Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 11990–11994, December 1993
Biochemistry

The SRY high-mobility-group box recognizes DNA by partial intercalation in the minor groove: A topological mechanism of sequence specificity

(testis-determining factor/transcription factors/DNA-binding proteins/two-dimensional-NMR spectroscopy)

CHIH-YEN KING* AND MICHAEL A. WEISS*†‡

*Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115; and †Department of Medicine, Massachusetts General Hospital, Boston, MA 02114

Communicated by Elkan Blout, September 24, 1993 (received for review July 18, 1993)

ABSTRACT **SRY, a putative transcription factor encoded by the sex-determining region of the human Y chromosome, regulates a genetic switch in male development. Impairment of this switch leads to intersex abnormalities of the newborn and is observed in association with mutations in the SRY DNA-binding domain [the high-mobility-group (HMG) box]. Here we show that the SRY HMG box exhibits a novel mechanism of DNA recognition: partial intercalation of a nonpolar side chain in the DNA minor groove. Base stacking (but not base pairing) is interrupted at the site of insertion. Sequence specificity reflects topological requirements of partial intercalation rather than direct readout of base-specific functional groups. Our results predict that the SRY HMG box inserts an α-helix into a widened minor groove at the center of a sharp DNA bend. A similar mechanism may underlie binding of SRY and homologous HMG proteins to four-way junctions (Holliday intermediates) and other noncanonical DNA structures.**

Protein–DNA recognition is mediated by direct or indirect readout of bases in a DNA sequence (1). Direct readout consists of specific chemical interactions between protein and nucleic-acid bases (2). Indirect readout, inferred from the crystal structure of the *trp* repressor (TrpR)–operator complex, refers to recognition of sequence-dependent variations in the structure or deformability of the DNA backbone (3). Here we describe a third mechanism of protein–DNA recognition, partial intercalation of a nonpolar protein side chain through the DNA minor groove. Our studies focus on SRY, a putative transcription factor that initiates male development in mammalian embryogenesis (4). SRY contains a high-mobility-group (HMG) box, a newly recognized motif (5, 6) also implicated in recognition of noncanonical DNA structures (7). A combination of biochemical and ^{1}H NMR methods is used to define a contact between the SRY HMG box and a bent DNA site. This contact generates rules of specificity based on topological requirements of partial intercalation.

MATERIALS AND METHODS

Experimental design is based on identification of high-affinity SRY target sites in promoters of sex-specific genes (8). The SRY HMG box (designated SRY-p; 85 residues) was expressed in *Escherichia coli* and purified as described (8). Oligonucleotides used in gel mobility-shift assays were obtained from Oligos, Etc. (Wilsonville, OR). ^{32}P-labeled duplex DNA and SRY-p were incubated at 20°C in 10 mM potassium phosphate, pH 7.4/50 mM KCl/2 mM $MgCl_2$/0.5 mM dithiothreitol/5% glycerol (vol/vol). Each reaction mixture contained a variant 15-bp duplex (the test site) and 36-bp duplex containing a native target sequence (the reference site); complexes were resolved by polyacrylamide gel electrophoresis as described in Fig. 1. Quantitation was obtained with a digital fluorescence scanner (Molecular Dynamics).

For NMR study, complementary 15-base oligonucleotides were purchased from Pharmacia; purity was >98% as assessed by gel electrophoresis. DNA and the SRY-p/DNA complex were made 2 mM in 50 mM KCl/10 mM potassium phosphate, pH 6.0, in 10% ^2H_2O; spectra of SRY-p were also obtained in 50 mM sodium deuterioacetate (pH 3.7). Samples in ^2H_2O were prepared following lyophilization. ^{1}H NMR spectra were obtained at 500 MHz at 15°C and 40°C. Nuclear Overhauser effect spectroscopy (NOESY) experiments (mixing times, 75, 150, and 200 ms) were performed in H_2O, using shaped pulses for selective excitation (ssNOESY; ref. 9). Total correlation spectroscopy (TOCSY) spectra (mixing time, 55 ms) were obtained in H_2O. NOESY, TOCSY, and double-quantum-filtered correlated spectroscopy (DQF-COSY) spectra were also obtained in ^2H_2O. Imino resonances were assigned on the basis of NOEs between successive base-pairs; adenine H2 protons were assigned by NOE from the T imino proton of the same A·T base-pair (10). Resonance assignment in the complex was corroborated by observation of chemical-exchange crosspeaks between free and bound DNA under conditions of slow exchange.

STATEMENT OF THE PROBLEM: THE SPECIFICITY PARADOX

Is SRY a sequence-specific DNA-binding protein? Although DNase protection experiments demonstrate the presence of specific target sites in human genes (8), sequence-independent binding to four-way junctions (Holliday intermediates) is also observed (11). The dissociation constant is 1–100 nM in each case. How can nucleic acid binding be both sequence-specific and sequence-independent? The SRY specificity paradox poses a general question of HMG-box structure and function.

RESULTS

Overview of Experimental Design. DNA target sites of SRY-p and homologous HMG domains, although not intrinsically bent, exhibit anomalous electrophoretic mobility on protein binding; in an SRY complex the apparent bend angle is 90° (11, 12). The present study employs a 15-bp DNA duplex containing a 9-bp target site (boldface).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: NOESY, nuclear Overhauser effect spectroscopy; HMG, high mobility group; SRY-p, 85-mer peptide containing the human SRY HMG box; ssNOESY, NOESY with gaussian-shaped pulses for selective excitation; TBP, TATA-binding protein.

‡To whom reprint requests should be addressed.

```
        5         10        15
5'-GGGGTGATTGTTCAG-3'
3'-CCCCACTAACAAGTC-5'
        5'        10'       15'
```

The affinity of this site for SRY-p (K_{app}) is ≈10 nM at 20°C. Of particular interest is the central TT element (A·T base pairs 8 and 9; boxed), which is invariant among related sites. Comparative gel-retardation studies demonstrate that specificity for T^8 and T^9 is stringent (Fig. 1 and Table 1, transitions and transversions). No specific binding is detectable following transition $T^9 \rightarrow C$ (lane j) or any transversion (lanes b, d, h, and k); transition $T^8 \rightarrow C$ reduces specific binding by a factor of 5 (lane c). Nevertheless, concerted substitution of either T·A base pair by cytidine·inosine (C·I) is well tolerated (12, 13). Because C·I substitution alters functional groups only in the major groove, its tolerance implies that recognition is effected in the minor groove. Here we focus on the molecular mechanism of such recognition.

A Protein Side Chain Inserts Between Base Pairs. Fig. 2 shows ^{1}H NMR spectra of SRY-p (spectrum A), target DNA (C), and a 1:1 protein–DNA complex (B). Large upfield shifts are observed in protein aliphatic resonances on DNA binding. Particularly prominent are the $C^{\gamma}H^{1,2}$ and δ-CH_3 resonances of an isoleucine spin system (1.06, 0.49, and −1.22 ppm, respectively; asterisk in Fig. 2, spectrum B). Intermolecular NOEs are observed from the isoleucine δ-CH_3 resonance to imino resonances of T^8 and T^9 (Fig. 3) and from isoleucine $C^{\gamma}H^{1,2}$ and δ-CH_3 resonances to $A^{9'}$ H2 in the minor groove (crosspeaks a–c in Fig. 4). An NOE is also observed from isoleucine δ-CH_3 to $A^{8'}$ H2 but is 10-fold weaker (arrow in Fig. 4). These NOEs demonstrate that the isoleucine side chain enters the minor groove (in the absence of Hoogsteen base pairing; see below) and partially penetrates between A·T base pairs 8 and 9. Its upfield complexation shift presumably reflects the interior ring current of DNA bases (14).

Assignment of the isoleucine has not been obtained; the two possibilities are Ile^{13} and Ile^{37} (consensus HMG numbering scheme; ref. 5). In the solution structure of HMG1, homologous residues Phe^{13} and Val^{37} are in α-helices 1 and 2, respectively (6). Extension of this model to SRY predicts insertion of an α-helix into the minor groove. Because the

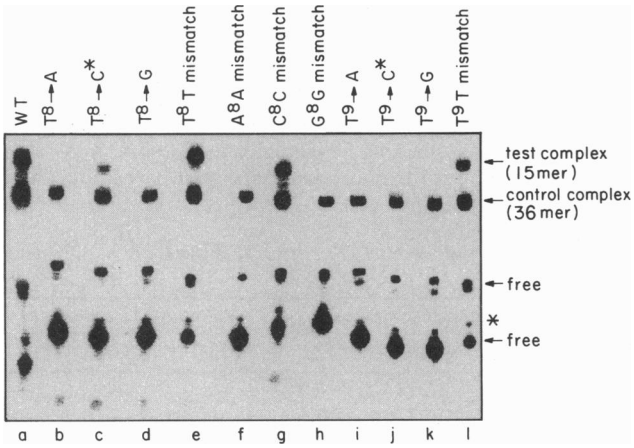

FIG. 1. Relative protein–DNA dissociation constants are determined by gel mobility-shift assay. In each lane the test site is a 15-mer duplex (5′-GGGGTGANNGTTCAG-3′) and the reference site is a 36-mer duplex (5′-CATACTGCGGGGGTGATTGTTCAGGATCATACTGCG-3′) (arrows at right). The test site used in lanes a–l is indicated at top; WT denotes wild-type test site, $T^8 \rightarrow A$ denotes transversion of T·A to A·T at position 8, etc. Asterisks at top (lanes c and j) highlight contrasting effect of T·A → C·G transitions. Each experiment was run in triplicate; the protein concentration was 10 nM. The figure is a composite of multiple gels. Asterisk at right indicates trace formation of 15-mer/36-mer heteroduplex. Complexes were incubated overnight to allow for the possibility of slow binding to mutant sites.

Table 1. Relative affinities of DNA analogs

Base pair	Change	Relative affinity, %	Base pair	Change	Relative affinity, %
	Transitions and Transversions				
8	TA → CG	20	9	TA → CG	ND
8	TA → AT	ND*	9	TA → AT	<1
8	TA → GC	<1	9	TA → GC	ND*
	Mismatches†				
7	AT → TT	5	7	AT → AA	20
8	TA → TT	130	8	TA → GG	ND
8	TA → CC	140	8	TA → AA	ND
9	TA → TT	100	9	TA → GG	ND
9	TA → CC	40	9	TA → AA	ND
9	GC → CC	ND	9	GC → GG	ND

Values shown represent the average of three independent determinations with variance <10% of the value shown (e.g., 20 ± 2%). ND, no specific binding detectable (relative K_d << 0.5%).
*Concerted mutations $T^8 \rightarrow A$ and $T^9 \rightarrow A$, which recreate a TTT site in the lower strand, also eliminate specific binding.
†Mismatches were studied at pH 7.4.

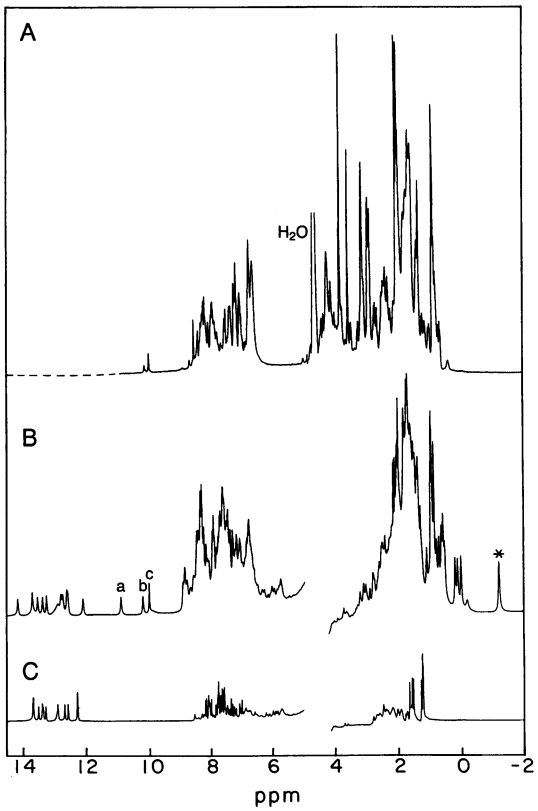

FIG. 2. ^{1}H NMR spectra of SRY-p (spectrum A), 1:1 protein–DNA complex (spectrum B), and free DNA duplex (5′-GGGGTGATTGTTCAG-3′) (spectrum C). Perturbations are observed in the chemical shifts of both DNA and protein resonances; DNA imino (12–14 ppm) and adenine H2 (7–8 ppm) resonance assignments are given in Table 2. Of particular interest are protein aliphatic resonances in the upfield region (−2 to 1 ppm), including an isoleucine δ-CH_3 resonance (asterisk in spectrum B). Tryptophan indole NH resonances are labeled a–c in spectrum B and also exhibit large changes. Spectrum A was acquired at 25°C in 50 mM sodium deuterioacetate (pH 3.7) following solvent presaturation; the residual solvent resonance is labeled in spectrum A; the dashed line in spectrum A represents an extrapolated baseline. Spectra B and C were acquired at 40°C in 50 mM KCl/10 mM potassium phosphate, pH 6.0, following selective excitation (9).

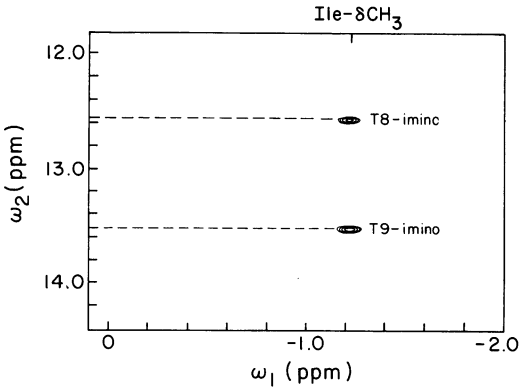

FIG. 3. Region of ssNOESY spectrum in H_2O showing NOEs between an upfield-shifted isoleucine δ-CH_3 resonance (ω_1, horizontal axis) and T^8 and T^9 imino resonances (ω_2, vertical axis). Mixing time was 75 ms and temperature was 40°C.

minor groove is too narrow in B-DNA to accommodate an α-helix, the data further predict that the minor groove is widened at the site of partial intercalation.

The DNA Structure Is Altered by Protein Binding. *Imino resonances.* Imino resonances in B-DNA ordinarily occur in discrete A·T and G·C spectral regions due to differences in intrinsic ring currents (10, 14). Such separation is observed in the ^1H NMR spectrum of the free SRY-binding site (Fig. 5, spectrum A, and Table 2). With successive addition of SRY-p a second set of imino resonances is observed in slow exchange (arrows in Fig. 5, spectrum B), leading to the appearance of a distinct bound-state spectrum at 1:1 stoichiometry (Fig. 5, spectrum D). No further changes are observed at higher stoichiometry. Retention of downfield imino resonances in the bound state demonstrates that base pairing is maintained at each step, including the site of isoleucine insertion.

Complexation shifts in imino resonances (Table 2) suggest sites of structural distortion. The largest such shift occurs at the site of isoleucine insertion: the T^8 imino resonance exhibits an upfield perturbation of 0.93 ppm. Its resonance position in the complex (12.58 ppm) is anomalous for an A·T base pair. Although its structural basis is not clear, the upfield T^8 chemical shift is reminiscent of those observed at intercalation sites of drug–DNA complexes and ascribed to the ring current of the intercalator (15, 16). Since isoleucine is not aromatic, the T^8 complexation shift must reflect a different

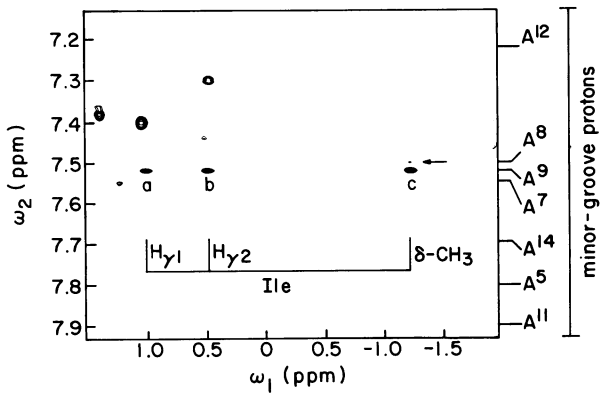

FIG. 4. Region of NOESY spectrum in ^2H_2O showing NOEs between an upfield isoleucine spin system (ω_1, horizontal axis) and $A^{8'}$ and $A^{9'}$ H2 resonances (ω_2, vertical axis). Chemical shifts of adenine H2 resonances are indicated at right. Strong NOEs to $A^{9'}$ H2 are labeled a–c; arrow indicates weak NOE between isoleucine δ-CH_3 and $A^{8'}$ H2. Mixing time was 75 ms and temperature was 40°C.

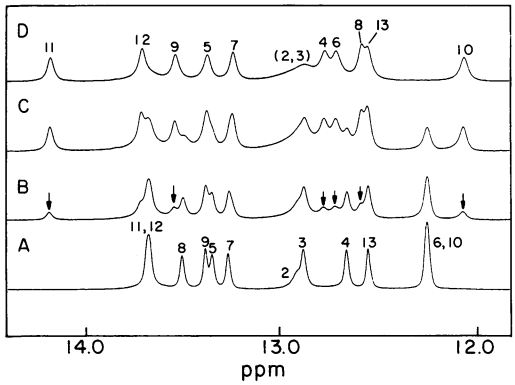

FIG. 5. ^1H NMR spectrum of thymine N_3H and guanine N_1H imino resonances of SRY-binding site (5′-GGGGTGATTGTTCAG-3′) at 40°C and pH 6.0: Spectra: A, free DNA; B, 1:4 molar ratio of SRY-p to DNA; C, 1:2 molar ratio of SRY-p to DNA; D, equimolar SRY-p/DNA complex. Arrows in spectrum B indicate emergence of bound-state spectrum in slow exchange. Assignments are as indicated in Table 2. Terminal and penultimate imino resonances are broadened at 40°C and pH 6.0 by fraying.

mechanism, such as weakening of the T^8·$A^{8'}$ imino hydrogen bond.

A corresponding anomaly is observed at T^8 in the pattern of NOEs between imino protons. Such NOEs, which ordinarily occur between successive base pairs, are readily observed in the free DNA and are retained in the complex with the exception of the T^8-T^9 step. The latter is ≈8-fold weaker than that of the other T-T step (T^{11}-T^{12}) and 4-fold weaker than those of neighboring base pairs $T^{7'}$-T^8 and T^9-G^{10} (mixing times of 75 and 150 ms at 15°C). This observation indicates an increase in spatial separation by ≈1 Å (4.0–5.0 Å): base stacking is interrupted at the T^8-T^9 step in accord with partial insertion of a protein side chain. Retention of a weak NOE is inconsistent with complete intercalation, which would separate successive imino protons by >6.5 Å. (Attenuation of an imino NOE can also arise from an increase in its rate of exchange with H_2O. This mechanism is not operative here, as the rate of imino–solvent exchange is reduced in the complex.)

Adenine H2 resonances. Retention in the complex of strong thymidine imino–adenine H2 NOEs excludes Hoogsteen base pairing. Adenine H2 chemical shifts are sensitive to base stacking, especially at 5′ A-A or 5′ A-G steps. The 5′ H2 proton overlies the six-membered portion of the 3′ flanking purine, giving rise to a large upfield ring

Table 2. Imino and adenine H2 chemical shifts (ppm at 40°C and pH 6.0)

Base pair	Imino resonances			Adenine H2 resonances		
	Free	Bound	Δ	Free	Bound	Δ
3	12.88	12.87	−0.01	—	—	—
4	12.66	12.77	0.11	—	—	—
5	13.35	13.37	0.02	7.59	7.81	0.22
6	12.25	12.71	0.46	—	—	—
7	13.27	13.24	−0.03	7.61	7.56	−0.05
8	13.51	12.58	**−0.93**	7.53	7.50	−0.03
9	13.39	13.53	0.14	6.96	7.53	**0.57**
10	12.25	12.06	−0.19	—	—	—
11	13.68	14.18	0.50	7.55	7.88	0.33
12	13.68	13.70	0.02	7.26	7.22	−0.04
13	12.55	12.55	0.00	—	—	—

Chemical shifts are relative to trimethylsilylpropionate at 0.00 ppm. Differences (Δ) shown are bound − free. Respective upfield and downfield shifts in T^8 imino and A^9 H2 resonances are shown in boldface.

current (>0.5 ppm); such overlap is not expected with the 5′ flanking base (Fig. 6*A*; ref. 17). The SRY-binding site contains two 5′ A-A (and no 5′ G-A) steps, 5′ $A^{9'}$-$A^{8'}$ and 5′ $A^{12'}$-$A^{11'}$. As expected, the H2 resonances of $A^{9'}$ and $A^{12'}$ in the free DNA are upfield (6.96 and 7.26 ppm, respectively) of the remaining five H2 resonances (including those of $A^{8'}$ and $A^{11'}$; Table 2). Interruption in $A^{8'}$-$A^{9'}$ stacking in the complex predicts a large downfield shift in the 5′ adenine H2 ($A^{9'}$) but not in the 3′ adenine H2 ($A^{8'}$). This is indeed observed. Smaller downfield shifts are also observed in the H2 resonances of $A^{5'}$ and $A^{11'}$ whose structural origins are not presently understood.

Partial Intercalation Rationalizes Sequence Specificity. *"Wedge intercalation" model.* Our ^{1}H NMR observations demonstrate partial insertion of an aliphatic side chain through the minor groove. Such insertion occurs at only one step and separates flanking base pairs (Fig. 6*B*). Although structural details await determination of a complete three-dimensional structure, it is likely that the flanking base pairs must also tilt (i.e., increase in their roll angle) and partially unwind to accommodate the inserted side chain. This mechanism (designated "wedge intercalation") predicts a kink in the DNA site—i.e., that protein-induced bending [as monitored by gel electrophoresis (11, 12)] is discontinuous. We postulate that sequence specificity at the kinked site has a topological basis: that only T^8-T^9 (or less well, C^8-T^9) permits accommodation of the protein side chain within a distorted DNA structure (Fig. 6*B*).

The wedge-intercalation model rationalizes the relative DNA-binding affinities of variant sites (Table 1, transition and transversions). The structures of A·T and G·C base pairs differ in the minor groove only by the absence (A) or presence (G) of a 2-amino group. Absence of specific binding following transition $T^9 \rightarrow C$ presumably reflects steric restriction: strong NOEs between $A^{9'}$ H2 and the isoleucine side chain suggest that the corresponding $G^{9'}$ 2-amino group would block side-chain penetration. In contrast to position 9, the NOE between $A^{8'}$ H2 and the isoleucine is weak and involves only the δ-CH_3. This difference suggests that analogous placement of a guanine 2-amino group would be better tolerated. This is also observed (Table 1, transitions and transversions). Transversions are presumably disallowed because a base pair is asymmetric: interchange of purines and pyrimidines shifts the junction between bases across the helical axis and alters allowed patterns of base-pair roll and propeller twist (17).

One feature of this model is that specificity for the native T^8-T^9 sequence is attributed not to favorable base-specific contacts (direct readout; refs. 1–3) but instead to exclusion of alternative sequences. The native sequence may be distinguished not by its *affinity* for isoleucine (in an appropriate structural context) but by its incurring the *least possible penalty* for side-chain intercalation. This model predicts that SRY-p will bind well to DNA analogs that differ in sequence but share the ability to accommodate side-chain penetration without additional penalty. Pyrimidine–pyrimidine mismatches at positions 8 or 9, for example, would be expected to be well-tolerated regardless of sequence. This surprising prediction is indeed observed (Table 1, mismatches, and Fig. 1, lanes e and g). In contrast, G·G and A·A mismatches at either position prevent specific binding (Table 1, mismatches, and Fig. 1, lanes f and h). The structures of these mismatches in the complex have not been investigated and are of future interest. Because in the native complex base pairing and stacking are maintained at the 7-8 and 9-10 steps, extrahelical displacement of the pyrimidine mismatches is unlikely. Although T·T mismatches can form a T(anti)·T(anti) wobble pair, analogous C(anti)·C^+(anti) pairing would require protonation (19).

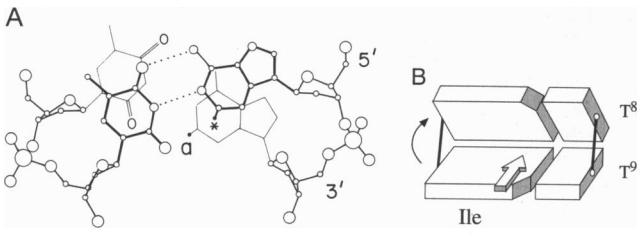

FIG. 6. (*A*) Crystal structure of 5′ A-A step in B-DNA dodecamer (17). Base stacking positions the 5′ adenine H2 (asterisk) over the 3′ purine (thin line); in contrast, the 3′ adenine H2 (a) is askew of the 5′ purine (thick line). (*B*) Model of isoleucine insertion site. Although base pairing is maintained, partial intercalation requires widening of the minor groove and an increase in roll between successive base pairs. The latter is predicted to induce a sharp bend in the DNA. Adapted from ref. 18.

DISCUSSION

Directed DNA bending can regulate gene expression by altering spatial relationships among DNA-bound transcription factors (20, 21). Of particular interest are HMG-box proteins (5), whose α-helical structure (6) preferentially binds to altered DNA topologies (four-way junctions, supercoiled DNA, and cisplatin–DNA adducts; refs. 7 and 22). Unlike classical transcription factors, the HMG proteins UBF and mtTF, for example, bind at specific promoter positions but not at specific promoter sequences (23, 24). SRY exhibits both sequence-specific binding to duplex DNA and sequence-independent binding to oligonucleotide models of four-way junctions (11, 25, 26).

The SRY-p/DNA complex exhibits unusual ^{1}H NMR features. Complexation shifts are larger than those observed in other protein–DNA complexes, such as the *lac* repressor headpiece (27), Antaennapedia homeodomain (28), the λ Cro repressor (29), and GAL4 binuclear cluster (29). The latter exhibit specific contacts between side chains and edges of base pairs, primarily in the major groove of B-DNA (1–3). In contrast, biochemical studies have established that the SRY HMG box binds primarily in the minor groove and induces a 90° bend (11–13). The present ^{1}H NMR analysis demonstrates partial intercalation of a nonpolar side chain in the minor groove, disrupting base stacking but not base pairing (Fig. 6*B*). The minor groove is presumably widened and displaced of hydration spines (17). An analogous mechanism has been proposed to mediate DNA binding by 3β,17β-dipyrrolidin-1′-yl-5α-andostane (dipyrandium iodide; ref. 30).

We have also investigated SRY-p binding to DNA sites containing transitions, transversions, and base mismatches flanking the intercalation site (A·T base pairs 8 and 9). Our results suggest that specificity is based not on hydrogen bonds to the central base pairs but on a topological criterion: that the central sequences of high-affinity DNA sites incur the least free-energy penalty in accommodating wedge intercalation (Figs. 6*B* and 7*A*). This penalty must be paid by favorable interactions elsewhere in the complex, presumably including salt bridges and hydrogen bonds to a bent DNA backbone. Topological recognition of an altered DNA structure may resolve the seeming paradox of sequence- and structure-specific recognition. A four-way junction (31) exhibits structural similarities to the structure of duplex DNA in a specific complex (Fig. 7*A*) and, in particular, may mimic a specific spatial distribution of backbone charges (Fig. 7*B*). Partial unstacking of base pairs at the center of the junction may facilitate analogous side-chain insertion. Like the SRY-p complex, a four-way junction retains continuous base pairing but is predicted to exhibit a widened minor groove (31).

Induction and recognition of noncanonical DNA structures may be a general feature of transcription. An outstanding

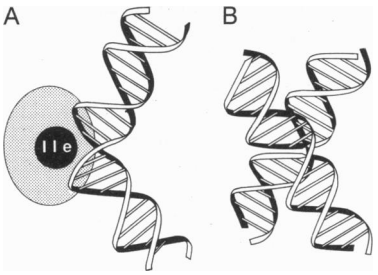

FIG. 7. (*A*) Wedge-intercalation model of a bent SRY–DNA complex. An isoleucine side chain in the DNA minor groove partially inserts between tilted base pairs (A·T base pairs 8 and 9). By analogy to HMG1 (6), the isoleucine occurs in an α-helix. (*B*) A four-way junction (31) mimics the structure of a bent DNA site in a specific SRY complex, permitting sequence-independent recognition (11).

example is provided by the TATA-binding protein TBP (TFIIDτ). Although unrelated to the HMG-box family in sequence or structure (5, 6), TBP also recognizes A·T base pairs in the minor groove and bends its target site (32). Mutations that alter sequence specificity (33) involve nonpolar residues in its putative DNA-binding surface (34). By analogy to SRY-p, we speculate that nonpolar or aromatic residues partially insert between base pairs of a widened minor groove. Partial intercalation may facilitate isomerization of the preinitiation complex and recruitment of eukaryotic RNA polymerases (see *Note Added in Proof*).

Concluding Remarks. The present studies establish a third mechanism of protein–DNA recognition. The SRY HMG box recognizes the central TT element of a bent DNA site by partial intercalation of a nonpolar side chain (isoleucine). Minor-groove intercalation disrupts base stacking but not base pairing. Although the isoleucine has not been assigned, the two possible assignments each correspond to α-helices in the solution structure of HMG1 (6). This homology predicts that SRY inserts an α-helix in a widened minor groove, extending the role of the α-helix in DNA recognition (1, 2).

What is the role of "wedge intercalation" in specification of the male body plan? Genetic analysis of a human intersex abnormality (46,XY pure gonadal dysgenesis) has demonstrated that mutation of either isoleucine in the SRY HMG box—*including therefore the intercalating side chain*—is associated with phenotypic sex reversal (26, 35). We propose that failure of DNA penetration by the variant side chain alters the specificity or DNA-bending properties of the mutant SRY. This hypothesis may be tested by comparative biochemical and structural studies of mutant SRY–DNA complexes. Such studies will provide an opportunity to define the molecular mechanism of a genetic switch in mammalian development.

Note Added in Proof. The predicted analogy between partial side-chain intercalation in SRY and TBP has recently been demonstrated in cocrystal structures of TBP–DNA complexes (36, 37).

We thank Prof. P. K. Donahoe for support, discussion, and communication of results prior to publication; C. M. Haqq for advice and assistance with DNA-binding studies; J. P. Lee and Q. X. Hua for assistance with NMR measurements; A. S. Stern and J. C. Hoch for ring-current calculations; D. C. Page for SRY cDNA; and C. O. Pabo for discussion. This work was supported by the National Institutes of Health (HD26465) and Council for Tobacco Research.

1. Steitz, T. A. (1990) *Q. Rev. Biophys.* **23,** 205–280.
2. Pabo, C. O. & Sauer, R. T. (1992) *Annu. Rev. Biochem.* **61,** 1053–1095.
3. Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A. & Sigler, P. B. (1988) *Nature (London)* **335,** 321–329.
4. Koopman, P., Munsterberg, A., Capel, B., Vivian, N. & Lovell-Badge, R. (1990) *Nature (London)* **348,** 450–452.
5. Laudet, V., Stehelin, D. & Clevers, H. (1993) *Nucleic Acids Res.* **21,** 2493–2501.
6. Weir, H. M., Kraulis, P. J., Hill, C. S., Raine, A. R. C., Laue, E. D. & Thomas, J. O. (1993) *EMBO J.* **12,** 1311–1319.
7. Lilley, D. M. J. (1992) *Nature (London)* **357,** 282–283.
8. Haqq, C. M., King, C.-Y., Donahoe, P. K. & Weiss, M. A. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 1097–1101.
9. Smallcombe, S. H. (1993) *J. Am. Chem. Soc.* **115,** 4776–4785.
10. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids* (Wiley, New York).
11. Ferrari, S., Harley, V. R., Pontiggia, A., Goodfellow, P. N., Lovell-Badge, R. & Bianchi, M. E. (1992) *EMBO J.* **11,** 4497–4506.
12. Giese, K., Cox, J. & Grosschedl, R. (1991) *Genes Dev.* **5,** 2567–2578.
13. van der Wetering, M. & Clevers, H. (1992) *EMBO J.* **11,** 3039–3044.
14. Bell, R. A., Everett, J. R., Hughes, D. W., Coddington, J. M., Alkema, D., Hader, P. A. & Neilson, J. (1985) *J. Biomol. Struct. Dyn.* **2,** 693–707.
15. Zhang, X. & Patel, D. J. (1991) *Biochemistry* **30,** 4026–4041.
16. Gilbert, D. E. & Feigon, J. (1991) *Biochemistry* **30,** 2483–2494.
17. Dickerson, R. E. & Drew, H. R. (1981) *J. Mol. Biol.* **149,** 761–786.
18. Calladine, C. R. & Drew, H. R. (1992) in *Molecular Structure and Life: Molecular Recognition of Nucleic Acids*, eds. Kyogoku, Y. & Nishimura, Y. (Japan Sci. Soc. Press, Tokyo), pp. 43–55.
19. Kouchakdjian, M., Li, B. F. L., Swann, P. F. & Patel, D. J. (1988) *J. Mol. Biol.* **202,** 139–155.
20. Ptashne, M. (1986) *Nature (London)* **322,** 697–701.
21. Kerppola, T. K. & Curran, T. (1991) *Cell* **66,** 317–326.
22. Pil, P. M. & Lippard, S. J. (1992) *Science* **256,** 234–237.
23. Bell, S. P., Pikaard, C. S., Reeder, R. H. & Tjian, R. (1989) *Cell* **59,** 489–497.
24. Fisher, R. P., Parisi, M. A. & Clayton, D. A. (1989) *Genes Dev.* **3,** 2202–2217.
25. Nasrin, N., Buggs, C., Goebl, M. & Alexander-Bridges, M. (1991) *Nature (London)* **354,** 317–320.
26. Harley, V., Jackson, D. I., Hextall, P. J., Hawkins, J. R., Berkovitz, G. D., Sockanathan, S., Lovell-Badge, R. & Goodfellow, P. N. (1992) *Science* **255,** 453–456.
27. Lamerichs, R. M. J. N., Boelens, R., van der Marel, G. A., van Boom, J. H., Kapstein, R., Buck, F., Fera, B. & Ruterjan, H. (1989) *Biochemistry* **28,** 2985–2991.
28. Otting, G., Qian, Y. Q., Billeter, M., Müller, M., Affolter, M., Gehring, W. & Wüthrich, K. (1990) *EMBO J.* **9,** 3085–3092.
29. Shirakawa, M., Matsuo, H., Serikawa, Y., Suzuki, N. & Kyogoku, Y. (1992) in *Molecular Structure and Life: Molecular Recognition of Nucleic Acids*, eds. Kyogoku, Y. & Nishimura, Y. (Japan Sci. Soc. Press, Tokyo), pp. 129–140.
30. Patel, D. J. & Canuel, L. L. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 24–28.
31. Lilley, D. M. J. & Clegg, R. M. (1993) *Annu. Rev. Biophys. Biomol. Struct.* **22,** 299–328.
32. Starr, D. B. & Hawley, D. K. (1991) *Cell* **67,** 1231–1240.
33. Strubin, M. & Struhl, K. (1992) *Cell* **68,** 721–730.
34. Nikolov, D. B., Hu, S.-H., Lin, J., Gasch, A., Hoffman, A., Horikoshi, M., Chua, N.-H., Roeder, R. G. & Burley, S. K. (1992) *Nature (London)* **360,** 40–46.
35. Vilain, E., Jaubert, F., Fellous, M. & Mclreavey, K. (1993) *Differentiation* **52,** 151–159.
36. Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. (1993) *Nature (London)* **365,** 512–520.
37. Kim, J. L., Nikolov, D. B. & Burley, S. K. (1993) *Nature (London)* **365,** 520–527.