The Molecular Action and Regulation of the Testis-Determining Factors, SRY (Sex-Determining Region on the Y Chromosome) and SOX9 [SRY-Related High-Mobility Group (HMG) Box 9]

VINCENT R. HARLEY, MICHAEL J. CLARKSON, AND ANTHONY ARGENTARO

Prince Henry’s Institute of Medical Research, Clayton 3168, Victoria, Australia; and Medical Research Council Molecular Haematology Unit, Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DS, United Kingdom

Despite 12 yr since the discovery of SRY, little is known at the molecular level about how SRY and the SRY-related protein, SOX9 [SRY-related high-mobility group (HMG) box 9], initiate the program of gene expression required to commit the bipotential embryonic gonad to develop into a testis rather than an ovary. Analysis of SRY and SOX9 clinical mutant proteins and XX mice transgenic for testis-determining genes have provided some insight into their normal functions. SRY and SOX9 contain an HMG domain, a DNA-binding motif. The HMG domain plays a central role, being highly conserved between species and the site of nearly all missense mutations causing XY gonadal dysgenesis. SRY and SOX9 are architectural transcription factors; their HMG domain is capable of directing nuclear import and DNA bending. Whether SRY and SOX9 activate testis-forming genes, repress ovary-forming genes, or both remains speculative until downstream DNA target genes are identified. However, factors that control SRY and SOX9 gene expression have been identified, as have a dozen sex-determining genes, allowing some of the pieces in this molecular genetic puzzle to be connected. Many genes, however, remain unidentified, because in the majority of cases of XY females and in all cases of XX males lacking SRY, the mutated gene is unknown. (Endocrine Reviews 24: 466–487, 2003)
a Y chromosome but are phenotypically female (XY females). Molecular analysis of the genomes of these so-called sex-reversed patients led to the isolation of the sex-determining region on Y gene, SRY (Sry in the mouse). Before this discovery, candidate Y-located genes had been isolated that subsequently failed to meet established criteria (5). SRY met the criteria, as discussed in Section I.A, and its discovery 12 yr ago was envisaged as the entry point through which the molecular genetic basis of mammalian sex determination would be quickly unraveled. However, progress has been slow, hampered by the seemingly intractable nature of the SRY protein. On a positive note, advances have been made on other fronts in the last decade: reverse genetic approaches on human sex-reversal syndromes and mouse gene knockout studies have led to the identification of sex-determining and gonad-formation genes such as those encoding the transcription factors SOX9 [SRY-type high-mobility group (HMG) box 9], DMRT1, GATA4, DAX1, SF1, WT1, LHX9, and cell-signaling molecules AMH, WNT4, FGF9, and DHH. Together with the output of a large number of genes from recent microarray and mouse mutagenesis screens, the genetic and biochemical interactions between these genes and proteins are beginning to reveal the nature of sex determination.

This review focuses on the regulation, function, and molecular interactions of two key mammalian testis-determining factors: SRY and the SRY-related protein, SOX9, allowing us to propose a model of how sex determination is initiated in mammals. Excellent reviews cover other genes in the pathways of two key mammalian testis-determining factors: SRY and the SRY-related protein, SOX9, allowing us to propose a model of how sex determination is initiated in mammals. Excellent reviews cover other genes in the molecular genetic basis of mammalian sex determination (9).

II. SRY, the Testis-Determining Factor

A. The discovery of SRY

SRY was discovered by analysis of the small fragments of the Y chromosome that had translocated to the X chromosome in the genomes of XX males and true hermaphrodites (10). Studies in humans and mice over the next couple of years provided evidence, both circumstantial and direct, which proved that SRY was indeed the long-sought-after TDF gene. In humans, SRY mutations were identified in patients with XY gonadal dysgenesis (11). The protein encoded by SRY showed sequence-specific DNA binding activity, which was absent or reduced in SRY from certain XY females with gonadal dysgenesis (12, 13). In mice Sry was expressed in the right place at the right time, just before the overt differentiation of the bipotential gonad and for a brief period in the somatic cells, to initiate Sertoli cell differentiation (14). The demonstration that XX mice transgenic for Sry develop as males with normal testis and testicular cords indicated that Sry was the only Y-encoded gene necessary for testes formation and the consequent male phenotype (15, 16).

B. Gonadal expression of SRY

Analysis of chimeric mouse gonads constructed by aggre-

gation of male and female embryos showed that almost all Sertoli cell precursors bear a Y chromosome (17), thus suggesting that these cells are the site of Sry expression (18). In the mouse, the adreno-genital primordium arises from the celomic epithelial cells and mesenchymal cells in the mesonephros. The cells of the adreno-primordium are identified as a group of SF1-expressing cells (19). Germ cells migrate to the primordium and cluster with somatic cells to form the genital ridge, first visible as a thickening of the mesonephros around 10 d post coitum (dpc). The timing of Sry expression in the mouse is entirely consistent with a role in sex determination. Sry is detected in the genital ridge at 10.5 dpc, and at 11.5 dpc, but by 12.5 dpc, Sry is detectable only at low levels (14). Sex-specific differences are apparent at about 11.5 dpc when the male gonad takes on a striped appearance, probably due to cells of the supporting cell lineage differentiating into Sertoli cells and aligning into testis cords, the presumptive seminiferous tubules. The ovary forms 2 or 3 d later in female embryos, behaving as a default pathway, initiated in the absence of specification of the male pathway forming granulosa or theca cells. The Sertoli cells and granulosa cells surround the germ cells in their respective organs. Germ cells are not required for testis determination because mice homozygous for the W" (white spotting) mutation lack germ cells, yet form somatically normal testes (20). If germ cells migrate to an ectopic tissue, they go into meiotic arrest, which is the fate of germ cells in the ovary (20).

Sex determination can be defined as the proliferation, migration, and differentiation of supporting cells to become Sertoli cells, committing the fate of the gonad to the testis pathway. Sry expression correlates with proliferation of bipotential supporting cells, some of which are the precursors of Sertoli cells in the XY gonad (21). Sry also induces the migration of cells into the XY gonad from the adjacent mesonephros (22). Sry expression progresses from the anterior to the posterior end of the testis over 2 d, with any one cell expressing Sry for no more than a few hours (23). The transient expression of Sry suggests that in mice, whereas it is responsible for determining Sertoli cell fate, Sry is not involved in maintenance of the Sertoli cell phenotype (6). Sertoli cells act as organizers, causing surrounding cells to differentiate into either Leydig or myoid cells. Additionally, Sertoli cells induce mitotic arrest in germ cells, conferring a spermatogenic fate upon them. Few effector molecules involved in these processes are known. A recent advance, however, is the observation that XY mice lacking the secreted fibroblast growth factor Fgf9 develop as phenotypic females (24); these mice show defects in testis cord formation and organization as a consequence of reduced and abnormal Sertoli cells and/or reduced migration and proliferation of interstitial cells.

In humans, as in mice, the onset of SRY mRNA expression defines testis determination. The human gonadal ridge forms around 33 d gestation, and SRY is detected at 41 d in XY embryos. SRY levels peak at 44 d, when the testis cords are first visible (25). Unlike in the mouse, human SRY is not switched off in the gonads and continues into adulthood. By 52 d gestation the germ cells are surrounded by Sertoli cells, which continue to express SRY at a low level (25), suggesting other roles for SRY, possibly in spermatogenesis. SRY protein has been detected in the nuclei of Sertoli cells in the embry-
ronic testis, consistent with a role as a transcription factor. RNA expression data suggest SRY is widely expressed in adult tissues and cell lines as well as in germ cells and preimplantation blastocysts (26, 27). However, these RT-PCR experiments have not been confirmed by a quantitative technique. In summary, human and mouse SRY are expressed just before the differentiation of the bipotential gonad into a testis.

C. The SRY protein

The open reading frame (ORF) of human SRY is contained within a single exon and encodes a 204-amino-acid protein. The protein can be divided into three regions (Fig. 1). The central 79 amino acids encode the HMG domain, which functions as a DNA-binding and DNA-bending domain and also contains two nuclear localization signals. The human SRY C-terminal domain has no obvious or conserved structure, except for the final seven amino acids, which interact in vitro with a PDZ domain protein (28). The N-terminal region of the protein also has no obvious structure, but phosphorylation of a sequence within this domain enhances DNA binding activity (29). Comparison of the amino acid sequence of the SRY HMG domains from human, mouse, rabbit, wallaby, marsupial mouse, and sheep reveal 70% identity. In contrast, there is no sequence conservation outside the HMG domain. Mouse Sry contains a glutamine-rich region absent from other mammalian species (and some mouse subspecies). The SRY gene has evolved rapidly (30) and the lack of homology outside the HMG domain may reflect the evolution of additional functions for SRY outside of sex determination, e.g., in spermatogenesis or in the brain.

The SRY HMG box is a conserved motif for minor groove DNA recognition. In vitro analysis of recombinant SRY protein suggests that its HMG domain has the ability to recognize the DNA binding motif A/TAACAAT/A with highest affinity (31), enhanced through phosphorylation (29). Sequence-specific DNA binding appears to be a property of the entire class of SOX (SRY type HMG BOX) proteins, which bind the core consensus sequence AACAAT. The affinity of SRY for double-stranded DNA varies with DNA sequence. SRY has been demonstrated to bind quite strongly to a number of sequences in vitro, with dissociation constants of 20 nM for human SRY box, GAAAACAG (32), and 3 nM for mouse Sry HMG box, AACAATG. How is binding specificity achieved, if the binding site is so short? In vitro selection assays of different SOX proteins suggest that the DNA footprint is longer, with different SOX proteins demonstrating different preferences for the two nucleotides flanking the AACAAT core motif (33).

Because the in vivo DNA targets of many SOX proteins are known, it is likely that SRY also binds to DNA in vivo to regulate gene expression. It is intriguing that an RNA binding activity for SOX6 and SRY has been reported (34).

The nuclear magnetic resonance structure of the HMG domain of SRY in complex with DNA has been determined (35) (Fig. 2). It consists of three α-helices forming an L-shape. Helix 1 and 2 run antiparallel and form the short arm. The longer helix 3 runs antiparallel to the amino-terminal strand and forms the long arm at a right angle to the short arm. Conserved aromatic amino acids from each helix pack together to form the hydrophobic core of the protein, stabilizing the structure. The DNA is severely unwound with a widened shallow minor groove, and the bound SRY causes the DNA to bend about 80 degrees, which is in agreement with biochemical studies. A number of amino acids make DNA contacts to specific bases (Asn65, Ile68, Ser88, Ile90, Gln117, Lys129, Tyr129), the sugar-phosphate backbone (Arg59, Arg62, Met64, Arg72, Arg75, Lys 92, Trp98, Arg121) or both (Phe67).

DNA bending is associated with the insertion of a non-polar cantilever side chain of Ile68 between the base pairs. Insertion disrupts base stacking but not base pairing. Substitution of Ile68 for a Thr, which occurs in a patient with XY gonadal dysgenesis, results in a 50-fold reduction in DNA binding. The precise function of protein-directed DNA bending is uncertain, but it is proposed to establish a particular DNA architecture in chromatin. Indeed, studies with chromatin have shown that SRY is capable of displacing a nucleosome at the fra-2 promoter (36). The angle to which DNA is bent by SRY in vitro depends upon the flanking residues

![Image](https://example.com/image.png)

**Fig. 2.** A model of the structure of the SRY HMG box bound to DNA. The nuclear magnetic resonance structure of the SRY HMG domain, showing the three α-helices and the L-shaped conformation and binding to the minor groove of DNA causing it to bend and severely unwind (35).
of the DNA sequence used (GAACAATG, 85° vs. CAACAATG, 60°). The degree of bending induced also differs between species (37). SRY and, indeed, all HMG1-type domains are also able to bind unusual structures such as DNA cruciforms (38, 39). The affinity of SRY HMG box for cruciform DNA is comparable with that for linear DNA and is thought to be facilitated by the angles formed between the arms of the four-way junction, the displacement of the central DNA base pairs, and the increased width of the minor groove 5' to the site of strand exchange (39, 40). Consequently, whereas the Ile68Thr SRY protein variant binds DNA less than 50-fold of wild-type affinity, it binds cruciform DNA with normal affinity. The cantilever residue I68 is not required for the displacement of the central base pairs of the cruciform DNA site, which leads to the partial unwinding of the helix and DNA bending (35, 40).

D. Clinical mutations in SRY

Mutations in SRY are associated with human male-to-female sex reversal in XY females. DNA sequencing has demonstrated that SRY is found in most XX males, and that SRY point mutations or deletions are found in only about 15% of XY females (41). This supports SRY being TDF and also suggests the existence of mutations in sex-determining genes other than SRY. Histologically, the gonads of XY females can be classified into two groups: 1) pure or complete gonadal dysgenesis, in which patients have dysgenic or male sex reversal in XY females. DNA sequencing has provided insight into the normal functions of SRY and, indeed, almost all XY female patients with SRY mutations show complete gonadal dysgenesis, consistent with a critical role for SRY early in testis formation (41). These patients develop as normal females with female internal and external genitalia and, due to the complete gonadal dysgenesis and lack of ovarian function, these XY females present at clinics with primary amenorrhea. In about 50% of cases gonadal tumors (gonadoblastoma or dysgerminoma) are associated with the gonadal dysgenesis, with the result that gonads are routinely removed by surgery.

The positions of SRY point mutations causing XY sex reversal are shown in Fig. 3. Although nonsense mutations are scattered throughout the SRY gene, missense mutations tend to cluster in the central region of the gene, which encodes the HMG domain. This strongly suggests that the DNA binding motif is essential in vivo. However, not all SRY mutations cause complete sex reversal. There are cases of familial SRY mutations in which the fertile father of an XY sex-reversed individual also carries the SRY mutation. In such cases, it is thought that genetic background can compensate for the mutation and that this is not occurring in the affected individual.

E. DNA binding activity of SRY from an XY female

Patients with XY gonadal dysgenesis carry point mutations, which substitute a single amino acid in their SRY ORF. Analysis of the biochemical activity of these mutant SRY proteins has provided insight into the normal functions of SRY. A number of mutations have been characterized and, in some cases, DNA binding or DNA bending activity are reduced (13, 32, 37, 42–46). Some mutants show reduced or abolished DNA binding activity but normal DNA bending (V60L, R62G, I68T, R75N, L94P, G95R, K106I, Y127C); some show moderate affinity (M64I, 190M, S91G, P125L, A113T); and some show wild-type affinity (S18N, M64R, M64T, F67V, F109S, R133W). The latter groups are presumably defective for an unknown biochemical activity of SRY. Some SRY mutants showed abnormal DNA bending and reduced DNA binding (R62G, M78T). In a third case (M64I), DNA bending was affected but DNA binding was largely unaffected; a second mutation at the same position M64R abolished DNA bending and reduced binding only about 4-fold, suggesting that the SRY determinants of DNA binding and DNA bending activity are partly independent. Taken together, these studies suggest that the DNA binding and DNA bending activity of SRY is essential for testis determination.

It is difficult to draw correlations between the activities of the mutated SRY proteins in vitro and the clinical severity of the mutation in the patient. A spectrum of effects are observed in vitro, from a small reduction in DNA binding or bending activity, to a total loss of DNA binding activity. Most have the same outcome in vivo: complete gonadal dysgenesis. The fact that relatively small reductions in DNA binding activity are sufficient supports the idea that SRY is expressed just above a threshold activity level.

Although most mutations are de novo, first arising in the affected individual and causing complete gonadal dysgenesis, there are a few mutations with a familial inheritance (90M, F109S) that show near-wild-type DNA binding activities. Surprisingly, a number of XY females carry mutations, which do not affect their DNA binding or DNA bending activities in vitro; this suggests that other essential activities of SRY must exist. Also, the results confirm our earlier suggestion that SRY functions at a biochemical threshold and that familial mutations are close to this threshold level and manifest in certain genetic backgrounds (13).

We have recently shown that one of these mutants, R133W, is defective in its interaction with importin β, a protein that SRY requires for nuclear import (46a). It is quite likely that other
interacting proteins remain to be discovered and that some of these SRY mutants will be abnormal in their interaction with them. Alternatively, the in vivo DNA site bound by SRY could vary substantially from the consensus SRY/SOX site used in in vitro binding assays.

Two individuals carry missense mutations outside the HMG domain in the N-terminal and C-terminal domains; both have unusual phenotypes. The 4-yr-old boy who was affected by the N-terminal domain mutation S18N had partial gonadal dysgenesis, had retained both Wolffian and Mullerian ducts, and had ambiguous external genitalia. However, both his father and adult brother carry the mutation and appear to be normal males (47). A patient carrying the K43 stop mutation failed to develop a uterus (48). In vitro DNA binding and DNA bending activity of the S18N mutant was normal (46). In each case, SRY function is affected in an unusual manner, suggesting that the domains flanking the HMG domain influence SRY function in vivo.

F. SRY: activator or repressor?

Evidence that human SRY is a transcriptional activator or repressor is equivocal with reporter gene studies in transfected cell cultures showing activation or repression. The main limitations of these studies has been the lack of a relevant DNA target and/or appropriate cell lines. One study suggested the AMH gene (also known as MIS) as a potential target of SRY (49). SRY was transfected into a rat gonadal ridge cell line with a reporter construct containing a fragment of the AMH promoter. Activation of the reporter construct was approximately 15-fold higher than basal levels, and the response was reduced when a variant of SRY with a 50-fold reduced DNA binding affinity was used. However, these effects were not due to a direct interaction between SRY and the AMH promoter because the removal of the single SRY binding site in the AMH promoter did not affect the levels of SRY-mediated activation of AMH. In a second study, SRY was shown to activate the fra-1 promoter (fos-related antigen) in Chinese hamster ovary cells about 20-fold higher than basal levels (50), suggesting that SRY has the capacity to activate gene expression in cells, although a SOX protein other than SRY is likely to be involved in fra-1 regulation because fra-1 is not expressed in the genital ridge (51).

On the other hand, it has been suggested that SRY represses the activity of a gene that would otherwise repress genes required for the development of the male sex (52). This model could explain why XX males with SRY rarely have ambiguous genitalia, whereas XX males lacking SRY commonly have ambiguous genitalia or are true hermaphrodites. It is argued that there would be varying degrees of repression of male sexual development genes in the absence of SRY, depending on which of these unidentified genes contained mutations. In support of this theory, there has been a reporter gene study in transfected cells demonstrating that SRY can repress transcription. A construct containing multiple SRY binding sites upstream of a reporter gene (thymidine kinase promoter) was transfected into monkey epithelial (COS7) cells with an SRY expression vector (29). The 2-fold level of repression was increased with phosphorylation, which enhances DNA binding activity, and no repression was detected when the phosphorylated residue of SRY was mutated. However, in a similar study in the same cell line but using a different promoter (adenovirus E1b promoter) a 3-fold activation was observed (P. Tang and V. R. Harley, unpublished observation). Thus, it is becoming increasingly apparent that until the in vivo target of SRY has been identified, we will not know whether SRY is an activator or repressor—or indeed both!

In contrast to studies of human SRY, studies on mouse Sry show that it strongly activates transcription through its C-terminal glutamine-rich region (53), which is completely absent in SRY from other mammals and certain mouse sub-species (Fig. 1). It is hard to envisage a different mechanism operating through the poorly conserved non-HMG domain regions in different mammalian species. More likely, the SRY HMG domain itself carries the necessary information for transcription through structure- and sequence-specific DNA recognition together with coactivator proteins to establish the correct architecture in chromatin, analogous to the HMG domain protein LEF-1 (54). Non-HMG domain regions may subserve the HMG domain function.

G. The SRY HMG box: all you need?

Arguing against the suggestion that the HMG box is all you need, the C-terminal domain of mouse Sry can act as a transcription activator in cultured cells (53). A translation stop engineered just before the glutamine region prevented the Sry transgene from causing female-to-male sex reversal in XX mice (55), but a negative result in transgenic experiments must be viewed with caution. In humans, the SRY K43 stop mutation in a human XY female might represent a similar situation, where the C-terminal domain had been deleted, resulting in loss of function.

In favor of the suggestion that the HMG box is all you need, when the human and mouse ORFs are compared, there is no homology outside the HMG box (Fig. 1). Of the 32 known SRY missense mutations, only two are located outside the HMG box; more would be expected in these domains, which represent 120 amino acids of the 204-amino-acid protein. Recent experiments indicate that the human SRY ORF can cause sex reversal of XX transgenic mice (56). This suggests that there is no requirement for the C-terminal glutamine-rich domain of mouse Sry for sex determination, although these regions may play a role in processes other than sex determination, e.g., spermatogenesis or regulating male behavior. These studies suggest that “a” C-terminal domain is required for Sry/SRY function, if only for protein stability. Supporting this suggestion, the HMG domain is never located at the C-terminal end of an SRY protein among mammalian species.

Clearly, the HMG domain is essential, but can any HMG domain do the job? In vitro studies suggested that SRY/SOX proteins differ in their intrinsic DNA sequence specificity (33), probably reflecting a difference in DNA binding specificity in vivo. Although SRY and SOX9 are capable of binding the same target sequences in vitro, their affinities are different (33). In vivo support for this observation came from work by Eicher and colleagues, which showed that only when over-expressed could SOX9 or SOX3 HMG domains substitute for
that of SRY to cause sex reversal in XX transgenic mice (57). It was noted that greater amounts of transgene were required, presumably because the protein dose needed to be greatly to override the weaker affinities of the SOX3 and SOX9 HMG domains for the target of SRY, currently unknown. However, swapping does not always work. When the HMG domain of SOX9 was swapped with that of SOX1, it greatly affected the transactivation potential of SOX9 (58), confirming that SOX9 HMG domain appears to possess some DNA target specificity not present in other SOX proteins.

H. In vivo targets of SRY

The essential nature of the SRY HMG box and the numerous examples of SOX proteins binding in vivo to DNA targets suggest a likely role for SRY as a DNA binding protein to a promoter/enhancer/silencer element within a downstream gene in Sertoli cell precursors in vivo. The early report of SRY binding to the AMH promoter is unlikely given that their expression profiles in the genital ridge are nonoverlapping (59). However, a number of genes are expressed soon after the induction of Sry, e.g., Sox9, Fgf9, Dhh, and Vnn1. Furthermore, a number of recent microarray screens hold much promise in revealing more candidate genes (60–62). Clearly, a direct demonstration of SRY binding to a candidate gene DNA in vivo will rely upon difficult procedures such as chromatin immunoprecipitation. Only then will the issue of whether SRY acts as a repressor or activator be more easily resolved.

It is likely that whatever SRY directly acts upon, the ultimate function of SRY is the up-regulation of Sox9. Shortly after the induction of Sry, Sox9 becomes activated, moving from the cytoplasm to the nucleus, and up-regulated in the male gonad (63). To address the precise function of Sox9 in the gonad, transgenic mice ectopically expressing Sox9 driven by the Wil promoter were produced (64). XX transgenic mice developed testes with apparently normal Sertoli cells and Leydig cells. This suggests that Sox9 can replace Sry and implies that Sry’s only function is to up-regulate Sox9.

Up-regulation of Sox9 is a key phenomena in all vertebrates, regardless of the switch mechanism controlling sex determination, i.e., SRY in mammals [except for the mole vole (65)], ZW chromosome gene/s in birds (66), and temperature sensitivity of egg incubation in turtles and crocodiles (67, 68).

Support for SRY being a repressor of a repressor comes from the creation of the Odsex (ocular degeneration with sex reversal) mouse. Sex reversal is caused by the insertion site of a transgene 1–2 megabases (Mb) upstream of Sox9, causing a 150-kb deletion. XX mice transgenic for Odsex develop as males with levels of Sox9 transcripts identical with those found in wild-type XY males at 11.5 and 14.5 dpc. The insertion/deletion did not affect the expression of Sox9 in cartilage or bone, which suggests that the transgene insertion deleted a regulatory element that is only required for the repression of Sox9 in the gonads. It was proposed that SRY normally disrupts the binding or function of this repressor (69). A candidate for the role of repressor of Sox9 is Dax1.

I. SRY interacting proteins

Two proteins implicated in nuclear import function have been identified as interacting with SRY. Nuclear localization signals (NLSs) are responsible for targeting proteins for recognition by cytoplasmic proteins, which facilitate translocation of proteins generally larger than 45 kDa through the nuclear pore complex. SRY protein appears to be localized in the nucleus of pre-Sertoli cells consistent with a transcriptional function (25). SRY, although only 27 kDa, is exclusively nuclear and therefore actively transported. SRY contains a bipartite NLS (61 KRpmnafivwsRdqRRK) at the N-terminal end and an simian virus 40-type NLS (126KyRppKKK) at the C-terminal end of the HMG box (Fig. 4). The N-terminal NLS forms part of a calmodulin (CaM)-binding domain, and a role for CaM in nuclear import of other proteins (72) suggested that nuclear import by this NLS is mediated by CaM. Recent studies using Sox9 and CaM antagonists demonstrate this to be the case (72a). Binding studies of NLSs with importins reveal that the C-terminal NLS of CaM and importin β-binding domains of SRY HMG domain. An alignment of the HMG box of SOX proteins (representative from each group), the HMG box of lymphoid enhancer factor 1 (LEF1), and the A box of HMG1. The region in gray represents the CaM-binding domain mapped in SRY. The N- and C-terminal NLS sites are underlined. Note lack of conservation of CaM and importin β-binding domains between SOX and non-SOX proteins. [Reproduced with permission from A. Argentaro et al. : J Biol Chem Epub ahead of print M302078200, 16 June 2003, (72a).]
SRY interacts with importin β. The SRY mutations in the C-terminal NLS, R133W, showed reduced nuclear import due to reduced binding to importin β (46a) (Fig. 5). This provides a molecular explanation for sex reversal in this variant, which also showed wild-type DNA binding and DNA bending activity.

A yeast two-hybrid assay revealed an interaction between SRY and SRY interacting protein, SIP1, via a PDZ-type domain with the seven C-terminal residues of SRY. SRY contains two PDZ domains, each of which can interact with SIP1 and with the same affinity, although not simultaneously, suggesting that SIP1 could act as a bridge between SRY and another protein with a PDZ recognition motif (28). It remains to be demonstrated whether these in vitro binding partners are functionally relevant. Similarly, it has been speculated that the C-terminal domain of mouse Sry, which lacks a PDZ recognition motif but consists of an unusual glutamine-rich repeat region, might also act as a bridge between Sry and other molecules in a complex (55).

III. SOX9, the Campomelic Dysplasia/Autosomal Sex-Reversal Gene

A. The discovery of SOX9

Apart from mutations in the SRY gene that cause XY gonadal dysgenesis, several autosomal and one X-linked loci have also been associated with the inability to develop a testis. An autosomal locus, termed sex reversal autosomal 1 on chromosome 17, in addition to causing sex reversal also causes a severe skeletal malformation syndrome known as campomelic dysplasia (CD). CD is a dominant lethal disorder characterized by congenital bowing of the long bones, narrow ilia, cleft palate, absence of olfactory bulbs and tracts, heart and renal malformations, hypoplastic lungs, narrow thoracic cage, and delayed bone age (73, 74). Screening of patients with CD identified up to 75% of XY individuals with testicular dysgenesis leading to sex reversal. CD patients generally die from respiratory defects. The analysis of five patients with chromosomal rearrangements associated with CD localized the gene to the long arm of chromosome 17 (75, 76). High-resolution mapping refined the location to 17q24.3–25.1. An SRY-related gene, termed SOX9, was located adjacent to the breakpoint in one patient. Mutational analysis and DNA sequencing confirmed that mutations in SOX9, a gene that shares homology to SRY through its HMG box, causes CD and sex reversal.

B. Clinical mutations in SOX9

SOX9 mutations occur throughout the ORF, unlike those in SRY that generally cluster within the HMG box. Mutations in the SOX9 ORF detected outside the HMG box are nonsense and frameshift mutations that disrupt the C-terminal domain of the protein and alter the ability of SOX9 to efficiently activate transcription of target genes (77–79). Mutations in SOX9 include splice acceptor/donor changes and missense, nonsense, translocation, and frameshift mutations (Figs. 6 and 7) (75, 76, 79). There is no correlation between severity of the disease or associated sex reversal and mutation type with many of the same mutations causing varying degrees of gonadal dysgenesis and bone malformations (81). For example, a mutation that causes a frameshift and subsequent premature stop at codon 254 was identified in three siblings. One sibling was a 46,XY true hermaphrodite with ovotestes and ambiguous genitalia; another was 46,XY with bilateral ovarian tissue and normal female genitalia; and the third sibling was a 46,XX with normal ovaries and female genitalia (83). In another case, two unrelated individuals carried a mutation that resulted in a premature stop (Y440X) (Fig. 7). One of these individuals exhibited testicular dysgenesis with sex reversal, whereas the other had unambiguous male genitalia (81, 82). There is no correlation between degree of expressivity of the SOX9 mutation, as well as differences in genetic background, rather than the specific type of mutation. Although most mutations occur in one
allele of SOX9, compound heterozygosity has also been detected in one patient with CD with a different mutation in each allele of SOX9 (76). Recently, an SRY-negative female-to-male sex reversal patient with a duplication of chromosome band 17q23–24 including the SOX9 gene was isolated (84). Thus, an extra dose of SOX9 may be sufficient to initiate testis differentiation in the absence of SRY.

No cases of sex reversal without CD, due to mutations in SOX9, have been detected (85).

Nine translocation breakpoints or inversions that leave the SOX9 ORF intact have been reported (86, 87). These breakpoints map at a distance of 50–950 kb upstream of the gene, and one breakpoint is as far as 850–950 kb proximal (86, 87). Generally, CD patients with translocation and inversion mutations are less severely affected and usually have a longer life expectancy than gene mutations located in the ORF (87).

The cause of CD in these cases could be due to the removal of cis-regulatory elements or the disruption of transcription or silencer elements (87). The 1.2 Mb of DNA upstream of the SOX9 gene contains no genes, suggesting that inversions and/or translocation most likely remove cis-regulatory elements (86).

C. SOX9 protein structure and activity

The human SOX9 gene encodes a protein of 509 amino acids consisting of the HMG box, which shares 70% amino acid homology to the HMG box of SRY. In addition, the SOX9 protein contains additional protein domains, including two transcriptional activation domains, downstream of the HMG box. The SOX9 protein, unlike SRY, is very highly conserved through vertebrate evolution (Fig. 8).

1. DNA binding and bending. The SOX9 HMG domain, like that of SRY, acts by binding to (and bending) specific DNA sites to activate transcription of target genes. A large majority of mutations in SOX9 are located within the HMG box, suggesting that its ability to efficiently bind to target sites is a critical factor. The HMG box of SOX9, like all SOX proteins tested, binds to the SRY high-affinity binding site, AACAAT, and the TCF1 site, AACAAAG (77, 78). The high-affinity binding site for human SOX9 selected in vitro is AGAA-CAATGG (33). SOX9 shows a higher DNA binding affinity for residues flanking the core SRY high-affinity binding site compared with SOX5, SOX17, and SRY, suggesting that SOX9 may contain different DNA binding specificity compared with other SOX proteins (33). The SOX9 consensus site shares some homology with the Col2a2 sequence, TCGATGAAATGG, located in the type II collagen enhancer and to the SOX sequence, GTCTCAAAGG, located adjacent to the Mullerian inhibiting substance response element (MISRE)1 site in the AMH promoter, both of which are known targets of SOX9.

Fig. 6. Mutations in SOX9. A, Diagram of the entire SOX9 ORF showing the different domains of the protein and the mutations that have been identified in patients with CD. Mutations identified in patients with sex reversal are indicated below and those in XX females (F) or XY males (M) are indicated above. Underlined are the transactivation domains identified (77, 79). B, Diagram of the SOX9 HMG box, showing amino acid substitution mutations that have been identified in patients with CD. Those identified in sex-reversed patients are shown above. SA, Splice acceptor mutations; fs, frameshift; X, stop codon. [Reproduced with permission from P. Bernard et al. (188). © Oxford University Press.]

Fig. 7. Mutations in SOX9 that truncate the C terminus. Those highlighted in black are similar to deletion mutants investigated to define the transactivation domain of human SOX9 (B). SOX9 mutant proteins encoded by CD patients are represented as a rectangle of length corresponding to their ORFs produced as a consequence of mutations shown in panel A. An asterisk indicates SOX9 mutant proteins identified in a CD patient with compound heterozygosity (individual with two different mutant alleles at SOX9 locus). F, Female; M, male; fs, frameshift; X, stop codon. [Reproduced with permission from P. Bernard et al. (188). © Oxford University Press.]
In addition to DNA binding specificity, interaction with specific partner proteins may also confer additional specificity. This is illustrated by the experiments performed by Kamachi et al. (58), who investigated the mechanisms of target selection by comparing the group B SOX proteins, SOX1, SOX2, and SOX3, which activate the 6-crystallin minimal enhancer, DC5, with SOX9, which activates the type II collagen minimal enhancer, Col2a1. SOX proteins have been classified according to the protein sequence homology of their HMG domains (90). The replacement of the HMG box of SOX1 with that of SOX9 did not alter transcriptional activity of the DC5 enhancer in lens cells (58). However, the reciprocal experiment resulted in a reduction in the ability of SOX9 protein (containing the SOX1 HMG box) to activate the type II collagen minimal enhancer, Col2a1 (58). Hence, the SOX9 HMG box confers some degree of specificity that may relate to interactions with specific partner proteins. The presence of promoter and/or enhancer sites with varying SOX9 DNA binding affinities may also contribute to differential regulation of target genes.

All known missense mutations in SOX9 occur in the HMG box (75, 76, 79–83). The majority of SOX9 missense mutations show altered DNA binding compared with wild type, although there are a few exceptions (79, 81). The A119V mutation showed near-wild-type DNA binding and bending; and another mutation, P170R, while showing near-wild-type DNA binding and bending, had altered DNA binding specificity (79). This suggests that other essential biochemical activities, in addition to DNA binding and bending, are impaired in these SOX9 mutations. In addition, no missense mutations identified have shown normal binding with altered DNA bending.

2. SOX9 transcriptional activity. Several studies have shown that the SOX9 protein acts as a potent transcriptional activator both in vitro and in vivo. The SOX9 protein is known to activate transcription of the type II collagen gene and anti-Mullerian hormone (91). The removal of the last 108 amino acids from the C terminus of SOX9, which is rich in proline, glutamine, and serine residues, abolished transcriptional activity (77–79). The region adjacent to the proline-glutamine-serine (PQS) domain, which consists entirely of prolines, glutamines, and alanine residues (PQA), was also found to be required for maximal transactivation; unlike the PQS domain, the PQA domain varies in length between different species and is completely absent in rainbow trout Sox9 (92) (Fig. 8). Species that lack the PQA domain would be predicted to have a weaker transactivation potential. Interestingly, this domain is only present in vertebrates and may relate in some way to organisms with an SRY sex-determining mechanism.

Some SOX9 mutations that encode proteins with C-terminal truncations showed decreased transactivation capabilities compared with wild-type SOX9. Two truncation mutations, Y440X and the 507fs frameshift mutation, while leaving the majority of the transactivation domain intact, reduced transactivation activity (81). The 507fs mutation encodes a mutant SOX9 protein that has an extended ORF creating an unstable mRNA and/or protein product (81). The transactivation potential of other SOX9 mutations that truncate the

![Fig. 8. Amino acid comparisons of SOX9 from different vertebrates.](#)
C-terminal domain or those that cause an alteration in DNA binding affinity have not been investigated but are predicted to have reduced transcriptional activity compared with wild-type SOX9. Hence, the biochemical defect in these mutants is likely to be an inability to adequately activate transcription.

Apart from sharing a high level of homology between orthologs, the C-terminal domain of SOX9 also shares homology with other Group E SOX proteins, in particular, the transactivation domain of SOX10 (Fig. 9). The C-terminal 23 amino acids are highly conserved in Sox8, SOX9, and SOX10 although the PQA domain is not present in either Sox8 or SOX10 (Fig. 9). It is not clear how the C-terminal domain of SOX9 may confer transcriptional activity; it is likely that both the PQA and PQS domains interact with as-yet-uncharacterized partner proteins to activate transcription. These proteins may include tissue-specific transcription factors and/or components of the basal transcription machinery.

3. Phosphorylation of SOX9. A major posttranslational modification of proteins arises by phosphorylation. The SOX9 protein contains a number of putative phosphorylation sites including seven sites for casein kinase II, five protein kinase C sites, two protein kinase A (PKA) sites, and a tyrosine kinase site (Fig. 8). These phosphorylation sites, as in SRY, may be involved in regulating SOX9 DNA binding, nuclear import, and transcriptional activity.

SOX9 was shown to interact specifically with the catalytic subunit of PKA (95). This interaction between SOX9 and PKA suggests that phosphorylation directly regulates SOX9 activity. SOX9 contains two putative PKA sites, flanking the HMG box, which are completely conserved in SOX9 (Fig. 4). PKA phosphorylation of SOX9 increased DNA binding activity and enhanced the activity of SOX9-dependent Col2al enhancer activity in RCS cells (95). In addition, PKA-phosphorylated Sox9 was detected in the prehypertrophic zone of the growth plate in chondrocytes (95). This suggests that phosphorylation of SOX9 by PKA in prehypertrophic chondrocytes is regulated by parathyroid hormone-related peptide and BMP-2 signaling. BMP-2 increased the expression of Sox9 in chondrocytes (97). Fibroblast growth factors (FGFs) enhance Sox9 expression in C3H10T1/2 cells (98). FGF increased the activity of Sox9-dependent Col2al activity, consistent with the increase in Sox9 expression in primary chondrocyte cells (98). This increase is mediated by the MAPK [MAP kinase (MEK)-MAPK] pathway because the addition of specific inhibitors of MEK-MAPK to primary chondrocytes inhibited FGF-induced expression of Sox9 (98).

The regulation of Sox9 activity in chondrocytes is complex, with phosphorylation playing a major role. A number of signaling pathways act during chondrogenesis to regulate the expression, DNA binding, and transactivation activities of Sox9. Although these studies demonstrate that cAMP-mediated pathways regulate the activity of Sox9 in chondrocytes, the identification of several casein kinase II and protein kinase C sites in SOX9 suggest the existence of other phosphorylation pathways that also regulate the activity of SOX9 in vivo. In particular, the role of phosphorylation in the Sertoli cell lineage, including the regulation of anti-Mullerian hormone (AMH) expression, has not been investigated.

4. Nuclear transport of SOX9. In order for transcription factors to act, they must be transported to the nuclei of cells to bind to and activate (or repress) transcription of target genes. In the developing mouse gonad, in pre-Sertoli cells, Sox9 transcripts are very low and perinuclear. Upon expression of Sry at 10 dpc, the level of Sox9 is unregulated and the protein is localized in the nuclei of cells (63). The temporal and spatial expression of Sox9 and SRY is critical for proper gonadal development.

The NLSs in the C and N termini of the HMG box of SRY are highly conserved throughout SOX proteins, and thus the mechanism of nuclear import is likely to be conserved (Fig. 4). The SOX9 HMG box, similar to SRY, was found to interact directly with the transport receptor importin β and, although shown to induce chondrocyte differentiation and activate PKA in cultured chondrocytes (96). BMP-2 increased the expression levels of Sox9 and a number of other chondrocyte markers, including aggregan and cartilage oligomeric matrix protein, in C3H10T1/2, a mesenchymal progenitor cell line. Hence, both parathyroid hormone-related peptide and BMP-2 signaling may cause phosphorylation of Sox9 by PKA in chondrocyte cells (97). Fibroblast growth factors (FGFs) enhance Sox9 expression in C3H10T1/2 cells (98). FGF increased the activity of Sox9-dependent Col2al activity, consistent with the increase in Sox9 expression in primary chondrocyte cells (98). This increase is mediated by the MAPK [MAP kinase (MEK)-MAPK] pathway because the addition of specific inhibitors of MEK-MAPK to primary chondrocytes inhibited FGF-induced expression of Sox9 (98).

Fig. 9. Amino acid comparison of the C-terminal domain of human SOX9, SOX10, and SOX8. Highly conserved regions are shaded in gray with percent amino acid identity indicated (homology in parentheses). Amino acid identity between all three SOX proteins is indicated with asterisks, conserved residues by dots.
not tested individually, importin β binding is likely to occur through the carboxy-terminal nuclear localization signal (C-NLS) (99). Unlike SRY, no clinical mutations have been identified in the NLSs of the SOX9 HMG box; however, one mutant (A158T), which lies close to the C-NLS, was found to have altered nuclear accumulation and may disrupt the function of the C-NLS (99). The fact that A158T bound with wild-type affinity to importin β suggests that, whereas this recognition step is normal, other components of the importin β-mediated pathway could be affected. The demonstration that a mutation that lies outside the NLS regions affects nuclear localization raises the possibility that a large number of clinical mutations in SOX9 and SRY could affect nuclear import, in addition to or distinct from DNA binding and bending.

Other mechanisms of nuclear import have been described, including a well characterized pathway involving CaM (72). CaM, an intracellular calcium receptor, plays roles in numerous processes in the cell including DNA replication, mitosis, DNA repair, and other nuclear functions such as the regulation of the condensation/relaxation of chromatin (100). Like the SRY protein, SOX9 also interacts with CaM, in a calcium-dependent manner, through a basic region that is well characterized (100, 101). The SOX9 protein, SOX9 also interacts with CaM, in a calcium-dependent manner, through a basic region that acts as an NLS and is highly conserved among SOX proteins (102) (Fig. 4). A reduction in SOX9 in vitro CaM binding was observed on native gels upon addition of CaM antagonists (72a). In addition, significant reductions in transcriptional activity and nuclear accumulation of SOX9 were observed upon treatment with specific CaM antagonists (72a). This suggests that CaM is involved in the nuclear transport of SOX9 (and all SOX proteins) in a process likely to involve a direct interaction (72a). It is likely that a highly specific mechanism acts to regulate SOX9 nuclear transport because the subcellular localization of SOX9 changes from cytoplasmic to nuclear upon onset of SRY expression in the developing mouse XY gonad (63). Components of both the importin/Ran-GTP and CaM-dependent pathways may act to regulate nuclear accumulation of SOX9 in Sertoli cells to allow a sufficient amount of SOX9 protein to activate the AMH gene (Fig. 5).

Although a number of groups have investigated nuclear import of SOX proteins, little is known about SOX nuclear export. Apart from the NLSs with the HMG box, SOX9 (as well as a number of other SOX proteins) also contains leucine-rich motifs that are putative nuclear export signals (103). In relation to nuclear import, clinical mutations located with the HMG box, in particular those located within the NLS and/or CaM binding domain, may reduce the efficiency of SOX9 protein being transported to the nucleus, which is critical for the proper regulation of the gonadal and chondrocyte pathways.

D. SOX9 function in gonad and bone formation

1. Role of SOX9 in gonadogenesis. The expression profile of SOX9 was shown to be sexually dimorphic in a number of species including humans, mouse, chicken, turtle, and alligator (63, 68, 104, 105). Specifically, mouse Sox9 transcripts and protein were detected within Sertoli cells and not within cells of any other lineage, both during development and in the adult testis (63). Human SOX9 shows a similar specificity of expression in the developing testis (106). Ectopic expression of SOX9 in undifferentiated gonads resulted in all XX transgenic animals being phenotypically normal with normal male reproductive ducts and functional Sertoli cells (64). This indicates that SOX9 plays an important role in the differentiation and function of Sertoli cells. Furthermore, a duplication of the genomic region containing the SOX9 gene was identified in a patient with female-to-male sex reversal (84). The constitutive expression of SOX9 appears to be sufficient to promote testis determination. The exact role of SOX9 in Sertoli cell differentiation, however, is not known.

One gene in the Sertoli cell lineage, the expression of which SOX9 was shown to directly regulate, is AMH. AMH is an essential component of the male sexual differentiation pathway, secreted by Sertoli cells, causing the regression of the female (Mullerian) reproductive tract (107, 108). Deoxynucleosilase 1 footprint analysis showed that SOX9 binds to a SOX-like site in the AMH proximal promoter, adjacent to the MISRE1 site and, together with SF1 (steroidogenic factor 1), WT1 (Wilms tumor gene 1), and GATA4, directly activates expression of AMH (108–110). Heterozygote mutations in the SF1 gene have been associated with complete XY sex reversal and adrenal failure in humans, suggesting that SF1 regulates the regression of the Mullerian ducts (111, 112). Several key experiments have shown that regulation of the AMH gene requires cooperative interaction between SOX9 and SF1. SF1 binds to the MISRE1 site, and reporter gene assay experiments have shown a cooperative increase in transcriptional activity with the cotransfection of SF1 and SOX9 compared with SF1 alone (108). A comprehensive study of Amh promoter mutations in transgenic mice (109) showed that mice with targeted mutations in the SOX site within the Amh promoter resulted in complete retention of Mullerian duct-derived organs with a complete absence of Amh transcript (109). This indicates that Sox9 is essential for Amh transcription. The complete absence of Amh transcripts in XY mice homozygous for the mutant SOX binding sites also suggests that Sox9 is required to initiate Amh transcription (109). In the case of SF1, mutations in the MISRE1 site caused partial decrease in Amh levels and partial regression of the Mullerian ducts, indicating that SF1 is required for the up-regulation of Amh transcription (109). GATA4, which is abundantly expressed in Sertoli cells, interacts through its zinc finger domain with SF1 to synergistically activate AMH transcription (110). Similarly WT1, also required for male gonadal development, interacts with SF1 to up-regulate expression of Amh transcription (113).

Recently, SOX9 has been shown to interact with HSP70 (heat shock protein 70) in testicular and chondrocyte cell lines (114). The interaction involves the C-terminal domain of HSP70 with a 100-amino-acid region of SOX9 between the HMG box and the PQA domain, hitherto of unknown function. HSP70 plays numerous roles in transcriptional regulation and interacts strongly with WT1 in vivo (115). Although binding sites to both SOX9 and SF1 are conserved within the AMH promoter, WT1 has no conserved binding site and it interacts with SF1 only weakly. Considering that WT1 strongly interacts with HSP70 in vivo (115), it is possible to speculate that WT1 binding at the AMH promoter is stabi-
lized by the formation of a SOX9-HSP70-WT1 protein complex.

Clearly, numerous protein-protein interactions exist at the Amh promoter. The ability of the SOX9 HMG box to bend DNA may bring SF1 and GATA4 in closer proximity to each other and along with WT1 and HSP70 form a tightly associated protein complex that activates transcription of the AMH gene (Fig. 10).

The observations that SOX9 mutations in CD patients block Sertoli cell differentiation yet Amh-deficient mice exhibit normal Sertoli cell differentiation suggest that SOX9 must play key roles in the testis additional to Amh

At present, it is not known what factor(s) initiate the Sox9 promoter in Sertoli cells. The up-regulation of Sox9 in Sertoli cells upon expression of Sry is consistent, however, with a role for Sry in the activation of Sox9 expression (63, 106). A minimal interval located between −193 and −73 bp from the transcription start site of Sox9 was sufficient to direct maximal promoter activity in male and female gonadal somatic cells (116). This region contains putative SOX-like binding sites, which are conserved in human, mouse, and chicken (116). Luciferase activity from this minimal region was significantly higher in testicular cell lines than in other cell lines tested with the difference abolished when this region was deleted (116). This suggests that the −193 to −73-bp region contributes, but is not sufficient, to direct testis-specific Sox9 expression. To further delineate the regulatory elements in the Sox9 promoter, mice transgenic for large fragments of human SOX9 in yeast artificial chromosomes were generated (117). Several regulatory elements required for SOX9 expression in chondrocyte development were identified in a 350-kb region upstream of SOX9 (117). However, testis-specific SOX9 expression was not detected in any of the transgenic mice tested, indicating that additional testis-specific regulatory elements are located outside this region (117). A translocation breakpoint identified in a CD patient, which maps more than 950 kb from SOX9, supports this observation (87). A 150-kb gonad-specific regulatory element, 1 Mb upstream of Sox9, which mediates the repression of Sox9 expression in XX fetal gonads, was mapped (69).

Transgenic mice with an insertional mutation, odsx (ods), which lack this 150-kb region developed as sterile XX males lacking Sry and showed no skeletal defects (69). Hence, regulation of Sox9 expression in the testis is likely to be complex and involves proteins bound at distant regulatory elements at the Sox9 promoter and/or enhancer.

2. Role of SOX9 in chondrogenesis. Sox9 is expressed at high levels at all sites where cartilage is being laid down with expression most abundant in mesenchymal condensation just before overt chondrocyte differentiation (78, 89). Sox9 is expressed in the first and second branchial arches, in the sclerotomes, and the lateral plate mesoderm that gives rise to the appendicular skeleton (78, 89). The expression of Sox9 and type II collagen (Col2a1), an early marker of chondrocyte differentiation, coincide and peak at 11.5–14.5 dpc (78, 89). The severe skeletal malformations seen in patients with CD, which include bowing of the femora and tibiae, hypoplastic scapulae, pelvic malformations, and bilateral clubfeet, and the expression pattern of Sox9 in cartilage are consistent with a role of Sox9 in chondrogenic development (75, 76).

In an effort to identify transcription factors that control chondrocyte differentiation, a minimal sequence was identified in intron 1 of the Col2a1 gene, which contains SOX-like binding sites and was able to direct chondrocyte-specific expression in transgenic mice (118). Sox9 binds to this enhancer sequence and activates transcription of a reporter construct in COS7 cells (78, 89). Other SOX proteins tested, however, failed to activate transcription, suggesting that type II collagen is a direct target for Sox9 (78, 89). In addition, mutation of these sequences abolished chondrocyte-specific expression of a Col2a1-driven reporter gene in transgenic mice (88).

A number of in vivo experiments performed illustrate the importance of SOX9 in chondrogenesis. In transgenic mice, ectopic expression of Sox9 resulted in ectopic expression of endogenous Col2a1 in a number of tissues (88). However, in the genital ridge of transgenic mice, Col2a1 is not expressed, although there is a high level of endogenous Sox9 present in this tissue. Therefore, Sox9 alone is not sufficient to direct Col2a1 expression in vivo and suggests that Sox9 interacts with other chondrocyte-specific enhancer binding proteins. Sox9 is known to form a complex with Sox6 and L-Sox5, which also binds to the Col2a1 enhancer (119). The lack of expression of Col2a1 in nonchondrogenic tissues in transgenic mice that express Sox9 could be due to the lack of or insufficient expression of Sox6, L-Sox5, or additional chon-
droocyte-specific enhancer-binding proteins. Chimeric mice were produced that are mosaic for a targeted deletion of Sox9, replaced with the lacZ gene (120). At 15.5 dpc when chondrogenesis is occurring, Sox9+−/− cells express Col2a1, Col9a2, Col11a2, and aggrecan. These Sox9+−/− cells are interspersed with wild-type cells (120). Sox9+−/− embryonic stem cells did not differentiate into chondrocytes as per wild-type cells. In contrast, Sox9+−/− cells were not localized to the cartilage and lacked both morphological characteristics of chondrocytes and expression of early chondrocyte markers (120). Earlier in development, at 9.5 dpc when mesenchymal condensation is occurring, neither were Sox9+−/− cells interspersed with wild-type cells nor did the cells express Col2a1. This indicates that Sox9 is an early marker of chondrocyte development and is required for the activation of specific collagen genes.

Mice heterozygous null for Col2a1 show only mild skeletal defects, suggesting that the more severe skeletal defects observed in CD are likely to be due to inadequate activation by Sox9 of additional genes (other than Col2a1) that are required for chondrogenesis such as Col11a2, Col9a2, and aggrecan. This is supported by the observation that Sox9 can bind to the promoters and activate transcription of the Col11a2, cartilage-derived retinoic acid-sensitive (CD-RAP), and aggrecan (agg) genes and is therefore likely to be involved in regulating the expression of these genes (121, 122).

IV. Regulation of SRY Expression in the Developing Gonad

A. SRY promoter analysis

Comparison of the SRY 5′-regulatory region and promoter in 10 different mammals has revealed that, although there are a number of potential regulatory elements that are present in all the species examined, their location relative to the promoter is often very different (123, 124). In support of the implication from these studies, i.e., that regulation of SRY transcription varies between species, it has been shown that a human SRY transgene cannot sex reverse mice but a mouse SRY transgene with the human coding sequence can (56). Due to species differences, SRY regulation has been examined concurrently in both mouse and human systems. In humans, information about regulatory regions has been obtained from tissue culture studies and characterization of sex-reversal mutations that lie outside the SRY ORF. In mice, much of what is known about Sry regulation has come from analysis of transgenes containing nested deletions of the Sry gene.

When it was first demonstrated that an Sry transgene was capable of inducing male gonad development in XX mice, a 14-kb genomic fragment containing the Sry coding region and 8 kb of 5′-sequence was used (15). Expression of Sry from the original 14-kb construct is more widespread and persists for longer than the endogenous Sry gene (19, 57, 125), indicating that the transgene is subject to position effects or lacks cis-acting repressors that temporally and spatially limit its expression. Either way, it would seem that tight positional regulation of Sry expression is not necessary for sex determination in the mouse. Surprisingly, constructs with progressive deletions of 5′-sequences to within 57 bp of the transcription start site were still capable of inducing testis development (125).

Conventionally, promoter elements are referred to in terms of their location relative to the site of transcript initiation; however, the initiation codon has been used as a reference point for discussing the location of SRY elements because multiple initiation sites have been detected in humans and mice. In characterizing the human transcript, one study identified two SRY transcript start sites at −78 bp and −136 bp (126), and other studies have identified a major initiation site at −91 (26, 127) and a minor initiation site at least 410 bp further upstream (26) (Fig. 11). In the mouse, three major and one minor transcription initiation site have been identified between −269 bp and −256 bp (59) (Fig. 11).

Analysis of the SRY gene from sex-reversed individuals has revealed that there are three cases in which the ORF is normal but there are changes in the region that surrounds SRY. In two, there are changes in the 5′-region: one is a single-base change at −2027 (128), and the other is a deletion extending upstream from approximately −1.8 kb for at least 33 kb but not for more than 60 kb (129). In the 3′-region there is a deletion that starts between 2 kb and 3 kb downstream from the polyadenylation site and extends away from the SRY gene for approximately 500 bp (130). Although these results imply the existence of SRY regulatory elements more than 2 kb 5′ and 3′, there are other possibilities. The single nucleotide change at −2027 could be a very rare polymorphism, and the deletions could create position effects that attenuate SRY expression. Because these regions have not been examined further, the issue of whether they contain SRY regulatory elements remains to be resolved.

Initial characterization of the human SRY promoter showed that a 313-bp fragment (−408 bp to −95 bp), which contained only one of the identified transcriptional start sites (−136 bp), could direct reporter gene expression in Ltk− cells (126). Subsequently, it has been shown that this region also functions as an activator in TM4 and NT2D1 cell lines (126, 131). TM4 is a mouse Sertoli cell line, and NT2D1 is a human pluripotent clonal cell line derived from a teratocarcinoma. Characterization of these cell lines has demonstrated that they contain most of the genes known to be involved in sex determination (26, 108). Although the minimal SRY promoter identified in Ltk− cells can function as an enhancer in NT2D1 cells, additional upstream sequences further influence expression of reporter constructs in the NT2D1 cell line. In NT2D1 cells, basal activity of a minimal SRY promoter (−461 bp to −13 bp) produces an 18-fold increase in production of reporter compared with vector with no insert. Promoters that include 916 bp and 727 bp of sequence 5′ to the initiation codon produce a 37- and 54-fold increase in reporter, respectively (132). Thus, in NT2D1 cells, promoter elements 727 bp 5′ of the initiator codon are capable of directing high levels of reporter gene expression, and sequences between 727 bp and 916 bp may contain a repressor element. Within the human SRY promoter region, the binding of three transcription factors has been shown to influence expression of reporter constructs, namely, SF1, Sp1, and WT1 (below).
B. Regulation of SRY transcription by SF1 and Sp1

SF1 is a member of the nuclear hormone receptor superfamily and contains a Zn finger DNA binding domain at the N terminus and a ligand-binding domain at the C terminus (133). Sp1 is a ubiquitous transcription factor that binds a GC-rich consensus sequence and has been implicated in the regulation of many genes (134). Binding studies using NT2D1 extracts have demonstrated that Sp1 can bind consensus elements in the SRY promoter at −150 and −130 bp (135), and SF1 can bind to the NHR1 consensus site at −315 bp (132). Cotransfection experiments in the Sp1-deficient Drosophila Schneider cell line, SL2, showed that Sp1 was able to activate transcription of a reporter construct containing SRY sequences from −175 to +53. Mutagenesis of the consensus Sp1 site at −150 reduced reporter gene expression by 90%, and additional mutagenesis of the site at −130 bp did not further reduce reporter gene expression (135). In NT2D1 cells, mutation of either the SF1 site or Sp1 site at −150 reduced basal activity of a reporter construct by 50%, and mutation of both of these sites reduced promoter activity 90% (132). Thus, it seems that the main factors required for basal SRY expression in NT2D1 cells are SF1 and Sp1.

SRY expression is abolished by cAMP treatment of NT2D1 cells (132) and cultured Sertoli cells from postnatal testis (136). Both SF1 and Sp1 (−150) sites are required for this effect in NT2D1 cells (132). Coimmunoprecipitation analysis of NT2D1 cells showed that these two factors interact (132). cAMP treatment directs PKA-mediated phosphorylation of the SF1 ligand-binding domain. This phosphorylation event disrupts SF1 DNA binding and its interaction with Sp1 (132). These results imply that SF1 and Sp1 might cooperate to activate the SRY promoter through direct protein-protein interactions and suggest a mechanism for repressing SRY expression.

It seems possible that SF1 could be required for human SRY expression because a mutation in the gene has been found in a sex-reversed individual (111). It is difficult to determine whether this is a direct effect on SRY expression, however, because SF1 is also involved in the expression of other genes involved in sex determination (91, 107, 108, 113, 137).

In the mouse, the expression of both Sf1 (21) and Sp1 (134) in early embryogenesis is consistent with a role for these proteins in regulating mouse Sry expression in vivo. Careful analysis of the phenotype of Sf1 knockout mice has also shown that the gonad develops only for a short time before it arrests and then regresses (138). Further to this, Sf1 expression in the bipotential gonad is initially widespread and then becomes restricted to pre-Sertoli cells just before the expression of Sry (21). Although they may contribute, the Sf1 and Sp1 binding sites in the mouse Sry promoter lie outside the 57-bp region of the promoter that is capable of inducing sex reversal (139). In addition, it has been reported that Sry transcripts are still detectable in Sf1 knockout mice (140).

C. Regulation of SRY transcription by WT1

In mammals, the WT1 gene can give rise to 24 different protein isoforms through a combination of alternative splicing, alternative translational start sites, and RNA editing (141, 142). All isoforms contain four C-terminal C_{2}H_{2} Zn fingers and an N-terminal Pro/Glu-rich region (141, 142). There are two alternate splicing events of the 10-exon WT1 transcript resulting in the incorporation or omission of exons 5 and 9. Alternate splicing of exon 5 is specific to placental mammals, whereas the alternate splicing of exon 9 is observed in all vertebrates (141, 143). Variable splicing of exon 9 results in proteins that are distinguished by incorporation.
or omission of just three amino acids, Lys, Thr, and Ser (KTS) between Zn fingers 3 and 4. Analysis of the −KTS and +KTS isoforms has revealed that they have very different functions. WT1(−KTS) appears to act as a classical transcription factor because it can bind a consensus element in DNA and regulate the expression of reporter and endogenous genes in transient transfection assays. Although it was initially thought that WT1(−KTS) could function as an activator or repressor of transcription (144), more recent studies have shown that its main function is likely to involve activation (141, 145, 146). The WT1 (+KTS) isoform is thought to be involved in RNA message production or processing because it binds RNA and colocalizes with splicing speckles in the nucleus.

WT1 is involved in a number of human disorders associated with kidney and urogenital abnormalities including Wilms’ tumor, WAGR syndrome (Wilms’ tumor, aniridia, genitourinary abnormalities, and mental retardation),Denys-Drash syndrome (DDS), and Frasier syndrome. Wilms’ tumors are kidney tumors that occur in embryogenesis or early childhood, indicating that they involve a disruption in normal developmental processes. WAGR syndrome is associated with chromosomal deletions that encompass the WT1 gene and surrounding sequences including PAX6. DDS is associated with gonadal and urogenital abnormalities including sex reversal, and Wilms tumors are common in DDS families. DDS most frequently results from mutations in the Zn finger region that disrupts WT1 DNA binding. Frasier syndrome is not associated with Wilms’ tumors but patients exhibit gonadal dysgenesis and sex reversal. In this syndrome, mutations in the exon 9 splicing region abolish synthesis of the +KTS isoforms.

To investigate the role of WT1 in sex determination and differentiation, Hossain and Saunders (131) examined the effect that WT1 had on the expression of SRY in NT2D1 cells. Expression of endogenous SRY expression was up-regulated by transient transfection of expression constructs encoding WT1(−KTS) but not by constructs encoding WT1(+KTS). In reporter assays, activation of the SRY promoter by WT1(−KTS) required a consensus binding element (−78 to −87 bp from the initiating ATG codon). Other studies have shown that WT1(−KTS) also up-regulates expression of a number of endogenous genes including BCL2 (apoptosis inhibitor), CDKN1A (cyclin-dependent kinase inhibitor), amphiiregulin (an epidermal growth factor), and DAX1 (antitestis gene).

Inactivation of the mouse WT1 gene by gene knock out completely abolishes both renal and gonadal development (147, 148). Selective inactivation of WT1 isoforms has shown that mice lacking WT1(−KTS) had reduced gonad size but expressed male-specific markers, whereas WT1(+KTS)-deficient mice exhibited complete male-to-female sex reversal (149). In both cases, the knockouts died shortly after birth due to kidney defects. The WT1(+KTS) knockout phenotype closely parallels Frasier syndrome, which arises from mutations that disrupt WT1(+KTS) production. The isoform-specific knockout results indicate that both WT1(+KTS) and WT1(−KTS) are required for gonad development but only WT1(+KTS) is required for male sex determination. This hypothesis is complicated by two issues. First, WT1(−KTS) is required for expression of Dax1 in vitro and in vivo (149, 150). Dax1 is an antagonist of male development, and therefore it is possible that expression of male-specific markers in WT1(−KTS)-deficient mice could be due to derepression of the male developmental pathway in these mice. Second, in the isoform-specific knockout it was shown that disruption of one isoform led to a doubling in the levels of the other isoform. Although it is possible that some of the phenotype could be due to overexpression, it seems unlikely (151, 152). How the WT1(+/KTS) up-regulates Sry expression is not known, but considering the colocalization of this protein isoform with splicing centers in the nucleus has led to the proposal that it could function by increasing Sry message processing or stability. In support of a role for posttranscriptional regulatory mechanisms in the regulation of Sry, consensus transcript instability motifs have been identified in the mouse Sry transcript (59). In addition, replacing the 3′-untranslated region (UTR) with artificial UTRs abolished the ability of Sry transgenes to cause sex reversal (125). Interestingly, the antitestis factor DAX1 (dosage-sensitive sex reversal, adrenal hyperplasia, X-linked), like WT1(+KTS), also binds RNA. This raises the possibility that DAX1 could antagonize male development by disrupting the function of WT1(+/KTS), resulting in a decrease in the abundance of Sry message.

The demonstration that WT1(+/KTS) is required for male-specific gene expression in mice seems at odds with the situation in humans where there is support for the function of WT1(−KTS) in sex determination. In tissue culture studies, WT1(−KTS) is capable of activating endogenous SRY expression (131) and in DDS, disruption of the WT1 DNA binding domain, which is required for the activating function of WT1(−KTS), causes sex reversal. In WT1(−KTS) knockout mice, male-specific markers are present, but at reduced levels compared with wild type (149), suggesting that WT1(−KTS) could contribute to the full expression of SRY in vivo. These results indicate that the two isoforms of WT1 could have independent roles in sex determination whereby WT1(−KTS) contributes to the full expression of SRY and WT1(+KTS) is required for SRY message processing or stability.

The other alternate splicing site in WT1, which has been detected only in mammals, results in the incorporation or omission of exon 5, which encodes a 17-amino-acid sequence between the transcripational activation and DNA-binding domains. Because this region is found only in placental mammals, it was thought that this region might be specific to reproductive development in these animals. In some experimental systems, this 17-amino acid domain augments the function of the WT1 activation domain (153, 154), and in others systems it has been reported to act as a repression domain (155, 156). This region interacts with PAR4, a transcriptional cofactor that is expressed in the same tissues as WT1 (155). The WT1-PAR4 interaction has also been shown to both positively (157) and negatively (155) regulate WT1 transcriptional activity. A mutation within this alternatively spliced region (G253A) that does not bind PAR4 has been isolated from Wilms’ tumor specimens (158). To test the role of this domain, an allele-specific knockout has recently been produced. Interestingly, deletion of this region did not seem to affect viability or fertility (159).
In summary, it seems that SRY transcription is likely to be regulated by the coordinated action of a number of transcription factors including SF1, Sp1, and both the +KTS and −KTS isoforms of WT1. Dissecting the interplay between these factors has not only provided information about the elements in the sex determination cascade, it has also led to the formulation of a number of hypothesis about the mechanisms of DSS. DSS occurs in males that have a duplication of the short arm of the human X chromosome. Positional cloning identified the gene responsible for this syndrome as Dax1. Dax1 has been classified as an atypical member of the nuclear hormone receptor family because it contains the conserved ligand-binding domain but lacks the Zn finger DNA binding motif found in other family members. Mice containing additional copies of Dax1 exhibit male-to-female sex reversal. On the basis of studies that have shown that the Amh expression is reduced in these mice, it has been proposed that Dax1 antagonizes the function of Sry. Interestingly, it has been shown that Dax1 can antagonize the transcriptional activation function of SF1 by recruiting the nuclear receptor corepressor. Thus, in humans, it is possible that this suppression of male development could arise as a direct result of DAX1 inhibiting SF1-mediated expression of SRY. It seems unlikely that this would be the case in mice because Sry is expressed in SF1 knockout mice (140). The action of DAX1 could also be indirect. Because SF1 is required for expression of genes downstream of SRY in the sex determination cascade (91, 107, 108, 113, 137), DAX1 might inhibit events in testsis development after the sex determination stage. It has also been shown that DAX1 can bind RNA (160). Considering that mouse knockout studies showed production of Sry message required the RNA binding molecule, WT1 (−KTS), it is possible to speculate that DAX1 could antagonize sex determination by disrupting SRY message processing or stability. DAX1, like HMG box proteins, can bind the minor groove of DNA and stem loop structures (161); therefore, it could also disrupt SRY function through direct competition for sites involved in male sex determination. This array of possibilities could be narrowed through further examination of the DSS mouse model, specifically by looking at other markers of male sex determination that are expressed earlier than AMH. Because only AMH expression has been examined in this model, it is not possible to determine whether the DSS phenotype arises from disrupted SRY expression or inhibition of other factors required for the activation of AMH.

V. Regulation of SOX9 Expression in the Developing Gonad

Results obtained from mapping breakpoints that occur in the SOX9 promoter in CD patients have implied that the regulatory region of this gene could be very large. One translocation breakpoint has been identified in a CD patient (that did not result in sex reversal) that maps 950 kb upstream from SOX9 (86). Other CD breakpoints that do cause sex reversal have been mapped up to 350 kb from SOX9 (75, 86, 87, 117, 162–166) (Fig. 12). Information derived from the breakpoint mapping has shown that there is no apparent correlation between the position of the breakpoint and phenotype, e.g., there are two patients with breakpoints 110–142 kb upstream from SOX9, one of which exhibits sex reversal, whereas the other does not (86). To further delineate the regulatory elements in the SOX9 promoter, mice transgenic for large fragments of human SOX9 in yeast artificial chromosomes were generated (117). Several regulatory elements required for SOX9 expression in chondrocytes and neuroectoderm were identified in a 350-kb region upstream of SOX9 (117). However, testis-specific SOX9 expression was not detected in any of the transgenic mice tested, indicating that additional testis-specific regulatory elements could be located outside this region (117). In the mouse, a 150-kb gonad-specific regulatory element that mediates repression of Sox9 in XX fetal gonads has been mapped, 1 Mb upstream of SOX9 (69). Transgenic mice with an insertional mutation, odsex (ods), which lack this 150-kb region, develop as sterile XX males lacking Sry and showing no skeletal defects (69). Hence, regulation of SOX9 expression in the testis is likely to be complex and involves proteins bound to distant regulatory elements at the SOX9 promoter and/or enhancer.

Analysis of a 6.8-kb fragment of the mouse Sox9 promoter in testis and ovary cells isolated from 13.5 dpc embryos showed that a minimal interval located between −193 and −73 from the Sox9 transcription start site was sufficient to direct maximal promoter activity in male and female gonadal somatic cells (116). Luciferase assays from reporters contain-

![Figure 12: Structure of SOX9 promoter. Schematic of the human SOX9 promoter illustrating the relative positions of chromosomal breakpoints found in CD and putative regulatory elements conserved between Fugu, mouse, and human. The position of the breakpoints and conserved elements are given in kilobases from the site of SOX9 transcript initiation. Below each breakpoint, details regarding the phenotypic sex and karyotype are given (m, XY male; f, XX female; f*, sex-reversed female). LHX9, Lim homeobox gene 9; LM1, Lin-2, Mec3; GATA4, GATA binding protein 4; DMRT1, doublesex male abnormal-3 related testis gene 1; EMX2, empty spiracles type homeobox gene; DHH, desert Hedgehog. Modified with permission from S. Bagheri-Pam et al.: Genomics 78:73–82, 2001 (168). © Elsevier.](https://edrv.endojournals.org/...)

This figure shows the structure of the SOX9 promoter, with schematic representations of the human SOX9 promoter illustrating the relative positions of chromosomal breakpoints found in CD and putative regulatory elements conserved between Fugu, mouse, and human. The figure includes details regarding the phenotypic sex and karyotype at each breakpoint, with symbols for male (m), female (f), and sex-reversed female (f*). The regulatory elements are marked with labels such as LHX9, LM1, GATA4, DMRT1, EMX2, and DHH. The breakpoints and conserved elements are given in kilobases from the site of SOX9 transcript initiation. Below each breakpoint, specific information about the regulatory region is provided, including the phenotypic sex and karyotype details.
ing this minimal region were significantly higher in testicular cell lines and the difference was abolished when this region was deleted (116). Although there was a difference in reporter gene expression between ovary and testis, this difference was substantially less (4-fold) than that of endogenous Sox9 in cells from these tissues (116). Thus, the −93 to −73 region contributes, but is not sufficient, to direct testis-specific Sox9 expression. This region contains putative Sox-like binding sites, which are conserved in human, mouse, and chicken (116). The role of these sites is consistent with two theories of Sox9 regulation. The up-regulation of Sox9 in Sertoli cells immediately after expression of SRY suggests a role for SRY in the activation of Sox9 expression (63, 106). Persistent expression of Sox9 after down-regulation of Sry in the mouse supports a role for Sox9 in positively regulating its own expression (125). In the tissue culture studies the latter mechanism seems more likely because cells for the analysis were isolated from 13.5 dpc embryos and by this time Sry has been turned off in mice (14, 167).

Recently, some progress has been made in delineating the gonad-specific elements of the Sox9 promoter. Bagheri-Fam et al. (168) used genomic sequencing information to compare the promoters of mouse human and Japanese pufferfish (Fugu rubripes) over large regions to identify conserved elements. Fugu was chosen for this comparison because it contains a compact genome with minimal intragenic material and has an evolutionary distance from mammals of 450 million years. This analysis identified eight conserved regions (E1–E8) (Fig. 12). Five of these elements were scattered through 290 kb, 5’ to Sox9 in the region where most of the translocation breakpoints responsible for CD are found. Three elements were found 3’ to Sox9, with the most distant being 452 kb downstream. In addition, two conserved elements within the Sox9 3’-UTR were identified. A mouse line carrying a transgene reporter construct with the E3, E4, and E5 elements in front of a 200-bp proximal promoter driving lacZ was active in the chords of testis at 13.5 dpc (169). A similar reporter that included E1 and E2 was not expressed in testis (169). Analysis of the E3, E4, and E5 elements has revealed that the E3 contains consensus binding sites for Sox, WT1, and Prx2; the E4 contains no known transcription factor binding consensus; and the E5 region contains a putative Sox binding site (168). Prx2 is a paired like homeobox transcription factor that is expressed along with Prx1 in the mesenchymal condensations of facial, limb, and vertebral column skeletal precursors. Prx1/2 double mutants exhibit skeletal defects that are similar to those found in CD patients (170, 171); however, Sox9 expression is not affected in the mutants (171). The presence of the WT1 consensus site in E3 is consistent with this protein having a role in sex determination. The Sox sites in E3 and E5 fit well with both the proposed role of Sry in activating Sox9 expression and Sox9 maintaining its own expression through a positive feedback loop.

In other studies, it has been reported that Sox9 expression can be reproduced in mice containing a lacZ transgene reporter with 70 kb of 5’- and 30 kb of 3’-sequence flanking Sox9 (56). The only conserved element between human, mouse, and Fugu in this region, E1, does not direct gonadal expression from the Sox9 promoter when in conjunction with E2 (169). It is likely that elements required for mammalian sex determination are not conserved in Fugu because the sex determination pathway is not the same in these organisms. Fugu lack Sry but fish express Sox9 in Sertoli cells (92, 172). One hypothesis that ties the results obtained from the transgenic studies with current theories on Sox9 regulation is that the conserved Sox sites identified in E3 and/or E5 are involved in Sox9 autoregulation, whereas an activator in 100 kb of Sox9 flanking sequence is mammal, and possibly Sry, specific. This model is highly speculative and, considering the complexity of the Sox9 regulatory region, this model will undoubtedly turn out to be oversimplistic.

VI. Initiating Sex Determination: A Molecular Genetic Model

A model of the initial events in mammalian sex determination can be speculated based upon experimental evidence from human and mouse studies described in this review (Fig. 10). All the factors shown in the figure act within the Sertoli cell precursor lineage, and some presumably signal to other lineages. In the bipotential gonad, before the onset of sex determination, triggered by the expression of Sry, a number of early genes are expressed. Mice lacking Sf1 lack gonads and adrenals have abnormal gonadotropic function (173). Another essential early gene is Wnt1, expressed around 9 dpc in the undifferentiated genital ridge of mice. Wnt1 knockout mice at 11 dpc the celomic epithelium of the urogenital ridge does not thicken and become the prominent gonadal component (147). Other factors required for gonadal development include Lmx1, Lhx9, and Gata4, as null mutations in each of these lead to failure of gonadal development in both sexes. In the Lhx9 knockout, there is evidence that Sf1 expression depends upon Lhx9 (174). Sf1 expression is positively regulated in part by Lhx9 but Sf1 protein activity is negatively regulated by Dax1 (possibly via nuclear receptor corepressor, which occurs in steroidogenesis). Sf1 also positively regulates its own antagonist, Dax1 (175). Wnt4 is essential for ovarian development (176) and Wnt4 signals the suppression of Leydig cell production. Wnt4 activates Dax1 but must also activate other genes required for ovary formation because Dax1 is not required for testis formation. Clinically, overexpression due to gene duplication in either Wnt4 (177) or Dax1 (178) antagonizes testis formation probably by Dax1 interfering with Sf1 protein activity, giving rise to XY females. Sry expression is regulated positively by the DNA binding (−KTS) form of WT1 and also by an Sf1/P1 complex. The WT1 (+KTS), an RNA binding protein, up-regulates Sry mRNA levels possibly through enhancing stability or processing. Sox9 is expressed in the bipotential gonad but is present at low levels and is predominantly cytoplasmic (25, 63).

Once SRY is expressed in the XY gonad, Sox9 is strongly up-regulated, and its subcellular localization changes from cytoplasmic to nuclear, events possibly involving signals between SRY and importin β (99) or CaM (102). Sox9 up-regulation is triggered by SRY, directly or indirectly. At the time Sry is turned on, other genes such as Vnn1, Dhh, and Fgf9 are switched on—possible targets of SRY action or mediators of Sox9 expression. A day after Sry expression, both Wnt4 and Dax1 are down-regulated (14, 23, 176). SRY might inhibit
DAX1 directly or WNT4 activation of DAX1, possibly by competing for TCF/SOX DNA binding sites in the WNT signaling pathway (179). SOX9 does participate with SFI, WT1, and GATA4 proteins in the regulation of AMH promoter. This complex might be stabilized through HSP70-WT1 (115) and HSP70-SOX9 interactions (114). Because AMH is not required for testis formation, SOX9 must be involved in switching on testis determining genes. ATRX is thought to be expressed before AMH, given the high incidence of cryptorchidism and ambiguous genitalia in individuals with β-thalassemia/mental retardation syndromes (180). DMRT1 is also expressed early and required for Sertoli cell maturation (181).

VII. Conclusions and Future Directions

Since the discovery of SRY 12 yr ago the key question still remains—what is the molecular function of SRY? In the intervening years, a dozen or so players in the molecular network have been identified. The fact that we have only identified mutations in a minority of human sex reversal cases indicates that we have a long way to go. Other sex-determining loci include 9p24, a region deleted in some XY females that contain transcription factor genes DMRT1 and DMRT2 (182). In a case of 1p duplication in an XY female, a double dose of WNT4 has been proposed as the candidate (177). Other loci for which candidate genes have not been ascribed include deletion in XY females at 10q, and a duplication in an XX male at 22q (179). In addition, several sex-determining autosomal loci have been identified in the mouse (183, 184).

Large pedigrees of sex reversal have also been described in specific breeds of horses, dogs, pigs, and goats. Gonadal dysgenesis is associated with other syndromes such as Smith-Lemli-Opitz syndrome (185).

Several approaches hold much promise in identifying more sex-determining genes. Microarray screens of genital ridge genes comparing male vs. females and different developmental stages (60, 62, 186) are likely to yield many new candidates. Large-scale ethyl nitrosourea mutagenesis screens for sex-reversed phenotypes in mice are now underway in several laboratories (56). Biochemical approaches such as chromatin immunoprecipitation will help identify the genes directly acted upon by the transcription factors in the pathway. Also, newer approaches, such as proteomics, will no doubt enter the foray soon. Almost no ovariadetermining gene has been identified and our understanding of ovary formation is sadly deficient but new technologies should facilitate these endeavors.

Acknowledgments

We thank Paisu Tang, Louisa Ludbrook, and Kevin Knover for reading of the manuscript, Mandy Curd for secretarial assistance, and Sue Panckridge for artwork.

Address all correspondence and requests for reprints to: Dr. Vincent R. Harley, Prince Henry’s Institute of Medical Research, Monash Medical Center, P.O. Box 5152, Clayton 3168, Australia. E-mail: Vincent.Harley@med.monash.edu.au

This work was supported by National Health and Medical Research Council (NHMRC) Grant 198713. V.H. is supported by a NHMRC Senior Research Fellowship, M.J.C. by a NHMRC Peter Doherty postdoctoral fellowship, and A.A. by a NHMRC C. J. Martin postdoctoral fellowship.

References

20. Albrecht KH, Eicher EM 2001 Evidence that Sry is expressed in pre-Sertoli cells and Sertoli and granulosa cells have a common precursor. Dev Biol 240:92–107
35. Werner MH, Bianchi ME, Gronenborn AM, Clore GM 1995 NMR spectroscopic analysis of the DNA conformation induced by the human testis determining factor SRY. Biochemistry 34:11998–12004
36. Ng KW, Ridgway P, Cohen DR, Tremethick DJ 1997 The binding of a Fos/Jun heterodimer can completely disrupt the structure of a nucleosome. EMBO J 16:2072–2085
46. Mitchell C, Harley VR 2002 Biochemical defects in eight SRY

Harley et al. • Molecular Action and Regulation of SRY and SOX9

missense mutations causing XY gonadal dysgenesis. Mol Genet Metab 77:217–225


53. Dubin RA, Ostrer H 1994 Sry is a transcriptional activator. Mol Endocrinol 8:1182–1192


60. Wertz K, Herrmann BG 2000 Large-scale screen for genes involved in gonad development. Mech Dev 98:51–70


84. Huang W, Zhou X, Lefebvre V, de Crombrugghe B 2000 Phosphorylation of SOX9 by cyclic AMP-dependent protein kinase A enhances SOX9's ability to transactivate a Col2a1 chondrocyte-specific enhancer. Mol Cell Biol 20:4149–4158
86. Zehentner BK, Leser U, Bartscher H 2000 BMP-2 and sonic hedgehog have contrary effects on adipocyte-like differentiation of C3H10T1/2 cells. DNA Cell Biol 19:275–281
96. Shen WH, Moore CC, Ikeda Y, Parker KL, Ingraham HA 1994 Nuclear receptor steroidogenic factor 1 regulates the Mullerian

Endocrine Reviews, August 2003, 24(4):466–487 485

inhibiting substance gene: a link to the sex determination cascade. Cell 77:651–661


119. Lefebvre V, Li P, de Crombrugghe B 1998 A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are co-expressed in chondrogenesis and cooperatively activate the type II collagen gene. EMBO J 17:5718–5733


tumor 1 gene have distinct functions during sex determination and nephron formation. Cell 106:319–329


Hastie ND 2001 Life, sex, and WT1 isoforms—three amino acids can make all the difference. Cell 106:391–394


Wang ZY, Qiu QQ, Gurrieri M, Huang J, Deuel TF 1995 WT1, the Wilms’ tumor suppressor gene product, represses transcription through an interactive nuclear protein. Oncogene 10:1243–1247


Dimerization of SOX9 is required for chondrogenesis, but not for sex determination or fertility. Mol Cell Biol 22:4433–4438

Lalli E, Ohe K, Hindelang C, Sassone-Corsi P 2000 Orphan receptor DAX-1 is a shuttling RNA binding protein associated with polyriboasomes via mRNA. Mol Cell Biol 20:4910–4921


Ninomiya S, Isomura M, Narahara K, Seino Y, Nakamura Y 1996 Isolation of a testis-specific cDNA on chromosome 17q from a region adjacent to the breakpoint of t(12;17) observed in a patient with acampomelic campomelic dysplasia and sex reversal. Hum Mol Genet 5:69–75


Gibbons RJ, Higgs DR 1996 The a-thalassemia/mental retardation syndromes. Medicine (Baltimore) 75:45–52

Raymond CS, Murphy MW, O’Sullivan MG, Bardwell VJ, Zarkower D 2000 Dmrt1, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. Genes Dev 14:2578–2595


