

Molecular Basis of Mammalian Sexual Determination: Activation of Müllerian Inhibiting Substance Gene Expression by SRY

Christopher M. Haqq,* Chih-Yen King,* Etsuji Ukiyama,
Sassan Falsafi, Tania N. Haqq, Patricia K. Donahoe,†
Michael A. Weiss†‡

The pathway of male sexual development in mammals is initiated by *SRY*, a gene on the short arm of the Y chromosome. Its expression in the differentiating gonadal ridge directs testicular morphogenesis, characterized by elaboration of Müllerian inhibiting substance (MIS) and testosterone. *SRY* and *MIS* each belong to conserved gene families that function in the control of growth and differentiation. Structural and biochemical studies of the DNA binding domain of *SRY* (the HMG box) revealed a protein-DNA interaction consisting of partial side chain intercalation into a widened minor groove. Functional studies of *SRY* in a cell line from embryonic gonadal ridge demonstrated activation of a gene-regulatory pathway leading to expression of *MIS*. *SRY* molecules containing mutations associated with human sex reversal have altered structural interactions with DNA and failed to induce transcription of *MIS*.

Sexual dimorphism in mammalian embryogenesis provides a model of a developmental switch. The signal event in the male pathway, differentiation of an indifferent (or bipotential) gonad into a testis (1), is regulated in early embryogenesis by the Y chromosome (2). A regulatory gene has been mapped to the short arm of the Y chromosome (interval 1A1) by molecular analysis of human sex reversal (XX males and XY females) (3). This gene encodes *SRY*, a DNA binding protein that is conserved among mammalian Y chromosomes and is similar to members of the HMG box family of transcription factors (4). Whereas mouse *Sry* and some other HMG box proteins contain a recognized transcription activation domain, other mammalian *SRY* proteins (including human) do not (4). The mechanism of action of *SRY* is unknown.

The mouse homolog of *SRY* is selectively expressed in the common gonadal primordia of male gonads just before morpho-

logic differentiation (3). Its transgenic expression in XX mice is sufficient to induce a cascade of downstream regulatory events that lead to formation of male-specific structures and regression of female primordia (5). Although the steps of this pathway and their regulatory relationships are incompletely delineated, testicular morphogenesis is characterized by secretion of *MIS* by Sertoli cells (6) and of testosterone by Leydig cells (1). *MIS* induces regression of the Müllerian duct (the anlagen of the uterus, fallopian tube, and upper vagina); androgenic steroids induce masculinization of the external genitalia and differentiation of the Wolffian duct (the anlagen of the vas deferens, seminal vesicles, and epididymis).

Progress in reproductive biology has traditionally been stimulated by interdisciplinary analysis of patients with altered or ambiguous sexual development. Together, observations of clinicians, geneticists, and biochemists have demonstrated that expression of the complete male phenotype requires a complex series of steps. Examples of clinical disorders and their underlying mechanisms include complete sex reversal caused by mutation or translocation of the *SRY* gene (7), retained Müllerian duct syndrome resulting from mutations in *MIS* (8), and male pseudohermaphroditism associated with defects in the androgen receptor (9) or testosterone 5 α -reductase (10).

We and others have used mutations in *SRY* in patients with 46,XY pure gonadal dysgenesis (7) to investigate initial molecular events in testicular differentiation. The components identified in this developmental pathway can define the mechanisms un-

derlying the molecular events that lead to sexual differentiation. Correlated studies of structure and function illustrate how such mutations affect this decision at the levels of protein folding, DNA recognition, and transcriptional regulation. In complementary studies of human or rat *MIS* promoters transfected with human *SRY* into an immortalized urogenital ridge cell line we investigated how *SRY* activates a male-specific transcription pathway. These investigations define a mechanism by which *SRY* recognizes DNA, demonstrate an *SRY*-dependent pathway of *MIS* gene expression, and provide evidence for an intervening factor or factors (designated *SRYIF*) interposed between *SRY* and activation of the *MIS* promoter.

The HMG Box: DNA Recognition by Side Chain Intercalation

SRY contains a conserved DNA-binding domain, the HMG box (11). Binding occurs primarily in the DNA minor groove (12) and induces a sharp bend (13). The structure of an HMG box contains three α helices with L-shaped orientation (14, 15); its angular protein surface presumably provides a template for DNA bending. Because the structure of *SRY* has not been determined, we have constructed a model of its DNA-binding domain on the basis of nuclear magnetic resonance (NMR) solution structure of rat HMG domain 1B (14, 16). With conservative assumptions, distance-geometry and simulated annealing (DG-SA) (17) yielded a single family of structures (Fig. 1A) that contained both similar secondary structure and hydrophobic core (18, 19). The models retained an angular tertiary structure, defining convex and concave surfaces (14). Mutations in *SRY* associated with human sex reversal (7) are clustered in the HMG box; these mutations (Fig. 1B) are predicted to occur both in the hydrophobic core and on the protein surface. Several of these mutations decrease specific DNA binding (7).

The inferred structure of the HMG box of *SRY* offers two possible surfaces for DNA recognition—convex (Model I) or concave (Model II) (Fig. 2A). The models predict

C. M. Haqq, E. Ukiyama, T. N. Haqq, and P. K. Donahoe are in the Pediatric Surgical Research Laboratories and Department of Surgery, Massachusetts General Hospital, Boston, MA 02114, and Harvard Medical School, Boston, MA 02115, USA. C.-Y. King and S. Falsafi are in the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA. M. A. Weiss is in the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, and Department of Medicine, Massachusetts General Hospital, Boston, MA 02114, USA.

*The first two authors contributed equally to this manuscript.

†To whom correspondence should be addressed.

‡Present address: Center for Molecular Oncology and Departments of Biochemistry and Molecular Biology and of Chemistry, The University of Chicago, Chicago, IL 60637, USA.

different associations between the DNA and protein. In Model I the DNA would bend toward the protein [as observed with the catabolite activator protein (CAP) from *Escherichia coli* (20) and the eukaryotic nucleosome core particle (21)]. In Model II the DNA would bend away from the protein [as observed with the TATA-binding protein (Fig. 1C) (22, 23)]. To distinguish between these models, we used a clinical mutation (I68T) (7) to identify a direct contact between the HMG box of SRY and the bent DNA site. This mutation caused sex reversal in a female patient with a 46XY

karyotype. The DNA used was a variant of a high-affinity SRY-binding site in the human MIS promoter (24–26).

```

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
5' G G G G T G A T T G T T C A G
3' C C C C A C T A A C A A G T C

```

NMR studies comparing complexes of wild-type and mutant proteins with DNA (27) revealed that an isoleucine side chain contacts the DNA minor groove and inserts between specific AT base pairs (Fig. 2C).

The distal methyl resonance of the isoleucine (δ -CH₃ in Fig. 2D) is shifted to high field, presumably by DNA ring currents, and exhibits nuclear Overhauser effect (NOE) contacts with base protons in the DNA interior (N³ imino protons of T8 and T9, circled in the central panel of Fig. 2B). Insertion of an isoleucine side chain by wild-type SRY disrupts base stacking (as indicated by attenuation of NOEs and loss of neighboring ring-current effects); base pairing is maintained (27). This interaction would be expected to alter torsion angles in the DNA backbone and the global direction of the double-helical axis.

To identify the inserted side chain and thus the DNA-binding surface, substitutions were introduced into each of the possible DNA-binding isoleucines [I68T and I90V in full-length human SRY (7)]. The mutant domains were properly folded, and their thermodynamic stabilities were similar to that of the wild-type protein (Fig. 2E). The NMR spectrum of the I90V complex retained an isoleucine spin system shifted to high field whereas this spin system was absent in the spectrum of the I68T complex (Fig. 2D). When the model was compared to similar NMR structures (14, 15), it was apparent that residue 68 projects from α helix 1 to define the inner crux of the convex surface (Figs. 1 and 2A); by contrast, residue 90 packs into the hydrophobic core. Assignment of I68 as the inserted side chain excludes Model I and supports Model II. A schematic model of an SRY-DNA complex (Fig. 1D), juxtaposes the concave surface of the HMG box with a bent DNA site. Docking of helix 1 and insertion of I68 is expected to widen the minor groove relative to that of B-DNA.

The inserted isoleucine is critical for DNA recognition. Specific binding to the ATTGTT site was tested in a gel-retardation assay (Fig. 2E). Whereas the I90V substitution did not affect specific DNA binding, the binding affinity of the I68T mutant was less than one-fiftieth that of the wild-type protein (28). The assay detected complexes with a broad range of electrophoretic mobilities (Fig. 2E, part a, lane 5), which presumably represent dissociating I68T protein-DNA complexes. Kinetic instability of the variant complex was directly demonstrated by ¹H-NMR spectroscopy. On addition of the I68T protein, DNA resonances exhibited fast exchange between free and bound chemical shifts, indicating an exchange lifetime of <5 ms at 25°C; in contrast, the wild-type complex exchanged slowly (lifetime >200 ms). We imagine that the inserted isoleucine serves to lock the DNA into a specific structure. Because the I68T substitution is associated with human sex reversal (46,XY pure gonadal dysgenesis), a correlation is obtained

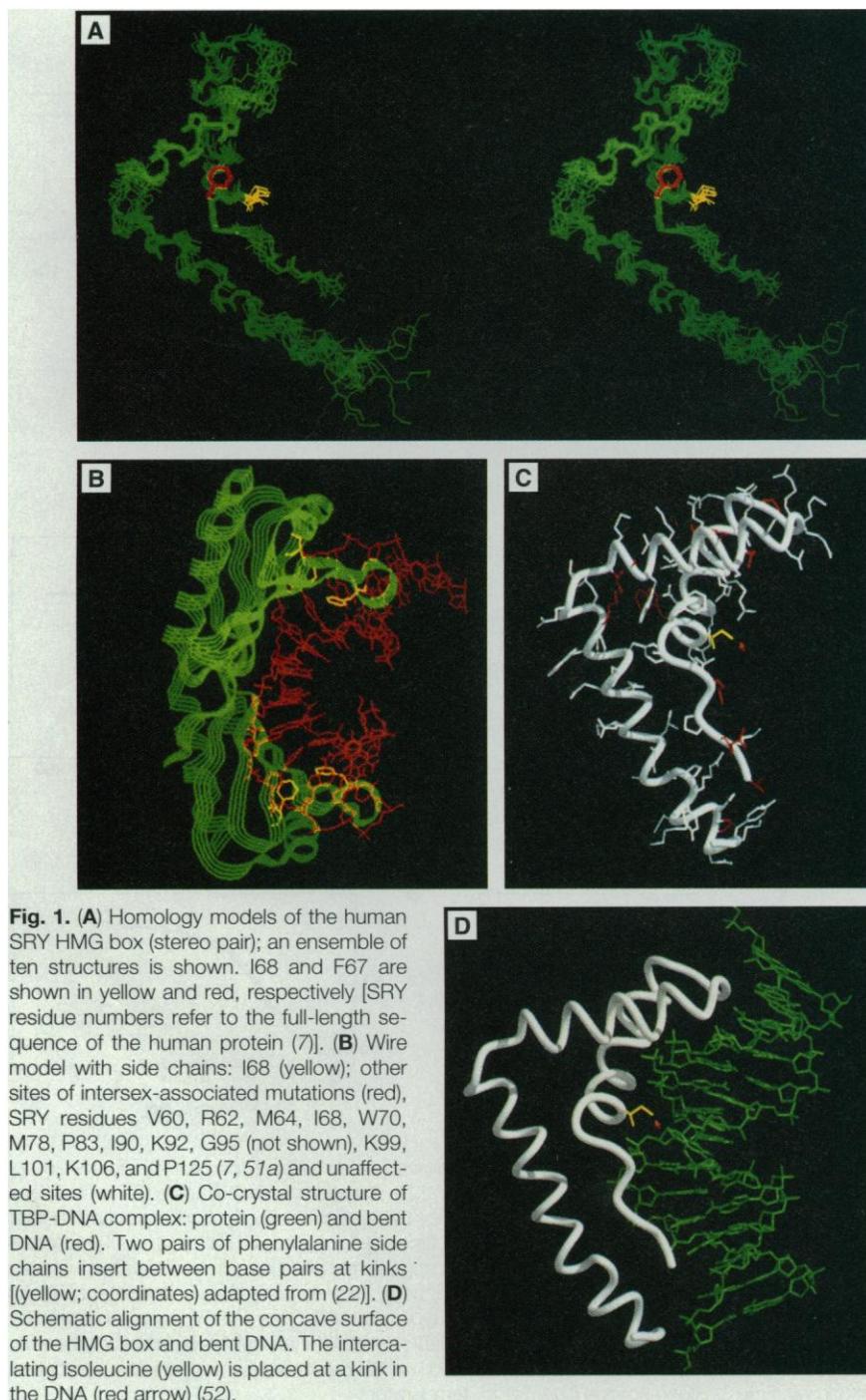


Fig. 1. (A) Homology models of the human SRY HMG box (stereo pair); an ensemble of ten structures is shown. I68 and F67 are shown in yellow and red, respectively [SRY residue numbers refer to the full-length sequence of the human protein (7)]. (B) Wire model with side chains: I68 (yellow); other sites of intersex-associated mutations (red), SRY residues V60, R62, M64, I68, W70, M78, P83, I90, K92, G95 (not shown), K99, L101, K106, and P125 (7, 51a) and unaffected sites (white). (C) Co-crystal structure of TBP-DNA complex: protein (green) and bent DNA (red). Two pairs of phenylalanine side chains insert between base pairs at kinks [(yellow; coordinates) adapted from (22)]. (D) Schematic alignment of the concave surface of the HMG box and bent DNA. The intercalating isoleucine (yellow) is placed at a kink in the DNA (red arrow) (52).

between a molecular mechanism and a developmental phenotype. However, position 90 can also be a site of clinical mutation [I90M (7)], presumably as a result of altered protein folding or stability (29).

Systematic variation of oligonucleotides around a protected site TTTGTG (24) of SRY-DNA interaction (Fig. 3A) led to identification of an optimal ATTGTT site (26, 27). Indeed, almost all single base changes from the ATTGTT site resulted in decreased

binding (Fig. 3B). Specificity was most stringent in the central TT (Fig. 2C), the site of penetration by the nonpolar I68 side chain (30). We also performed complementary studies of phosphate contacts (Fig. 3C). Site-specific interference with SRY binding was observed after incorporation of methylphosphonate, a neutral and nearly isosteric analog of the phosphodiester linkage (31). The results demonstrated that each phosphate in the ATTGTT site contributes to the pro-

tein-DNA interaction. This pattern—contiguous phosphate contacts across both strands—is different than that observed with major groove DNA-binding motifs (32). To align the extensive SRY-phosphate contacts along a single face would require partial unwinding of the double helix. The inferred DNA distortions (bending, partial unwinding, and widening of the minor groove) predict marked dispersion of ^{31}P -NMR resonances (33). This was indeed observed in the

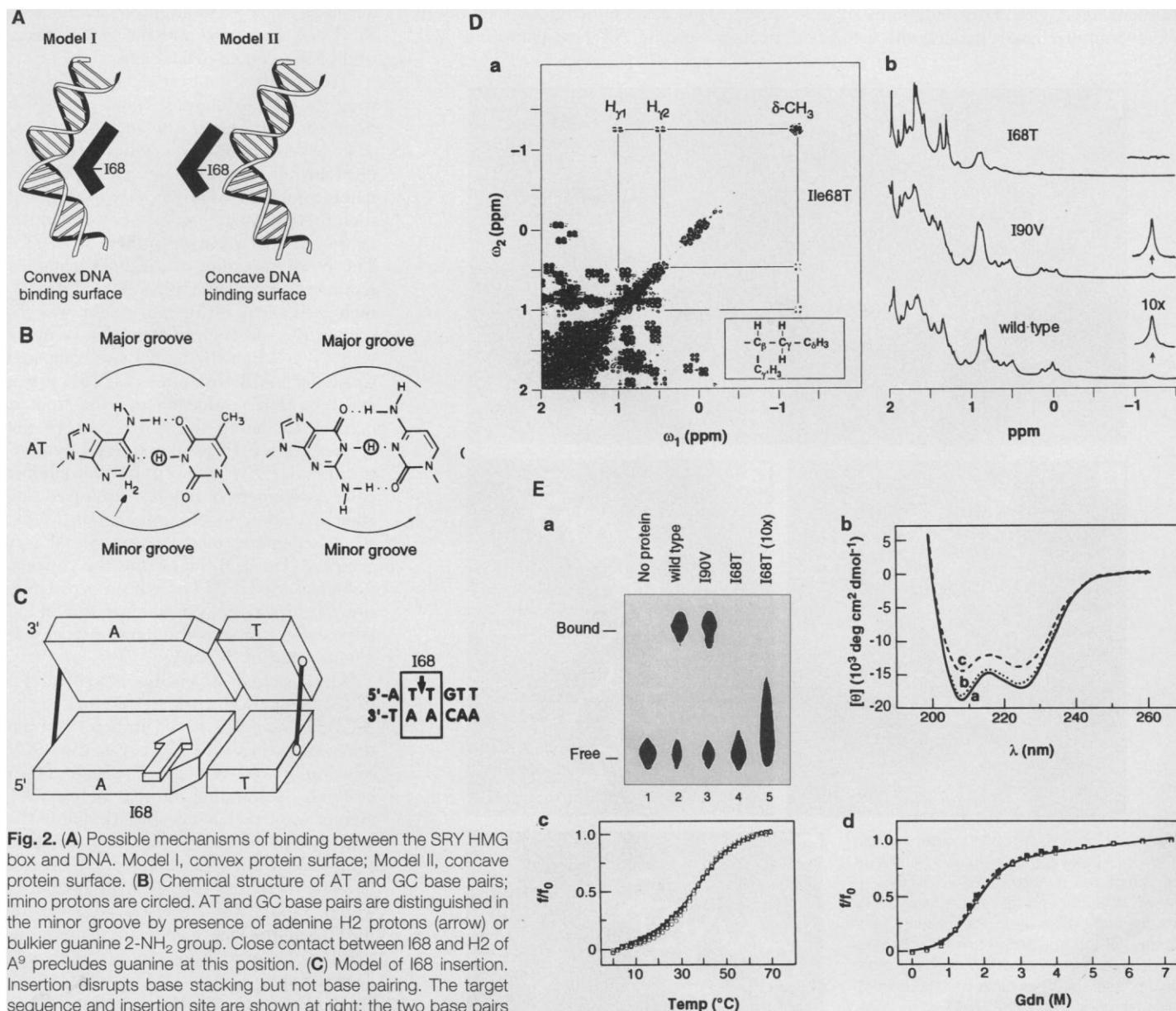


Fig. 2. (A) Possible mechanisms of binding between the SRY HMG box and DNA. Model I, convex protein surface; Model II, concave protein surface. (B) Chemical structure of AT and GC base pairs; imino protons are circled. AT and GC base pairs are distinguished in the minor groove by presence of adenine H2 protons (arrow) or bulkier guanine 2-NH₂ group. Close contact between I68 and H2 of A⁹ precludes guanine at this position. (C) Model of I68 insertion. Insertion disrupts base stacking but not base pairing. The target sequence and insertion site are shown at right; the two base pairs flanking the insertion site are boxed. Partial intercalation requires a widened minor groove (27). (D) (a) 500 MHz DQF-COSY ¹H-NMR spectrum of specific 1:1 complex between SRY HMG box and duplex DNA site 5'-GGGGTGATTTGTCAG-3'. The upfield δ-methyl and γ-methylene spin system of I68 is outlined. The pattern of the through-bonded ¹H-¹H correlation (solid line) is unique to the isoleucine side chain (53). (Inset) structure of Ile side chain with protons labeled. (b) Assignment of inserted isoleucine to SRY position 68 by site-directed mutagenesis (54–56). Aliphatic regions of 1D ¹H-NMR spectra of native (bottom), I90V (middle), and I68T (top) complexes; the amplitude (y-axis) of the upfield I68 δ-CH₃ resonance is enlarged tenfold as indicated in lower two spectra (arrows). (E) (a) Gel mobility-shift assay. Free

and bound bands are as indicated at left. Lane 1, no protein; lane 2, 25 nM wild-type SRY-p2; lane 3, 25 nM I90V variant; lane 4, 25 nM I68T variant; lane 5, 250 nM I68T variant (57). (b) CD spectra at 25°C of wild-type and mutant SRY domains: wild type [solid line (a)], I68T [dotted line (b)], and I90V [dashed line (c)]. The I68T substitution appears to be structurally conservative whereas I90V is associated with an apparent 25% decrease in α helix content (58). (c) Thermal denaturation of wild type (□), I68T (○), and I90V (Δ) as monitored by CD at 222 nm (59, 60). (d) Guanidine-HCl denaturation studies of wild type (Δ; solid line) and I68T (□; dashed line) SRY domains as monitored by tryptophan fluorescence (61). Temp, temperature.

function as a transcription factor, SRY-dependent activation of the MIS promoter in an embryonic gonadal cell line is indirect. Therefore, we postulate the existence of an intervening SRY-induced factor or factors, designated SRYIFs, that transduce the SRY signal to the responding MIS promoter.

The 114-bp MIS promoter contains regulatory regions conserved in the promoter of MIS from distantly related species. One such region (M2), that contains the steroidogenic factor-1 (SF-1) MIS-RE-1 site, can bind tissue-specific and developmentally regulated factors (42). SF-1 is a candidate SRYIF on the basis of observations that (i) SF-1 activates MIS transcription in 20-day primary postnatal Sertoli cells (43) and (ii) transgenic male homozygously deleted of SF-1 lack gonadal development (44). In the CH34 cell line, however, SF-1 failed either alone or in combination with SRY to elicit expression of the MIS reporter (Fig. 4B). In fact, SF-1 decreased basal expression from the MIS promoter to 35% of that of the antisense or empty vector controls (Fig. 4B). This inhibition and the ability of SF-1 to block MIS induction by SRY indicates that SF-1 may be a transcriptional repressor in this context.

Further evidence that SF-1 is a transcriptional repressor in the urogenital ridge cell line came from mutagenesis of the MIS-RE-1 site. Mutation of the TC at positions -95 and -96 to CT caused a decrease in binding of SF-1 to the MIS target site in gel shift experiments (43). With a 114-bp MIS reporter containing this mutation (40), however, SRY enhanced transcription 39-fold rather than the 15-fold, as seen with the wild-type promoter.

SRY activates a pathway that leads indirectly to expression of the MIS gene promoter. Such activation occurs in an appropriate developmental context, a male urogenital ridge cell line, through a small region of the proximal MIS promoter. Mutational analysis of that region should further delineate the mechanism of SRY action.

Summary and Perspectives

Molecular and cellular model systems have been used to address mechanisms by which SRY recognizes specific sites in DNA and activates downstream gene expression, and to examine how structure and function can be correlated in testis determination. The information currently available supports the following molecular principles: (i) SRY binds and bends DNA by means of the concave surface of an α -helical structure (Fig. 2A, Model II). Sequence specificity requires insertion of a nonpolar side chain through the DNA minor groove, disrupting base stacking but not base pairing. In human SRY, mutation of this isoleucine is associat-

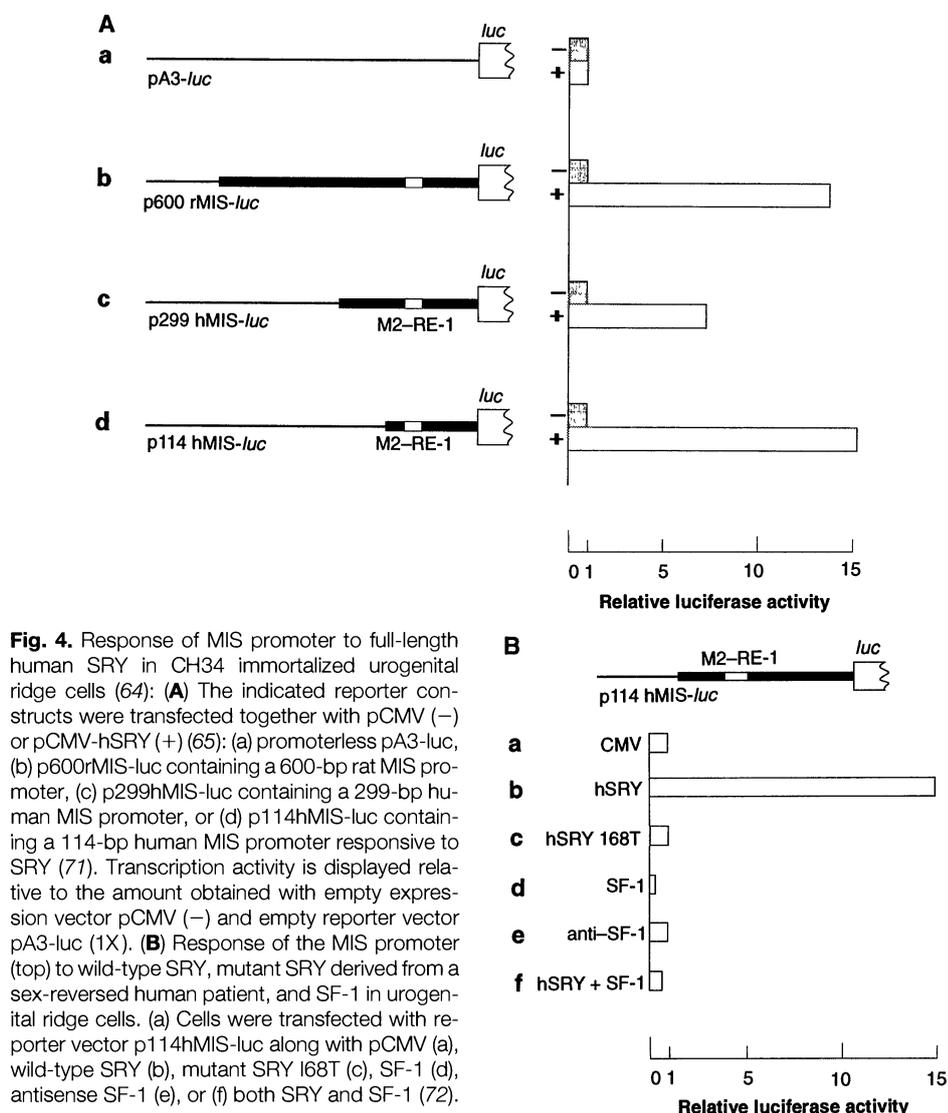


Fig. 4. Response of MIS promoter to full-length human SRY in CH34 immortalized urogenital ridge cells (64): **(A)** The indicated reporter constructs were transfected together with pCMV (-) or pCMV-hSRY (+) (65): (a) promoterless pA3-luc, (b) p600rMIS-luc containing a 600-bp rat MIS promoter, (c) p299hMIS-luc containing a 299-bp human MIS promoter, or (d) p114hMIS-luc containing a 114-bp human MIS promoter responsive to SRY (71). Transcription activity is displayed relative to the amount obtained with empty expression vector pCMV (-) and empty reporter vector pA3-luc (1X). **(B)** Response of the MIS promoter (top) to wild-type SRY, mutant SRY derived from a sex-reversed human patient, and SF-1 in urogenital ridge cells. (a) Cells were transfected with reporter vector p114hMIS-luc along with pCMV (a), wild-type SRY (b), mutant SRY I68T (c), SF-1 (d), antisense SF-1 (e), or (f) both SRY and SF-1 (72).

ed with clinical sex reversal. (ii) Expression of SRY in an embryonic testicular cell line initiates transcription of MIS indirectly through the 114-bp promoter. (iii) Mutation of full-length SRY at position 68, the site of partial side chain intercalation, abolishes its ability to induce transcription of MIS. (iv) Analysis of regulation through the proximal MIS promoter indicates that intervening transcription factors exist that may directly interact with the 114-bp region of the MIS promoter (Fig. 5) or influence the basal transcription machinery.

Sexual differentiation in mammals requires a precise choreography of molecular and cellular events. The initial step—establishment of gonadal sex in the previously undifferentiated urogenital ridge—provides a general model of a genetic switch in organogenesis. Identification of SRY as the testis determining factor (3) permits biochemical characterization of this switch. Its expression is appropriately specific for sex (male), tissue (urogenital ridge), and stage

of development (just prior to morphologic differentiation of the testis) (3). Although its biochemical mechanism is not well understood, SRY belongs to a family of HMG transcription factors and can itself activate transcription (4, 45). Selective binding of SRY to responsive promoters by the mechanisms we have observed may be further enhanced by local changes in DNA structure (as influenced by superhelical density and nucleosome phasing) and by cooperative binding of other transcription factors (46, 47). Sharp DNA bends induced by SRY binding are proposed to organize the higher order structure of promoters (12, 48), thereby modulating assembly of more complex protein-protein interactions.

Molecular and cellular model systems can provide insight into mechanisms of gene regulation, but do not capture the complexities of organogenesis. How do such molecular pathways, once defined, orchestrate cell-cell interactions in testicular differentiation? The mature testis is a dynamic

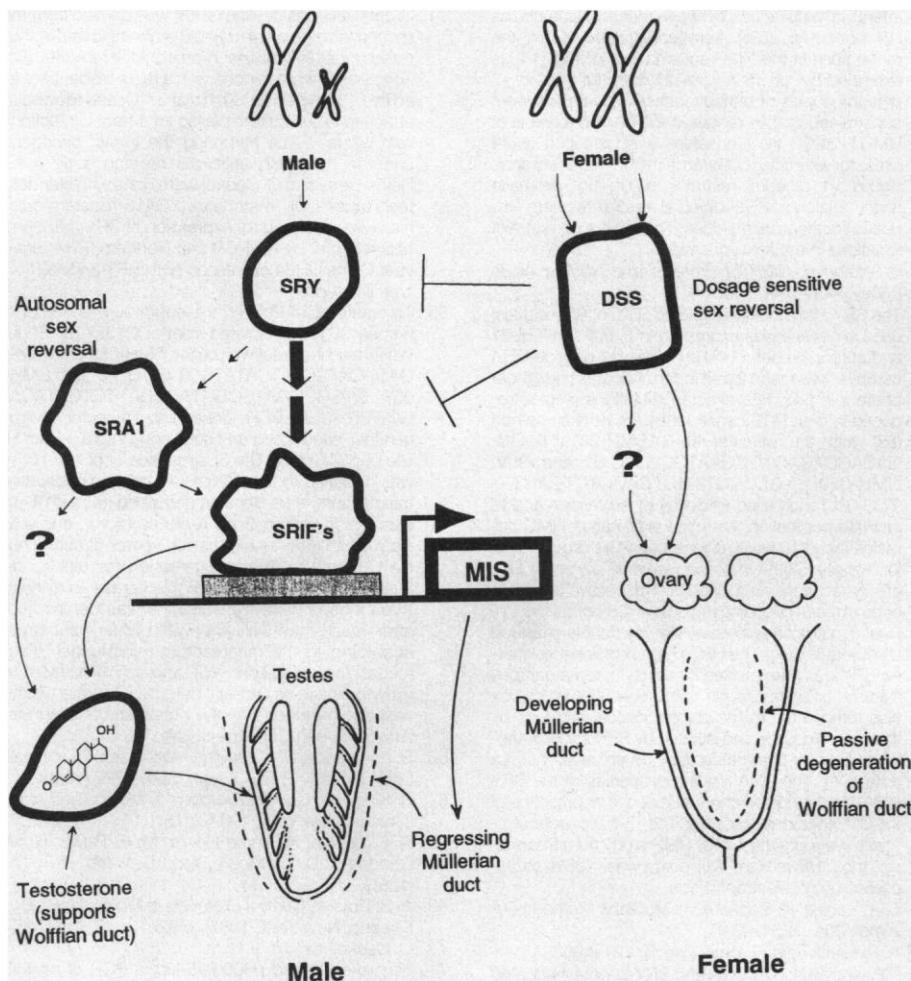


Fig. 5. A hypothetical scheme for sex determination in mammals. The male regulatory pathway is initiated by the sex regulator on the Y chromosome, SRY, which acts to influence gonadal morphogenesis leading from an indifferent pattern of support cells surrounding germ cells to a male pattern of dispersed support cells surrounding germ cells in seminiferous tubules. A factor or factors (SRYIFs) mediates induction of the MIS gene by SRY. Although we have illustrated SRYIF interacting with the MIS promoter, an indirect interaction is also possible, for example, through modification of the basal transcription machinery. MIS, in turn, plays a key role in male sexual development as a diffusible substance causing regression of the female Müllerian duct derivatives—the uterus, fallopian tubes, and vagina.

matrix of Sertoli, Leydig, myeloid, and germ cells that provides an environment for elaboration of growth and differentiating factors and for spermatogenesis in adult life. Future dissection of testicular morphogenesis will require identification of those genes intermediary between SRY and MIS in the male pathway (Fig. 5). SRY is proposed to directly activate target genes that may include, for example, the autosomal sex reversal factor SRA-1 [linked to campomelic dysplasia on chromosome 17q in humans (49)], or the canine autosomal sex reversal gene (50). Complementary analysis of sex-reversed females with intact SRY has led to identification of a locus on the short arm of the X chromosome at which duplication is associated with female differentiation. This locus, designated DSS (dosage-sensitive sex reversal), may act either by repressing the SRY-induced male pathway or by activating

expression of genes required for formation of female structures (51). SRYIF may be one of several factors participating with SRY in the determination of gonadal sex, as suggested by the phenomenon of sex reversal associated with abnormalities on mouse chromosome 17 T (51). Combined application of mammalian genetics, biochemistry, and molecular biology promises to define more fully the initial, intermediary, and distal downstream steps of the male and female developmental programs. Such information may have broad clinical implications for infertility, contraception, and control of gonadal tumors.

REFERENCES AND NOTES

1. A. Jost, *Arch. Anat. Microsc. Morphol. Exp.* **36**, 271 (1947); *ibid.* **61**, 415 (1962); T. Taketo, J. Saeed, Y. Nishiola, P. Donahoe, *Dev. Biol.* **146**, 386 (1991).
2. P. A. Jacobs and J. A. Strong, *Nature* **183**, 302

- (1959); A. McLaren, *ibid.* **346**, 216 (1990); A. de la Chappelle, *Am. J. Hum. Genet.* **24**, 71 (1972).
3. D. C. Page et al., *Cell* **51**, 1091 (1987); M. S. Palmer et al., *Nature* **342**, 937 (1990); A. H. Sinclair et al., *ibid.* **346**, 240 (1990); J. Gubbay et al., *ibid.* p. 245; P. Koopman et al., *ibid.* **348**, 450 (1990).
4. V. Laudet et al., *Nucleic Acids Res.* **21**, 2493 (1993); R. A. Dubin and H. Oster, *Mol. Endocrinol.* **8**, 1182 (1994); K. Giese and R. Grosschedl *EMBO J.* **12**, 4667 (1993).
5. P. Koopman, J. Gullay, N. Vivian, P. Goodfellow, R. Lovell-Badge, *Nature* **351**, 117 (1991).
6. R. L. Cate, P. K. Donahoe, D. T. MacLaughlin, in *Handbook of Experimental Pharmacology*, M. B. Sporn and A. B. Roberts, Eds. (Springer-Verlag, Berlin, 1990), vol. 95/II, chap. 25; M. M. Lee and P. K. Donahoe, *Endocr. Rev.* **14**, 152 (1993).
7. V. R. Harley et al., *Science* **255**, 453 (1992); P. Berta et al., *Nature* **348**, 448 (1990); J. R. Hawkins et al., *Am. J. Hum. Genet.* **51**, 979 (1992); J. R. Hawkins et al., *Hum. Genet.* **88**, 471 (1992); P. N. Goodfellow and R. A. Lovell-Badge, *Annu. Rev. Genet.* **27**, 71 (1993); E. Villain, F. Jaubert, M. Fellous, K. McElreavey, *Differentiation* **52**, 151 (1993); J. Maller, M. Schwartz, N. E. Skakkebaek, *J. Clin. Endocrinol. Metab.* **75**, 331 (1992); R. J. Jager, V. R. Harley, R. A. Pfeiffer, P. N. Goodfellow, G. Scherer, *Hum. Genet.* **90**, 350 (1992). Abbreviation code for amino acids: I, isoleucine; K, lysine; R, arginine; W, tryptophan; L, leucine; V, valine; T, threonine; M, methionine; and P, proline.
8. D. Guerrier et al., *J. Clin. Endocrinol. Metab.* **68**, 46 (1989); B. Knebelmann et al., *Proc. Natl. Acad. Sci. U.S.A.* **88**, 3767 (1991).
9. D. B. Lubahn et al., *Science* **240**, 327 (1988); C. Chang, J. Kokontis, S. Liao, *ibid.* **240**, 324 (1988); W. D. Tilley et al., *Proc. Natl. Acad. Sci. U.S.A.* **86**, 327 (1989).
10. S. Andersson, R. W. Bishop, D. W. Russel, *J. Biol. Chem.* **264**, 16249 (1989).
11. The HMG motif identifies two classes of nuclear proteins. Class I members, such as HMG1 and UBF1, contain two or more HMG boxes and bind to sites of altered DNA structure. Class II proteins are sequence-specific and contain a single box; an example is the lymphoid-specific transcription factor LEF-1.
12. K. Geise, J. Box, R. Grosschedl, *Cell* **69**, 185 (1992).
13. S. Ferrari et al., *EMBO J.* **11**, 4497 (1992).
14. H. M. Weir et al., *ibid.* **12**, 1311 (1993).
15. D. N. M. Jones et al., *Structure* **2**, 609 (1994); C. M. Read et al., *Nucleic Acids Res.* **21**, 3427 (1993).
16. SRY and HMG1 box b share 25% sequence identity and 45% sequence similarity.
17. T. F. Havel and M. E. Snow, *J. Mol. Biol.* **217**, 1 (1991).
18. Correspondence of α -helical segments is supported by neural net analysis of secondary structure (19) and circular dichroism (Fig. 2C). Internal helix-helix contacts in the NMR structure of HMG1 box B correspond to nonpolar residues in the SRY sequence.
19. L. H. Holley and M. Karplus, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 10 (1989).
20. S. C. Schultz, G. C. Shields, T. A. Steitz, *Science* **253**, 1001 (1991).
21. T. J. Richmond, J. T. Finch, B. Rushton, D. Rhodes, A. Klug, *Nature* **311**, 532 (1984).
22. J. L. Kim, D. B. Nikolov, S. K. Burley, *ibid.* **365**, 520 (1993).
23. Y. Kim, J. H. Geiger, S. Hahn, P. B. Sigler, *ibid.*, p. 512.
24. C. M. Haqq, C.-Y. King, P. K. Donahoe, M. A. Weiss, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1097 (1993).
25. The ATGTGT target site has also been identified by random binding-site selection (26).
26. V. R. Harley, R. Lovell-Badge, P. N. Goodfellow, *Nucleic Acid Res.* **22**, 1500 (1994).
27. C.-Y. King and M. A. Weiss, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 11990 (1993).
28. R. Peters, C.-Y. King, E. Ukiyama, S. Falsafi, P. K. Donahoe, M. A. Weiss, unpublished results.
29. The 190M allele exhibits variable penetrance and thus may be inherited (7). The mutant domain exhibits temperature-dependent unfolding between 35 and 39°C (F. Tao and M. A. Weiss, unpublished results).
30. Specificity for TT is presumably mediated in part by

- DNA distortability and in part by nonpolar contacts in the minor groove.
31. P. S. Miller and P. O. P. Ts'o, *Anti-Cancer Drug Des.* **2**, 117 (1987); M. C. Botfield and M. A. Weiss, *Biochemistry* **33**, 2349 (1994).
 32. S. C. Harrison and A. K. Aggarwal, *Annu. Rev. Biochem.* **59**, 533 (1990).
 33. D. G. Gorenstein, *Methods Enzymol.* **177**, 295 (1989).
 34. S. Hahn, S. Buratowski, P. A. Sharp, L. Guarente, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5718 (1989); V. L. Singer, C. R. Wobbe, K. Struhl, *Genes Dev.* **4**, 636 (1990); P. W. J. Rigby, *Cell* **72**, 7 (1993).
 35. I68 in human SRY is replaced by phenylalanine in other primate sequences (36), yielding a Phe pair analogous to those in TBP (Fig. 1C).
 36. L. S. Whitfield, R. Lovell-Badge, P. N. Goodfellow, *Nature* **364**, 713 (1993).
 37. We used the oncogene *v-myc* to immortalize cells harvested from embryonic rat urogenital ridges early on day 14 at a time appropriate for SRY expression in this species. Individual clones were examined for expression of SRY by reverse transcription-polymerase chain reaction (RT-PCR) and northern (RNA) analysis, and one positive line, CH34, was selected for use in subsequent cotransfection assays.
 38. R. Cate *et al.*, *Cell* **45**, 685 (1986); J.-Y. Picard, R. Benarous, D. Guerrier, N. Josso, A. Kahn, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5464 (1986).
 39. A. Munsterberg and R. Lovell-Badge, *Development* **13**, 613 (1991); T. Zwingman, R. P. Erickson, T. Boyer, A. Ao, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 814 (1993); S. Hirobe, W. W. He, M. M. Lee, P. K. Donahoe, *Endocrinology* **131**, 854 (1992).
 - 39a. C. M. Haqq and P. K. Donahoe, unpublished observations.
 40. Mutant plasmids were constructed as described (Fig. 4A) with new PCR primers. For mutation of nucleotides -95 and -96, primers CMH-141 5'-GGTACCGGGCACTGTCCCAAGGCT-GCGGCA GAGG-3' and CMH-107 were used. To construct the SRYe mutation, the promoter was rebuilt from two PCR products ligated at the Apa I restriction site. The 5' fragment used primers CMH-128 with CMH-119 5'-GATGGGCCAGGACAGACCCCTATCTCTCTGC-3' and the 3' fragment was made from annealing oligomers CMH-111 5'-CGCGGGCCCCACCCACCTTCCACTCGGCTCACCTAAGGCAGGCAGCCAGCCCTGG-CAGCAACCAAGCTTCCG-3' and CMH-112 5'-CGGAAGCTTGGGTGCTGCCAGGGCTGGGCTGCCTGCCTTAAGTGAGCCGAGTGAAGGTGGGGTGGGGCCCGCG-3'. The resulting promoter was subcloned into the Hind III site of pA3-luciferase (41).
 41. I. H. Maxwell *et al.*, *Biotechniques* **7**, 276 (1989); W. M. Wood *et al.*, *J. Biol. Chem.* **264**, 14840 (1989).
 42. C. Haqq, R. Cate, P. Donahoe, *J. Cell. Biochem.* **14C**, 301 (1990).
 43. W.-H. Shen, C. Moore, Y. Ikeda, K. L. Parker, H. A. Ingraham, *Cell* **77**, 651 (1994).
 44. X. Luo, Y. Ikeda, K. L. Parker, *Cell* **77**, 481 (1994).
 45. D. R. Cohen, A. H. Sinclair, J. D. McGovern, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4372 (1994).
 46. As observed for HMG-1 in the cooperative assembly of a preinitiation complex on the β -interferon promoter (47).
 47. W. Du, D. Thanos, T. Maniatis, *Cell* **74**, 887 (1993).
 48. M. van de Wetering and H. Clevers, *EMBO J.* **11**, 3039 (1992).
 49. B. Bardoni *et al.*, *Nature Genet.* **7**, 497 (1994).
 50. V. Meyers-Wallen *et al.*, personal communication.
 51. L. L. Washburn, B. K. Lee, E. M. Eicher, *Genet. Res.* **56**, 185 (1990); N. Tommerup *et al.*, *Nature Genetics* **4**, 170 (1993).
 - 51a. D. C. Page, personal communication.
 52. Figures were generated with the Quanta (Molecular Simulations, Inc.) and Mitus (UCSF Graphics Laboratory) software packages. The DG-SA homology model was built with the DGI program (T. F. Havel, Harvard Medical School). We used 22,005 distance and 226 dihedral restraints as follows: (i) d_{NN} , ($i, i+3$), ($i, i+4$), and (ϕ, ψ) restraints (total 349) imposed corresponding α -helical segments. Distances were constrained to within $\pm 10\%$ of their ideal values; dihedral angles were constrained to within $\pm 20^\circ$. (ii) Interatomic distances between identical side chains (19 positions; 25%) were constrained within the range seen in the NMR ensemble of HMG1 (14) as extended by $\pm 0.15 \text{ \AA}$ (total 21,698). (iii) χ_1 and χ_2 dihedral angles of buried identical side chains were constrained to their ranges in the NMR ensemble of HMG1 $\pm 20^\circ$. (iv) Conservative substitution (eight sites; for example, L69V and Y109F) received analogous χ_1 dihedral restraints $\pm 20^\circ$. For aromatic rings, analogous χ_2 dihedral restraints were imposed for identical residues. There were no restraint violations in the final ensemble.
 53. K. Wüthrich, *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986).
 54. The SRY HMG box (designated SRY-p; 85 residues) used in panel a was expressed in *E. coli* strain BL21 and purified as described (27). The 1:1 peptide-DNA complex was made 2 mM in 10 mM potassium phosphate (pH 6.0). Site-directed mutations were introduced into an M13 construct by the Kunkel method (55) with the primers: I68T, CMH-081a 5'-GC-GAGACCACACGGCGAATCGTTC-3'; and I90V, CMH-079 5'-GCTGCTTGTCTACCTCTGAGTTCG-3'. To increase efficiency of expression and to facilitate purification, wild-type and mutant HMG domains (panel b) were recloned by PCR into pTSN53 [J. Markley (56)]. A thrombin cleavage site and a His₆ affinity tag were introduced into respective 5' and 3' ends of the SRY coding region by PCR primer design. Final constructions were verified by double-stranded DNA sequencing. The His₆ fusion proteins (designated SRY-p2) were purified by affinity chromatography (Ni-NTA column; Qiagen, Inc.). The SRY HMG box was removed from the staphylococcal nuclease by thrombin digestion and isolated by SP-Fast Flow cation exchange chromatography (Pharmacia, Inc., La Jolla, CA). The DNA binding properties of the SRY HMG box are identical in the absence or presence of the COOH-terminal His₆ tag. The 15-base oligonucleotides were purchased for NMR study from Pharmacia, Inc., (Milwaukee, WI); purity was >98% as assessed by gel electrophoresis.
 55. T. A. Kunkel, K. Bebenek, T. McClary, *Methods Enzymol.* **204**, 125 (1991).
 56. A. P. Hinck *et al.*, *Protein Eng.* **6**, 221 (1993).
 57. ³²P-labeled duplex DNA and SRY-p were incubated at 20°C in 10 mM potassium phosphate (pH 7.4), 50 mM KCl, 2 mM MgCl₂, 10% glycerol for more than 2 hours. Complexes were resolved on a polyacrylamide gel (6%). A 15-mer DNA site 5'-GGGGTGAT-TGTTGAG was used.
 58. CD spectra were obtained at 25°C with an Aviv spectropolarimeter; the protein was made 25 μ M in 200 mM NaCl, 10 mM sodium phosphate (pH 6.0). A 0.1-cm path-length cell was used and incubated for approximately 5 min before data acquisition. Final spectra were obtained by averaging 6 multiple-scan spectra acquired with a scan speed of 10 s at each wavelength.
 59. For thermal unfolding studies, samples in 0.1-cm CD cell were equilibrated for 10 min at 0°C followed by 2-min equilibration, and 60-s data acquisition. The procedure was repeated successively at 2° temperature increments. Theoretical curve (solid line) was fit by the least squares procedure to the equation, $\Delta G = \Delta H - (T/T_m)(\Delta H_1 - \Delta G_1) + \Delta C_p [T - T_m - T \times \ln(T/T_m)]$ (60). In each case $\Delta H_1 = 27 \pm 3 \text{ kcal/mol}$ and the T_m is approximately 39°C.
 60. J. M. L. Johnson and S. G. Frasier, *Methods Enzymol.* **117**, 301 (1985).
 61. Excitation wavelength 295 nm (slit width 5 nm) and emission wavelength 350 nm (slit width 10 nm) were used. Wild-type and mutant proteins were made 1 μ M in 50 mM NaCl, 10 mM NaH₂PO₄ (pH 6) in various concentrations of guanidine HCl. Experimental curves were fitted to the equations: $\Delta G = \Delta G_U + m \times [\text{Gdn}] = -RT \times \ln K_p$. An extinction coefficient of $E_{290\text{nm}} = 2.11 \times 10^4 \text{ mol}^{-1} \text{ cm}^{-1}$ was used for both wild-type and mutant samples. In the two-state model (60) ΔG_U for both wild type and I68T are $\approx 2.3 \text{ kcal/mol}$.
 62. Gel mobility-shift assays was performed as described in (57).
 63. Both samples were made 0.625 mM in D₂O with 2.33 mM potassium phosphate (pH 6.4) and 11.66 mM potassium chloride. The spectra were taken at 40°C with a 10-mm probe.
 64. Urogenital ridge cell line CH34 was derived from the urogenital ridge of 14-day-old embryonic rats euthanized under institutional protocol MGH #A3596-01. Forty ridges were minced with a razor blade, triturated in 0.1% trypsin for 30 min at 37°C, and allowed to attach to tissue culture plastic for 4 hours. Infections with $\psi 2$ retrovirus harboring the *v-myc* oncogene and the neomycin antibiotic resistance gene (S. Fields-Berry and C. Cepko) were done with standard techniques (70). Immortalized G418-resistant colonies were analyzed for expression of SRY, MIS, and MIS receptor by RT-PCR and Northern (RNA) analysis. Clone CH34 expressed both SRY and the MIS type II receptor.
 65. To create pCMV-hSRY, full-length 204-amino acid human SRY was isolated from pDP1327 (D. Page, Whitehead Institute) with oligos CMH-066 5'-ACCG-GATCCATGCAATCATATGCTTCTGC-3' and CMH-067 5'-GGCGGATCCGGTACCGATTGCTCTACAGCTTTG-3' for PCR, cloned into TA vector (Invitrogen Inc), sequenced and subcloned into the Bam HI site of pCMV (66). The 3' sequence of pCMV-hSRY was designed to alter T202P. All calcium phosphate transfections were done as described (67) with 1- μ g expression vector, 2- μ g reporter vector, and 0.5- μ g pXGH5 (68), an expression vector directing human growth hormone expression for use as an internal standard. Average Luciferase activity of three independent transfections was determined, after freeze-thaw lysis (69), with Luciferin substrate according to the manufacturers instruction (Promega, Inc, Madison, WI) and normalization to growth hormone activity detected by chemiluminescent immunoassay (Nichols Institute Diagnostics, San Juan Capistrano, CA).
 66. S. Andersson, D. N. Davis, H. Dahlback, H. Jornvall, D. W. Russell, *J. Biol. Chem.* **264**, 8222 (1989).
 67. R. Kingston, *Current Protocols in Molecular Biology* (Greene, New York, 1987), p. 9.1.1.
 68. R. F. Selden, K. Burke Howie, M. E. Rowe, H. M. Goodman, D. D. Moore, *Mol. Cell. Biol.* **6**, 3173 (1986).
 69. A. R. Brasier, *Current Protocols in Molecular Biology* (Greene, New York, 1987), p. 9.6.12.
 70. C. Cepko, *ibid.*, p. 9.10.1.
 71. To make plasmid p60rMIS-luc, a PCR of rat MIS phase DNA (24) with primers CMH-115 5'-CTAAAGCTTCTGCTCTGGAAACCCCTGTGGCCAG-3' and CMH-116 5'-CTAAAGCTTGGCAGGGTCTGGGTG-GTCCCTGC-3' was subcloned into TA vector, sequenced and subcloned into the Hind III site of pA3-luc (41). Using an identical strategy, primers CMH-106 5'-GGGAAGCTTGGCCGCTACTCCCA-CCTG-3' and CMH-107 5'-CGGAAGCTTGGGT-GCTGCCAGGGGCTG-3' were used to make human promoter construct p299hMIS-luc, and CMH-128 5'-GGTACCGGGCACTGTCCCAACCCCAAGG-3' and CMH-107 were used to construct p114hMIS-luc.
 72. SRY mutation I68T was made by site directed mutagenesis (55) with primer I68T CMH-081 5'-GC-GAGACCACACGGTGAATCGTTC-3'. PCR from mutant phage using CMH-066 and CMH-067 and subcloning into TA vector, sequencing, and subcloning into the Bam HI site of pCMV were as described for pCMV-hSRY, including the T202P design at the 3' terminus (Fig. 4A). SF-1 sense and antisense expression vectors were from K. Parker.
 73. We thank D. Russell for pCMV; A. Hinck and J. Markley for pTSN53; W. Wood for pA3-luc and -180 RSV-luc; S. Fields-Berry and C. Cepko for retrovirus $\psi 2$ -*v-myc* packaging cells; D. C. Page for pDP1327 and helpful advice; D. Moore for helpful advice; S. Nasserri for expert assistance; R. Kingston, C. Miller, W. Walker, and M. Lachenmann for advice; S. Burley for coordinates of TBP; J. Thomas for coordinates of HMG1; D. Moore, J. Habener, D. MacLaughlin, and L. Perkins for critically reviewing the manuscript; F. Tao for experiments in Fig. 2E; R. Peters for assisting in protein purification and in experiments in Figs. 2E and 3A; J. Lee and C. Pabo for NMR advice; T. Havel for software and advice; G. Waneck and M. Kurian for help with mutagenesis; and D. Wiley for use of spectropolarimeter. Supported by NIH grants HD30812 (P.K.D.), GM51558 (M.A.W.), and Reproductive Endocrine Sciences Center P30HD28138 (P.K.D.).