# Sry-Directed Sex Reversal in Transgenic Mice Is Robust with Respect to Enhanced DNA Bending: Comparison of Human and Murine HMG Boxes†

Nelson B. Phillips,[‡,§] Tatiana Nikolskaya,[‡,§,||] Agnes Jancso-Radek,[‡,§,⊥] Varda Ittah,[@] Fashun Jiang,[‡,#] Rupinder Singh,[‡] Elisha Haas,[@] and Michael A. Weiss*[,‡]

*Department of Biochemistry, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106-4935, and Faculty of Life Sciences, Bar Ilan University, Ramat Gan, Israel 52900*

*Received January 9, 2004; Revised Manuscript Received March 19, 2004*

ABSTRACT: The testis-determining factor SRY contains an HMG box DNA-bending domain. Human and murine factors (hSRY and mSRY, respectively) exhibit marked sequence divergence and are reported to differ markedly in DNA bending properties. Surprisingly, the combined application of time-resolved fluorescence resonance energy transfer (tr-FRET) and permutation gel electrophoresis demonstrates that the hSRY−DNA complex is more sharply bent than the murine complex and not less bent as previously reported. tr-FRET-based analyses of the distribution of end-to-end distances in the bent DNA−protein complexes further suggest that a broader range of DNA bend angles is populated in the murine ensemble than in the human ensemble. The two domains and their respective DNA complexes nevertheless exhibit similar thermodynamic stabilities. $^1$H NMR spectra indicate analogous intercalation of distinct "cantilever" side chains (isoleucine or methionine) with subtle differences in induced DNA structure. Interchange of cantilevers does not affect DNA bending. That transgenic expression of either human or murine *Sry* in XX mice can confer a male somatic phenotype suggests that SRY-directed transcriptional regulation is robust to enhanced DNA bending and to changes in the precision of DNA bending. We propose that male-specific gene regulation requires DNA bending above a critical threshold set by architectural requirements of enhanceosome assembly.

Testis determination in eutherian mammals is directed by *Sry*, a gene on the Y chromosome (*1*, *2*). Mutations in SRY[1] are associated with human sex reversal (46, XY pure gonadal dysgenesis) and cluster in its DNA-binding domain (*3*, *4*), a high-mobility group (HMG) box (*5*). The HMG box defines a class of DNA-bending proteins of central importance in organogenesis (*6*−*9*). Designated architectural transcription factors, such proteins exhibit an L-shaped structure (Figure 1A; *10*−*12*) and novel mechanism of DNA recognition (Figure 1B−D; *4*, *13*, *14*). The structure of a specific SRY−DNA complex demonstrates that the domain binds in an expanded minor groove and induces a sharp DNA bend (*4*, *15*).[2] Bending is associated with partial unwinding and insertion of a "cantilever" side chain between base pairs as shown in schematic form in Figure 1B, disrupting base stacking but not base pairing (I13 in Figure 1C; *4*, *13*). DNA-binding surfaces are extended by a C-terminal basic tail, which remains in the minor groove at one end of the bent DNA site (Figure 1C,D; *15*).

Human and murine SRY (hSRY and mSRY, respectively) exhibit marked sequence divergence both within and extrinsic to the HMG box (*16*). Substitutions occur at 28 of 85 positions in the HMG box (Figure 2), spanning its DNA-binding surface (red boxes), other surfaces (green), and the hydrophobic core (purple). The structure of mSRY has not been characterized. Grosschedl and co-workers reported that the hSRY domain bent a consensus SRY target site (5′-

† This work was supported in part by an Equipment Grant 553-99 from the Israel Science Foundation to E.H., Grant 98-362 from the United States/Israel Binational Foundation (E.H. and M.A.W.), and the National Institutes of Health Grant GM051558 (M.A.W.). V.I. and E.H. were supported in part by the Damadian Center for Magnetic Resonance Research at Bar Ilan University.

* To whom correspondence should be addressed. Telephone: (216) 368-5991. Fax: (216) 368-3419. E-mail: michael.weiss@case.edu.

‡ Case Western Reserve University.

§ These authors contributed equally to this work.

|| Present address: GeneGo, Inc., 227 S. Berrien St., New Buffalo, MI 49117.

⊥ Present address: Epicentre Technologies Corp., 726 Post Rd., Madison, WI 53713.

@ Bar Ilan University.

# Present address: 386 S. Crown Court, Palatine, IL 60074.

[1] Abbreviations: CD, circular dichroism; F−B, DNA bending model of Bianchi and colleagues (*37*); DTT, dithiothreitol; Fl, 5-acetamidofluorescein; FRET, fluorescence resonance energy transfer; fwhm, full width at half-maximal intensity; HMG, high-mobility group; GMSA, gel mobility shift assay; LEF-1, lymphoid enhancer factor 1; MIS, Mullerian-Inhibiting Substance (also termed Anti-Mullerian Hormone); NLS, nuclear localization signal; NOE, nuclear Overhauser effect; *orf*, open reading frame; PGE, permutation gel electrophoresis; RP-HPLC, reverse-phase high-performance liquid chromatography; $^1$H NMR, proton nuclear magnetic resonance; SOX, SRY-related HMG box; SRY, protein encoded by the sex-determining region of the Y chromosome; TBP, TATA-binding protein; TCF-1α, T-cell factor 1α and homologue of LEF-1; T−L, DNA bending model of Thompson and Landy (*66*); tr, time-resolved; TMR, tetramethylrhodamine. Amino acids are designated by the standard single-letter code. Unless otherwise stated, residue numbers in SRY or LEF-1 refer to the HMG box consensus sequence as defined by Clore and co-workers (*4*). Consensus position 1 corresponds to residue 56 of intact human SRY. Genes and DNA sites are in italics and proteins in capital letters.

[2] A similar DNA structure is induced upon binding of lymphoid enhancer factor (LEF-1; *14*) and non-sequence-specific HMG boxes (HMG1, NHP6A, and HMG-D; *38*−*40*). Unlike SRY, the tails of LEF-1 and NHP6A bind across a compressed major groove.

10.1021/bi049920a CCC: $27.50 
Published on Web 05/14/2004

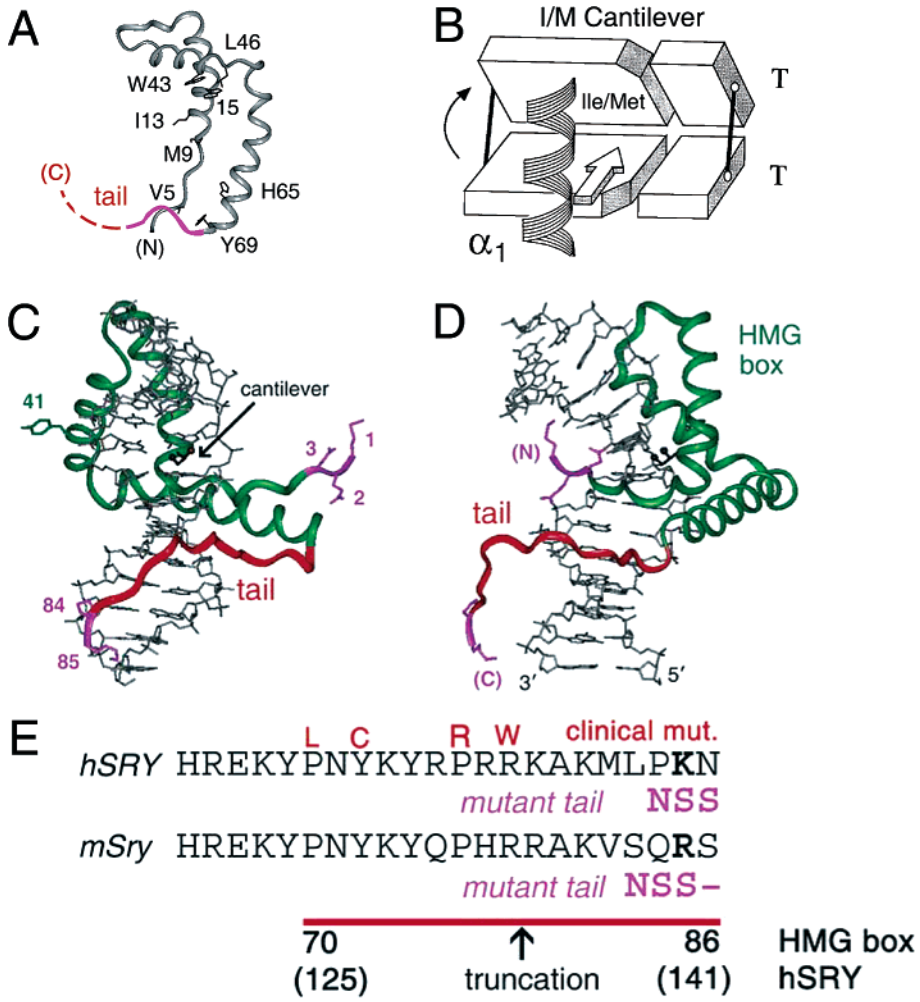


FIGURE 1: Human SRY−DNA complex and tail sequences. (A) Ribbon model of the SRY HMG box with selected side chains labeled. The position of the disordered tail is shown as a dashed red line. (B) Schematic model of insertion of the cantilever side chain between AT base pairs, disrupting base stacking but not base pairing. (C and D) Ribbon models of the NMR structure (*15*; Protein Data Bank entry 1J46) giving front and side views related by a 90° rotation about the vertical axis. Purple highlights positions of vector-derived substitutions in SRY fragments employed in an earlier comparison of hSRY and mSRY (*17*). The C-terminal tail is shown in red (residues 70−83) or purple (residues 84 and 85); N-terminal residues 1−3 distant from DNA are shown in purple. The HMG box is otherwise shown in green, and DNA is shown as gray sticks. Cantilever side chain I13 is shown in black (arrow in panel C). Y41, a nonconserved site on the back surface of human SRY, is substituted with cysteine in murine SRY (see Figure 2) and with serine in the model murine domain studied here. (E) Tail sequences of hSRY and mSRY indicating inadvertent vector-derived NSS substitutions (purple) in the hSRY fragment employed in the prior study (*17*). Sites of human sex-reversal mutations in the tail are shown above the sequence in red [P70L (*62*), Y72C (*63*), P76R (*64*), and R78W (*65*)]. The red bar delimits the tail. Residue numbers refer to the HMG box consensus or (in parentheses below) full-length hSRY. The arrow denotes the truncation site in SRY employed in prior studies by others (*4*, *18*, *37*).

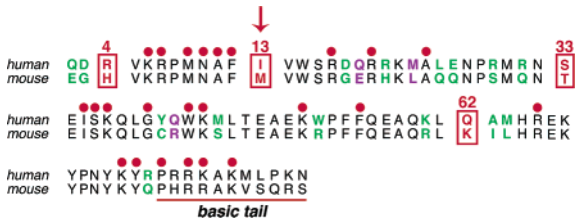


FIGURE 2: Sequences of human and murine SRY HMG boxes. Conserved residues are shown in black, substitutions at the DNA-binding surface in red (circles and boxes), other surface substitutions in green, and substitutions in the adjoining hydrophobic core in purple. The arrow (red) denotes cantilever side chain I13 or M13. The basic tail is underlined in red. In ref *15*, R21 is also described as contacting the DNA backbone, but analysis of the coordinates suggests it is too far from the DNA; consequently, R21 (human) and H21 (murine) are shown in green.

ATTGTT-3′ and complement[3]) 25−30° less sharply than did the mSRY domain (*17*). These findings rationalized the seeming inability of hSRY to cause sex reversal in XX transgenic mice (*6*). The reported correlation between impaired testicular differentiation in transgenic mice and inhibited DNA bending appeared to be consistent with a case of human sex reversal associated with an SRY variant defective in DNA bending (*18*).

Comparison between the DNA bending properties of hSRY and mSRY is of renewed biological interest in light of a recent reappraisal of the transgenic experiments (*19*, *20*). In the original study, transgenic expression of the mouse allele, but not human *Sry*, was found to be sufficient to induce male somatic differentiation (*6*). The *mSRY* open reading frame (*orf*) is contained within a 14 kb DNA segment (Figure 3, top), which contains appropriate *cis*-acting regulatory

[3] The consensus DNA target site for SRY has been defined by random binding site selection (*67*), in accord with mutagenesis of candidate target elements uncovered by footprinting promoter and enhancer elements (*3*, *17*, *36*).

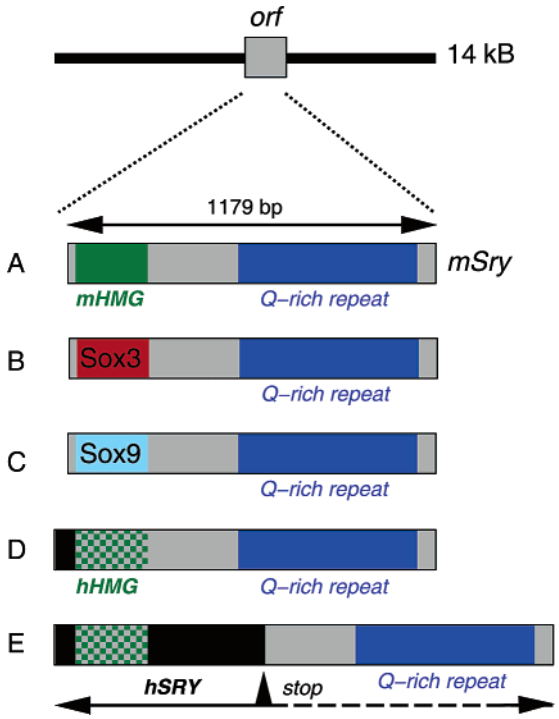


FIGURE 3: Transgenic studies of Sry-directed sex reversal in XX mice. The male somatic phenotype is induced by a 14 kb genomic fragment of the Y chromosome containing the *Sry* open reading frame (*orf*, top panel; *6*). The domain structure of mSRY is delineated in panel A: N-terminal HMG box (green) and C-terminal glutamine-rich region due to CAG repeat (blue). (B and C) Chimeric *mSry* constructions encoding the domain swap of the mSRY HMG box with those of mSOX3 (B, red) or mSOX9 (C, powder blue; *19*). In chimeric construction D, the N-terminal portion of the *mSry* HMG box is replaced with that of *hSry*, including coding regions for N-terminal nonbox sequences (black) and the HMG box (green checkerboard; *20*). Chimeric construction E contains intact the *hSry orf*, including the coding region for the C-terminal nonbox sequences (black) and the 3′ stop codon (arrowhead), instead of the N-terminal segment of the *mSry orf*. The glutamine-rich domain of mSRY is thus not expressed (*20*). Chimeric constructions in each panel function in XX transgenic mice to induce testicular differentiation and male somatic development (*19*, *20*). Sequences of SOX3 and SOX9 HMG boxes are identical in human and murine genes.

elements for directing expression of *mSry* (Figure 3A) in the differentiating gonadal ridge. In the recent studies, "domain swap" of the HMG box of mSRY with that of mSOX3 or mSOX9 (Figure 3B,C) was shown not to impair male differentiation, indicating their functional equivalence in the context of mSRY (*19*).[4] Swap of murine and human HMG boxes and amino-terminal nonbox sequences (Figure 3D) likewise yields a chimeric protein that is able to direct testicular differentiation (*20*). Indeed, an identical phenotype was observed following replacement of the 5′-portion of the *mSry orf* by the *hSry* coding region, including its stop codon (Figure 3E). Expression of *hSry* under the regulatory control of murine *cis*-acting elements is thus sufficient for directing testicular differentiation.

This investigation of hSRY and mSRY employs time-resolved fluorescence resonance energy transfer (tr-FRET) and permutation gel electrophoresis (PGE) to demonstrate that the human protein−DNA complex is more sharply bent than the murine complex. tr-FRET-based analyses of end-to-end distances in the bent DNA complexes further demonstrate a salient difference in dynamics: a broader range of DNA bend angles is populated in the murine ensemble than in the human ensemble. The two domains nevertheless exhibit similar secondary structures, as inferred from circular dichroism (CD), and similar thermodynamic stabilities. Proton nuclear magnetic resonance ($^1$H NMR) spectra of free and bound DNA sites suggest that induced DNA structures share overall features, including analogous sites of cantilever insertion, but differ in structural details. To test whether the divergent cantilevers of hSRY and mSRY (I13 and M13, respectively; arrow in Figure 2) are responsible for the observed differences in the DNA bend angle, we constructed and investigated an M13I analogue of the mSRY HMG box. The substitution did not affect DNA bending or the stringency of sequence recognition. In light of the ability of hSRY to direct male development of XX transgenic mice (*20*), our results suggest that SRY-directed sex reversal in the mouse is robust with respect to enhanced DNA bending and to changes in the precision of DNA bending. The extent of enhancement (5−10° as inferred from PGE) is similar to the range of DNA bend angles observed among primate SRY−DNA complexes (*21*).[5] Accordingly, we propose that DNA bending within an allowed range of sharp angles can function to direct assembly of a male-specific enhanceosome.

## MATERIALS AND METHODS

*Protein Purification.* The murine SRY HMG box (residues 2−86 in the HMG box) and the corresponding human domain (*22*) were expressed in *Escherichia coli* as a thrombin-cleavable fusion protein as described previously (*22*, *23*). Fragments contained two additional N-terminal residues ($G_{-1}S_0$) derived from the thrombin site in the fusion protein. Initial purification utilized an N-terminal $His_6$ tag for $Ni^{2+}$ affinity chromatography. Further purification of the SRY HMG box fragment was accomplished by FPLC using a MonoS column (Pharmacia, Inc.) and reverse-phase HPLC. The purity was >98% as assessed by SDS−polyacrylamide gel electrophoresis. Results of mass spectrometry were in accord with expected values.

*Site-Directed Mutagenesis.* Single-amino acid substitutions were introduced into the SRY coding region in phage M13mp19RF by oligonucleotide-directed mutagenesis by polymerase chain reaction (PCR) and cloned into an expression plasmid (*22*, *23*). All constructions were verified by

[4] Rodent *Sry* coding regions contain a CAG DNA repeat encoding a novel glutamine-rich domain C-terminal to the HMG box (*68*; Figure 3A). Induction of sex reversal in XX transgenic mice by hSRY lacking this domain (Figure 3E) suggests that the glutamine-rich domain is not required for male-specific gene regulation (*20*). This conclusion is not in accord with transgenic studies of truncation mutants of *mSry* transgenes in which an intact glutamine-rich domain was apparently required for XX sex reversal (*69*). The possibility that such truncations led to proteolyic instability or aggregation of the variant proteins, thus preventing SRY-directed gene regulation, could not be excluded, however. It remains formally possible that mSRY functions through its glutamine-rich domain whereas hSRY functions through recruitment of an unrelated set of interacting proteins (*60*), coincidentally leading to identical phenotypes.

[5] Using SRY fragments truncated at HMG residue 81 (residue 135 in intact human SRY), Goodfellow and colleagues reported an 8° range in DNA bend angles among primates, from 82.3° in the pigmy chimpanzee to 90.3° in the marmoset. The hSRY complex yielded an angle of 89°; the lower value in the chimpanzee is due in part to substitution of the I13 cantilever with phenylalanine (*22*).

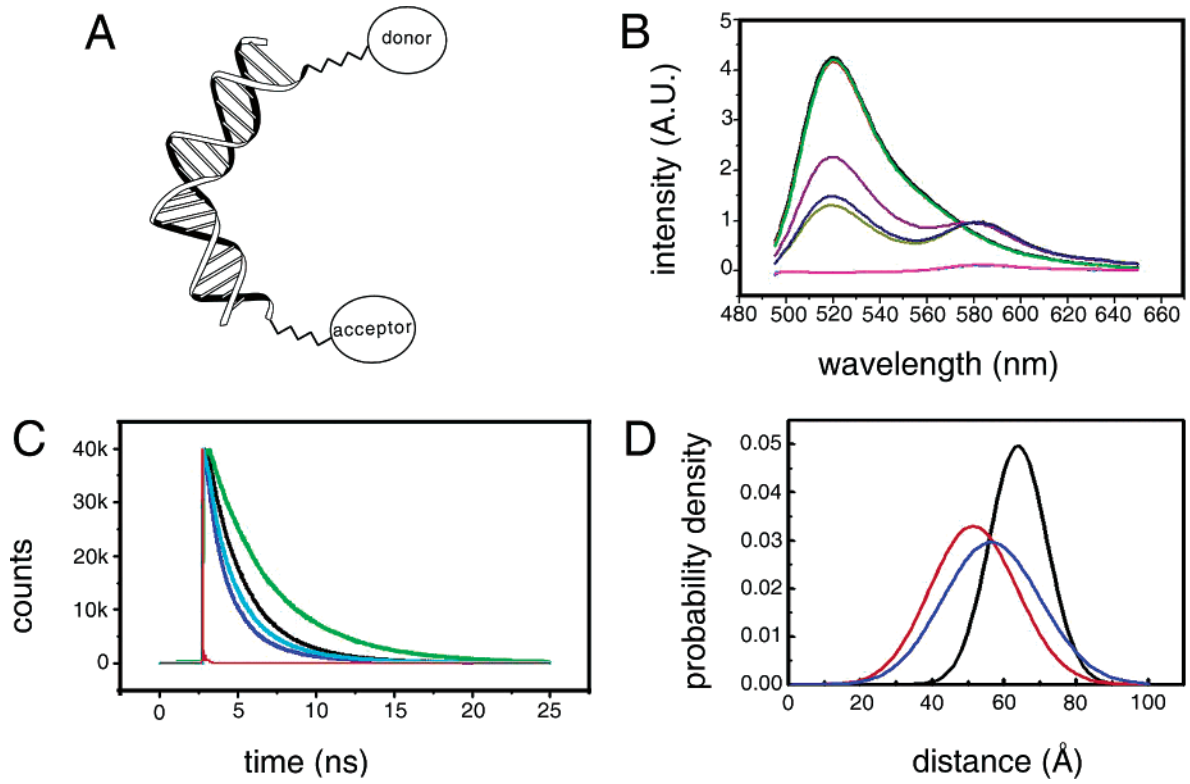


FIGURE 4: Design of FRET studies and steady-state spectra. (A) Schematic illustration of DNA probe design. A central DNA bend is expected to reduce the distance between respective 5′ ends of the DNA duplex. One end is labeled with fluorescein (donor) and the other with tetramethylrhodamine (acceptor). (B) Emission spectra of the free and bound DNA following excitation at 465 nm. Singly labeled DNA: black for emission of the donor in free DNA, red for the complex with hSRY, and green for the complex with mSRY (overlapping spectra). Doubly labeled DNA: purple for free DNA, olive for the hSRY complex, and blue for the mSRY complex. The emission spectra of the DNA singly labeled with the acceptor (free, complexed with hSRY, and complexed with mSRY) are shown in pink and underlays. The correspondence between these data and the tr-FRET confirms the analysis. (C) Fluorescence decay of the donor (fluorescein) bound to the 5′ end of the model DNA duplex under 490 nm excitation detected at 520 nm (8 nm bandwidth, channel width of 0.0061 ns/channel). Singly labeled DNA in the absence of an acceptor is shown in green; the donor emission decays in the doubly labeled DNA in the free form (black) and complexes with mSRY (cyan) and hSRY (blue) are shown. The reduction of the donor fluorescence lifetime corresponds to the changes of the emission spectra and is clearly a result of the distance-dependent FRET. (D) Gaussian models of end-to-end distances in the free and bound DNA sites. The end-to-end distance distributions of doubly labeled model DNA free (black) and in the complexes with mSRY (blue) and hSRY (red). The bending of the DNA by the complexation with the SRY domain is clearly visible in the reduction of the mean of the distributions.

DNA sequencing. Some murine SRY domains contain a C41S substitution on the back surface of the HMG box to avoid cysteine oxidation. Control studies of DNA binding and DNA bending indicate that the substitution is nonperturbing (see panels A and B of Figure 5).

*Electrophoretic DNA Binding Assay.* The 36 bp complementary oligonucleotides (5′-CATACTGCGGGGGT-GATTGTTCAGGATCATACTGCG-3′) containing the consensus ATTGTT sequence were purchased from Oligos, Etc. (Wilsonville, OR). One strand was labeled with [$\gamma$-$^{33}$P]ATP (ICN Biomedicals, Inc.). The duplex probe was annealed and analyzed using a gel retardation assay as described previously (*13*, *24*). Each reaction mixture contained 1−32 nM protein and <1 nM labeled DNA in 10 mM potassium phosphate (pH 7.0), 30 $\mu$g/mL bovine serum albumin (BSA), 50 mM KCl, 2.5 mM $MgCl_2$, 5% glycerol, and 4 mM dithiothreitol (DTT); the reaction mixture was incubated in a volume of 12.5 $\mu$L for 1 h on ice. Only specific binding is observed under these conditions at the protein concentrations that were tested. Previously prepared 0.45× TBE polyacrylamide gels (8%) were prerun for 45−60 min at 10 V/cm and 4 °C. After the sample was loaded, the gel was run at 15 V/cm for 1 h at 4 °C. The gel was dried and exposed overnight, and the gel image was visualized using a Molecular Dynamics phosphorimager. Relative band intensities were quantitated using ImageQuant version 5.2 (Molecular Dynamics).

*Electrophoretic DNA Bending Assay.* Double-stranded synthetic oligonucleotides (5′-GTGATTGTTCAG-3′) containing the consensus 5′-ATTGTT-3′ SRY binding site (underlined) were cloned between the *Xba*I and *Sal*I sites of the circular permutation vector pBend2 (*25*). Probes of equal length (bending probes 1−6), with the binding site located at varying distances from the ends, were generated by PCR, and 5′-labeled with [$^{33}$P]ATP and T4 polynucleotide kinase. The 10 $\mu$L binding reaction mixtures contained 50 mM KCl, 20 mM Tris-HCl (pH 7.4), 5 mM $MgCl_2$, 50 ng of poly(dI-dC), 10% glycerol, approximately 1 nM $^{33}$P-labeled DNA probe, and purified SRY HMG box protein (60 nM). After incubation on ice for 1−2 h, samples were run on polyacrylamide gels in 0.5× (0.045 M) TBE buffer at ≈10 V/cm.

*Circular Dichroism.* Spectra were recorded using an Aviv spectropolarimeter (Proterion Corp., Lakewood, NJ) equipped with thermister temperature control for automated analysis of thermal unfolding. The samples were concentrated to 30 $\mu$M in 140 mM KCl and 10 mM potassium phosphate (pH 7.4) and observed in a quartz cuvette with a path length of 1 mm.

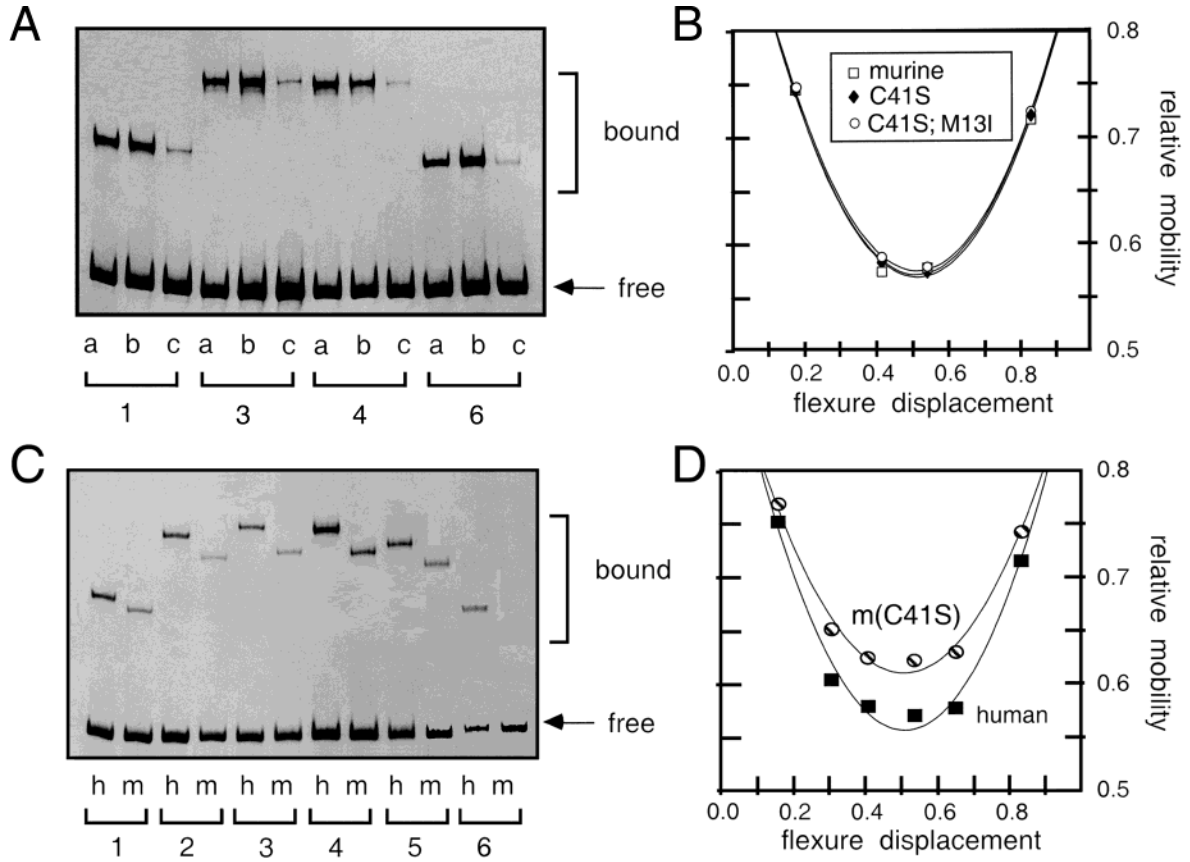


FIGURE 5: DNA bending studies of the murine HMG box and analogues. (A and C) Permutation gel electrophoresis (PGE) studies indicating similar specific DNA bending by wild-type mSRY HMG box and analogues C41S and C41S/M13I. (B and D) PGE studies indicating sharper DNA bending by hSRY than by the C41S analogue of mSRY. (A) PGE using bending probes 1, 3, 4, and 6 (see Materials and Methods): wild-type mSRY HMG box (lane a), C41S analogues (lane b), and C41S/M13I analogue (lane c). Studies employed DNA site 5′-AGGGCACCCATTGTTCTCTCC-3′ and its complement (site a in Figure 6B; core consensus sequence underlined) as described by Goodfellow and co-workers (*18*, *36*). Inferred DNA bend angles are indistinguishable (71.5 ± 0.5° in this gel system; 10% polyacrylamide with a 29:1 acrylamide:bisacrylamide ratio at 4 °C). (B) Plot of the relative electrophoretic mobility as function of flexure displacement in DNA probes: wild-type mSRY HMG box [murine (□)], C41S analogues (◆), and C41S/M13I analogue (○). (C) PGE using bending probes 1−6: hSRY HMG box (lane h) and C41S mSRY HMG box (lane m). Studies employed DNA site 5′-GTGATTGTTCAG-3′ and its complement (core consensus sequence underlined) as described by Ukiyama and co-workers (22). Inferred DNA bend angles are 79 ± 1° (human) and 71 ± 1° (murine) in this gel system. (D) Plot of the relative electrophoretic mobility vs flexure displacement: hSRY HMG box (■) and C41S mSRY HMG box (striped circles). Assays employed 10% polyacrylamide with a 58:1 acrylamide:bisacrylamide ratio at 4 °C.

*Fluorescence Spectroscopy.* Spectra were recorded using an Aviv model ATF 105 fluorometer (Proterion Corp., Lakewood, NJ) using bandwidths of 0.5 and 3 nm at the excitation and emission monochromator, respectively. For FRET studies, the excitation and emission wavelengths were 490 and 520 nm, respectively, for the donor and 490 and 585 nm, respectively, for the acceptor. The quantum yield of the donor was determined for each sample using a reference solution of fluorescein in 0.1 M NaOH as a standard (0.85; *26*). The calculated value for the quantum yield of the Fl probe attached to the 5′ end of the oligonucleotide was 0.83 at 4 °C for the singly labeled free DNA and varied between 0.8 and 0.9 for all other samples. The Förster critical distance, $R_o$, was calculated accordingly (Table 2; 58.3 Å at 4 °C for the free DNA). Fluorescence-detected guanidine titrations were monitored at 4 °C by intrinsic Trp fluorescence as measured at an emission wavelength of 390 nm (slit width of 5 nm) following excitation at 270 nm (slit width of 2 nm). For such studies, human and murine domains were concentrated to 1 $\mu$M in 140 mM KCl and 10 mM potassium phosphate (pH 7.4) in the titrating cuvette. The same concentration of SRY was used in the titrant reservoir containing 7.2 M guanidine-HCl in the buffer described above. Experimental curves were fitted by nonlinear least squares to the equations $\Delta G = \Delta G_u + m[\text{guanidine}] = -RT \ln K$, where $\Delta G_u$ is the free energy of unfolding extrapolated to zero denaturant concentration and $K$ is the equilibrium constant between the native and unfolded states of the protein (*22*, *27*).

Table 1: Thermodynamic Stabilities of mSRY and hSRY Domains[a]

| species | $C_{mid}$ (M) | $m$ (kcal mol$^{-1}$ M$^{-1}$) | $\Delta G_u$ (kcal/mol) | $T_m$ (°C) |
|---|---|---|---|---|
| human | 2.2 ± 0.1 | 1.8 ± 0.1 | 4.0 ± 0.1 | 39.0 ± 0.5 |
| murine | 1.6 ± 0.1 | 2.3 ± 0.1 | 3.7 ± 0.1 | 36.0 ± 0.5 |
| murine C41S | 1.7 ± 0.1 | 2.3 ± 0.1 | 3.8 ± 0.1 | 37.5 ± 0.5 |

[a] Steady-state fluorescence and automated guanidine titrations were used to determine thermodynamic parameters, except for $T_m$ (obtained by CD). $C_{mid}$ is defined as that concentration of guanidine at which 50% of the protein is unfolded. The $m$ value, defined as the derivative of $\Delta G_u$ with respect to denaturant concentration, is related to the difference in solvent-accessible surface area between unfolded and native states. The statistical confidence of fitting is in each case characterized by an $R$ value of greater than 0.9998. $T_m$ was obtained by temperature denaturation experiments monitored by CD at 222 nm.

Table 2: Time-Resolved FRET Analysis of End-to-End Distances[a]

| sample | peak (Å) | mean (Å) | fwhm (Å)[b] | $\chi^2$ [c] |
|---|---|---|---|---|
| free DNA | 63.9 (63.8−64.1) | 63.9 (63.8−64.1) | 18.2 (17.9−20.5) | 1.75 |
| hSRY complex | 51.5 (50.8−51.9) | 51.8 (51.2−52.2) | 28.9 (27.9−30.0) | 1.66 |
| mSRY complex | 56.4 (55.8−57.0) | 56.8 (56.2−57.3) | 32.9 (30.0−29.7) | 1.77 |

[a] Measurements were taken at 4 °C. Each analysis included two independent repeats of each experiment analyzed simultaneously, a procedure which reduced the statistical uncertainties of the parameters. The peak of the end-to-end distribution function and its mean are provided. The values in parentheses represent possible errors in parameters calculated by the detailed rigorous analysis procedure as described above (see Materials and Methods; *28*). The average $R_o$ was 58.3 Å for the free DNA at 4 °C, yet the specific $R_o$ value suitable for each experiment was introduced via eq 3, using the radiative lifetime to account for the variation of the values of $R_o$ with the lifetime components of the donor. [b] fwhm is the full width at half-maximum of the distribution. [c] The global $\chi^2$ is derived from joint analysis of four donor fluorescence decay curves; its value is increased by the large number of experiments forced into joint analysis. The quality of the fit of the calculated decay curves based on the distance distribution model and the experimental decay curves was judged by several indicators (see Materials and Methods). The relatively high $\chi^2$ values presented in this table are global values obtained for the simultaneous fit of at least four decay curves obtained in two independent sets of experiments. For that reason, even at the best fit values of the parameters, the global $\chi^2$ values were relatively high.

*FRET Sample Preparation.* Fluorescence studies of DNA bending employed a 15 bp DNA duplex of sequence 5′-TCGGTGATTGTTCAG-3′ ("upper strand") and complement (consensus target site underlined). The site of cantilever insertion is therefore between base pairs 8 and 9, i.e., near the center of the oligonucleotide. Three samples of labeled DNA were prepared: (i) one containing 5′-fluorescein tethered to the 5′-phosphate of the upper strand by a hexanyl linker (6-FAM), (ii) one with tetramethylrhodamine tethered to the 5′-phosphate of the lower strand by an analogous linker (TAMRA), and (iii) a doubly labeled derivative. HPLC-purified oligonucleotides were purchased from Qiagen Operon (Alameda, CA) and purified further by HPLC after annealing using an ion exchange semipreparative HPLC column (DNA PAC PA-100, Dionex Corp., Sunnyvale, CA). For the FRET study, the duplex DNA probes were concentrated to 5 $\mu$M in 10 mM potassium phosphate and 10 mM Tris-HCl (pH 8.4) containing 140 mM KCl, 1 mM EDTA, and 1 mM DTT. The ratio of labeling in the doubly labeled sample is close to 1:1 (see the Results).

*Time-Resolved Fluorescence Measurements.* A home-built system at Bar Ilan University employs time-correlated single-photon counting. The excitation source has a picosecond mode-locked Ti:sapphire laser (Tsunami, Spectra-Physics, Santa Clara, CA) pumped by a high-power diode laser (Millenia XsJ, Spectra-Physics) and equipped with broadband optics. The output frequency of the laser was multiplied by a flexible second- and third-harmonics generator (GWU, Spectra-Physics). The laser pulse width was 1.6 ps before doubling. A pulse selector (model 3980, Spectra-Physics) was used to reduce the basic 80 MHz pulse rate to 4 MHz with energies of up to 0.5 nJ/pulse (2 mW at 4 MHz). The second harmonic was used for excitation at 490 nm. The emission was collected with a polarizer at a magic angle relative to the excitation polarizer at 520 nm. Measurements were taken at 4 °C. The emission wavelength was selected by a double $^1/_8$ m subtractive monochromator, with an emission slit width of 8 nm (DIGIKROM CM112), and directed to the surface of cooled microchannel plate PMT (MCP-PMT, Hammamatsu 3809U-50) biased at −3200 V. The acceptor emission was collected at 585 nm (emission bandwidth of 32 nm). A single-photon counting board (SPC 630, Becker & Hickel GmbH) fed via a preamplifier (HFAC-26DB 0.1UA) and triggered by a photodiode (PHD-400N) was used. The response of the system yielded a pulse width of 35 ps.

*Global Analysis of tr-FRET.* Distance distribution functions were obtained from simultaneous global analysis of four experimental fluorescence decay curves. These were recorded for (i) samples containing only the donor, (ii) samples containing only the acceptor, and (iii) donor and acceptor decay curves from samples containing both fluorophores. Global analyses of the fluorescence decay curves were performed using the Marquardt nonlinear least-squares method (*28*). Four (or two) theoretical decay curves were calculated for each set of FRET experiments and fitted to the corresponding experimental curves. Calculated decay curves were prepared by numerical solution of a modified version of the second-order differential equation (*28*, *29*):

$$\frac{\partial p_i(r,t)}{\partial t} = D\frac{\partial}{\partial r}\left\{\exp\left[-\frac{U(r)}{k_B T}\right]\frac{\partial}{\partial r}\left[\exp\frac{U(r)}{k_B T}p_i(r,t)\right]\right\} - k_i(r)p_i(r,t) \quad (1)$$

where $p_i(r,t)$ is the probability density for finding an excited-state donor (of the fluorescence lifetime component $\tau_i$ and normalized preexponential factor $\alpha_i$) with a distance $r$ from the acceptor at time $t$ after excitation, so the total probability density is

$$p(r,t) = \sum_{i=1}^{n} \alpha_i\, p_i(r,t) \quad (2)$$

In eq 1, $D$ is the intramolecular segmental diffusion constant. This was fixed to zero (i.e., no diffusion was assumed), since in the first set of experiments the analyses gave zero diffusion. $U(r)$ is a distance-dependent potential of mean force; $k_B$ is the Boltzmann constant, and $T$ is the absolute temperature. This potential is usually taken to be harmonic (thus implying a Gaussian distribution for the distance between the fluorescent probes). $U(r) = a(r - b)^2$, where $a$ and $b$ are parameters. $k_i(r)$ is the reaction term, including the spontaneous emission rate and the Förster energy transfer rate:

$$k_i(r) = \frac{1}{\tau_i} + \frac{1}{\tau_r}\frac{9\ln 10}{128\pi^5 n^4 N_A}\kappa^2 J\frac{1}{r^6} \quad (3)$$

where $\tau_r$ is the radiative lifetime of the donor, $n$ is the refractive index of the medium between the donor and acceptor, $N_A$ is Avogadro's number, $\kappa^2$ is the orientational factor, and $J$ is the normalized spectral overlap integral.

Evaluation of each analysis and the significance of inferred parameters were based on four indicators: the global $\chi^2$ values, the distributions of the residuals, the autocorrelation of the residuals, and the error intervals of the calculated parameters (*28*, *30*). The error intervals were obtained by a rigorous analysis procedure carried out for each set of experiments.

*NMR Spectroscopy*. Spectra were observed at 500 and 600 MHz using Varian spectrometers (Varian Instruments, Inc., Palo Alto, CA). Spectra of exchangeable resonances were obtained in $H_2O$ using laminar-shifted shaped pulses in the absence of solvent presaturation as described previously (*13*). Spectra of complexes were obtained at 15 and 25 °C in 10 mM potassium phosphate (pH 6.0) and 50 mM KCl. Assignments of DNA imino and adenine H2 resonances were obtained as described previously (*13*). Assignment of M13 resonances in the murine complex was based on comparison with the spectrum of the M13I variant as described in studies of hSRY (*31*).

## RESULTS

This study employs human and murine SRY domains consisting of the HMG box and C-terminal basic region similar in length to that employed in structural studies of hSRY by Clore and colleagues (Figure 1C; *15*). The extent of specific DNA bending by the corresponding human domain is similar to that of full-length hSRY (Supporting Information). Proteins were freshly reduced with DTT and repurified by RP-HPLC to prevent oxidation of cysteine or methionine. Whereas the human domain lacks cysteine, the murine domain contains a single cysteine on the back surface of the HMG box (C41; Figures 1C and 2). To avoid artifacts due to its oxidation, studies in part employ murine domains containing the conservative C41S substitution. Control studies indicated that the substitution does not affect specific DNA binding (Supporting Information) or DNA bending (see below). Position 41 in the HMG box is not conserved among SRY or SOX sequences. As expected, the hSRY and mSRY domains exhibit similar secondary structures as inferred from comparison of far-ultraviolet CD spectra (Supporting Information). These domains and the C41S murine analogue exhibit similar thermodynamic and thermal stabilities (Table 1).

*Time-Resolved Fluorescence Resonance Energy Transfer*. Bending of a consensus DNA target site was characterized by tr-FRET analysis (*29*, *32*−*35*). End-to-end distances in the DNA site are expected to be sensitive to the degree of DNA bending (Figure 4A). DNA probes consist of a 15 bp duplex labeled with a fluorescent donor (fluorescein) at one 5′ end and/or a fluorescent acceptor (tetramethylrhodamine) at the other 5′ end as described previously (*22*; see Materials and Methods). To ensure the rotational mobility of the probes, the fluorophores were tethered to the DNA by a flexible linker. Comparison of absorption spectra demonstrates that the donor-to-acceptor ratio of labeling in the doubly labeled duplex is close to 1:1. Fluorescence emission spectra of the DNA probes and representative protein−DNA complexes following excitation at 490 nm indicate that protein binding has no effect on the shape of the spectrum of the donor (Figure 4B). Comparison of the emission spectrum of the doubly labeled DNA and emission spectra obtained for the two singly labeled sites demonstrates a decrease in the emission intensity of the donor accompanied by a corresponding increase in the emission intensity of the acceptor. The acceptor emission was very low in the absence of the donor (pink trace in Figure 4B). As expected for the Förster-type FRET mechanism, the shapes of the emission spectra of the two probes were unchanged. FRET studies were conducted at 4 °C to permit comparison with PGE experiments at the same temperature (see below).

Control steady-state spectroscopic experiments were undertaken to test the suitability of these samples for tr-FRET analysis of distributions of end-to-end distances. The possibility of intermolecular FRET was excluded in control studies of solutions containing equimolar mixtures of donor-labeled DNA and acceptor-labeled DNA (or their protein−DNA complexes). Further, intramolecular FRET efficiencies of the free and bound DNA sites are independent of the concentration of DNA (or the protein−DNA complex) in the range of 1−15 $\mu$M. Steady-state anisotropies of the donor and acceptor attached to the DNA duplex were measured for the free and protein-bound sites. For the singly labeled free DNA, the respective anisotropies of the donor and acceptor were 0.10 and 0.23 at 4 °C. As expected, in the doubly labeled DNA site, the anisotropy of the donor increased to 0.13 due to the reduction of the fluorescence lifetime; the anisotropy of the acceptor was unchanged. Formation of the human and murine complexes in each case caused a ca. 30% increase in donor emission anisotropy and a 20% increase in acceptor emission anisotropy (see the Supporting Information). Such effects presumably reflect the further reduction of the donor lifetime and changes in the DNA conformation on protein binding. These steady-state fluorescence experiments demonstrate that the reduction of the fluorescence intensity of the donor in the doubly labeled DNA site is indeed a result of FRET interaction and not of a spurious quenching process (such as end-to-end stacking of DNA). Further, the spectroscopic properties of the probes are not perturbed by the bound protein. The low steady-state emission anisotropy of the donor and the partial anisotropy of the acceptor emission in the free and bound DNA site support the determination of intramolecular distances based on averaged orientation factors (*34*). Qualitative inspection of the steady-state spectra suggests that the murine complex is less bent than the human complex. Fluorescence spectra of the mSRY and C41S mSRY complexes are essentially identical (data not shown).

Time-resolved measurements of the decay of donor fluorescence at 4 °C (Figure 4C) demonstrate a significant change in FRET on protein binding. This effect is manifest by the reduction of the lifetime of the fluorescein in the doubly labeled DNA site on binding of either mSRY or hSRY. Because the reduction is more marked in the human complex, qualitative inspection of these data reveals that binding of hSRY shortens the mean end-to-end distance in the DNA significantly more than does binding of mSRY. Global analysis of the fluorescence decay of the donor in the singly and doubly labeled samples yields an estimate of the end-to-end distance distribution in the free and bound states (Figure 4D). Parameters are given in Table 2. When the hSRY HMG box binds, the mean end-to-end distance of the DNA was reduced from 63.9 ± 0.2 to 51.8 ± 0.5 Å. By contrast, when the mSRY HMG box binds, the mean distance was reduced to 56.8 ± 0.5 Å. Reduction of end-to-end

Table 3: DNA Bend Angles in Human or Murine SRY-p Complexes[a]

| DNA site | protein | gel | angle (deg) | $\Delta\theta$ (deg) |
|---|---|---|---|---|
| 1 | human | 10%/58:1 | 79 | − |
| 1 | murine | 10%/58:1 | 71 | 8 |
| 2 | human | 10%/58:1 | 87 | − |
| 2 | murine | 10%/58:1 | 78 | 9 |
| 3 | human | 10%/58:1 | 84 | − |
| 3 | murine | 10%/58:1 | 74 | 10 |
| 4 | human | 10%/58:1 | 83 | − |
| 4 | murine | 10%/58:1 | 77 | 6 |

[a] Studies employed the C41S analogue of the mSRY domain. DNA sites are as defined in the text. Site 1 is as employed in FRET studies. Sites 2 and 3 are designated *TCR-11* and *TCR-15*, respectively, in studies by Grosschedl and co-workers in Table S2 (*17*). Site 4 is obtained from Goodfellow and co-workers (*18*). Protein domains contain intact C-terminal basic tails. The murine domain contains the C41S substitution, which does not affect DNA binding properties; the position of Y41 at the back surface of hSRY is shown in Figure 1C. DNA bend angles were obtained using the F−B model (*37*). The precision of PGE-based estimates of DNA bend angles is in general ±1°, implying uncertainties in $\Delta\theta$ values of ±2°.

distances reflects both DNA bending and DNA unwinding. Comparison of the distributions also reveals a statistically significant broadening of the end-to-end distribution in both complexes. The full widths of the Gaussian distributions at half-maximal height (fwhm; column 4 in Table 2) increased from ca. 18 Å in the free DNA to ca. 30 Å in the complexes. This broadening suggests that the bound DNA structures exhibit conformational flexibility and in turn a range of occupied DNA bend angles in the ensemble. The breadth of the distribution in the free DNA is consistent with the allowed range of flexible linker configurations (A. R. Srinivasan and W. Olson, personal communication). Whereas macromolecular assembly often damps conformational fluctuations, broadening of end-to-end distance distributions in the present bent DNA complexes suggests that binding of a protein can enhance the flexibility of DNA.

*Permutation Gel Electrophoresis.* The specific affinities of the mSRY and hSRY HMG boxes for a 36 bp DNA duplex containing a consensus target site are in the range of 5−10 nM in accord with past studies (*22*). Studies of DNA bending were extended by PGE analysis of the 5′-ATTGTT-3′ core consensus site in four different sequence contexts. Site 1 (5′-**GTG**ATTGTT**CAG**-3′ and complement), employed in the FRET studies described above, is derived by optimizing a SOX binding site in the promoter of the Mullerian-Inhibiting Substance (MIS) gene as described previously (*13*). Representative electrophoretic data are shown in panels A and C of Figure 5. Sites 2 and 3 (5′-CCCATTGTT**CTC**-3′ and 5′-CCCATTGTT**GCT**-3′, respectively; differences in bold) were employed in the original studies of Grosschedl and co-workers (designated *TCR-11* and *TCR-15*, respectively; *17*). Site 4 (5′-C**G**CATTGTT-**ATC**-3′ and complement) is derived from the *CD3ϵ* enhancer employed in the DNA bending studies of Goodfellow and co-workers (*18*, *36*). Analysis of electrophoretic mobilities as a function of flexure displacement in the permuted DNA probes by the F−B model (*37*; Figure 5B,D) provides estimates of induced DNA angles as given in Table 3. Whereas previous comparison of human and murine domains indicated that the mSRY−DNA complex was 25−30° more sharply bent than the hSRY−DNA complex (*17*), the results presented here indicate the opposite: the human complex is 6−10° more acutely bent (column 5 in Table 3).[6] Because absolute calibration of PGE results is model-dependent, this range of bend angles (79−83° for the human complex) may overestimate the actual bend angle.[7] Significantly, with respect to differences in bending between human and murine complexes, the PGE results presented here are in qualitative accord with tr-FRET analysis. The PGE estimate of $\Delta\theta$ is smaller than that implied by a two-dimensional triangle model of DNA bending (Figure 4A; see the Discussion).

*Differences in DNA Stringency and Specificity.* In accord with previous studies (*17*), the hSRY and mSRY domains each recognize a consensus DNA target site (5′-CATTGTTC-3′ and complement; core bases underlined) with similar affinities. An electrophoretic screen for changes in specificity and stringency of sequence recognition (Figure 6A) was performed on the basis of variant DNA probes (Figure 6B). The screen employed a constant protein concentration of 50 nM. As reported by Grosschedl and co-workers (*17*), the murine domain discriminates more effectively between the SRY site and the related LEF-1 site (5′-**T**TTG**AA** and complement; changes in bold face). Whereas the difference in the percent shift between SRY and LEF-1 sites is only 2-fold for hSRY (42% shift vs 18%; lanes a and d at the left in Figure 6A), the difference in the percent shift is 20-fold for mSRY (40% vs 2%; lanes a and d in the middle panel of Figure 6A). Similarly, whereas a core 3′ T → A transversion (5′-ATTGT**A**-3′ and complement; site b in Figure 6B) slightly impairs binding of the human domain in accord with past studies (*13*), the same transversion slightly enhances binding of the murine domain. Binding in each case is enhanced by changing the base 3′ to the core (5′-CATTGTA**G**-3′ and complement; site c in Figure 6B). The 3′ two base pairs are contacted by the C-terminal tail in the NMR structure of the SRY−DNA complex, which employed variant site 5′-GTTTGTG**C**-3′ (and complement; *15*). The phenotypic interchangeability of murine and human SRY HMG boxes in transgenic mice (*20*) indicates that these differences in specificity are unlikely to be biologically significant.

*NMR Studies Demonstrate Analogous Side Chain Intercalation.* Inspection of the sequences of mSRY and hSRY indicates four differences in the DNA-binding surface of the HMG box (red boxes in Figure 2). Of these, the most prominent contact is made by the cantilever side chain at position 13 (arrow in Figure 2). In the hSRY−DNA complex, I13 inserts between AT base pairs at the center of the DNA bend to disrupt base stacking but not base pairing (*13*, *15*). The corresponding residue in mSRY is methionine, which

[6] The apparent effect of 3′ flanking DNA sequences on the bend angle of a variant hSRY peptide−DNA complex (*17*) is almost completely attenuated in the tailed complexes [$\Delta\theta = 3°$ (Table S2 of the Supporting Information)].

[7] NMR studies of an SRY−DNA complex yielded a DNA bend angle of 54 ± 2° (*15*). These studies employed the nonconsensus DNA site 5′-GGTG**TTTGTG**CAGG-3′ and complement; the two nonconsensus bases are underlined. PGE analysis of this variant site yields an apparent DNA angle of 67.5 ± 1°. Thus, the PGE values in Table 3 may overestimate the degree of bending by 13.3 ± 3°. This may not be a general feature of the PGE method as studies of the LEF-1 HMG box by Crane-Robinson and co-workers found that PGE-inferred bend angles could vary from 78° (using 141 bp DNA probes; *37*) to 130° (using 440 bp DNA probes; *42*), whereas the bend angle observed in the NMR structure of this complex is ca. 117° (*14*).

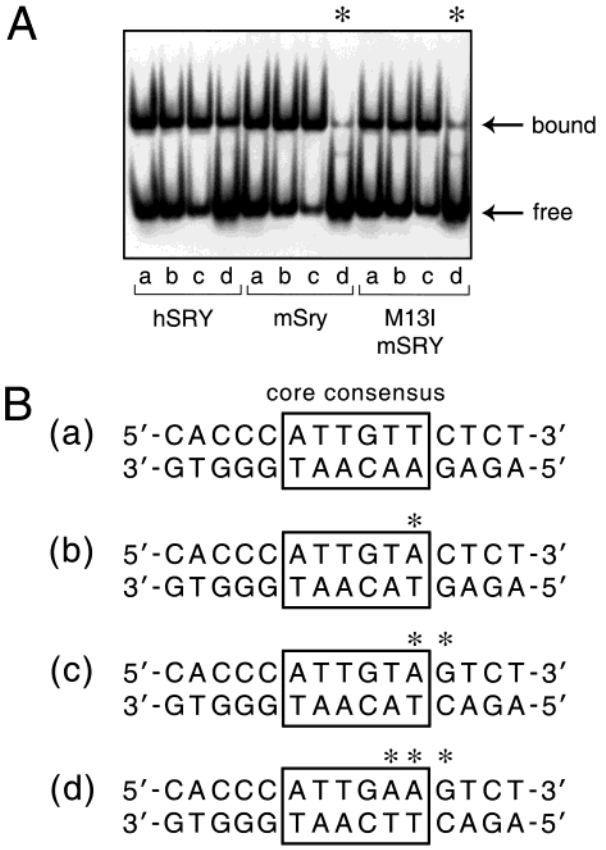


FIGURE 6: Murine SRY domain which exhibits more stringent sequence specificity. (A) GMSA screening of consensus and variant DNA sites and (B) sequences. Lanes a−d in panel A refer to DNA sites a−d, respectively, defined in panel B; boxes denote the core consensus sequence. The mSRY HMG box and M13I analogue demonstrate more effective discrimination (asterisks in panel A) against site d (5′-ATTGAA-3′, nonconsensus bases underlined) than the hSRY domain, in accord with results of ref *17*. For each protein, a change of the 3′ flanking base pair from C to G (upper strand) in the context of the 5′-ATTGTAX-3′ site slightly enhances binding. Percent shifts are as follows: (a) 42, (b) 36, (c) 59, and (d) 18 for the hSRY HMG box, (a) 40, (b) 50, (c) 80, and (d) 2 for mSRY, and (a) 21, (b) 24, (c) 48, and (d) 4 for mSRY M13I. The murine domain and M13I analogue contain the C41S substitution.

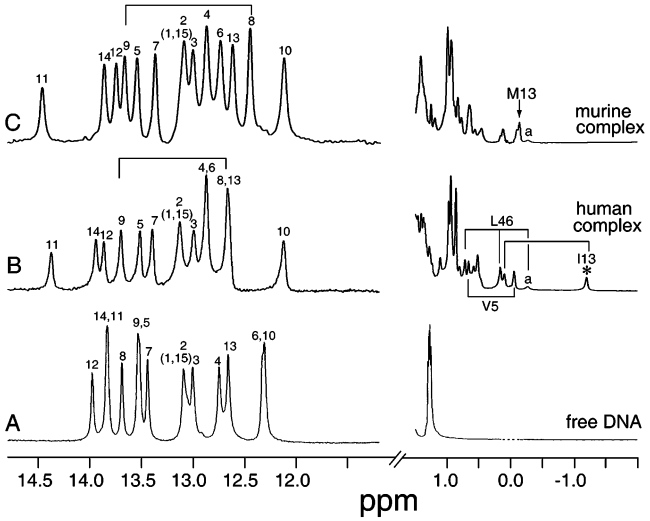


FIGURE 7: NMR studies of human and murine SRY−DNA complexes. At the left are shown the downfield regions of 600 MHz $^1$H NMR spectra containing DNA imino resonances (12−14.5 ppm) and at the right upfield regions containing thymidine methyl and protein aliphatic resonances: (A) free 15 bp DNA (5′-GGGGT-GATTGTTCAG-3′ and its complement; inset in Figure 8A), (B) 1:1 complex with the hSRY HMG box, and (C) 1:1 complex with the C41S analogue of the mSRY HMG box. Assignments of DNA imino resonances (left) and selected protein aliphatic spin systems (right) are as indicated. Resonance a in panel B is the upfield $\beta$ resonance of L46. Brackets in panel B (left) indicate the site of cantilever insertion between base pairs 8 and 9. Broadening of G1 and G15 terminal resonances is due to fraying of the double helix. The asterisk in panel B denotes the cantilever side chain I13 in hSRY. The arrow in panel C denotes the corresponding side chain M13 in the murine complex. Complexation shifts of I13 and M13 are similar. Spectra were obtained at 25 °C in 10 mM potassium phosphate (pH 6.5) and 50 mM KCl in 90% $H_2O$ and 10% $D_2O$.

differs from isoleucine in both shape and length. Because homologous HMG boxes can utilize side chains at nonhomologous sites to insert between DNA base pairs (*38*−*40*), we sought to determine by two-dimensional $^1$H NMR methods whether I13 and M13 play similar roles in respective human and murine complexes. These studies employed a 15 bp DNA duplex containing a central consensus 5′-ATTGTT-3′ sequence (site 1) in DNA bending studies (see above). Successive titration of the mSRY domain into a solution containing the DNA duplex reveals slow exchange between free and bound DNA resonances (Supporting Information) in accord with past studies of the human domain (*13*) and the nanomolar affinities of these complexes (above). One-dimensional $^1$H NMR spectra of the imino and methyl resonances of the free DNA are shown in Figure 7A (left- and right-hand panels, respectively). Assignment of resonances is as previously described (*13*). As is characteristic of B-DNA, guanine imino resonances (base pairs 1−4, 6, 10, 13, and 15) are clustered between 12.0 and 13.2 ppm whereas thymidine imino resonances are clustered between 13.4 and 14.0 ppm. Terminal imino resonances are broad due to fraying of the double helix. The methyl region contains a poorly resolved group of thymidine resonances near 1.2 ppm.

Addition of the hSRY or C41S mSRY HMG box results in marked changes in the chemical shifts of imino resonances, providing intrinsic probes for DNA structure (panels B and C of Figure 7, respectively). The pattern of changes in chemical shifts ("complexation shifts") exhibits qualitative similarities and detailed differences (Supporting Information). In each case, for example, G10 is shifted upfield and T11 downfield. Similarly, in each case, T5, T7, and G13 are changed little. Conversely, differences between complexation shifts of corresponding DNA imino resonances in the two complexes presumably reflect subtle changes in DNA structure. Although divergent amino acids at the protein−DNA interface (boxes in Figure 2) can also affect such complexation shifts, NMR studies of an engineered murine domain containing a "humanized" interface indicate that the contribution of DNA structure is predominant (unpublished results). These findings in turn suggest that induced DNA bend angles (whether human or murine) are specified not only by particular interfacial side chains, but also by the overall framework of the respective HMG boxes.

Previous studies have established that I13 in hSRY inserts between T8 and T9, leading to a dramatic upfield shift of T8 into the guanidine imino region (*13*). This is accompanied by an upfield shift of the I13 spin system (asterisk in Figure 7B, right), presumably due to ring currents of adjoining DNA bases. Intermolecular nuclear Overhauser effects (NOEs) are observed from the imino protons of T8 and T9 to the

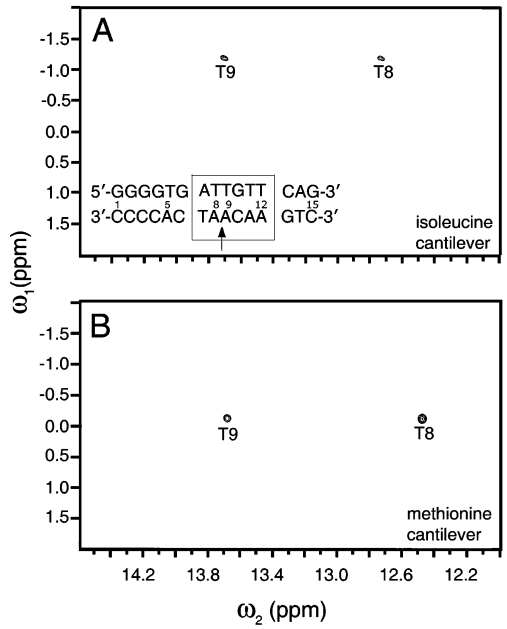


FIGURE 8: Two-dimensional $^1$H NMR NOEs diagnostic of SRY−DNA intercalation. NOESY spectra of human (A) and murine (B) protein−DNA complexes in 90% $H_2O$ and 10% $D_2O$. Corresponding intermolecular NOEs are observed involving I13 (human) and M13 (murine) in the respective DNA complexes. The sequence of the duplex DNA site is the inset in panel A; the arrow denotes the site of insertion of the cantilever side chain. The mixing time was in each case 150 ms. Spectra were obtained as described in the legend of Figure 7. The murine domain contains the C41S substitution.

$\delta$-methyl resonance of I13 (Figure 8A). The mSRY−DNA complex exhibits a similar pattern of chemical shifts and NOEs at these sites. Whereas the imino resonance of T9 is similar in the two complexes, T8 is shifted even further upfield in the murine complex than in the human complex (Figure 7C, right). These bases are in contact with a methionine $\epsilon$-methyl resonance (Figure 8B). A candidate tyrosine hydroxyl resonance corresponding to that assigned in the human complex was not observed (*16*). The intercalating methionine was identified as residue 13 by a comparative study of an M13I variant domain (see below). The $^1$H NMR spectrum of the M13I mSRY−DNA complex lacks an upfield-shifted methionine methyl resonance but instead exhibits an upfield-shifted isoleucine spin system in NOE contact with T8 and T9, similar to that observed in the human complex. A methionine cantilever has been observed in the structure of a LEF1−DNA complex (*14*).

The extreme upfield shift of the T8 imino resonance observed in spectra of both the murine and human SRY−DNA complexes (12.51 and 12.71 ppm, respectively) was not observed in NMR studies of a nonconsensus SRY−DNA complex containing variant MIS target site 5′-TTTGTG-3′ and its complement (nonconsensus bases underlined; *15*). Whereas the corresponding T8 resonance in the nonconsensus site is also shifted upfield, its complexation shift is attenuated: the T8 resonance remains within the canonical T region of the imino $^1$H NMR spectrum. PGE studies indicate that the variant hSRY complex is 10° less bent than the consensus 5′-ATTGTT-3′ hSRY complex (site 1; see footnote 7). This difference suggests that the structure of the (more sharply bent) hSRY−DNA complex investigated here differs from the reported hSRY−DNA structure (Figure 1C and Supporting Information; *15*) at the cantilever insertion site. Possible origins of the extreme T8 upfield shift are weakening of the A8···T8 Watson−Crick imino hydrogen bond and/or more marked disruption of base stacking between base pairs 8 and 9, leading to attenuation of downfield ring currents within the double helix. Evidence for the latter mechanism is provided in both human and murine complexes by the absence of a base stacking-related NOE between the imino protons of T8 and T9 (Supporting Information). In the reported structure (*15*), this distance is 3.4 Å and so would be readily detectable (Supporting Information). The absence of this NOE at the observed signal-to-noise ratio indicates that the T8−T9 distance is greater than 3.9 Å in these complexes. The pattern of $^1$H NMR complexation shifts among adenine H2 resonances, like that among imino resonances, is similar but not identical between the human and murine complexes (Supporting Information). In these complexes, the imino protons of T8 and T9 retain strong NOE contacts with respective adenine H2 protons of the same base pair (permitting their assignment) but not with H2 protons of flanking base pairs [7.67 (A7), 7.61 (A8), and 7.55 ppm (A9) at 15 °C]. The distal methyl resonances of the I13 and M13 cantilevers exhibit NOE contacts to the H2 resonances of both A8 and A9. Possible weak NOEs between the latter DNA protons could not be evaluated due to the similarity of their chemical shifts. A table of interproton distances at the cantilever insertion site of the reported hSRY−DNA (5′-TTTGTG-3′) complex is provided in the Supporting Information.

*Interchange of Cantilevers Does Not Alter Specificity or Bending.* Previous studies have shown that foreshortened or polar substitutions at position 13 of the hSRY HMG box significantly impair DNA binding (*3*, *31*).To test whether differences between isoleucine or methionine cantilevers are responsible for the distinct specificities and DNA bend angles of hSRY and mSRY, we constructed an M13I variant of the C41S mSRY domain. The variant murine domain recapitulates the specificity (Figure 6A, right) and DNA bend angle (Figure 8A,B) of the wild-type domain. The absence of a significant effect of the cantilever shape on the DNA bend angle is consistent with the previous studies of a variant hSRY domain in which I13 was substituted with phenylalanine as in the chimpanzee (*22*). Together, these studies suggest that the cavity induced between splayed AT base pairs must be filled by an aliphatic or aromatic side chain, presumably to avoid a cavity "penalty" (*41*), but that the detailed configuration of the cantilever does not specify DNA affinity or bend angle.

## DISCUSSION

The present studies have investigated the DNA bending properties of human and murine SRY domains. These studies employed a combination of time-resolved FRET and gel-based PGE methods. Careful comparison of these domains was motivated by recent studies of chimeric *mSry* transgenes in XX mice (Figure 3; *19*, *20*). Our results demonstrate that the hSRY domain bends a consensus DNA target site more sharply than does the homologous murine domain. This conclusion is the opposite of that reached in a previous study (*17*) but is in overall accord with a range of other observations (Supporting Information; *21*, *37*, *42*, *43*). Although

hSRY and mSRY each induce sharp bends in DNA, the relative extent of DNA bending has been uncertain as a broad range of bend angles have been inferred from PGE. This range is due in part to differences among studies with respect to protein constructions, DNA target sites, conditions of electrophoresis, length of DNA probes, and methods of interpretation. Such variability highlights the importance of side-by-side comparisons of the two proteins by independent biochemical and biophysical methods.

The original studies of *Sry* transgenes in XX mice by Lovell-Badge and colleagues found that a 14 kb genomic segment of the murine Y chromosome containing the *mSry orf* and flanking control elements could direct male somatic differentiation (*6*). This pioneering observation demonstrated that mSRY was sufficient for initiating a genetic program of male organogenesis in an otherwise female genetic background. Surprisingly, however, an analogous 24 kb segment of the human Y chromosome containing *hSry* could not function in a similar fashion. The seemingly "defective" DNA binding properties of hSRY relative to mSRY (*17*) were proposed to rationalize these results: failure of hSRY-directed male gene expression in XX mice was proposed to reflect impaired DNA bending or less stringent sequence specificity (*17*). Other possibilities, such as divergence of *cis*-acting control elements between human and murine genomic segments or the absence of a glutamine-rich repeat hSRY, could not be excluded. Differences in DNA bending have provided an attractive mechanism in light of the conservation of DNA bend angles observed among primate SRY complexes (*21*). Further, a decrement in the angle of DNA bending similar in magnitude to that reported between mSRY and hSRY was characterized in a variant hSRY−DNA complex associated with *de novo* sex reversal (*18*).

Recent studies of transgenic XX mice by the laboratories of Eicher (*19*) and Lovell-Badge (*20*) have demonstrated that "domain swap" of the mSRY HMG box with the boxes of hSRY, mSOX3, or mSOX9 does not impair *Sry*-directed male sex determination (Figure 3). Because the mSOX3 domain exhibits significantly weaker specific DNA binding affinity than the mSRY domain (*44*) and because the hSRY domain is less specific than the mSRY domain (ref *17* and results presented here), the transgenic studies indicate that such biochemical differences do not impair the function of the chimeric proteins in directing native sex-, stage-, and tissue-specific gene expression. It is not known whether position effects enhancing levels of expression of the transgene can compensate for lower intrinsic affinities or specificities, an uncertainty compounded by the variable phenotypes of the transgenic mice (*6*, *19*). Nonetheless, of particular importance, the functional interchangeability of hSRY and mSRY (*20*) implies that the multiple divergent residues in hSRY (Figure 2) may be functionally regarded as neutral polymorphisms in the mouse. Comparison of mSRY and hSRY thus provides an opportunity to investigate what range of biochemical properties is compatible with SRY-directed testicular differentiation. Such studies would complement investigation of the perturbed features of clinical variants that are apparently incompatible with testicular differentiation (*18*). One such variant (I13T in the hSRY HMG box; residue 68 in the intact protein) exhibits defective sequence-specific DNA recognition but retains the ability to bind the sharp angles of four-way DNA junctions (*31*). Although hSRY is less stringent in its specificity than mSRY, sex reversal in that patient strongly suggests that some degree of sequence recognition is required for testis determination.

*Comparison with Previous Results.* Why might these results conflict with the earlier biochemical study by Grosschedl and co-workers (*17*). The answer may lie in inadvertent differences between the amino acid sequences of the respective "human" domains. In the earlier study, the HMG box fragments were expressed in *E. coli* as factor Xa-cleavable fusion proteins with glutathione *S*-transferase. The mSRY polypeptide (85 residues) consisted of 80 SRY-derived amino acids (residues 3−82 in the HMG box consensus) flanked by two additional residues at the amino terminus (GI; derived from the factor Xa site) and additional residues at the carboxy terminus (NSS, because of the translational stop codon in the vector; *42*). Similarly, the hSRY polypeptide (93 residues) consisted of 88 SRY-derived amino acids (residues 1−83 in the HMG box consensus and five additional residues from the preceding nonbox sequence; *17*) flanked by the same vector-derived amino- and carboxy-terminal extensions (GI and NSS, respectively; *42*). The mSRY and hSRY polypeptides thus differed from the native proteins and from each other at each end of the HMG box and in particular contained variant distal tail sequences (Figure 1E). The contribution of the C-terminal tail to DNA bending has been emphasized in studies of LEF-1 and TCF-1α (*45*).

The previous side-by-side PGE studies of DNA bending appeared to show that the hSRY-related DNA complex was less sharply bent than the mSRY-related DNA complex ($\Delta\theta = 25-30°$). Whereas the target site 5′-CCCATTGTTCTC-3′ (and complement; site *TCR-11* in ref *17*) was bent by 85° by the mSRY-derived fragment, it was bent by only 60° by the hSRY-derived fragment. DNA bending was blunted in each case (to angles of 60° and 30°, respectively) using variant target site *TCR-15* (5′-CCCATTGTT**GCT**-3′ and complement, in which variant bases are bold; *17*). Gel compositions were not reported. We have not investigated whether the reported DNA bend angles were influenced by vector-derived substitutions in the expressed domains or other possible confounding factors, such as differential methionine oxidation or inadvertent proteolytic truncation of the human domain. It is possible that the reported decrease of 25−30° between hSRY and mSRY complexes might have been an unappreciated consequence of the variant polypeptides previously employed. In the study presented here, we have undertaken a side-by-side comparison of corresponding mSRY and hSRY domains, each containing intact C-terminal tails. Although the reported differences in discrimination between SRY consensus target sites (5′-ATTGTT-3′ and complement) and LEF-1 target sites (5′-TTTGAA-3′ and complement) are verified (Figure 6), we find that hSRY complexes are bent somewhat more sharply than are mSRY complexes (Table 3). Further, the marked impairment of DNA bending previously reported between sites *TCR-11* and *TCR-15* ($\Delta\theta = 25-30°$ for both mSRY and hSRY) is not observed. Since these sites differ only in the tail-binding region (5′-CCCATTGTT**CTC**-3′ vs 5′-CCCATTGTT**GCT**-3′, in which differences are in bold), we speculate that an intact tail largely "rescues" a bending defect (observed $\Delta\theta < 6°$ for both mSRY and hSRY in a 10%/58:1 gel system) otherwise associated with the 3′-**GCT** flanking sequence. Partial truncation of the tail after residue 78 in the HMG

box consensus decreases the DNA bend angle of both mSRY and hSRY domains and to similar extents ($\Delta\theta = 6-10°$).[8]

*Comparison of NMR, FRET, and PGE Results.* Comparative $^1$H NMR studies of the human and murine SRY−DNA complexes indicate differences in the complexation shifts of imino resonances. These studies employed the 5′-ATTGTT-3′ consensus sequence (distinct from the nonconsensus 5′-TTTGTG-3′ site seen in the available NMR structures; *15*, *46*). These $^1$H NMR spectra provide intrinsic probes of structural distortions in the double helix (*13*). In each case, a similar overall pattern of complexation shifts is observed, i.e., which resonances are shifted upfield or downfield on protein binding. Differences are nonetheless observed in the detailed appearance of the spectrum, suggesting local adjustment of structural parameters (such as the relative orientation of bases and the strengths of Watson−Crick hydrogen bonds) associated with the inequivalent DNA bend angles. It would be of interest in the future to determine the three-dimensional structure of an mSRY−DNA complex. Such a comparison would enable the same DNA sequence to be visualized at two different bend angles.

Whereas NMR chemical shifts and NOEs are local parameters, a significant difference in the global structure of the protein−DNA complexes, as probed by the distance between respective 5′ ends of the two strands, is demonstrated by FRET. In the free DNA, this distance is ca. 64 Å (Table 2), consistent with the canonical structure of B-DNA (including the effects of the linkers). Binding of the human domain leads to a dramatic reduction in the end-to-end distance (ca. 52 Å). By contrast, the effect of the murine domain is less marked (ca. 57 Å). Thus, the reduction in the murine complex ($\Delta l = 7.1 \pm 0.7$ Å) is only ca. 60% of that seen in the human complex ($\Delta l = 12.1 \pm 0.7$ Å). Because end-to-end distances are influenced by both DNA bending and DNA unwinding, it is not possible to rigorously interpret these distances in relation to DNA bend angles. A simple triangle model, employed by us and others as a first-order approximation (*22*, *47*), implies respective DNA bend angles of 72° (hSRY) and 55° (mSRY). This model neglects effects of DNA unwinding but may provide qualitative insight.

PGE analysis of the same DNA site also indicates sharper bending by the human domain, but the increment is only 5−10°, smaller than what would naively be predicted on the basis of FRET distances. We suggest that this apparent discrepancy is due to less unwinding of the DNA in the murine complex. It is not known how PGE estimates may be influenced by conformational flexibility in the bent protein−DNA complex. An interesting and unexpected feature of the FRET-derived distribution of end-to-end distances is its broadening in both complexes. Whereas the free DNA exhibits a range of flexibility consistent with the mobility of the two linkers tethering the fluorophores, the fwhm values in the two complexes suggest that a range of DNA bend angles are occupied in solution. Since NMR spectra contain a single set of protein and DNA resonances (*46*), transitions between distinct bent conformers must be fast on the time scale of NMR chemical shifts (<1 ms) and slow on the time scale of the lifetime of the excited state of the donor, i.e., slower than 5 ns. It should be noted that the FRET experiment showed that fluctuations of the end-to-end distances were slower than the lifetime of the excited state, and therefore, no intramolecular diffusion could be detected. Such fast exchange also rationalizes why electrophoretic methods give rise to a single well-defined band and not multiple shifted 1:1 complexes or diffuse bands at the trailing edge. These observations suggest that DNA bend angles observed in crystal structures or inferred from distance geometry models may represent only one of several possible structures in solution.

Although binding of proteins to B-DNA is ordinarily thought to damp conformational fluctuations (a general feature of macromolecular assembly), sharp DNA bending by hSRY or mSRY is associated with enhanced flexibility. We speculate that the breadth of the tr-FRET-derived distance distributions reflects the topography of the energy landscape of DNA far from its ordinary ground state: a penalty in free energy is incurred on DNA bending and unwinding. That a range of DNA bend angles must be populated also highlights the architectural plasticity of the HMG box. It would be of interest in the future to investigate whether mutations in the HMG box can perturb the precision of DNA bending as well as its mean bend angle. Such studies might employ residual dipolar couplings in isotopically labeled DNA as a laboratory-frame-based probe of DNA bending and conformational fluctuations.

*Role of the Cantilever Side Chain.* Human and murine SRY differ in the identity of the amino acid employed as a cantilever side chain (residue 13 in the HMG box consensus). In the structure of the human complex, I13 inserts between TA base pairs (arrow at the intercalation site in 5′-A**T**↓**T**GTT-3′) to disrupt base stacking but not base pairing (*4*, *13*). The present NMR studies have established, as expected, that the corresponding methionine (M13) in mSRY inserts at the same site. A methionine cantilever has previously been observed in the LEF-1 structure (*14*). In each case, the terminal methyl resonance of the cantilever exhibits a large upfield shift, presumably due to the large ring currents of the neighboring DNA bases.

To investigate whether this difference in cantilevers is responsible for the overall differences in DNA bending or stringency of DNA recognition, we have investigated the properties of an M13I analogue of the murine domain. The substitution is conservative, indicating that the distinct DNA binding properties of mSRY and hSRY are due to other sites of sequence differences between the two domains. Because only four other sites in the predicted protein−DNA interface differ [three in the HMG box (red boxes in Figure 2) and K73R in the tail; *15*], it is possible that their respective DNA binding properties are distinguished by a small number of residues. It is also possible that the multiple sequence variations extrinsic to the actual protein−DNA interface affect the structural framework of the HMG box and indirectly modulate DNA binding properties. The similar DNA bending properties of structurally diverse cantilevers

[8] The DNA bending properties of a truncated murine domain ending at position 75 of the HMG box (arrow in Figure 1E) were investigated in the *TCR-11* site (site 2 in Table 3). Side-by-side comparison of truncated murine and human complexes yields a decrease in the bend angle ($\Delta\theta = 10°$) similar to that observed upon binding of the intact domains in the same gel system ($\Delta\theta = 8°$; lines 3 and 4 in Table 3). Partial truncation of the human or murine tails impairs DNA bending to similar extents, indicating that their minimal HMG boxes differ in the intrinsic set point for DNA bending.

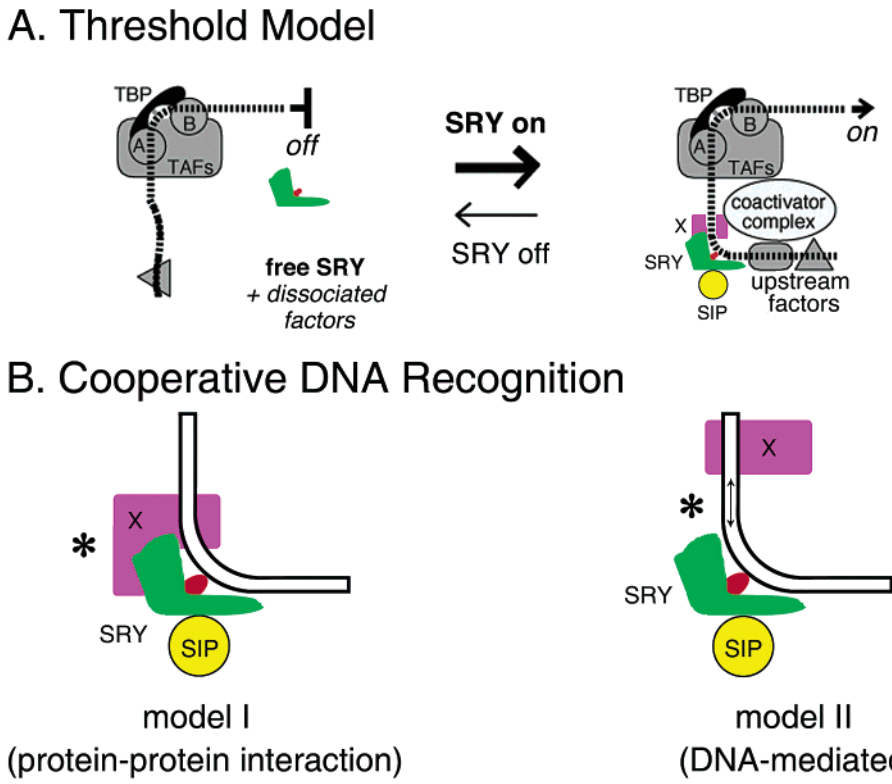


FIGURE 9: Proposed mechanism of SRY-directed architectural gene regulation. (A) Schematic model of SRY-directed assembly of a male-specific transcriptional preinitiation complex through sharp DNA bending. At the right, activated transcription occurs in the presence of bound SRY (green) as a specific DNA bend permits assembly of the stable activator−coactivator−basal preinitiation complex (enhanceosome; *50*). At the left, activated transcription is off in the absence of bound SRY due to disassembly of the DNA−multiprotein complex and dissociation of activator−coactivator complexes. Putative factor X (at right) is proposed to bind cooperatively with SRY to an adjoining DNA site. (B) Enlargement of the proposed SRY multiprotein−DNA complex. Cooperative binding of factor X, a putative sequence-specific DNA-binding protein (purple), is proposed to bind at a DNA site adjoining that of SRY. Cooperativity may be mediated either via protein−protein interactions (model I, e.g., by an SRY-interacting protein at the left) or via changes in the local DNA structure induced by SRY (model II at the right).

(isoleucine, methionine, and phenylalanine) suggest that the active "cantilever" metaphor (*48*) may be a misnomer: a side chain of suitable length and hydrophobicity fills a cavity between base pairs created by the overall HMG framework. Such tolerance appears to differ from the rigid constraints on cantilever shape imposed by the unrelated LacI/GalR class of major groove DNA-bending proteins (*49*). The dimensions of the potential cavity in the SRY−DNA complex may nonetheless exclude other side chains on the basis of length (alanine), shape (leucine), or size (tryptophan).

*Biological Implications.* How SRY functions to initiate testicular differentiation is not well understood. Its putative function as an architectural transcription factor is consistent with the presence of a specific HMG box but does not exclude alternative biochemical mechanisms (*20*). A variety of evidence, including analysis of human sex-reversal mutations, nonetheless supports the hypothesis that specific DNA bending plays an essential role in male-specific gene regulation (for a review, see ref *70*). We discuss the biological implications of our results in the context of this model.

These results suggest that SRY-directed sex reversal is robust with respect to enhanced DNA bending (in the $\Delta\theta$ range of ≈6−10°; Table 3), whereas a comparable decrease ($\Delta\theta \approx 13°$ as visualized by NMR; *15*) has been associated with XY gonadal dysgenesis (*18*). We propose that sharp DNA bending above a critical threshold is necessary for directing male-specific transcriptional regulation (Figure 9A). This model envisages that SRY-directed DNA bending within a permissive range of angles facilitates assembly of DNA−multiprotein preinitiation complexes ("enhanceosomes" and "repressosomes"; *50*, *51*), giving rise to architectural gene regulation (*9*, *52*). The notion of a threshold in DNA bending giving rise to a range of functional DNA bend angles is supported by studies of transcriptional activation in *E. coli*. Integration host factor (IHF), a specific DNA-bending protein that is able to induce 180° bends (*53*), facilitates open complex formation at downstream promoters through structural reorganization of DNA (*54*). An IHF DNA "U-turn" can functionally be replaced *in vivo* with a LEF-1-directed DNA bend of 117° (*54*). It is possible that the LEF-1-directed bend remains above the critical threshold but that further decreases in the angle of DNA bending would not be tolerated. Indeed, mutant IHF proteins have been identified in random genetic screens that have enhanced transcriptional activation activity but induce DNA bend angles of <180° (*55*).

A decreased angle of DNA bending below a critical threshold may impair transcriptional regulation. Functional analysis of nucleotide substitutions in the LEF-1-responsive enhancer element TCRa in T cells has shown that variant target sites bent to a reduced extent (90°) exhibit impaired transcriptional activation uncorrelated with effects of the substitutions on protein binding (*56*). Similar studies of amino acid substitutions in the HMG box of SOX2 and nucleotide substitutions in its 5′-TTTGTTT-3′ (and comple-

ment) target site demonstrated that spatially precise DNA bending is essential for its transcriptional regulation of the *fgf4* target gene (fibroblast growth factor 4; *57*). Variant SOX2-binding site 5′-TTTG**G**TT-3′ (and complement) exhibits a DNA bend angle of 42° rather than an angle of 80° as in the native complex with no change in phase orientation. Although SOX-2 binds well to the variant site, transactivation of the *fgf4* enhancer is blocked (*57*). Surprisingly, substitution of N48 in SOX-2 (residue 10 in the HMG box consensus; Figure 2) with glutamine enhances transcriptional activation despite impairing specific DNA binding by 8-fold and DNA bending by 34° (to a bend angle of 46°). However, the N48Q SOX2−DNA complex exhibits a change in phase orientation, which is proposed to rescue transactivation by optimizing other architectural features of the enhanceosome (*57*). Testing of this model awaits the biochemical reconstitution of a SOX2-directed enhanceosome.

The protocol developed in transgenic XX mice for phenotypic analysis of chimeric *mSry* genes containing "swapped" HMG boxes (*19*, *20*) provides a powerful approach for analysis of structure−function relationships in the murine SRY HMG box. The diversity of HMG boxes functionally tolerated in chimeric *mSRY* transgenes contrasts with the subtle biochemical deficits observed or inferred in a subset of sex-reversal human SRY variants (*18*), including mutations or deletions outside of the HMG box (*58*, *59*). It is possible that biochemical differences occur between molecular mechanisms of sex determination in the two species, such as might result from polymorphisms among putative SRY-interacting proteins (*60*), phosphorylation sites (*61*), expression levels, and rates of degradation. Irrespective of such possible differences, it would be of interest in the future to investigate how a wide range of DNA binding properties are compatible with male sex determination in the mouse. It may be possible, for example, to define upper and lower bounds for allowed DNA bend angles through the use of chimeric SRY−LEF-1 HMG boxes. When target genes for SRY are characterized, a combination of genetic and biochemical approaches may enable dissection of this and related architectural requirements of SRY-directed enhanceosome assembly. Such a multiprotein−DNA complex would provide a model for the integration of sex- and tissue-specific signals in a developmental switch.

## ACKNOWLEDGMENT

This is a contribution from the Cleveland Center for Structural Biology. We thank W. Olson and A. R. Srinivasan for analysis of DNA probe and linker configurations in FRET studies, G. Reddy for mass spectrometry, P. K. Donahoe, and C.-Y. King for discussion, and K. Hallenga, Q. X. Hua, and D. N. Jones for advice on NMR methods.

## SUPPORTING INFORMATION AVAILABLE

Seven figures showing DNA binding properties for native murine domain and C41S analogue, CD spectra of HMG boxes, PGE bending studies of intact human SRY in relation to the isolated HMG box, slowly exchanging bound and free DNA sites, additional NOESY spectra of human and murine SRY−DNA complexes, and a molecular model of cantilever insertion and four tables giving selected interproton distances at the cantilever insertion site, $^1$H NMR imino chemical shifts, a summary of previous PGE-based DNA bending studies, and the steady-state anisotropies of the free DNA and the hSRY− and mSRY−DNA complexes. This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES

1. Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Munsterberg, A., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1990) A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes, *Nature 346*, 245−250.
2. Berta, P., Hawkins, J. R., Sinclair, A. H., Taylor, A., Griffiths, B. L., Goodfellow, P. N., and Fellous, M. (1990) Genetic evidence equating SRY and the testis-determining factor, *Nature 348*, 448−450.
3. Haqq, C. M., King, C. Y., Ukiyama, E., Falsafi, S., Haqq, T. N., Donahoe, P. K., and Weiss, M. A. (1994) Molecular basis of mammalian sexual determination: activation of Mullerian inhibiting substance gene expression by SRY, *Science 266*, 1494−1500.
4. Werner, H. M., Huth, J. R., Gronenborn, A. M., and Clore, G. M. (1995) Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex, *Cell 81*, 705−714.
5. Ner, S. S. (1992) HMGs everywhere, *Curr. Biol. 2*, 208−210.
6. Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1991) Male development of chromosomally female mice transgenic for Sry, *Nature 351*, 117−121.
7. van Genderen, C., Okamura, R. M., Farinas, I., Quo, R. G., Parslow, T. G., Bruhn, L., and Grosschedl, R. (1994) Development of several organs that require inductive epithelial-mesenchymal interactions is impaired in LEF-1-deficient mice, *Genes Dev. 8*, 2691−2703.
8. Verbeek, S., Izon, D., Hofhuis, F., Robanus-Maandag, E., te Riele, H., van de Wetering, M., Oosterwegel, M., Wilson, A., MacDonald, H. R., and Clevers, H. (1995) An HMG-box-containing T-cell factor required for thymocyte differentiation, *Nature 374*, 70−74.
9. Grosschedl, R. (1995) Higher-order nucleoprotein complexes in transcription: analogies with site-specific recombination, *Curr. Opin. Cell Biol. 7*, 362−370.
10. Weir, H. M., Kraulis, P. J., Hill, C. S., Raine, A. R., Laue, E. D., and Thomas, J. O. (1993) Structure of the HMG box motif in the B-domain of HMG1, *EMBO J. 12*, 311−319.
11. Read, C. M., Cary, P. D., Crane-Robinson, C., Driscoll, P. C., and Norman, D. G. (1993) Solution structure of a DNA-binding domain from HMG1, *Nucleic Acids Res. 21*, 3427−3436.
12. Jones, D. N., Searles, M. A., Shaw, G. L., Churchill, M. E., Ner, S. S., Keeler, J., Travers, A. A., and Neuhaus, D. (1994) The solution structure and dynamics of the DNA-binding domain of HMG-D from *Drosophila melanogaster*, *Structure 2*, 609−627.
13. King, C. Y., and Weiss, M. A. (1993) The SRY high-mobility-group box recognizes DNA by partial intercalation in the minor groove: a topological mechanism of sequence specificity, *Proc. Natl. Acad. Sci. U.S.A. 90*, 11990−11994.
14. Love, J. J., Li, X., Case, D. A., Giese, K., Grosschedl, R., and Wright, P. E. (1995) Structural basis for DNA bending by the architectural transcription factor LEF-1, *Nature 376*, 791−795.
15. Murphy, E. C., Zhurkin, V. B., Louis, J. M., Cornilescu, G., and Clore, G. M. (2001) Structural basis for SRY-dependent 46-X,Y sex reversal: modulation of DNA bending by a naturally occurring point mutation, *J. Mol. Biol. 312*, 481−499.
16. Whitfield, L. S., Lovell-Badge, R., and Goodfellow, P. N. (1993) Rapid sequence evolution of the mammalian sex-determining gene SRY, *Nature 364*, 713−715.
17. Giese, K., Pagel, J., and Grosschedl, R. (1994) Distinct DNA-binding properties of the high mobility group domain of murine and human SRY sex-determining factors, *Proc. Natl. Acad. Sci. U.S.A. 91*, 3368−3372.
18. Pontiggia, A., Rimini, R., Harley, V. R., Goodfellow, P. N., Lovell-Badge, R., and Bianchi, M. E. (1994) Sex-reversing mutations affect the architecture of SRY-DNA complexes, *EMBO J. 13*, 6115−6124.
19. Bergstrom, D. E., Young, M., Albrecht, K. H., and Eicher, E. M. (2000) Related function of mouse SOX3, SOX9, and SRY HMG domains assayed by male sex determination, *Genesis 28*, 111−124.

20. Lovell-Badge, R., Canning, C., and Sekido, R. (2002) in *The genetics and biology of sex determination* (Chadwick, D., and Good, J., Eds.) pp 4−22, John Wiley, West Sussex, U.K.
21. Pontiggia, A., Whitfield, S., Goodfellow, P. N., Lovell-Badge, R., and Bianchi, M. E. (1995) Evolutionary conservation in the DNA-binding and -bending properties of HMG-boxes from SRY proteins of primates, *Gene 154*, 277−280.
22. Ukiyama, E., Jancso-Radek, A., Li, B., Milos, L., Zhang, W., Phillips, N. B., Morikawa, N., King, C. Y., Chan, G., Haqq, C. M., Radek, J. T., Poulat, F., Donahoe, P. K., and Weiss, M. A. (2001) SRY and architectural gene regulation: the kinetic stability of a bent protein-DNA complex can regulate its transcriptional potency, *Mol. Endocrinol. 15*, 363−377.
23. Hinck, A. P., Walkenhorst, W. F., Westler, W. M., Choe, S., and Markley, J. L. (1993) Overexpression and purification of avian ovomucoid third domains in *Escherichia coli*, *Protein Eng. 6*, 221−227.
24. Haqq, C. M., King, C. Y., Donahoe, P. K., and Weiss, M. A. (1993) SRY recognizes conserved DNA sites in sex-specific promoters, *Proc. Natl. Acad. Sci. U.S.A. 90*, 1097−1101.
25. Kim, J., Zwieb, C., Wu, C., and Adhya, S. (1989) Bending of DNA by gene-regulatory proteins: construction and use of a DNA bending vector, *Gene 85*, 15−23.
26. Dawson, W. R., and Windsor, M. W. (1968) Fluorescence yields of aromatic compounds, *J. Phys. Chem. 72*, 3251−3260.
27. Sosnick, T. R., Fang, X., and Shelton, V. M. (2000) Application of circular dichroism to study RNA folding transitions, *Methods Enzymol. 317*, 393−409.
28. Beechem, J. M., and Haas, E. (1989) Simultaneous determination of intramolecular distance distributions and conformational dynamics by global analysis of energy transfer measurements, *Biophys. J. 55*, 1225−1236.
29. Haas, E. (1997) The problem of protein folding and dynamics: time-resolved dynamic non-radiative excitation energy transfer measurements, *IEEE-JSTC 2*, 1088.
30. Grinvald, A., and Steinberg, I. Z. (1974) On the analysis of fluorescence decay kinetics by the method of least squares, *Anal. Biochem. 59*, 583−598.
31. Peters, R., King, C. Y., Ukiyama, E., Falsafi, S., Donahoe, P. K., and Weiss, M. A. (1995) An SRY mutation causing human sex reversal resolves a general mechanism of structure-specific DNA recognition: application to the four-way DNA junction, *Biochemistry 34*, 4569−4576.
32. Förster, T. (1948) Zwischen Molekulare Energie Wanderung und Fluoreszenz, *Ann. Phys. 2*, 55−75.
33. Steinberg, I. Z. (1971) Long-Range Nonradiative Transfer of Electronic Excitation Energy in Proteins and Polypeptides, *Annu. Rev. Biochem. 40*, 83−114.
34. Haas, E., Katchalski-Katzir, E., and Steinberg, I. Z. (1978) Effect of the orientation of donor and acceptor on the probability of energy transfer involving electronic transitions of mixed polarization, *Biochemistry 17*, 5064−5070.
35. Van Der Meer, B. W., Coker, G. I., and Chen, S.-Y. (1994) *Resonance energy transfer: Theory and data*, VCH Publishers, New York.
36. Harley, V. R., Jackson, D. I., Hextall, P. J., Hawkins, J. R., Berkovitz, G. D., Sockanathan, S., Lovell-Badge, R., and Goodfellow, P. N. (1992) DNA binding activity of recombinant SRY from normal males and XY females, *Science 255*, 453−456.
37. Ferrari, S., Harley, V. R., Pontiggia, A., Goodfellow, P. N., Lovell-Badge, R., and Bianchi, M. E. (1992) SRY, like HMG1, recognizes sharp angles in DNA, *EMBO J. 11*, 4497−4506.
38. Ohndorf, U. M., Rould, M. A., He, Q., Pabo, C. O., and Lippard, S. J. (1999) Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins, *Nature 399*, 708−712.
39. Murphy, F. V., Sweet, R. M., and Churchill, M. E. (1999) The structure of a chromosomal high mobility group protein-DNA complex reveals sequence-neutral mechanisms important for non-sequence-specific DNA recognition, *EMBO J. 18*, 6610−6618.
40. Allain, F. H., Yen, Y. M., Masse, J. E., Schultze, P., Dieckmann, T., Johnson, R. C., and Feigon, J. (1999) Solution structure of the HMG protein NHP6A and its interaction with DNA reveals the structural determinants for non-sequence-specific binding, *EMBO J. 18*, 2563−2579.
41. Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Matthews, B. W. (1992) Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect, *Science 255*, 178−183.
42. Giese, K., Cox, J., and Grosschedl, R. (1992) The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures, *Cell 69*, 185−195.
43. Chow, C. S., Whitehead, J. P., and Lippard, S. J. (1994) HMG domain proteins induce sharp bends in cisplatin-modified DNA, *Biochemistry 33*, 15124−15130.
44. Collignon, J., Sockanathan, S., Hacker, A., Cohen-Tannoudji, M., Norris, D., Rastan, S., Stevanovic, M., Goodfellow, P. N., and Lovell-Badge, R. (1996) A comparison of the properties of Sox-3 with Sry and two related genes, Sox-1 and Sox-2, *Development 122*, 509−520.
45. Carlsson, P., Waterman, M. L., and Jones, K. A. (1993) The hLEF/TCF-1a HMG protein contains a context-dependent transcriptional activation domain that induces the TCRa enhancer in T cells, *Genes Dev. 7*, 2418−2430.
46. Werner, M. H., Bianchi, M. E., Gronenborn, A. M., and Clore, G. M. (1995) NMR spectroscopic analysis of the DNA conformation induced by the human testis determining factor SRY, *Biochemistry 34*, 11998−12004.
47. Parkhurst, K. M., Brenowitz, M., and Parkhurst, L. J. (1996) Simultaneous binding and bending of promoter DNA by the TATA binding protein: real time kinetic measurements, *Biochemistry 35*, 7459−7465.
48. Bianchi, M. E., and Lilley, D. M. (1995) DNA-protein interactions. Applying a genetic cantilever, *Nature 375*, 532.
49. Arvidson, D. N., Lu, F., Faber, C., Zalkin, H., and Brennan, R. G. (1998) The structure of PurR mutant L54M shows an alternative route to DNA kinking, *Nat. Struct. Biol. 5*, 436−441.
50. Maniatis, T., Falvo, J. V., Kim, T. H., Kim, T. K., Lin, C. H., Parekh, B. S., and Wathelet, M. G. (1998) Structure and function of the interferon-$\beta$ enhanceosome, *Cold Spring Harbor Symp. Quant. Biol. 63*, 609−620.
51. Adhya, S., Geanacopoulos, M., Lewis, D. E., Roy, S., and Aki, T. (1998) Transcription regulation by repressosome and by RNA polymerase contact, *Cold Spring Harbor Symp. Quant. Biol. 63*, 1−9.
52. Bewley, C. A., Gronenborn, A. M., and Clore, G. M. (1998) Minor groove-binding architectural proteins: structure, function, and DNA recognition, *Annu. Rev. Biophys. Biomol. Struct. 27*, 105−131.
53. Rice, P. A., Yang, S., Mizuuchi, K., and Nash, H. A. (1996) Crystal structure of an IGF-DNA complex: a protein-induced DNA U-turn, *Cell 87*, 1295−1306.
54. Parekh, B. S., and Hatfield, G. W. (1996) Transcriptional activation by protein-induced DNA bending: evidence for a DNA structural transmission model, *Proc. Natl. Acad. Sci. U.S.A. 93*, 1173−1177.
55. Engelhorn, M., and Geiselmann, J. (1998) Maximal transcriptional activation by the IHF protein of *Escherichia coli* depends on optimal DNA bending by the activator, *J. Mol. Microbiol. 30*, 431−441.
56. Giese, K., Pagel, J., and Grosschedl, R. (1997) Functional analysis of DNA bending and unwinding by the high mobility group domain of LEF-1, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12845−12850.
57. Scaffidi, P., and Bianchi, M. E. (2001) Spatially precise DNA bending is an essential activity of the sox2 transcription factor, *J. Biol. Chem. 276*, 47296−47302.
58. Tajima, T., Nakae, J., Shinohara, N., and Fujieda, K. (1994) A novel mutation localized in the 3′ non-HMG box region of the SRY gene in 46,XY gonadal dysgenesis, *Hum. Mol. Genet. 3*, 1187−1189.
59. Assumpcao, J. G., Benedetti, C. E., Maciel-Guerra, A. T., Guerra, G. J., Baptista, M. T. M., Scolfaro, M. R., and de Mello, M. P. (2002) Novel mutations affecting SRY DNA-binding activity: the HMB box N65H associated with 46,XY pure gonadal dysgenesis and the familial non-HMG box R301 associated with variable phenotypes, *J. Mol. Med. 80*, 782−790.
60. Poulat, F., Girard, F., Chevron, M. P., Goze, C., Rebillard, X., Calas, B., Lamb, N., and Berta, P. (1995) Nuclear localization of the testis determining gene product SRY, *J. Cell Biol. 128*, 737−748.
61. De Santa Barbara, P., Bonneaud, N., Boizet, B., Desclozeaux, M., Moniot, B., Sudbeck, P., Scherer, G., Poulat, F., and Berta, P. (1998) Direct interaction of SRY-related protein SOX9 and steroidogenic factor 1 regulates transcription of the human anti-Mullerian hormone gene, *Mol. Cell. Biol. 18*, 6653−6665.
62. Schmitt-Ney, M., Thiele, H., Kaltwasser, P., Bardoni, B., Cisternino, M., and Scherer, G. (1995) Two novel SRY mis-

sense mutations reducing DNA binding identified in XY females and their mosaic fathers, *Am. J. Hum. Genet. 56*, 862−869.
63. Poulat, F., Soullier, S., Goze, C., Heitz, F., Calas, B., and Berta, P. (1994) Description and functional implications of a novel mutation in the sex-determining gene SRY, *Hum. Mutat. 3*, 200−204.
64. Lundberg, Y., Ritzén, M., Harlin, J., and Wedell, A. (1998) Novel missense mutation (P131R) in the HMG box of SRY in XY sex reversal, *Hum. Mutat. Suppl. 1*, S328.
65. Veitia, R., Ion, A., Barbaux, S., Jobling, M. A., Souleyreau, N., Ennis, K., Ostrer, H., Tosi, M., Meo, T., Chibani, J., Fellous, M., and McElreavey, K. (1997) Mutations and sequence variants in the testis-determining region of the Y chromosome in individuals with a 46,XY female phenotype, *Hum. Genet. 99*, 648−652.
66. Thompson, J. F., and Landy, A. (1988) Empirical estimation of protein-induced DNA bending angles: applications to lambda site-specific recombination complexes, *Nucleic Acids Res. 16*, 9687−9705.
67. Harley, V. R., Lovell-Badge, R., and Goodfellow, P. N. (1994) Definition of a consensus DNA binding site for SRY, *Nucleic Acids Res. 22*, 500−501.
68. Dubin, R. A., and Ostrer, H. (1994) Sry is a transcriptional activator, *Mol. Endocrinol. 8*, 1182−1192.
69. Bowles, J., Cooper, L., Berkman, J., and Koopman, P. (1999) SRY requires a CAG repeat domain for male determination in *Mus musculus*, *Nat. Genet. 22*, 405−408.
70. Weiss, M. A. (2004) Molecular mechanisms of male sex determination: the enigma of SRY, in *DNA Conformation and Transcription* (Ohyama, T., Ed.) Landes Bioscience, Georgetown, TX (in press).

BI049920A