Supplemental Information

for

SRY and the Enigma of Inherited Human Sex Reversal
FRUSTRATED INDUCED FIT IN A BENT PROTEIN-DNA COMPLEX

Nelson B. Phillips, Joseph Racca, Yen-Shan Chen, Rupinder Singh, Agnes Jancso-Radek,

James T. Radek, Nalinda P. Wickramasinghe, Elisha Haas, & Michael A. Weiss*

Table of Contents

1

Purpose of Supplement

This Supplement presents a more detailed description of the methods, including pertinent equations employed in modeling, and contains 15 figures and 3 tables. For clarity, references cited in the Supplement are independently numbered at the end of this document.

Supplemental figures are intended to provide both additional structural perspectives and additional supporting data. In Figure S1 is shown a ribbon model of the HMG box (front view and side view) to clarify the locations of the major and minor wings in relation to mutation site V60 (position 5 of the box). In Figure S2 is shown a molecular model of the wild-type SRY HMG box/DNA complex with variant core target site 5'-TTTGTG-3' (and complement) in which the hydrophobic wedge (including cantilever side chain I68; box position 13) and role of the basic tail are highlighted. Structural relationships in the SRY minor wing and at its interface with a specific DNA site are outlined in Figure S3.

Additional data are provided in Figures S4-S13. Figure S4 provides a gel-mobility-shift assay (GMSA) showing activity of the V60L and V60A domains in relation to the wild-type HMG box. In Figure S5 is shown a Monte Carlo simulation of the fluorescent donor (FAM) and fluorescent acceptor (TAMRA) flexibly tethered to the respective 5' ends of a 15-bp B-DNA site. Additional 1D and 2D NMR spectra are provided in Figures S6-S12. Of particular note is evidence of corresponding cantilever insertion sites (Fig. S10) and corresponding solvent exchange by imino DNA protons (Fig. S12) in the wild-type and V60L variant DNA complexes. Figure S13 contains the results of control cell-culture studies of *Sox9* transcriptional activation as a function of the amount of plasmid DNA employed in the transient transfections.

Whereas the "front" surface of the SRY HMG box contacts DNA, it is possible that part or all of the "back" surface may contact other proteins in an SRY-containing enhanceosome and transcriptional pre-initiation complex. Figure S14 illustrates the role of V60 (box position 5) or its homolog (Ile) in the HMG box of LEF-1 at the back surface of the minor wing. DNA bending by HMG boxes of wild-type SRY and V60L SRY with a variant DNA site is illustrated in Figure S15.

Results of the application of a two-state model to the guanidine-induced unfolding transitions of wild-type or variant SRY HMG boxes are given in Table S1. The dependence of the free energy of unfolding (ΔG_u) on pH and salt concentration is given in Table S2. Ring-current simulations of the SRY HMG box are provided in Table S3 to highlight the predicted upfield shift in chemical shift of the resonances of V60 on DNA-dependent folding of the minor wing; such shifts are due to insertion of the aliphatic side chain into an "aromatic box" (see main text).

Detailed Methods

DNA-Binding Assay—The DNA fragments were purchased from Oligos, Etc. (Wilsonville, OR). One strand was labeled with [γ-^{33}P]-ATP (ICN Biomedicals, Inc.). The 36-bp duplex probe was annealed and analyzed by GMSA as described (1,2). Each reaction contained 0-80 nM protein and less than 1 nM labeled DNA in 10 mM potassium phosphate (pH 7.0), 30 μg/ml bovine serum albumin, 50 mM KCl, 2.5 mM MgCl$_2$, 5% glycerol and 4 mM dithiothreitol; the reaction was incubated in 12.5 μl reactions for one hour on ice. Only specific binding is observed under these conditions at the protein concentrations tested. Polyacrylamide gels (8%) containing 0.45X Tris-borate-ethylenediamine-tetra-acetic acid (TBE) were pre-run for 45-60 minutes at 10 V/cm at 4 °C. After sample loading, the gels were run at 15 V/cm for one hour at 4

^{o}C.   The temperature of the gel during the course of electrophoresis was kept constant by a circulating refrigerated water bath.

DNA-Bending Assay— Double-stranded synthetic oligonucleotides containing consensus SRY binding site 5'-ATTGTT-3' were cloned between the *XbaI* and *SalI* sites of the circular permutation vector pBend2 (3). Probes of equal length (147 base pairs), with the binding site at varying distance from the ends (120, 95, 79, 51, 47, and 27 base pairs) were generated by PCR, and 5'-labeled with [γ-^{33}P]ATP and T4 polynucleotide kinase. 10 μl binding reactions contained 50 mM KCl, 20 mM Tris-HCl pH 7.4, 5 mM MgCl$_2$, 50 ng poly (dI-dC), 10% glycerol, approximately 1 nM ^{33}P-labeled DNA probe and purified SRY-HMG box protein (60 nM). After 1 - 2 hr incubation on ice, samples were run on polyacrylamide gels in 0.5X (0.045 M) TBE buffer at ≈ 10 V/cm. To assess reproducibility of PGE estimates of DNA bend angle in different gels of same composition, the wild-type SRY domain was assayed six times in the 10%/29:1 system. Inferred angles were 77.8°, 78.5°, 78.6°, 79.2°, 79.2°, and 80.5°, giving mean value of 79° and standard deviation of ± 1°. Prior studies of SRY have established that PGE estimates of induced DNA bend angle are in accord with independent estimates of end-to-end distances in the DNA duplex based on phase-modulation fluorescence resonance energy transfer (4). Absolute calibration of PGE-based estimates of DNA bend angles is not well established in relation to three-dimensional structures. A DNA bend angle of 54.0^o ± 1.7^o was observed in the NMR structure of a specific complex with a variant DNA site (5'-GTG**TTTGTG**CA-3' and complement; core site underlined with non-consensus bases in bold (5)).

Transient Transfection—Transfections were performed using reagent Fugene6 (Hoffman LaRoche, Ltd., Nutley, NJ). After two washes with phosphate-buffered saline (PBS), constructs were transfected into CH34 cells as described by the vendor. The cell lines were recovered in fresh Dulbecco's modified Eagle medium with 5% heat-inactivated fetal calf serum after 24 h in serum-free medium. Expression plasmid pCMX-SRY was introduced at an initial plasmid concentration of 1 μg per 3 μl (Fugene6), as added to a well containing 10^6 cells. Serial dilutions of the wild-type or variant plasmid were undertaken (in the presence of a compensating dose of the empty plasmid to maintain a constant total DNA concentration) to reduce the level of SRY expression in transfected cells. Transfection efficiency was determined using a pCMX-GFP construct. Cells were co-transfected with SRY and pCMX-GFP at equal amounts and efficiency measured as a ratio of GFP positive cells. Extent of wild-type or variant SRY expression was assessed by western blot using an antiserum directed against the hemaglutinin tag (Sigma-Aldrich, Inc., St. Louis, MO). At each dilution the level of expression of the variant SRY proteins was similar to that of the wild-type.

Protein Stability. Fluorescence-detected guanidine titrations were monitored at 4 ^{o}C as measured at an emission wavelength of 390 nm (slit width 5 nm) following excitation at 270 nm (slit width 2 nm). For such studies wild type and variant domains were made 1 μM in 140 mM KCl and 10 mM potassium phosphate (pH 7.4) in the titrating cuvette. In CD- and fluorescence-based assays the same concentration of SRY domain was used in the titrant reservoir containing 7.2 M guanidine-HCl in the above buffer. CD- and fluorescence-based denaturation transitions were observed to be coincident.

Thermodynamic Modeling—Experimental denaturation curves (as monitored by CD or fluorescence) were in each case fitted by non-linear least squares to the equations $\Delta G = \Delta G_u + m \times [\text{guanidine}] = -RT \ln K$, where ΔG_u is the free energy of unfolding extrapolated to zero denaturant concentration, and K is the equilibrium constant between native and unfolded states

of the protein (4,6) In brief, the data, $\theta(x)$ where x indicates the concentration of denaturant, were fitted by a nonlinear least-squares program according to

$$\theta(x) = \frac{\theta_A + \theta_B e^{(-\Delta G^o_{H_2O} - mx)/RT}}{1 + e^{-(\Delta G^o_{H_2O} - mx)/RT}}$$

where x is the concentration of guanidine hydrochloride and where θ_A and θ_B are baseline values in the native and unfolded states. These baselines were approximated by pre- and post-transition lines $\theta_A(x) = \theta_A^{H_2O} + m_A x$ and $\theta_B(x) = \theta_B^{H_2O} + m_B x$. Fitting the original CD or fluorescence data and baselines simultaneously circumvents artifacts associated with linear plots of ΔG as a function of denaturant according to $\Delta G^o(x) = \Delta G^o_{H_2O} + m^o x$ (7).

FRET Sample Preparation—Fluorescence studies of DNA bending employed a 15-bp DNA duplex of sequence 5'-TCGGTGATTGTTCAG-3' ("upper strand") and complement (consensus target site underlined). The site of cantilever insertion is therefore between base pairs 8 and 9, i.e., near the center of the oligonucleotide. Use of a 15-bp DNA site restricts binding to the high-affinity 1:1 complex as verified by GMSA Three samples of labeled DNA were prepared: (i) containing 5'-fluorescein tethered to the 5'-phosphate of the upper strand by a hexanyl linker (6-FAM); (ii) tetramethylrhodamine tethered to the 5'-phosphate of the lower strand by an analogous linker (TAMRA); and (iii) a double-labeled derivative. HPLC-purified oligonucleotides were purchased from Qiagen Operon, (Alameda, CA) and purified further by HPLC after annealing using an ion-exchange semi preparative HPLC column (DNA PAC PA-100, Dionex Corp., Sunnyvale, CA). For FRET study, the duplex DNA probes were made 5 µM in 10 mM potassium phosphate and 10 mM Tris-HCl (pH 8.4) containing 140 mM KCl, 1 mM EDTA and 1 mM DTT. The ratio of labeling in the double-labeled sample is close to 1:1 (see Results).

FRET-based Dissociation Constants—K_d measurements were made at 15 °C. Varying SRY concentrations were titrated in the FRET buffer described above at a constant DNA concentration of 25 nM. Emission spectra were recorded from 500-650 nm after excitation at 490 nm. K_d values were determined by plotting change in fluorescence intensity at 520 nm against total protein concentration. Data were fit with a single-site ligand-binding equation as implemented in *Origin 8.0* (OriginLab Corp., Northampton, MA).

$$\Delta F = \Delta F_o\{0.5(1 + S/D_o + K_d/D_o)$$

$$- [0.25(1 + S/D_o + K_d/D_o)^2 - S/D_o]^{0.5}\}$$

In this formalism ΔF is the change in donor fluorescence observed on addition of the SRY domain to DNA relative to the DNA fluorescence in the absence of the protein; ΔF_0 is the maximum fluorescence change obtained when the DNA is fully complexed; K_d is the dissociation constant; D_0 is the concentration of DNA (25nM); and S is the concentration of SRY domain.

Time-resolved fluorescence measurements—A home-built system at Bar Ilan University employs time-correlated single-photon counting. The excitation source has a picosecond mode-locked Ti:sapphire laser (Tsunami, Spectra-Physics, Santa Clara, CA) pumped by a high-power diode laser (Millenia XsJ Spectra-Physics, Santa Clara, CA) and equipped with broadband optics. The output frequency of the laser was multiplied by a flexible second- and third harmonics generator (GWU, Spectra-Physics). The laser pulse width was 1.6 ps before doubling. A pulse selector (Model 3980 Spectra-Physics) was used to reduce the basic 80 MHz pulse rate to 4 MHz with energies of up to 0.5 nJ/pulse (2 mW at 4 MHz). The second harmonic

was used for excitation at 490 nm. The emission was collected with a polarizer at a magic angle relative to the excitation polarizer at 520 nm. Measurements were done at 4 $^{\circ}$C. The emission wavelength was selected by a double 1/8 meter subtractive monochromator, the emission slit width was 8 nm (DIGIKROM CM112) and directed to the surface of cooled micro-channel plate PMT (MCP-PMT, Hammamatsu 3809U-50) biased at -3200 V. The acceptor emission was collected at 585 nm (emission bandwidth 32 nm). A single-photon counting board (SPC 630, Becker & Hickel GmbH) fed via a preamplifier (HFAC-26DB 0.1UA) and triggered by a Photodiode (PHD-400N) was used. The response of the system yielded a pulse width of 35ps.

*Global Analysis of Time-Resolved FRET—*Distance distribution functions were obtained from simultaneous global analysis of four experimental fluorescence decay curves: these were recorded for (i) samples containing only the donor, (ii) samples containing only the acceptor, (iii) donor- and acceptor decay curves from samples containing both fluorophores. Global analyses of the fluorescence decay curves were performed using the Marquardt nonlinear least-squares method (8). Four (or two) theoretical decay curves were calculated for each set of FRET experiments and fitted to the corresponding experimental curves. Calculated decay curves were prepared by numerical solution of a modified version of the second-order differential equation (8; for review see Haas 1996 ref 47 in main text)

$$\frac{\partial p_i(r,t)}{\partial t} = D\frac{\partial}{\partial r}\{\exp(-\frac{U(r)}{k_BT})\frac{\partial}{\partial r}[\exp\frac{U(r)}{k_BT}p_i(r,t)]\} - k_i(r)p_i(r,t)$$

In this equation $p_i(r,t)$ is the probability density to find an excited state donor (of the fluorescence lifetime component τ_i and normalized preexponential factor α_i) with distance r from the acceptor at time t after excitation, so that the total probability density is

$$p(r,t) = \sum_{i=1}^{n}\alpha_i\,p_i(r,t)$$

In partial differential equation D is the intramolecular segmental diffusion constant. This was fixed to zero (*i.e.,* no diffusion was assumed), since in the first set of experiments the analyses gave zero diffusion. $U(r)$ is a distance-dependent potential of mean force, k_B is the Boltzmann constant, and T is the absolute temperature. This potential is usually taken to be harmonic (thus implying a Gaussian distribution for the distance between the fluorescent probes): $U(r) = a(r-b)^2$, where a and b are parameters. $k_i(r)$ is the reaction term including the spontaneous emission rate and the Förster energy transfer rate:

$$k_i(r) = \frac{1}{\tau_i} + \frac{1}{\tau_r} \times \frac{9\ln10}{128\pi^5 n^4 N_A} \times \kappa^2 \times J \times \frac{1}{r^6},$$

where τ_r is the radiative lifetime of the donor, n is the refractive index of the medium between the donor and acceptor, N_A is Avogadro's number, κ^2 is the orientation factor, and J is the normalized spectral overlap integral. Evaluation of each analysis and the significance of inferred parameters were based on four indicators: the global χ^2 values; the distributions of the residuals; the autocorrelation of the residuals and the error intervals of the calculated parameters (8,9). The error intervals were obtained by a rigorous analysis procedure carried out for each set of experiments.

*Photophysical Control Studies—*Control steady-state spectroscopic experiments were undertaken to test the suitability of the present samples for tr-FRET analysis of distributions of end-to-end distances as described (10). In brief, the possibility of intermolecular FRET was excluded in control studies of solutions containing equimolar mixtures of donor-labeled DNA

6

and acceptor-labeled DNA (or the wild-type protein-DNA complexes). Further, intramolecular FRET efficiencies of the free and bound DNA sites are independent of concentration of DNA (or wild-type protein-DNA complex) in the range 1-15 μM. Steady-state anisotropies of the donor and acceptor attached to the DNA duplex were measured for the free and protein-bound sites. For the singly-labeled free DNA the respective anisotropies of the donor and acceptor were 0.10 and 0.23 at 4 $^{\circ}$C. As expected, in the double-labeled DNA site the anisotropy of the donor increased to 0.13 due to the reduction of the fluorescence lifetime; the anisotropy of the acceptor was unchanged. Formation of the human and V60L variant complexes in each case caused a *ca.* 30% increase in donor emission anisotropy and 20% increase in acceptor emission anisotropy, presumably reflecting the further reduction of the donor lifetime and changes of DNA conformation on binding of the wild-type SRY domain. The reduction of the fluorescence intensity of the donor in the double-labeled DNA site is thus a result of FRET not a spurious quenching process (such as end-to-end DNA stacking). Further, the spectroscopic properties of the probes are not perturbed by the bound protein. Determination of intramolecular distances based on averaged orientation factors (11) was justified by the low steady-state emission anisotropy of the donor and the partial anisotropy of the acceptor emission in the free and bound DNA site.
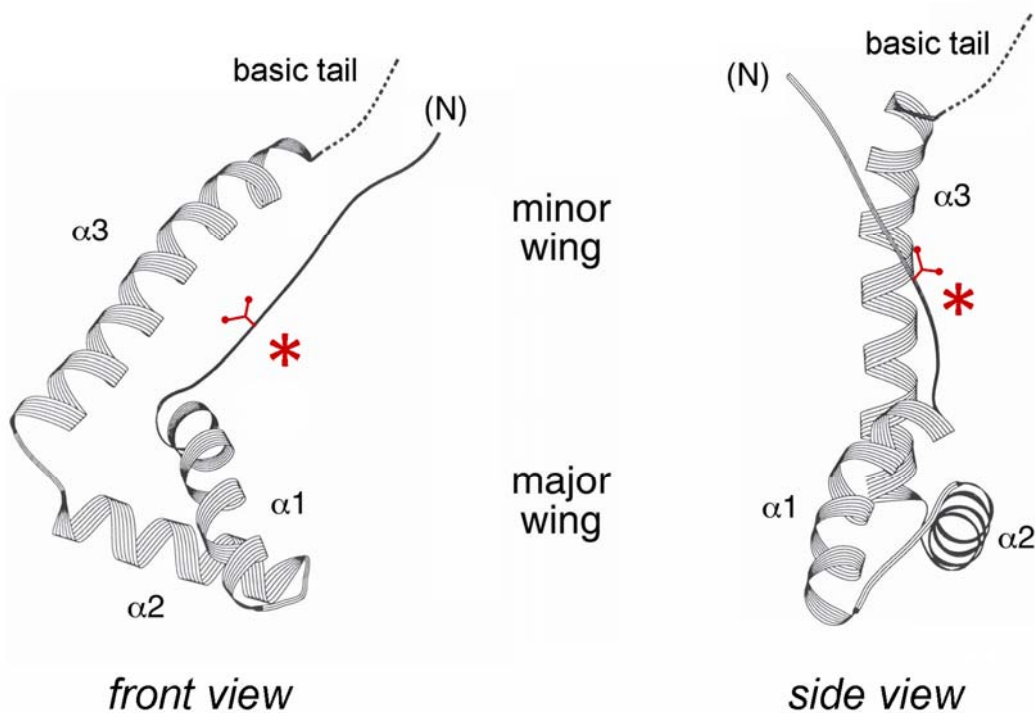
Figure S1. **Structure of HMG box and position of V60 at consensus position 5**.   Ribbon model showing L-shaped structure of HMG box. The motif contains a major wing (consisting of α-helix 1, α-helix 2, the first turn of α-helix 3, and connecting loops) and minor wing (consisting of the N-terminal α-strand, the remainder of α-helix 3, and the C-terminal segment). Dotted line indicates basic tail, which is disordered in the free domain but which extends the protein-DNA interface within DNA minor groove. A front view is shown at left, and side view at right. Asterisks indicate position of V60 (red side chain).

Figure S2. **Structure of SRY HMG box/DNA complex**. SRY-DNA complex (49 in main text; protein databank accession code 1J46). Ribbon model of protein is shown in blue (HMG box) and red (tail; residues 70-85). DNA is shown as black sticks. Cantilever side chain I68 (consensus position 13) and wedge side chains F67 and W98 (consensus positions 12 and 43, respectively) are shown in green at center of complex; asterisk indicates M64 (consensus position 9). *De novo* mutation M64I; (12) impairs nuclear import and, in the context of a truncated HMG box lacking part of the basic tail, impairs specific DNA bending (see main text for reference: (13)).
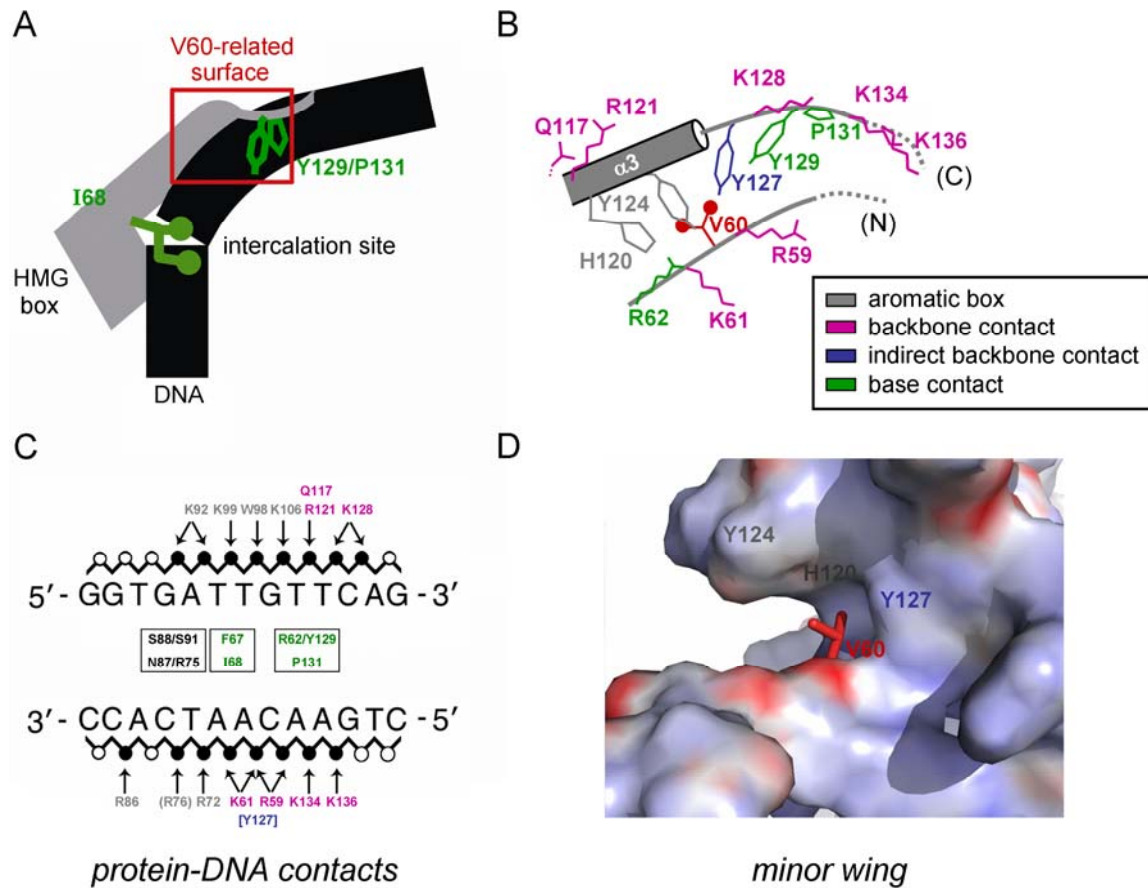
Figure S3. **Overview of SRY minor wing and role in DNA binding and DNA bending**. (*A*) Schematic model of the L-shaped HMG box (gray) and bent DNA site (black). The two principal DNA-bending surfaces (*green*) are provided by a hydrophobic wedge surrounding the cantilever side chain I68 (box residue 13) and the V60-related surface of the minor wing (*red box*). (*B*) Schematic depiction of the packing of V60 within an aromatic box defined by the side chains of H120, Y124, and Y127. Side chains are color-coded as defined in the inset box. The color code is the same as in panels *A* and *B*. (*C*) Summary of protein-DNA contacts as predicted by the NMR structure of a related specific complex (5); see main text). (*D*) Pymol representation of the minor wing in absence of DNA. The side chain of V60 is shown (red sticks) relative to the electrostatic surface of the minor wing. Positive and negative electrostatic potential are represented by shades of red and violet, respectively; gray indicates nonpolar surface. A major portion of the V60-related environment comprises the aromatic side chains of H120, Y124, and Y127 (box residues 65, 69, and 72, respectively). This image was calculated using the coordinates provided by Protein Databank entry 1J46.
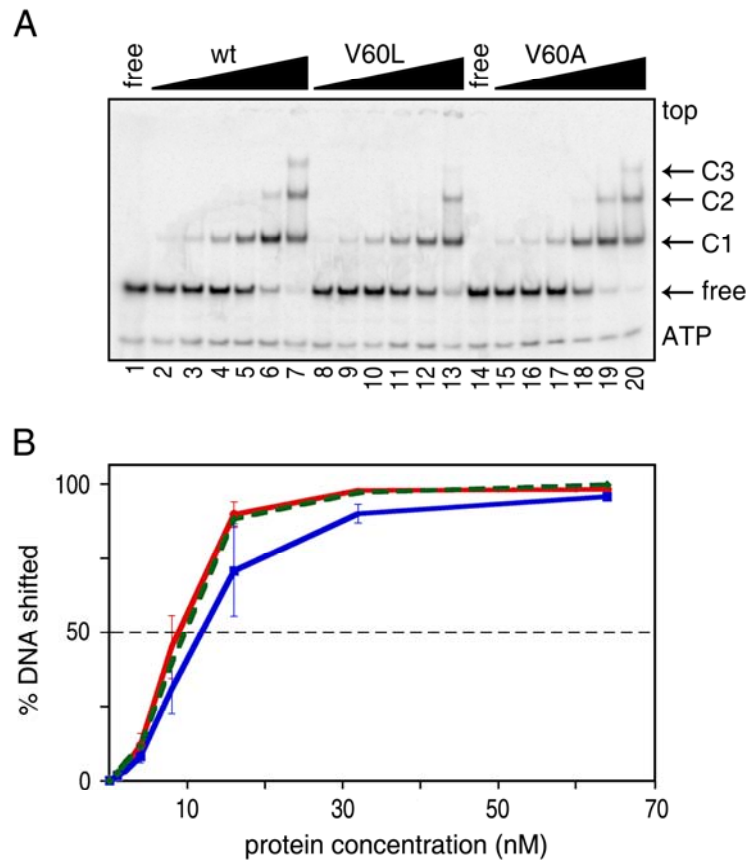
Figure S4. **Specific DNA-Binding affinities of native and variant domains.** (*A*) Representative GMSA gel at 4 °C showing specific binding of the wild-type (wt) and variant domains to 32 base-pair DNA duplex containing a central 5'-ATTGTT consensus DNA target site. The specific 1:1 complex is labeled C1; higher-order complexes C2 and C3 are observed at higher protein concentrations. Successive protein concentrations in lanes 2-7, 8-13, and 15-20 are in each case 2.5, 5, 10, 20, 40 and 80 nM. The free probe is shown in lanes 1 and 14. (*B*) Quantification of GMSA gels providing mean and standard error of three determinations. Dashed line indicates 50% shift. Affinities of wild-type (green dashed line) and V60A (red solid line) domains are indistinguishable; that of the V60L domain (blue solid line) is reduced by less than twofold. The temperature of the gel during electrophoresis was regulated by a recirculating chilled water bath.
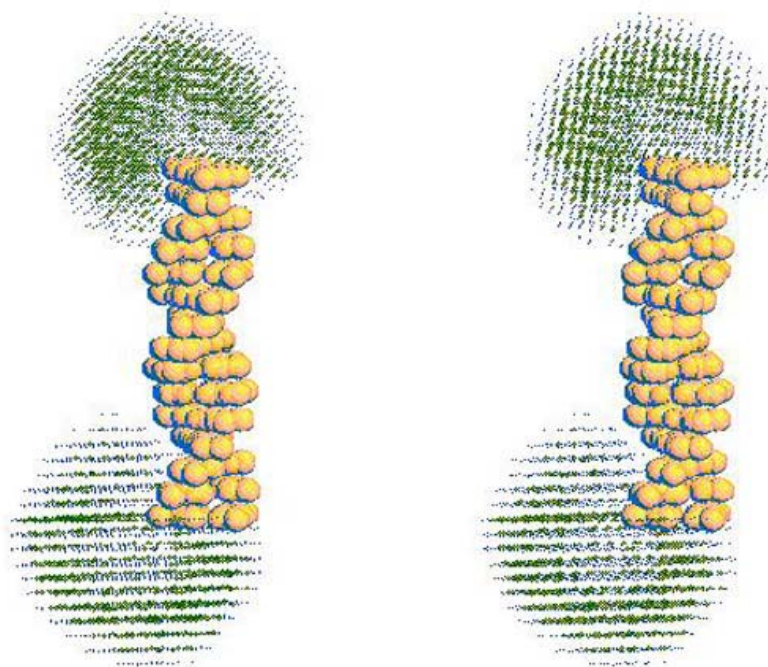
Figure S5. Monte Carlo simulation of ensembles of donor (top) and acceptor (bottom) positions at respective 5'-ends of a 15-base-pair B-DNA site. Because the periodicity of β-DNA is about 10.5 bp/turn of the double helix, a duplex of 15-bp is predicted to position respective 5'-ends on the same side. The model employed the actual linker structures and fluorophores employed in the present study (FAM and TAMRA; see main text). The DNA site was modeled as canonical (unbent) B-DNA. The respective ensembles of probe positions leads to baseline broadening of the inferred distance distribution in the free DNA site (see Figure 8*B* in main text). This figure was kindly provided by Dr. A. R. Srinivasan and Prof. Wilma K. Olson.
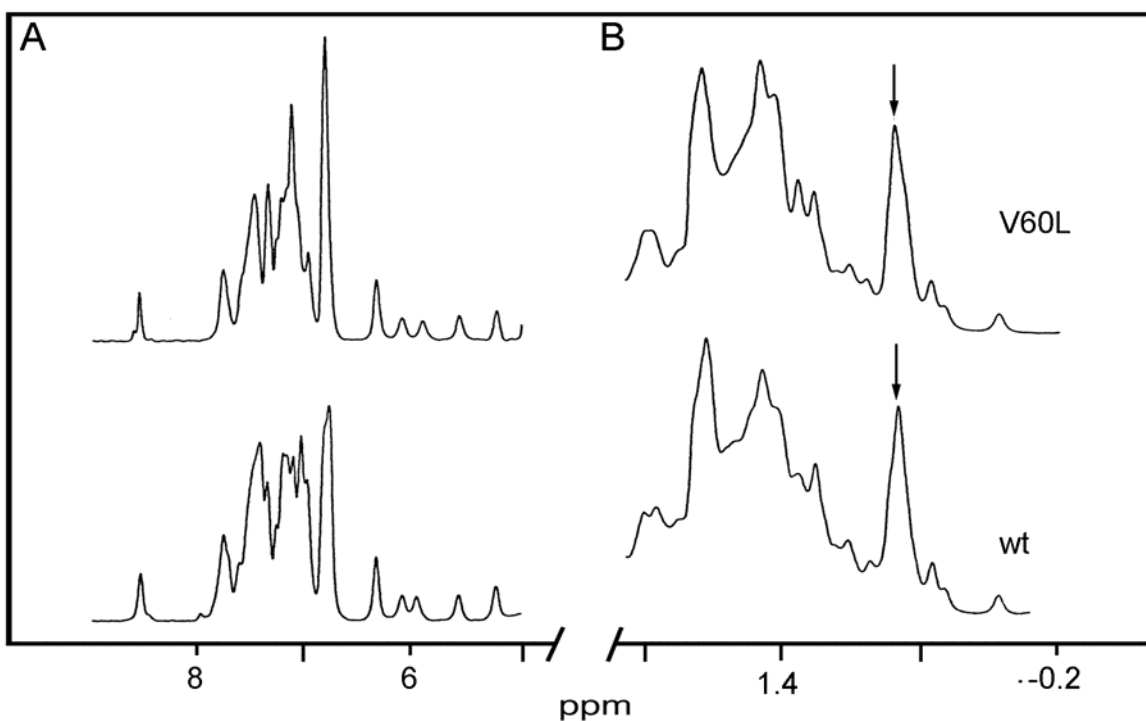
Figure S6. **Wild-type and mutant domains exhibit similar ^{1}H-NMR spectra at pH 4.5 and 4 °C.** 600 MHz ^{1}H-NMR spectra of the wild-type SRY HMG box (bottom) and V60L variant (top). In each case the aromatic region of the spectrum is shown in panel *A* and the upfield aliphatic spectrum in panel *B*. Arrow indicates envelope of methyl resonances containing the γ-CH_3 resonances of V60 and δ-CH_3 resonances of L60 in the free domains. The upfield methyl resonances belong to L101 (box position 46 in the major wing; see Supplemental Figure S8). The protein concentration was in each case 1 mM.
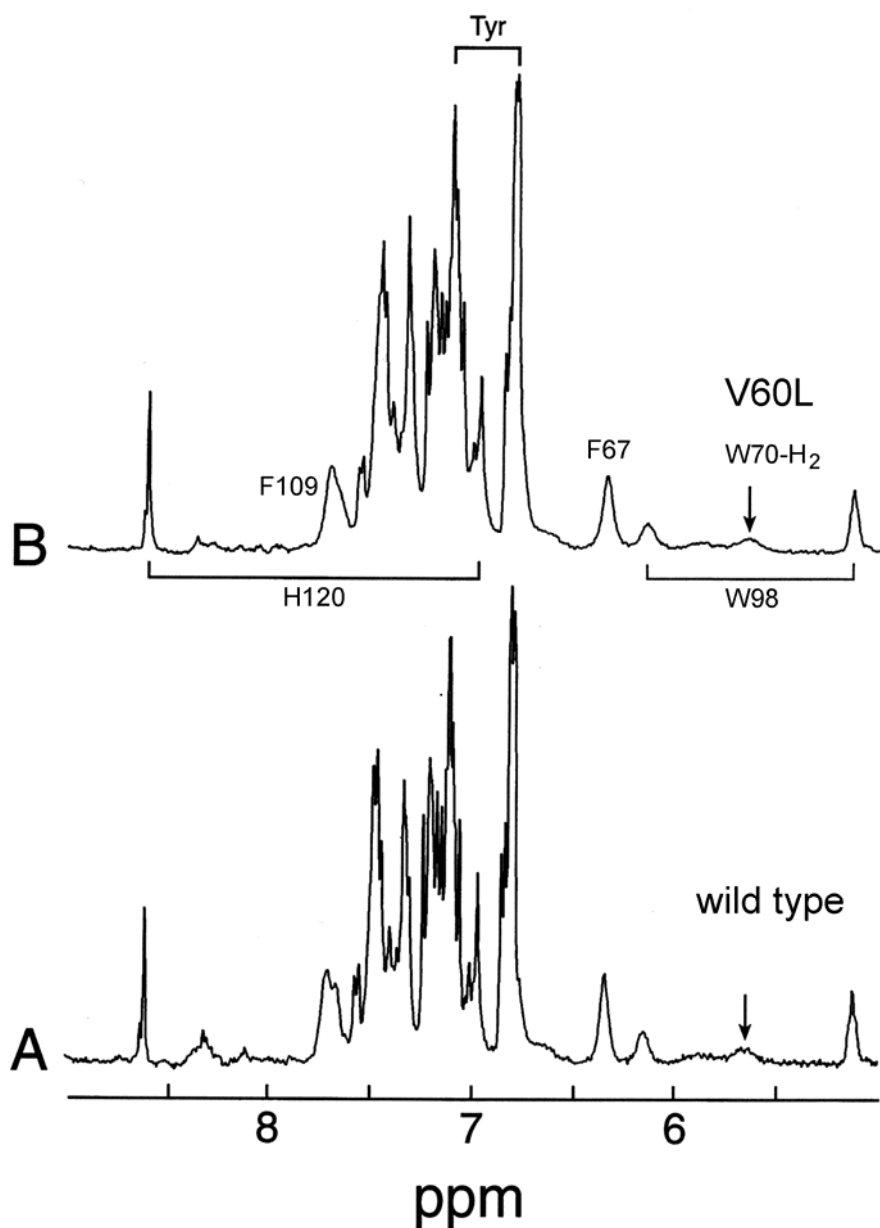
Figure S7. **Wild-type and mutant domains exhibit similar ^{1}H-NMR spectra at pD 7.6 at 25 ^{o}C.** Respective aromatic regions of 600 MHz ^{1}H-NMR spectra (*A*) wild-type SRY HMG box and (*B*) V60L domain. Conformational resonance broadening occurs within the core of the major wing as indicated by the resonances of W70 (consensus position 15; arrows). Narrowing of major-wing core resonances on specific DNA binding indicate a dynamic conformational selection. Resonances of residues F67 and F109 (consensus residues F12 and F54) are also labeled. Samples were prepared in 140 mM KCl and 10 mM potassium phosphate. The protein concentration was in each case ca. 1.0 mM.
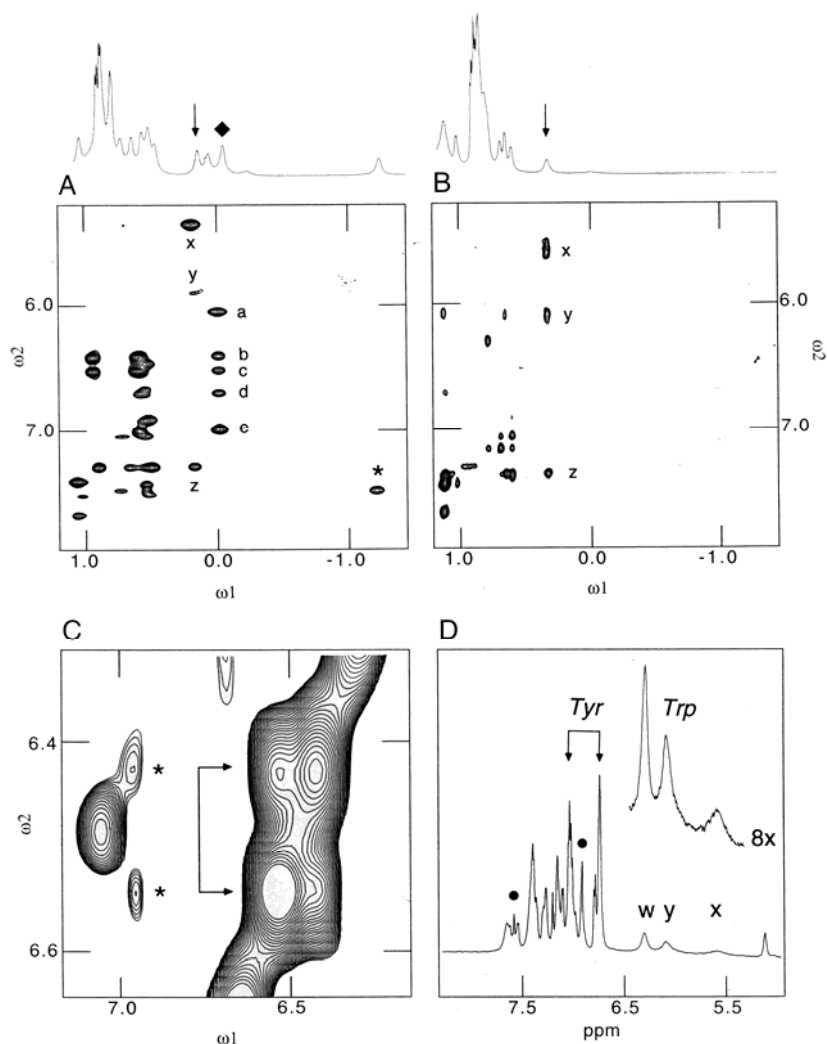
Figure S8. (*A*) NOESY spectrum of specific complex between SRY HMG box and 15 base-pair DNA site contains cross peaks between aromatic side chains and upfield-shifted methyl resonances of V5 (♦ in 1D spectrum at top) and L46 (↓ at top). Cross-peak assignments: (a) V60-H120 C_4H, (b) V60-Y124 *meta*, (c) V60-Y124 *ortho*, (d) V60-Y127 *meta*, (e) V60-Y127 *ortho*, (x, y, and z) L101-W98 ring. Asterisk indicates NOE between the δ-CH3 resonance of I13 and the adenine H2 of A8 in DNA minor groove. The mixing time was 75 ms. (B) NOESY spectrum of isolated SRY HMG box. Contacts between W98 and L101 are retained (x, y, z; arrow in 1D spectrum at top). The secondary shifts of V60, H120 and Y124 are negligible. The mixing time was 200 ms. (*C*) NOE in SRY-DNA complex between H120-H_ε (asterisk) and the aromatic resonances of Y124 (arrows). The mixing time was 75 ms. (*D*) 1D ^{1}H-NMR aromatic spectrum of isolated SRY HMG box. The upfield spin system of W98 is broad. The chemical shifts of H120 and the four tyrosines are near their random-coil values. The complex was made 1.5 mM in 50 mM KCl, 10 mM potassium phosphate (pD 6.0) at 40 °C; the free protein was made 1.0 mM in the same buffer at pD 7.6 and 25 °C.
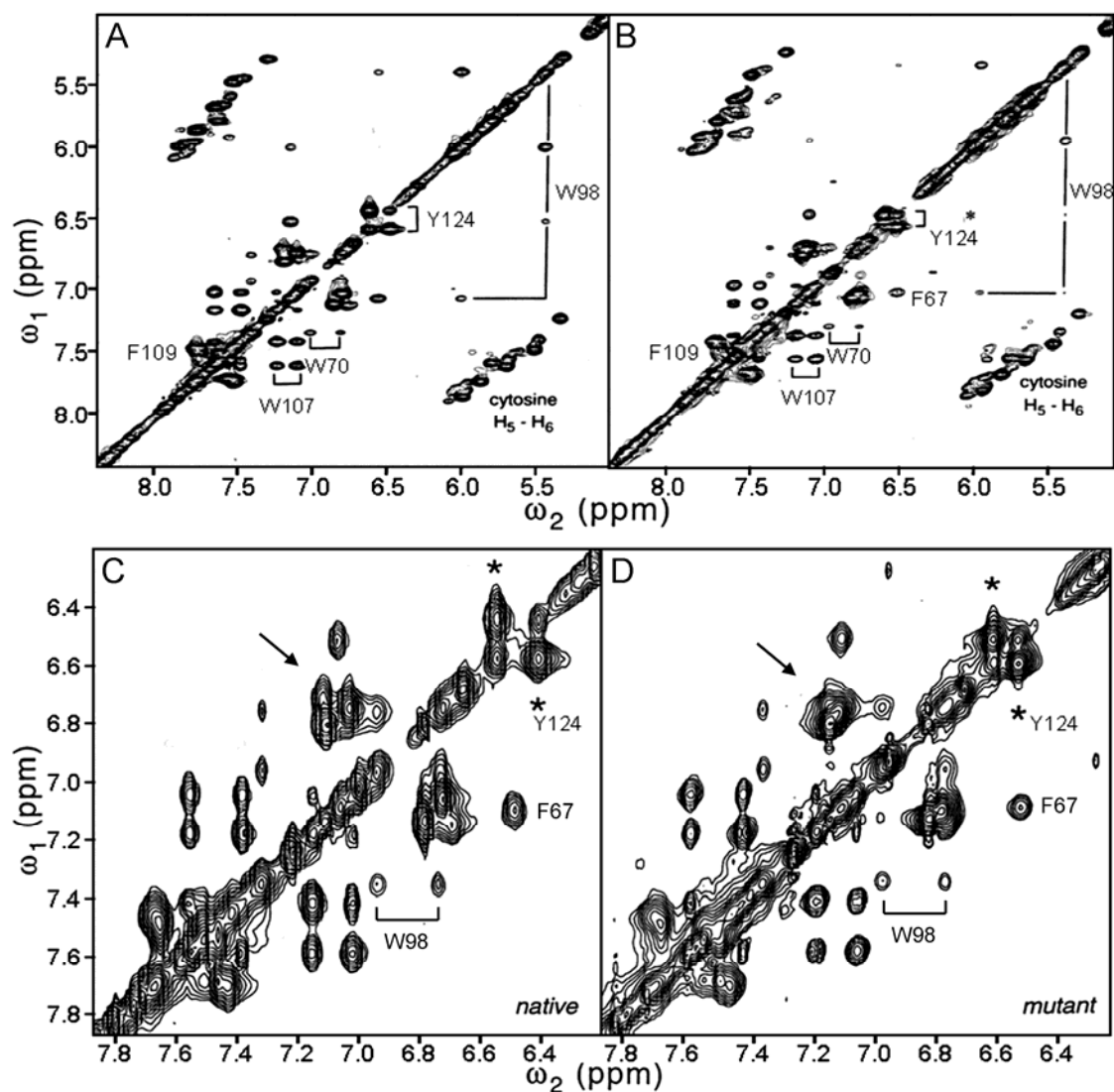
15

Figure S9. **Comparison of 2D 1H-NMR spectra of wild-type and mutant SRY-DNA complexes.** Total correlation (TOCSY) spectra of (*A*) wild-type complex and (*B*) V60L complex at 40 $^{\circ}$C in 10 mM potassium phosphate and 50 mM KCl. (*C* and *D*) expansions of key region of 2D spectra showing altered aromatic chemical shifts of Y124 (asterisks; position 69 of HMG-box consensus). Chemical shifts of wedge side chains F67 and W98 (consensus positions 12 and 43, respectively) are by contrast unperturbed by the mutation. NOEs involving these aromatic resonances are shown in Figure 9 in the main text.
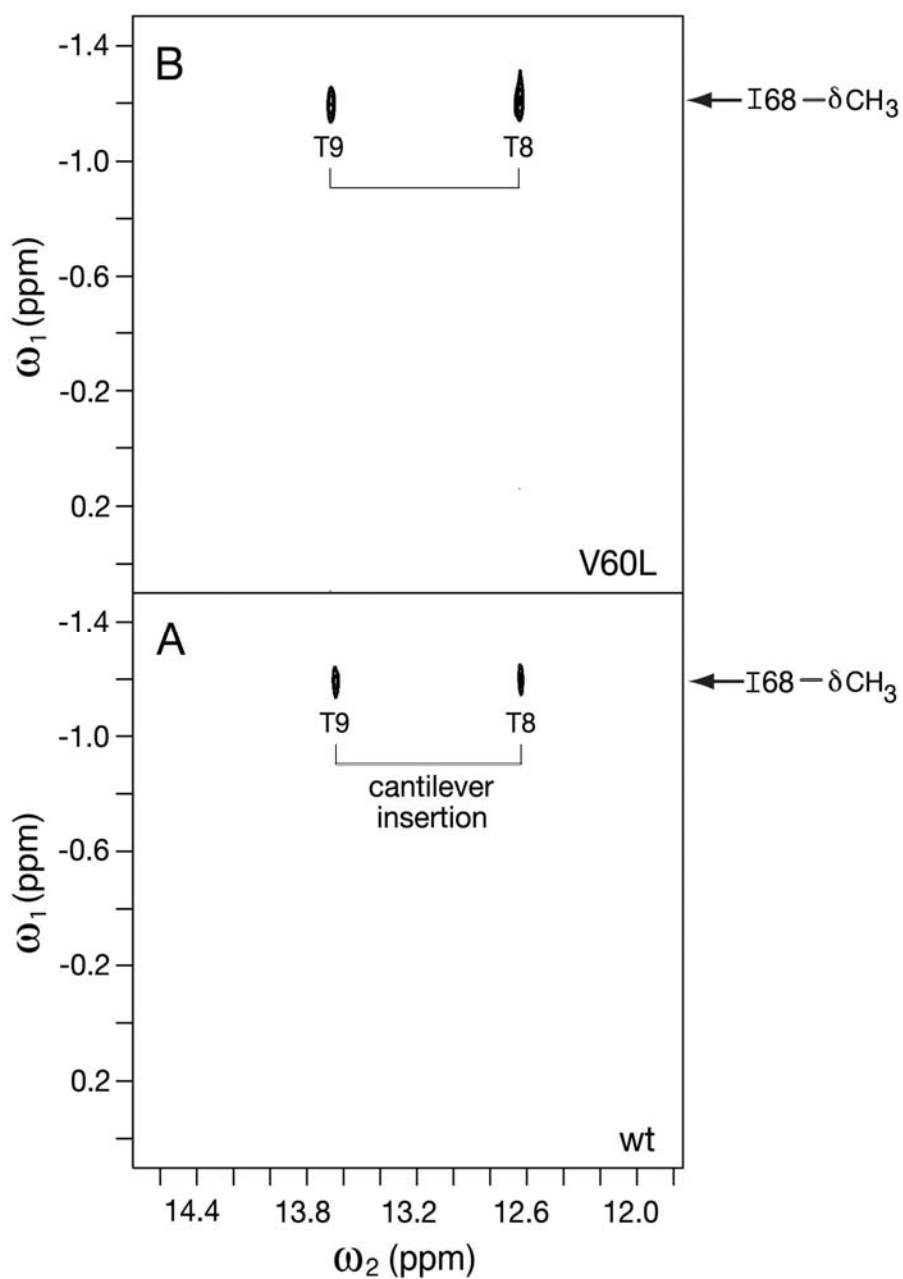
Figure S10. **Wild-type and variant protein-DNA complexes exhibit corresponding cantilever insertion between AT base pairs**. (*A*) Wild-type SRY HMG box/DNA complex and (*B*) V60L variant complex. The horizontal axis (W₂) contains imino DNA resonances whereas the vertical axis (W₁) contains upfield-shifted methyl resonances from SRY. The cantilever side chain of human SRY is I68 (box position 13). The spectra were acquired at 600 MHz at 25 °C and pH 6.0 with mixing time 120 ms.
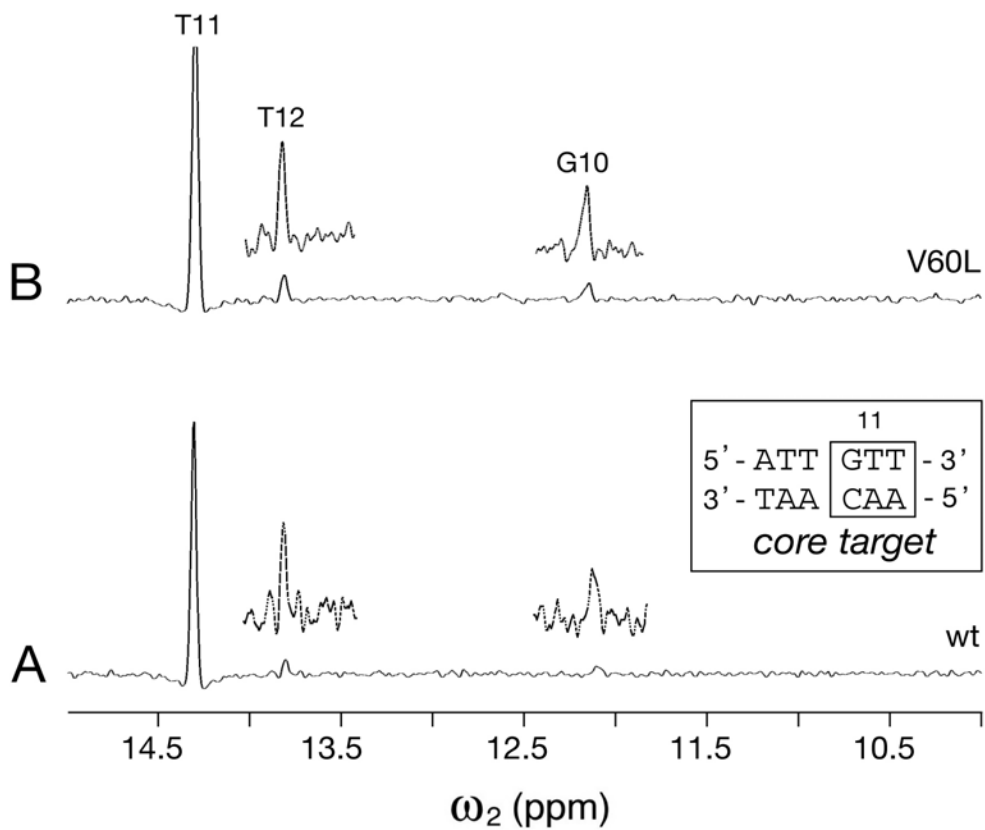
Figure S11. **Wild-type and variant protein-DNA complexes exhibit similar NOE patterns between imino protons within respective bent DNA sites**. Slices from NOESY spectra shown in Figure 9*C* in main text with resonance assignments as indicated. In each case the T_{11}-T_{12} NOE is stronger than the G10-T11 NOE. *Inset*: core DNA Target site with 5'-GTT-3' subsite in inner box. The spectra were each acquired at 600 MHz and 25 °C with mixing time 120 ms.
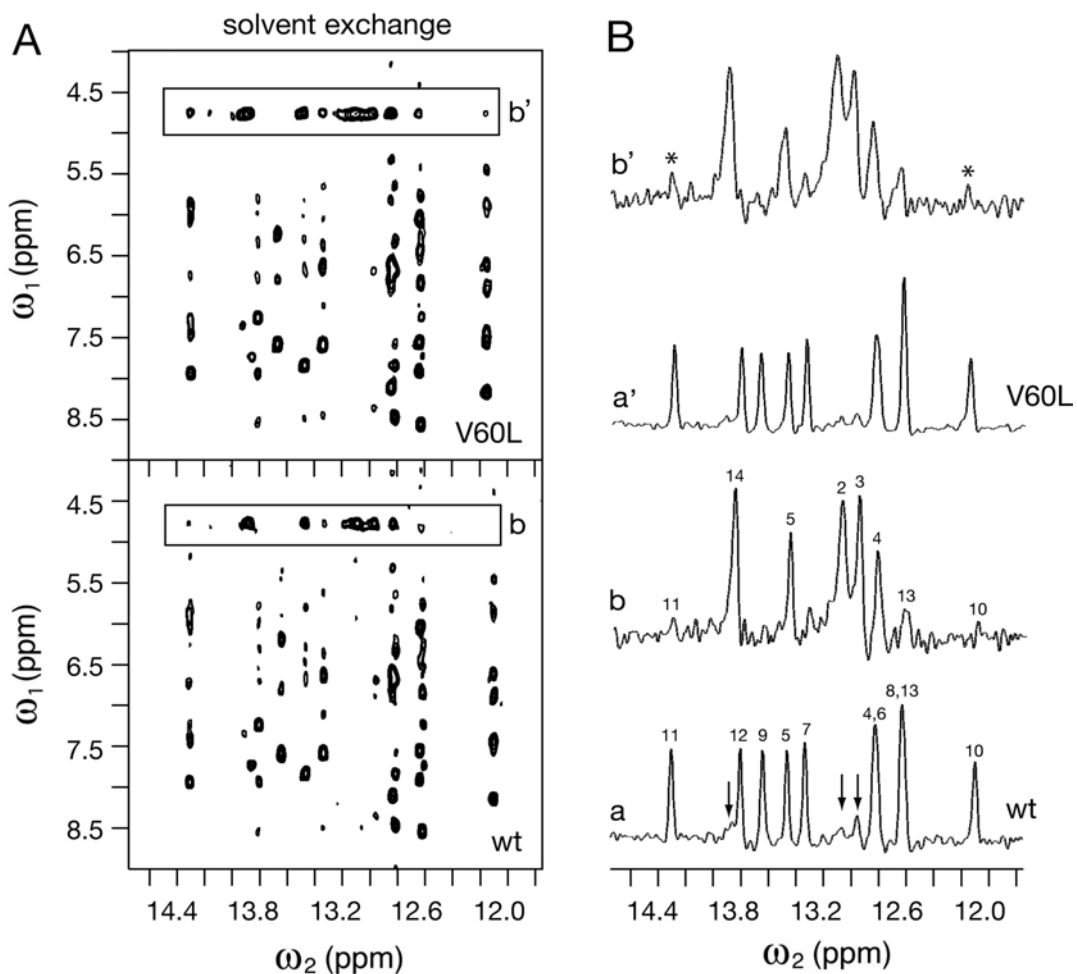
Figure S12. **NMR assay of solvent exchange at Watson-Crick imino hydrogen bonds**. (*A*) NOESY spectra of wild-type (*bottom*) and variant (*top*) protein-DNA complexes (mixing time 120 ms at 25 °C). Boxes indicate solvent exchange cross peaks at the W_1 frequency of H_2O. (B) Extracted 1D spectra: (*a* and *a'*) initial spectra of 2D acquisition with $t_1 = 0$ ms. (*b* and *b'*) Slice of 2D spectrum extracted at the W_1 solvent frequency (*boxes* in panel *A*).
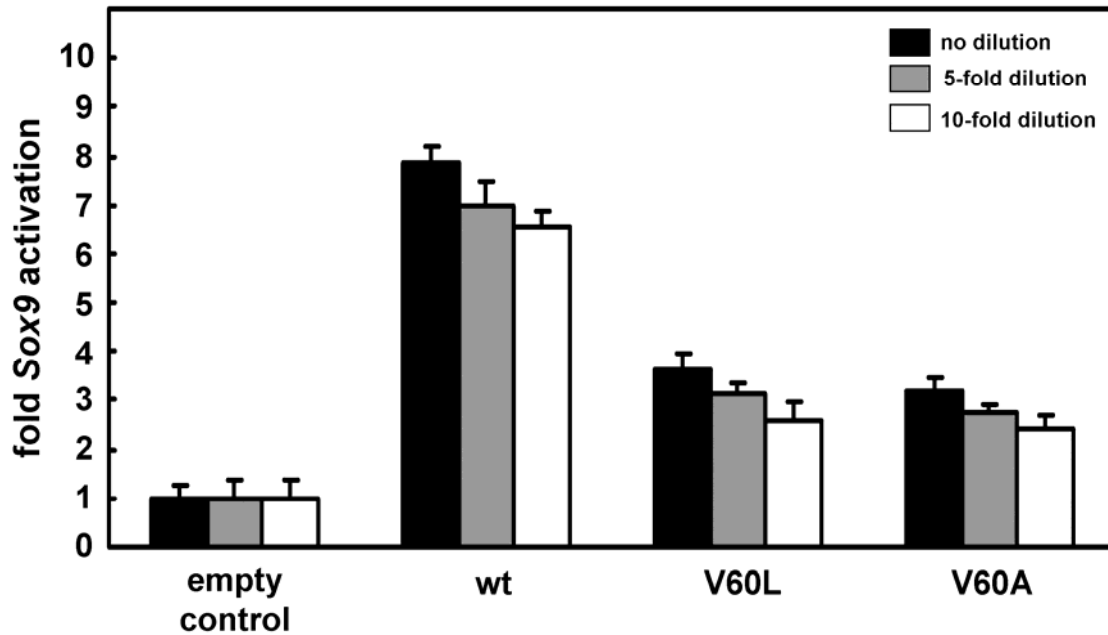
Figure S13. **Dependence of *Sox9* transcriptional activation on dose of transfected plasmid DNA**. Cell-culture studies shown in Figure 10 of main text employed a 1 mg of plasmid DNA per transfection (see Experimental Procedures); these data are reshown here as *black bars*. Successive 5-fold and 10-fold dilutions of the plasmids (*gray* and *white bars*, respectively) yielded similar estimates of the partial activities of the V60L and V60A variants relative to the wild-type SRY. In these studies compensating amounts of the empty plasmid were added to maintain a constant total concentration of DNA. Differences between the three SRY proteins relative to the empty-vector controls (at *left*) are in each case significant with p-values < 0.05. Differences between the mutants and wild-type SRY are also significant at this level.
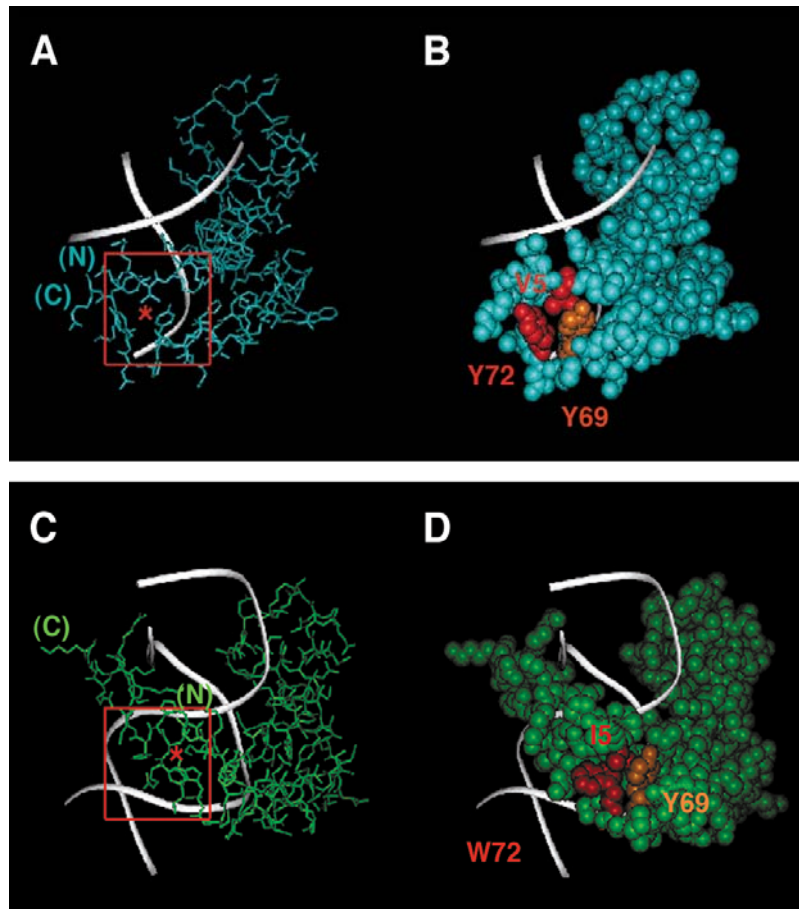
Figure S14. **Back surfaces of SRY-DNA complex (*A* and *B*) and homologous Lef-1 complex (*C* and *D*).** Mutations at V5 are predicted to alter the topography of this surface, proposed to function as a binding site for accessory proteins in an SRY-directed transcriptional preinitiation complex (enhanceosome) or SRY-specific repressor complex (repressosome). The proteins are shown as sticks (*A* (turquoise) and *C* (green)) and space-filling models (*B* and *D*). Red boxes in panels A and C indicate hydrophobic mini-cores containing the side chain of residue 5. Residues 5, and 72 are shown in red, and residue 69 in orange. In each panel the backbone of the DNA duplex is shown as a white ribbon.

A

5'-GTGTTTGTGCAG-3'

complex

← free DNA

wt V60L wt V60L wt V60L wt V60L wt V60L wt V60L

B

70.9°

73.8°

relative mobility

flexure displacement

Figure S15. **DNA bending by the HMG boxes of wild-type SRY and V60L SRY using a variant DNA site** (5'-GTG**TTTGTG**CAG-3' and complement; non-consensus bases in bold). (*A*) PGE: the gel system was 10% polyacrylamide with 58:1 ratio of acrylamide:bis-acrylamide at 4 °C. (*B*) Plot of relative electrophoretic mobility as function of flexure displacement in DNA probes. Inferred DNA bend angles with wild-type SRY (*black squares*) and V60L SRY (*green squares*) are 73.8° and 70.9°, respectively. Samples were loaded side by side on the same gel; for clarity, lanes containing unrelated samples were deleted from the image.

Table S1. Thermodynamic Stabilities of SRY Domains[a]

Protein	ΔGu (kcal/mol)	Cmid (M)	m (kcal mol^{-1} M^{-1})
SRY-wt	3.8 ± 0.1	2.2 ± 0.02	1.8 ± 0.02
SRY-V60L	3.7 ± 0.1	2.1 ± 0.05	1.8 ± 0.04
SRY-V60A	3.9 ± 0.1	2.2 ± 0.03	1.8 ± 0.03

[a]Free energies of unfolding (ΔG_u) are given in kcal/mole as inferred from a two-state model extracted to zero denaturant concentration (see Methods in main text). C_{mid} is that concentration of denaturant (M) associated with half-fractional unfolding; the m value correlates with extent of exposure of nonpolar surfaces on denaturation. Protein unfolding was monitored by intrinsic Trp fluorescence at 4 ^{o}C in a buffer consisting of 10 mM potassium phosphate (pH 7.4) and 140 mM KCl. Within experimental error similar estimates of ΔG_u were obtained when protein unfolding was probed by CD at 222 nm.

Table S2. Thermodynamic Stabilities

pH	[NaCl]	ΔGu (kcal/mol)	pH	[NaCl]	ΔGu (kcal/mol)
2.0	100 mM	1.9	6.0	100 mM	1.7
3.0	100 mM	2.0	6.0	130 mM	1.8
4.0	100 mM	1.9	6.0	300 mM	1.8
5.0	100 mM	1.8	6.0	600 mM	2.0
6.0	100 mM	1.7	6.0	1000 mM	2.3
7.4	100 mM	1.8			

Column 1, Effects of pH were studied by urea denaturation in 100 mM NaCl and 10 mM sodium phosphate at 25 ^{o}C. *Column 2*, Effects of ionic strength were studied by urea denaturation in 10 mM sodium phosphate (pH 6.0).

Table S3. Influence on V5 of Predicted Aromatic Ring Currents[a,b]

Proton	F12	W15	Y41	W43	W52	F54	F55	H65	Y69	Y72	Y74	*sum*
V60 H$_\alpha$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	**0.27**	0.12	0.05	**0.46**
V60 H$_\beta$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.36**	0.01	0.16	**0.53**
V60 γ_1CH$_3$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	**1.67**	0.10	0.06	**1.84**
V60 γ_2CH$_3$	0.00	0.00	0.00	0.00	0.000	0.00	0.00	0.00	**0.36**	**0.29**	0.14	**0.78**

---

[a]Values shown represent the average of individual calculations for each member of the published ensemble. For methyl resonances the value is further averaged among the three component protons. Influence of the Y69 ring current (helix 3) on V5 (N-terminal strand), reflecting long-range is shown in boldface; a smaller influence is predicted due to the Y72 ring current. Calculations were based on the NMR-derived structure of the SRY domain-DNA complex as determined in 1995 by G. M. Clore, A. Gronenborn and colleagues (Protein Databank accession code 1HRY; see main text for reference (15). Ring-current shifts were calculated as described by (15).
[b]Residue numbers refer to HMG box consensus; for numbering in intact human SRY, 55 is added to the residue number.

REFERENCES

1. Haqq, C. M., King, C. Y., Donahoe, P. K., and Weiss, M. A. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 1097-1101
2. King, C. Y., and Weiss, M. A. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 11990-11994
3. Kim, J., Zwieb, C., Wu, C., and Adhya, S. (1989) *Gene* **85**, 15-23
4. 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) *Mol. Endocrinol.* **15**, 363-377
5. Murphy, E. C., Zhurkin, V. B., Louis, J. M., Cornilescu, G., and Clore, G. M. (2001) *J. Mol. Biol.* **312**, 481-499
6. Sosnick, T. R., Fang, X., and Shelton, V. M. (2000) *Methods Enzymol.* **317**, 393-409
7. Pace, C. N., and Shaw, K. L. (2000) *Proteins* **Suppl 4**, 1-7
8. Beechem, J. M., and Haas, E. (1989) *Biophys. J.* **55**, 1225-1236
9. Grinvald, A., and Steinberg, I. Z. (1974) *Anal. Biochem.* **59**, 583-598
10. Phillips, N. B., Nikolskaya, T., Jancso-Radek, A., Ittah, V., Jiang, F., Singh, R., Haas, E., and Weiss, M. A. (2004) *Biochemistry* **43**, 7066-7081
11. Haas, E., Katchalski-Katzir, E., and Steinberg, I. Z. (1978) *Biochemistry* **17**, 5064-5070
12. Li, B., Phillips, N. B., Jancso-Radek, A., Ittah, V., Singh, R., Jones, D. N., Haas, E., and Weiss, M. A. (2006) *J. Mol. Biol.* **360**, 310-328
13. Phillips, N. B., Jancso-Radek, A., Ittah, V., Singh, R., Chan, G., Haas, E., and Weiss, M. A. (2006) *J. Mol. Biol.* **358**, 172-192
14. Love, J. J., Li, X., Case, D. A., Giese, K., Grosschedl, R., and Wright, P. E. (1995) *Nature* **376**, 791-795
15. Weiss, M. A., and Hoch, J. C. (1987) *J. Magn. Reson.* **72**, 324-333