Structure-Function Relationships in Human Testis-determining Factor SRY

AN AROMATIC BUTTRESS UNDERLIES THE SPECIFIC DNA-BENDING SURFACE OF A HIGH MOBILITY GROUP (HMG) BOX**

Received for publication, July 16, 2014, and in revised form, September 23, 2014. Published, JBC Papers in Press, September 25, 2014, DOI 10.1074/jbc.M114.597526

Joseph D. Racca, Yen-Shan Chen, James D. Maloy, Nalinda Wickramasinghe, Nelson B. Phillips, and Michael A. Weiss

From the Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106

Background: Testis-determining factor SRY provides a model of bent protein-DNA complex.

Results: Mutation of an invariant Trp perturbs multiple biochemical, cellular, and transcriptional activities.

Conclusion: Folding and function of a sequence-specific HMG box require a core “aromatic buttress.”

Significance: Mutations in SRY causing human sex reversal probe the architecture and evolution of a DNA-bending motif.

Male development in therian mammals is initiated by Sry, a Y-encoded architectural transcription factor. Mutations in SRY cause 46 XY gonadal dysgenesis with female somatic phenotype (Swyer syndrome) and confer a high risk of malignancy (gonadoblastoma). Such mutations cluster in the SRY high mobility group (HMG) box, a conserved motif of specific DNA binding and bending. To explore structure-function relationships, we constructed all possible substitutions at a site of clinical mutation (W70L). Our studies thus focused on a core aromatic residue (position 15 of the consensus HMG box) that is invariant among SRY-related HMG box transcription factors (the SOX family) and conserved as aromatic (Phe or Tyr) among other sequence-specific boxes. In a yeast one-hybrid system sensitive to specific SRY-DNA binding, the variant domains exhibited reduced (Phe and Tyr) or absent activity (the remaining 17 substitutions). Representative nonpolar variants with partial or absent activity (Tyr, Phe, Leu, and Ala in order of decreasing side-chain volume) were chosen for study in vitro and in mammalian cell culture. The clinical mutation (Leu) was found to markedly impair multiple biochemical and cellular activities as respectively probed through the following: (i) in vitro assays of specific DNA binding and protein stability, and (ii) cell culture-based assays of proteosomal degradation, nuclear import, enhancer DNA occupancy, and SRY-dependent transcriptional activation. Surprisingly, however, DNA bending is robust to this or the related Ala substitution that profoundly impairs box stability. Together, our findings demonstrate that the folding, trafficking, and gene-regulatory function of SRY requires an invariant aromatic “buttress” beneath its specific DNA-bending surface.

2 The abbreviations used are: TF, transcription factor; 6-FAM, 6-carboxyfluorescein; AMH/MIS, anti-Müllerian hormone/Müllerian inhibiting substance; C/EBP, CCAAT enhancer binding protein; co-IP, co-immunoprecipitation; GRN, gene regulatory network; HMG, high mobility group; HSQC, heteronuclear single-quantum coherence; NLS, nuclear localization signal; qPCR, quantitative real time PCR; SOX, SRY-related HMG box; TESCO, core SRY-binding sites within the TES element of Sox9; TAMRA, tetramethylrhodamine; DSD, disorders of sexual differentiation; Y1H, yeast one-hybrid. Amino acids are designated by standard three- and one-letter codes.

3 In the original clinical report, the patient was described as phenotypically female with unilateral gonadoblastoma, a potential precursor lesion to a second class germ cell tumor, thought to be caused by a 50% decrease in nuclear accumulation of SRY due to defective nuclear localization.

© 2014 by The American Society for Biochemistry and Molecular Biology, Inc. Published in the U.S.A.
ing proteins (14). The discovery of Sry in therian mammals enabled identification of metazoan TFs containing Sry-related HMG boxes (designated Sox) (15). Broadly involved in the regulation of development (16), this family is of interest in relation to organogenesis (17), stem-cell biology (18), patterning of the brain (19), diverse human birth defects (20), and mechanisms of transcriptional deregulation in cancer (21, 22). Structure-function relationships in human SRY are likely to generalize to the Sox family and thus provide insight into the evolution of an ancestral eukaryotic DNA-bending motif.

The HMG box contains major and minor wings, which together define an angular DNA-binding and DNA-bending surface (13, 23–27). Structures of specific DNA complexes containing the human SRY box or homologous SOX domains have been determined by multidimensional NMR spectroscopy (25, 28) and x-ray crystallography (29). In such complexes, DNA bending is associated with partial intercalation of a protein side chain (designated the “cantilever”) (30–32) at the crux of this angular surface (33). The cantilever (Ile-68 in human SRY; consensus position 13 of the HMG box as defined by Werner et al. (25)) is part of a conserved “hydrophobic wedge” of major wing side chains (Met-64, Phe-67, and Trp-98; respective HMG box positions 9, 12, and 43) that insert into a widened DNA minor groove (28). Additional determinants of DNA bending are provided by the minor wing and its basic C-terminal tail (34). These DNA-binding and -bending elements are mutational hot spots for DSD-associated mutations in SRY and SOX9 (35). Corresponding mutations in other SOX factors cause diverse developmental abnormalities reflecting their individual biological roles (16, 22).
Aromatic Buttress in an HMG Box

This study focused on the in vitro properties of the variant SRY domains (the HMG box) in relation to the following: (a) its wild-type (WT) structure (25, 28) and (b) the biological activities of the intact protein (38). The apparent lack of organized structure within the non-box segments of SRY (12), their general lack of conservation (11), and the similar DNA-binding and DNA-bending properties of intact protein and the isolated domain (39, 40) support the logic of such a modular approach.4

Our experimental design thus employed systematic mutagenesis in a yeast one-hybrid (Y1H) system (41), followed by complementary investigation of representative variants in vitro and in mammalian cell culture. Because box position 15 is part of its hydrophobic core, priority was given to nonpolar and aromatic substitutions with particular attention to DSD-associated mutation W70L. These studies culminated in analysis of transcriptional regulatory activities in an embryonic rat pre-Sertoli cell line (42), previously characterized as a model of the bipotential gonadal ridge at a stage of development just prior to its morphological differentiation (38). This model enables dissection of distinct cellular processes (including TF expression, proteasome-mediated turnover, and nuclear import) underlying the operation of a sex-specific GRN. Because SRY functions as a nuclear protein (43–46), effects of the substitutions were evaluated with respect to the machinery of nuclear import (candidate mediators Exportin-4 (Exp4 (47)) and calmodulin (CaM (48))).5

Remarkably, our results have demonstrated that Trp-70 contributes to multiple biophysical properties and diverse biochemical activities of SRY. Each of the 19 possible amino acid substitutions at this site impairs, to a greater or lesser extent, the function of the SRY DNA-binding domain in our Y1H system. Studies of representative variants demonstrated that the core tryptophan both contributes to the folding of the HMG box and enables critical cellular processes pertinent to biological function. Whereas significant residual function was observed among aromatic variants (W70F and W70Y) in accordance with sequence patterns in the HMG box superfamily (36), the clinical variant W70L and related aliphatic variant W70A were associated with pleiotropic molecular and cellular

4 Structural and biophysical studies of SRY and homologous SOX TFs by this laboratory (31, 41, 90) and other laboratories (28, 91, 92) have been restricted to their respective HMG box domains. In our hands, the intact human SRY readily aggregates in vitro leading to progressive loss of specific DNA binding activity (40). Such loss may underlie the higher sequence- or structure-specific DNA-binding affinities observed by Lipardi and co-workers (39) in studies of the SRY HMG box relative to the intact protein. Such focus on isolated DNA-binding domains is broadly representative of an extensive literature concerning modular TFs containing conserved motifs (93).

5 Exp4 is a member of a class of proteins that mediate the export or import of cargo proteins (such as SRY) from the nucleus to the cytoplasm or from the cytoplasm to the nucleus. Although Exp4 was originally designated as an "exportin" in relation to its Ran-dependent role in the nuclear export of eukaryotic translational initiation factor 5A (94), Demmers and co-workers (47) have shown that Exp4 mediates nuclear import of SRY via interactions with the N-terminal segment of the HMG box (NLS residues 59–77 in intact human SRY); an homologous interaction contributes to nuclear import of SOX9 (69). Nuclear exports of SRY and SOX9 are instead mediated by an interaction of CRM1 with a classical nuclear export signal (38, 95). The HMG boxes of SRY and other SOX factors also contain a C-terminal basic NLS (43, 45) recognized by importin-β (96, 97).
defects that together block SRY-directed transcriptional activation of Sox9. Although the ability of the SRY variants to direct sharp DNA bending in vitro was in each case retained, altered core packing led to mild (W70F and W70Y), moderate (W70A), or severe (W70L) perturbations to the lifetime of the bent protein-DNA complex.

Deciphering the informational content of protein sequences can provide key insight into their function and evolution. Among the canonical families of eukaryotic DNA-binding motifs, aromatic side chains often provide conserved structural supports. The respective DNA recognition α-helices of homeodomains and zinc fingers, for example, are anchored by such core residues (49, 50). The eukaryotic superfamily of HMG boxes is likewise remarkable for the broad conservation of an aromatic buttress beneath an angular DNA-binding surface (13, 23–27). Our findings, exploiting the growing human genome database, illuminate how this buttress contributes to the biophysical properties and biological activities of an ancestral DNA-bending motif. DSD-associated mutation W70L in SRY thus represents a molecular “perfect storm”6 (51) that impairs gonadogenesis leading to a human malignancy (9).

**EXPERIMENTAL PROCEDURES**

**Y1H Screening**—Reporter strains of *Saccharomyces cerevisiae* (derived from strain YM4271) were as described previously (41). Expression of an SRY HMG box fusion protein (Fig. 3A, gray oval) was directed by plasmid pGAD-T7 (Clontech). Transformants were selected on minimal medium lacking Leu and uracil. Colonies were isolated and grown in liquid medium under the same selection. The extent of SRY-dependent reporter activation (β-galactosidase (β-gal)) was determined as follows: (i) a qualitative assay on selective plates containing a chromogenic substrate (80 μg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) to which 5–7-μl aliquots of overnight broth were spotted), and (ii) a quantitative assay in solution using ortho-nitrophenyl-β-galactoside as described by the vendor (Clontech).

**Protein Purification**—WT and variant SRY domains of human SRY were expressed in *Escherichia coli* strain T1LY (New England Biolabs, Inc., Ipswich, MA) and purified as described (52). Purity was determined in each case to be >98% by SDS-PAGE. Results of matrix-assisted laser-desorption ionization time-of-flight mass spectrometry were in agreement with the expected values.

**Circular Dichroism**—Far- and near-ultraviolet (UV) CD spectra were obtained at 4 and 37 °C in a 1-mm path length quartz cuvette using an Aviv spectropolarimeter equipped with a titrating unit (Aviv Biomedical, Lakewood, NJ). The domains were made 25 μM in 140 mM KCl and 10 mM potassium phosphate (pH 7.4; “standard buffer”). Thermal unfolding was monitored at a helix-sensitive wavelength of 222 nm. CD difference spectra were calculated as the buffer-corrected difference between the observed spectrum of a protein-DNA complex and the sum of the spectra of the free protein and free DNA site.

**Tryptophan Fluorescence Spectroscopy**—Intrinsic tryptophan (Trp) fluorescence of free SRY and DNA complex domains was determined in standard buffer at 15 °C at a concentration of 5 μM. Emission spectra were acquired from 300 to 500 nm following excitation at 295 nm. In studies of protein-DNA complexes, the specific DNA site was included in the buffer for baseline correction. Extent of an inner filter effect under these conditions was estimated through control studies of the WT domain in the presence or absence of the free deoxynucleotides (dAMP, dCMP, dGMP, and dTMP) with a molar ratio equivalent to that in the

---

The metaphoric use of this term in the biomedical literature (51) arose in relation to the 1997 account of the loss of the New England fishing boat “Andrea Gail” as described by Junger (98).
Aromatic Buttress in an HMG Box

specific DNA site (which contains eight AT bp and seven GC bp) and at a concentration such that absorbance at 295 nm was equal to that of the DNA site at 5 μM. Standard errors in the measurement of integrated fluorescent intensities were estimated in five independent replicates of the spectrum of the WT domain.

Fluorescence Resonance Energy Transfer (FRET)—Protein-directed DNA bending was probed by steady-state FRET as described (34). This protocol employed a 15-bp DNA duplex of sequence 5′-TCCGGTGATTGTTCAG-3′ (“upper strand”) and complement; consensus target site is underlined. Use of a 15-bp DNA site restricted protein binding to the 1:1 high affinity complex (41). To provide a fluorescent donor, the upper strand was extended at its 5′ terminus by 6-carboxyfluorescein (6-FAM); the dye was flexibly linked to the DNA through a hexanyl linker. To provide a corresponding acceptor, the lower strand was extended at its 5′-end by tertamethylrhodamine (TAMRA), also via a hexanyl linker. The labeled DNA strands were purchased from Oligos, Etc., Inc. (Wilsonville, OR). Photophysical control studies verified the mobilities of the probes and excluded non-FRET-related mechanisms of donor quenching (34) for FRET studies, duplex DNA sites were made 3 μM at pH 8.4 in 10 mM potassium phosphate, 10 mM Tris-HCl, 140 mM KCl, 1 mM EDTA, and 1 mM dithiothreitol (“FRET buffer”).

In this formalism, ΔF is the change in donor fluorescence observed on addition of the SRY domain relative to the baseline DNA fluorescence; ΔF₀ is the maximum fluorescence change obtained in a 1:1 protein-DNA complex; K_D is the dissociation constant; D₀ is the concentration of DNA (25 nM); and S is the concentration of SRY domain. In the case of the W70L domain, whose specific DNA-binding affinity was significantly weaker than those of the other variants, the protein concentration range was extended to 3 μM to achieve saturation.

Non-specific DNA Binding—To test whether the binding of variant SRY domains to the consensus FRET DNA target site (above) represented a specific or non-specific mode of DNA binding (31), the FRET-labeled protein-DNA complexes were challenged by addition of equimolar unlabeled DNA target site (15 bp) or a 10-fold molar excess of an unrelated 15-bp DNA site derived from the left operator site of bacteriophage λ (5′-ATCACGCGTGTG-3′ and complement; site O₁₁) (54). The WT SRY domain exhibited no binding to this DNA site as assessed by a gel mobility shift assay.

Stopped-flow Kinetic FRET Assay—Rates of protein-DNA dissociation were measured with an Aviv double-mixing stopped-flow apparatus at a fixed temperature (6, 15, 25, or 37 °C); the instrument contained a thermostatic temperature controller. Fluorescence emission was monitored at 520 nm following excitation at 490 nm using an Aviv ATF 105 spectrofluorometer (34). In brief, a 20-fold molar excess (final stoichiomteries) of unlabeled DNA in FRET buffer was employed to sequester the WT or variant SRY domain on rapid mixing. Estimates of dissociation rate constants (k_off) were obtained by fitting the traces to a mono-exponential equation; values represent the mean and standard error of 4–5 replicates. Control studies of the WT domain indicated that similar time-dependent recovery of donor emission was observed irrespective of unlabeled DNA in the molar excess range 10–50-fold relative to the specific protein-FRET-labeled DNA complex.

NMR Spectroscopy—15N-Labeled WT and variant SRY domains were prepared by growth of an overexpression strain of E. coli (34) in minimal medium containing [15N]ammonium sulfate as sole nitrogen source. The proteins were dissolved in a nitrogen-purged buffer containing 10 mM potassium phosphate buffer (10% D₂O, pH 6.5) and 50 mM NaCl and placed in a 280-μl Shigemi NMR tube. To obtain a spectrum of the unfolded state, the variant 15N-labeled domain was also dissolved in phosphate-buffered saline solution containing 5.4 mM urea (10% D₂O, pH 7.4). Two-dimensional 1H-15N heteronuclear single-quantum coherence (HSQC) spectra were acquired at 25 °C using a BRUKER AVANCE 700 MHz spectrometer.

Mammalian Plasmids—Plasmids expressing full-length human SRY or variants were constructed by polymerase chain reaction (PCR) (41). Following the initiator Met, the cloning site encoded a hemagglutinin (HA) tag in triplicate. In selected constructs, an element encoding a heterologous nuclear localization signal (NLS; sequence PRRRKV as derived from the large T antigen of simian virus 40 (SV40)) was inserted after the HA-related codons (38). Mutations in SRY were introduced using QuikChange™ (Stratagene). Constructions were verified by DNA sequencing.

Mammalian Cell Culture—CH34 cells (kindly provided by Dr. P. K. Donahoe, Massachusetts General Hospital, Boston) (38, 42) were cultured in Dulbecco’s modified Eagle’s medium (DMEM) containing 5% heat-inactivated fetal bovine serum at 37 °C under 5% CO₂. For proteosome-inhibitor studies, transfected cells were maintained for 24 h in serum-free conditions and...
then treated with the proteasome inhibitor MG132 for 6 h followed by 18 h of incubation in 5% serum-containing medium.

**Transient Transfection**—Transfections were performed using FuGENE6 (Hoffmann-La Roche) as described by the vendor. After 24 h in serum-free medium, cells were recovered using fresh DMEM containing 5% heat-inactivated fetal bovine serum. Transfection efficiencies were determined by ratio of green fluorescent protein (GFP)-positive cells to untransfected cells following co-transfection with pCMX-SRY and pCMX-GFP in equal amounts (38). Subcellular localization was visualized by immunostaining 24 h post-transfection following treatment with 0.01% trypsin (Invitrogen) and plating on 12-mm coverslips. SRY expression was monitored by Western blot via its triplicate HA tag (see above).

**Cycloheximide Chase Assay and Western Blot**—24-h post-transfection, cells were split evenly into 6-well plates and treated with cycloheximide to a final concentration of 20 μg/ml in DMEM for the indicated times; cells were then lysed by radioimmunoprecipitation (RIPA) buffer (Hoffmann-La Roche). After protein normalization, cell lysates were subjected to 12% SDS-PAGE and Western blot using anti-HA antiserum (Sigma) at a dilution ratio of 1:5000 with α-tubulin as loading control. Quantification was performed by ImageJ software (National Institutes of Health, website rsbweb.nih).

**Transcriptional Activation Assay**—Following transient transfection (see above), SRY-mediated transcriptional activation of Sox9 was measured in triplicate by quantitative real time Q-rtPCR (qPCR) as described (41). Cellular RNA was extracted using RNeasy (Qiagen, N.V., Hilden, Germany). The transfection protocols were performed with the following: (i) SRY expression plasmid only, which contains 1 μg of WT or variant SRY-encoded plasmid per million cells (“1 g”) and target SRY-encoded plasmid (0.02 μg), with the overall transfected mass of plasmids retained at 1 μg (“50×” conditions). Such dilution provided a control for potential artifacts of TF overexpression (38).

**Immunocytochemistry**—Transfected cells were plated evenly on 12-mm coverslips, fixed with 3% para-formaldehyde in phosphate-buffered saline (PBS at pH 7.4) on ice for 30 min, treated with cold permeability buffer solution (PBS containing 10% goat serum and 1% Triton X-100; Sigma) for 10 min, blocked with 0.1% Tween 20 in cold PBS (Sigma), and incubated overnight at 4 °C with FITC-conjugated anti-HA antibody (diluted to 1:400 ratio; Santa Cruz Biotechnology, Santa Cruz, CA). After washing and 4,6-diamidino-2-phenylindole (DAPI) staining, cells were visualized by fluorescence microscopy. Nuclear localization was evaluated by the ratio of cells exhibiting nuclear HA-tagged SRY to the total number of GFP-positive cells.

**Chromatin Immunoprecipitation (ChIP)**—CH34 cells were transiently transfected with WT or variant SRY constructs, exposed to MG132, and subjected to ChIP. In brief, recovered cells were cross-linked in wells by formaldehyde, collected, and lysed after quenching the cross-linking reaction. Chromatin lysates were sonicated to generate 300–400-bp DNA fragments and immunoprecipitated with anti-HA antisera ( Sigma) coupled with protein A slurry (Santa Cruz Biotechnology) after pre-clearing; a nonspecific antisera (Santa Cruz Biotechnology) served as a control. After reversal of cross-linking at 65 °C overnight, fragments were treated with proteinase K and RNase (Hoffmann-La Roche), followed by extraction with 1:1 phenol chloroform/isoamyl alcohol (CIAA) solution. A high fidelity PCR protocol was provided by the vendor.

**Co-immunoprecipitation (Co-IP) Assays**—CH34 cells expressing HA-tagged SRY variants were treated with MG132 and lysed using complete Lysis-M buffer containing a protease inhibitor mixture as described by the vendor (Hoffmann-La Roche). Two co-IP targets were investigated. (i) In SRY-CaM studies, lysates were precipitated with monoclonal anti-HA-agarose beads ( Sigma). Following 12% SDS-PAGE, Western blots employed anti-CaM antisera (Abcam, Cambridge, MA). Equal CaM loading was verified by Western blot. (ii) In SRY-Exp4 studies, transfected cells were co-transfected with pCMX-FLAG-human Exp4. MG132-treated cell lysates were immunoprecipitated with monoclonal anti-FLAG agarose using the vendor’s protocol ( Sigma). After analysis by 10% SDS-PAGE and electroblocting, hybridization solutions containing horseradish peroxidase (HRP)-conjugated anti-HA antisera (Hoffmann-La Roche) was used to investigate Exp4-bound SRY. Anti-FLAG antisera was used to monitor Exp4 expression; respective antisera against HA tag and α-tubulin (Sigma) provided SRY input and general loading controls.

**RESULTS**

Our study had three parts. We first screened all 19 possible substitutions at SRY residue 70 (box position 15) in a Y1H system in which expression of β-gal was dependent on high affinity binding of the SRY HMG box to consensus DNA target sites (triplicate 5′-ATTGTTT-3′ and complement) (41). We next characterized the biochemical and biophysical properties of four representative variant domains. Because box position 15 (helix α1) contributes to the hydrophobic core of the HMG box, these in vitro studies focused on nonpolar side chains of high α-helical propensity (55) but progressively smaller size as follows: Tyr, Phe, Leu, and Al a.

To relate structure to biological function, our final studies employed an embryonic pre-Sertoli cell line (rodent XY cell line CH34 (33, 42) to probe, following transient transfection, SRY-directed transcriptional activation of the endogenous Sox9 gene (38, 41). Employing qPCR, this assay measured the time-dependent accumulation of mRNAs encoded by Sox9, the principal target gene of murine Sry in vivo (4). This cellular model also permitted analysis of protein turnover, nuclear localization, and occupancy of the testis-specific core enhancer of Sox9 (TESCO (4)) as described (38).

**Y1H Screening Distinguished between Aromatic and Non-aromatic Substitutions**—The yeast system employed the following: (i) plasmid-based expression of WT or variant SRY fusion proteins, and (ii) a chromosomally integrated reporter constructed to express β-gal under the transcriptional control of triplicate SRY target sites (Fig. 3A). An initial screen of all 19 possible amino acid substitutions was performed on X-Gal indicator plates ( Fig. 3B). Whereas the WT fusion protein gave rise to a deep blue colony, the variant colonies were either light blue (Phe and Tyr) or white (the remaining 17 substitutions,
including Leu and Ala). Quantitative assessment of selected β-gal enzymatic activities was undertaken in cell lysates (Fig. 3C). In this system, the low enzymatic activities associated with the W70L and W70A (consensus box position 15) fusion proteins were indistinguishable from that of an empty vector control (em in Fig. 3C). In accordance with visual inspection of the colonies, W70F and W70Y fusion proteins gave rise to enzymatic activities reduced by 50–70% relative to WT.

Substitutions Perturbed Structure and Impair Stability—Structures of the free domains were investigated by far-UV CD (as a probe of secondary structure) and intrinsic Trp fluorescence (a probe of tertiary structure and quenching). The free W70L domain was also probed by heteronuclear two-dimensional 1H-15N NMR spectroscopy.

CD spectra of the free domains exhibited two classes of perturbations (Fig. 4A). Whereas the W70L and W70A domains were largely disordered at 4 °C and essentially without organized secondary structure at 37 °C (Fig. 4B), at 4 °C the W70F and W70Y domains exhibited only a subtle attenuation of α-helix-associated features with decreased ratio of ellipticities at 222 and 208 nm (Fig. 4A); this perturbation was more marked at 37 °C (Fig. 4C). Mean residue ellipticities at 37 °C are given in Table 2. The temperature dependence of mean residue ellipticity at 222 nm indicated partial (W70F and W70Y) or marked (W70A and W70L) loss of thermal stability (Fig. 4, A–C).

Stabilities were further assessed by guanidine-induced unfolding at 4 °C as monitored by CD. Whereas variants W70F and W70Y (Fig. 4D, blue and green lines) were only moderately less stable (leftward shift in denaturation transition) than WT (black), the W70L and W70A domains (red and gray) exhibited noncooperative unfolding transitions characteristic of nascent α-helical peptides or molten globules (Fig. 4D) (56, 57). The latter deviation from classical two-state unfolding was especially prominent at low concentrations of guanidine HCl (0.015 to 0.9 M; inset in Fig. 4D) as no pre-transition baseline could be defined. Application of a two-state model to the aromatic variants yielded estimates of ΔG° (Table 3). Relative to the baseline stability of the WT domain (ΔG° 3.0(±0.1) kcal/mol), decre-

### TABLE 1

| Variant | Kd (nM) 15 °C | Kd (nM) 37 °C | Tm (complex) (°C) |
|---------|--------------|--------------|-------------------|
| WT      | 14 ± 3       | 14 ± 2       | 62                |
| W70L    | 53 ± 6       | 259 ± 37     | 44                |
| W70A    | 330 ± 30     | 466 ± 46     | 42                |
| W70F    | 49 ± 5       | 24 ± 2       | 59                |
| W70Y    | 32 ± 3       | 26 ± 4       | 59                |

* Tm is given at a protein-DNA concentration of 25 μM.

### FIGURE 4. Folding and stability of WT and variant SRY domains.

A, far-UV CD spectra of WT (black) and Trp-70 variants. Aromatic containing domains (W70F and W70Y, blue and green, respectively) resemble that of WT, compared with the diminished helical spectra (highlighted by the orange arrow at 222 nm) of the nonaromatic variants (W70A, W70L; gray and red, respectively). B, far-UV CD spectra of aliphatic Trp-70 variants at 37 °C. Relative to WT, W70A and W70L display reduced helical signal similar to that observed at 4 °C in A. C, aromatic variants W70F and W70Y (blue and green, respectively) display less attenuation in helical content at 37 °C relative to the aliphatic variants. D, change in mean residue ellipticity of WT and W70 variants at 222 nm as a function of concentration of guanidine HCl at 4 °C. Color codes as in A. Unfolding patterns of W70F and W70Y (blue and green solid lines, respectively) are consistent with a two-state model with lowered ΔG° values (Table 3). The aliphatic variants W70A and W70L (gray and red lines, respectively) exhibit noncooperative unfolding transitions. Inset, at low concentrations of guanidine HCl (0.015 to 0.9 M), no pre-transition baseline (as that observed for WT) could be defined for W70L. E–H, thermal denaturation studies of WT and variant domains. Change in mean residue ellipticity of WT (E) and Trp-70 variants (G) at 222 nm as a function of temperature is shown. The apparent midpoint unfolding temperature, Tm, was found to be 40 °C (arrow in F) for WT as determined from the derivative of the thermal unfolding curve, whereas W70Y and W70F yielded lower Tm values of 30 and 34 °C respectively (arrows in H). Color code, as in A. I, NMR spectroscopic structural studies. 1H-15N HSQC spectral overlay of WT SRY (black), W70L mutant (blue), and W70L mutant in 5.4 M urea (red).
Aromatic Buttress in an HMG Box

The integrated intensity of the Trp fluorescence emission spectrum of the WT complex at 15 °C was 2-fold less than that of the free domain, providing evidence for the quenching of one or more Trp side chains; such quenching is presumably due to stabilization of the domain on specific DNA binding leading to more complete desolvation of its hydrophobic core. In accordance with this interpretation, the maximum of the native WT spectrum (355 nm) was blue-shifted relative to its $\lambda_{\text{max}}$ in 8 M guanidine (364 nm, data not shown).

Comparative studies of Trp fluorescence demonstrated that the integrated intensities of the spectra of the free variant domains were each larger than that of the WT domain, despite the added presence of Trp-70 in the latter. Relative intensities followed the order W70L > W70A > W70Y = W70F > WT (Fig. 5A). With the exception of W70L > W70A, this trend correlates with side-chain volume. A red shift in $\lambda_{\text{max}}$ was observed for aliphatic variants W70L and W70A (359 ± 0.2 and 357.8 ± 0.2 nm, respectively; $\Delta \lambda$ 4 ± 0.4 nm relative to WT), whereas aromatic substitutions W70Y and W70F were associated with blue shifts ($\lambda_{\text{max}}$ 342.0 ± 0.2 nm and 344.2 ± 0.2 nm; $\Delta \lambda$ = 15 ± 0.4 nm and $\Delta \lambda$ = 11 ± 0.4 nm relative to WT). In control experiments (data not shown) Trp emission spectra of the denatured variant domains (unfolded in 8 M guanidine) were attenuated by one-third relative to WT in accordance with the number of Trp residues (two versus three); each exhibited the same emission maximum ($\lambda_{\text{max}}$ 362 nm).

On binding to the specific DNA site, Trp emission spectra of the W70F and W70Y domains exhibited changes similar to those observed in the WT complex (Fig. 5B). Despite the absence of Trp-70, their bound-state integrated amplitudes were only slightly larger than that of the WT complex, suggesting that the residue-specific contribution of Trp-70 to the emission spectrum of the WT complex is <7% of the total fluorescence intensity (Table 4). The spectra of the aliphatic variants retained the highest overall intensities. Unlike the free domains, however, relative intensities in the variant DNA complexes correlated with side-chain volume (W70A > W70L). In all cases, the domain.

The integrated intensity of the Trp fluorescence emission spectrum of the WT complex at 15 °C was 2-fold less than that of the free domain, providing evidence for the quenching of one or more Trp side chains; such quenching is presumably due to stabilization of the domain on specific DNA binding leading to more complete desolvation of its hydrophobic core. In accordance with this interpretation, the maximum of the native WT spectrum (355 nm) was blue-shifted relative to its $\lambda_{\text{max}}$ in 8 M guanidine (364 nm, data not shown).

Comparative studies of Trp fluorescence demonstrated that the integrated intensities of the spectra of the free variant domains were each larger than that of the WT domain, despite the added presence of Trp-70 in the latter. Relative intensities followed the order W70L > W70A > W70Y = W70F > WT (Fig. 5A). With the exception of W70L > W70A, this trend correlates with side-chain volume. A red shift in $\lambda_{\text{max}}$ was observed for aliphatic variants W70L and W70A (359 ± 0.2 and 357.8 ± 0.2 nm, respectively; $\Delta \lambda$ 4 ± 0.4 nm relative to WT), whereas aromatic substitutions W70Y and W70F were associated with blue shifts ($\lambda_{\text{max}}$ 342.0 ± 0.2 nm and 344.2 ± 0.2 nm; $\Delta \lambda$ = 15 ± 0.4 nm and $\Delta \lambda$ = 11 ± 0.4 nm relative to WT). In control experiments (data not shown) Trp emission spectra of the denatured variant domains (unfolded in 8 M guanidine) were attenuated by one-third relative to WT in accordance with the number of Trp residues (two versus three); each exhibited the same emission maximum ($\lambda_{\text{max}}$ 362 nm).

On binding to the specific DNA site, Trp emission spectra of the W70F and W70Y domains exhibited changes similar to those observed in the WT complex (Fig. 5B). Despite the absence of Trp-70, their bound-state integrated amplitudes were only slightly larger than that of the WT complex, suggesting that the residue-specific contribution of Trp-70 to the emission spectrum of the WT complex is <7% of the total fluorescence intensity (Table 4). The spectra of the aliphatic variants retained the highest overall intensities. Unlike the free domains, however, relative intensities in the variant DNA complexes correlated with side-chain volume (W70A > W70L). In all cases, the domain.

7 The contribution of Trp-107 to the WT fluorescence spectrum was inferred from comparison of the WT domain and a W107D variant (data not shown). Whereas the latter exhibited a native CD spectrum and native specific DNA-binding affinity below 30 °C, its Trp emission spectrum was decreased by ~30% relative to WT. In 6 M urea, the spectrum of the variant was decreased by one-third as expected (two Trp residues versus three).

**TABLE 2**

| Molecule | Free (37 °C) | Specific complex (37 °C) | Difference spectra* |
|----------|-------------|-------------------------|---------------------|
|          | $[\theta]_{208}$ | $[\theta]_{132}$ | $[\theta]_{208}$ | $[\theta]_{132}$ | $\Delta[\theta]$ |
| Free DNA | $-0.6 \pm 0.09$ | $2.0 \pm 1.2$ | IA* | IA | $-1.6 \pm 0.4$ |
| WT | $-15.6 \pm 0.06$ | $-12.5 \pm 0.4$ | $-16.5 \pm 0.8$ | $-16.2 \pm 0.4$ | $-0.7 \pm 0.7$ |
| W70A | $-9.7 \pm 0.8$ | $-4.2 \pm 1.5$ | $-12.0 \pm 2.5$ | $-9.0 \pm 0.9$ | $-2.3 \pm 1.6$ |
| W70L | $-11.7 \pm 0.7$ | $-4.9 \pm 0.4$ | $-13.0 \pm 0.9$ | $-10.0 \pm 0.6$ | $-1.3 \pm 0.5$ |
| W70F | $-13.0 \pm 0.6$ | $-8.0 \pm 0.3$ | $-15.3 \pm 0.8$ | $-13.4 \pm 0.5$ | $-2.1 \pm 0.5$ |
| W70Y | $-13.6 \pm 1.1$ | $-7.7 \pm 1.5$ | $-17.8 \pm 1.3$ | $-13.0 \pm 0.4$ | $-4.2 \pm 0.9$ |

* This is defined as the difference between the observed spectrum of the protein-DNA complex and the sum of the spectra of the free DNA and free protein (see Fig. 7).

**TABLE 3**

| Thermodynamic stabilities and thermal unfolding temperatures |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Variant         | $\Delta G_m$    | $m^o$           | $C_{\text{mid}}$ | $T_m$          |
|                 | kcal/mol         | kcal mol $^{-1}$ | m                |                |
| WT               | 3.0 ± 0.1        | 1.4 ± 0.01      | 2.2 ± 0.04       | 40°C           |
| W70A             | ND              | ND              | ND               | ND             |
| W70L             | ND              | ND              | ND               | ND             |
| W70F             | 2.4 ± 0.1        | 1.4 ± 0.03      | 1.7 ± 0.04       | 34°C           |
| W70Y             | 2.0 ± 0.1        | 1.3 ± 0.03      | 1.5 ± 0.04       | 30°C           |

* $\Delta G_m$ is apparent change in free energy of unfolding as extrapolated to zero denaturant concentration by a two-state model.

**Note:** Includes all Trp residues in the SRY domain.

- W70A: Trp-70, Trp-98, and Trp-107.
- W70L: Trp-70 and Trp-107.
- W70F: Trp-70, Trp-98, and Trp-107.
- W70Y: Trp-70, Trp-98, and Trp-107.

**Note:** The excitation wavelength (295 nm) and path length (0.3 cm) were chosen to minimize the DNA-dependent inner filter effect (see “Experimental Procedures”). Emission maxima ($\lambda_{\text{max}}$) and integrated amplitudes are given in Table 4.

**The Trp fluorescence emission spectrum of the WT domain contains a major contribution from the exposed and mobile indole ring of Trp-107 (box position 52) on the back surface of the domain.**
and variant domain clinical variant W70L (Table 4, 4th and 5th columns). The WT the integrated amplitude of the bound spectrum was decreased

**FIGURE 5.** Aromatic Buttress in an HMG Box

**TABLE 4**

| Molecule | λ_max (nm) | Integral (I; relative units) | iPer residue (I/Trp) |
|----------|-----------|-----------------------------|---------------------|
| WT       | 355       | 100                         | 33.3                |
| W70A     | 357.8     | 193                         | 96.5                |
| W70L     | 359       | 232                         | 116                 |
| W70F     | 344.2     | 150                         | 75                  |
| W70Y     | 340.2     | 148                         | 74                  |

Spectra were recorded with a step size of 0.2 nm, and so formally values are ±0.2 nm. Five independent control spectra of the WT domain acquired with a step size of 0.2 nm yielded an estimated λ_max = 354.6 (±0.2) nm.

Integrated amplitudes of the Trp emission spectrum (see Fig. 5, A and B) were normalized using five independent spectra of the free WT domain, with an average integral of 46.6 ± 0.7, set to a value of 100 (±1.5)% (bold in column 4).

This value represents the integrated amplitude of the Trp emission spectrum, normalized relative to the spectrum of the WT domain, and divided by the number of Trp residues in the polypeptide; the WT value is thus 100/33.3 (33.3, bold in column 6). Standard errors were estimated to be ±1.5% of the relative values shown based on five replicates of the WT spectrum.

**Variant Domains Exhibited Specific DNA-dependent Protein Folding**—At 4 °C far-UV CD spectra of the WT domain in its free and bound states exhibited similar α-helical features (Fig. 5C); the relative contribution of the DNA site to the spectrum in this region is smaller and of opposite sign (positive) (200–225 nm subregion) (58). The spectrum of the WT complex thus resembled the sum of the spectra of the free domain and free DNA, leading to a difference spectrum that was without significant features (see Fig. 7A, below). By contrast, in the near-UV region (320–250 nm) the DNA duplex makes the predominant contribution; its CD spectrum undergoes marked changes that resemble those associated with the canonical B → A transition (bracket in Fig. 5C) (59). Relative contributions from protein aromatic residues in this region, as inferred from the spectrum of the free domain (open circles in Fig. 5C), were negligible. Indeed, whereas the near-UV CD spectrum of the free DNA site (solid line in Fig. 5C) was typical of a mixed sequence DNA duplex (i.e. not a homopolymer), the spectrum of the bound DNA site (filled circles) exhibited A-like features, including a blue shift in band maximum (from 280 to 270 nm; Table 2). Such a spectral change has been observed in other SOX-DNA complexes (60) in accordance with the A-like structure of the bound DNA, which is under-wound and exhibits a widened minor groove relative to canonical B DNA (59).

At 37 °C the CD spectrum of the WT complex provided evidence for partial α-helical stabilization relative to the free domain (Fig. 5D) as observed in a previous study (41). Ellipticity at 222 nm is more negative in the complex than in the free domain (vertical gold arrow in Fig. 5D) despite negligible ellipticity at this wavelength in the spectrum of the free DNA site (solid line in Fig. 5D). Although a more significant contribution from the novel SRY-bound DNA conformation cannot be excluded, this possibility seems unlikely as reference CD spectra of canonical A, B, C, and Z DNA are weak in this region relative to that of an equimolar α-helical domain (59). Further-
more, the difference spectrum (Fig. 7B, below) exhibited an overall α-helical signature in the 200–250-nm region. Specific DNA-dependent accentuation of α-helical features in the CD spectrum of the WT complex at 37 °C may in principle reflect either induced fit or conformational selection (62).

The above DNA-dependent protein folding reflected partial thermal unfolding of the free WT domain at a temperature at which the specific DNA complex is stable. Relative thermal stabilities of the WT and variant DNA complexes were thus probed by monitoring ellipticity at 222 nm (Fig. 5E). Apparent midpoint unfolding temperatures \( T_m \) are given in Table 1; these values depend on the concentration of the complex (25 μM in this study). Under these conditions, the \( T_m \) of the WT complex was 62 °C (black dashed line in Fig. 5E). \( T_m \) values of the W70L and W70A complexes were markedly reduced (42 and 44 °C, respectively; red and gray dashed lines in Fig. 5E), whereas those of the W70F and W70Y complexes were similar to that of the WT complex (59 °C; blue dashed line in Fig. 5E).

Differential perturbations to the thermal stabilities of the variant domains and their DNA complexes motivated assessment of specific DNA-induced protein folding in each case. Observed CD spectra are shown in Fig. 6 and calculated difference spectra in Fig. 7. α-Helix stabilization was especially pronounced in the W70L complex at 4 °C (Fig. 6A). Its far-UV CD spectrum (red filled circles in Fig. 6A) and its calculated difference spectrum (open black triangles in Fig. 7C) exhibited a prominent α-helical signature at 208 and 222 nm. In the near-UV region, an A-like CD feature of the bound DNA (red filled circles in Fig. 6A) was observed, similar to that of the WT complex (black filled circles in Fig. 5C). In each case, the band maximum was blue-shifted relative to the spectrum of the free DNA site (brackets in each panel and Table 3).

At 37 °C, each of the variant domains exhibited DNA-dependent induction or stabilization of α-helix (Fig. 6, B–E). Although DNA binding is weaker at this temperature (Table 1), CD studies were performed under conditions in which the variant domains were >95% bound as inferred from their respective dissociation constants. The W70L complex appeared to contain less α-helix than the W70A complex as inferred from the magnitude of the negative bands at 208 and 222 nm (Fig. 6, B and C). Far-UV CD spectra of the W70L and W70A complexes were in turn attenuated relative to spectra of the W70F and W70Y complexes (Fig. 6, D and E) in accordance with their respective \( T_m \) values (Fig. 5E). As expected, based on disproportionate unfolding of the WT and variant free domains above 30 °C, CD difference spectra at 37 °C (Fig. 7, B, D, and F) in general exhibited more marked features than at 4 °C (Fig. 7, A, C, and E). The near-UV CD signature of an altered DNA structure was attenuated in the W70L complex at 4 °C (red line in Fig. 7G) relative to the WT or W70F complexes (black and blue lines in Fig. 7G). At 37 °C, both the W70L and W70F complexes exhibited attenuated near-UV signatures relative to that of the WT complex (Fig. 7H). Such attenuation, more marked than could be attributed to partial dissociation of the variant complexes, was not accompanied by a shift in band maximum toward that of the free DNA site.

**Aromatic Buttress in an HMG Box**

**FIGURE 6.** Specific DNA-induced folding of Trp-70 variants. Free and DNA-bound domains are indicated by open and closed circles, respectively. Increases in helical signal are indicated by gold arrows at the helix-sensitive wavelength of 222 nm (B–E). A, CD spectra of the W70L domain in free (open circles) and bound (closed circles) states at 4 °C. The spectrum of the free DNA site is shown as a black line. The near-UV band maximum of the variant domain-DNA complex is shifted relative to the free DNA site (bracket). B and C, the W70A and W70L are more structured when bound to DNA. D and E, DNA-induced increase in helical structures of the aromatic containing variants W70F (D) and W70Y (E) at 37 °C.

Specific DNA Bending Was Retained in the Variant SRY Domain-DNA Complexes—To evaluate SRY-directed DNA bending, equimolar solutions of the domain and the 15-bp FRET-labeled DNA duplex were prepared at a sufficiently high concentration (5 μM) to ensure at least 98% binding of the Tyr, Phe, and Ala variants and 90% binding of the Leu variant at 15 °C. Binding of the WT domain led to an increase in FRET efficiency from 10% (free DNA) to 47% (complex) in accordance with past studies (34). Similar increases in FRET efficiency were observed on binding of the W70Y and W70F domains (Fig. 8A). The variant W70L domain-DNA complex at a nominal concentration of 3 μM exhibited an apparent FRET efficiency of 34% (Fig. 8B); however, because 10% of the DNA would be unbound at this concentration (given \( K_d \) 0.3 mM; Table 1), this value represents an underestimate. To ensure at least 98% binding, steady-state FRET was re-measured at an effective concentration of 30 μM as accomplished by addition of the corresponding unlabeled DNA-domain complex to a concentration of 27 μM. Under these conditions, the FRET efficien-
Aromatic Buttress in an HMG Box

Substitutions Reduced the Lifetime of the Bent Domain-DNA Complex—Disassociation rates of the protein-DNA complexes were determined using a stopped-flow FRET-based assay. Preformed bent protein-DNA complexes were rapidly mixed with 20-fold excess of unlabeled target DNA (Fig. 8D). On mixing with the labeled complex, an increase in 6-FAM (donor) fluorescence was observed, reflecting the dissociation mixing with the labeled complex, an increase in 6-FAM (donor) fluorescence due to the addition of the nonspecific DNA duplex. An increase in 6-FAM (donor) fluorescence due to the addition of the nonspecific DNA site would signify competing nonspecific DNA binding. As shown in Fig. 9 (A, C, and E), for the aliphatic variants the increase (relative to WT) was negligible, providing evidence that the bent FRET-labeled complexes represent specific binding. In control studies employing an equimolar unlabeled specific DNA site, 6-FAM FRET emission was increased by 50% as would be predicted by re-equilibration of the domains between 6-FAM/TAMRA-labeled and -unlabeled specific DNA sites (Fig. 9, B, D, and F).

Cellular Studies Demonstrated Perturbations in Multiple Biological Processes—Comparative CH34 cell-based studies of epitope-tagged human SRY and its variants (W70F, W70Y, W70A and W70L) were undertaken following transient transfection of full-length constructs. Transfections were performed with either the SRY expression plasmid alone (1 μg of DNA per million cells; 1 × protocol) or a 1:50 mixture of the expression plasmid and its empty parent (combined DNA mass of 1 μg; 50×); such dilution was designed to mitigate potential effects of protein overexpression.8 Whereas undiluted transient transfection of the WT SRY expression plasmid led, under these conditions, to accumulation of ~1 million SRY molecules per cell, its 1:50 dilution led to accumulation of ~1000–10,000 molecules per cell9 (38) as is typical of TFS that regulate cell-fate decisions (63). Because variability in levels of WT SRY expres-

8Use of a strong viral promoter to drive expression of SRY in transient transfection studies (as employed here and by others (99)) typically leads to accumulation of ~108 protein molecules per cell, which on nuclear localization corresponds to a concentration of ~1 μM. Such overexpression can alter the effects of TF mutations on the transcriptional regulation of target gene or reporter genes (38).

9The number of WT or variant epitope-tagged SRY molecules per cell was estimated as described (38) by Western blot in reference to housekeeping control α-tubulin, assumed to be expressed at a level of 5 × 107 molecules per cell (61).
sion could occur when different aliquots of CH34 cells were employed at different dates (presumably reflecting differences in passage number since their original derivation), our comparative studies were restricted to side-by-side analysis of WT and variant SRY constructs on transient transfection of the same batch of CH34 cells on the same date. Under these conditions, the relative extent of variability of WT SRY accumulation (or that of a given mutant SRY), as evaluated following undiluted transfection by Western blot (directed against the HA tag of the expressed SRY) with normalization based on an α-tubulin housekeeping control, was less than 10%. Such reproducibility required stringent attention to the timing of transfection, length of cell culture, and their conditions.

Mutations at position 70 (HMG consensus position 15) can markedly impair SRY expression and nuclear localization. Whereas aromatic substitutions W70Y and W70F were well marked by partial dissociation of FRET-labeled W70A and W70L complexes (red, C and D, respectively) on addition of an equimolar specific unlabeled DNA site (black, E). Dissociation rates and lifetimes of protein–DNA complexes were determined according to (1/λ): 

\[
\lambda = k_{\text{off}}
\]

where \( k_{\text{off}} \) is the dissociation rate constant. The data in Table 5 were obtained from a mono-exponential fitting of the fluorescence donor emission on partial dissociation of FRET-labeled W70A and W70L complexes at a nominal concentration of 3 μM (B, Fig. 9A). At the latter concentration (100-fold higher than the perturbed \( k_{\text{off}} \)) FRET efficiency is indistinguishable from that of WT (C). D, stopped-flow FRET assay. Stopped-flow apparatus coupled to the fluorimeter and experimental design to measure dissociation of the SRY-DNA complex. One syringe contained preformed SRY-labeled DNA complex; a second syringe contained unlabeled DNA in 20-fold excess. Overlapping spectra were observed except for the W70L domain-DNA complex (red) due to its partial dissociation under these conditions. 

![Figure 8. FRET-based studies of specific DNA binding.](image)

![Figure 9. Sequence-specific DNA binding.](image)

### Table 5

| Protein complex | Dissociation rate constants \((k_{\text{off}}, s^{-1})\) for specific and non-specific DNA binding. Dissociation rates and lifetimes of protein–DNA complexes were determined according to (1/λ): |
|-----------------|--------------------------------------------------|
|                | 37 °C, lifetime | 25 °C, lifetime | 15 °C, lifetime | 6 °C, lifetime |
| WT              | 0.083 ± 0.001 | 0.046 ± 0.001 | 0.033 ± 0.001 | 0.032 ± 0.001 |
| W70A            | 12.4 ± 10^3  | 21.7 ± 10^3  | 30.3 ± 10^3  | 31.2 ± 10^3  |
| W70L            | >0.7 ± 10^3  | >0.7 ± 10^3  | 5.0 ± 10^3  | 7.14 ± 10^4  |
| W70F            | >0.7 ± 10^3  | >0.7 ± 10^3  | 5.0 ± 10^3  | 7.14 ± 10^4  |
| W70Y            | 0.1 ± 0.001  | 0.09 ± 0.001  | 0.065 ± 0.003 | 0.06 ± 0.003  |
|                 | 10^3         | 10^3         | 10^3         | 10^3         |

* ND means not determined due to rapid dissociation (lifetime < 7 × 10^-3 ms).
tolerated in both regards (Figs. 10A and 11A), aliphatic substitutions W70L and W70A led to an 8-fold reduction in the intensity of the corresponding HA-immunoblotted signal (Fig. 11B, lanes 2 and 3) and partial impairment of nuclear localization (Fig. 10A, panels f and g). Native-like levels of expression of W70L and W70A SRY could be regained on treatment of the cells with proteasome inhibitor MG132 (lanes 6 and 8 in Fig. 11B relative WT control in lane 4) and maintained on modification of the constructs by inclusion of an exogenous NLS (NLS; green in Fig. 11B as defined in Fig. 11C). Although following rescue of expression by MG132, the unmodified W70L and W70A SRY constructs exhibited impaired nuclear localization at a level intermediate between those of control NLS-defective constructs R62G (47) and R133W (45) (Figs. 10A, panels d and e; and 10B), on exogenous NLS fusion the W70L and W70A variants exhibited levels of nuclear localization (versus pan-cyttoplasmic staining) similar to that of WT SRY (Fig. 10B; illustrated in the case of W70L in panel j of Fig. 10A). In contrast, aromatic substitutions of W70Y and W70F have no significant effect on the pattern of subcellular localization (panels h and i in Fig. 10A and histogram in 10B).

The above patterns of subcellular localization correlate with the relative strength of binding to epitope-tagged Exp4 (co-IP assay in Fig. 10C), whereas binding to endogenous CaM is unaffected by the substitutions at position 70 (Fig. 10D). Together, these findings suggest that an aromatic residue at position 70 contributes to the function of bipartite NLS1 in the HMG box and also contributes to the resistance of SRY to proteosomal degradation. Additional evidence that the W70L and W70A variants (but not the W70Y and W70F variants) undergo accel-
erated degradation was provided by analysis of time-dependent protein turnover following inhibition of ribosomal translation by cycloheximide (Fig. 11)

The transcriptional regulatory activities of the SRY variants were assessed by quantitative rTPCR in relation to transcriptional activation of the endogenous Sox9 gene (38). Correspondence between this cell-based assay and in vivo function (i.e. in situ activation of a male-specific GRN in the bipotential XY gonadal ridge (64)) has previously been validated (38). Initial studies employed transient transfection of the undiluted WT and variant SRY expression plasmids in the absence of MG132 and exogenous NLS fusion (black bars in Fig. 11E). Under these conditions, the activities of W70Y and W70F SRY were indistinguishable from that of WT SRY, whereas W70L and W70A variants were completely without detectable activity relative to an empty vector control. To exclude the possibility that the high activities of W70L and W70F SRY variants represented an artifact of TF overexpression (38), their potencies were reassessed on 50X dilution (white bars in Fig. 11E, left). Such re-investigation confirmed the native-like activities of these variants; results were not affected by inclusion of an exogenous NLS in the construct (green-filled and open bars at right in Fig. 11E).

As expected, the inactive W70L and W70A SRY variants remained inactive on 50X dilution of these respective expression plasmids. To test whether such complete inactivity was due solely to impaired expression and nuclear entry or represented an intrinsic molecular defect of the mutant protein structure, we “rescued” nuclear accumulation by combination of MG132-mediated proteosomal inhibition and exogenous NLS fusion. Under these conditions, the extent of expression and nuclear localization under 1X conditions was similar to that of WT SRY (under 50X conditions respective cellular levels of expression were likewise similar to the 50X WT control, but nuclear localization could not be evaluated). Unlike the defective activity of inherited DSD-associated variant V60L as described previously (41), which was almost completely rescued by these manipulations (38), the W70L and W70A SRY variants remained completely inactive in accordance with their marked biophysical perturbations in vitro at 37 °C (see above).

ChIP studies enabled evaluation of the occupancy of the WT or variant epitope-tagged SRY at TESCO regulatory sites of the endogenous Sox9 gene in rat chromosome 10; primer sets a and c (Fig. 12A) served as probes to monitor SRY occupancy, and set b (flanking a neighboring DNA segment lacking detectable
Aromatic Buttress in an HMG Box

SRY-binding sites) provided a negative control. As in the above qPCR studies, SRY variants were constructed with an in-frame heterologous NLS (+NLS in Fig. 12B), and the transfected cells were treated with MG132 to equalize nuclear expression of the WT or variant SRY proteins. TESCO occupancy by W70F and W70Y SRY was indistinguishable (within error) from WT (Fig. 12A and histogram in C). By contrast, the W70L and W70A variant proteins exhibited no detectable enhancer binding activity (Fig. 12B). In these studies, a control was provided by an unrelated inherited mutation (V60L), previously shown to exhibit partial TESCO occupancy under these conditions (38).

DISCUSSION

This study has focused on a de novo mutation in the HMG box of SRY associated with somatic sex reversal and gonadoblastoma (9). Because the mutation affects a core Trp residue that is invariant among Sry and Sox domains (consensus position 15 (36, 37)) and is conserved as an aromatic residue (Phe or Tyr) among other families of HMG boxes (14), including Schizosaccharomyces pombe mating factor spMATA-mc (Fig. 13) (65), we exploited this clinical case report as a starting point for analysis of structure-activity relationships. The mutation (W70L) preserves the nonpolar character of this site but not its aromaticity. Although hydrophobic cores of globular proteins in general (such as globins (66) and lysozymes (67)) often tolerate interchange of aromatic and aliphatic side chains as alternative nonpolar packing schemes, we hypothesized that possible core packing in the HMG box is markedly restricted due to its small size, unusual angular shape, and functional role (as a DNA-bending motif).

Our approach began with an SRY-directed Y1H system to screen all 19 possible substitutions at box position 15 (position 70 of human SRY). Remarkably, only Trp conferred a blue colony on X-Gal plates, indicative of robust β-gal expression. Powder-blue colonies were conferred by aromatic substituents Tyr and Phe; the remaining 17 substitutions gave rise to white colonies. These results were corroborated by quantitative assays of enzymatic activity in yeast extracts. Based on this screen, we purified representative variant domains for biophysical study; in order of Y1H activity and side-chain volume, these were the WT domain, aromatic variants Phe and Tyr, and inactive variants Leu and Ala. Although marked differences were observed in the biophysical properties of the free domains, each variant exhibited detectable specific DNA-binding and near-native DNA bending activity as probed by FRET.

Biological Implications—SRY provides a model for a genetic switch that regulates an organ-specific GRN. Because a Y1H system is artificial and unlikely to reflect mechanisms of architectural gene regulation through long range mammalian enhancer elements (such as the TESCO site of the mammalian Sox9 gene (4)), we extended our studies to the effects of the SRY mutations in the intact protein on transient transfection of an embryonic pre-Sertoli cell line (38). Previous studies have shown that this cell line recapitulates essential aspects of the SRY-dependent GRN in the differentiating XY gonadal ridge, including immediate transcriptional activation of Sox9, downstream activation of embryonic testis-specific Sox9 target genes (encoding fibroblast growth factor 9 and prostaglandin D2 synthase; Fgf9 and Ptgsd) (38).

Our studies employed the endogenous Sox9 gene of this pre-Sertoli cell line as a transcriptional target of the transfected TF. Use of an endogenous target gene offered the advantages of providing a native enhancer element (TES) with core elements (TESCO) amenable to ChIP analysis (4); its presumed chromatin structure and epigenetic marks are presumably similar to those found in the site and stage of SRY function in embryogenesis (42). This protocol also avoided potential artifacts associated with co-transfected reporter genes, such as nonphysiological promoter structure and elevated copy number. Although

FIGURE 12. ChIP assays of Sox9 TESCO occupancy for human SRY variants. A, schematic model of rat Sox9 with testis-specific enhancer elements, including regions of TES and TESCO. TESCO fragments with SRY-binding sites are highlighted with gray boxes. Primer sets a and c associated with the fragments 4 and 8 (each contains two consensus SRY-binding elements highlighted with red dots) were probed for SRY occupancy; primer set b serves as negative control. B, representative gel shows PCR products of ChIP of SRY variants with fused exogenous SV40 NLS (green) and MG132 treatment. At left are nonspecific IgG controls; a negative control is provided by inactive human SRY variant I68A. C, histogram showing relative TESCO occupancies in each set by SRY variant proteins; the WT SRY signal is defined as 100%. Horizontal brackets designate statistical comparisons as in Fig. 10.
Our interests focused on the gene regulatory activities of the variant SRYs, potential confounding perturbations were assessed with respect to protein expression and subcellular localization. Indeed, robust cellular expression of the variant SRYs required an aromatic residue at position 70 (consensus position 15). Variants W70L and W70A by contrast exhibited diminished steady-state protein levels resolvable by proteasome inhibitor MG132; accelerated turnover of the latter mutant proteins was verified by cycloheximide pulse-chase assays. Furthermore, the W70L and W70A variants exhibited decreased binding to Exp4 (47), resulting in impaired nuclear localization.

To assess the intrinsic gene regulatory activities of these mutant TFs in a mammalian cell, we first defined conditions such that the levels of WT and mutant proteins were similar (via MG132 proteome inhibition) and the extent of nuclear localization was similar (via a fused SV40 NLS). Under these conditions, SRY-dependent transcriptional activation of the endogenous Sox9 gene was probed by qPCR at high and low levels of TF overexpression. The latter (corresponding to 1000–10,000 protein molecules per cell (38)) was accomplished by dilution of the expression vector by its empty parent plasmid (thus preserving the total amount of DNA per transfection). Through these experimental maneuvers, we found that the aromatic variants W70Y and W70F retain native gene regulatory activity, whereas the aliphatic variants W70L and W70A are completely inactive even when robustly expressed and properly localized in the nucleus. Such inactivity correlated with negligible accumulation of the mutant proteins at TESCO sites as monitored by ChIP.

This study has also provided mechanistic insight into the bipartite NLS in human SRY (residues 60–77 in intact human SRY; box positions 5–22) (68). This motif binds Exp4 and is conserved among SOX domains (47, 69). Although such Exp4 complexes have not been characterized, analogous bipartite NLS motifs have been dissected in studies of importin-α (70) as exemplified by the NLS of cargo protein nucleoplasmin (71). The crystal structure of an Imp-α-cargo peptide complex demonstrated interactions not only with the two clusters of basic residues but also with intervening nonpolar residues (72). In SRY and SOX domains, Trp-70 is located at position 8 of a 12-residue linker between basic clusters. Our findings that its substitution by Leu or Ala markedly impairs binding to Exp4 in association with defective nuclear import strongly suggest that this conserved aromatic ring contacts Exp4. Although it is formally possible that the effects of the Leu and Ala substitutions are indirect (i.e. mediated via transmitted structural perturbations associated with the partial unfolding of the free domains), this possibility seems unlikely given that partial unfolding of the WT domain may be required in the Exp4-cargo complex. Impaired Exp4 binding and nuclear entry have also been described on substitution of Val-60 and Met-64 (respective box positions 5 and 9) by other nonpolar side chains (38, 40). Together, these observations suggest that the bipartite NLS of SRY presents a specific signature of nonpolar side chains as well as canonical clusters of basic residues as a motif for Exp4 recognition. It would be of future interest to determine the crystal structure of a complex between Exp4 and a cargo peptide derived from an SRY or SOX domain.

Despite their perturbed efficiency of nuclear import, the present substitutions did not impair binding of SRY to CaM, previously proposed to mediate nuclear localization (48, 73). It is possible that CaM and Exp4 are each required for nuclear entry of SRY such that its impaired binding to either protein could reduce its nuclear import and accumulation, in turn attenuating transcriptional activation of Sox9.

Biophysical Implications—Each substitution at position 70 (box position 15) creates a cavity in the HMG domain. Cavity-creating mutations in the cores of globular proteins are in general destabilizing (74); the extent of destabilization depends the volume of the cavity, presence of polar groups, and potential for compensating changes in neighboring core positions (75). The potential cavity created at position 70 (when modeled with Gly in the absence of structural compensation) has a volume of 164 Å³; in such rigid models, each of the variants would leave a packing defect in core of the HMG box ranging from 153 Å³ (Ala) to 32 Å³ (Tyr). Studies of the free domains demonstrated that although aromatic substituents Tyr and Phe enable near-native folding and stability, aliphatic substituents Leu and Ala were associated with attenuated helical content and marked reductions in thermodynamic and thermal stabilities. Matthews and co-workers (76) have estimated that cavities in globular domains impose a free-energy penalty of 24–33 cal/mol/
A surprising outcome of this study was the finding that SRY-dependent DNA bending is robust to mutations (such as W70L and W70A) that profoundly impair the thermodynamic stability of the HMG box. Under conditions in which these variant free domains exhibit little organized secondary or tertiary structure (as probed by CD, fluorescence, and $^1$H-$^1$N NMR spectroscopy), native-like bending of a specific DNA target site is maintained as probed by FRET. This assay, which employs a 15-bp DNA duplex labeled at its respective 5'-ends by 6-FAM (donor) and TAMRA (acceptor), has previously been shown to correlate with the results of permutation gel electrophoresis (52). We imagine that specific DNA-dependent protein folding restores a well organized bent protein-DNA interface. It would be of future interest to investigate by time-resolved FRET and distance-distribution analysis (34) whether the precision of such sharp DNA bending is likewise unaffected.

The counterintuitive robustness of protein-directed DNA bending to near-complete mutational unfolding of the SRY domain raises the intriguing possibility that the HMG box superfamily may contain members that are natively unfolded (89) and yet functional in gene regulation. In the case of SRY, two factors may have favored the evolution of an autonomous fold. First, destabilization of the HMG box at 37 °C is associated with more rapid proteasome-mediated degradation of the intact variant protein in our rodent embryonic cell model. Such accelerated degradation may impair a genetic switch (by limiting enhancer occupancy) as exemplified by the mutation W70L clinical mutation. Second, native unfolding of a free polypeptide would be expected to reduce its specific DNA-binding affinity; the free-energy cost of specific DNA-dependent protein folding would reduce the otherwise favorable binding free-energy of a prefolded motif. Such an offsetting cost may contribute to the reduced affinities of the present variant domains, especially at 37 °C. Because DNA-dependent protein folding can enhance specificity despite lower affinities (89), it is nonetheless possible that this biophysical strategy may be utilized elsewhere in the HMG box superfamily.

Concluding Remarks—Specific residues in a protein may be conserved due to key roles in structure and function. In small motifs, such as the HMG box, a given side chain may play multiple roles, reinforcing its broad conservation among a family or superfamily of related folds. Trp-70 in human SRY (box position 15) provides a compelling example of such a residue. Clinical mutation W70L, originally described in association with gonadal dysgenesis and neoplasia (9), leads to a perfect storm of perturbations affecting domain folding, domain stability, resistance to proteasomal degradation, cellular trafficking, specific DNA-binding affinity, and enhancer occupancy. Amid this storm, selective preservation of specific DNA bending was unexpectedly observed. Although not sufficient to confer gene regulatory activity, we suggest that such coupling between DNA-dependent protein folding and sharp DNA bending may extend the repertoire of natively unfolded proteins to include anomalous members of the HMG box superfamily.

Clinical mutations in SRY and other TFs associated with birth defects and cancer provide valuable experiments of nature. Our studies of representative substitutions at box position 15 in SRY have thus highlighted the contribution of a conserved aromatic
residue to the thermodynamic and kinetic stability of a bent protein-DNA complex. Analyses of these mutations in a cellular model of the differentiating gonadal ridge have illuminated the relationship between these biophysical features and the impaired transcriptional induction of the male program in a patient with Swyer syndrome and a gonadal malignancy.

Acknowledgments—We thank C. M. Haqq and Prof. P. K. Donahoe for cell line CH34; Y. Yang for assistance with NMR spectroscopy; and Profs. B. Baker, P. DeHaseth, E. Haas, H.-Y. Kao, M. Montano, D. Samols, D. Schlesinger, and V. Yee for discussion. M. A. W. thanks J. Radek and A. Jansco-Radek for their participation in initial studies of SRY and is grateful to G. M. Clore, P. Koopman, V. Lefebvre, R. Lovell-Badge, R. Sekido, and D. Wilhelm for advice and encouragement.

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) A gene with homology to a conserved DNA-binding motif. Nature 346, 240–244
2. Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1991) Male development of chromosomally female mice transgenic for Sry. Nature 351, 117–121
3. Sekido, R. (2010) SRY: a transcriptional activator of mammalian testis determination. Int. J. Biochem. Cell Biol. 42, 417–420
4. Sekido, R., and Lovell-Badge, R. (2008) Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. Nature 453, 930–934
5. Lovell-Badge, R., Canning, C., and Sekido, R. (2002) Sex-determining Genes in Mice: Building Pathways, pp. 4–22, John Wiley & Sons Ltd., West Sussex, UK
6. Josso, N., Picard, J. Y., Rey, R., and di Clemente, N. (2006) Testicular anti-Mullerian hormone: history, genetics, regulation and clinical application. Pediatr. Endocrinol. Rev. 3, 347–358
7. McClelland, K., Bowles, J., and Koopman, P. (2012) Male sex determination: insights into molecular mechanisms. Asian J. Androl. 14, 164–171
8. Wilhelm, D., and Koopman, P. (2006) The makings of maleness: towards an integrated view of male sexual development. Nat. Rev. Genet. 7, 620–631
9. Hersmus, R., de Leeuw, B. H., Stoop, H., Bernard, P., van Doorn, H. C., Looijenga, L. H. (2009) A novel SRY missense mutation affecting nuclear export in a 46, XY female patient with bilateral gonadoblastoma. Eur. J. Hum. Genet. 17, 1642–1649
10. Ner, S. S. (1992) HMGs everywhere. Curr. Biol. 2, 208–210
11. Whitfield, L. S., Lovell-Badge, R., and Goodfellow, P. N. (1993) Rapid sequence evolution of the mammalian sex-determining gene SRY. Nature 364, 713–715
12. Zhao, L., and Koopman, P. (2012) SRY protein function in sex determination: thinking outside the box. Chromosome Res. 20, 153–162
13. Weir, H. M., Kraulis, P. J., Hill, C. S., Raine, A. R., Lause, E. D., and Thomas, J. O. (1993) NMR structure of an HMG box. EMBO J. 12, 1311–1319
14. Murphy, E. C., Zhurkin, V. B., Louis, J. M., Cornilescu, G., and Clore, G. M. (2001) Structural basis for SRY-dependent 46-X,Y sex reversal: modulation of DNA bending by a naturally occurring point mutation. J. Mol. Biol. 312, 481–499
15. Murphy, E. C., Zhurkin, V. B., Louis, J. M., Cornilescu, G., and Clore, G. M. (2001) Structural basis for SRY-dependent 46-X,Y sex reversal: modulation of DNA bending by a naturally occurring point mutation. J. Mol. Biol. 312, 481–499
16. Weiss, M. A., Ukiyama, E., and King, C. Y. (1997) The SRY cantilever motif discriminates between sequence- and structure-specific DNA recognition: alanine mutagenesis of an HMG box. J. Biomol. Struct. Dyn. 15, 177–184
17. Jauch, R., Ng, C. K., Narasimhan, K., and Kolatkar, P. R. (2012) The crystal structure of the Sox4-HMG domain-DNA complex suggests a mechanism for positional interdependence in DNA recognition. Biochem. J. 443, 39–47
18. Arvidson, D. N., Lu, F., Faber, C., Zalkin, H., and Brennan, R. G. (1998) The structure of PurR mutant L54M shows an alternative route to DNA kinking. Nat. Struct. Biol. 5, 436–441
19. King, C. Y., and Weiss, M. A. (1993) The SRY high mobility group box recognizes DNA by partial intercalation in the minor groove: a topological mechanism of sequence specificity. Proc. Natl. Acad. Sci. U.S.A. 90, 11990–11994
20. Kim, J., Zwieb, C., Wu, C., and Adhya, S. (1989) Bending of DNA by gene-regulatory proteins: construction and use of a DNA bending vector. Gene 85, 15–23
21. Haqq, C. M., King, C. Y., Ukiyama, E., Falsafi, S., Haqq, T. N., Donahoe, P. K., and Weiss, M. A. (1994) Molecular basis of mammalian sexual determination: activation of Mullerian inhibiting substance gene expression by SRY. Science 266, 1494–1500
22. Phillips, N. B., Jansco-Radek, A., Ittah, V., Singh, R., Chan, G., Haas, E., and Weiss, M. A. (2006) SRY and human sex determination: The basic tail of the HMG box functions as a kinetic clamp to augment DNA bending. J. Mol. Biol. 358, 172–192
23. Cameron, F. J., and Sinclair, A. H. (1997) Mutations in SRY and SOX9: testis-determining genes. Hum. Mutat. 9, 388–395
24. Lauter, V., Stiehlein, D., and Clevers, H. (1993) Ancestry and diversity of the HMG box superfamily. Nucleic Acids Res. 21, 2493–2501
25. Liang, T., Hou, C. C., She, Z. Y., and Yang, W. X. (2012) The SOX gene family: function and regulation in testis determination and male fertility maintenance. Mol. Biol. Rep. 316, 359–370
26. Chen, Y. S., Racca, J. D., Phillips, N. B., and Weiss, M. A. (2013) Inherited human sex reversal due to impaired nucleocytoplasmic trafficking of SRY defines a male transcriptional threshold. Proc. Natl. Acad. Sci. U.S.A. 110, E3567–E3576
27. Trimmer, E. E., Zamble, D. B., Lippard, S. J., and Essigmann, J. M. (1998)
Aromatic Buttress in an HMG Box

Human testis-determining factor SRY binds to the major DNA adduct of cisplatin and a putative target sequence with comparable affinities. Biochemistry 37, 352–362.

40. Li, B., Phillips, N. B., Janesco-Radek, A., Ittah, V., Singh, R., Jones, D. N., Haas, E., and Weiss, M. A. (2006) SRY-directed DNA bending and human sex reversal: reassessment of a clinical mutation uncovers a global coupling between the HMG box and its tail. J. Mol. Biol. 360, 310–328.

41. Phillips, N. B., Racca, J., Chen, Y. S., Singh, R., Janesco-Radek, A., Radek, J. T., Wickramasinghe, N. P., Haas, E., and Weiss, M. A. (2011) Mammalian testis-determining factor SRY and the enigma of inherited human sex reversal. J. Biol. Chem. 286, 36877–36807.

42. Haqq, C. M., and Donohoe, P. K. (1998) Regulation of sexual dimorphism in mammals. Physiol. Rev. 78, 1–33.

43. Südböck, P., and Scherer, G. (1997) Two independent nuclear localization signals are present in the DNA-binding high-mobility group domains of SRY and SOX9. J. Biol. Chem. 272, 27848–27852.

44. de Santa Barbara, P., Moniot, B., Poulat, F., and Berta, P. (2000) Expression and subcellular localization of SF-1, SOX9, WT1, and AMH proteins during early human testicular development. Dev. Dyn. 217, 293–298.

45. Li, B., Zhang, W., Chan, G., Janesco-Radek, A., Liu, S., and Weiss, M. A. (2001) Human sex reversal due to impaired nuclear localization of SRY. A clinical correlation. J. Biol. Chem. 276, 46480–46484.

46. Angertaro, A., Sim, H., Kelly, S., Preiss, S., Clayton, A., Jans, D. A., and Harley, V. R. (2003) A SOX9 defect of calmodulin-dependent nuclear import in campomelic dysplasia/autosomal sex reversal. J. Biol. Chem. 278, 33839–33847.

47. Gontan, C., Güttler, T., Engellen, E., Demmers, J., Fornerod, M., Grosveld, F., Tibboel, D., Görlich, D., Poot, R. A., and Rottier, R. J. (2009) Exportin 4 mediates a novel nuclear import pathway for Sox family transcription factors. J. Cell Biol. 185, 27–34.

48. Sim, H., Rimmer, K., Kelly, S., Ludbrook, L. M., Clayton, A. H., and Harley, V. R. (2005) Defective calmodulin-mediated nuclear transport of the sex-determining region of the Y chromosome (SRY) in XY sex reversal. Mol. Endocrinol. 19, 1884–1892.

49. Gehring, W. J., Müller, M., Affolter, M., Percival-Smith, A., Billetter, M., Qian, Y. Q., Otting, G., and Wuthrich, K. (1990) The structure of the homeodomain and its functional implications. Trends Genet. 6, 323–329.

50. Wolfe, S. A., Nekludova, L., and Pabo, C. O. (2000) DNA recognition by Cys2His2 zinc finger proteins.

51. de Ferranti, S., and Mozaffarian, D. (2008) The perfect storm: obesity, adipocyte dysfunction, and metabolic consequences. 183–212.

52. Weiss, M. A., Patel, D. J., Sauer, R. T., and Karplus, M. (1984) Two-dimensional 1H NMR study of the dynamics of the DNA binding of the HMG box from mouse sox-5.

53. Sudbeck, P., and Scherer, G. (1997) Two independent nuclear localization factors.

54. Haqq, C. M., and Donohoe, P. K. (1998) Regulation of sexual dimorphism in mammals. Physiol. Rev. 78, 1–33.

55. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1997) A direct comparison of factors.

56. Qian, Y. Q., Otting, G., and Wuthrich, K. (1990) The structure of the homeodomain and its functional implications.

57. Kuwajima, K. (1996) The molten globule state of a- helix.

58. Hiller, G., and Weber, K. (1978) Radioimmunoassay for tubulin: a quantitative comparison of the tubulin content of different established tissue culture cells and tissues. Cell 14, 795–804.

59. Goentoro, L., Shoval, O., Kirschner, M. W., and Alon, U. (2009) The incoherent feedforward loop can provide fold-change detection in gene regulation. Mol. Cell 36, 894–899.

60. Clore, G. M. (2014) Interplay between conformational selection and induced fit in multimeric protein-ligand binding probed by paramagnetic relaxation enhancement. Biophys. Chem. 186, 3–12.

61. Dooijes, D., van de Wetering, M., Knippsel, L., and Clevers, H. (1993) The Schizosaccharomyces pombe mating-type gene mat-Mc encodes a sequence-specific DNA-binding high mobility group box protein. J. Biol. Chem. 268, 24813–24817.

62. Tsuchiya, M., Ogawa, H., Suzuki, T., Sugiyama, N., Haraguchi, T., and Hiraoka, Y. (2011) Exportin 4 interacts with Sox9 through the HMG box and inhibits the DNA binding of Sox9. PLoS One 6, e25694.

63. Prieve, M. G., Guttridge, K. L., Munguia, I., and Waterman, M. L. (1998) Differential importin-α recognition and nuclear transport by nuclear localization signals within the high mobility group DNA binding domains of lymphoid enhancer factor 1 and T-cell factor 1. Mol. Cell. Biol. 18, 4819–4832.

64. Robbins, J., Dilworth, S. M., Laskey, R. A., and Dingwall, C. (1991) Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence.

65. Stewart, M. (2006) Structural basis for the nuclear protein import cycle. Biochem. Soc. Trans. 34, 701–704.

66. Luke, D., and Scherer, G. (1997) Proteins.

67. Redfield, C., Schulman, B. A., Milhollen, M. A., Kim, P. S., and Dobson, C. M. (1999) a-Lactalbumin forms a compact molten globule in the absence of disulfide bonds. Nat. Struct. Mol. Biol. 6, 948–952.

68. Hua, Q.-X., Jia, W., Frank, B. H., Phillips, N. F., and Weiss, M. A. (2002) A protein caught in a kinetic trap: structures and stabilities of insulin disulfide isomers. Biochemistry 41, 14700–14715.
83. Harrison, S. C. (1991) A structural taxonomy of DNA-binding domains. *Nature* **353**, 715–719
84. Weiss, M. A., Ellenberger, T., Wobbe, C. R., Lee, J. P., Harrison, S. C., and Struhl, K. (1990) Folding transition in the DNA-binding domain of GCN4 on specific binding to DNA. *Nature* **347**, 575–578
85. Ellenberger, T. (1994) Getting a grip on DNA recognition: structures of the basic region leucine zipper, and the basic region helix-loop-helix DNA-binding domains. *Curr. Opin. Struct. Biol.* **4**, 12–21
86. Weiss, M. A. (2001) Floppy SOX: mutual induced fit in hmg (high-mobility group) box-DNA recognition. *Mol. Endocrinol.* **15**, 353–362
87. Love, J. J., Li, X., Chung, J., Dyson, H. J., and Wright, P. E. (2004) The LEF-1 high-mobility group domain undergoes a disorder-to-order transition upon formation of a complex with cognate DNA. *Biochemistry* **43**, 8725–8734
88. Burley, S. K., and Petsko, G. A. (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science* **229**, 23–28
89. Liu, Z., and Huang, Y. (2014) Advantages of proteins being disordered. *Protein Sci.* **23**, 539–550
90. Ukiyama, E., Jancso-Radek, A., Li, B., Milos, L., Zhang, W., Phillips, N. B., Morikawa, N., King, C. Y., Chan, G., Haqq, C. M., Radek, J. T., Poulat, F., Donahoe, P. K., and Weiss, M. A. (2001) SRY and architectural gene regulation: the kinetic stability of a bent protein-DNA complex can regulate its transcriptional potency. *Mol. Endocrinol.* **15**, 363–377
91. Crane-Robinson, C., Read, C. M., Cary, P. D., Driscoll, P. C., Dragan, A. I., and Privalov, P. L. (1998) The energetics of HMG box interactions with DNA. Thermodynamic description of the box from mouse Sox-5. *J. Mol. Biol.* **281**, 705–717
92. Jauch, R., Ng, C. K., Saikatendu, K. S., Stevens, R. C., and Kolatkar, P. R. (2008) Crystal structure and DNA binding of the homeodomain of the stem cell transcription factor Nanog. *J. Mol. Biol.* **376**, 758–770
93. Weinrich, M., and Hughes, T. R. (2011) in *A Handbook of Transcription Factors* (Hughes, T. R., ed) pp. 25–73, Springer, Netherlands
94. Lipowsky, G., Bischoff, F. R., Schwarzaier, P., Kraft, R., Kostka, S., Hartmann, E., Kutay, U., and Görlich, D. (2000) Exportin 4: a mediator of a novel nuclear export pathway in higher eukaryotes. *EMBO J.* **19**, 4362–4371
95. Gasca, S., Canizares, J., De Santa Barbara, P., Mejean, C., Poulat, F., Berta, P., and Boizet-Bonhoure, B. (2002) A nuclear export signal within the high mobility group domain regulates the nucleocytoplasmic translocation of SOX9 during sexual determination. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11199–11204
96. Forwood, J. K., Harley, V., and Jans, D. A. (2001) The C-terminal nuclear localization signal of the sex-determining region Y (SRY) high mobility group domain mediates nuclear import through importin β1. *J. Biol. Chem.* **276**, 46575–46582
97. Harley, V. R., Layfield, S., Mitchell, C. L., Forwood, J. K., John, A. P., Briggs, L. J., McDowall, S. G., and Jans, D. A. (2003) Defective importin β recognition and nuclear import of the sex-determining factor SRY are associated with XY sex-reversing mutations. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7045–7050
98. Junger, S. (1997) *The Perfect Storm: A True Story of Men Against the Sea*, W. W. Norton and Co., New York
99. Knowler, K. C., Kelly, S., Ludbrook, L. M., Bagheri-Fam, S., Sim, H., Bernard, P., Sekido, R., Lovell-Badge, R., and Harley, V. R. (2011) Failure of SOX9 regulation in 46XY disorders of sex development with SRY, SOX9, and SF1 mutations. *PLoS One* **6**, e17751