Mammalian Testis-determining Factor SRY and the Enigma of Inherited Human Sex Reversal

FRUSTRATED INDUCED FIT IN A BENT PROTEIN-DNA COMPLEX

Received for publication, May 11, 2011, and in revised form, August 1, 2011. Published, JBC Papers in Press, August 17, 2011, DOI 10.1074/jbc.M111.260091

Nelson B. Phillips§, Joseph Racca§, Yen-Shan Chen†, Rupinder Singh§, Agnes Jancko-Radek§, James T. Radek§,³, Naïlinda P. Wickramasinghe§, Elisha Haas§, and Michael A. Weiss‡

From the §Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106 and the ‡Faculty of Life Sciences, Bar Ilan University, Ramat Gan 52900, Israel

Mammalian testis-determining factor SRY contains a high mobility group box, a conserved eukaryotic motif of DNA bending. Mutations in SRY cause XY gonadal dysgenesis and somatic sex reversal. Although such mutations usually arise de novo in spermatogenesis, some are inherited and so specify male development in one genetic background (the father) but not another (the daughter). Here, we describe the biophysical properties of a representative inherited mutation, V60L, within the minor wing of the L-shaped domain (box position 5). Although the stability and DNA binding properties of the mutant domain are similar to those of wild type, studies of SRY-induced DNA bending by subnanosecond time-resolved fluorescence resonance energy transfer (FRET) revealed enhanced conformational fluctuations leading to long range variation in bend angle. 1H NMR studies of the variant protein-DNA complex demonstrated only local perturbations near the mutation site. Because the minor wing of SRY folds on DNA binding, the inherited mutation presumably hinders induced fit. Stopped-flow FRET studies indicated that such frustrated packing leads to accelerated dissociation of the bent complex. Studies of SRY-directed transcriptional regulation in an embryonic gonadal cell line demonstrated partial activation of downstream target Sox9. Our results have demonstrated a nonlocal coupling between DNA-directed protein folding and protein-directed DNA bending. Perturbation of this coupling is associated with a genetic switch poised at the threshold of activity.

The male phenotype in eutherian mammals is determined by Sry, a gene on the Y chromosome (Fig. 1A) (1). SRY contains a high mobility group (HMG) box (2), a conserved motif of DNA binding and DNA bending (3). This central domain (black in Fig. 1B) and its basic tail (gray) are conserved among an extensive family of SRY-related transcription factors (designated SOX) broadly involved in developmental decisions. Of particular interest, SOX9 functions in pre-Sertoli cells downstream of SRY (Fig. 1A) (4). The SOX9-dependent program of testicular differentiation is initiated by SRY-responsive enhancer elements (5, 6).

Assignment of SRY as the testis-determining factor is supported by transgenic murine models (7) and studies of human intersex abnormalities (46, XY gonadal dysgenesis) (8–13). Clinical mutations in SRY lead to failure of testicular differentiation in embryogenesis and in turn to a female somatic phenotype (14). Such mutations are subject to genetic classification (Fig. 1B). The majority arise de novo as meiotic errors in spermatogenesis (auburn triangles in Fig. 1B) (14) and markedly impair DNA binding and/or DNA bending (6, 15–17). Variant SRY alleles may also be inherited. Whereas similar molecular mechanisms underlie inheritance of dysfunctional alleles from mosaic fathers (open red triangles in Fig. 1B), an enigma is posed by the compatibility of the same SRY allele with male or female development (closed red triangles). A model for such inheritance is proposed by intersex mouse phenotypes associated with the compatibility or incompatibility of the divergent Y chromosomes of Mus musculus domesticus and Mus domesticus poschiavinus (YDOM and YPOS) in Black 6 (C57BL/6) strains (18). Such background-dependent incompatibility has been proposed to reflect differences in SRY expression levels (19), the timing of expression in the differentiating gonadal ridge (19, 20), and/or coding region polymorphisms (21).

© 2011 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.
The HMG box is an L-shaped domain containing an N-terminal β-strand (consensus residues 1–11) and three α-helices (α1, α2, and α3; Fig. 1C and supplemental Fig. S1) (22). Packing of the β-strand against the C-terminal segment of α3 defines the minor wing; α1, α2, and the N-terminal portion of α3 comprise the major wing. Whereas the architecture-specific DNA-binding domains of HMG-1 and HMG-D are well ordered in the absence of DNA (23, 24), isolated SRY/SOX domains exhibit induced fit; in the free domains the N-terminal segment may in part be disordered, whereas α3 folds with only loose tertiary contacts (25, 26). Structures of SRY-DNA and SOX-DNA complexes have been determined by multidimensional NMR spectroscopy (27, 28) and x-ray crystallography (29, 30). Major and minor wings of the HMG box bind in a widened...
minor groove; sharp DNA bending is effected in part by partial side-chain intercalation at the crux of the angular domain (supplemental Fig. S2) (31, 32). The protein-DNA interface is extended by a basic tail C-terminal to /H9251 (28, 33) whose binding enhances kinetic stability (34).

This study focused on two clinical mutations in the mini-core of the minor wing, V60L and V60A (red and black asterisks in Fig. 1B) (8, 11, 35, 36). This site (motif position 5 in the HMG box) is of structural interest. Val-60 packs on specific DNA binding within an “aromatic box” defined by His-65, Tyr-69, and Tyr-72 in the HMG box; boxed) are also conserved. Residue numbers in human SRY are provided above the sequences; residue numbers in an HMG box consensus are provided below.

FIGURE 2. Sequence conservation of the minor wing of the SRY/SOX HMG box. Val is invariant (with the exception of buffalo SRY) at position 5 of the HMG box (asterisks) among mammalian Sry alleles (3) and conserved as either Val or Ile among SRY-related HMG boxes (Sox genes) (107). Side chains adjoining V5 in the mini-core of the minor wing but distant in the sequence (His-65, Tyr-69, and Tyr-72 in the HMG box; boxed) are also conserved. Residue numbers in human SRY are provided above the sequences; residue numbers in an HMG box consensus are provided below.

minor groove; sharp DNA bending is effected in part by partial side-chain intercalation at the crux of the angular domain (supplemental Fig. S2) (31, 32). The protein-DNA interface is extended by a basic tail C-terminal to \( \alpha_3 \) (28, 33) whose binding enhances kinetic stability (34).

This study focused on two clinical mutations in the mini-core of the minor wing, V60L and V60A (red and black asterisks in Fig. 1B) (8, 11, 35, 36). This site (motif position 5 in the HMG box) is of structural interest. Val-60 packs on specific DNA binding within an “aromatic box” defined by His-65, Tyr-69, and Tyr-72 (Fig. 1C, left). Conserved as Val among mammalian SRY domains (with the exception of Ile in buffalo SRY) and as Val or Ile among SOX domains (Fig. 2), Val-60 buttresses a protein surface that in turn contacts the bent DNA site (boxed in Fig. 3A) at its 3’-end (5’-TTTTGTGCA-3’ in the NMR structure determined by Clore and co-workers (33); 3’-subsite in the top strand is underlined in figure caption). The bound structure suggests that mutations at Val-60 (red asterisks) could lead to transmitted perturbations of the adjoining DNA-binding surface (Fig. 3, B and C; see also supplemental Fig. S3). Indeed, V60L was originally reported by Harley et al. (11) to abolish detectable specific DNA binding activity as probed by gel
mobility shift assay (GMSA). Although V60L and other substitutions at this site might be well tolerated in the flexible free domain, on DNA binding the larger Leu side chain would be expected to incur steric clash within the aromatic box; its accommodation within a reorganized minor wing might therefore perturb, like a string of dominos, key protein-DNA contacts (33). The molecular properties of V60A SRY have not previously been reported.

Our re-investigation was motivated by a seeming paradox; the inheritance of V60L SRY implies that this variant allele is compatible with two alternative developmental outcomes as follows: testicular differentiation leading to virilization and spermatogenesis (in the fertile father) or gonadal dysgenesis leading to a female somatic phenotype (in the sterile daughter). Whereas the reported loss of specific DNA binding activity (11) could rationalize the phenotype of the daughter, what might account for the phenotype of the father? Indeed, the reported biochemical perturbations would (if validated) challenge current models of SRY-dependent transcriptional regulation based on specific DNA binding and DNA bending (6, 15–17). These anomalies suggested either that testicular differentiation may occur in the absence of SRY-DNA interactions or that the original GMSA studies were technically misleading.

Although the possible inheritance of V60A SRY has not been ascertained, this related mutation was also associated with two developmental fates within the proband, i.e. distinct regions of testicular and ovarian differentiation (‘ovotestes’) with features of true hermaphroditism (37).9

To resolve the father-daughter paradox of V60L SRY, we undertook its multidisciplinary study in relation to wild-type SRY and V60A. Because GMSA studies may in principle be confounded by changes in the kinetic properties of protein-DNA complexes (leading to dissociation of complexes during electrophoresis), we sought to distinguish (a) between kinetic and equilibrium features of protein-DNA binding and (b) between structural and dynamic features of protein-directed transcriptional regulation.

9 The father of the intersexual patient bearing V60A was not available for genetic testing (36). Nonetheless, the V60A phenotype differs from classical Swyer syndrome associated with de novo mutations in SRY by the presence of a large phallus (partial external virilization), absence of a cervix (partial Müllerian regression), mixed gonadal histology (development of ovotestes), and asymmetric descent of the gonads toward or in the labioscotal fold. These anomalous features suggest an intermediate phenotype. By contrast, the V60L patient lacked features of external virilization, and on histopathological examination, her gonads were cystic but otherwise consistent with 46, XY pure gonadal dysgenesis (ovarian stroma lacking in germ cells).
DNA bending. Such distinctions may be drawn through the application of fluorescence and NMR spectroscopy. Of particular utility, fluorescence resonance energy transfer (FRET) provides both a kinetic probe (when coupled to stopped-flow detection) and a dynamic probe (when coupled to distance-distribution analysis) (34). To relate such probes to developmental outcomes, we further sought to correlate biophysical properties with changes in transcriptional regulation in a rodent cellular model of the downstream pathway of testicular differentiation in utero (38). Remarkably, this assay yielded intermediate values of transcriptional activity reminiscent of those observed in studies of murine Y chromosome/autosome incompatibility (19).

Patients with inherited sex reversal and their families define experiments of nature that offer a potential richness of biological insight. V60L and V60A SRY are partially active; altered packing in the minor wing of the HMG box leads to only subtle biochemical and biophysical perturbations. Yet to our knowledge, this study provides the first demonstration of a structural coupling between the frustrated induced fit of a DNA-binding motif, nonlocal dynamic features of DNA bending, and in turn the kinetic stability of the bent protein-DNA complex. Our studies of the father-daughter paradox have thus defined a developmental switch poised at the threshold of function.

**EXPERIMENTAL PROCEDURES**

**Plasmids**—Bacterial and yeast plasmids expressing human SRY or its fragments were constructed as described previously (39, 40). Consensus position 1 corresponds to residue 56 of intact human SRY. Genes and DNA sites are indicated in italics, and proteins in capital letters. For use in a yeast one-hybrid system (Y1H), DNA encoding the HMG box of human SRY was also subcloned into plasmid pGAD-T7 between NdeI and BamHI restriction sites (Clontech). The construction encodes a fusion protein containing an N-terminal nuclear localization signal (NLS), central activation domain derived from GAL4 (Saccharomyces cerevisiae), and C-terminal SRY fragment. Constructions were verified by DNA sequencing.

**Mammalian Expression Plasmids**—The SRY coding region was amplified by PCR and ligated into vector pCMX under the control of a CMV promoter (41). Following the initiator Met codon, the cloning site encodes in triplicate a hemagglutinin (HA) tag. Mutations were introduced using the QuikChange™ site-directed mutagenesis kit (Stratagene, San Diego). Constructions were verified by DNA sequencing.

**Y1H Reporter Strains**—26-bp oligonucleotides containing SRY consensus binding sites in triplicate (boldface, 5′-AATTTCGCAATTGTATTGTTATTGTT-3′ and complement) were annealed. Containing 5′-AA overhangs, the duplex was ligated into EcoRI/Xhol-digested plasmid pLacZi; this site lies upstream of lacZ. Ligation and insert orientations were verified by DNA sequencing. A control plasmid was likewise constructed bearing a triplicate repeat of a nontarget site (5′-ATCGAT-3′ and complement); the substituted bases (boldface) markedly impair specific binding of SRY (31). The plasmids were digested with NcoI, and the linear DNA was employed to create integrated reporter strains of *Saccharomyces cerevisiae* with YM4271 background (Clontech).

**Y1H Screening**—Reporter strains bearing integrated SRY target or nontarget sites were tested for nonspecific activation of lacZ on plates deficient in uracil and containing colorimetric indicator 5-bromo-4-chloro-3-indolyl-β-d-galactoside (X-Gal). Colonies with low levels of β-galactosidase expression were isolated and transformed by pGAD-T7-derived plasmids expressing wild-type or variant NLS-AD-SRY fusion proteins; transformants were selected on minimal media lacking uracil and Leu. Colonies were isolated and grown in liquid minimal medium under the same selection. Extent of SRY-dependent expression of β-galactosidase was evaluated by the following: (i) inspection of colony color following spotting the overnight culture (5 μl) on X-Gal plates (80 μg/ml agar medium) and (ii) quantitative assay of enzyme activity in liquid culture using ortho-nitrophenyl-β-galactoside as described by the vendor (Clontech).

**Mammalian Cell Culture**—Rodent cell line CH34 (38), kindly provided by T. R. Clarke and P. K. Donahoe (Massachusetts General Hospital, Boston), was employed for studies of SRY-regulation gene activation (48). Cells were cultured in Dulbecco’s modified Eagle’s medium (DMEM) containing 5% heat-inactivated fetal bovine serum at 37 °C under 5% CO₂.

**Transient Transfection**—Mammalian expression plasmids were transfected into CH34 cells using reagent FuGENE 6 as described by the vendor (Hoffmann-La Roche). Transfection efficiency was determined using a pCMX-GFP construct. Cells were co-transfected with SRY and pCMX-GFP at equal amounts, and efficiency was measured as a ratio of GFP-positive cells. Extent of wild-type or variant SRY expression was assessed by Western blot using an antisemur directed against the hemagglutinin tag (Sigma). At each dilution, the level of expression of the variant SRY proteins was similar to that of the wild type. Further details are provided in the supplemental material.

**Transcriptional Activation Assay**—SRY-directed transcriptional activation of endogenous Sox9 was evaluated in CH34 cells (14). The fold increase in abundance of Sox9 mRNA was probed by real time quantitative (RT-Q) PCR as described previously (40). In brief, the Sox9 assay employed 50-μl PCR mixtures, containing 25 μl of SYBR Green (Bio-Rad), 500 nM of each primer, 1 μg of RNA extract, and 1 μl of iScript™ reverse transcriptase. The first real time PCR step was 30 min at 50 °C and 5 min at 95 °C, followed by 45 cycles of denaturation for 10 s at 95 °C and annealing/extension for 30 s at 57 °C. Fluorescent intensities were recorded and analyzed during PCR in an ABI Prism 7700 sequence detector system (Applied Biosystems) using the SDS (version 1.91) software. Dissociation curves for Sox9 were generated after each run to confirm that the increased fluorescence intensities were not attributable to non-specific signals (primer-dimers). CT values within the linear exponential increase phase were used to measure original mRNA template copy numbers and construct calibration curves. The concentration of plasmid CDNA for each target mRNA was estimated by measuring the A_{260} in triplicate. Copy number was calculated according to the following formula: copies/ml = 6.023 × 10^{23} × C × A_{260}/(molecular mass). Assays were performed in triplicate; Western blots and RT-PCR were
performed in each case to normalize the quantity of expressed SRY.

**Protein Purification**—Intact SRY and its domains were expressed in *Escherichia coli* strain BL21(DE3)pLysS and purified as described previously (42, 43). Purity was >98% as assessed by SDS-PAGE; results of mass spectrometry were in accordance with expected values.

**Gel-based DNA Binding and DNA Bending Assays**—A gel mobility shift assay (GMSA) employed a 36-bp consensus DNA site (5′-CATACTGCGGGGTATGTTCCAGGATCAT-3′ and complement) as described previously (31, 43, 44). DNA bending assays were performed by permutation gel electrophoresis (PAGE) using 140-bp DNA fragments as described previously (34) (see also supplemental Methods). SRY-directed bending of the consensus site and two variant sites was investigated (see Table 1).

**Circular Dichroism**—Spectra were obtained in a 1-mm path length quartz cuvette using an Aviv spectropolarimeter (Aviv Biomedical, Lakewood, NJ). Thermal unfolding of the free domains was monitored at 222 nm at 2 °C increments from 4 to 80 °C in 140 mM KCl and 10 mM potassium phosphate (pH 7.4) as described previously (43). Thermal unfolding of equimolar protein-DNA complexes (made 25 μM in the same buffer) was likewise monitored using a 12-bp consensus DNA duplex (5′-GTTGATTTGTCAGGAC-3′ and complement); CD spectra (200–320 nm) were measured from 4 to 90 °C at 2.5 °C increments.

**Protein Stability**—Fractional unfolding of the wild-type or variant SRY domains was evaluated as a function of the concentration of guanidine hydrochloride or urea by means of CD (at helix-sensitive wavelength 222 nm) or intrinsic Trp fluorescence as described in the supplemental material. CD and fluorescence-based denaturation were observed to be coincident.

**Fluorescence Spectroscopy**—For FRET studies, the respective excitation and emission wavelengths were 490 and 520 nm (donor) and 560 and 585 nm (acceptor). The donor quantum yield was determined for each sample using a reference solution of fluorescein in 0.1 M NaOH (0.85; see Ref. 45). The calculated value for the quantum yield of the donor as attached to the 5′-end of the DNA duplex was 0.83 at 4 °C and varied between 0.8 and 0.9 for all other samples. The Förster critical distance \( R_0 \) was calculated accordingly (58.3 Å for the free DNA).

**FRET-based Dissociation Constants**—Steady-state FRET was employed to determine the protein-DNA dissociation constant \( K_d \) for the wild-type SRY HMG box and its variants on binding to a 5′-FAM/TAMRA-labeled 15-bp DNA duplex containing a consensus SRY target site as described previously (34). Sample preparations were described in the supplemental material. \( K_d \) values were determined by plotting change in fluorescence intensity at 520 nm against total protein concentration. Data were fit to a single-site ligand-binding equation as implemented in Origin 8.0 (OriginLab Corp., Northampton, MA) as described in the supplemental material.

**Time-resolved FRET and Global Analysis**—Fluorescence measurements and analysis were as described previously (34). In brief, distribution functions were obtained from simultaneous global analysis of four experimental fluorescence decay curves; these were recorded for (i) samples containing only the donor, (ii) samples containing only the acceptor, (iii and iv, respectively) donor- and acceptor decay curves from samples containing both fluorophores. Global analyses were performed using the Marquardt nonlinear least squares method (46). Calculated decay curves were prepared by numerical solution of a modified version of the second-order differential equation (46, 47) as described in the supplemental material.

**NMR Spectroscopy**—Spectra of free domains were observed at 600 MHz in (i) 10 mM deuterated acetic acid (pD 4.5) and 140 mM KCl and (ii) 10 mM potassium phosphate (pD 7.6) and 140 mM KCl (4 and 25 °C). Spectra of specific protein-DNA complexes were obtained at 5, 15, 25, and 40 °C in 10 mM potassium phosphate (pH 6.0) and 50 mM KCl. The DNA site contained 15-bp (5′-GGGTTATTTGTTGAC-3′ and complement); core consensus target site is in boldface) (31). Analysis of solvent exchange by imino DNA resonances employed a NOESY experiment with selective excitation (flip-back pulse sequence) (49). Resonance assignment of the wild-type complex was obtained by heteronuclear multidimensional NMR methods as described previously (28, 33). NOESY mixing times were 120 and 200 ms; TOCSY mixing times were 55 ms. Partial assignment of the variant complex was obtained by analogy.

**Molecular Modeling**—Energy minimization and models of variant SRY HMG boxes were obtained using CHARMM (50) based on the solution structure of the domain as bound to 14-bp DNA site 5′-GGTGTTCAGGAC-3′ and complement (core target site in boldface; Protein Data Bank code 1J4E). The protocol employed the isolated protein domain. Only box residues 4–6, 64–66, and 68–73 (residues 59–61, 119–121, and 123–128 in human SRY, respectively) were allowed to move; the remaining main-chain and side-chain atoms were constrained to positions in the mean NMR model (28). Energy minimizations were performed using steepest descent with tolerance of 0.001 kcal/mol between steps. Models were visualized with InsightII (Accelrys, Inc, San Diego).

**RESULTS**

**Y1H Screening**—Re-investigation of V60L SRY and its extension to V60A were motivated by the unexpected results of Y1H screening of clinical mutations. The system employed an integrated reporter constructed to express β-galactosidase under the control of triplicate consensus SRY-binding sites (5′-ATT-GTT-3′ and complement; Fig. 4A). Expression of the reporter was regulated by a plasmid-encoded fusion protein containing the transcriptional activation domain of Gal4 linked to the wild-type or variant SRY HMG box (Fig. 4B). Whereas the wild-type fusion protein gave rise to blue colonies on an indicator plate containing X-Gal, the deletion of the SRY HMG box (Fig. 4C, empty) or its substitution by the inactive variant I68T (Fig. 4C, right) led to loss of β-galactosidase expression. Whereas previous GMSA studies of V60L SRY had shown no detectable specific DNA binding (11), in this Y1H system the V60L fusion protein gave rise to powder-blue colonies (Fig. 4D, 10). Clinical mutation I68T markedly impairs specific DNA binding because of perturbation of partial side-chain intercalation by the foreshortened and polar cantilever; structure-specific DNA binding to the preformed sharp angles of a four-way DNA junction is retained (72).
Inherited Human Sex Reversal

FIGURE 4. Design and application of yeast one-hybrid system. A, integrated Y1H reporter containing three adjoining SRY target sites (5'-ATTGTT-3' and complement; red, black, and blue) inserted upstream of reporter lacZ (black rectangle) and downstream of selectable marker ura3 (gray rectangle). Not shown: analogous control construction containing a nonspecific insert (5'-ATCGTAATCGTAATCGTA-3') lacking specific SRY-binding affinity. B, expression plasmid pGAD-T7 encoding a fusion protein consisting of an N-terminal nuclear localization signal (nis, aquamarine), the transcriptional activation domain (AD; blue) of yeast regulator Gal4, and the SRY HMG box (black) under the control of the ADH1 promoter (red). C, left to right, yeast colonies grown on X-Gal plates respectively expressing the wild-type SRY HMG box fusion protein, nis-activation domain in the absence of an SRY domain ("empty"), or the I68T SRY domain (cantilever position 13 in the HMG box consensus). The latter variant is associated with a de novo case of Swyer syndrome because of impaired specific DNA binding (53, 57). Only the wild-type fusion protein gave rise to a blue colony with this colorimetric indicator. D, analogous comparison of empty and wild-type (wt) colonies (left) with variant fusion proteins expressing V60A and V60L fusion proteins (right). The variants gave rise to blue colonies of hue similar to (V60A) or lighter than (V60L) than that of the wild type. E, quantitative analysis of β-galactosidase activity in liquid culture in Miller units. Each expression plasmid (left to right, empty, wild-type, V60A, and V60L) was tested in the 5'-ATTGTT-dependent reporter strain and the 5'-ATCGAT-containing control strain. ortho-Nitrophenyl-β-galactoside values obtained in the specific reporter strain are labeled "S" (left-hand bar in each pair); ortho-nitrophenyl-β-galactoside values obtained using the nonspecific reporter strain are labeled "N" (gray; right-hand bar). The V60L fusion protein (green) gave rise to half the level of enzyme activity as the wild-type fusion protein (black); V60A (red) gave rise to similar enzymatic activities. Assays were performed in triplicate; error bars represent standard deviations. Values were normalized according to cell density but not according to level of protein expression. Asterisks indicate statistical significance with p value < 0.001 (Student’s t test).

right) whose hue was distinct from that of the empty vector (Fig. 4D, left); the color of V60A colonies was similar to that of wild-type (Fig. 4D, center). Trends observed by inspection of X-Gal plates were verified by quantitative assessment of β-galactosidase activity in cell extracts as normalized by cell density (histogram in Fig. 4E). Thus, the Y1H results strongly suggested that specific DNA binding activity was retained by variant SRY HMG boxes V60L and V60A. Control studies of the four expression constructs (Fig. 4E, empty, wt, V60A, and V60L) in a reporter strain containing an integrated nonspecific lacZ construct (obtained by integration of an 18-bp variant DNA site in triplicate; see "Experimental Procedures") did not result in an SRY-dependent increase in β-galactosidase activity (gray bars).

Fluorescence Studies of DNA Binding and DNA Bending—Because GMSA can be confounded by dissociation of protein-DNA complexes during electrophoresis, SRY-DNA binding was re-investigated by steady-state FRET studies of a 15-bp DNA duplex (5'-TCGGTGATTGTCTCAG-3' and complement; target site in boldface) as described previously (34). Changes in the distance between respective 5'-ends of the upper and lower strands on protein binding were monitored by changes in steady-state FRET efficiency between FAM (donor) and TAMRA (acceptor) (Fig. 5A). The donor and acceptor were flexibly linked to respective 5'-ends of the DNA, justifying the customary assumption that κ2 (the orientation factor) is 2/3 and leading to an ensemble of fluorophore positions (supplemental Fig. S5).

Equimolar solutions of the DNA and SRY domain were prepared at mutual concentrations of 1 μM. In accordance with PGE (39) and the sharply bent NMR structure of the SRY-DNA complex (28), binding of the wild-type SRY HMG box to this donor/acceptor-labeled DNA probe led to an increase in FRET efficiency at 15 °C from 51 (±1)% in the free DNA to 69 (±1)% in the wild-type complex (Fig. 5B). Similar increases in FRET efficiency were observed on binding of the V60L and V60A SRY HMG boxes, indicating that under these conditions the variant domains are capable of binding and bending a consensus SRY target site with native-like DNA bend angles (see below; time-resolved FRET). Dissociation constants were determined based on titration of the donor/acceptor-labeled DNA probe (25 nm) with successive amounts of SRY domain (1 nm to 2 μM) at 15 °C (Fig. 6, A–C). Curve fitting yielded respective Kd values of 14 ± 3 nm (wild-type domain), 32 ± 2 nm (V60L), and 27 ± 5 nm (V60A). The mutations thus each impair specific DNA binding by about 2-fold. Because Val-60 does not itself contact DNA (28), these modest changes reflect transmitted effects of the mutations.

Dissociation rate constants were characterized using a FRET-based kinetic assay (Fig. 6D). Wild-type or variant domains were prebound to the same double-labeled DNA duplex as above. On rapid mixing of the preformed complex with a 20-fold excess of an unlabeled DNA target site, an increase in donor fluorescence was observed with time, which reflects the dissociation rate as the protein redistributes among DNA-binding sites (Fig. 6E). Observed rate constants (kobs) were determined in each case by fitting 5−7 individual traces to a single exponential equation and averaging the results. Under

1 Although reductions of end-to-end distance can reflect both DNA bending and correlated changes in DNA unwinding, the predominant contribution is due to DNA bending (43, 97, 98).
In these conditions, V60L and V60A complexes were observed to dissociate markedly more rapidly ($k_{off}$ values 0.24($\pm$0.008) and 0.17($\pm$0.005) s$^{-1}$, respectively) than did the wild-type domain ($k_{off}$ = 0.033 ($\pm$0.001) s$^{-1}$). Given the dissociation constants determined above, respective association rate constants ($k_{on}$) were inferred to be 2.4($\pm$0.5) $\times$ 10$^6$ M$^{-1}$s$^{-1}$ (wild-type domain), 7.5($\pm$0.5) $\times$ 10$^6$ M$^{-1}$s$^{-1}$ (V60L), and 6.3($\pm$1.2) $\times$ 10$^6$ M$^{-1}$s$^{-1}$ (V60A). Mutation of the minor wing thus accelerates both association and dissociation. Modest changes in affinity thus mask more profound changes in kinetics.

The marked kinetic instability of the variant protein-DNA complexes may account for the original failure to detect such complexes by GMSA in cellular extracts (see “Discussion”) (11). Accordingly, we re-investigated the GMSA approach using a modified protocol designed to minimize kinetic effects during electrophoresis; relative to the original studies, the temperature of gel was reduced from room temperature to 4 °C and maintained by a circulating refrigerated water bath during electrophoresis; the percent acrylamide in the gel was increased from 4 to 8%; the radiolabeled DNA probe was increased in length from 29 to 36 bp; and because purified proteins were employed, non-specific DNA (e.g., pre-addition of poly(dI-dC) and salmon sperm DNA to cellular extracts) was not included. Under these conditions GMSA yielded similar estimates of protein-DNA dissociation constants (supplemental Fig. S4). Whereas wild-type and V60A domains gave rise to indistinguishable patterns of shifted bands, apparent binding of the V60L domain was slightly reduced; the fractional change in $K_d$ was less than 2-fold. The sequence specificities of the variant domains, defined by the GMSA affinity of the 5'-ATTGTT consensus target site relative to a set of variant sequences containing single transitions or transversion (38), were indistinguishable from that of the wild-type domain (data not shown).

The feasibility of GMSA enabled assessment of protein-directed specific DNA bending by PGE. In accordance with steady-state FRET studies (see above), no significant differences were observed in patterns of electrophoretic mobility as a function of flexure displacement on binding of the wild-type or variant domains to the consensus DNA site (5'-ATTGTT-3' and complement; Fig. 5, C and D). Decrements in DNA bend

---

**FIGURE 5. Studies of specific DNA bending.** A, schematic illustration of DNA probe design, a central bend reduces distance between respective 5' ends. One 5'-end is labeled with fluorescein (FAM; donor), and the other 5’ end with TAMRA (acceptor). The bent DNA site was depicted based on the co-crystal structure of the Sox17-DNA complex (29). B, emission spectra of free and bound DNA following excitation at 465 nm. DNA singly labeled is as follows: red, emission of fluorescein donor in free DNA; purple, complex with wild-type (wt) SRY domain; and powder blue, complex with V60L domain (overlapping spectra). Double-labeled DNA is as follows: black triangle, free DNA; green filled triangle, wild-type SRY complex; and green open triangle, V60L SRY complex. Emission spectra of DNA singly labeled with rhodamine acceptor (free, complexed with wild-type SRY and complexed with V60L SRY) are shown in turquoise and underlays. C and D, permutation gel electrophoresis comparing the wild-type and variant SRY domains. The DNA fragments (147 bp) contain consensus site 5'-ATTGTT-3' (and complement) respectively positioned at 120, 95, 79, 51, 47, and 27 bp from one end. Within experimental error, no differences in bend angles were inferred (see Table 1). In each case the samples (2x for V60A) were loaded side by side on the same gel; for clarity, lanes containing unrelated samples were deleted from the images. See supplemental Fig. S15 for PGE studies of a nonconsensus site.
angles ($\Delta \theta$) were detected, however, on binding of the domains to nonconsensus DNA sites of lower affinity (Table 1 and supplemental Fig. S15). Interestingly, bend angles in the variant complexes were more strongly perturbed by the changes in DNA sequence, including the target site employed in the NMR studies of Clore and co-workers (column 3 in Table 1) (28). On binding to a variant DNA site of lower affinity (5'-ATTGAT-3'), for example, the PGE-defined bend angles of the wild-type and V60L complexes differed by 4° ± 2° (column 4 in Table 1). This change in DNA sequence adjoins the minor wing near the site of mutation (Fig. 3B), suggesting that the nonconsensus DNA site unmasks a potential DNA-bending defect not observed in a consensus complex.

**Structure and Stability of the Variant HMG Boxes**—The L-shaped DNA-bound structure of the SRY HMG box is similar to that of isolated nonsequence-specific HMG boxes (15). Conserved residues in each case pack to form two clusters of nonpolar and aromatic side chains. The first and larger cluster occurs at the confluence of helices $\alpha$1-$\alpha$3 (major wing); the second is formed by an extended N-terminal segment and $\alpha$3 (minor wing). Whereas the structure of the SRY-DNA complex predicts that V60L and V60A would perturb this "mini-core" within the minor wing (33), the induced fit model implies that the mutations would be well tolerated within a disordered N-terminal segment (Fig. 1C).

To test this prediction, we investigated the structure of the free domains by far-ultraviolet CD (a probe of secondary structure; Fig. 7A) and intrinsic Trp fluorescence (a probe of quenching because of side-chain desolvation in the major wing$^{12}$) (Fig. 7B). In each case the wild-type spectrum (solid line in Fig. 7, A and B) coincides with the spectra of the V60L domain (green) and V60A (red) domains. $^1$H NMR studies of V60L further demonstrates an absence of detectable perturbation (supplemental Figs. S6 and S7). Whereas the methyl resonances of Val-60 and Leu-60 are each shifted upfield in the protein-DNA

---

**TABLE 1**

| Substitution | 5'-ATTGTT-3' | 5'-TTTGTA-3' | 5'-ATTGAT-3' |
|--------------|--------------|--------------|--------------|
| Wild type    | $-5^\circ$   | $-5^\circ$   | $-16^\circ$  |
| V5L          | $<0.5^\circ$ | $-3^\circ$   | $-20^\circ$  |
| V5A          | $-0.5^\circ$ | $-7^\circ$   | ND           |

$^a$ The native bend angle using the DNA site 5'-GGGTATTTGTGCAGG-3' (and complement) is estimated to be 80° from PGE data obtained in 10% (29:1) gels and 72° from data obtained in 8% (29:1) gels. Despite such variation, the precision of PGE-based estimates of DNA bend angles under a given set of conditions is in general ±1°, implying uncertainties in $\Delta \theta$ values of ±2° ND, not determined.

$^b$ The T→A transversion, which adjoins the minor wing of SRY near Val-60, leads to a marked loss of SRY binding affinity (31).

$^c$ No qualitative differences were observed by inspection of PGE patterns as defined side-by-side in the same gel (see Fig. 5, C and D); such relative patterns (i.e., whether a mutation leads to increased or decreased DNA bending) are robust to gel composition.

---

12 The SRY HMG box contains three Trp residues (Trp-70, Trp-98, and Trp-107; consensus positions 15, 43, and Trp-52, respectively). Trp-70 and Trp-98 are largely buried within the major wing, whereas Trp-107 projects from its surface. The edge of Trp-98 is exposed and contacts the bent DNA site as part of a wedge of nonpolar residues adjoining the cantilever side chain (54).
complex (see below), no such upfield shifts were observed in the free domains, presumably because of the lack of stable packing of the respective aliphatic side chains within the aromatic box of the minor wing as posited in the induced fit model (Fig. 1C).

$^1$H NMR probes of the major wing are likewise unperturbed; the chemical shift dispersion of Trp and Phe side chains and their patterns of conformational broadening are indistinguishable in the wild-type and V60L domains. On specific DNA binding, $^1$H NMR chemical shifts and NOE patterns within the minor wing are in accordance with the canonical structure of an HMG box (supplemental Fig. S8).

The thermal stabilities of the native and mutant domains are indistinguishable ($T_m$ 40$\pm$1°C; Fig. 7, C and D). Estimates of the free energies of unfolding ($\Delta G_u$) were obtained from studies...
of guanidine-induced protein denaturation as probed at 4 °C by CD (at 222 nm) or Trp fluorescence (at 390 nm emission wavelength after excitation at 270 nm). These probes were found to be concordant, indicating coincident loss of α-helical structure and exposure of internal Trp side chains. Application of a two-state model yielded similar estimates of ΔG° as extrapolated to zero denaturant concentration (supplemental Table S1) as follows: 3.8(±0.1) kcal/mol (wild-type), 3.7(±0.1) kcal/mol (V60L), and 3.9(±0.1) kcal/mol (V60A). Such unperturbed thermal and thermodynamic stabilities would be inconsistent with mutational induction of steric clash (V60L) or cavity formation (V60A) in a pre-stabilized minor wing. As expected, the wild-type and variant domains exhibited enhanced thermal stabilities on specific DNA binding (Fig. 7, E and F). Protein unfolding and DNA melting occurred in the same temperature range (55–65 °C). Although the extent of thermal stabilization would depend on the concentration of the complex (as a bimolecular association) and on DNA length (as DNA melting supervenes), qualitative inspection of the denaturation transitions revealed an order of apparent thermal stabilities (wild-type > V60A > V60L; respective black, red, and green lines in Fig. 7F) that mirrors their order of specific DNA affinities as determined by FRET.

Dynamic Perturbation of Bent V60L Complex—A subtle difference between wild-type and V60L protein-DNA complexes was detected by subnanosecond time-resolved measurements of the decay of donor fluorescence (Fig. 8A). Such time domain data make possible distance distribution analysis using skewed gaussian models of end-to-end distances in the bent DNA site (Fig. 8B) (46, 47). In previous such studies, we have reported that such time-resolved FRET measurements are remarkable for a significant change in donor lifetime in the double-labeled DNA site on binding SRY domains. Global analyses yielded estimates of distance distributions in the free and bound states (Table 2). On binding of the wild-type domain, the mean end-to-end distance of the DNA at 15 °C was reduced from 61.7(±0.8) Å (row 3 in Table 2 and thick black line in Fig. 8B) to 51.3(±1.0) Å (row 1 and thin black line). Similarly, on binding of the V60L domain, the mean distance was reduced to 51.9(±0.7) Å (row 2 in Table 2 and green distribution in Fig. 8B), which is equivalent (with error) to the mean distance in the wild-type complex.

Skewed gaussian modeling of each protein-DNA complex indicated a significant broadening of end-to-end distributions in the protein-DNA complexes relative to free DNA. Whereas the breadth of the distribution in the free DNA (−14.5(±2.4) Å) is consistent with the allowed range of flexible linker configurations, the widths increased to 19.5(±0.9) Å in the wild-type complex and 24.3(±0.7) Å in the corresponding V60L complex (Table 2). Such broadening suggests the following: (a) the wild-type SRY complex exhibits a distribution of populated DNA bend angles, and (b) on mutation of the minor wing, a perturbation in dynamic stability leads to further excursions in bend angle. Time-resolved FRET has thus provided evidence for a nonlocal coupling between induced fit of the minor wing and the dynamics of overall DNA bending (model II at right in Fig. 8C). Transmitted effects of V60L are thus distinct from those of mutations that lead to an altered but well defined DNA bend angle (model I, center, in Fig. 8C).

NMR Features of the V60L Complex—Intrinsic probes of the bent DNA structure are provided by 1H NMR imino resonances (at left in Fig. 9A). In the free DNA guanine imino resonances (bp 2–4, 6, 10, and 13) are clustered between 12 and 13 ppm, whereas thymidine resonances (bp 5, 7–9, 11 and 12) are clustered between 13.4 and 14.0 ppm (spectrum i in Fig. 9A). Addition of either wild-type or V60L to form a specific 1:1 complex results in marked changes in chemical shifts (spectra ii and iii in Fig. 9A). These complexes exhibit similar patterns of changes in chemical shift (formation shifts; Δδ). In each case insertion of cantilever side chain Ile-68 between T8 and T9 is associated with an anomalous upfield shift of T8 into the guanidine region (31). Whereas the downfield shift of T11 is essentially identical in the two complexes (vertical line between spectra ii and iii at left in Fig. 9A), small differences (<0.05 ppm) were observed in the extent to which imino resonances T9 and G10 are shifted upfield (broken vertical lines between spectra ii and iii in Fig. 9A). These base pairs are closest to the site of mutation in the minor wing.

Induced fit of the minor wing on specific DNA binding results in upfield shifts in the methyl resonances of Val-60 (0.7–0.0 ppm region at right in spectrum ii in Fig. 9A). These shifts presumably reflect the aromatic ring currents of side chains His-120, Tyr-124, and Tyr-127 (the latter two are depicted in Fig. 1C; see also supplemental Table S3). Indeed, the NOESY spectrum of the wild-type complex contains prominent cross-peaks between the Val-60 γ-methyl resonances and these aromatic spin systems (spectrum iv in Fig. 9B and supplemental Fig. S8). The minor wing of the V60L domain undergoes analogous induced fit. The δ-methyl resonances of Leu-60 also exhibit upfield shifts (although to a lesser extent than do the γ-methyl resonances of Val-60; at right in spectrum iii in Fig. 9A) and share NOEs with the same aromatic side chains across the minor wing (asterisks in spectrum v in Fig. 9B, right). The variant minor wing nonetheless exhibits a subtle NMR difference: whereas the wild-type spectrum contains an unusual protected Tyr para-OH resonance at the DNA interface (assigned to Tyr-129; position 74 in the HMG box (34), the curved arrow near 9.8 ppm in spectrum ii in Fig. 9A), this resonance is broadened at the variant interface (asterisk in spectrum iii). We ascribe such broadening to accentuated millisecond conformational fluctuations and/or enhanced exchange of the protected Tyr-129 para-OH proton with bulk solvent.

The nonlocal change in distribution of DNA bend angles in the variant complex, as inferred above from time-resolved FRET analysis, stands in contrast to the remarkable similarity of NMR features in the major wing and at the major wing/DNA interface. (a) The upfield δ-CH3 resonance of cantilever side chain Ile-68 is near −1.2 ppm in both complexes and exhibits intermolecular NOEs to the H2 resonances of the flanking adenine bases at the site of partial intercalation (spectra iv and v; at upper left in Fig. 9B). (b) The chemical shifts of the three Trp indole Hα resonances are identical in the wild-type and variant complexes (10–11 ppm region in spectra ii and iii in Fig. 9A); these include Trp-98 whose indole NH resonance is in each case shifted downfield at the DNA interface. (c) The methyl...
resonances of Leu-101 (position 46 in the HMG box) are shifted upfield in both the free and bound domains due to aromatic ring currents in the core of the major wing. These secondary shifts are essentially identical in the wild-type and V60L complexes (spectra ii and iii in Fig. 9A); the corresponding inter-residue NOEs are observed between the methyl resonances of Leu-101 and aromatic resonances of Trp-98 (spectra iv and v in Fig. 9B). The aromatic regions of TOCSY spectra of the wild-type and variant complexes are shown in supplemental Fig. S9.

Despite the global change in the width of the V60L end-to-end distance distribution relative to that of the wild-type complex (above), local $^1$H NMR features of the bound DNA sites are
remarkable for their similarity. Prominent intermolecular NOEs were observed in each case between the imino protons of T8 and T9 at the classical site of cantilever insertion and the δ-CH3 resonances of Ile-68 (supplemental Fig. S10). Furthermore, NOESY spectra of Watson-Crick imino protons (whose cross-peaks reflect stacking of successive base pairs) exhibit similar patterns of NOE intensity or attenuation (spectra vi and vii in Fig. 9C). In each case no NOE is observed between the N3H protons of T8 and T8 across the site of Ile-68 insertion (dashed lines in Fig. 9C), and in each case the T7–T8 NOE is markedly attenuated (boxes). Subtle differences in relative cross-peak intensity were likewise maintained in the variant complex (illustrated by the G10–T11 and T11–T12 steps in supplemental Fig. S11). The extent of solvent exchange by the DNA imino protons and its site-specific damping by protein binding (as probed by NOESY exchange cross-peaks (49)) was also found to be similar in the wild-type and V60L complexes (supplemental Fig. S12). Together, native-like 1H NMR features of the Watson-Crick imino protons, including NOE patterns, demonstrate that the increased tr-FRET-derived width of the V60L distance distribution is not a consequence of either (a) breakage of hydrogen bonds within base pairs, (b) perturbed intercalation by the cantilever side chain, or (c) disruption of base stacking at other DNA sites.

**V60L and V60A SRY Retain Partial Gene Regulatory Activity—**

A physiological model of SRY-directed transcriptional activation has been developed based on a real time PCR (RT-quantitative-rtPCR) assay for transcriptional activation of the endogenous Sox9 genes in a male rat embryonic gonadal cell line (CH34 cells; kindly provided by Prof. P. K. Donahoe, Massachusetts General Hospital, Boston (14, 40)) following transient transfection of human SRY or variant SRY alleles (see “Experimental Procedures”). Design of this assay recapitulates the role of the SRY-SOX9 pathway in the overall program of testis determination (Fig. 10A). Transient transfection of wild-type SRY (efficiency ~17% based on control co-transfection of GFP in the same plasmid backbone as that of the SRY expression plasmid) was found to activate transcription of Sox9 and Müllerian inhibiting substance (Mis; also designated Amh) but not of a control housekeeping gene (encoding β-actin) or other Sox genes (Sox2, -3, -4, -5, and -10) (51). The CH34 cell culture model probes a key regulatory step in testicular differentiation (38).

Transient transfection of wild-type SRY (black bar in Fig. 10B) activates Sox9 by ~8-fold relative to an empty vector control (open bar), whereas such activation is attenuated on transient transfection of V60L or V60A SRY (green and red bars in Fig. 10B). Whereas these minor wing variants retained almost half of the wild-type activity in CH34 cells, *de novo* mutations I68T and R133W were found to be without significant activity in this functional assay (gray bars in Fig. 10B). Application of Student’s *t* test indicated that the difference observed between the variants at position 60 and wild type is in each case significant with a *p* value of less than 0.05 (eight replicates); the difference between the position 60 variants and each *de novo* control is also significant at this level. The relative transcriptional activities of V60L and V60A SRY in the CH34 cell culture model were robust to serial reduction of the amount of expression plasmid employed in transient transfection (supplemental Fig. S13).

**DISCUSSION**

Conservation of the SRY HMG box as a sequence-specific DNA-bending motif (52) suggests that the protein functions as an architectural transcription factor (2, 53). Indeed, the marked effects of SRY on DNA structure (54) are proposed to contribute to assembly of sex-specific transcriptional pre-initiation complexes (15, 16, 55, 56). This proposal is supported by clinical observations that almost all mutations causing sex reversal occur in the HMG box and impair specific DNA binding (53, 57). The central contribution of DNA bending to developmental gene regulation is supported by studies of chimeric murine Sry (*mSry*) transgenes in XX mice (58, 59).

Testicular differentiation provides a model for a genetic program of organogenesis. Although direct targets of SRY are incompletely characterized, expression profiling of murine pre-Sertoli gonadal cells has defined an extensive network of candidate downstream genes, including Sox9, Desert hedgehog (*Dhh*), fibroblast growth factor 9 (*Fgf9*), and prostaglandin D synthase (*Ptgds*) (60). The importance of this network is demonstrated by human genetics (61–64) and studies of knock-out mice (65–67). A growing body of evidence suggests that SRY
Inherited Human Sex Reversal

directly activates SOX9 at the transcriptional level and that such activation is central to testicular differentiation (4, 5). Binding sites for SRY have been identified in a Sox9 enhancer element by chromatin immunoprecipitation and shown to be necessary for its tissue- and stage-specific transcriptional regulation (68). The regulatory linkage between SRY and SOX9 provides a unified framework for the understanding of diverse human genetic syndromes (14, 69).
Inherited Human Sex Reversal—This study has addressed an anomaly in human genetics that has challenged the above paradigm (17), i.e. an inherited mutation reported to block detectable specific DNA binding activity (11). The mutation (V60L) is of dual structural and biological interest. On the one hand, the native side chain was observed in NMR structures to pack in the minor wing of the HMG box (i.e. between the N-terminal β-strand and C-terminal α-helix (28, 33)) and so participate in its induced fit on specific DNA binding (26). Substitutions at this site probe the bidirectional coupling between protein folding and DNA bending. On the other hand, the biological implications of impaired coupling seemed enigmatic as the reported complete loss of specific DNA binding activity posed a father-daughter paradox. Although commonly reported among de novo mutations, an inactive SRY variant, within a DNA-centered transcriptional paradigm (17), could not account for the sex of the fertile father.

A molecular understanding of inherited sex reversal could in principle have invoked cryptic genetic mosaicism in the father (such that his testes and germ line contained a subset of cells containing a wild-type Y chromosome (70)) or potential non-DNA-dependent mechanisms of gene regulation by SRY (such as regulation of pre-mRNA splicings (71)). The importance of DNA-dependent mechanisms of gene regulation by SRY (such that his testes and germ line contained a subset of cells containing a wild-type Y chromosome (70)) or potential non-DNA-dependent mechanisms of gene regulation by SRY (such as regulation of pre-mRNA splicings (71)). The importance of DNA-dependent mechanisms of gene regulation by SRY (such that his testes and germ line contained a subset of cells containing a wild-type Y chromosome (70)) or potential non-DNA-dependent mechanisms of gene regulation by SRY (such as regulation of pre-mRNA splicings (71)).

FIGURE 9. 1H NMR studies of wild-type and V60L SRY-DNA complexes. A, binding of wild-type SRY domain to 15-bp consensus DNA site (5′-GCGGGTATTGTTGAC-3′) and complement; core target site in boldface); Left, downfield DNA imino and protein Trp N1 resonances. Spectrum i, free DNA imino resonances (12–14 ppm) at 15 °C; assignments are as indicated. Broadening of G1 and G15 (not well resolved) terminal resonances is due to DNA fraying. Spectra ii and iii, marked changes in imino chemical shifts in specific 1:1 complexes with wild-type domain (spectrum ii) and V60L domain (spectrum iii). Although changes are similar in the two complexes, a smaller complexation shifts is observed at base pair 10 in variant complex (Δδ 0.1 ppm; broken vertical line near 12 ppm). This base pair (position 5′-GCGGGTATTGTTGAC-3′ in top strand; underlined) is closest to residue 60 (see Fig. 38). N1 resonances T8 and T9 at site of ile-68 insertion are indicated at top by filled circles; the former is shifted upfield into GC region of spectrum. Trp indole NH resonances in complexes (10–11 ppm) exhibit essentially identical chemical shifts and line widths. Asterisk indicates attenuation of protected para-OH resonance of Tyr-74 in variant complex, which is broadened in the variant complex. Right, upfield region of 1H NMR spectra containing aliphatic resonances from the protein and selected DNA resonances (thymidine methyl groups and 2′-deoxyribose protons). Upfield methyl resonances of cantilever side chain ile-68 and major wing core resonances of Leu-101 are essentially identical in the wild-type and variant spectra (spectra ii and iii). The respective methyl resonances of Val-60 and Leu-60 exhibit marked differences in chemical shift. B, comparison of NOESY spectra of protein-DNA complexes in D2O; wild-type (spectrum iv) and V60L variant (spectrum v) at 25 °C (mixing times 120 ms). Resonances assigned to side chains in the major and minor wings are as indicated. Corresponding cantilever-specific NOEs were observed between the side chain of ile-68 and the H2 proton of A9 in the minor groove (as upper left of each box). The mixing times were 200 ms. C, comparison of NOESY spectra of imino DNA resonances in specific protein-DNA complexes in H2O: wild-type (spectrum vi) and V5L variant (spectrum vii) at 25 °C. Boxes indicate position of weak NOE between the N2H resonances of T7 and T8. Dashed line indicates resonance position of T9 adjoining cantilever side chain ile-68. Spectra were in each case acquired at 600 MHz.
sion or the influence of autosomal polymorphisms influenced the divergent developmental fates of the father and daughter. Because V60A led to ovoestes and not pure gonadal dysgenesis (36), we predict that genetic analysis of the father, if he could be identified and would consent to genetic studies, would reveal the same variant Y chromosome and so document a pattern of inheritance similar to that of V60L.

2-Fold changes in the biochemical properties of proteins may seem subtle in vitro but can have profound biological consequences. A fundamental example in genetics is provided by chromosome counting in diverse phyla; the ratio of X chromosomes to autosomes in Dipteran insects, for example, regulates a sex-determining hierarchy and RNA-splicing cascade leading to divergent developmental outcomes (male or female) (73).

This study has provided a compelling example of a 2-fold change in transcriptional activation associated with an ambiguity in the outcome of a developmental switch, i.e. the regulation of gonadal differentiation by SRY-initiated expression of Sox9.

FIGURE 10. SRY-regulated testicular differentiation and transcriptional activation of Sox9. A, central dogma of SRY-SOX9-driven male-specific gene regulatory network. Genetic interactions during male sex determination: the SRY-SOX9 pathway (red box) is essential for testis formation. SRY binds to upstream enhancer elements of Sox9 (5). These factors each act within a supporting cell lineage to direct Sertoli cell differentiation and signal to the other lineages (such as the Leydig cell pathway; vertical dotted line in center). Shown in schematic form are repressive pathways (.), oblique dashed and solid lines), which ensure that the central SRY pathway (horizontal at top) does not promote growth and differentiation of female primordial. This panel is adapted from Ref. 59 (with permission of the authors) with updated information (6). The abbreviations used are as follows: BMP8b, bone morphogenetic protein 8B; DAX1, dosage-sensitive sex reversal, adrenal hypoplasia critical region, on chromosome X, gene 1; DMRT1, double-sex and mab-3 related transcription factor 1; FGF9, fibroblast growth factor 9; GATA4, GATA-binding protein 4; LHX9, LIM homeobox 9; LIM1, homeobox protein Lhx 1; MIS, Müllerian inhibiting substance (also designated AMH, anti-Müllerian hormone); PTGDS, prostaglandin D2 synthase; SF1, steroidogenic factor 1; WT1, Wilms tumor 1; Wnt, wingless-type murine mammary tumor virus integration site family, member 4. B, transient-transfection transcriptional assays in a rat embryogenic gonadal cell line: left to right, empty-vector control (open bar), wild-type (wt) SRY (black), V60L (green), V60A (red), and de novo controls R133W and I68T (gray). Histogram provides results of real time RT-PCR assays in CH34 cells; error bars indicate standard deviation of eight replicates. Results indicate that V60L and V60A SRY exhibit partial activity, whereas de novo mutations I68T and R133W essentially abolish Sox9 transcriptional activation (40). Statistical significance is indicated by asterisks as follows: * indicates p < 0.05; ** indicates p < 0.01.
Inherited Human Sex Reversal

Coupled Landscape of Protein-DNA Bending—Mutations V60L and V60A preserve the hydrophobic character of the wild-type side chain but differ in size and shape. Although in a pre-folded protein these substitutions would be expected to perturb the structure and stability of the minor wing, we have found that the native and variant domains exhibit similar CD, fluorescence, and 1H NMR spectra. Furthermore, the domains exhibit similar thermal and thermodynamic stabilities. These results support a model in which the minor wing of the SRY HMG box is disordered and folds on specific DNA binding (26).

Although the structure of the free SRY HMG box has not been well characterized, observation of limited chemical-shift dispersion and a paucity of minor wing-specific NOEs (relative to the DNA complex) suggest that its tertiary structure is stabilized on specific DNA binding (Fig. 1C) (26). Additional evidence for the absence of stable tertiary packing within the unbound minor wing is provided by the pH dependence of the His-120 1H NMR aromatic chemical shifts. Stable packing of the mini-core (as in non-specific HMG boxes) would be expected to be disrupted by protonation of the imidazole ring, leading to a shift in its pKₐ below 6.0. The observed pKₐ was 6.3, however, typical of histidines in unstructured peptides. Furthermore, the stability of the free domain, as inferred by guanidine titration, is similar in the pH range 2.0–7.4 at 25 °C (supplemental Table S2).

We speculate that mutual induced fit (DNA-directed protein folding and protein-directed DNA bending) provides an economical mechanism whereby a single protein may specify a broad range of DNA bend angles depending on subtle changes in DNA target sequence (Table 1). It is not known whether or how such a code may contribute to the gene regulatory properties of SRY or other transcription factors containing sequence-specific HMG boxes. Nonetheless, the kinetic effects of V60L and V60A are likely to reflect perturbation of mutual induced fit. In structures of SRY-DNA, SRY-POU-DNA, and SOX-DNA complexes (27–29, 54), packing of the wild-type side chain within a co-conserved aromatic box appears to buttress the DNA-binding surface of the minor wing (54) and may also orient its C-terminal basic tail (28). Because truncation of the tail has likewise been shown to reduce the kinetic lifetime of the protein-DNA complex (34), we envisage that steric perturbation of aromatic box (by V60L) or introduction of a packing defect (by V60A) would in turn perturb the adjoining DNA-binding surface to reduce its specific DNA affinity and promote its dissociation. It is also possible that substitutions at this site could alter the topography of the back surface of the HMG box and hence its recognition by another factor in an SRY-dependent transcriptional preinitiation complex (supplemental Fig. S14).

It is not apparent why Leu-60 gives rise to nonoptimal induced fit as molecular models of the variant HMG box may...
Inherited Human Sex Reversal

readily be constructed and, following energy minimization, yield empirical energies similar to or lower than that of the native minor wing. Steric clash by the variant side chain within the aromatic box can be relieved with only small adjustments in side-chain dihedral angles leading to changes in the positions of the aromatic side chains of less than 0.25 Å (data not shown). The modesty of the decrease in affinity (2-fold as determined by FRET-monitored titration and less than 2-fold as monitored by GMSA) reflect such straightforward structural accommodation. Overall changes in affinity may also reflect entrophy-enthalpy compensation (88); the entropic penalty imposed by DNA-directed protein folding and protein-directed DNA bending would presumably be mitigated in a variant complex characterized by a less well ordered minor wing and protein-DNA interface. Models based on energy minimization in the absence of solvent, DNA, and counterions may suggest the plausibility of possible structures but are unlikely to rationalize small changes in binding constants and activation barriers as observed here.

The present spectroscopic studies provide an example of how a single amino acid substitution in the core of a protein can cause transmitted changes at a protein-DNA interface. The core SRY target site (5′-ATTGTTTT-3′ and complement) contains a major wing-binding element (5′-ATTGTTTT-3′ and complement; boldface) and minor wing-binding element (5′-ATTGTTTT-3′ and complement; underlined). Each undergoes sharp DNA bending (54). Comparison of the downfield 1H NMR spectra of DNA imino protons (which participate in Watson-Crick hydrogen bonding) revealed only subtle differences between wild-type and V60L complexes, and these are at the base pairs closest to the site of mutation. Val-60 and Leu-60 otherwise exhibit analogous inter-residue NOEs between their methyl resonances and neighboring aromatic side chains in the minor wing (His-120, Tyr-124, and Thr-127). Changes in such aromatic chemical shifts (Δδ −0.12 and −0.05, respectively, for ortho and meta resonances of Tyr-124 and Δδ −0.08 and −0.01 for Tyr-127; see Fig. 9B) nonetheless suggest that the aromatic box reorganizes in response to the steric challenge of the V60L substitution. These negative values of Δδ (wild-type minus mutant) are in the direction of random-coil values. Chemical shifts and NOEs within the major wing and at the site of partial intercalation by cantilever side chain Ile-68 are by contrast essentially identical in the wild-type and variant complexes. These observations suggest that local perturbations are subtle and confined to the 3′ DNA subsite contacted by the minor wing. Similarity of local 1H NMR parameters is in accordance with similarity in Raman-active vibrational frequencies in the bound DNA sites (42).

Despite the restricted nature of 1H NMR changes in local parameters, a global change in the dynamics of DNA bending was elucidated by subnanosecond tr-FRET. Experimental design focused on end-to-end distances in the DNA as probed by interactions between a 5′-fluorescence donor (FAM; upper strand) and 5′-fluorescence acceptor (TAMRA; lower strand). The flexible linkages between these probes and the DNA duplex simplified tr-FRET interpretation in terms of distance distributions (89) but also led to a broad distribution of implied donor-acceptor distances even in the free DNA. Because the length of the DNA duplex (15 bp) is far below the persistence length of DNA, we ascribed this base-line width (full width at half-maximal value of 14.5±2.4 Å; Table 2) to local reorientation of the probes. The increased width of the wild-type protein-DNA complex (full width at half-maximal value of 19.5±0.9 Å) may in part reflect geometric factors (15) and in part a range of populated DNA bend angles. Because the mean bend angle of the variant complex is similar to that of the wild-type complex, we ascribe the further broadening in the variant complex (full width at half-maximal value of 24.3±0.6 Å) exclusively to excursions in DNA bend angle. It would be of future interest to develop a computational model of such ensembles from which estimates of bend angle distributions could be extracted.

Global differences in the width (but not the mean) of the end-to-end distance distribution in the variant complex, as extracted from tr-FRET data, would be difficult to validate by conventional protocols for calculation of NMR-derived solution structures. Such protocols typically utilize large numbers of restraints to generate models simultaneously consistent with each restraint. The resulting ensemble may be precise but not a meaningful representation of the physical ensemble (90). A promising route of exploration may exploit side-by-side comparison of residual dipolar couplings (15N-1H in the bound protein and 31P-1H in the bound DNA) in the wild-type and variant complexes in an anisotropic medium (27); such studies might reveal inequivalent averaging of such couplings at corresponding sites because of global differences in the dynamics of DNA bending (91, 92). Such probes may be combined with molecular dynamics simulations of the bent protein-DNA complexes to illuminate the structural origins of how “molten” DNA bending may lead to long range fluctuations in end-to-end distances. Such combined experimental and computational approaches might establish rules relating sequence conservation (or divergence) among mammalian HMG boxes to the functional role of protein-directed DNA bending in gene regulation and development (43).

Dynamic perturbation of the V60L SRY-DNA complex is formally unrelated to its accelerated rates of association and dissociation. Yet we speculate that these two sets of findings have a common physical origin, frustration in the coupling between DNA-directed protein folding and protein-directed DNA bending. Insertion of the side chain of Val-60 into the aromatic box formed by residues His-120, Tyr-124, and Tyr-127 appears pivotal to the stabilization of the overlying protein-DNA interface. Although molecular modeling suggests that Ile could readily be accommodated without perturbation of the box (in accordance with its frequent occurrence among SOX domains; Fig. 2B), subtle steric frustration (V60L) or introduction of a packing defect (V60A) may lead to no unique optimum to the problem of bidirectional induced fit. In this scheme interconversion among multiple near-isoenenergetic bound-state conformations would reduce in tandem kinetic barriers to asso-

15 Monte Carlo simulations of donor and acceptor orientations flexibly linked to straight or bent 15-bp DNA sites indicate that at bend angles similar to that of the SRY-DNA complex (50–80°) geometric factors lead to broadening of the distribution of predicted donor-acceptor distances. (A. R. Srinivasan and W. Olson, unpublished results (100) (supplemental Fig. 55).)
SRY-DNA association is slower than would be predicted by the diffusion limit, presumably due to a choreography of conformational changes both in the HMG box and bent DNA site. Mutations V60L and V60A accelerate inferred rates of association; as a seeming paradox, perturbation of the minor wing enhances the apparent efficiency of a conformational search. In a related context, Sanchez et al. (93) have developed a theoretical formalism for visualizing frustration based on coupled energy landscapes. By analogy to theories of nascent protein folding, the essential idea invokes a minimally frustrated free energy landscape for sequence recognition. This view encompasses the range of kinetic mechanisms (multiple kinetic intermediates defining parallel routes of association and dissociation) underlying the deceptive simplicity of two-state models (free and bound). The coupled landscapes represent native and non-native ensembles of complexes (Fig. 8D). Non-native protein-DNA contacts in the (Fig. 8D, funnel at left) lead to its ruggedness. Application of this model to variant SRY-DNA complexes highlights possible global changes to the following: (a) the dimensions of a funnel along both entropic and enthalpic dimensions (horizontal and vertical axes in Fig. 8E); (b) the transition states between coupled funnels and to the unbound state, and (c) the breadth or ruggedness of the native-like ground state. Whereas Sanchez et al. (93) envisaged that frustration is widespread even in wild-type protein-DNA complexes, our studies of V60L SRY suggest that introduction of a single newly frustrated site can give rise to nonlocal changes in DNA bending.

It is not known whether the subtle changes in the DNA bending properties of V60L SRY, as indicated by tr-FRET-based distance distribution analysis, may influence the transcriptional regulatory activity of the variant protein. Two perspectives are pertinent. On the one hand, electrophoretic studies suggest that differences between the wild-type and variant DNA complexes may be more accentuated at nonconsensus target sites (Table 1 and supplemental Fig. S15). We speculate that changes at the 3′-end of the target site (boldface in 5′-ATTGTT-3′) more markedly perturb folding of the variant minor wings than the wild type. Because SRY-responsive genomic elements may contain both consensus and nonconsensus DNA sites, it is possible that biophysical studies of consensus complexes may underestimate the range of DNA bending defects present in vivo. On the other hand, whereas a de novo mutation in SRY associated with gonadal dysgenesis (M64I; HMG box position 9) was originally reported to impair specific DNA binding but not DNA binding (94), re-investigation of this mutation in the context of full-length SRY demonstrated that native DNA bending was restored by the basic tail (48).

Further insight may be provided from an evolutionary perspective. Although the SRY-DNA bend angle is conserved among primates (as inferred by PGE (95)), the extent and precision of SRY-directed DNA bending varies among other mammals, and the extent of variation encompasses the broadened end-to-end distance distribution characteristic of V60L SRY (43). Studies of this and other SRY variants in CH34 cells may provide an opportunity to dissect this and other potential structural determinants of activity in a model of the SRY-FOX9 transcriptional axis.

Concluding Remarks—Testis determination in mammals is regulated by SRY, a specific DNA-bending protein encoded by the Y chromosome. Whereas de novo mutations generally impair DNA binding or bending leading to pure gonadal dysgenesis, a seeming paradox was posed by certain inherited alleles; complete loss of specific DNA binding was reported in a family bearing the variant V60L Y chromosome despite divergent developmental outcomes (fertile father or sterile daughter) (11). Although the reported inactivity of V60L SRY would have rationalized nascent ovarian differentiation with subsequent degeneration (because of loss of germ cells), how could the islands of testicular differentiation be understood, and more pointedly, why was the father male? This question was made pointed by the further association between features of true hermaphroditism and related mutation V60A (36). These substitutions have not been observed in surveys of neutral polymorphisms.

This study has demonstrated that the specific DNA binding affinities and gene regulatory function of V60L SRY and the related variant V60A are intermediate between the base-line properties of wild-type SRY and the impaired properties of de novo variants. The compatibility of V60L SRY (and presumably of V60A SRY) with male or female somatic phenotypes, both between generations and within the same gonad, may reflect stochastic effects in an otherwise canaliized developmental pathway and/or the interplay of unlinked polymorphisms (X-linked or autosomal) separating the parental genotypes. It will be of future interest to explore what evolutionary constraints have enjoined the invariance of Val-60 (position 5 of the HMG box) among mammalian SRY alleles. The enigma of inherited mutations in SRY provides an opportunity to characterize a genetic switch poised at the threshold of biological activity.

Acknowledgments—We thank Prof. P. K. Donahoe (Massachusetts General Hospital, Boston) for providing rodent cell line CH34; V. Utah and D. Amir for assistance with FRET studies; A. R. Srinivasan and W. K. Olson for Monte Carlo simulations and helpful discussion; C. Haqq and E. Ukiyama for advice regarding cell culture studies; A. Stern and J. C. Hoch for aromatic ring current shifts; W. Jia for preparation of NMR figures; H. L. Woodcock for advice regarding molecular modeling and molecular mechanics calculations; and G. Chan, H.-J. Chen, C.-Y. King, B. Li, and F. Tao for participation in the early stages of this work.

REFERENCES
1. Sinclair, A. H., Berta, P., Palmer, M. S., Hawkins, J. R., Griffiths, B. L., Smith, M. J., Foster, J. W., Frischauf, A. M., Lovell-Badge, R., and Goodfellow, P. N. (1990) Nature 346, 240–244
2. Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Münsterberg, A., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1990) Nature 346, 245–250
3. Ner, S. S. (1992) Curr. Biol. 2, 208–210
4. Sekido, R., Bar, I., Narváez, V., Penny, G., and Lovell-Badge, R. (2004) Dev. Biol. 274, 271–279
5. Sekido, R., and Lovell-Badge, R. (2008) Nature 453, 930–934
