doi:10.1016/j.jmb.2006.01.060 *J. Mol. Biol.* (2006) **358**, 172–192

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

SRY and Human Sex Determination: The Basic Tail of the HMG Box Functions as a Kinetic Clamp to Augment DNA Bending

Nelson B. Phillips[1]†, Agnes Jancso-Radek[1]†, Varda Ittah[2] Rupinder Singh[1], Ging Chan[1], Elisha Haas[2] and Michael A. Weiss[1]*

[1]*Department of Biochemistry Case School of Medicine, Case Western Reserve University Cleveland, OH 44106-4935 USA*

[2]*Faculty of Life Sciences, Bar Ilan University, Ramat Gan 52900, Israel*

**Corresponding author*

Human testis-determining factor SRY contains a high-mobility-group (HMG) box, an α-helical DNA-binding domain that binds within an expanded minor groove to induce DNA bending. This motif is flanked on the C-terminal end by a basic tail, which functions both as a nuclear localization signal and accessory DNA-binding element. Whereas the HMG box is broadly conserved among otherwise unrelated transcription factors, tails differ in sequence and mode of DNA binding. Contrasting examples are provided by SRY and lymphoid enhancer factor 1 (LEF-1): whereas the SRY tail remains in the minor groove distal to the HMG box, the LEF-1 tail binds back across the center of the bent DNA site. The LEF-1 tail relieves electrostatic repulsion that would otherwise be incurred within the compressed major groove to enable sharp DNA bending with high affinity. Here, we demonstrate that the analogous SRY tail functions as a "kinetic clamp" to regulate the lifetime of the bent DNA complex. As in LEF-1, partial truncation of the distal SRY tail reduces specific DNA affinity and DNA bending, but these perturbations are modest: binding is reduced by only 1.8-fold, and bending by only 7–10°. "Tailed" and truncated SRY complexes exhibit similar structures (as probed by NMR) and distributions of long-range conformational substates (as probed by time-resolved fluorescence resonance energy transfer). Surprisingly, however, the SRY tail retards dissociation of the protein–DNA complex by 20-fold. The marked and compensating changes in rates of association and dissociation observed on tail truncation, disproportionate to perturbations in affinity or structure, suggest that this accessory element functions as a kinetic clamp to regulate the lifetime of the SRY-DNA complex. We speculate that the kinetic stability of a bent DNA complex is critical to the assembly and maintenance of a sex-specific transcriptional pre-initiation complex.

Keywords: gene regulation; protein structure; intersex abnormalities; gonadal dysgenesis; human development

† N.B.P. & A.J.-R. contributed equally to the work.

Present address: A. Jancso-Radek, Epicentre Technologies, Madison, WI 53713, USA.

Abbreviations used: 6-FAM, 6-isomer of aectamido fluorescein; FRET, fluorescence resonance energy transfer; HMG, high mobility group; GMSA, gel mobility-shift assay; LEF-1, lymphoid enhancer factor 1; MIS, Mullerian inhibiting substance (also designated Anti-Mullerian Hormone); NLS, nuclear localization signal; NOESY, nuclear Overhauser effect spectroscopy; PGE, permutation gel electrophoresis; rp-HPLC, reverse-phase high-performance liquid chromatography; SOX, SRY-related HMG box; SRY, protein encoded by the sex-determining region of the Y chromosome; SRY-p and SRY-p_{Δ}, domains containing residues 2–86 and 2–78, respectively; TAMRA, tetramethyl-rhodamine; TBP, TATA-binding protein; TCF-1α, T-cell factor 1α and homologue of LEF-1; tr, time resolved. Amino acids are designated by standard single-letter code. Residue numbers in SRY refer to an HMG box consensus unless otherwise indicated. Consensus position 1 corresponds to residue 56 of intact human SRY. Genes and DNA sites are indicated in italics, and proteins in capital letters.

E-mail address of the corresponding author: michael.weiss@case.edu

Introduction

Sexual dimorphism provides a model of a genetic switch between alternative programs of development.[1] The male phenotype in eutherian mammals is determined by *Sry* (sex-determining region of the Y chromosome),[2] a gene on the short arm of the Y chromosome. SRY contains a sequence-specific HMG box,[3] a conserved motif of DNA binding and DNA bending.[4] Assignment of SRY as the testis-determining factor (TDF) is supported by studies of transgenic murine models[5] and human intersex abnormalities.[6] SRY belongs to an extended family of SRY-related HMG-box proteins, designated SOX.[7] Members of this family play critical but diverse roles in mammalian organogenesis, including development of the heart[8] and central nervous system.[9] Autosomal gene *Sox9* functions downstream of *Sry* in the testis-determining pathway.[10,11] Clinical mutations in either *Sry*[6,12] or *Sox9*[13,14] are associated with pure gonadal dysgenesis leading to XY sex reversal. The HMG boxes of SRY, SOX9, and other members of the SOX family exhibit similar specific DNA-binding and DNA-bending properties.[15–18]

How SRY regulates testicular differentiation is not well understood. Comparison of *Sry* alleles has demonstrated marked divergence outside of the HMG box.[19] Selective conservation of a specific DNA-bending motif led to the proposal that SRY functions as an architectural transcription factor.[3,20] This proposal is supported by clinical observations that almost all point mutations causing human sex reversal occur *de novo* in the HMG box and impair specific DNA binding[20,21] or DNA bending.[22] That the primary function of SRY is site-specific DNA bending is suggested by studies of chimeric murine *Sry* (*mSry*) transgenes in XX mice.[23,24] Participation of SRY-interacting proteins in multi-protein-DNA complexes[25–28] and splicing-related RNA complexes[29] has also been proposed.[30] Although target genes for SRY-mediated transcriptional regulation have not been defined, the dramatic effects of HMG boxes on DNA structure (Figure 1(a) and (b)) are proposed to contribute to assembly of sex-specific transcriptional pre-initiation complexes (Figure 1(c)).[30,31] Studies of chimeric SRY transgenes have shown that SRY-directed sex reversal in XX mice readily tolerates modest changes in the molecular properties of the HMG box,[23,24] including subtle changes in sequence specificity and extent of DNA bending.[32,33]

The HMG box exhibits an L-shaped structure in which three α-helices and an N-terminal β-strand pack to form major and minor "wings" (Figure 1(a)).[34,35] The motif was originally described in the non-sequence-specific DNA-binding domains of abundant non-histone chromosomal proteins.[31] Two classes of HMG boxes are distinguished by their DNA-binding properties. Whereas non-histone chromosomal proteins (such as HMG1 and HMG2) typically contain two or more boxes that recognize distorted DNA structures with weak or absent sequence specificity, sequence-specific architectural transcription factors (such as lymphoid enhancer factor 1; LEF-1)[36,37] contain one HMG box that recognizes specific DNA sequences. Specific and non-specific domains each dock within a widened minor groove: the angular wings provide a template for DNA bending.[37–41] The extent of bending varies but in each case the protein binds on the outside of the DNA bend to compress the major groove. Structural models of sequence-specific HMG box-DNA complexes are provided by SRY (Figure 1(a))[38,42] and LEF-1 (Figure 1(b)).[37] DNA bend angles in these structures are 54° and 117°, respectively. The reported precision of the DNA bend angle in the SRY complex is ±2°. The physical significance of such precision is not well characterized.

The structure of an SRY–DNA complex has provided a basis for analysis of mutations associated with sex reversal (46, XY pure gonadal dysgenesis).[38] The present study focuses on the role of a conserved basic region C-terminal to the HMG box (red segment in Figure 1(a) and Figure 2(a)). Comprising residues 125–140 in human SRY (residues 70–85 in an HMG box consensus sequence; Figure 2(b)), this basic C-terminal tail also contains a nuclear localization signal[43,44] and sites of clinical mutations.[30] Although SRY and LEF-1 contain analogous tails, the two proteins exhibit different patterns of basic residues, proline and glycine residues (upper panel and last line of Figure 2(b), respectively). These divergent features correspond to distinct structural roles. Whereas the SRY tail remains in the DNA minor groove distal to the HMG box and is well ordered (Figure 1(a)),[42] the LEF-1 tail crosses back across the major groove in the center of the DNA site and is not well ordered (Figure 1(b)).[37] The basic side-chains of the LEF-1 tail (KKRK; green in Figure 2(b)) are proposed to alleviate electrostatic repulsion that would otherwise occur between neighboring DNA phosphodiester charges in the compressed major groove.[37] Truncation of the LEF-1 tail markedly impairs both DNA bending (from 130° to 52°) and specific DNA affinity (by at least 100-fold).[45–48]

To define the contribution of the SRY tail to DNA binding and DNA bending, we have investigated SRY domains containing either intact or truncated tails. These domains (designated "tailed" or truncated, respectively) correspond to the polypeptides employed in two successive NMR studies of the SRY–DNA complex.[38,42] We have characterized tailed and truncated domains in otherwise identical contexts and by multiple independent biochemical and biophysical assays. As expected,[49] the structure and stability of the free domain are not perturbed by truncation of its disordered tail; respective DNA-bound structures are also similar. As in LEF-1,[45–48] the SRY tail augments specific DNA binding and bending, but in SRY these effects are modest: DNA binding is reduced by less than twofold and DNA bending by only 7–10°. Further, binding of the tail does not damp long-range fluctuations in DNA