MOLECULAR AND CELLULAR BIOLOGY, May 2010, p. 2411–2423 Vol. 30, No. 10
0270-7306/10/$12.00 doi:10.1128/MCB.01508-09

Identification of SNAPc Subunit Domains That Interact with Specific Nucleotide Positions in the U1 and U6 Gene Promoters[▽]

Mun Kyoung Kim,[1,2]†‡ Yoon Soon Kang,[3]† Hsien-Tsung Lai,[1,2]¶ Nermeen H. Barakat,[3]§ Deodato Magante,[3]# and William E. Stumph[1,3]*

Molecular Biology Institute,[1] Department of Biology,[2] and Department of Chemistry and Biochemistry,[3] San Diego State University, San Diego, California 92182-1030

Received 18 November 2009/Returned for modification 15 December 2009/Accepted 25 February 2010

The small nuclear RNA (snRNA)-activating protein complex (SNAPc) is essential for transcription of genes coding for the snRNAs (U1, U2, etc.). In *Drosophila melanogaster*, the heterotrimeric DmSNAPc recognizes a 21-bp DNA sequence, the proximal sequence element A (PSEA), located approximately 40 to 60 bp upstream of the transcription start site. Upon binding the PSEA, DmSNAPc establishes RNA polymerase II preinitiation complexes on U1 to U5 promoters but RNA polymerase III preinitiation complexes on U6 promoters. Minor differences in nucleotide sequence of the U1 and U6 PSEAs determine RNA polymerase specificity; moreover, DmSNAPc adopts different conformations on these different PSEAs. We have proposed that such conformational differences in DmSNAPc play a key role in determining the different polymerase specificities of the U1 and U6 promoters. To better understand the structure of DmSNAPc-PSEA complexes, we have developed a novel protocol that combines site-specific protein-DNA photo-cross-linking with site-specific chemical cleavage of the protein. This protocol has allowed us to map regions within each of the three DmSNAPc subunits that contact specific nucleotide positions within the U1 and U6 PSEAs. These data help to establish the orientation of each DmSNAPc subunit on the DNA and have revealed cases in which different domains of the subunits differentially contact the U1 versus U6 PSEAs.

The *Drosophila melanogaster* small nuclear RNA (snRNA)-activating protein complex (DmSNAPc) is a heterotrimeric transcription factor (21) that is required for the synthesis of the U1, U2, U4, U5, and U6 spliceosomal snRNAs (2, 22, 31). Homologous protein complexes are required for snRNA gene expression in humans (1, 11, 12, 18, 33–35) and for spliced leader RNA synthesis in trypanosomes (6, 7, 30). This indicates that SNAPc appeared early in eukaryotic evolution and continues in contemporary times to be utilized for the transcription of important noncoding RNA molecules in diverse organisms. DmSNAPc binds sequence-specifically to an essential, conserved ~21-bp promoter element termed the proximal sequence element A (PSEA) that is located approximately 40 to 60 bp upstream of the transcription start site of fly snRNA genes (8, 13, 19, 23).

In animals, the U1 to U5 snRNA genes are transcribed by RNA polymerase II (Pol II), but U6 snRNA genes are transcribed by RNA Pol III (4, 5, 10, 14, 19, 24, 29, 31). Surprisingly, the primary determinant of the RNA polymerase specificity of the *D. melanogaster* snRNA genes is the precise sequence of the PSEA. A few conserved nucleotide differences in the 3′ half of the 21-bp PSEA are sufficient to determine the polymerase specificity of the fly snRNA genes *in vitro* and to restrict the polymerase specificity *in vivo* (2, 19, 20, 26).

The three subunits of DmSNAPc are termed DmSNAP190, DmSNAP50, and DmSNAP43 so that the names correspond to the most widely used nomenclature for the three orthologous human SNAPc subunits, SNAP190, SNAP50, and SNAP43. These three human subunits are also known as PTFα, PTFβ, and PTFγ, respectively (1, 11, 12, 33–35). Although all three of these subunits of human and fly SNAPc are required for sequence-specific binding of the complex to DNA, only one of the subunits (SNAP190) contains a canonical DNA-binding domain (33).

In earlier studies, we used site-specific protein-DNA photo-cross-linking to map the nucleotide positions within the U1 and U6 PSEAs that are contacted by each of the three subunits of DmSNAPc (32). Those results revealed that the DmSNAP190 subunit could be cross-linked to nucleotide positions extending over the entire length of the PSEA. In contrast, DmSNAP50 and DmSNAP43 cross-linked to nucleotides in the 3′ half of the 21-bp PSEA but not to nucleotides in the 5′ half of the PSEA.

Interestingly, the cross-linking patterns of the subunits, particularly of DmSNAP50 and DmSNAP43, were found to be significantly different depending upon whether DmSNAPc was bound to a U1 versus a U6 PSEA (32). Most strikingly, a later study (22) revealed that DmSNAP43 cross-linked to DNA extending 20 bp downstream of the U1 PSEA, but it cross-linked to DNA only 5 bp downstream of a U6 PSEA. Based

* Corresponding author. Mailing address: Department of Chemistry and Biochemistry, San Diego State University, 5500 Campanile Dr., San Diego, CA 92182-1030. Phone: (619) 594-5575. Fax: (619) 594-4634. E-mail: wstumph@sciences.sdsu.edu.

‡ Present address: Program in Molecular Biology, 15 North 2030 East, EIHG Bldg. 533, Rm. 1400, University of Utah, Salt Lake City, UT 84112.

¶ Present address: 5323 Harry Hines Blvd. North Campus, ND2.114, Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center at Dallas, Dallas, TX 75390.

§ Present address: Division of Hematology/Oncology, Department of Medicine, Case Western University, 10900 Euclid Ave., Cleveland, OH 44106.

Present address: Diazyme-GA, P.O. Box 85608, San Diego, CA 92186-5608.

† These two authors have contributed equally to this work.

▽ Published ahead of print on 8 March 2010.

upon the cross-linking data as well as functional data relating to transcription, we have proposed a model in which the DNA sequence of the PSEA acts as a differential allosteric effector of DmSNAPc conformation that in turn leads to RNA polymerase specificity (2, 13, 20, 22, 32).

At this time, there is little or no knowledge about the three-dimensional structure of SNAPc or its subunits and how they bind to DNA. Thus, to better understand the mode of SNAPc binding to DNA, we have developed a novel protocol that combines the photo-cross-linking technique with subsequent site-specific chemical cleavage of the protein. This has allowed us to localize domains or regions of each of the three DmSNAP subunits that interact with specific individual nucleotide positions in the U1 and U6 PSEAs. Furthermore, novel regions of DmSNAP50 and of DmSNAP43 that contact the DNA have been mapped, and our results reveal for the first time the orientation of each subunit on the DNA. Finally, the data reveal that distinct domains of DmSNAP50 and of DmSNAP43 interact closely with the DNA depending upon whether DmSNAPc is bound to a U1 versus a U6 PSEA.

MATERIALS AND METHODS

DmSNAPc constructs, expression, and purification. The preparation of untagged and N- and C-terminally FLAG-tagged constructs encoding each of the three DmSNAPc subunits has been recently described (16). Constructs with point mutations to eliminate or introduce hydroxylamine cleavage sites (asparaginyl-glycyl peptide bonds) were prepared by using Stratagene's QuikChange II site-directed mutagenesis kit. To minimize the possibility of harmful changes, conservative mutations (e.g., N to Q) were chosen for eliminating NG peptide bonds. In addition, new NG cleavage sites were introduced at residues that were not conserved in the orthologous proteins of other *Drosophila* species based upon genomic sequencing information available from Flybase (http://flybase.org/; data not shown).

Various forms of the three subunits of DmSNAPc were coexpressed in stably transfected S2 cells by using Invitrogen's *Drosophila* expression system as previously described (16, 21, 22). For mapping domains within DmSNAP190, N- or C-terminally tagged DmSNAP190 was coexpressed with untagged DmSNAP50 and DmSNAP43, whereas untagged DmSNAP190 was coexpressed with untagged DmSNAP50 and N-terminally tagged DmSNAP43 (to allow FLAG purification). For mapping DmSNAP50, N- or C-terminally tagged DmSNAP50 was coexpressed with untagged DmSNAP190 and DmSNAP43, whereas untagged DmSNAP50 was coexpressed with N-terminally tagged DmSNAP190 and untagged DmSNAP43. To map DmSNAP43, N- or C-terminally tagged DmSNAP43 was coexpressed with untagged DmSNAP190 and untagged DmSNAP50, whereas untagged DmSNAP43 was coexpressed with N-terminally tagged DmSNAP190 and untagged DmSNAP50. The DmSNAP50 construct used in the DmSNAP43 mapping experiments contained three point mutations that eliminated all NG sites within the coexpressed DmSNAP50. This was done to prevent cross-linked DmSNAP50 cleavage products from interfering with identification of the DmSNAP43 digestion products.

Subunit expression was induced with copper sulfate, and DmSNAPc was purified by FLAG immunoaffinity chromatography as recently described (16). Electrophoretic mobility shift assays were carried out to confirm the DNA-binding activity of the FLAG-purified DmSNAPc variants (16). Although the results were not quantitative, no significant differences in the DNA-binding activity of the point mutants or of the tagged variants were apparent compared to that of wild-type DmSNAPc (data not shown). The samples were dialyzed against 25 mM HEPES K^+ (pH 7.6), 12.5 mM $MgCl_2$, 0.01 mM $ZnCl_2$, 0.1 mM EDTA, 100 mM KCl, 1 mM dithiothreitol (DTT), and 10% (by volume) glycerol to remove 3× FLAG peptide in preparation for the photo-cross-linking reactions.

Site-specific protein-DNA photo-cross-linking and hydroxylamine digestion. (i) Preparation of double-stranded site-specific DNA photo-cross-linking probes. Double-stranded DNA probes that contained a photo-cross-linking agent (azidophenacyl group) attached at specific individual phosphate positions within or downstream of the U1 or U6 PSEA sequence were prepared as previously described in detail (32), with one major exception. In previous works (20, 22, 32), two shorter oligonucleotides (usually 20 to 30 nucleotides long) were utilized. One contained an azidophenacyl group attached to a phosphorothioate located 5′ of the third nucleotide from the 5′ end. The other served as an upstream primer. These were annealed to a longer oligonucleotide (about 70 nucleotides in length) and extended with T4 DNA polymerase to form the double-stranded probe. In the current work, longer versions of the "short" oligonucleotides were synthesized so that the double-stranded probe could be formed without the DNA polymerase step. As an example, to make a DNA probe with an azidophenacyl group at phosphate position 12 of the template strand of the U1 PSEA, the 67-mer of the nontemplate strand, 5′-ACGAATTCATTCTTATAATTCCCAACTGGTTTTAGCGGTACCGCCATGGAAAGGTATGGGATCCTCA-3′ (underlined nucleotides corresponding to the U1 PSEA), was annealed to the following oligonucleotides of the template strand: 3′-TGCTTAAGTAAGAATATTAAGGGTTxGA*-5′ (where x indicates the position of a phosphorothioate derivatized with azidophenacyl bromide and the asterisk indicates a ^{32}P) and 3′-CCAAAATCGCCATGGCGGTACCTTTCCATACCCTAGGAGT-5′. After annealing, the nick was sealed by using T4 DNA ligase, and the double-stranded probe was purified by gel electrophoresis as previously reported (32).

Three additional long oligonucleotides were used for probe preparation depending upon the strand chosen to contain the cross-linker or the U6 versus the U1 PSEA. These were the U1 PSEA template strand 67-mer, 5′-TGAGGATCCCATACCTTTCCATGGCGGTACCGCTAAAACCAGTTGGGAATTATAAGAATGAATTCGT-3′; U6 PSEA nontemplate strand 67-mer, 5′-ACGAATTCATTCTTATAATTCTCAACTGCTCTTTCCGGTACCGCCATGGAAAGGTATGGGATCCTCA-3′; and U6 PSEA template strand 67-mer, 5′-TGAGGATCCCATACCTTTCCATGGCGGTACCGGAAAGAGCAGTTGAGAATTATAAGAATGAATTCGT-3′. The position of the cross-linking agent in the probe was determined by appropriately increasing or decreasing the lengths of the two shorter oligonucleotides used in the annealing reaction so that a fully double-stranded probe would be generated. In each case, the phosphate positions containing cross-linker and ^{32}P radiolabel were separated by a single unmodified phosphate. All oligonucleotides were purchased from Integrated DNA Technologies.

(ii) Photo-cross-linking of DmSNAPc to the DNA probes. Photo-cross-linking reactions were carried out as previously described (32) but were scaled up 10-fold. Briefly, 40 μl of FLAG-purified DmSNAPc was incubated with 1,000,000 cpm DNA probe in the dark for 30 min in a volume of 80 μl in a final buffer composition of 12.5 mM HEPES (pH 7.6), 50 mM KCl, 6.25 mM $MgCl_2$, 0.05 mM EDTA, and 5% (by volume) glycerol. The samples were then irradiated with 313 nm UV light for 333 s. Following irradiation, cross-linked samples were subjected to DNase I and S1 nuclease digestion to remove all but 2 or 3 nucleotides of the DNA probe. The success of the cross-linking reaction was verified by analyzing 1/10 of each sample on denaturing polyacrylamide gels as previously described (32).

(iii) Chemical digestion of cross-linked protein and gel electrophoresis. The remaining 9/10 of each reaction (volume, ~160 μl) was placed into dialysis tubing (Spectra/Por 1, 6,000 to 8,000 molecular weight cutoff; catalog no. 132 645) and dialyzed at 45°C against 400 ml of 1.8 M hydroxylamine (Sigma catalog no. 159417) solution (pH 9.0) freshly prepared as described in *Current protocols in protein science* (3). After 5 h of dialysis to allow chemical cleavage, the samples were transferred and dialyzed against 6 mM Tris-HCl (pH 6.8), 0.25 M urea, 0.2% sodium dodecyl sulfate, and 0.5% β-mercaptoethanol for 4 h at 25°C to remove the hydroxylamine salts in preparation for gel electrophoresis. The samples were then evaporated to dryness in a Savant SpeedVac concentrator and then redissolved in an amount of water that was 1/10 the volume prior to the drying step. Then a solution of 50% glycerol, 25% β-mercaptoethanol, and 0.125% bromophenol blue was added to produce final concentrations of 10% glycerol, 5% β-mercaptoethanol, and 0.025% bromophenol blue. Samples were then run on 11 to 15% denaturing polyacrylamide gels. Following electrophoresis, gels were dried and subjected to autoradiography.

Immunoblots of chemically digested DmSNAPc. FLAG affinity-purified DmSNAPc samples (200 μl of each) were treated exactly as described above for digestion with hydroxylamine, except that the incubation with DNA photo-cross-linking probe and UV irradiation was omitted. Immunoblotting was carried out as recently described (16). N-terminal FLAG-tagged fragments of the DmSNAP subunits were detected by using alkaline phosphatase-conjugated monoclonal antibody against the FLAG epitope (Sigma catalog no. A9469). C-terminal fragments of DmSNAP190 were detected by using a polyclonal rabbit antibody generated against a 14-amino-acid peptide from the C terminus of DmSNAP190 (22) as the primary antibody. Antibodies that were generated by injecting rabbits with full-length bacterially expressed DmSNAP50 or DmSNAP43 (2) were uti-

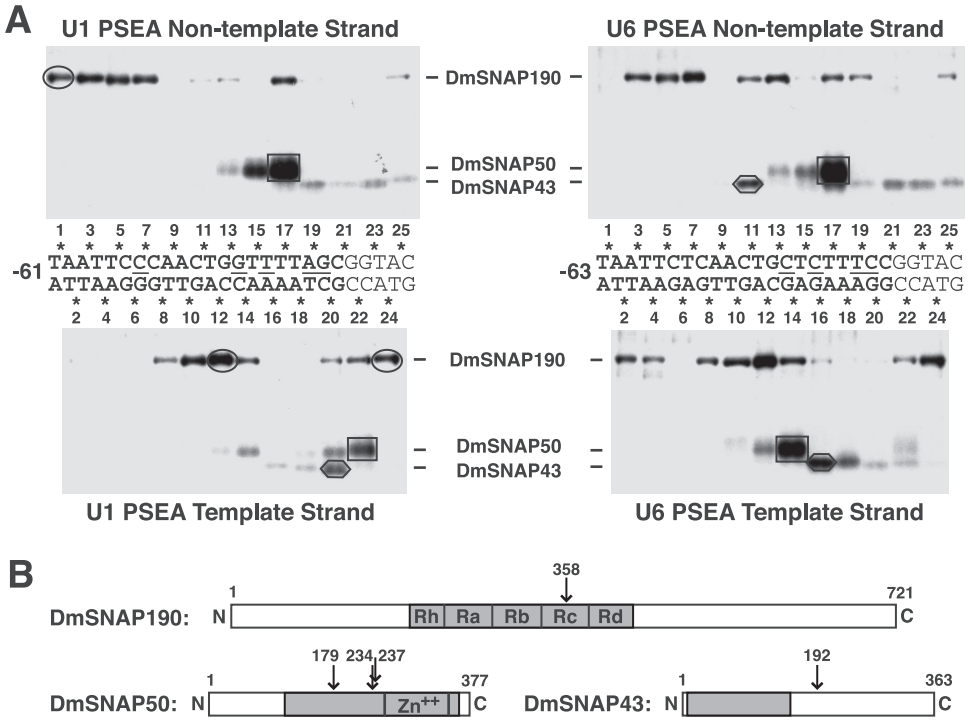

FIG. 1. (A) Site-specific protein-DNA photo-cross-linking data taken from Wang and Stumph (32). Cross-linking positions utilized in the current study are indicated by ovals, squares, and hexagons for DmSNAP190, DmSNAP50, and DmSNAP43, respectively. The sequences of the PSEAs are from the *D. melanogaster* U1:95Ca gene (left) and U6:96Ab gene (right) and differ at the five positions indicated by underlining (13). The 21-bp PSEA sequences (in bold type) extend from position −61 to −41 (left to right) relative to the U1 transcription start site and from position −63 to −43 (left to right) relative to the U6 start site. (B) Structural features of DmSNAP190, DmSNAP50, and DmSNAP43 which consist of 721, 377, and 363 amino acid residues, respectively. The shaded areas indicate the regions that are greater than 26% identical and 42% similar to the human SNAP orthologs. The illustration also shows that SNAP190 contains 4.5 Myb repeats designated Rh, Ra, Rb, Rc, and Rd. SNAP50 contains a unique zinc binding domain as indicated. Downward arrows indicate the locations of NG peptide bonds in the wild-type proteins.

lized to detect cleavage fragments from those two subunits. The population of anti-DmSNAP50 polyclonal antibodies readily detected both N- and C-terminal fragments when DmSNAP50 was cleaved at position 234 but detected the N-terminal fragment very weakly when cleavage was at position 179. The N-terminal cleavage fragment that contained only 102 amino acid residues could not be detected by these antibodies. The population of DmSNAP43 polyclonal antibodies strongly detected the C-terminal fragments when cleavage of DmSNAP43 occurred at either position 192 or 272. The N-terminal fragments in each case yielded very weak signals, but the corresponding bands were clearly present with longer development times. In all cases, the secondary antibody was alkaline phosphatase-conjugated goat anti-rabbit IgG(Fc) (Promega catalog no. S3731).

RESULTS

Figure 1A shows previously published photo-cross-linking data that initially defined the interaction of the three DmSNAPc subunits with specific nucleotide positions within the U1 and U6 PSEAs (32). For those experiments, photo-cross-linking reagent (together with an adjacent ^{32}P radiolabel) was incorporated into the phosphate backbone individually at every other nucleotide position (indicated by the asterisks above and below the DNA sequences in the middle of Fig. 1A) in the nontemplate (upper) strand or in the template (lower) strand of the DNA. This was done for both the U1 (left) and U6 (right) PSEAs (21-bp sequences shown in bold type). Following incubation of the DNA probe with DmSNAPc, subunits of DmSNAPc that cross-linked at each position were identified following UV irradiation, nuclease digestion, denaturing polyacrylamide gel electrophoresis, and autoradiography. Phosphate positions that were chosen for the protein-mapping studies in the current report are enclosed within ovals, rectangles, or hexagons for DmSNAP190, DmSNAP50, and DmSNAP43, respectively. Although not shown in Fig. 1A, for the current studies, we also employed positions 28 and 40 of the template strand downstream of the U1 PSEA, because positions 28 and 40 cross-linked exclusively to DmSNAP43 when DmSNAPc was bound to a U1 (but not U6) PSEA (22).

To localize domains within the individual DmSNAPc subunits that cross-linked to the chosen phosphate positions, we carried out site-specific protein-DNA photo-cross-linking and followed this with cleavage of the proteins with hydroxylamine reagent. This allowed the specific radiolabeled fragment(s) of the hydroxylamine-digested protein that cross-linked to the DNA to be identified by relative size on a denaturing polyacrylamide gel.

Hydroxylamine cleaves polypeptides preferentially at NG peptide bonds (3). The positions of NG peptide bonds in the wild-type DmSNAP subunits are shown in Fig. 1B. DmSNAP190 and DmSNAP43 each contain a single NG bond, whereas DmSNAP50 contains three NG bonds, two of which are located within three residues of each other. Site-directed mutagenesis was used to eliminate NG bonds and/or to introduce NG bonds at new positions in order to provide novel hydroxylamine cutting sites. This allowed us to increase the resolution of the mapping

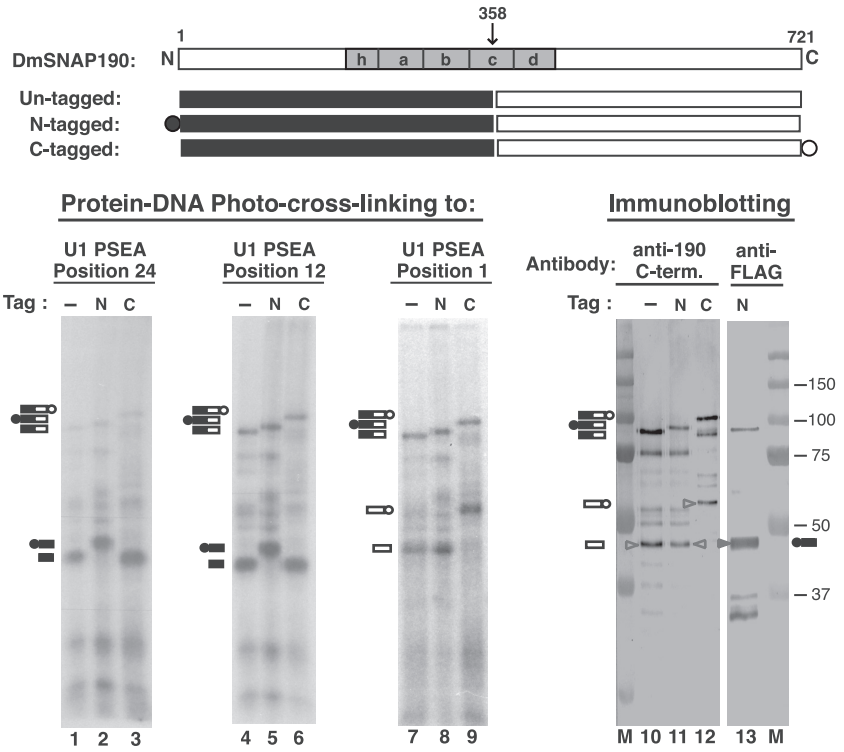

FIG. 2. U1 PSEA positions 24 and 12 cross-link to the N-terminal half of DmSNAP190, but position 1 cross-links to the C-terminal half of DmSNAP190. The upper line in the figure represents wild-type DmSNAP190 and the location of the 4.5 Myb repeats. Below this, the two halves of DmSNAP190 (darkly shaded and unshaded) that result from hydroxylamine cleavage following amino acid residue 358 in the untagged, N-tagged, and C-tagged DmSNAP190 constructs are shown. The circles at the N terminus or at the C terminus represent the tags on the DmSNAP190 constructs. The lower part of the figure shows autoradiograms of denaturing gels of the cross-linked protein fragments after hydroxylamine digestion when cross-linking was carried out to U1 PSEA position 24 (lanes 1 to 3), position 12 (lanes 4 to 6), or position 1 (lanes 7 to 9). The symbols alongside the gel indicate the origin of each band and correspond to the symbols in the upper part of the figure. Lanes 10 to 13 show immunoblots of cleaved protein fragments detected by using antibodies against a C-terminal peptide of DmSNAP190 or against the FLAG epitope. Lane M contained molecular mass markers (in kDa). Arrowheads (unshaded for C-terminal and shaded for N-terminal fragments) indicate the specific cleavage products of interest. Bands corresponding to residual undigested full-length DmSNAP190 (untagged, N-tagged, and C-tagged) are observed in the upper part of all panels.

experiments within the primary amino acid sequence of the proteins. For reasons indicated below, most DmSNAP constructs were expressed as untagged, N-terminally FLAG-tagged, and C-terminally FLAG-tagged variants (16). *In vivo*-assembled DmSNAPc was prepared by co-overexpressing tagged or untagged wild-type or mutant variants of all three subunits in stably transfected *D. melanogaster* S2 cells and purified by FLAG affinity chromatography (16, 22).

Mapping regions of DmSNAP190 that contact nucleotide positions 24, 12, and 1 of the U1 PSEA. As shown in Fig. 1A, DmSNAP190 could be cross-linked to nucleotides stretching over the entire length of the U1 PSEA. We wished to localize the regions of DmSNAP190 that cross-linked to the 3′ end, to the middle, and to the 5′ end of the U1 PSEA. For this, we selected positions 24, 12, and 1, respectively, because DmSNAP190 is the only subunit of DmSNAPc that cross-linked to each of these three positions (Fig. 1A). DmSNAP190 contains 4.5 Myb repeats termed Rh, Ra, Rb, Rc, and Rd (Fig. 1B), and this Myb domain is the region that is most evolutionarily conserved relative to human SNAP190 (22). The Myb domain of human SNAP190, in particular the Rc and Rd repeats, has been implicated in DNA binding by human SNAPc (15, 25, 28, 33). The orientation of the 4.5 Myb repeats on the PSEA, however, is not known.

The NG site at position 358 in wild-type DmSNAP190 is located near the middle of the Rc Myb repeat; this is also very near the middle of the entire 721-amino-acid polypeptide chain (Fig. 1B). Thus, cleavage with hydroxylamine should yield two fragments of nearly equal size that would be indistinguishable on denaturing gels. To positively identify whether the N-terminal half or the C-terminal half of DmSNAP190 cross-linked to each of the PSEA positions 24, 12, or 1, we expressed three different versions of DmSNAP190: untagged, N-terminally tagged, and C-terminally tagged (Fig. 2, upper section). The N-terminal tag increased the size of the N-terminal fragment by 2.2 kDa, and the C-terminal tag increased the size of the C-terminal fragment by 5.7 kDa.

Figure 2 shows results from the cross-linking of DmSNAPc to positions 24, 12, and 1 of the U1 PSEA. When cross-linked to position 24, DmSNAP190 tagged at the N terminus (lane 2) produced a cross-linked radiolabeled fragment with a decreased mobility relative to the cross-linked fragments obtained from untagged or C-terminally tagged DmSNAP190 (lanes 1 and 3, respectively). From these results, we conclude

that position 24 of the U1 PSEA cross-linked to the N-terminal fragment of DmSNAP190 when the protein was cleaved at position 358 within the polypeptide chain. The same pattern was obtained when the cross-linking agent was placed at position 12 of the U1 PSEA (Fig. 2, lanes 4 to 6).

However, a substantially different result was obtained when photo-cross-linking was carried out to position 1 of the U1 PSEA. In this case (Fig. 2, lanes 7 to 9), the fragments from the untagged and N-terminally tagged DmSNAP190 migrated to the same position (lanes 7 and 8), but the fragment obtained from the C-terminally tagged DmSNAP190 migrated with a slower mobility. This indicated that nucleotide position 1, at the 5′ end of the U1 PSEA, cross-linked to the C-terminal half of DmSNAP190 cleaved at position 358.

To further confirm the identities of the bands on the autoradiograms, immunoblots were carried out on the hydroxylamine-digested proteins. Lanes 10 to 12 of Fig. 2 show that antibodies directed against a peptide at the C terminus of DmSNAP190 produced the same pattern of bands observed in the cross-linking experiments in lanes 7 to 9, consistent with the evidence that position 1 cross-linked to the C-terminal fragment of DmSNAP190. Furthermore, the FLAG antibodies detected a band (lane 13) that migrated similarly to the cross-linked amino-terminally tagged bands observed in lanes 2 and 5. (This band also comigrated with the untagged C-terminal fragment.)

From the data in Fig. 2, we conclude that residues located within the N-terminal half of DmSNAP190 cross-linked to positions 24 and 12 of the U1 PSEA. On the other hand, residues located within the C-terminal half of DmSNAP190 cross-linked to U1 PSEA position 1.

To map the locations of the cross-linked residues of DmSNAP190 with greater precision (i.e., with a resolution of approximately one Myb repeat), additional DmSNAP190 constructs that each contained a single hydroxylamine cleavage site near the center of either the Ra, Rb, or Rd repeat, or just N-terminal or C-terminal to the Myb domain, were prepared (Fig. 3A). In each of the mutant constructs, the normal cleavage site in the wild-type sequence at position 358 was eliminated by changing asparagine 358 to glutamine. Schematic diagrams of these constructs are shown in Fig. 3B. Each of these constructs was prepared with the FLAG tag present at the N terminus. Cleavage at the single NG site in each construct H through E was expected to yield a pattern of C-terminal fragments of decreasing size and N-terminal fragments of increasing size. Following hydroxylamine cleavage of the protein, the C- and N-terminal fragments of DmSNAP190 were detected on immunoblots by using antibodies specific for the C terminus or for the FLAG-tagged N terminus (Fig. 3C, lanes 1 to 12). Arrowheads point out the position of the relevant band in each lane.

Figure 3C (lanes 13 to 18) shows the photo-cross-linking pattern obtained when the cross-linking agent was at position 24 of the U1 PSEA. Although the N- and C-terminal fragments of wild-type DmSNAP190 (construct C) did not resolve in lane 16, results shown in Fig. 2 had already demonstrated that position 24 cross-linked to the N-terminal fragment of DmSNAP190. When the cleavage site was moved farther C terminal (constructs D and E), the N-terminal fragment, as expected, continued to cross-link (Fig. 3C, lanes 17 and 18). However, when the cleavage site was moved to the center of repeat Rb (construct B), it was the C-terminal fragment that cross-linked (lane 15). Together, these results indicated that position 24 of the U1 PSEA cross-linked to amino acids of DmSNAP190 located between residues 306 and 358. When cross-linking was carried out to position 12 of the U1 PSEA (Fig. 3C, lanes 19 to 24), an identical pattern was observed. Thus, position 12 also cross-linked to DmSNAP190 residues located between residues 306 and 358.

Lanes 25 to 30 of Fig. 3C show that a different pattern of results was obtained when DmSNAPc was cross-linked to position 1 of the U1 PSEA. The pattern was similar to that observed in lanes 13 to 24 of Fig. 3C but with two important exceptions. Data in Fig. 2 had indicated that the band observed in lane 28 (construct C) arose from cross-linking to the C terminus of wild-type DmSNAP190 (in contrast to the N-terminal cross-linking shown in lanes 16 and 22). When the cleavage site was moved to the center of the Rd repeat (construct D), the band of strongest intensity (lane 29) corresponded also to the C-terminal fragment. Thus, amino acids C-terminal of residue 409 cross-linked most strongly to position 1 of the U1 PSEA. However, a weaker band that corresponded to the N-terminal fragment of construct D was also visible in lane 29 (compare its position to the corresponding fragment in lanes 17 and 23), suggesting that amino acid residues between 359 and 409 were also able to cross-link to position 1 of the U1 PSEA, but with lower intensity. When the protein cleavage site was moved C-terminal to the Myb domain (construct E), cross-linking could only be detected to the N-terminal fragment (lane 30). Thus, when DmSNAPc bound the U1 PSEA, the region of DmSNAP190 between positions 359 and 483 mapped in close proximity to position 1 of the U1 PSEA.

The results of the DmSNAP190 mapping experiments are summarized schematically in Fig. 3D. Amino acids localized to the C-terminal half of Rb and/or the N-terminal half of Rc contact positions 24 and 12 of the U1 PSEA. In contrast, amino acids localized to Rd and/or the C-terminal half of Rc closely approach position 1 of the U1 PSEA. These data further suggest that the DNA-binding domain of DmSNAP190 on the U1 PSEA is oriented with its N terminus proximal to the U1 transcription start site and its C terminus more distal from the start site.

Mapping domains of DmSNAP50 that contact nucleotide positions 22 and 17 of the U1 PSEA and positions 14 and 17 of the U6 PSEA. When DmSNAPc binds to the U1 or U6 PSEAs, the DmSNAP50 subunit cross-links most strongly to phosphate positions 17 and 22 of the U1 PSEA but to positions 17 and 14 of the U6 PSEA (Fig. 1A). We therefore chose to map regions within DmSNAP50 that contact these four nucleotide positions. DmSNAP50 contains no canonical DNA-binding domain, but it does contain an unorthodox zinc-binding domain near its C terminus (Fig. 1B) that has been termed the SNAP finger (17). Although this unique region has no homology to other types of DNA-binding zinc fingers, it has nevertheless been implicated in DNA binding by human SNAPc (17).

DmSNAP50 contains three NG peptide bonds (Fig. 1B). However, for the purposes of domain-mapping experiments, the two nearly adjacent sites at residues 234 and 237 can effectively be treated as a single hydroxylamine cleavage site. Initial mapping experiments were attempted with the wild-type

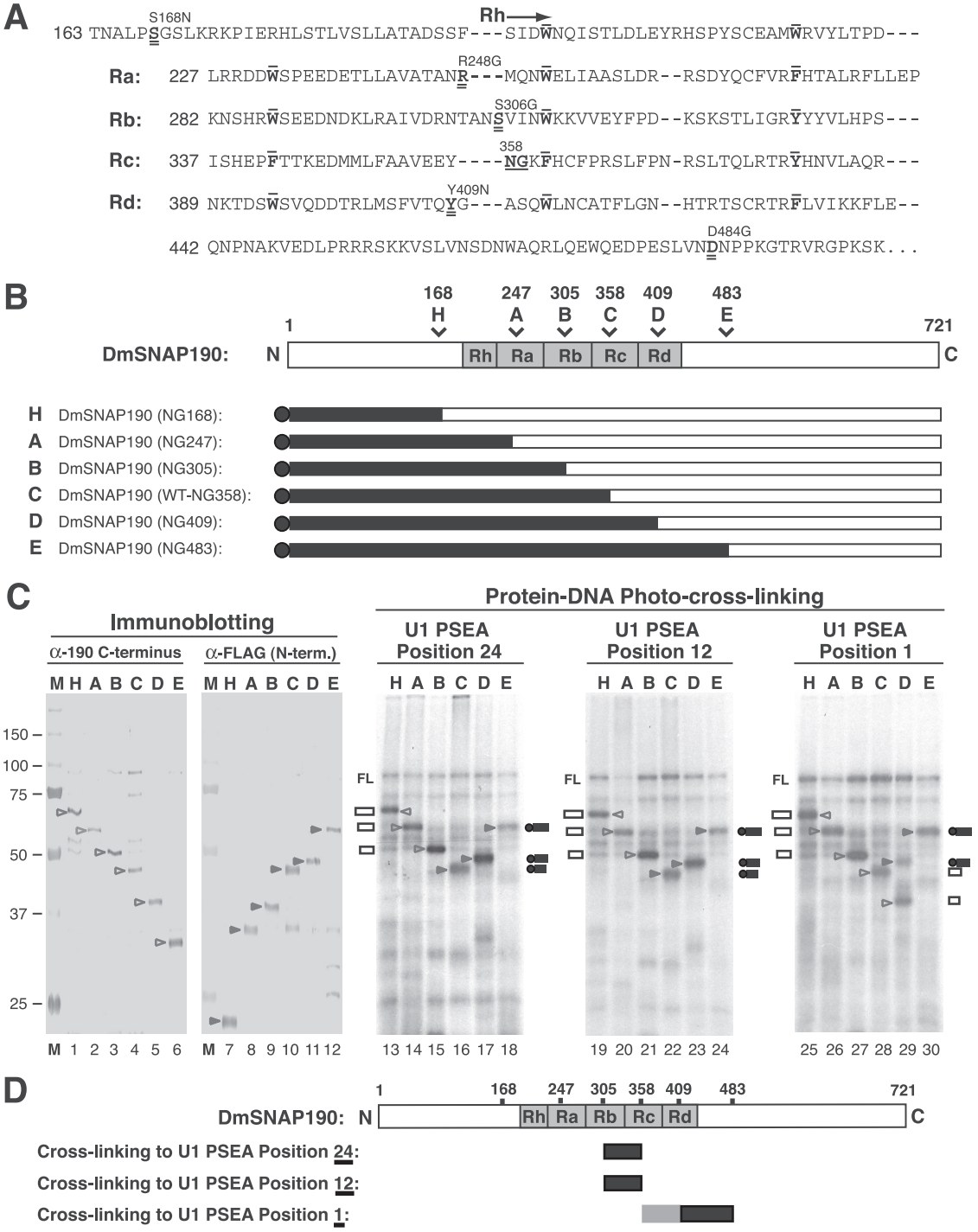

FIG. 3. Localization of domains of DmSNAP190 that cross-link to U1 PSEA positions 24, 12, and 1. (A) Alignment of the Myb repeat sequences and flanking amino acids of DmSNAP190. The conserved aromatic amino acids characteristic of Myb repeats are shown in bold and are overlined. The NG hydroxylamine cleavage site at position 358 in the wild-type protein is bold and underlined. The positions where mutations were introduced to generate new NG cleavage sites are indicated above the sequence with the affected residues shown in bold and double underlined. (B) Schematic representation of the hydroxylamine cleavage patterns of six N-tagged DmSNAP190 constructs. The shaded and unshaded regions indicate the two cleavage fragments in each construct. (C) All gel results shown are subsequent to hydroxylamine digestion of DmSNAPc from cell lines separately expressing the six DmSNAP190 variants represented by constructs H to E shown in panel B. The first two panels show immunoblots using antibodies against a C-terminal peptide of DmSNAP190 (lanes 1 to 6) or antibodies against the FLAG epitope at the N terminus of each construct (lanes 7 to 12). Arrowheads (unshaded for C-terminal fragments and shaded for N-terminal fragments) point out the specific cleavage products of interest. The third, fourth, and fifth panels show results of protein-DNA photo-cross-linking of the six DmSNAP190 variants to U1 PSEA position 24 (lanes 13 to 18), position 12 (lanes 19 to 24), or position 1 (lanes 25 to 30). Symbols representing the fragments present in the first three lanes of each panel are shown to the left of each panel, and symbols representing the fragments in the last three lanes are shown to the right. In lane 29, the N- and C-terminal fragments both cross-linked to position 1. FL refers to the full-length undigested DmSNAP190. (D) Schematic drawing indicating the region of DmSNAP190 that cross-links to each of the positions 24, 12 and 1 of the U1 PSEA. The dark shading indicates the regions of strongest cross-linking, and the light shading indicates a region of weaker cross-linking. WT, wild type.

protein, but an excessive number of bands on the autoradiograms of the gels made it difficult to interpret the data.

Therefore, new constructs that lacked NG bonds at either position 179 or at positions 234 and 237 were made. We also made a third construct that lacked all three wild-type NG sites but contained a new NG bond at position 102. These constructs, with their expected hydroxylamine fragmentation patterns, are illustrated in Fig. 4A.

Figure 4B shows immunoblots of the various constructs after hydroxylamine digestion and detection with polyclonal antibodies produced against full-length recombinant DmSNAP50 (lanes 1 to 3, 5 to 7, and 9 to 11) or with monoclonal antibodies against the FLAG epitope (lanes 4, 8, and 12). In lanes 1 to 3, both the N- and C-terminal fragments were clearly detected by the anti-DmSNAP50 antibodies. In lanes 5 to 7, the C-terminal fragment gave a strong signal, but the signal from the N-terminal fragment, although still visible, was barely detectable. In lanes 9 to 11, only the large C-terminal fragments could be detected. In lanes 4, 8, and 12 of Fig. 3B, the anti-FLAG antibodies clearly revealed and confirmed the position of the FLAG-tagged N-terminal fragment for each of the three constructs.

Figure 4C shows the results of photo-cross-linking these constructs to either position 22 of the U1 PSEA (lanes 1 to 9) or position 14 of the U6 PSEA (lanes 10 to 18). When the construct that could be cleaved only at positions 234/237 was cross-linked to position 22 of the U1 PSEA, only N-terminal fragments cross-linked (Fig. 4C, lanes 1 to 3). (Compare with the N-terminal fragments in the immunoblots in Fig. 4B, lanes 1 to 3.) When cleavage of DmSNAP50 occurred at position 179, again only the N-terminal fragments cross-linked to the radiolabeled probe (Fig. 4C, lanes 4 to 6). (Compare to the weakly detected N-terminal bands in lanes 5 to 7 of Fig. 4B.)

When the DmSNAP50 cleavage site followed residue 102, the pattern changed (Fig. 4C, lanes 7 to 9). In this case, the cross-linked fragments from the untagged and the N-terminally tagged constructs migrated to identical positions, but the fragment from the C-terminally tagged construct migrated more slowly, matching the pattern of C-terminal fragments seen in Fig. 4B (lanes 9 to 11). This indicates that cross-linking occurred C-terminal to position 102.

By combining the data from lanes 1 to 9 of Fig. 4C, we conclude that position 22 of the U1 PSEA is contacted by amino acid residues located between positions 103 and 179 of DmSNAP50. Initially, this was an unexpected finding, because work on human SNAP50 had implicated the noncanonical zinc-binding domain (i.e., the SNAP finger) at the C terminus of the protein as being involved in DNA binding (17).

Besides the major bands that corresponded to the expected cleavage products and to the undigested full-length tagged and untagged DmSNAP50 polypeptides, a number of other bands were also visible in the autoradiograms. Many of these bands could be accounted for in a specific manner. The solid horizontal tick on the right side of each panel indicates the position of weak bands visible in lanes 2, 3, 5, 6, and 8 (but obscured in lane 9). These bands undoubtedly represent hydroxylamine digestion products of DmSNAP190, as this subunit could also cross-link to the U1 PSEA position 22 (Fig. 1). This band was not observed in lanes 1, 4, and 7, because DmSNAP190 in those lanes had an N-terminal tag to allow purification of DmSNAPc that contained untagged DmSNAP50. If the N-terminal fragment of DmSNAP190 cross-linked to position 22, it should run with a slightly slower mobility and would thus be obscured by the full-length DmSNAP50 in lanes 1, 4, and 7. Thus, these data provide a "bonus" indication that the N-terminal half of DmSNAP190 cross-linked to position 22 of the U1 PSEA.

The asterisk to the right of lane 9 in Fig. 4C indicates a minor unexpected band. This band in lane 9 appears to run at a position that is just slightly above the C-terminal fragments that lacked tags (Fig. 4C, lanes 7 and 8). Although we do not understand the exact origin of this band, it appears to represent a C-terminal fragment from which most of the C-terminal tag has been removed. A weak band was also normally seen at this position in the immunoblots (migrating slightly slower than the untagged C-terminal fragments) as indicated by asterisks to the right of lanes 3, 7, and 11 in Fig. 4B. Perhaps the unnatural C-terminal tag is to some extent subject to *in vitro* proteolysis following FLAG purification or during hydroxylamine treatment. This would explain why this "starred" band very commonly appeared in the autoradiograms whenever cross-linking to a C-terminal fragment occurred (for further examples, see Fig. 4C, lanes 12, 15, and 18, as discussed below).

Next, we investigated the region of DmSNAP50 that contacted position 14 of the U6 PSEA. These data are shown in Fig. 4C (lanes 10 to 18). Lanes 10 to 12 demonstrate that the DmSNAP50 fragment C-terminal to residue 237 cross-linked to U6 PSEA position 14. Thus, amino acid residues located between positions 238 and 377, a region that includes the SNAP finger, are in close proximity to position 14 of the U6 PSEA. This cross-linking pattern and result were in sharp contrast to that obtained when the same construct was cross-linked to position 22 of the U1 PSEA (Fig. 4C, compare lanes 10 to 12 with lanes 1 to 3). Thus, position 22 of the U1 PSEA and position 14 of the U6 PSEA each cross-linked to a distinct domain of DmSNAP50.

When the other two DmSNAP50 constructs were used in the assay, it was again the C-terminal fragment that cross-linked to U6 PSEA position 14 (Fig. 4C, lanes 13 to 18). This was consistent with the conclusion drawn from the data in lanes 10 to 12 that DmSNAP50 residues C-terminal to residue 237 contacted position 14 of the U6 PSEA.

The results of cross-linking DmSNAP50 to position 22 of the U1 PSEA and to position 14 of the U6 PSEA are summarized diagrammatically in Fig. 4D. An important conclusion of these data is that these two nucleotide positions in the U1 and U6 PSEAs are contacted by two distinct regions of DmSNAP50.

There were again some unexpected bands in the gel. The bands marked with asterisks in lanes 12, 15, and 18 are similar to the band in lane 9, as previously discussed. Additional minor bands in lanes 13 to 18 are marked with dotted lines, and these bands migrate to the same positions as the major cleavage products in lanes 10 to 12. These unexpected bands therefore must arise from residual hydroxylamine cleavage at positions 234/237. Because the NG sites at these positions were changed to QG sites, the appearance of these bands suggests that hydroxylamine is capable of cleaving at the QG peptide bonds but with low efficiency. The appearance of this residual cleav-

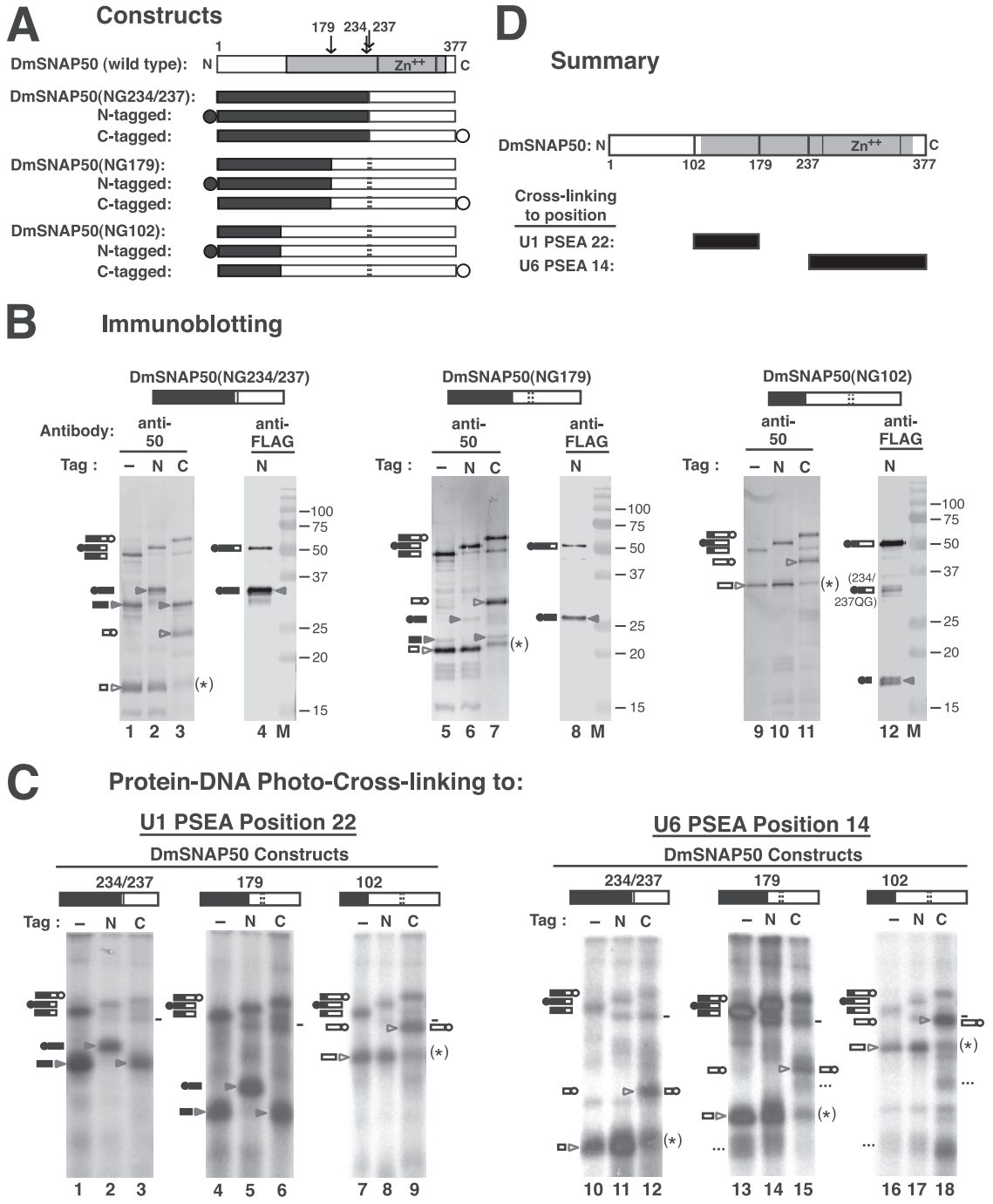

FIG. 4. Two distinct domains DmSNAP50 contact different nucleotides in the U1 and U6 PSEAs. (A) The top line shows the structure of wild-type DmSNAP50, including the evolutionarily conserved region (shaded), the hydroxylamine cutting sites (vertical arrows), and the zinc finger region. Below this are diagrammed nine DmSNAP50 constructs used for photo-cross-linking. (For one set of constructs, a new cutting site was introduced at position 102.) The vertical dashes at positions 234/237 indicate a point of weak residual cutting in the lower six constructs that arises from the two nearly adjacent sites that were mutated from NG to QG. (B) Immunoblots showing the locations of N-terminal and C-terminal fragments arising from hydroxylamine digestion of the constructs shown in panel A. Symbols are as described in the legends of Fig. 2 and 3. The asterisks in parentheses indicate the position of a band (in lanes 3, 7, and 11) that appears to correspond to a C-terminal fragment that has lost its C-terminal tag. See the text for further discussion. (C) Autoradiograms of chemical digestion products of the nine constructs cross-linked to U1 PSEA position 22 (lanes 1 to 9) or to U6 PSEA position 14 (lanes 10 to 18). The symbols alongside the gels designate the fragments responsible for the bands observed, and the arrowheads further clarify the relevant bands. The horizontal tick alongside the right of each panel indicates a digestion product of DmSNAP190. (D) The dark bars indicate the two different regions of DmSNAP50 that cross-linked to position 22 of the U1 PSEA or to position 14 of the U6 PSEA.

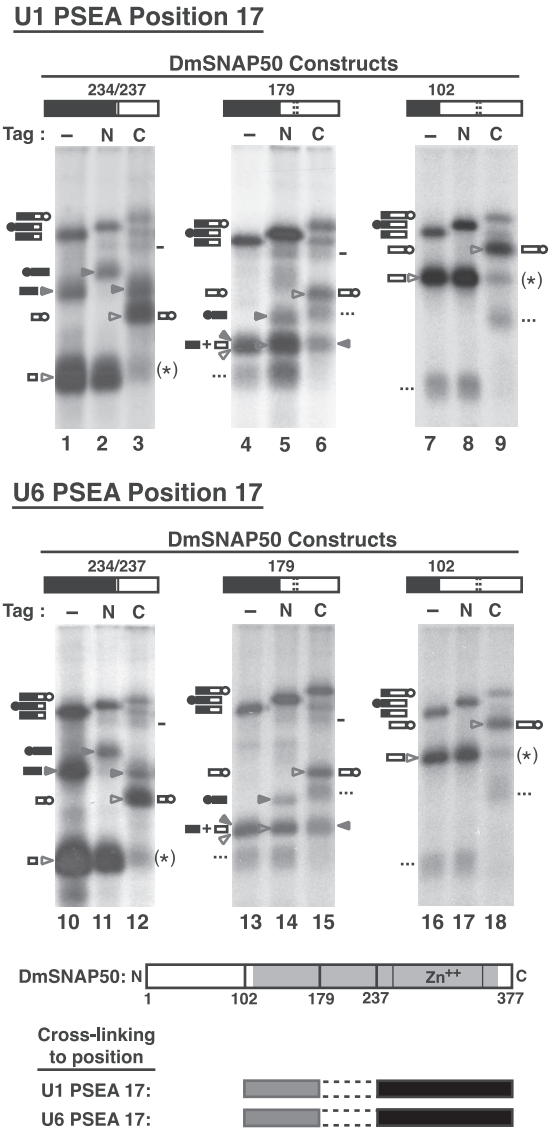

FIG. 5. Two domains of DmSNAP50 contact position 17 of both the U1 and U6 PSEAs. DmSNAPc separately containing nine different constructs of DmSNAP50 (shown in Fig. 4) was cross-linked to position 17 of the U1 PSEA or to position 17 of the U6 PSEA, digested with hydroxylamine, run on denaturing gels, and subjected to autoradiography. Symbols are the same as those described in previous figure legends. The bottom of the figure summarizes the regions of DmSNAP50 that cross-linked to position 17 of each PSEA. Dark shading indicates a region that cross-linked strongly, and lighter shading indicates a region of somewhat weaker cross-linking. It is undetermined whether residues 180 to 237 cross-link to position 17.

age is enhanced in this case because there are two QG sites nearly adjacent to each other.

Next, we investigated the cleavage patterns of the DmSNAP50 NG variants after cross-linking to position 17 of the U1 PSEA or to position 17 of the U6 PSEA (Fig. 5). When DmSNAP50 was cleaved at position 234/237, the N-terminal fragment and the C-terminal fragment both cross-linked to U1 position 17 (Fig. 5, lanes 1 to 3; compare to the immunoblot pattern in Fig. 4B, lanes 1 to 3). In fact, the autoradiography pattern obtained was a sum of the bands previously observed when cross-linking was carried out to U1 position 22 and to U6 position 14 (Fig. 4C, lanes 1 to 3 and 10 to 12).

When the DmSNAP50 construct that had the hydroxylamine cleavage site at position 179 was used for cross-linking to U1 position 17 (Fig. 5, lanes 4 to 6), the pattern was initially difficult to interpret. However, upon further study, it became evident that this pattern also represented the sum of the bands observed when this construct was cross-linked to U1 position 22 and U6 position 14 (compare with Fig. 4C, lanes 4 to 6 and 13 to 15; also see the immunoblot in Fig. 4B, lanes 5 to 7). In Fig. 5, lane 4, the untagged N-terminal fragment and the untagged C-terminal fragment did not resolve on the autoradiogram. Besides these unresolved untagged N- and C-terminal fragments, a fragment with slower mobility due to its N-terminal tag was seen in Fig. 5, lane 5, and a fragment with still slower mobility due to its C-terminal tag was seen in lane 6. These correspond to the N-tagged and C-tagged fragments identified in the immunoblots of Fig. 4B, lanes 6 to 8.

When DmSNAP50 was fragmented at position 102, only the C-terminal fragment cross-linked (Fig. 5, lanes 7 to 9). This finding was consistent with the results shown in Fig. 4C, lanes 7 to 9 and 16 to 18. Additional bands marked in Fig. 5 with asterisks or dotted lines arose from the previously discussed artifactual phenomena.

Lanes 10 to 18 of Fig. 5 show the cross-linking patterns obtained when the DmSNAPc NG variants were cross-linked to position 17 of the U6 PSEA. In each case, the gel patterns were identical to those observed for position 17 of the U1 PSEA (compare each panel to the one directly above it). Thus, at the resolution obtained by these experiments, phosphate 17 is in a similar environment relative to the DmSNAP50 polypeptide when DmSNAPc binds either to a U1 PSEA or to a U6 PSEA. We can conclude from the data in Fig. 5 that position 17 of both PSEAs cross-linked to DmSNAP50 at residues located between 238 and 377 and with less intensity to residues between 103 and 179. From these data, however, we were unable to exclude the possibility that position 17 might additionally cross-link to residues located between 180 and 237. A schematic summary of these findings is shown at the bottom of Fig. 5.

Mapping regions of DmSNAP43 that contact nucleotide positions 40, 28, and 20 of the U1 PSEA and positions 16 and 11 of the U6 PSEA. DmSNAP43 has no known canonical DNA-binding domains; furthermore, it contacts the U1 and U6 PSEAs very differently. When DmSNAPc binds to a U6 PSEA, DmSNAP43 cross-links most strongly to positions 11 and 16 of the U6 PSEA (but not when it binds to a U1 PSEA) (Fig. 1A). On the other hand, when DmSNAPc binds to a U1 PSEA, DmSNAP43 cross-links much more strongly to position 20 and to positions 28 and 40 downstream of the U1 PSEA (Fig. 1A and reference 22). We therefore mapped domains of DmSNAP43 that cross-link to each of these five positions within the U1 or U6 PSEAs.

In initial experiments, the single NG peptide bond present at position 192 in wild-type DmSNAP43 was used for protein mapping. Figure 6, lanes 1 to 3, shows an immunoblot of hydroxylamine-cleaved untagged, N-tagged, and C-tagged DmSNAP43 detected using antibodies produced in rabbits against full-length recombinant DmSNAP43. These polyclonal antibodies strongly reacted with the C-terminal fragments but

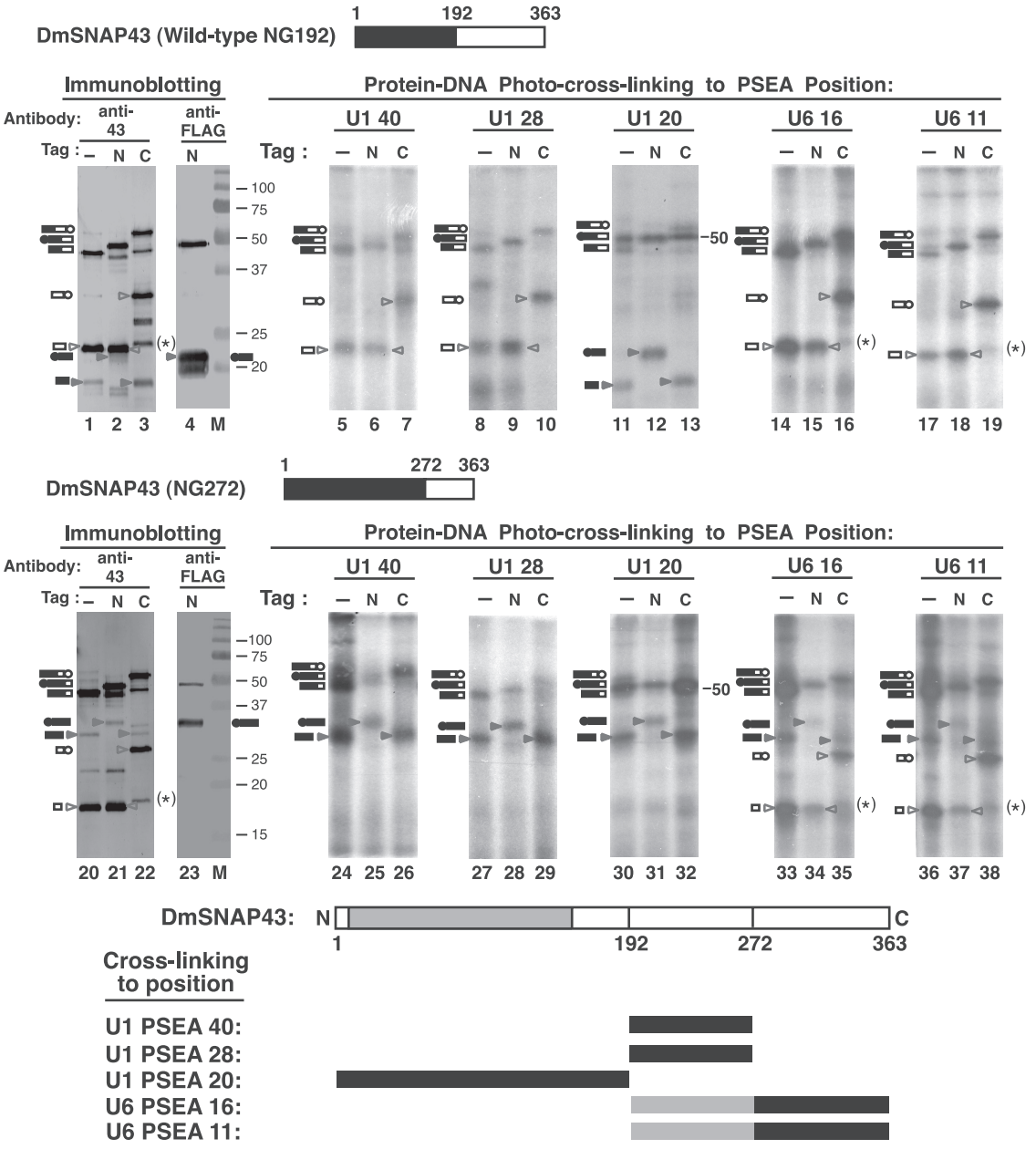

FIG. 6. Three distinct regions of DmSNAP43 differentially contact the U1 and U6 PSEAs. DmSNAPc was cross-linked to three phosphate positions in or downstream of the U1 PSEA and to two positions in the U6 PSEA. The DmSNAPc contained DmSNAP43 with a hydroxylamine cutting site either at the wild-type position (upper panels) or at position 272 (lower panels). Immunoblots shown in lanes 1 to 4 and 20 to 23 reveal the positions of the C-terminal and N-terminal fragments. The symbols that mark the band positions and identify the fragments in each panel are similar to those used in Fig. 2 to 5. The bottom of the figure summarizes the regions of DmSNAP43 that cross-link to the indicated positions in the U1 and U6 PSEAs. Dark shading indicates regions that cross-link strongly to the indicated nucleotide positions, and lighter shading indicates a region of weaker cross-linking.

detected the N-terminal fragments poorly. Nonetheless, the bands corresponding to the N-terminal fragments were visible when the blots were overdeveloped relative to the C-terminal fragment bands. Furthermore, an immunoblot with FLAG antibody (lane 4) confirmed the position of the tagged N-terminal fragment.

Figure 6, lanes 5 to 19, shows the results of photo-cross-linking to various nucleotide positions within the U1 and U6 PSEAs. When the cross-linking agent was at U1 PSEA position 40, only the C-terminal fragment cross-linked (lanes 5 to 7). The same result was obtained with the cross-linker positioned at nucleotide 28 of the U1 PSEA (lanes 8 to 10). However, when the cross-linker was at position 20, the pattern changed and revealed that the N-terminal fragment cross-linked to U1 PSEA position 20 (lanes 11 to 13).

Somewhat surprisingly, when cross-linking was carried out to positions 16 and 11 of the U6 PSEA, it was again the C-terminal half of DmSNAP43 that cross-linked to these posi-

tions (Fig. 6, lanes 14 to 19). Thus, the N-terminal half of DmSNAP43 cross-linked to position 20 of the U1 PSEA, but the C-terminal half cross-linked to positions 28 and 40 of the U1 PSEA and to positions 11 and 16 of the U6 PSEA.

To better localize the contact points within the C-terminal domain of DmSNAP43 for the U1 and U6 PSEAs, we next prepared a construct that contained an NG peptide bond only at position 272. The immunoblots in Fig. 6, lanes 20 to 23, show the positions of the N- and C-terminal fragments following hydroxylamine cleavage (the signals from the N-terminal fragments are weak in lanes 20 to 22). When this mutant construct was cross-linked to U1 PSEA positions 40 or 28 (lanes 24 to 29), only the N-terminal fragment (residues 1 to 272) cross-linked. In combination with the results shown in lanes 5 to 10, we can conclude that DmSNAP43 amino acid residues between 193 and 272 are in close proximity to nucleotide positions 40 and 28 downstream of the U1 PSEA. Figure 6 (lanes 30 to 32) shows that U1 position 20 cross-linked to the N-terminal fragment of the mutant construct; this was consistent with U1 position 20 cross-linking to the N-terminal half of DmSNAP43, as previously seen in lanes 11 to 13.

Importantly, and in contrast, positions 16 and 11 of the U6 PSEA cross-linked primarily to the most C-terminal fragment (residues 273 to 363) (Fig. 6, lanes 33 to 38). Less-intense bands of cross-linking to the N-terminal fragment were also visible in each of these lanes. From these results, we conclude that U6 PSEA positions 11 and 16 cross-linked most strongly to the last 91 amino acids of DmSNAP43, but residues between 193 and 272 could also be cross-linked with a lower efficiency.

A summary of the conclusions from the DmSNAP43 mapping experiments is shown at the bottom of Fig. 6. The N-terminal half of DmSNAP43 cross-linked to position 20 of the U1 PSEA, but positions 28 and 40 cross-linked to residues located between 193 and 272. Finally, positions 16 and 11 of the U6 PSEA cross-linked most strongly to residues 273 to 363 of DmSNAP43 but also cross-linked weakly to residues 193 to 272.

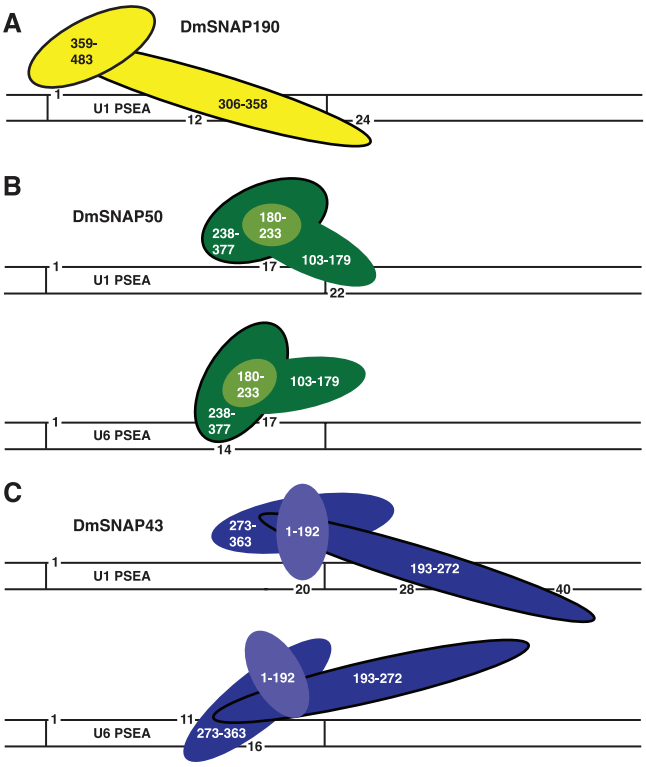

FIG. 7. Domains of DmSNAP190, DmSNAP50, and DmSNAP43 that interact with specific nucleotide positions within and downstream of the U1 and U6 PSEAs. Ovals represent amino acid domains within each of the subunits as labeled. (A) Domains of DmSNAP190 interacting with nucleotide positions 1, 12, and 24 of the U1 PSEA. (B) Domains of DmSNAP50 that interact with positions 17 and 22 of the U1 PSEA or with positions 14 and 17 of the U6 PSEA. (C) Domains of DmSNAP43 that interact with positions 20, 28, and 40 of the U1 PSEA or with positions 11 and 16 of the U6 PSEA.

DISCUSSION

Previous studies had identified nucleotide positions contacted by each of the three subunits of DmSNAPc (22, 32). However, the domains of DmSNAPc that interact directly with the DNA had not been determined. In this study, we have been able to localize regions within the DmSNAPc subunits that are involved not only in contacting DNA but also in contacting specific nucleotide positions within the U1 and U6 PSEAs. From these data, we have mapped different domains of DmSNAP50 and of DmSNAP43 that contact different regions of the U1 and U6 PSEAs. In the absence of an atomic structure for SNAPc, these cross-linking results provide a basic understanding of significant structural differences of DmSNAPc on the U1 and U6 PSEAs. The results are summarized schematically in Fig. 7.

Domains of DmSNAP190 that interact with the U1 PSEA. Work with human SNAP190 revealed that the RcRd repeats alone were capable of binding to DNA but with relatively low sequence specificity (15, 25, 28, 33). In those studies, no DNA-binding activity was noted for the RhRaRb region (28, 33). Our results, within the resolution of the experiments shown in Fig. 3, are consistent with those findings. We found that a region of the protein encompassed by residues 306 to 358 (located from the middle of the Rb repeat to the middle of the Rc repeat) interacted with positions 12 and 24 of the U1 PSEA (Fig. 3D and 7A). In contrast, a region encompassed by residues 359 to 483 (including the C-terminal half of the Rc repeat, the Rd repeat, and 42 amino acids C-terminal of the Rd repeat) was in proximity to position 1 of the U1 PSEA (Fig. 3D and 7A). This indicates that the C-terminal region of the Myb domain interacts with the 5′ end of the U1 PSEA and that the central region of the Myb domain interacts with the 3′ portion of the U1 PSEA. In a simplistic interpretation, this orientation of binding would place the N terminus of DmSNAP190 nearer the transcription start site and the C terminus farther from the start site. Whether this holds true as the protein folds in three dimensions is not yet clear. However, this orientation is consistent with the fact that the N-terminal domain of human SNAP190 interacts with the TATA-binding protein (TBP) (25), and TBP would most likely localize downstream of the PSEA toward the transcription start site. This orientation also makes sense from the perspective that the C-terminal domain of human SNAP190 interacts with Oct-1 protein bound to an upstream enhancer element (9).

Recent work from our laboratory indicated that a region of

DmSNAP190 C-terminal to the Myb domain was required for DmSNAPc to bind to DNA (16). However, in the studies reported here, we did not observe cross-linking of the C-terminal domain (residues 484 to 721) to the DNA. It is thus possible that the C-terminal domain of DmSNAP190 may contact nucleotides other than the three examined in the current study; alternatively, this region of DmSNAP190 may be required for DmSNAPc to adopt a proper DNA-binding conformation. Additional work will be required to answer this question.

Domains of DmSNAP50 interacting with the U1 and U6 PSEAs. DmSNAP50 and DmSNAP43 are the two subunits of DmSNAPc whose cross-linking patterns exhibit the most diversity when DmSNAPc binds to a U1 versus a U6 PSEA (22, 32). Moreover, the cross-linking of these two subunits is restricted to the 3′ half (and to sequences downstream) of the PSEAs. Therefore, we concentrated most of our effort on identifying the domains of these subunits that are closely associated with the DNA when DmSNAPc binds to either the U1 or U6 PSEA.

In the case of DmSNAP50 on the U6 PSEA, the SNAP finger at the C terminus of DmSNAP50 contacts phosphate position 14 (Fig. 4 and 7B), consistent with the finding that the SNAP finger is required for DNA binding by human SNAPc (17). However, when DmSNAPc binds to a U1 PSEA, a distinct region closer to the N terminus of DmSNAP50 (i.e., residues 103 to 179) associates closely with position 22 of the U1 PSEA (Fig. 4 and 7B).

Interestingly, both DmSNAP50 DNA-binding domains contact position 17 of the U1 and U6 PSEAs (Fig. 5 and 7B). This raises the possibility that position 17 may act as a fulcrum about which DmSNAP50 rotates so that residues 103 to 179 are in close proximity to position 22 of the U1 PSEA, but residues 237 to 377 are in close proximity to position 14 of the U6 PSEA (Fig. 7B).

Recent work from our laboratory found that deletion of the nonconserved N-terminal domain of DmSNAP50 (residues 1 to 91) eliminated the DNA-binding activity of DmSNAPc while having no effect on the ability of the three subunits to associate with each other. This N-terminal region of the protein did not cross-link to any of the four phosphate positions we examined (U1-22, U1-17, U6-14, and U6-17). As discussed above regarding DmSNAP190, it may be that this N-terminal domain of DmSNAP50 contacts nucleotides other than the four studied herein; alternatively, this region of DmSNAP50 may be required for DmSNAPc to adopt a proper DNA-binding conformation.

Domains of DmSNAP43 interacting with the U1 and U6 PSEAs. DmSNAP43 cross-links with significantly different patterns on U1 and U6 PSEAs (Fig. 1A and references 22 and 32) but contains no obvious sequence similarity to any characterized DNA-binding domains. Our earlier site-specific protein-DNA photo-cross-linking studies indicated that DmSNAP43 cross-linked strongly to the phosphate at position 20 in the U1 PSEA but only very weakly to that position when DmSNAPc was bound to the U6 PSEA (Fig. 1A). The new data indicate that this contact to the U1 PSEA at position 20 occurs through the N-terminal domain encompassed by residues 1 to 192 of DmSNAP43 (Fig. 6 and 7C).

On the other hand, phosphate positions 28 and 40, located downstream of the U1 PSEA, cross-linked to residues located within an 80-residue domain stretching from residue 193 to 272 of DmSNAP43 (Fig. 6 and 7C). Interestingly, on a U6 PSEA, this same region of DmSNAP43 cross-linked weakly to phosphates 11 and 16 of the U6 PSEA (Fig. 6 and 7C). Thus, this domain of the protein (residues 193 to 272) may be in radically different positions when DmSNAPc binds to a U6 versus a U1 PSEA.

The strongest cross-linking contacts with phosphates 11 and 16 of the U6 PSEA, however, were made by the most C-terminal domain of DmSNAP43 (residues 273 to 363) (Fig. 6 and 7C). These contacts are nonexistent or extremely weak when DmSNAPc binds to a U1 PSEA (Fig. 1A). Thus, when the C-terminal domain of DmSNAP43 (residues 273 to 363) closely approaches positions 11 and 16 of the U6 PSEA, it probably also brings residues 193 to 272 into closer proximity to these nucleotides as well (Fig. 7C).

In work recently published by our laboratory (16), a C-terminal truncation of DmSNAP43 that deleted amino acids beyond residue 172 eliminated the DNA-binding activity of DmSNAPc even though it had no effect on the ability of DmSNAP43 to form a complex with DmSNAP50 and DmSNAP190. It is now clear from the photo-cross-linking data that the C-terminal domain of DmSNAP43, although not conserved in sequence during evolution, can make extensive contacts with the DNA that can extend over a 30-bp region. Furthermore, the pattern is different depending upon whether DmSNAPc is bound to a U1 or U6 PSEA.

The mutant constructs used in the structural investigations reported here were not tested for activity in a transcription assay. Despite this caveat, each of the mutant constructs continued to bind to DNA and to cross-link strongly to the same positions as the wild-type protein. Thus, it seems unlikely that the point mutations resulted in significant structural alterations to the DmSNAP complex.

In conclusion, we have localized domains within the DmSNAPc subunits that contact specific nucleotides in the U1 and U6 snRNA gene promoters. These data help to establish, to a first approximation, the orientation of the subunits on the DNA. Furthermore, the results reveal that different domains of the DmSNAP50 and DmSNAP43 subunits are in close proximity to different nucleotide positions when DmSNAPc binds to a U1 versus a U6 PSEA. These results support a model in which different conformations of DmSNAPc, induced by relatively minor DNA sequence differences, are important for determining the RNA polymerase specificity of snRNA gene promoters. In a similar vein, recent work established that the structure and the activity of the glucocorticoid receptor depend upon the specific DNA sequence that it recognizes and binds to as a homodimer (27).

ACKNOWLEDGMENTS

We thank Kathleen McNamara-Schroeder for excellent technical assistance.

This work was supported by National Science Foundation grant MCB-0842770 and in part by the California Metabolic Research Foundation. N.H.B. was a recipient of an Arne N. Wick predoctoral research fellowship from the California Metabolic Research Foundation.

REFERENCES

1. **Bai, L., Z. X. Wang, J. B. Yoon, and R. G. Roeder.** 1996. Cloning and characterization of the β subunit of human proximal sequence element-

binding transcription factor and its involvement in transcription of small nuclear RNA genes by RNA polymerases II and III. Mol. Cell. Biol. **16:** 5419–5426.
2. **Barakat, N. H., and W. E. Stumph.** 2008. TBP recruitment to the U1 snRNA gene promoter is disrupted by substituting a U6 proximal sequence element A (PSEA) for the U1 PSEA. FEBS Lett. **582:**2413–2416.
3. **Crimmins, D. L., and S. M. Mische.** 1996. Chemical cleavage of proteins in solution, p. 11.4.1-11.4.8. *In* V. B. Chanda, vol. 2, suppl. 4. (ed.), Current protocols in protein science. John Wiley & Sons, Inc., New York, NY.
4. **Dahlberg, J. E., and E. Lund.** 1991. How does III x II make U6? Science **254:**1462–1463.
5. **Dahlberg, J. E., and E. Lund.** 1988. The genes and transcription of the major small nuclear RNAs, p. 38–70. *In* M. L. Birnstiel (ed.), Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer Verlag KG, Heidelberg, Federal Republic of Germany.
6. **Das, A., and V. Bellofatto.** 2003. RNA polymerase II-dependent transcription in trypanosomes is associated with a SNAP complex-like transcription factor. Proc. Natl. Acad. Sci. U. S. A. **100:**80–85.
7. **Das, A., Q. Zhang, J. B. Palenchar, B. Chatterjee, G. A. Cross, and V. Bellofatto.** 2005. Trypanosomal TBP functions with the multisubunit transcription factor tSNAP to direct spliced-leader RNA gene expression. Mol. Cell. Biol. **25:**7314–7322.
8. **Das, G., D. Henning, and R. Reddy.** 1987. Structure, organization, and transcription of *Drosophila* U6 small nuclear RNA genes. J. Biol. Chem. **262:**1187–1193.
9. **Ford, E., M. Strubin, and N. Hernandez.** 1998. The Oct-1 POU domain activates snRNA gene transcription by contacting a region in the SNAPc largest subunit that bears sequence similarities to the Oct-1 coactivator OBF-1. Genes Dev. **12:**3528–3540.
10. **Henry, R. W., E. Ford, R. Mital, V. Mittal, and N. Hernandez.** 1998. Crossing the line between RNA polymerases: transcription of human snRNA genes by RNA polymerases II and III. Cold Spring Harbor Symp. Quant. Biol. **63:** 111–120.
11. **Henry, R. W., B. C. Ma, C. L. Sadowski, R. Kobayashi, and N. Hernandez.** 1996. Cloning and characterization of SNAP50, a subunit of the snRNA-activating protein complex SNAPc. EMBO J. **15:**7129–7136.
12. **Henry, R. W., C. L. Sadowski, R. Kobayashi, and N. Hernandez.** 1995. A TBP-TAF complex required for transcription of human snRNA genes by RNA polymerases II and III. Nature **374:**653–656.
13. **Hernandez, G., F. Valafar, and W. E. Stumph.** 2007. Insect small nuclear RNA gene promoters evolve rapidly yet retain conserved features involved in determining promoter activity and RNA polymerase specificity. Nucleic Acids Res. **35:**21–34.
14. **Hernandez, N.** 2001. snRNA genes: a model system to study fundamental mechanisms of transcription. J. Biol. Chem. **276:**26733–26736.
15. **Hinkley, C. S., H. A. Hirsch, L. P. Gu, B. LaMere, and R. W. Henry.** 2003. The small nuclear RNA-activating protein 190 Myb DNA binding domain stimulates TATA box-binding protein-TATA box recognition. J. Biol. Chem. **278:**18649–18657.
16. **Hung, K.-H., M. Titus, S. C. Chiang, and W. E. Stumph.** 2009. A map of *Drosophila melanogaster* small nuclear RNA-activating protein complex (DmSNAPc) domains involved in subunit assembly and DNA binding. J. Biol. Chem. **284:**22568–22579.
17. **Jawdekar, G. W., A. Hanzlowsky, S. L. Hovde, B. Jelencic, M. Feig, J. H. Geiger, and R. W. Henry.** 2006. The unorthodox SNAP50 zinc finger domain contributes to cooperative promoter recognition by human SNAPc. J. Biol. Chem. **281:**31050–31060.
18. **Jawdekar, G. W., and R. W. Henry.** 2008. Transcriptional regulation of human small nuclear RNA genes. Biochim. Biophys. Acta **1779:**295–305.
19. **Jensen, R. C., Y. Wang, S. B. Hardin, and W. E. Stumph.** 1998. The proximal sequence element (PSE) plays a major role in establishing the RNA polymerase specificity of *Drosophila* U-snRNA genes. Nucleic Acids Res. **26:** 616–622.
20. **Lai, H.-T., H. Chen, C. Li, K. J. McNamara-Schroeder, and W. E. Stumph.** 2005. The PSEA promoter element of the *Drosophila* U1 snRNA gene is sufficient to bring DmSNAPc into contact with 20 base pairs of downstream DNA. Nucleic Acids Res. **33:**6579–6586.
21. **Lai, H.-T., Y. S. Kang, and W. E. Stumph.** 2008. Subunit stoichiometry of the *Drosophila melanogaster* small nuclear RNA activating protein complex (SNAPc). FEBS Lett. **582:**3734–3738.
22. **Li, C., G. A. Harding, J. Parise, K. J. McNamara-Schroeder, and W. E. Stumph.** 2004. Architectural arrangement of cloned proximal sequence element-binding protein subunits on *Drosophila* U1 and U6 snRNA gene promoters. Mol. Cell. Biol. **24:**1897–1906.
23. **Lo, P. C. H., and S. M. Mount.** 1990. *Drosophila melanogaster* genes for U1 snRNA variants and their expression during development. Nucleic Acids Res. **18:**6971–6979.
24. **Lobo, S. M., and N. T. Hernandez.** 1994. Transcription of snRNA genes by RNA polymerases II and III, p. 127–159. *In* R. C. Conaway and J. W. Conaway (ed.), Transcription: mechanisms and regulation. Raven Press, New York, NY.
25. **Ma, B. C., and N. Hernandez.** 2002. Redundant cooperative interactions for assembly of a human U6 transcription initiation complex. Mol. Cell. Biol. **22:**8067–8078.
26. **McNamara-Schroeder, K. J., R. F. Hennessey, G. A. Harding, R. C. Jensen, and W. E. Stumph.** 2001. The *Drosophila* U1 and U6 gene proximal sequence elements act as important determinants of the RNA polymerase specificity of snRNA gene promoters in vitro and in vivo. J. Biol. Chem. **276:**31786–31792.
27. **Meijsing, S. H., M. A. Pufall, A. Y. So, D. L. Bates, L. Chen, and K. R. Yamamoto.** 2009. DNA binding site sequence directs glucocorticoid receptor structure and activity. Science **324:**407–410.
28. **Mittal, V., B. C. Ma, and N. Hernandez.** 1999. SNAPc: a core promoter factor with a built-in DNA-binding damper that is deactivated by the Oct-1 POU domain. Genes Dev. **13:**1807–1821.
29. **Parry, H. D., D. Scherly, and I. W. Mattaj.** 1989. 'Snurpogenesis': the transcription and assembly of U snRNP components. Trends Biochem. Sci. **14:**15–19.
30. **Schimanski, B., T. N. Nguyen, and A. Gunzl.** 2005. Characterization of a multisubunit transcription factor complex essential for spliced-leader RNA gene transcription in *Trypanosoma brucei*. Mol. Cell. Biol. **25:**7303–7313.
31. **Su, Y., Y. Song, Y. Wang, L. Jessop, L. C. Zhan, and W. E. Stumph.** 1997. Characterization of a *Drosophila* proximal-sequence-element-binding protein involved in transcription of small nuclear RNA genes. Eur. J. Biochem. **248:**231–237.
32. **Wang, Y., and W. E. Stumph.** 1998. Identification and topological arrangement of *Drosophila* proximal sequence element (PSE)-binding protein subunits that contact the PSEs of U1 and U6 snRNA genes. Mol. Cell. Biol. **18:**1570–1579.
33. **Wong, M. W., R. W. Henry, B. C. Ma, R. Kobayashi, N. Klages, P. Matthias, M. Strubin, and N. Hernandez.** 1998. The large subunit of basal transcription factor SNAPc is a Myb domain protein that interacts with Oct-1. Mol. Cell. Biol. **18:**368–377.
34. **Yoon, J. B., S. Murphy, L. Bai, Z. Wang, and R. G. Roeder.** 1995. Proximal sequence element-binding transcription factor (PTF) is a multisubunit complex required for transcription of both RNA polymerase II- and RNA polymerase III-dependent small nuclear RNA genes. Mol. Cell. Biol. **15:**2019–2027.
35. **Yoon, J. B., and R. G. Roeder.** 1996. Cloning of two proximal sequence element-binding transcription factor subunits (γ and δ) that are required for transcription of small nuclear RNA genes by RNA polymerases II and III and interact with the TATA-binding protein. Mol. Cell. Biol. **16:**1–9.