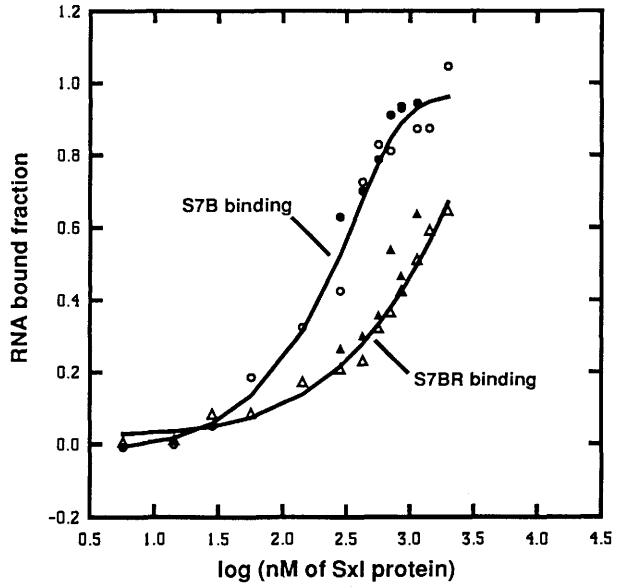


**Figure 4.** Comparison of filter binding and bandshift methods for Sxl RNA binding. Shown are the results of filter-binding assays with RNA substrates S7B and S7BR (see Fig. 3B). Circles indicate data points of binding with S7B (double-U RNA), and triangles with S7BR (single-U RNA); filled symbols show bandshift results; open symbols show filter binding. The curves showing the percentage of bound RNA vs. $\log_{10}$ of Sxl concentration (in nM) were generated by curve fitting of the filter-binding data using the computer program SigmaPlot 4.11 (Jandel Scientific). The bandshift data from Fig. 3B were included in the graph for comparison but not used for curve fitting.

moidal shape of the S7B binding curve supports the conclusion that binding to the double sites is cooperative.

### *The amino terminus of Sxl protein is required for cooperativity*

In a search for the structural elements of Sxl protein that are responsible for cooperativity, we made serial deletion mutants and tested them in binding assays. Initial results showed that the amino-terminal 125 amino acids before the first RNP-CS domain were essentially dispensable for the specificity and affinity of RNA binding. However, the previously observed cooperativity of Sxl protein was greatly reduced by deletions in this region. Here, we present our results with one of the amino-terminal deletion mutants, SxlN1, which lacks the first 38 amino acids of Sxl.

When bound to double-U RNA (S7B′, slightly different from S7B), SxlN1 differed from wild-type Sxl. SxlN1 showed two complex bands and a much slower decrease of unbound RNA relative to protein concentration (Fig. 5). Calculation indicates that the first complex is likely to be *PR* with dissociation constant $K_1 = 1.1 \pm 0.2 \times 10^{-6}$ M, and the second is likely to be $P_2R$ with $K_2 = 0.9 \pm 0.3 \times 10^{-12}$ M$^2$. Because $K_2$ is very close to $(K_1)^2$, the binding cooperativity observed with wild-type Sxl protein is almost completely destroyed by the SxlN1 deletion. When SxlN1 and Sxl proteins were mixed together in binding reactions with double-U S7B, only the complex of Sxl alone was formed (Fig. 5, lanes 15–17). This was true even when there was four times more SxlN1 than Sxl (data not shown). On the other hand, when the RNA had one U-run, two bands representing complexes formed by either SxlN1 or Sxl were readily observable (Fig. 5, lanes 18–20). This is best explained if only intact Sxl protein has the ability to cooperate upon binding to the double U-runs and drives the equilibrium toward the formation of $(Sxl)_2$–RNA complexes but not $(SxlN1)_n$–RNA complexes. When the RNA has only one binding site, the singly bound protein can no longer cooperate, so both Sxl–RNA and SxlN1–RNA complexes have an equal chance to form. This result clearly differs from that of the similar experiment combining Sxl and the GST–Sxl protein described previously (Fig. 3B, lanes 17–20). There, a new intermediate complex was formed on double-U RNA that presumably contained one of each protein, as they both have the cooperative amino-terminal domain.

In conclusion, the RNA-binding cooperativity of Sxl protein is mediated mainly, if not completely, through the first part of the amino terminus. This part of the protein is required on both interacting protein molecules for cooperation. We consider these studies with deletion mutants of Sxl as additional evidence in support of the idea that the wild-type Sxl protein can bind RNA cooperatively.

### *The Sxl amino terminus is required for alternative splicing regulation*

It was of interest to test whether this small amino-terminal region of Sxl protein is of significance for Sxl function in vivo. We chose the method of cotransfection into *Drosophila* tissue culture cells, which has been successfully used in studying *Sxl* regulation of itself and *tra* in vivo (Inoue et al. 1990; Sakamoto et al. 1992), and *tra* and *tra2* regulation of *dsx* (Hedley and Maniatis 1991; Hoshijima et al. 1991; Ryner and Baker 1991). Figure 6 shows the splicing product of a reporter construct that contains the alternatively spliced region of *Sxl* from exons 2 through 5 (~6 kb) fused to a *lacZ* sequence tag (Bell et al. 1991). The transcript of the reporter splices in a male-specific manner in Schneider 2 cells (Fig. 6A, Control), in agreement with the previous report that this cell line is primarily male (Ryner and Baker 1991). When the *Sxl* cDNA was cotransfected with the reporter, splicing was completely switched to the female pattern (Fig. 6A, Sxl). However, in cotransfection with *SxlN1*, only a small portion of the splicing products were of the female pattern (Fig. 6A, SxlN1). Notably, splicing in the *Sxl* cotransfection seemed much more efficient than with *SxlN1* or without any *Sxl*, as evidenced by the ratio of spliced to

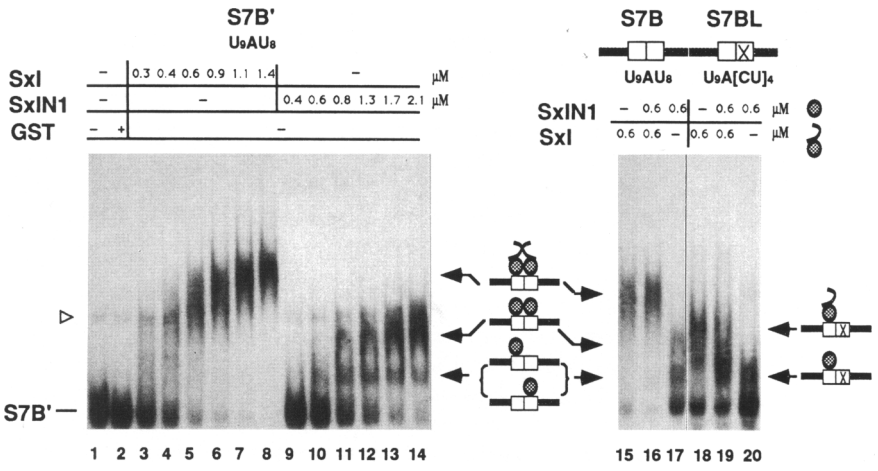


**Figure 5.** Effect of a Sxl amino-terminal deletion on Sxl-binding cooperativity. (*Left*) Binding of Sxl and the mutant SxlN1 with S7B′ RNA (see Fig. 2). The protein concentrations are shown above the lanes. The U-runs of S7B′ are the same as S7B ($U_9AU_8$). The faint band (open arrowhead) is a nonspecific band existing in all lanes, possibly encoded by the transcription vector. (*Right*) The combination of Sxl and SxlN1 in binding with S7B and S7BL. All complexes are drawn as in Fig. 3B.

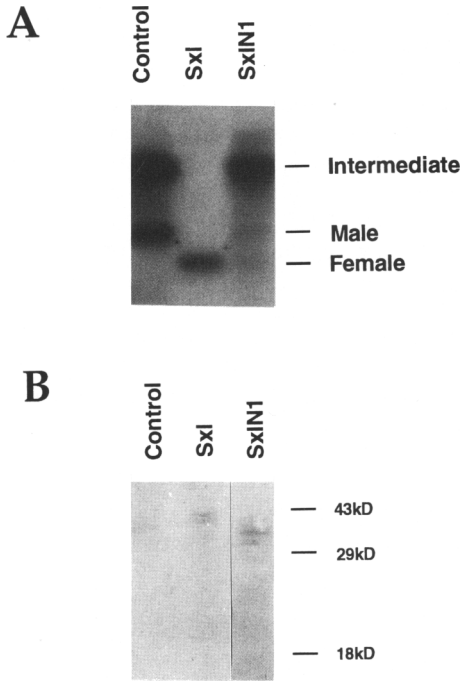


**Figure 6.** The amino-terminal part of Sxl protein is important for Sxl's function in autoregulation. (*A*) Southern blot of the products of reverse transcription and PCR from transfection experiments. DNAs cotransfected into Schneider 2 cells along with the reporter plasmid are shown above each lane. In the control lane, the cDNA transfection vector (hsp83–CaSpeR) was cotransfected with the reporter. Splicing patterns are indicated next to the bands. The intermediate is a partially spliced product containing one intron. (*B*) Western blot of Sxl proteins expressed in transfected cells. SxlN1 is a smaller protein compared with Sxl.

incompletely spliced RNAs. The large band of ~1.1 kb observed in the control and SxlN1 lanes is a partially spliced form that retains a single intron. The total amount of reporter RNA present was much greater for the *Sxl* cotransfection than for the others, and the amount of RNA used in PCR amplification had to be ~100 times less to give comparable Southern blot signals (see Materials and methods). This appears similar to the case in male *Drosophila*, where severalfold less *Sxl* RNA is observed in males than females (Salz et al. 1989). Instability of RNAs containing early nonsense codons might contribute to the low level of male RNA.

The results of a Western blot showed that both intact and mutant Sxl proteins were expressed in the transfected cells (Fig. 6B). The appearance of the doublets of these proteins is similar to that seen in Bopp et al. (1991). None of these bands were recognized by the secondary antibody alone (data not shown). Because this cell line is primarily male, the very low abundance proteins observed in the control lane, which cross-reacted with anti-Sxl antibodies, are likely to be the same as the 33- to 35-kD male proteins described in Bopp et al. (1991). The SxlN1 protein may be less stable than Sxl, because ~100-fold more plasmid was required to achieve equal levels of protein expression. It is also likely that much less wild-type Sxl was required, because it could activate the endogenous copy of *Sxl* in the tissue culture cells, whereas SxlN1 could not.

## Discussion

### *Multiple binding, cooperativity, and implications for exon selection*

We have found that Sxl can bind numerous distant sites surrounding the regulated *Sxl* male-specific exon, in addition to the $U_8$ sequence associated with the alternative 3′ splice site. Moreover, Sxl can interact cooperatively through its amino terminus when binding to adjacent sites. These results suggest that the splicing autoregulation of *Sxl* may be more complicated than the straightforward blockage model proposed for the regulation of *tra*, in which Sxl competes with U2AF at the 3′ splice site region (Sosnowski et al. 1989; Valcárcel et al. 1993).

Our results are consistent with various recent in vivo studies that strongly support the contention that the 3′ splice site region is not sufficient for Sxl regulation of its own pre-mRNA. Sakamoto et al. (1992) have shown that deletions of the various poly(U) sequences between *Sxl* exon 2 and exon 4 severely reduced the splicing regulation of a truncated *Sxl* pre-mRNA by Sxl protein in cultured *Drosophila* cells. These regions correlate with the U-runs discussed in this paper, despite minor discrepancies in sequence probably resulting from polymorphisms in the long monobase regions. These U-rich regions were further shown to UV–cross-link to Sxl protein, and stronger cross-linking was observed with longer U clusters. Protein dimers were postulated to explain an upper band of the cross-linking products seen with the long U-runs. Horabin and Schedl (1993a,b) have used transgenic reporter constructs in male and female *Drosophila* to show that deletions of U-rich regions from the introns surrounding the male-specific exon interfere with sex-specific splicing regulation. Moreover, disruption of the $U_8$ at the male-specific 3′ splice site had no effect on sex-specific splicing regulation.

In agreement with previous studies on the involvement of U-rich regions in Sxl binding, we have found that a stretch of about 8 U bases can be sufficient as a single Sxl-binding site. We have also observed that flanking sequences as well as the transcript length can influence Sxl-binding to both single and cooperative double sites. For example, the overall binding of the single-U RNA S5A seems strong in relation to the double-U RNA S7A, presumably because they derive from two separate regions of the *Sxl* transcript having different flanking sequences (Fig. 3A). Also, the artificial substrate made by inserting a $U_8$ into the pGEM4 vector showed an affinity ~10-fold higher than the more natural *Sxl* S7BL or S7BR substrates, which themselves differ somewhat in binding affinity with Sxl (Fig. 3B,C). The superficial requirement of 8 U bases in a row could also sometimes be relaxed depending on the sequence of the entire tran-

script, as may be the case with the $U_7GAU_7$ sequence in *Sxl* S6 and the $U_5CU_4CU$ in *Sxl* S3 (Fig. 2).

The finding that Sxl can bind cooperatively to multiple sites on its pre-mRNA raises several possibilities as to how the autoregulation of splicing might be achieved. This step is biologically critical because of the position *Sxl* occupies in the sex determination pathway. First, cooperative, multiple binding could provide additive binding strength, which might simply be required in some cases for the specific interaction between Sxl and its target RNA sequences. In addition, cooperativity could also cause an interaction system to respond more steeply to its effector. In this case, the splicing mode switch could be set in a reliable way such that a threshold level of Sxl protein is both necessary and sufficient for the female splicing pattern to be chosen, as naturally occurs in females. Slight fluctuations of Sxl protein level would then not alter the established female mode in females nor mistakenly initiate the female mode in males. This becomes important when we consider both the sensitivity and complexity of the control by the X/A ratio counting elements of early *Sxl* protein production, and the regulated splicing of the large and multiexon *Sxl* transcripts. In addition, *Sxl* is the central on/off switch of fly sex determination and controls a number of downstream pathways. Perturbation of the *Sxl* splicing pattern could cause a deadly switch in the sex determination state. The general strategies used in the initiation and maintenance of fly sex determination have been compared to the lytic–lysogenic control system of phage λ, which employs cooperative binding of repressors to the operators as a key step in an autoregulatory loop (Keyes et al. 1992).

Second, strong binding of Sxl at different places on the pre-mRNA in addition to 3′ splice sites hints that Sxl may antagonize splicing factors other than U2AF in autoregulation, although it has been reported that Sxl does compete with U2AF in the regulation of *tra* splicing (Valcárcel et al. 1993). This might explain the observation of Sakamoto et al. (1992) that the U-run in the polypyrimidine tract immediately upstream of the *Sxl* male exon, where the competitive binding of U2AF should be, is less important than the U-runs farther away. This is also consistent with the observation of Horabin and Schedl (1993a) that the same U-run fails to alter regulated splicing in transgenic flies and that the male-specific 5′ splice site is more important for regulation. In general, selection of each splice site involves a number of factors by means not well understood. Moreover, accumulating data support the notion that the exon rather than the intron may be the recognition unit during splicing (Hoffman and Grabowski 1992; Niwa et al. 1992). Thus, the importance of the U-runs downstream of the *Sxl* male exon could reflect the fact that both the 3′ and 5′ splice sites of the regulated *Sxl* exon may have to be blocked, whereas only a 3′ splice site needs to be excluded in the regulation of *tra* splicing.

A third possibility is that cooperative interactions among Sxl proteins bound at distant locations might loop out a large region that contains the *Sxl* male exon to cause its exclusion. Similarly, cooperative binding of multiple Sxl proteins may facilitate binding to nonspecific RNA sequences between specific sites to "coat" such a region and block out splicing factors. Indirectly supporting this idea, the observed cooperativity and nonspecific binding seemed to be somewhat enhanced when longer RNAs with multiple sites were used in binding assays (J. Wang and L.R. Bell, unpubl.).

### *The Sxl amino terminus and protein interactions*

The above discussion is basically an extension of the model that Sxl binds to specific sites on target pre-mRNA and passively blocks certain splicing sites. However, there seems to be no obvious reason why Sxl could not actively contact other splicing factors. Interactions with other factors might be mediated by either the Sxl amino terminus, the carboxyl terminus, or both. We have demonstrated the requirement of part of the amino terminus for correct female splicing in *Drosophila* tissue culture cells, and we have shown that Sxl monomers interact cooperatively through their amino termini when they bind RNA. It is possible that under some circumstances, the Sxl amino terminus might interact with other splicing proteins instead of another Sxl molecule. In different situations, the Sxl amino terminus might interact with other factors in addition to cooperating in the binding of Sxl, perhaps with a Sxl dimer forming the core of a multiprotein complex. This seems especially likely given the small size of the SxlN1 deletion used in these studies, which removes only a fraction of the entire amino-terminal region.

The entire amino terminus of Sxl is very rich in glycine, serine, asparagine, and proline, with 20%, 16%, 16%, and 8%, respectively, in this 125-amino-acid region (Bell et al. 1988). Clusters of these amino acids would be predicted to cause coils and turns and would result in overall flexibility of this region. This type of glycine-rich protein region is found in several hnRNP components, for instance, proteins A1, B2, L, and *Drosophila* A/B type Hrb87F and Hrb98DE (Cobianchi et al. 1988; Piñol-Roma et al. 1989; Buvoli et al. 1990a; Haynes et al. 1990, 1991). Of particular interest is hnRNP protein A1, whose carboxyl terminus has the highest similarity to the amino terminus of Sxl. The carboxyl terminus of A1 is also rich in glycine, asparagine, and serine, is flexible, and has been shown to be responsible for the cooperative binding of A1 to single-stranded polynucleotides by physical–chemical studies (Kumar et al. 1990; Casas-Finet et al. 1991; Nadler et al. 1991). It is also known that A1 has preferential affinity toward 3′ splice sites, but has highest affinity for a different but related sequence (Swanson and Dreyfuss 1988; Buvoli et al. 1990b; Burd and Dreyfuss 1994). It must be acknowledged that the carboxyl terminus of A1 has intrinsic RNA-binding affinity, possibly involving the periodically distributed aromatic amino acids (Cobianchi et al. 1988; Kumar et al. 1990; Casas-Finet et al. 1991; Nadler et al. 1991), whereas the Sxl amino terminus does not have as many aromatic amino acids and was not found to

influence the basal RNA-binding affinity of Sxl. Mayeda and Krainer (1992) have reported that A1 counteracts splicing factor SF2/ASF to promote the use of distal 5′ splice sites. It is plausible that Sxl could interact with such general splicing factors through their common interactive domains and thereby alter the default process for choosing proximal splice sites to cause the exclusion of alternative exons. Alternatively, Sxl could directly participate in spliceosome formation, for example, by replacing A1, and impede the joining of the regulated exon to common exons.

As a further comparison between Sxl and other splicing factors, both A1 and SF2/ASF have natural isoforms generated by alternative splicing, and tissue specificity has been suggested for some A1 isoforms (Biamonti et al. 1989; Buvoli et al. 1990a; Ge et al. 1991). Addition of 52 amino acids, 27 of which are glycines, to the glycine-rich carboxyl terminus of A1 by alternative splicing may generate the splicing factor B2, which shows significantly higher affinity toward single-stranded polynucleotides (Buvoli et al. 1990a). Similarly, Sxl also has many isoforms that appear to display some spatial and temporal specificities (Bopp et al. 1991; Samuels et al. 1991). Most of the Sxl isoforms are generated by the alternative use of exons encoding the carboxyl terminus. Like the Sxl amino terminus, the SxlcF1 carboxyl terminus is also rich in certain amino acids, such as alanine (14.5% in the region) and proline (13%), that could cause flexibility of the structure.

One model based on our initial studies with various deletions of Sxl is that the carboxy-terminal region could interact with the cooperative amino terminus to stabilize the monomeric protein in solution. Upon binding, the two RNP-CS domains might interact with specific sites on the RNA and, consequently, alter the protein structure so that the two termini are freed to extend for intermolecular interaction with other proteins, such as other Sxl molecules and general and tissue-specific or stage-specific splicing factors. It will be important to determine whether products of other genes known to affect *Sxl* splicing, including *snf* and *fl(2)d* (Granadino et al. 1990; Salz 1992; Albrecht and Salz 1993; Flickinger and Salz 1994), can interact with Sxl or the Sxl–RNA complex. Few synergistic interactions among regulatory splicing factors have been described, but cooperativity is often seen among specific and general transcription factors, even among those bound at distant sites. More RNA-binding proteins involved in splicing regulation, as well as RNA stabilization and localization, should be found to interact in such ways.

## Materials and methods

### *Overexpression and purification of the Sxl proteins in* Escherichia coli

The GST expression system was used to produce Sxl fusion proteins (Smith and Johnson 1988). To generate the GST–SxlcF1 fusion, a *Dde*I site 7 bp upstream of the *Sxl* cDNA cF1 translation start codon was cut, filled with Klenow, and the *Eco*RI at the 3′ end of the cF1 clone (Bell et al. 1988) was digested. This *Dde*I–*Eco*RI fragment was ligated into the *Sma*I–*Eco*RI sites, in-frame, of the pGEX-2T vector (AMRAD) to make plasmid pGTcF1. For making the GST–SxlN1 expression plasmid, pGTN1, the *Bam*HI site of the vector and the *Bsp*MII of cF1 were cut and filled, then ligated to an 8-bp *Xho*I linker d(CCTCGAGG) so the reading frame was maintained.

Fusion proteins were produced in IPTG-induced *E. coli* cells following the procedures of Smith and Johnson (1988) with minor variations. The 5-ml supernatant of cell lysate from a 1-liter culture was passed through a 1.0- to 1.5-ml column composed of glutathione–agarose beads (Sigma), rinsed with about five column volumes of MTPBS-2 buffer (MTPBS buffer of Smith and Johnson 1988, plus 1 mM each of DTT, PMSF, EDTA at pH 8.0) and 0.6 M NaCl. The fusion proteins retained on the column were released with 4.5 mM reduced glutathione (Sigma) in 50 mM Tris-HCl (pH 8.0). About 95% of the proteins appeared to be the induced fusion proteins at this stage. A fraction of the fusion proteins appeared to migrate at ~30 kD after the affinity purification. (The GST part of the fusions is 27 kD.) These most likely resulted from a breakage at the front part of the flexible Sxl amino terminus and would be excluded from the full-size fusion protein fractions in the subsequent steps. The fusion proteins were concentrated by Centricon 30 columns (Amicon), digested by thrombin (Sigma) as described (Smith and Johnson 1988), and run through fresh glutathione–agarose columns to retain the GST part cut off from the fusions. Fractions containing the highest concentrations of Sxl protein, monitored by Bio-Rad assays, were collected and checked by SDS-PAGE. Active concentrations were determined by RNA saturation binding (see below). Proteins were stored in MTPBS-2 at −80°C.

### *Preparation of RNA substrates for binding assays*

The $^{32}$P-labeled RNAs were synthesized by in vitro transcription using either T7, T3, or SP6 RNA polymerase following the manufacturer's protocols (Promega). The DNA templates were made by linearizing genomic subclones (Salz et al. 1989; Samuels et al. 1991) with appropriate restriction enzymes. TCA precipitations were performed to determine the concentrations of radioactively labeled RNAs.

To make transcript *Sxl* S1, the *Pvu*II (6140 on *Sxl*, with 0 as the start site of the late transcript, as defined by Samuels et al. 1991) to *Xho*I (6825) genomic fragment was cloned into pGEM2 (Promega) and transcribed from the SP6 promoter. To make RNA S2, the *Xho*I (6825)–*Bgl*II (8124) fragment (g1 in Salz et al. 1989) inserted in pGEM2 was linearized at *Pvu*II (7500) and transcribed with T7 polymerase. To make transcript S3, the sequence between *Pvu*II (7500) and the *Eco*RI polylinker site outside the *Bgl*II (8124) was isolated, subcloned into pGEM4 (Promega), and transcribed with T7 polymerase. To make transcript S4, plasmid pGEM2-g2 that contains fragment *Bgl*II (8124) to *Pst*I (10139) (g2 in Salz et al. 1989) was linearized at the *Hin*cIII (8865) and transcribed from the SP6 promoter. To make RNA S5, the sequence between *Hin*cIII (8865) and the *Sma*I polylinker site outside the *Bgl*II (8124) of pGEM2-g2 was deleted; the new plasmid was cut at *Nco*I (9431) and transcribed with SP6 polymerase. Similarly, to make RNA S6, the sequence between *Nco*I (9431) and the *Sma*I polylinker site of pGEM2-g2 was deleted; the new plasmid was linearized at *Spe*I (9809), and transcribed with SP6 polymerase. To make transcript S7, the sequence after *Spe*I (9809) was joined to the same *Sma*I site of pGEM2-g2, and transcribed with SP6 polymerase. To make transcript S8, the fragment from *Pst*I (10139) to *Xho*I (11688) cloned in pGEM2 (g3 in Salz et al. 1989) was linearized at *Bsp*MII (10448) and transcribed with T7 RNA polymerase.

Smaller *Sxl* RNAs, with the exception of S7B and its variants, were either transcribed from the above plasmids after linearizing at more 5′ restriction sites or from new plasmids made by deleting part of the above constructs. Template for S4A was made by deleting the sequence in front of *Sty*I (8255) from the template for S4. S5A contains the sequence from *Pvu*II (9278) to *Afl*II (9373) of *Sxl*, cloned into the *Sma*I of pGEM2, and transcribed from the SP6 promoter. S7A covers *Sxl* sequence *Spe*I (9809) to *Afl*II (9892) and has the same region of pGEM2 as S5A. S7B′ has a shorter *Sxl* sequence, *Spe*I (9809) to *Rsa*I (9861), compared with S7A. S8A has the *Sxl* sequence *Eco*RV (10355) to *Bsp*MII (10448) and the same pGEM2 sequence.

To make S7B, both strands of *Sxl* sequence *Spe*I (9809) to *Rsa*I (9861), which encodes S7B′, plus 3 bp downstream and overhanging sequences of *Eco*RI and *Kpn*I, were synthesized, and the annealed dsDNA was ligated into the *Eco*RI–*Kpn*I sites of pGEM4. The construct, named pGEMS7B, was cut at *Sma*I downstream of *Kpn*I and transcribed from the SP6 promoter. The S7B variants were made from similarly constructed plasmids that contained appropriate changes to create U → C mutations of the transcripts. Sequences caa$T_8CT_8$ cg, caa$T_8C(TC)_4$ cg, or caa$(TC)_4(TC)_4$ cg, flanked by *Kpn*I and *Bam*HI overhanging ends, were inserted into pGEM4 to make plasmid pGEM-2U, pGEM-1U, and pGEM-0U, respectively. These were linearized at *Hin*cII of pGEM4 downstream of *Bam*HI and transcribed with SP6 polymerase.

A *Dra*I–*Cla*I clone of *tra* in pBSKS (Stratagene), a gift from M. McKeown (Salk Institute, San Diego, CA), was cut at *Fsp*I between the non-sex-specific and female-specific 3′ splice sites and transcribed from the T7 promoter to produce RNA traN. To make RNA traF, the *Fsp*I–*Xho*I fragment was transcribed, after the sequence 5′ of *Fsp*I was deleted from the above construct, with T7 polymerase.

The *ftz* sequence between *Sal*I and *Bgl*II (Laughon and Scott 1984), which includes the entire intron, was subcloned from the genomic clone into pGEM1 (Promega) and transcribed from the SP6 promoter. Polylinker regions of various cloning vectors were transcribed to produce some control RNAs.

### *In vitro RNA binding*

In vitro-transcribed RNA was resuspended in 10× binding buffer (500 mM Tris-HCl at pH 8.0, 144 mM BME, 1 M NaCl) and yeast tRNA (35 μg), and was then mixed with Sxl or control protein diluted in MTPBS-2 (the protein storage buffer) with or without BSA of the same concentration by weight. The 10-μl reaction mixture was incubated at room temperature with optional mild orbital shaking for ~10′–15′. Three microliters of loading buffer containing 0.075% BPB dye and 50% glycerol was added to the reaction prior to loading onto 4% polyacrylamide gel (80:1 acrylamide to bis-acrylamide) in Tris-glycine (pH 8.3) for reactions with an RNA probe longer than 200 bases or 4.3% gel in 0.5× TBE for smaller RNAs. The gel was run at constant 125 V for ~2.5 hr, depending on the size of the RNA. A 15′ to 30′ prerun at the same voltage was normally performed before loading the samples. The gel was dried and exposed to either X-ray film or PhosphorImager screen.

To calculate the binding constants, free RNA concentration, $[R]$, and complex concentration, $[P_nR]$, were represented by the relative intensity of corresponding bands determined using the PhosphorImager; free protein concentration, $[P]$, was approximated by the known total input protein concentration, as all the reactions were performed with ~100-fold molar excess of protein. Results at each protein concentration point, repeated an average of three times, were used in the dissociation constant equation $K_n = [P]^n[R]/[P_nR]$ with $n$ representing the number of proteins in the complex. The calculated $K_n$ values were subsequently averaged, and the sample standard variations were deduced. As mentioned in the text, there was a gradual increase in the apparent size of the complex formed between Sxl and double-U RNAs, especially at low Sxl concentrations. One possible reason among several was the existence of some $P_1R$ complex that was of low abundance and stability and not separated efficiently from $P_2R$ complex. To avoid its effect in deciding $[P_2R]$ when calculating $K_2$ for binding with double-U RNAs such as S7B, only the data with $[Sxl] \geq K_{1(average)}$ were used. This would keep $[P_1R] < \sim 7\%$ of $[P_2R]$, given $K_2 = \sim 0.07 \times (K_1)^2$. However, with or without this precautionary treatment, the calculated $K_2$ values were nearly the same.

The active concentrations of the Sxl proteins were determined by RNA saturation or titration assays. Protein at a fixed concentration near $K_{d(average)}$, as determined by primitive binding experiments, was used in binding reactions with $^{32}$P-labeled RNA S5A at concentrations spanning 1 out of 100 to 100-fold of $K_{d(average)}$. A curve of bound RNA versus total RNA was generated for each fixed protein amount picked for the titration assay. The RNA was considered in excess when such a curve leveled off and the moles of bound RNA at the plateau equaled the moles of active protein. Protein concentration was calculated, and results from different assays were averaged. Concentrations of GST and GST fusion proteins were determined by comparing Coomassie Brilliant blue-stained bands with the bands of standard proteins or Sxl.

The binding reactions were the same for filter-binding assays, but instead of analysis by electrophoresis, the mixture was passed through a Hoefer 0.45 μm nitrocellulose membrane presoaked in wash buffer (1× binding buffer without yeast tRNA). The filter was washed quickly with 0.5 ml wash buffer and dried before counting. To obtain the relative amount of the total input RNA substrates, aliquots of RNA the same as those used to bind Sxl were directly spotted onto the filter, counted, and adjusted according to the fraction of full-length substrate RNA. The fraction of full-length substrate RNA over total RNA was determined by counting appropriate bands from gels versus total input radioactivity. Binding was repeated two to four times at each protein concentration. The sample standard deviations were in a similar range as for the bandshift assays.

For competition experiments, unlabeled RNAs were synthesized as before but in a larger reaction. Poly(U), poly(UC), and poly(C) were purchased from Sigma. The concentrations were determined by $OD_{260}$ and checked in some cases by parallel reactions that contained a small amount of [$^{32}$P]GTP so that RNA concentrations could be obtained by TCA test. The desired amount of cold RNAs were included as dried RNA (transcribed) or as a small volume in water (purchased) in the regular binding reaction.

For reactions using combinations of different Sxl proteins, proteins were mixed in dilution buffer MTPBS-2 before addition to RNA redissolved in binding buffer.

To perform the order of addition test, the normal binding reaction was increased proportionally to 40 μl in duplicates: one tube with labeled RNA (reaction 1) and one without (reaction 2). Sxl stock protein was diluted into the two reaction mixtures, and the final concentrations were in the range that should bind a portion of the RNA substrates at equilibrium. After 2′ incubation at room temperature, reaction 2 was transferred into a tube with labeled RNA as in reaction 1. In this way, the order of adding RNA and diluting protein was reversed in one reaction compared with the other. The two reactions were further incubated, and complex formation was tested by filter binding at different time intervals. The time courses of both reactions were compared to determine whether the bound fraction of re-

action 2 increased with time while that of reaction 1 did not (Pirrotta et al. 1970).

### *Transfection*

*Sxl* cDNA cF1 (called *Sxl*) and its derivative *SxlN1* were moved from the pGEX-2T vector (see above) into the hsp83–CaSpeR vector (Bell et al. 1991) under the control of a consensus *Drosophila* translation signal (Cavener 1987) and a built-in ATG codon. The *lacZ* reporter construct was the same as in Bell et al. (1991), and consisted of *Sxl* exons 2–5 joined to *lacZ* and placed in hsp83–CaSpeR. Total DNA of ~16 μg was used to transfect Schneider 2 cells at $6\times10^5$ cells/ml using the Sigma $CaCl_2$ Transfection System. The ratio of *Sxl*, *SxlN1*, and vector DNA to reporter DNA was 10:1 for the experiment shown in Figure 6A. Repeats were conducted several times with ratios ranging from 0.1:1 to 15:1. For the experiment shown in Figure 6B, the ratio of *Sxl* to reporter was 0.1:1, other ratios were 10:1, and RNA patterns were identical to those shown in Figure 6A. One-tenth of the transfected cells was used for Western blot analysis using a mouse monoclonal antibody against Sxl protein as in Bell et al. (1991) and Bopp et al. (1991). Half of the total RNA from the rest of the cells were subjected to reverse transcription. For all transfections except with *Sxl*, half of the cDNA was amplified with primers on *Sxl* exon 2 and the *lacZ* sequence attached to exon 5, by 5 PCR cycles of 94°C for 1 min, 58°C for 2 min, 72°C for 2 min, followed by 10 cycles with the annealing temperature at 62°C. Except for the *Sxl* transfection, ~2% of the first PCR reaction was reamplified, using a nested primer on the *lacZ* sequence upstream of the first *lacZ* primer, by 15 cycles of 95°C for 45 sec, 61°C for 2 min, and 72°C for 2 min. Twenty percent of the reaction was then run on a 1.5% agarose gel, blotted to GeneScreen filter, immobilized, and probed with radioactively labeled *Sxl* cDNA. For the *Sxl* cotransfection, 10% of the cDNA was amplified, 0.2% of the first PCR was reamplified, and 8% of the reaction was loaded. Splicing patterns were confirmed by cloning and sequencing the PCR products.

## Acknowledgments

We are grateful to the M. Goodman laboratory for allowing us to use the PhosphorImager. We thank R. Deonier and Y. Luo for helping with the equilibrium studies, and members of the Moses, Pellegrini, Tower, Warrior, and Weber laboratories for encouraging discussion. This work was supported by grants from the American Cancer Society, the National Institutes of Health, and a Beckman Young Investigator Award to L.R.B.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

## References

Albrecht, E.B. and H.K. Salz. 1993. The Drosophila sex determination gene *snf* is utilized for the establishment of the female-specific splicing pattern of *Sex-lethal*. *Genetics* **134:** 801–807.

Baker, B.S. 1989. Sex in flies: The splice of life. *Nature* **340:** 521–524.

Bandziulis, R.J., M.S. Swanson, and G. Dreyfuss. 1989. RNA-binding proteins as developmental regulators. *Genes & Dev.* **3:** 431–437.

Bell, L.R., E.M. Maine, P. Schedl, and T.W. Cline. 1988. *Sex-lethal*, a Drosophila sex determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. *Cell* **55:** 1037–1046.

Bell, L.R., J.I. Horabin, P. Schedl, and W.T. Cline. 1991. Positive autoregulation of *Sex-lethal* by alternative splicing maintains the female determined state in Drosophila. *Cell* **65:** 229–239.

Bernstein, M. and T.W. Cline. 1994. Differential effects of *Sex-lethal* mutations on dosage compensation early in Drosophila development. *Genetics* **136:** 1051–1061.

Biamonti, G., M. Buvoli, M.T. Bassi, C. Morandi, F. Cobianchi, and S. Riva. 1989. Isolation of an active gene encoding human hnRNP protein A1; evidence for alternative splicing. *J. Mol. Biol.* **207:** 491–503.

Boggs, R.T., P. Gregor, S. Idriss, J.M. Belote, and M. McKeown. 1987. Regulation of sexual differentiation in D. melanogaster via alternative splicing of RNA from the transformer gene. *Cell* **50:** 739–747.

Bopp, D., L.R. Bell, T.W. Cline, and P. Schedl. 1991. Developmental distribution of female-specific *Sex-lethal* proteins in *Drosophila melanogaster*. *Genes & Dev.* **5:** 403–415.

Burd, C.G. and G. Dreyfuss. 1994. RNA binding specificity of hnRNP A1: Significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO J.* **13:** 1197–1204.

Buvoli, M., F. Cobianchi, M.G. Bestagno, A. Mangiarotti, M.T. Bassi, G. Biamonti, and S. Riva. 1990a. Alternative splicing in the human gene for the core protein A1 generates another hnRNP protein. *EMBO J.* **9:** 1229–1235.

Buvoli, M., F. Cobianchi, G. Biamonti, and S. Riva. 1990b. Recombinant hnRNP protein A1 and its N-terminal domain show preferential affinity for oligodeoxynucleotides homologous to intron/exon acceptor sites. *Nucleic Acids Res.* **18:** 6595–6600.

Casas-Finet, J.R., R.L. Karpel, A.H. Maki, A. Kumar, and S.H. Wilson. 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.

Cavener, D.R. 1987. Comparison of the consensus sequence flanking translational start sites in Drosophila and vertebrates. *Nucleic Acids Res.* **15:** 1353–1361.

Cline, T.W. 1978. Two closely linked mutations in *Drosophila melanogaster* that are lethal to opposite sexes and interact with daughterless. *Genetics* **90:** 683–693.

———. 1979. A male-specific lethal mutation in *Drosophila melanogaster* that transforms sex. *Dev. Biol.* **72:** 266–275.

———. 1984. Autoregulatory functioning of a Drosophila gene product that establishes and maintains the sexually determined state. *Genetics* **107:** 231–277.

———. 1988. Evidence that *sisterless-a* and *sisterless-b* are two of several discrete "numerator elements" of the X/A sex determination signal in Drosophila that switch *Sxl* between two alternative stable expression states. *Genetics* **119:** 829–862.

Cobianchi, F., R.L. Karpel, K.R. Williams, V. Notario, and S.H. Wilson. 1988. Mammalian heterogeneous nuclear ribonucleoprotein complex protein A1: Large-scale overproduction in Escherichia coli and cooperative binding to single-stranded nucleic acids. *J. Biol. Chem.* **253:** 1063–1071.

Flickinger, T.W. and H.K. Salz. 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.

Ge, H., P. Zuo, and J.L. Manley. 1991. Primary structure of the human splicing factor ASF reveals similarities with Drosophila regulators. *Cell* **66:** 373–382.

Gorman, M., M.I. Kuroda, and B.S. Baker. 1993. Regulation of

the sex-specific binding of the maleless dosage compensation protein to the male X chromosome in Drosophila. *Cell* **72:** 39–49.

Granadino, B., S. Campuzano, and L. Sanchez. 1990. The Drosophila melanogaster *fl(2)d* gene is needed for the female-specific splicing of *Sex-lethal* RNA. *EMBO J.* **9:** 2597–2602.

Green, M.R. 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Annu. Rev. Cell Biol.* **7:** 559–599.

Haynes, S.R., G. Raychaudhuri, and A.L. Beyer. 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, S.R., D. Johnson, G. Raychaudhuri, and A.L. Beyer. 1991. The Drosophila Hrb87F gene encodes a new member of the A and B hnRNP protein group. *Nucleic Acids Res.* **19:** 25–31.

Hedley, M.L. and T. Maniatis. 1991. Sex-specific splicing and polyadenylation of *dsx* pre-mRNA requires a sequence that binds specifically to *tra-2* protein. *Cell* **65:** 579–586.

Hoffman, B.E. and P.J. Grabowski. 1992. U1 snRNP targets an essential splicing factor, U2AF65, to the 3′ splice site by a network of interactions spanning the exon. *Genes & Dev.* **6:** 2554–2568.

Horabin, J.I. and P. Schedl. 1993a. *Sex-lethal* autoregulation requires multiple *cis*-acting elements upstream and downstream of the male exon and appears to depend largely on controlling the use of the male exon 5′ splice site. *Mol. Cell. Biol.* **13:** 7734–7746.

———. 1993b. Regulated splicing of the Drosophila *Sex-lethal* male exon involves a blockage mechanism. *Mol. Cell. Biol.* **13:** 1408–1414.

Hoshijima, K., I. Inoue, I. Higuchi, H. Sakamoto, and Y. Shimura. 1991. Control of *doublesex* alternative splicing by *transformer* and *transformer-2* in *Drosophila*. *Science* **252:** 833–836.

Inoue, K., K. Hoshijima, H. Sakamoto, and Y. Shimura. 1990. Binding of the Drosophila *Sex-lethal* gene product to the alternative splice site of *transformer* primary transcript. *Nature* **344:** 461–463.

Kenan, D.J., C.C. Query, and J.D. Keene. 1991. RNA recognition: Towards identifying determinants of specificity. *Trends Biochem. Sci.* **16:** 214–220.

Keyes, L.N., T.W. Cline, and P. Schedl. 1992. The primary sex determination signal of Drosophila acts at the level of transcription. *Cell* **68:** 933–943.

Kleinschmidt, C., K. Tovar, and W. Hillen. 1991. Computer simulations and experimental studies of gel mobility patterns for weak and strong non-cooperative protein binding to two targets on the same DNA: Application to binding of Tet repressor variants to multiple and single *tet* operator sites. *Nucleic Acids Res.* **19:** 1021–1028.

Kumar, A., J.R. Casas-Finet, C.J. Luneau, R.L. Karpel, B.M. Merrill, K.R. Williams, and S.H. Wilson. 1990. Mammalian heterogeneous nuclear ribonucleoprotein A1: Nucleic acid binding properties of the COOH-terminal domain. *J. Biol. Chem.* **265:** 17094–17100.

Lane, D., P. Prentki, and M. Chandler. 1992. Use of gel retardation to analyze protein-nucleic acid interactions. *Microbiol. Rev.* **56:** 509–528.

Laughon, A. and M.P. Scott. 1984. Sequence of a Drosophila segmentation gene: Protein structure homology with DNA-binding proteins. *Nature* **310:** 25–31.

Lucchesi, J.C. and T. Skripsky. 1981. The link between dosage compensation and sex differentiation in *Drosophila melanogaster*. *Chromosoma* **82:** 217–227.

Mattaj, I.W. 1993. RNA recognition: A family matter? *Cell* **73:** 837–840.

Mayeda, A. and A.R. Krainer. 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell* **68:** 365–375.

McKeown, M. 1992a. Alternative mRNA splicing. *Annu. Rev. Cell Biol.* **8:** 133–155.

———. 1992b. Sex differentiation: The role of alternative splicing. *Curr. Opin. Genet. Dev.* **2:** 299–303.

McKeown, M., J.M. Belote, and R.T. Boggs. 1988. Ectopic expression of the female *transformer* gene product leads to female differentiation of chromosomally male Drosophila. *Cell* **53:** 887–895.

Nadler, S.G., B.M. Merrill, W.J. Roberts, K.M. Keating, M.J. Lisbin, S.F. Barnett, S.H. Wilson, and K.R. Williams. 1991. Interactions of the A1 heterogeneous nuclear ribonucleoprotein 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.

Nagoshi, R.N., M. McKeown, K.C. Burtis, J.M. Belote, and B.S. Baker. 1988. The control of alternative splicing at genes regulating sexual differentiation in Drosophila melanogaster. *Cell* **53:** 229–236.

Niwa, M., C.C. MacDonald, and S.M. Berget. 1992. Are vertebrate exons scanned during splice-site selection? *Nature* **360:** 277–280.

Oliver, B., Y.-J. Kim, and B.S. Baker. 1993. *Sex-lethal*, master and slave: A hierarchy of germ-line sex determination in *Drosophila*. *Development* **119:** 897–908.

Piñol-Roma, S., M.S. Swanson, J.G. Gall, and G. Dreyfuss. 1989. A novel heterogeneous nuclear RNP protein with a unique distribution on the nascent transcripts. *J. Cell Biol.* **109:** 2575–2587.

Pirrotta, V., P. Chadwick, and M. Ptashne. 1970. Active form of two coliphage repressors. *Nature* **227:** 41–44.

Ptashne, M. 1992. *A genetic switch*, 2nd ed., Appendix One, pp. 177–183. Blackwell Scientific Publications, Palo Alto, CA.

Ruskin, F., P.D. Zamore, and M.R. Green. 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell* **52:** 207–219.

Ryner, L.C. and B.S. Baker. 1991. Regulation of *doublesex* pre-mRNA processing occurs by 3′-splice site activation. *Genes & Dev.* **5:** 2071–2085.

Sakamoto, H., K. Inoue, I. Higuchi, Y. Ono, and Y. Shimura. 1992. Control of Drosophila *Sex-lethal* pre-mRNA splicing by its own female-specific product. *Nucleic Acids Res.* **20:** 5533–5540.

Salz, H.K. 1992. The genetic analysis of *snf*: A Drosophila sex determination gene required for activation of *Sex-lethal* in both germline and the soma. *Genetics* **130:** 547–554.

Salz, H.K., E.M. Maine, L.N. Keyes, M.E. Samuels, T.W. Cline, and P. Schedl. 1989. The *Drosophila* female-specific sex-determination gene, *Sex-lethal*, has stage-, tissue-, and sex-specific RNAs suggesting multiple modes of regulation. *Genes & Dev.* **3:** 708–719.

Samuels, M.E., P. Schedl, and T.W. Cline. 1991. The complex set of late transcripts from the *Drosophila* sex determination gene Sex-lethal encodes multiple related polypeptides. *Mol. Cell. Biol.* **11:** 3584–3602.

Schüpbach, T. 1985. Normal female germ cell differentiation requires the female X chromosome to autosome ratio and expression of *Sex-lethal* in *Drosophila melanogaster*. *Genetics* **109:** 529–548.

Smith, C.W.J., J.G. Patton, and B. Nadal-Ginard. 1989. Alterna-

tive splicing in the control of gene expression. *Annu. Rev. Genet.* **23:** 527–577.

Smith, D.B. and K.S. Johnson. 1988. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene* **67:** 31–40.

Sosnowski, B.A., J.M. Belote, and M. McKeown. 1989. Sex-specific alternative splicing of RNA from the *transformer* gene results from sequence-dependent splice site blockage. *Cell* **58:** 449–459.

Steinmann-Zwicky, M., H. Schmid, and R. Nöthiger. 1989. Cell-autonomous and inductive signals can determine the sex of the germ line of Drosophila by regulating the gene *Sxl*. *Cell* **57:** 157–166.

Swanson, M.S. and G. Dreyfuss. 1988. RNA binding specificity of hnRNP proteins: A subset binds to the 3′ end of introns. *EMBO J.* **7:** 3519–3529.

Valcárcel, J., R. Singh, P.D. Zamore, and M.R. Green. 1993. The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of *transformer* pre-mRNA. *Nature* **362:** 171–175.

Zamore, P.D., J.G. Patton, and M.G. Green. 1992. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* **355:** 609–614.

# The Sex-lethal amino terminus mediates cooperative interactions in RNA binding and is essential for splicing regulation.

J Wang and L R Bell

*Genes Dev.* 1994, **8:**
Access the most recent version at doi:10.1101/gad.8.17.2072