Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 2043–2047, March 1996
Biochemistry

Purification and physical properties of the male and female double sex proteins of *Drosophila*

(shape/oligomerization/DNA binding/transcription factors)

SAYEON CHO AND PIETER C. WENSINK*

Department of Biochemistry and The Rosenstiel Center, Brandeis University, Waltham, MA 02254-9110

Communicated by Gregory A. Petsko, Brandeis University, Waltham, MA, November 17, 1995 (received for review June 15, 1995)

ABSTRACT **The double sex gene (*dsx*) encodes two proteins, DSX^M and DSX^F, that regulate sex-specific transcription in *Drosophila*. These proteins bind target sites in DNA from which the male-specific DSX^M represses and the female-specific DSX^F activates transcription of yolk protein (*Yp*) genes. We investigated the physical properties of these DSX proteins, which are identical in their amino-terminal 397 residues but are entirely different in their carboxyl-terminal sequences (DSX^F, 30 amino acids; DSX^M, 152 amino acids). DSX^M and DSX^F were overexpressed in cultured insect cells and purified to near homogeneity. Gel filtration chromatography and glycerol gradient sedimentation showed that at low concentrations both proteins are dimers of highly asymmetrical shape. The axial ratios are ≈18:1 (DSX^M, 860 × 48 Å; DSX^F, 735 × 43 Å). At higher concentrations, the proteins form tetramers. Through use of a novel, double crosslinking assay (protein–DNA plus protein–protein), we demonstrated that a DNA regulatory site binds to both monomers of the DSX dimer and to only two monomers of the tetramer. Furthermore, binding another DNA molecule to what we presume is the second and identical site in the tetramer dramatically shifts the equilibrium from tetramers to dimers. These oligomerization and DNA binding properties are indistinguishable between the male and female proteins.**

Sexual phenotype in *Drosophila melanogaster* is regulated by a pathway of genes that produce alternative, sex-specific proteins as a result of regulated RNA splicing (1, 2). Most of these alternative female- and male-specific proteins regulate the sex-specificity of RNA splicing that occurs within the pathway itself (3, 4). However, one pair of these alternative sex-specific proteins, the DSX^M and DSX^F proteins encoded by the double sex (*dsx*) gene, do not regulate splicing but instead regulate sex-specific transcription of target genes that encode the sexual phenotype of the body (5–7). Thus, these two DSX proteins are the crucial and direct links between the pathway and its target genes.

The best characterized targets of DSX^F and DSX^M regulation are the yolk protein (*Yp*) genes (8, 9). The DSX proteins regulate *Yp* genes through the o-r enhancer which consists of four protein binding sites, one of which, dsxA, is responsible for the female specificity of transcription. DSX^M binds to this site in males and represses transcription, whereas DSX^F binds to the same site in females and activates transcription. Both proteins act by influencing adjacent transcription factors that are bound to the o-r enhancer and those adjacent factors, in turn, regulate transcription more directly (8, 9).

The regulatory functions of DSX proteins require sequences that are identical between the proteins and also those that are different. The amino-terminal 397 amino acids of the proteins are identical and contain the DNA binding domain necessary to the function of each protein (10). The remaining carboxyl-terminal extensions of 152 amino acids for DSX^M and 30 amino acids for DSX^F are entirely different from each other in sequence (7). These sex-specific sequences are responsible for the sex-specific transcriptional effects of each protein (8, 11).

This paper describes physical characteristics of the DSX proteins. We have overexpressed them in an insect cell culture system, purified each to near homogeneity, and then examined their shapes, oligomeric structures, and binding interactions with DNA.

MATERIALS AND METHODS

Expression Vectors. pBlueBac-DSX^M and pBlueBac-DSX^F were constructed by inserting a blunt-ended *Nde* I–*Hin*dIII cDNA from pT7-7:DSX^M or a *Nde* I–*Cla* I cDNA from pT7-7:DSX^F (10) into an *Nhe* I site of the pBlueBac baculovirus vector (Invitrogen) after modification to produce full length DSX proteins with no extra amino acids.

Protein Purification. Recombinant (AcMNPV-DSX^M or AcMNPV-DSX^F) baculovirus plaques were isolated after cotransfection with *Autographa californica* nuclear polyhedrosis virus (AcMNPV) and either pBlueBac-DSX^M or pBlueBac-DSX^F into *Spodoptera frugiperda* (Sf9) insect cells (12). To produce the proteins, 2×10^8 Sf9 cells were infected with recombinant baculovirus at a multiplicity of infection of 10, incubated for 66 hr in a spinner flask, and then harvested by centrifugation. Harvested cells were washed in ice-cold 137 mM NaCl/2.7 mM KCl/10 mM Na_2HPO_4/1.8 mM KH_2PO_4, pH 7.4, and resuspended in 4 ml of buffer I (10 mM Hepes, pH 7.6/2 mM $MgCl_2$/1 mM dithiothreitol/0.01 M KCl) containing 1 mM phenylmethylsulfonyl fluoride, 5 μg of aprotinin per ml, and 1 μg each of leupeptin, antipain, and pepstatin A per ml. After 30 min of incubation on ice, cells were homogenized in a Dounce homogenizer (pestle B). All subsequent procedures were done at 4°C. Nuclei were collected by centrifugation at $1000 \times g$ for 5 min and then resuspended in 3 ml of buffer II (25 mM Hepes, pH 7.6/0.1 mM EDTA/1 mM dithiothreitol) containing 0.5 M NaCl, 10 μM $ZnCl_2$, 10% (vol/vol) glycerol, and protease inhibitors as above. Resuspended nuclei were incubated for 30 min and then homogenized with Dounce pestle A. Insoluble nuclear material was removed by centrifugation (for 1 hr at 40,000 rpm in a Beckman SW 50.1 rotor). To the supernatant 0.6 volume of 80% ammonium sulfate in the same buffer was added. After 30 min of incubation with gentle agitation, precipitated protein was collected by centrifugation ($12,000 \times g$ for 40 min) and resuspended in buffer II containing 0.25 M NaCl and 20% glycerol.

The complementary DNA oligonucleotides, 5′-TCG**ACACAACTACAATGTTGCAATCAGC**TAGCC-3′ and 5′-TC**GAGGCTAGCTGATTGCAACATTGTAGTTGTG**-3′ were annealed to produce the high-affinity dsxA binding site of *Yp1* (boldface nucleotides; refs. 6, 8, and 9), phosphorylated, ligated, and coupled to CNBr-activated Sepharose 4B (Phar-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: BSA, bovine serum albumin.

*To whom reprint requests should be addressed.

macia LKB) (13). The resuspended ammonium sulfate fraction was applied to a 5-ml column of this DNA affinity resin. The column was washed with 2 column volumes of buffer II containing 0.25 M NaCl and 20% glycerol and then with the same volume of buffer II containing 0.3 M NaCl and 20% glycerol. DSX protein was eluted with buffer II containing 1 M NaCl and 20% glycerol, concentrated (Centriprep-30, Amicon), and stored at −70°C.

Antibodies. The purified full-length DSX^M protein was further purified on SDS/PAGE and used to immunize female rats (14).

Mobility-Shift Assay. The dsxA DNA fragment was labeled with [α-^{32}P]dCTP by Klenow polymerase (15). Assays were performed by incubating DSX proteins with 0.1–0.2 ng of ^{32}P end-labeled dsxA DNA for 30 min at room temperature in 20 μl of buffer II, 0.1 M NaCl, 10% glycerol, 100 μg/ml BSA followed by electrophoresis on non-denaturing gels and autoradiography (16, 17). Upper strand of the double-stranded non-specific competitor was 5′-GTTACCCGATGGATACTTAATAACC-3′.

Gel-Filtration Chromatography. Purified DSX protein [0.2–1 μg in 200 μl of the running buffer (buffer II containing 0.1 M NaCl and 10% glycerol)] was applied to a 37 × 1 cm column of Sephacryl S-300 preequilibrated with the running buffer (18). Of each eluted fraction, 18 μl was added to 2 μl of buffer II containing 10% glycerol, 0.1 M NaCl, 0.1 ng of ^{32}P-labeled dsxA DNA, and 1 mg of bovine serum albumin (BSA) per ml. This small dilution minimized the oligomeric state change between the chromatography and binding reactions. The reaction mixtures were assayed by mobility shifts. Protein standards were chromatographed on the same column and assayed by SDS/PAGE and copper staining.

Glycerol-Gradient Sedimentation. Purified DSX protein (0.1–0.5 μg) and 20 μg of aldolase in 100 μl of the running buffer were layered directly on a 4.8-ml gradient of 15–30% glycerol in buffer II containing 0.1 M NaCl, 1 μg of both leupeptin and pepstatin A per ml, and 1 mM phenylmethylsulfonyl fluoride. Centrifugation was performed as described (18). Fractions (≈120 μl) were collected and assayed as for gel filtration.

Estimates of Native Molecular Weight and Shape. The molecular weight and frictional axial ratio of each DSX protein were estimated by using the Stokes radii and sedimentation coefficients determined from hydrodynamic experiments (19). Partial specific volumes for DSX^M (0.710 ml/g) and DSX^F (0.705 ml/g) used in determining the native molecular weights were calculated from the amino acid sequences (20).

Combined Glutaraldehyde and UV Crosslinking. 5′-TCGACACAACTACAATGTTGCAATCAGCTAGCC-3′ was hybridized to a 14-nucleotide primer (5′-GGCTAGCTGATTGC-3′) and labeled with bromodeoxyuridine (BrdUrd) and [α-^{32}P]dCTP as described (21). DNA binding reactions (0.1 μg of purified DSX protein) were performed as in the mobility-shift assay except without BSA. After binding, glutaraldehyde was added to the final concentrations indicated and incubated for 15 min at room temperature. After electrophoresis of reaction mixtures on non-denaturing gels, the gels were placed over Saran Wrap on the filter surface of Fotodyne 305-nm UV transilluminator at room temperature and irradiated for the times indicated. After autoradiography, shifted bands were excised, soaked in 2× SDS/PAGE sample dye for 10 min, boiled for 3 min, and analyzed by 5–10% linear gradient SDS/PAGE followed by autoradiography.

RESULTS

Expression and Purification of DSX^M and DSX^F. The two DSX proteins were expressed in Sf9 insect cells using a baculovirus expression vector. Over-expressed DSX^M and DSX^F proteins were detected only in nuclear fractions (Fig. 1), an observation in keeping with the role of the proteins in transcriptional regulation. Both proteins appear unusually sensitive to proteases (data not shown), leading us to optimize the yield of full-length protein by harvesting cells midway through the viral infection cycle and selectively precipitating full-length DSX protein from nuclear extracts with ammonium sulfate (*Materials and Methods*). The highly specific binding of DSX proteins to the dsxA DNA regulatory site from the *Yp* genes allowed DNA affinity chromatography to be used to purify the proteins essentially to homogeneity. Typically 0.1 mg of DSX^M or DSX^F protein was purified from 2×10^8 cells with yields of 30–40% (Fig. 1).

Oligomerization of Both Proteins Is a Function of DNA and Protein Concentration. Mobility-shift experiments that assayed binding to dsxA DNA indicated that several different oligomeric forms of DSX^M and DSX^F occur. Low concentrations of either purified protein yielded a predominant and apparently homogeneous complex bound to dsxA (Fig. 2, lanes 1 and 6). Higher concentrations eliminated this complex and produced instead several complexes of lower mobility, including a major low-mobility (slow) complex (Fig. 2, lanes 1–3 and 6–8).

Experiments using unlabeled DNAs in excess of the radiolabeled, specific dsxA showed that DNA concentration influences the equilibrium between slow and fast complexes. The slow complex of DSX^F or DSX^M dissociated into the fast complex in the presence of an excess of a nonspecific DNA molecule the same size as dsxA (Fig. 2, lanes 5 and 10), poly(dI–dC)·poly(dI–dC), salmon sperm DNA, or several other nonspecific DNAs (data not shown). The ratio of fast to slow complex also increased as the concentration of specific dsxA DNA was increased (data not shown).

Competition with unlabeled dsxA DNA showed that in each complex at least one binding site of the protein binds DNA specifically. Excess unlabeled dsxA competes with labeled dsxA, whereas the same excess of unlabeled, nonspecific DNA affects the oligomerization state of DSX protein but does not compete with dsxA for binding (Fig. 2, lanes 4, 5, 9, and 10). The length and sequence of nonspecific DNA competitor had no effect (data not shown). This suggests that when the slow, presumably higher order oligomer, binds additional DNA, whether specific or nonspecific, the specific binding site remains, but the equilibrium shifts toward the fast complex.

We conclude that the concentration-dependent oligomerization properties of the female and male proteins are very similar and are likely to be due to the 397-amino acid region that is common between them. Furthermore, binding interactions with specific and nonspecific DNA appear to be complex but very similar between the two proteins and, therefore, are also likely to be due to the common sequence region.

In the Absence of DNA, Both Proteins Form Dimers of Highly Asymmetric Shape at Low Concentration. To investi-

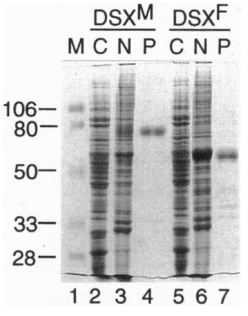

FIG. 1. DSX^M and DSX^F from baculovirus-infected Sf9 insect cells. Coomassie blue-stained protein after electrophoresis in SDS/10% PAGE gels. Lanes: 1, molecular masses (in kDa) of standards (shown to the left); 2–4, cytoplasmic (lane 2), nuclear (lane 3), and purified DSX^M (lane 4) protein from expressing cells; 5–7, cytoplasmic (lane 5), nuclear (lane 6), and purified DSX^F (lane 7) protein from expressing cells.

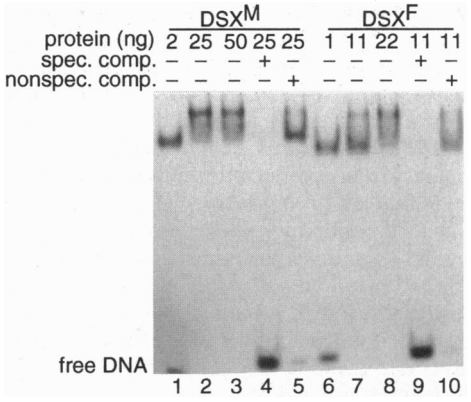

FIG. 2. Mobility-shift assay of binding to a specific DNA site. The indicated amounts of DSX^M (lanes 1–5) and DSX^F (lanes 6–10) were incubated with 0.1 ng of ^{32}P-labeled dsxA DNA in the absence of competitor (lanes 1–3 and 6–8) or in the presence of 50 ng of unlabeled dsxA (lanes 4 and 9) or nonspecific competitor DNA (lanes 5 and 10). Nonspecific competitor was a double-stranded molecule of the same length as the dsxA-containing molecule (*Materials and Methods*).

gate the structure of the fast complexes of DSX^M and DSX^F, we examined their sizes and shapes when no DNA was present. This was accomplished by gel-filtration chromatography and glycerol-gradient sedimentation at low protein concentrations. These concentrations produced only the fast complex with DNA when assayed by mobility-shift experiments and only one peak in the chromatographic and sedimentation assays when no DNA was present. Analysis of the gel filtration data yielded Stokes radii of 62.5 and 55.5 Å for DSX^M and DSX^F, respectively (Fig. 3 *Left*). Sedimentation coefficients of 5.1 and 4.3 S were determined for DSX^M and DSX^F, respectively (Fig. 3 *Right*).

The fast complexes are dimers. The native molecular weights of the fast complexes as estimated from the hydrodynamic data are 124 and 91 kDa for the male and female proteins, respectively (see *Materials and Methods*). To determine the subunit composition of these complexes, we estimated monomer molecular weights from the amino acid sequences. We chose this method rather than estimating from mobility in denaturing SDS/PAGE gels because, as shown previously, even when expressed in *E. coli*, both of these proteins migrate abnormally in such gels (6). The molecular weights of DSX^M and DSX^F monomers calculated from sequence are 58.5 and 45.8 kDa. Therefore, we conclude that the fast complexes detected in our assays are dimers.

Using these estimates of molecular weights and Stokes radii, we calculated frictional ratios of 1.91 for the DSX^M dimer and 1.88 for the DSX^F dimer, indicating that both are highly asymmetric (19). If these frictional ratios are used to model protein shapes as prolate ellipsoids of revolution, we calculate axial ratios of 18:1 for the DSX^M dimer and 17:1 for the DSX^F dimer (22). From these axial ratios and the molecular volumes based on Stokes radii, we estimate the dimensions of the DSX^M dimer as 860 × 48 Å and the DSX^F dimer as 735 × 43 Å. We conclude that both dimers are highly asymmetric and appear to have essentially the same shape—long cylinders with a diameter of ≈45 Å. The two proteins differ in cylinder length, with the higher molecular weight protein from males being the longer cylinder.

In the Absence of DNA, Both Proteins Form Tetramers and Higher Order Oligomers at High Concentrations. We examined the subunit stoichiometry of the slow complex by glutaraldehyde crosslinking in the absence of DNA. At protein concentrations yielding predominantly the slow complex with DNA when assayed by mobility-shift experiments (Fig. 2), each protein was incubated with different concentrations of glutaraldehyde and no DNA. The molecular weights of the resulting crosslinked molecules were assayed by Western blots after electrophoresis on SDS/PAGE gels. The molecular weights observed correspond to dimers, some intermediate forms and tetramers of DSX^M and DSX^F with an increasing proportion of the higher order oligomers as glutaraldehyde concentration was increased (Fig. 4). These oligomers appear to result from crosslinks within single DSX complexes, because adding an equal concentration of carbonic anhydrase yielded no detectable DSX–carbonic anhydrase complex (data not shown). These results strongly indicate that tetramers of each protein occur in the absence of DNA.

As an independent assay for the oligomeric structures of the DSX proteins, we performed pore exclusion-limit gel electrophoresis using ^{32}P-labeled DSX^M (ref. 23; data not shown). We detected two major and several minor lower mobility bands in these gels. This suggests that the protein forms an oligomeric series which yields two major complexes corresponding to the two oligomeric forms characterized above, and several higher order oligomers.

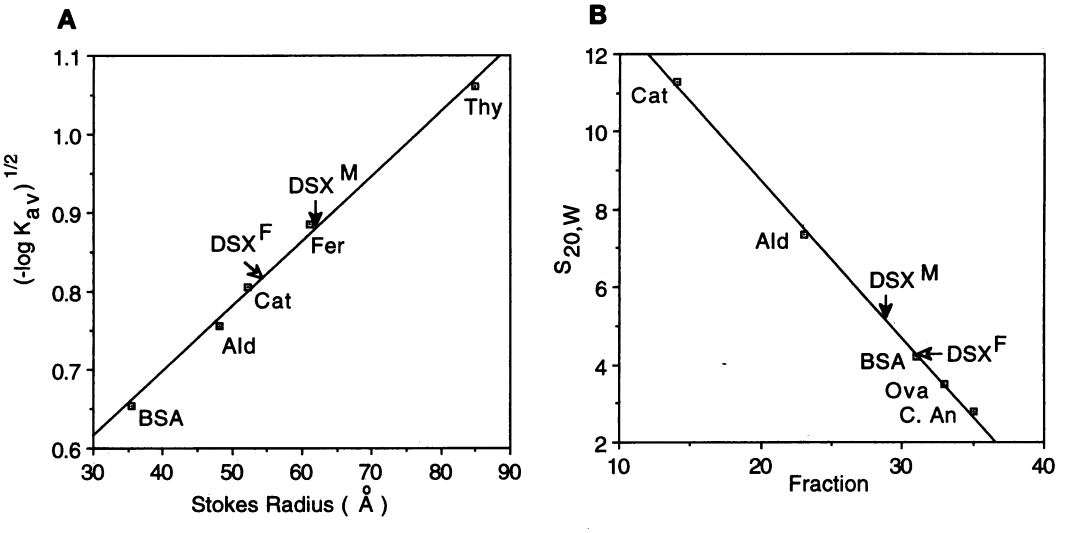

FIG. 3. Gel-filtration and glycerol-gradient sedimentation of DSX^M and DSX^F. (*Left*) Stokes radii of purified DSX^M and DSX^F determined by gel filtration. Standards run in parallel were as follows: Thy, thyroglobulin (85 Å); Fer, ferritin (61 Å); Cat, catalase (52.2 Å); Ald, aldolase (48.1 Å); BSA, (35.5 Å). Positions of DSX^M and DSX^F are indicated by arrows. (*Right*) Sedimentation coefficients of DSX^M and DSX^F determined by glycerol-gradient sedimentation. Standards sedimented in parallel were as follows: Cat, catalase (11.3 S); Ald, aldolase (7.35 S); BSA (4.2 S); Ova, ovalbumin (3.5 S); and C. An, carbonic anhydrase (2.8 S).

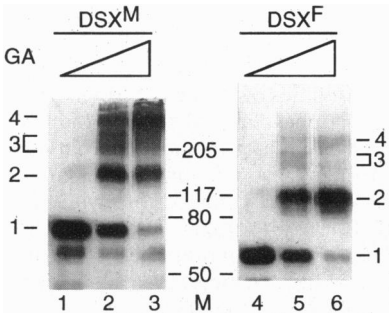

FIG. 4. Oligomerization of the DSX proteins. Western blot of SDS/5–10% PAGE gels of DSX^M (100 ng) (lanes 1–3) and DSX^F (30 ng) (lanes 4–6) crosslinked with 0% (lanes 1 and 4), 0.01% (lanes 2 and 5), or 0.1% (lanes 3 and 6) glutaraldehyde in the absence of BSA and DNA for 25 min at room temperature. DSX protein was detected with rat polyclonal antibody raised against DSX^M, peroxidase-conjugated anti-rat secondary antibody, and the enhanced chemiluminescence detection system (Amersham). Numbers to the left and the right indicate our interpretation of the number of crosslinked monomers. Molecular mass standards (in kDa) are shown between lanes 3 and 4. The lower concentration of DSX^F relative to DSX^M in this experiment reduces the concentration of DSX^F tetramers that can be crosslinked (lanes 3 and 6).

Based on results from these two experimental approaches, we conclude that in the absence of DNA ligand, DSX^M and DSX^F form dimers, tetramers, and higher order oligomers. In the concentration range we tested, the major complexes are dimers and tetramers.

Both Slow and Fast Complexes with DNA Appear to Have Only Two Protein Subunits in Direct Contact with a Single DNA Molecule. We examined the number of DSX monomers in contact with a single dsxA molecule, using mobility-shift/UV-crosslinking experiments. The dsxA DNA was labeled with BrdUrd and [α-^{32}P]dCTP and then incubated with DSX^M or DSX^F. Slow and fast protein–DNA complexes were separated in a nondenaturing polyacrylamide gel and then illuminated with UV light to crosslink protein to the BrdUrd-labeled DNA. The complexes were then excised from the gel and analyzed by SDS/PAGE.

The data shown in Fig. 5 indicate that only two monomers in the slow and fast complexes of DSX^M are in close contact with DNA. Brief UV crosslinking between dsxA and protein in the fast complex revealed only one crosslinked complex (Fig. 5 *Left*, lane 1). By comparison with protein standards, the molecular weight of this crosslinked complex is 80 kDa, the approximate sum of the molecular weights of the DSX^M monomer and one strand of dsxA DNA. Increased UV crosslinking of the fast complex produced a second crosslinked complex of ≈150 kDa (Fig. 5 *Left*, lanes 2–4). The highest level of crosslinking produced no additional complexes with either DSX^M or DSX^F, indicating that for each protein no more than two monomers are in contact with the same molecule of dsxA DNA (Fig. 5 *Center* and *Right*, lanes 1). Compared to the fast complexes, UV crosslinking within the slow complex of DSX^M or DSX^F gave essentially the same crosslinking pattern between dsxA and two monomers (Fig. 5 *Center* and *Right*, lanes 4), but a third and fourth monomer were also crosslinked to DNA, although at considerably lower efficiency than the first and second (Fig. 5 *Center* and *Right*, lanes 4). This suggests that only two monomers are in direct contact with a single molecule of DNA in both the slow and fast complexes. Further, it suggests that two other monomers in the slow complex are near that DNA molecule but are unlikely to be in direct contact.

Slow and Fast Complexes with DNA Correspond to Tetramers and Dimers of the Proteins. To examine the total number of monomers in each complex, we developed a method that introduces both protein–protein and protein–DNA crosslinks into a single complex. BrdUrd- and ^{32}P-labeled dsxA was incubated with either DSX^M or DSX^F and, after binding was complete, the resulting complexes were incubated with different concentrations of glutaraldehyde to produce different levels of crosslinking between protein monomers within a complex. After fractionating the resulting complexes by mobility shifts, the gel was illuminated with UV to crosslink DSX monomers to the dsxA DNA in the complex. Each protein–DNA complex was then excised and examined by SDS/PAGE.

When the UV crosslinking between dsxA and DSX protein is kept constant, increased protein crosslinking increases the number of monomers crosslinked to DNA up to a limit (Fig. 5 *Center* and *Right*). Without protein crosslinking, slow and fast complexes have predominantly one and two monomers crosslinked to DNA. As protein crosslinking is increased, the ratio between two and one monomer covalently linked to a DNA molecule in the fast complex increases and reaches a limit of two. The slow complex reaches a limit of four. This indicates that there are two monomers in the fast complex and four in the slow.

We conclude that the major oligomeric forms of the male and female proteins are the same whether free or bound to DNA. Furthermore, both dimers and tetramers of each protein can bind DNA specifically.

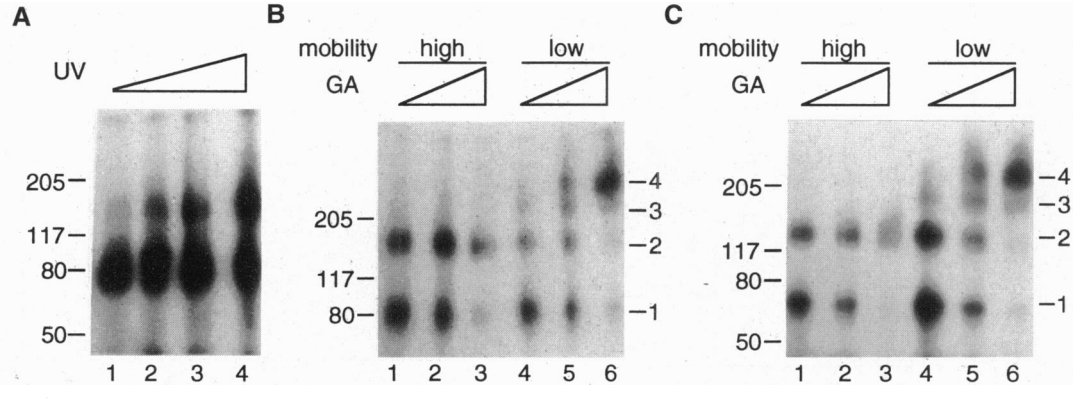

FIG. 5. Stoichiometry of complexes between DSX protein and DNA. (*Left*) An autoradiogram of UV-crosslinked molecules from high-mobility complexes between DSX^M and BrdUrd/^{32}P-labeled dsxA DNA. Complexes illuminated by UV light for 2 (lane 1), 5 (lane 2), 10 (lane 3), or 25 min (lane 4) are shown electrophoretically separated in a SDS/PAGE gel. (*Center* and *Right*) DSX^M (*Center*) and DSX^F (*Right*) are autoradiograms of UV- and glutaraldehyde-crosslinked complexes between BrdUrd/^{32}P-labeled dsxA and DSX protein. High (lanes 1–3) and low (lanes 4–6) mobility complexes crosslinked with 0% (lanes 1 and 4), 0.01% (lanes 2 and 5), or 0.1% (lanes 3 and 6) glutaraldehyde and then illuminated with UV light for 25 min are shown electrophoretically separated on SDS/PAGE gels. Molecular mass standards (in kDa) are shown on the left of each panel, and our interpretation of the number of monomers crosslinked is shown on the right.

DISCUSSION

We have investigated physical properties of two sex-specific transcription factors, DSX^M and DSX^F, alternative products of the double sex gene. The results indicate that the two proteins have very similar and often indistinguishable physical properties. First, each forms a concentration-dependent series of oligomers, with dimers and tetramers being the predominant complexes over the concentration range assayed, but with higher order oligomers also occurring. Second, the effect of protein concentration on equilibria between these forms appears to be indistinguishable between DSX^M and DSX^F. Third, the predominant DNA binding complexes of each protein also are dimers and tetramers. Fourth, when either male or female tetramer binds additional DNA, it dissociates into dimers. Finally, the shapes of the two proteins are essentially the same and highly asymmetric.

The general similarity between these two proteins is perhaps not surprising since they are identical in their amino-terminal 397 amino acids and differ only in their relatively short carboxyl termini (30 and 152 amino acids for DSX^F and DSX^M, respectively). The general similarity in physical properties sets limits on the mechanisms by which the two proteins and their smaller, sex-specific domains direct such extremely different transcriptional regulation. As cited above, these sex-specific domains alone are responsible for the sex-specific regulation of each protein (11). Based on results presented in this paper, we conclude that these domains have no effect on the overall physical properties of the entire protein—neither the general shape nor the ability to form oligomers or to bind a single DNA site. Instead, the sex-specific domains are likely to have effects independent of the remainder of the protein. These effects are more likely to be changes in the interactions with other regulatory proteins or in the cooperative binding of DSX protein to multiple DNA sites. The latter possibility may be relevant because multiple, closely spaced DSX binding sites have been observed in the only molecularly characterized target of DSX protein regulation (6).

The simplest interpretation of the oligomerization and DNA binding data and the previous observation that the consensus sequence for specific DNA ligands is palindromic (6, 9) is that the protein dimer forms a single DNA binding site with high affinity for specific regulatory DNA as well as lower affinity for nonspecific DNA. Further, in this interpretation, all DNA binding interactions and consequent effects on protein oligomerization are due to the single copy of this site in dimers and the pair of copies in tetramers. A DNA binding region of the DSX polypeptides that is a good candidate for this has previously been localized within the shared sequence (10). Presumably, two of these regions are necessary to bind the palindromic DNA sequence with useful affinity.

A hypothesis for the DNA binding and protein equilibria is diagrammed in Fig. 6. Dimers and tetramers of each protein bind to DNA with high DNA sequence specificity (ref. 9; also this paper). Tetramers also have at least one site that binds DNA in a way that is mutually exclusive with tetramer stability. Several explanations for the effect of this second site on protein oligomerization are likely. First, two identical DNA binding sites in a tetramer, one site in each of the component dimers, may be so close together that binding a second DNA fragment is sterically excluded, as has been proposed for p53 (24). Alternatively, binding the second identical site or some non-identical site may induce a conformational change in the protein that is mutually exclusive with tetramer stability.

The shapes of the two proteins leads us to speculate upon the arrangement of shared and different functions within the proteins. Our results show that dimers of DSX^M and DSX^F are very similar in hydrodynamic shape, both having unusual axial ratios of ≈18:1 but differing in the length of the long axis (860 × 48 Å and 735 × 43 Å for DSX^M and DSX^F, respectively). These shapes indicate that the proteins are highly extended. For comparison, the highly extended coiled-coil form of α-helical structure has an axial spacing of 1.5 Å per residue, similar to that we observe with DSX proteins (1.6 and 1.7 Å per residue for DSX^M and DSX^F, respectively). This suggests that the sex-specific, carboxyl-terminal domain is well separated from most of the shared sequence and, in particular, from the other known functional region of the protein, the DNA binding domain. This arrangement fits well with our observations that the sex-specific domains have little or no effect on the DNA binding and oligomerization properties and that the sex-specific domains are more likely to be involved in interactions with different regulatory proteins or with cooperative binding to multiple DNA sites.

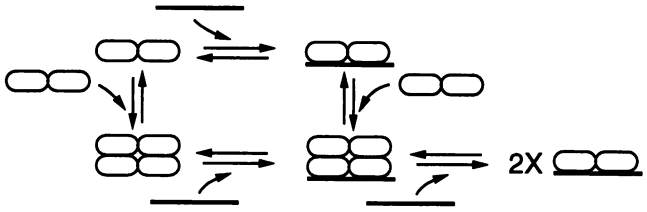

FIG. 6. Hypothesis for oligomerization and DNA-binding of DSX proteins. Dimers are symbolized by paired ovals and DNA by heavy lines. Equilibria are represented by oppositely oriented arrows.

We thank Hansraj Mahtani for help in learning baculovirus expression methods and for the gift of Sf9 cells. We thank Marie Lossky, Bob Kolouch, Jeff Gelles, Dagmar Ringe, and Yue Wang for helpful criticism of this manuscript. We gratefully acknowledge financial support by the National Institutes of Health (Research Grant GM R01 46327).

1. Belote, J. (1992) *Semin. Dev. Biol.* **3,** 319–330.
2. Burtis, K. C. (1993) *Curr. Opin. Cell Biol.* **5,** 1006–1014.
3. McKeown, M. (1992) *Curr. Opin. Genet. Dev.* **2,** 299–303.
4. Amrein, H., Hedley, M. L. & Maniatis, T. (1994) *Cell* **76,** 735–746.
5. Belote, J. M., Handler, A. M., Wolfner, M. F., Livak, K. J. & Baker, B. S. (1985) *Cell* **40,** 339–348.
6. Burtis, K. C., Coschigano, K. T., Baker, B. S. & Wensink, P. C. (1991) *EMBO J.* **10,** 2577–2582.
7. Burtis, K. C. & Baker, B. S. (1989) *Cell* **56,** 997–1010.
8. An, W. & Wensink, P. C. (1995) *Genes Dev.* **9,** 256–266.
9. An, W. & Wensink, P. C. (1995) *EMBO J.* **14,** 101–110.
10. Erdman, S. E. & Burtis, K. C. (1993) *EMBO J.* **12,** 527–535.
11. Coschigano, K. T. & Wensink, P. C. (1993) *Genes Dev.* **7,** 42–54.
12. Summers, M. D. & Smith, G. E. (1987) *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures* (Texas A&M, College Station, TX).
13. Kadonaga, J. T. & Tjian, R. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 5889–5893.
14. Harlow, E. & Lane, D. (1988) *Antibodies: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
15. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
16. Garner, M. M. & Revsin, A. (1981) *Nucleic Acids Res.* **9,** 3047–3060.
17. Carey, J. (1991) *Methods Enzymol.* **208,** 103–117.
18. Mitsis, P. G. & Wensink, P. C. (1989) *J. Biol. Chem.* **264,** 5195.
19. Siegel, L. M. & Monty, K. J. (1966) *Biochim. Biophys. Acta* **112,** 346–362.
20. Reynolds, J. A. & McCaslin, D. R. (1985) *Methods Enzymol.* **117,** 41–53.
21. O'Donnell, K. H. & Wensink, P. C. (1994) *Nucleic Acids Res.* **22,** 4712–4718.
22. Cantor, C. R. & Schimmel, P. R. (1980) *Biophysical Chemistry* (Freeman, San Francisco).
23. Sarge, K. D., Murphy, S. P. & Morimoto, A. B. (1993) *Mol. Cell. Biol.* **13,** 1392–1407.
24. Delphin, C., Cahen, P., Lawrence, J. J. & Baudier, J. (1994) *Eur. J. Biochem.* **223,** 683–692.