THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 279, No. 2, Issue of January 9, pp. 1449–1457, 2004
Printed in U.S.A.

Molecular Basis for Synergistic Transcriptional Activation by Oct1 and Sox2 Revealed from the Solution Structure of the 42-kDa Oct1·Sox2·*Hoxb1*-DNA Ternary Transcription Factor Complex*

Received for publication, September 3, 2003, and in revised form, October 2, 2003
Published, JBC Papers in Press, October 14, 2003, DOI 10.1074/jbc.M309790200

David C. Williams, Jr.‡, Mengli Cai, and G. Marius Clore§

From the Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda, Maryland 20892-0520

The Oct and Sox transcription factors control many different aspects of neural development and embryogenesis, often binding to adjacent sites on DNA, and interacting with one another through their DNA binding domains to regulate transcription synergistically. Oct proteins contain two DNA binding domains (POU_S and POU_{HD}) connected by a flexible linker, which interact with DNA in a bipartite manner. Residual dipolar coupling measurements on the binary Oct1·DNA complex reveal that the two domains are characterized by distinct alignment tensors in both phage pf1 and polyethylene glycol/hexanol liquid crystalline media. We show that this difference is due to a fast microscopic dissociation/association process involving alternative binding modes for the weaker binding POU_S domain in the binary complex. Upon binding of Sox2 to an adjacent site in the *Hoxb1* regulatory element, all components of the ternary Oct1·Sox2·DNA complex share a single alignment tensor. Thus ternary complex formation increases the site-specific affinity of Oct1 for DNA by effectively locking the POU_S domain in a single orientation on the DNA. The solution NMR structure of the ternary 42 kDa Oct1·Sox2·*Hoxb1*-DNA complex, determined by novel procedures based on orientational restraints from dipolar couplings and conjoined rigid body/torsion angle dynamics, reveals that Sox2 and POU_S interact through a predominantly hydrophobic interface, surrounded by a ring of electrostatic interactions. These observations suggest a mechanism of combinatorial control involving direct protein-protein interactions on the DNA whereby Oct1 in conjunction with a co-interacting transcription factor provide cell-specific transcription regulation.

Transcription regulation in eukaryotes involves the formation of protein-DNA complexes that can interact with and modulate the downstream transcriptional machinery (1, 2). Unlike prokaryotes that often use a single protein for this function, eukaryotes generally employ complexes of multiple proteins in what has been termed combinatorial control (2, 3). This mechanism effectively integrates many different signaling pathways to provide a more complex regulatory network based on a finite number of transcription factors. Biological and structural studies of these complexes indicate that combinatorial control is achieved by employing transcription factors with adaptable DNA and protein binding surfaces. This adaptability allows different combinations of these factors to interact on specific promoter elements to drive synergistic transcription regulation (2, 3).

Transcription regulation by the Oct and Sox families of transcription factors reflects many of the principles of combinatorial control. Different members of each family have been shown to interact on different promoter elements to regulate transcription during embryogenesis and neural development (4, 5). The Sox family is characterized by an HMG-box DNA binding domain that binds in the minor groove, bends DNA (50–90°) and specifically recognizes variations of the consensus sequence CTTTGTT (4–6). The DNA binding region of Oct proteins, known generically as POU, comprises two small, tethered domains, a specific domain (POU_S) and a homeodomain (POU_{HD}), that bind in the major groove, do not alter the DNA appreciably from straight B-form, and typically recognize the octamer consensus sequence ATGCTAAT (5, 7).

Previous studies have demonstrated that specific promoters are selectively responsive to different members of the Oct transcription factor family. For example, both Oc3/4 and Oct1 can bind adjacent to Sox2 on the *FGF4* and *UTF1* promoter elements, but only Oct3/4 (or chimeras containing the POU domain from Oct3/4) can synergistically drive transcription from each. Biological data indicate that this synergism derives from direct interaction between the DNA binding domains of each protein (8, 9). Interestingly, the *FGF4* element contains three additional base pairs between the Oct and Sox binding sites when compared with the *UTF1* element. The increased separation between the two proteins must lead to alternative interprotein binding surfaces. More recently, Di Rocco *et al.* (10) showed that the *Hoxb1* autoregulatory element contains contiguous Oct and Sox binding sites, with the same relative spacing as in the *UTF1* element. However, Oct1 instead of Oct3/4 synergistically drives transcription from the *Hoxb1* element. This latter regulatory element functions to selectively promote transcription of Hoxb1 in rhombomere 4 of the hindbrain during embryogenesis (10, 11).

To further our understanding of combinatorial control at the molecular level, we have solved the three-dimensional solution NMR structure of the biologically relevant ternary 42,000 M_r complex formed by the Oct1 (POU_S + POU_{HD}) and Sox2 DNA binding domains bound to a 19-base pair fragment of the *Hoxb1* element.

* This work was supported in part by the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (to G. M. C.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and experimental NMR restraints (code 1O4X) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

‡ Recipient of a Pharmacology Research Associate training postdoctoral fellowship from NIGMS, National Institutes of Health.

§ To whom correspondence should be addressed: Laboratory of Chemical Physics, Bldg. 5, Rm. B1-30I, NIDDK, National Institutes of Health, Bethesda, MD 20892-0520. Tel.: 301-496-0782; Fax: 301-496-0825; E-mail: mariusc@intra.niddk.nih.gov.

EXPERIMENTAL PROCEDURES

Sample Preparation—The POU region (POU_S + POU_{HD}) of human Oct1 (residues 275–439) and the HMG-box domain of human Sox2 (residues 40–123) were cloned as histidine tag fusions, expressed, purified by affinity chromatography, and the tags removed proteolytically using standard procedures. Each domain was numbered according to previously determined structures (6, 7). NMR samples (in 10 mM sodium phosphate, 10 mM dithiothreitol, and 0.02% NaN_3, pH 6.5) comprised stoichiometric binary and ternary complexes with the DNA unlabeled and the proteins in various isotopic-labeling combinations (unlabeled, ^{15}N, $^{15}N/^{13}C$, $^{15}N/^{13}C/^{2}H$, or $^{15}N/^{13}C/^{2}H$/VLI-methyl protonated, Ref. 12). The concentrations of the binary and ternary complexes in the various samples ranged from 0.5 to 1 mM.

NMR Spectroscopy—All NMR experiments were carried out at 30 °C on Bruker 500, 600, 750, and 800 MHz spectrometers. Spectra were processed using the NMRPipe package (13), and analyzed using the programs PIPP, CAPP, and STAPP (14). ^{1}H, ^{15}N, and ^{13}C backbone and side-chain resonance assignments of the proteins were carried out using TROSY[1] versions (15, 16) of three-dimensional triple resonance experiments (17). ^{1}H resonance assignments for the DNA were obtained from two-dimensional ^{12}C-filtered NOE and Hartmann-Hahn experiments, aided by prior assignments on the SRY·DNA binary complex (6). Intermolecular distance restraints were derived from three-dimensional ^{13}C-separated/^{12}C-filtered NOE experiments (17). Heteronuclear ^{3}J couplings were measured by quantitative J-correlation spectroscopy (18). Residual dipolar couplings ($^{1}D_{NH}$, $^{1}D_{NC'}$, $^{1}D_{C\alpha C'}$) were measured in 10 mg/ml phage pf1 (19, 20) using TROSY triple resonance experiments (21).

Structure Calculations—Structures were calculated from the experimental restraints by conjoined rigid body/torsion angle dynamics (22) using the Xplor-NIH macromolecular NMR structure determination package (23). Structure figures were generated with the programs VMD-XPLOR (24) and RIBBONS (25).

The starting coordinates for the 19-mer DNA were built as follows: the POU_S and POU_{HD} hemi-binding sites (base pairs 11–14 and 17–19, respectively) were derived from the 1.9 Å resolution crystal structure of the Oct1·*MORE*-DNA complex (26), the Sox2 binding site (base pairs 1–10) from the NMR structure of the related SRY-DNA complex (6), and the intervening sequences (base pairs 15–16) and regions containing substitutions (base pairs 1, 4, and 10) from classical B-DNA. The resulting model was then subjected to regularization.

The following strategy was used in the conjoined rigid body/torsion angle dynamics calculations (see "Results" for general justification of the approach employed). There were four rigid bodies: rigid body 1, backbone and non-interfacial side chains of POU_{HD} and base pairs 17–19 of the DNA (taken from the 1.9 Å resolution Oct1·*MORE* DNA complex (26); PDB accession code 1E3O); rigid body 2, backbone and non-interfacial side chains of POU_S (also taken from the Oct1·*MORE*-DNA complex (26); PDB accession code 1E3O); rigid body 3, backbone and non-interfacial side chains of Sox2 and base pairs 1–4 of the DNA (taken from the NMR structure of the related SRY·DNA complex (6), see "Results"; PDB accession code 1J46); rigid body 4, the axis of the dipolar coupling alignment tensor. Rigid bodies 1–3 have both rotational and translational degrees of freedom, while rigid body 4 is given only rotational degrees of freedom. The following interfacial side chains were given torsional degrees of freedom. For POU_{HD}: 10 residues at the POU_{HD}/DNA interface (residues 107, 108, 113, 144, 147, 148, 151, 154, 155, and 158) with 24 side chain torsion angles restrained to within a range of ±20° of values in crystal structures of Oct1·DNA binary complexes. For POU_S: 6 residues at the POU_S/Sox2 interface (residues 14, 17, 18, 21, 26, and 52); and 14 residues at the POU_S/DNA interface (residues 20, 27, 41, 42, 44–46, 48, 49, 54, 58, 59, 62, and 63) with 35 side chain torsion angles restrained to within a range of ±20° of values in crystal structures of Oct1·DNA binary complexes. For Sox2: 7 residues at the POU_S/Sox2 interface (residues 59, 62, 63, 66, 67, 71, 73) and 18 residues at the Sox2/DNA interface (residues 4, 6–10, 12, 13, 17, 31, 35, 43, 44, 51, 55, 76, 78, and 79) with 47 side chain torsion angles of conserved residues restrained to within a range of ±20° of values in solution structure of homologous SRY·DNA binary complex. Base pairs 5–16 of the DNA were given torsional degrees of freedom with 220 loose backbone phosphodiester torsion angle restraints to prevent local mirror images (6): $\alpha = 70 \pm 30°$; $\beta = 180 \pm 50°$; $\gamma = 60 \pm 35°$; $\delta = 145 \pm 15°$ (except for T5, T6, and A32, which were restrained to 75 ± 15°, see the SRY·DNA binary complex, Ref. 6); $\epsilon = 180 \pm 40°$, and $\zeta = -90 \pm 30°$. In addition, data base base-base positional (27) and torsion angle (28) potentials of mean force were employed to ensure that good stereochemistry and base-base interactions were preserved throughout the DNA. The following distance restraints were used to preserve protein-DNA interactions observed in the binary Oct1·DNA complex (7, 26). At the POU_S/DNA interface: 5 distances (3.0 ± 0.5 Å) involving hydrogen bonding interactions to bases from 3 residues; and 10 distances involving electrostatic interactions with phosphates (≤6.1 Å) from 10 residues. At the POU_{HD}/DNA interface: 3 distances (3.0 ± 0.5 Å) involving hydrogen-bonding interactions to bases from 2 residues; and 2 distances involving electrostatic interactions with phosphates (≤6.1 Å) from 2 residues.

The final force constants for the various terms in the target function are as follows: 30 kcal·mol^{-1}·Å^{-2} and 200 kcal·mol^{-1}·rad^{-2} for the square-well potentials describing the interproton distance and torsion angle restraints, respectively; 1.0 kcal·mol^{-1}·Hz^{-2} for the harmonic potential describing the $^{1}D_{NH}$ dipolar coupling restraints; 0.05 and 0.035 kcal·mol^{-1}·Hz^{-2} for the harmonic potentials describing the $^{1}D_{NC'}$ and $^{1}D_{C\alpha C'}$ dipolar coupling restraints normalized relative to the $^{1}D_{NH}$ dipolar couplings; 10 kcal·mol^{-1}·Å^{-2} for the harmonic Watson-Crick hydrogen bonding distance potentials (six per base pair; Ref. 29); 100 kcal·mol^{-1}·Å^{-2} for the harmonic base pair planarity restraints used to prevent undue buckling while permitting unrestricted propeller twisting (29); 4 kcal·mol^{-1}·Å^{-4} for the quartic van der Waals repulsion term with a van der Waals radius scale factor of 0.78; 1.5 for the torsion angle database potential of mean force; 0.3 for the positional base-base potential of mean force; and 0.1 kcal·mol^{-1}·Å^{-2} for the very weak harmonic non-crystallographic symmetry restraint employed to maintain the approximate translational separation between POU_S and POU_{HD} observed in the monomer unit of the Oct1·*PORE*-DNA complex (see "Results").

RESULTS AND DISCUSSION

Comparison of Domain Structures in Binary and Ternary Complexes—We approached the study of the Oct1·Sox2·*Hoxb1*-DNA complex by using a variety of isotope (^{15}N, ^{13}C, ^{2}H) labeling schemes to permit the observation of each component of the complex individually as well as to selectively detect intermolecular interactions between the components (17). The largest backbone $^{1}H_N/^{15}N$ chemical shifts differences ($\Delta_{H/N} \geq$ 150 Hz, at 600 MHz) between the ternary (Oct1·Sox2·DNA) and binary (Oct1·DNA and Sox1·DNA) complexes involve contiguous regions of POU_S and Sox2 adjacent to the binding site of the other transcription factor (Fig. 1*A*), consistent with the sequential arrangement of the DNA binding sites (Fig. 2*B*) and indicative of direct protein-protein interactions between Sox2 and POU_S. Titration experiments show that formation of the ternary complex from binary complexes is in slow exchange on the chemical shift scale.

Residual dipolar couplings yield unique orientational information, independent of the spatial proximity of the atomic bond vectors, that provides a powerful and rapid method of structure validation in relatively rigid systems such as proteins and nucleic acids (30, 31). Backbone residual dipolar couplings ($^{1}D_{NH'}$, $^{1}D_{NC'}$, $^{1}D_{C\alpha C'}$), measured on the ternary Oct1/Sox2/DNA complex dissolved in a dilute liquid crystalline medium of bacteriophage (19, 20), were in excellent agreement with the structures of the individual domains in existing binary complexes (Table I): the $^{1}D_{NH}$ dipolar coupling R-factors ($R^{NH.}_{dip}$; Ref. 20) for the POU_S and POU_{HD} domains fitted individually to the 1.9 Å resolution Oct1·DNA complex (26) are 16.2 and 17.5%, respectively; R^{NH}_{dip} for Sox2 best-fitted to the solution structure of the related SRY·DNA complex (6) is 16.0%; and the correlation coefficients for all fits to the $^{1}D_{NH}$ couplings are ≥0.97. These data indicate that the backbone and core structures of the domains do not change upon formation of the ternary complex.

The $^{1}D_{NH}$ dipolar coupling data indicate unambiguously that the backbone coordinates of Sox2 and the highly homologous SRY (68% sequence identity) in their respective ternary and

[1] The abbreviations used are: TROSY, transverse-optimized relaxation spectroscopy; HSQC, heteronuclear single quantum coherence; NOE, nuclear Overhauser effect; r.m.s., root mean-squared; PEG, polyethylene glycol.

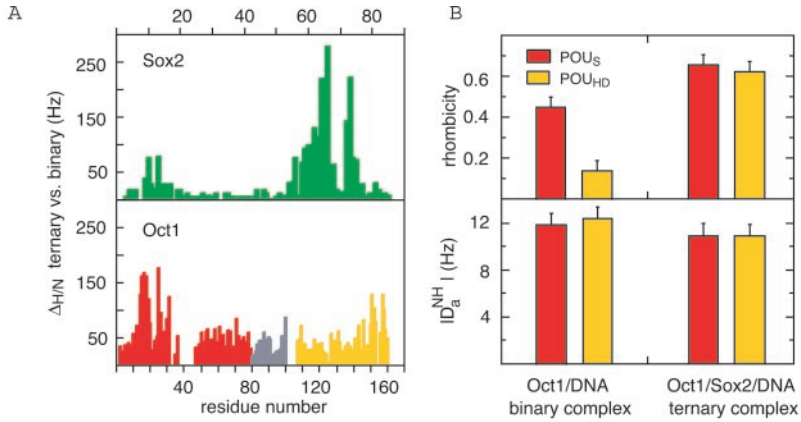

FIG. 1. **Dynamic configuration of the binary Oct1·*Hoxb1*-DNA and ternary Oct1·Sox2·*Hoxb1*-DNA complexes derived from residual dipolar coupling data.** *A*, chemical shift mapping of the Sox2 and Oct1 interaction surfaces. Backbone $^{1}H_N/^{15}N$ chemical shift perturbation ($\Delta_{H/N} = [\Delta(\delta^{1}H_N)^2 + \Delta(\delta^{15}N)^2]^{1/2}$ in Hz at a ^{1}H frequency of 600 MHz) associated with the formation of the ternary Oct1·Sox2·*Hoxb1*-DNA complex from binary Oct1·*Hoxb1*-DNA and Sox2·*Hoxb1*-DNA complexes are plotted against residue number for Sox2 (*upper panel*, *green*), POU_S (*lower panel*, *red*), POU_{HD} (*lower panel*, *gold*), and the linker region between POU_S and POU_{HD} (*lower panel*, *gray*). *B*, magnitude of the dipolar coupling alignment tensors obtained by singular value decomposition using the coordinates of the individual domains from the 1.9 Å resolution Oct1·*MORE*-DNA complex (26). The rhombicity (*top panel*) and absolute value of the axial component of the alignment tensor normalized to the NH dipolar couplings (D_a^{NH}, *lower panel*) are plotted for POU_S (*red*) and POU_{HD} (*gold*) in the binary (*left*) and ternary (*right*) complexes. *Error bars* indicate one S.D. calculated as described in Ref. 33. (Note that the values of D_a^{NH} for both POU_S and POU_{HD} are negative in the binary complex but positive in the ternary one).

TABLE I
Structure validation of individual domain components in the ternary Oct1·Sox2·DNA complex

The table displays the agreement between observed dipolar couplings measured in the ternary Oct1·Sox2·DNA complex and the crystal structures of the individual POU_s and POU_{HD} domains in three Oct1·DNA crystal structures and the solution NMR structure of the SRY·DNA complex obtained by singular value decomposition using individual alignment tensors for each domain. The dipolar coupling *R*-factor (R_{dip}) is defined as the ratio of the r.m.s. deviation between observed and calculated values to the expected r.m.s. deviation if the vectors were randomly oriented. The latter is given by $\{2D_a^2\{4+3\eta^2\}/5\}^{1/2}$, where D_a is the magnitude of the axial component of the alignment tensor and η the rhombicity (30). The dipolar coupling *R*-factors are correlated to the resolution at which the structures are solved. As expected, the lowest R_{dip} values are obtained for the highest resolution (1.9 Å) crystal structure of the Oct1·*MORE*-DNA complex (26). This is purely a reflection of higher coordinate accuracy as a function of crystal structure resolution. Note that the backbone for POU_S and POU_{HD} for the three crystal structures are still very similar with pairwise backbone rms differences ranging from 0.6 to 0.8 Å for POU_S and 0.4 to 0.7 Å for POU_{HD}. The number of measured dipolar couplings is as follows: for POU_S, 39 $^{1}D_{NH}$, 33 $^{1}D_{NC'}$ and 34 $^{1}D_{CaC'}$; for POU_{HD}, 39 $^{1}D_{NH}$, 34 $^{1}D_{NC'}$ and 27 $^{1}D_{CaC'}$; for SRY, 51 $^{1}D_{NH}$, 39 $^{1}D_{NC'}$ and 49 $^{1}D_{CaC'}$. The PDB accession codes for the Oct1·*H2B*-DNA (7), Oct1·*PORE*-DNA (26), Oct1·*MORE*-DNA (26), and SRY·DNA (7) binary complexes are 1OCT, 1HFO, 1E3O, and 1J46, respectively.

	R_{dip}		
	$^{1}D_{NH}$	$^{1}D_{NC}$	$^{1}D_{CaC}$
	%		
Oct1·H2B-DNA (3.0 Å resolution)			
POU_S	28.1	44.1	37.2
POU_{HD}	26.3	37.7	38.1
Oct1·*PORE*-DNA (2.7 Å resolution)			
POU_S	26.5	28.9	24.7
POU_{HD}	18.9	31.3	42.1
Oct1·*MORE*-DNA (1.9 Å resolution)			
POU_S	15.7	24.9	29.5
POU_{HD}	18.0	25.3	31.7
SRY·DNA (NMR)			
Sox2	16.0	29.5	32.7

binary complexes are identical within the limits of the NMR method. In terms of coordinate accuracy, a value of 16% for R_{dip}^{NH} corresponds approximately to that of a 1.5–2.0 Å resolution crystal structure (see Table I, Refs. 19, 30, 32–35). It should be also emphasized that the dipolar coupling data measured on Sox2 in the ternary complex provide excellent independent cross-validation for the accuracy of the solution NMR coordinates of SRY in the binary complex. Not only were the two data sets collected on homologous proteins bound to different DNA fragments, but they were also collected in different alignment media: the SRY·DNA complex was refined against dipolar couplings measured in a lipid bicelle liquid crystalline medium (6) in which the alignment tensor is determined by shape as opposed to both shape and charge distribution in the case of phage pf1 used in the present study (36). The structural similarity between Sox2 and SRY is not surprising given that the amino acid differences are localized exclusively to the protein surface. Further, both SRY and Sox2 bend DNA to the same extent, as judged by gel shift circular permutation assays (37) reflecting the fact that only 4 out of 27 DNA-contacting residues differ between the HMG-boxes of the two proteins (three conservative substitutions of residues that interact with backbone phosphates and an isoleucine to methionine substitution of the residue that intercalates between two AT base pairs).

Affinity Modulation and Cooperativity—Comparison of the dipolar coupling data between the POU_s and POU_{HD} domains measured on the Oct1·DNA binary complex revealed an unexpected finding. Residual dipolar couplings arise from partial alignment of proteins in a magnetic field, as a consequence, for example, of being dissolved in a dilute liquid crystalline medium (30). For a macromolecular complex in which the relative orientations of the different components are fixed, all components of the complex will share a single alignment tensor (30, 38). The dipolar coupling data measured in the negatively charged phage medium (Fig. 1*B*, *top panel*), however, clearly indicate that the POU_S and POU_{HD} domains in the binary complex are characterized by two distinct alignment tensors with very different rhombicities (η): highly rhombic for POU_S ($\eta = 0.46 \pm 0.05$) but close to axially symmetric for POU_{HD} ($\eta = 0.14 \pm 0.05$). Exactly the same phenomenon is observed in a liquid crystalline medium of neutral polyethylene glycol (PEG)/hexanol (39), which yields rhombicities of 0.42 ± 0.05 and 0.19 ± 0.05 for the POU_S and POU_{HD} domains, respectively (data not shown). Moreover, the $^{1}D_{NH}$ dipolar couplings measured in the two media are highly correlated with a correlation coefficient of 0.96.

The observation of two distinct alignment tensors for POU_S

and POU_{HD} in the binary Oct1·DNA complex indicates that for some portion of the time the two domains do not tumble as a single, unique rigid body. These data therefore demonstrate unambiguously that POU_S and POU_{HD} must experience interdomain dynamic fluctuations (38, 40) that are fast on the chemical shift time scale even while Oct1 is bound sequence specifically to DNA. Since the sequence-specific macroscopic affinity of Oct1 for DNA is very high ($K_{ass} \approx 1.5 \times 10^{10}$ M^{-1}; Ref. 40) and the binding sites for POU_S and POU_{HD} overlap making any large scale bending motions (within a cone of semi-angle >30°; Refs. 38 and 39) physically unrealistic for the short 8 base pair Oct1 DNA binding site, one can conclude that a microscopic dissociation/association process is being observed for at least one of the two domains in the binary complex. In short either two (or more) rapidly exchanging, fully bound configurations are being observed with different relative orientations of the POU_S and POU_{HD} domains, or a conformational state(s) is being detected in which the POU_S and POU_{HD} domains are not bound to DNA simultaneously.

From gel shift assays, the equilibrium association constants for sequence specific binding of the isolated (*i.e.* non-tethered) POU_S and POU_{HD} domains to the *H2B* octamer DNA site are $\sim 5 \times 10^4$ M^{-1} and 7×10^6 M^{-1}, respectively (41). In addition, the isolated POU_S and POU_{HD} domains bind cooperatively to the octamer *H2B* site (with a cooperativity factor of ~13) although there are no direct intermolecular protein-protein interactions between them (41). This result led to the proposal that cooperativity arises through shared contacts of the two domains with the DNA involving a thymine methyl of the reverse strand at position 5 of the octamer *H2B* binding site (41). (Note that the octamer binding sequence from the *Hoxb1* regulatory element differs at this position by a T to A substitution, which may partially diminish DNA binding cooperativity between the POU_S and POU_{HD} domains). The length of flexible linker that is invisible in the electron density maps of the binary Oct1·DNA complexes is 26 residues (7, 26). The average end-to-end distance for a random coil polypeptide of this length is ~50 Å (given by $(C_n n l^2)^{1/2}$ where n is the number of residues, C_n the characteristic ratio which has a value of 7.2 for 26 residues, and l = 3.8 Å, the average Cα-Cα distance in a polypeptide; Ref. 42). One can therefore deduce from simple geometric considerations that the approximate effective local concentration of the dissociated domain in the hemi-associated binary state is ~3 mM (*i.e.* 1 molecule in a sphere of radius ~50 Å). At the concentrations of protein and DNA employed in the NMR experiments (0.5–1 mM) and in the absence of any binding cooperativity between the tethered POU_S and POU_{HD} domains, the populations of the hemi-associated states with POU_S- or POU_{HD}-free would be ~1% and ~0.005%, respectively. In the presence of cooperativity, these populations would be reduced by approximately an order of magnitude (~0.1 and 0.0005%, respectively). For the predominant hemi-associated POU_S-free state to make any significant contribution to the observed alignment tensor would therefore require that the degree of ordering of free POU_S by both phage and PEG/hexanol be in the range 0.05–0.5 (the maximum degree of ordering for nonspecific binding) compared to approximately 10^{-3} for the binary complex (which corresponds to a $|D_a^{NH}|$ of ~10 Hz).

Free Oct1 (which is positively charged) is highly ordered by negatively charged phage resulting in extreme line broadening. However, no significant line broadening of free Oct1 is observed in the presence of the neutral PEG/hexanol medium and the degree of ordering is $\sim 10^{-3}$. Since the dipolar couplings in the phage and PEG/hexanol media are highly correlated, and the rhombicities of the POU_S and POU_{HD} domains are comparable in the two media, one can conclude that the binary Oct1·DNA complex consists of two (or more), comparably populated, alternative binding modes. Given that only a single set of resonances is observed for the bound Oct1, exchange between the two (or more) states must be fast on the chemical shift scale. Since the sequence specific DNA binding affinity of POU_S is 2 orders of magnitude weaker than that for POU_{HD} (41), which corresponds approximately to a 2-fold increase in the dissociation rate constant for POU_S relative to POU_{HD}, one can conclude that, in all likelihood, the POU_{HD} domain is bound in a single configuration in the binary complex, whereas the POU_S domain can adopt two (or more) rapidly interchanging alternative bound orientations.

TABLE II
Orientation of POU_S and POU_{HD} in the ternary Oct1·Sox2·DNA complex

The table displays the agreement between observed $^1D_{NH}$ dipolar couplings measured in the ternary Oct1·Sox2·*Hoxb1*-DNA complex and the crystal structures of the combined POU_S and POU_{HD} domains in three Oct1·DNA crystal structures obtained by singular value decomposition using a single alignment tensor for POU_S and POU_{HD}. To ensure that the comparison solely reflects differences in POU_S/POU_{HD} orientations, the coordinates of the individual POU_S and POU_{HD} domains of the 1.9 Å resolution Oct1·*MORE*-DNA (26) structure were best-fitted to the orientations present in the other two structures (7, 26). The difference in the relative orientation of POU_S and POU_{HD} between the Oct1·*PORE*-DNA and Oct1·*H2B*-DNA complexes is ~16°, while that between the Oct1·*PORE*-DNA and Oct1·*MORE*-DNA complexes is 174°.

	$^1D_{NH}$ R_{dip}	
	POU_S	POU_{HD}
	%	
Oct1·*H2B*-DNA	19.9	22.1
Oct1·*PORE*-DNA	17.6	18.7
Oct1·*MORE*-DNA	40.9	36.6

In the ternary Oct1·Sox2·DNA complex, on the other hand, the magnitude of the alignment tensor (rhombicity and D_a^{NH}) for POU_S and POU_{HD} (Fig. 1*B*), as well as for Sox2, are the same, indicative of a single alignment tensor. This result directly indicates that the incorporation of Sox2 into the protein-DNA complex must cooperatively increase the site-specific affinity of Oct1 for DNA by locking the tethered POU_S domain into a single bound configuration. Isolated POU_S interacts with *Hoxb1* DNA in fast exchange on the chemical shift scale in the binary POU_S·DNA complex but in intermediate exchange in the presence of Sox2 (data not shown). This change in exchange regime in the presence of Sox2 requires at least a 10-fold decrease in the dissociation rate constant of POU_S from DNA. From the biological perspective, a reduction in the dissociation of POU_S from its relevant DNA target site will result in a corresponding increase in the overall affinity of the Oct1·DNA complex and hence the occupancy of the relevant Oct1 DNA target site, thereby promoting transcription more efficiently.

Structure Determination—Previous crystallographic work has shown that POU_S and POU_{HD} adopt different orientations while bound to DNA depending on the promoter sequence and whether Oct1 binds as a monomer or dimer (7, 26). We compared the measured dipolar couplings in the ternary complex with the orientation of POU_S and POU_{HD} in three different crystal structures (Table II): the monomeric Oct1·*H2B*-DNA complex at 3.0 Å resolution (7); the dimeric Oct1·*PORE*-DNA complex at 2.7 Å resolution that uses the same interdomain orientation as the monomeric complex (15); and the dimeric Oct1·*MORE*-DNA complex at 1.9 Å resolution that uses a distinct interdomain orientation (15). (To ensure that the comparison solely reflects differences in POU_S/POU_{HD} orientations, the coordinates of the individual POU_S and POU_{HD} domains of the 1.9 Å resolution structure were best-fitted to the orientations present in the other two structures). The dipolar couplings indicate that POU_S and POU_{HD} adopt the same relative

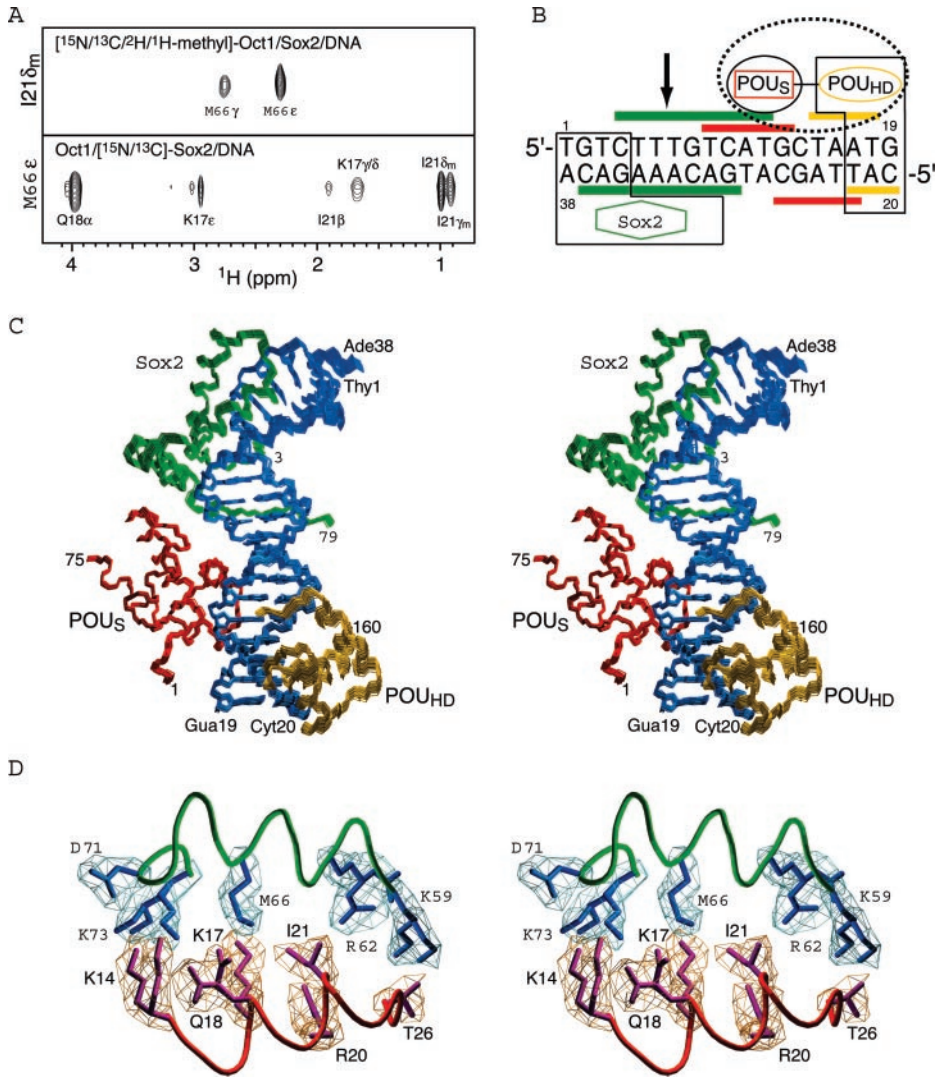

FIG. 2. **Quality of the three-dimensional structure of the ternary Oct1·Sox2·*Hoxb1*-DNA complex solved by NMR.** *A*, strips from three-dimensional ^{13}C-separated/^{12}C-filtered NOE spectra recorded at 600 MHz on [$^{15}N/^{13}C/^{2}H/^{1}H$-methyl]-Oct1·Sox2·*Hoxb1*-DNA (*upper panel*) and Oct1·[$^{15}N/^{13}C$]-Sox2·*Hoxb1*-DNA (*lower panel*) complexes, illustrating intermolecular NOEs between POU_S and Sox2. *B*, schematic illustration of the calculational strategy employed to solve the structure of the complex using conjoined rigid body/torsional angle dynamics. POU_S (*red box*), POU_{HD} (*yellow oval*), and Sox2 (*green polygon*) are shown next to their respective binding sites (*colored bars*, as described in Fig. 1). The *solid black lines* outline regions that were treated as rigid bodies (excluding interfacial side chains). The *dashed oval* indicates that the translational placement of POU_S and POU_{HD} was restrained by a weak non-crystallographic symmetry restraint with respect to the POU·*PORE*-DNA complex (26). The *arrow* indicates the intercalation site of Met-13 from Sox2. *C*, stereoview of a best-fit superposition of the final 100-simulated annealing structures (backbone of Sox2, POU_S, and POU_{HD} in *green*, *red*, and *gold*, respectively; *Hoxb1*-DNA in *blue*). *D*, stereoview of an isosurface of the reweighted atomic density map (49) (contoured at 25% maximum value) calculated from the final 100 simulated annealing structures illustrating side chains of Sox2 (*cyan*) and POU_S at the Sox2/POU_S interface. The backbone of Sox2 (*green*) and POU_S (*red*) are depicted as tubes, and residues of Sox2 are denoted in italic type. The coordinates are those of the restrained regularized mean structure; note that several long side chains are in multiple conformations.

orientation in the ternary Oct1·Sox2·DNA complex as that in the crystal structures of the Oct1·*H2B*-DNA complex and the monomer unit in the Oct1·*PORE*-DNA complex (overall R^{NH}_{dip} values of 21 and 18.5%, respectively, compared with 39% for the Oct1·*MORE*-DNA complex). The slightly better fit to the monomer unit of the Oct1·*PORE*-DNA complex compared with the Oct1·*H2B*-DNA complex reflects a small ~16° difference in the relative orientation of POU_S and POU_{HD} and a concomitant 1.3 Å backbone r.m.s. shift in the coordinates.

Orientational restraints from dipolar couplings in combination with translational information derived from limited intermolecular nuclear Overhauser enhancement (NOE) data can be used to rapidly and accurately determine the structures of large macromolecular complexes using conjoined rigid body/torsion angle dynamics (22, 34, 35, 43). Based on the excellent agreement between the dipolar couplings and the known structures of the binary complexes, the ternary complex was amenable to a variation of this approach (Fig. 2*B*). The structure of the ternary complex was solved on the basis of 345 backbone dipolar couplings, supplemented by the following intermolecular NOE data. 16 intermolecular NOEs between POU_S and Sox2 define the Oct1/Sox2 interface (see Fig. 2*A*); 48 intermolecular NOEs between Sox2 and DNA confirm that the location and binding of Sox2 to DNA is fully consistent with the known structure of the related SRY·DNA binary complex (6); and the 3 observed intermolecular NOEs between POU_{HD} and DNA confirm that the POU_{HD} binding site is the same as that in the binary complex (7, 26). Despite the absence of any observable intermolecular NOEs between POU_S and DNA (due to resonance line broadening of interfacial residues) the base pair separation between POU_S and POU_{HD} on the DNA is confirmed by both intermolecular NOEs with Sox2 and the excel-

lent agreement between the measured dipolar couplings and the crystal structure of Oct1·*PORE*-DNA complex (see above). Even a single base pair alteration in the separation between the POU_S and POU_{HD} binding sites on the DNA would necessarily change the relative orientation of POU_S and POU_{HD} by at least 30°, inconsistent with the dipolar coupling data. In addition, torsion angle restraints for residues at the POU_S/Sox2 interface were derived from heteronuclear ^3J couplings (18). The pattern of intramolecular NOEs within the DNA and the distribution of ^{31}P shifts (44) confirmed that the DNA was predominantly B-form.

On the basis of the preceding experimental data, the following conjoined rigid body/torsion angle dynamics calculational strategy was employed (Fig. 2*B*). Sox2 and POU_{HD} (excluding interfacial side chains) together with the terminal portions of their respective DNA binding sites (base pairs 1–4 and 17–19, respectively) were treated as rigid bodies. POU_S (excluding interfacial side chains) was also treated as a rigid body and a very weak non-crystallographic symmetry restraint was employed to maintain the approximate translational separation between POU_S and POU_{HD} observed in the monomer unit of the Oct1·*PORE*-DNA complex. In addition to giving side chains at the POU_S/Sox2 interface full torsional degrees of freedom, the side chains at the three protein-DNA interfaces were also allowed to move within torsion angle ranges of ±20° of those observed in the respective binary complexes. Loose distance restraints were employed to preserve the protein-DNA hydrogen bonding interactions observed in the binary Oct1·DNA complexes. The DNA from base pairs 4–16 was given full torsional degrees of freedom while being subjected to database torsion angle and base-base positional potentials of mean force (27, 28). The latter ensure that the DNA retains appropriate stereochemistry and base-base interactions while permitting the path of the DNA (including the extent of Sox2 induced DNA bending which is centered between base pairs 6 and 7 at the site of intercalation of Met-13) to be entirely determined by the relative orientation of Sox2, POU_S, and POU_{HD}, as dictated by the dipolar couplings.

Overall Structure of the Ternary Complex—The resulting ensemble of simulated annealing structures for the Oct1·Sox2·DNA ternary complex (Fig. 2*C*) is in excellent agreement with the experimental dipolar couplings (Table III). Thus, the $^1D_{NH}$, $^1D_{NC'}$, and $^1D_{C\alpha C'}$ dipolar coupling *R*-factors for the ensemble calculated with a single alignment tensor are essentially identical to those obtained when fitting each domain with individual alignment tensors (Table I). In addition, the intermolecular NOE-derived interproton distance restraints and torsion angle restraints are satisfied within their experimental uncertainties (Table III). The global orientation of the three domains and residues at the protein-protein interface are well defined (Fig. 2, *C* and *D*, respectively). Sox2 introduces a bend (~50°) in the DNA of the ternary complex (Fig. 2*C* and Fig. 3*A*) that is similar to that observed in the related SRY·DNA binary complex (~55°) (6). The interdomain orientation adopted by the POU_S and POU_{HD} domains differs from that in the Oct1·*H2B*-DNA (7) and Oct1·*PORE*-DNA (26) complexes by ~12 and 6°, respectively (with corresponding backbone r.m.s. shifts of 1.1 and 0.7 Å, respectively).

The POU_S/Sox2 Interface—The spacing of the Sox2 and Oct1 binding sites in the *Hoxb1* regulatory element brings the C-terminal halves of the first helix (residues 6–23) of POU_S and the third helix (residues 48–68) of Sox2 into close proximity (Fig. 3*A*). These two helices are oriented anti-parallel to one another and form a tight, well-packed interface over the phosphate backbone of DNA without having to distort either of the protein structures or the B-form DNA (Fig. 3, *A–C*). Despite its small size (~520 Å^2 of buried surface area, ~245 Å^2 originating from POU_S and 280 Å^2 from Sox2), the interface exemplifies a fairly common organization of protein-protein interactions (34, 35): a hydrophobic center surrounded by complementary electrostatic contacts. In particular, the hydrophobic side chains of Ile-21 of POU_S and Met-66 and Ala-63 of Sox2 (with residues of Sox2 indicated by italic type) form the core of the interface with additional hydrophobic interactions provided by the aliphatic portions of Lys-17 and Arg-62. Interestingly, Ile-21 also plays a pivotal role in the formation of Oct1 homodimers on the *PORE* B-cell-specific promoter (26). Edges of the interface include electrostatic interactions between the backbone carbonyl oxygen of Gly-24 and the amino group of Lys-59, and between the amino group of Lys-14, the carboxylate of Asp-71, and the backbone carbonyl oxygen of Pro-70 (Fig. 3*C*). Several residues at the interface participate in both protein-DNA and protein-protein contacts (*e.g.* Lys-17, Thr-26, and Arg-65). These bridging interactions likely contribute to the stabilization of the POU_S-DNA interaction upon ternary complex formation.

Comparison with the Oct1·Sox2·DNA Complex on the FGF4 Promoter Element—After we completed the determination of

TABLE III
Structural statistics for the ternary Oct1·Sox2·DNA complex[a]

	⟨SA⟩	$(\overline{SA})_r$
Number of experimental NMR restraints		
Residual dipolar couplings	345	
NOE-derived interproton distances[b]	67	
Side chain torsion angles[b]	21	
Dipolar coupling *R*-factors (%)[c,d]		
Sox2		
$^1D_{NH}$ (51)	17.7 ± 0.2	17.8
$^1D_{NC'}$ (39)	29.6 ± 0.1	29.5
$^1D_{CaC'}$ (49)	32.0 ± 0.4	32.3
POU_S		
$^1D_{NH}$ (39)	16.7 ± 0.1	16.8
$^1D_{NC'}$ (33)	24.5 ± 0.2	24.3
$^1D_{CaC'}$ (34)	29.7 ± 0.1	29.7
POU_{HD}		
$^1D_{NH}$ (39)	17.7 ± 0.1	17.5
$^1D_{NC'}$ (34)	25.0 ± 0.1	25.0
$^1D_{CaC'}$ (27)	30.4 ± 0.2	30.3
r.m.s. deviations from distance and torsion angle restraints[c,e]		
Distances (Å) (102)	0.051 ± 0.001	0.04
Torsion angles (347) (°)	0.51 ± 0.06	0.62
Coordinate precision (Å)		
Protein backbone + DNA[f]	0.29	
Side chains at POU_S/Sox2 interface	0.86	

[a] The notation of the NMR structures is as follows: ⟨SA⟩, the final 100 simulated annealing structures obtained by conjoined rigid body/torsion angle dynamics; $(\overline{SA})_r$, the restrained regularized mean coordinates derived from the ensemble of simulated annealing structures using the procedure described in Ref. 35.

[b] There are 16, 48, and 3 intermolecular, NOE-derived, interproton distance restraints at the POU_S/Sox2, Sox2/DNA, and POU_{HD}/DNA interfaces, respectively; and 18 and 3 torsion angle restraints at the POU_S/Sox2 and Sox2/DNA interfaces, respectively.

[c] The number of restraints is given in parentheses.

[d] The magnitude of the single alignment tensor is: D_a^{NH} = 10.9 Hz and η = 0.63.

[e] There are no violations in distance and torsion angle restraints ≥0.3 Å or ≥5°, respectively. In addition to the experimental restraints, the number of restraints included in the statistics comprise loose distance restraints to preserve the protein-DNA hydrogen bonding interactions observed in the binary Oct1·DNA complexes, side chain torsion angle restraints at the three protein-DNA interfaces set to within ranges of ±20° of those observed in the respective binary complexes, and loose torsion angle restraints for the sugar-phosphate backbone of the DNA (see "Experimental Procedures").

[f] Note that since the backbones of the three protein domains are treated as rigid bodies, the backbone precision does not take into account errors in the X-ray coordinates (26) of POU_S and POU_{HD} or the NMR coordinates of SRY (6). For reference, the percentage of residues in the most favored region of the Ramachandran plot for POU_S, POU_{HD}, and Sox2 is 95.5, 96.1, and 95.6%, respectively (from the program PROCHECK; Ref. 50).

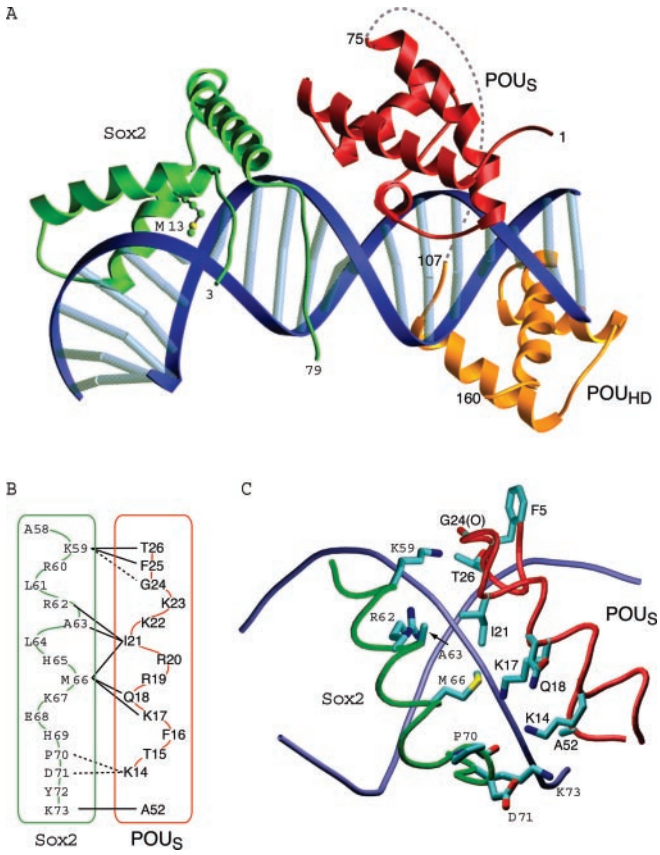

FIG. 3. **Overview of the ternary Oct1·Sox2·*Hoxb1*-DNA complex.** *A*, *ribbon diagram* of the ternary complex (Sox2, *green*; POU_S, *red*; POU_{HD}, *gold*; and *Hoxb1*-DNA, *blue*). The side chain of Met-13 that intercalates between base pairs 6 and 7 located at the center of the Sox2-induced DNA bend is also shown. A *dashed line* represents the disordered 32-residue linker that connects POU_S and POU_{HD}. (This region is disordered in solution in both the binary and ternary complexes, as well as in the binary crystal structures, Refs. 7 and 26). *B*, diagrammatic representation of POU_S/Sox2 interface in the ternary complex formed by two anti-parallel helices (Sox2, *green* and residues denoted in italic type; POU_S, *red*). *Solid* and *dashed lines* indicate van der Waals contacts and electrostatic interactions, respectively. *C*, detailed view of the POU_S/Sox2 interface. The side chain atoms are color coded according to atom type (oxygen, *red*; nitrogen, *blue*; carbon, *cyan*; sulfur, *yellow*). The backbone of Sox2, POU_S, and *Hoxb1*-DNA are depicted as tubes in *green*, *red*, and *blue*, respectively.

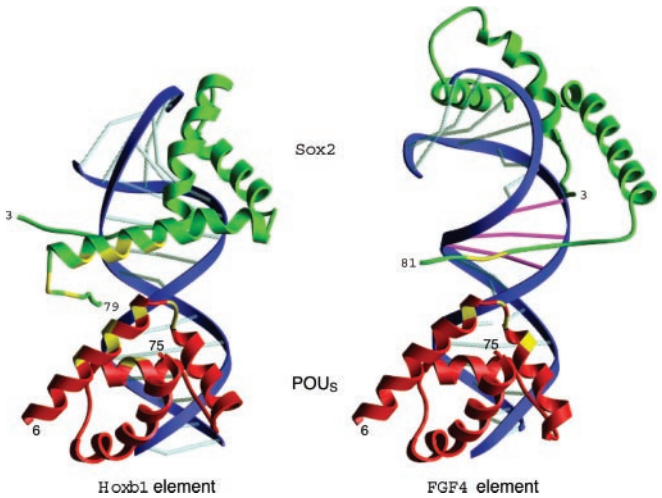

FIG. 4. **Comparison of the relative orientations of POU_S and Sox2 in the ternary Oct1·Sox2·*Hoxb1*-DNA (*left*) and Oct1·Sox2·*FGF4*-DNA (*right*) complexes.** The separation between the POU_S and Sox2 binding sites is increased by three base pairs (shown in *magenta*) in the *FGF-4* element relative to the *Hoxb1* element. This translates to a difference of 108° in the relative orientations POU_S and Sox2 in the two ternary complexes. Sox2 is displayed in *green*, Oct1 in *red*, and the DNA backbone in *blue*; the backbone of residues at the POU_S/Sox2 interface is highlighted in *yellow*. The coordinates of the Oct1·Sox2·*FGF4*-DNA complex are taken from Ref. 45 (PDB accession code 1GT0).

the solution structure of the Oct1·Sox2·*Hoxb1*-DNA ternary complex, a paper describing the 2.6 Å resolution crystal structure of the ternary complex formed by Oct1 and Sox2 bound to the *FGF4* element was published by Reményi *et al.* (45). A comparison of these two ternary complexes is of considerable interest since the separation between the Sox2 and Oct1 binding sites in the two complexes differ by three base pairs, and thus sheds light on multiple modes of interaction between Sox2 and the POU_S domain of Oct1.

The structures of the individual domains in the two complexes are very similar with backbone atomic r.m.s. differences of 1.0 Å for POU_S (residues 5–75), 0.5Å for POU_{HD} (residues 107–160) and 0.9 Å for Sox2 (residues 7–79). Despite the close similarity in the backbone coordinates, the residual dipolar coupling data do not fit the individual domains of the Oct1·Sox2·*FGF4*-DNA complex as well, reflecting the lower resolution (2.6 Å) of this crystal structure. Thus, the values of R^{NH}_{dip} for the POU_S and POU_{HD} domains in the Oct1·Sox2·*FGF4*-DNA complex are 28.7 and 30.8%, respectively, compared with 16.2 and 17.5%, respectively, for the 1.9 Å Oct1·*MORE*-DNA binary complex. Similarly, R^{NH}_{dip} for Sox2 in the Oct1·Sox2·*FGF4*-DNA complex is 42.7%, which improves to 28.4% after removing the data for Met-23 (which deviates by 25 Hz) compared with 16.0% for the solution structure of the SRY·DNA binary complex.

The dipolar coupling data, however, are in complete agreement with the interdomain orientation between the POU_S and POU_{HD} in the Oct1·Sox2·*FGF4*-DNA complex with an overall R^{NH}_{dip} of 17.2% (using the 1.9 Å resolution coordinates of the individual domains best-fitted to the coordinates of the *FGF4* ternary complex). The difference in relative orientation of the POU_S and POU_{HD} domains in the *Hoxb1* and *FGF4* ternary complexes is only 5°, which corresponds approximately to the uncertainty in the orientation of the alignment tensor (33).

By way of contrast, the dipolar coupling data are completely inconsistent with the relative orientation of Oct1 and Sox2 in the Oct1·Sox2·*FGF4*-DNA ternary complex with an overall R^{NH}_{dip} of 44% (using the coordinates of the individual POU_S, POU_{HD}, and Sox2 domains from the 1.9 Å resolution Oct1·DNA binary complex and the NMR SRY·DNA binary complex best-fitted to the coordinates of the *FGF4* ternary complex).

The *FGF4* element sequence differs from that of the *Hoxb1* autoregulatory element by a three base pair insertion between the binding sites of POU_S and Sox2 (Fig. 4). Accordingly, the relative orientation of Sox2 with respect to Oct1 changes by a 108.2° rotation around the axis of the DNA, which corresponds to a 36.1° rotation per base pair insertion, as expected for B-DNA (Fig. 4). This rotation leads to an interaction between the two domains that involves a similar, partially overlapping surface on POU_S but a very different surface on Sox2. In particular, the protein:protein interface on the *FGF4* element involves only four residues from the POU_S domain (Ile-21, Gly-24, Thr-26, and Asp-29) and two residues from Sox2 (Arg-75 and Arg-76). The buried accessible surface at the POU_S/Sox2 interface (~240 $Å^2$) on the *FGF4* element is just under half that on the *Hoxb1* element (~520 $Å^2$). Three contact residues from the POU_S domain (Ile-21, Gly-24, and Thr-26) are also involved in the *Hoxb1* ternary complex but neither of the contact residues from Sox2 is involved in the *Hoxb1* ternary complex. As can be seen in Fig. 4, the protein:protein interface in the *Hoxb1* ternary complex extends further along the first

helix of POU_S and makes use of the third helix of Sox2 instead of the extended flexible tail.

The structural differences between the two ternary complexes strongly suggest that stabilization of Oct1 binding to DNA in the presence of Sox2 will be significantly greater on the *Hoxb1* element than on the *FGF4* element. We suggest that this difference is the major contributory factor with regard to the biological selectivity for the POU domain of Oct3 on the *FGF4* element. The linker length between the POU_S and POU_{HD} domains is 7 residues shorter in Oct3 than in Oct1. This translates to a ~2-fold increase (given by $[(C_{Oct1}\cdot n_{Oct1})/(C_{Oct3}\cdot n_{Oct3})]^{3/2}$, where n_{Oct1} and n_{Oct3} are the flexible linker lengths in Oct1 and Oct3, respectively, and C_{Oct1} and C_{Oct3} are the corresponding characteristic ratios for those linker lengths (42)) in local concentration of the dissociated domain in the hemi-associated state, and hence a corresponding increase in the overall equilibrium association constant. Although the smaller protein-protein interface between Oct3 and Sox2 on the *FGF* promoter would lead to a reduction in binding cooperativity, the higher intrinsic sequence specific affinity of Oct3 for DNA due to its shorter linker length would allow ternary complex formation to increase the Oct3 affinity for DNA beyond a critical point necessary to synergistically drive transcription. Thus, we propose that the shorter linker length for Oct3 relative to Oct1 offsets the reduction in the POU_S/Sox2 interface on the *FGF4* element.

Reményi *et al.* (45) also derived a model of the Oct3/4·Sox2 ternary complex formed on the *UTF1* regulatory region from their crystal structure of the Oct1·Sox2·*FGF4*-DNA complex. The *UTF1* regulatory element contains variations of the Sox2 and Oct consensus binding sequences with the same relative spacing as that in the *Hoxb1* element sequence studied here. Based on their model, Reményi *et al.* (45) correctly predicted an important contact between Ile-21 of Oct3/4 and Met-66 of Sox2. The involvement of these two residues is further supported by mutagenesis studies which are consistent with both the NMR structure of the Oct1·Sox2·*Hoxb1*-DNA complex and the model of the Oct3/4·Sox2·*UTF1*-DNA complex: the Sox2 double mutants K59E/R62E and R62E/M66E, as well as the Oct3/4 double mutant I21Y/D29E prevent the formation of a ternary complex on the *UTF1* element (45). However, the POU_S/Sox2 interface in the Oct3/4·Sox2·*UTF1*-DNA model, while similar to that in the Oct1·Sox2·*Hoxb1*-DNA complex determined by NMR, appears to be less extensive. In particular, it was proposed (45) that the *UTF1* ternary complex involves the same residues of Oct1 employed in the *FGF4* ternary complex and a segment of helix 3 from Sox2 (residues Lys-59 to Met-66, using the current numbering scheme). Since the spacing between the Oct and Sox2 binding sites on the *UTF1* and *Hoxb1* regulatory elements is identical, we conclude that the POU_S/Sox2 interface on the *UTF1* element actually involves more residues from both proteins (Lys-14, Lys-17, Gln-18, and Ala-52 from POU_S; and Pro-70, Asp-71, and Lys-73 from Sox2; see Fig. 3*B*).

Nishimoto *et al.* (8) originally proposed that selectivity, in terms of transcriptional activation, for Oct3 on the *UTF1* regulatory element was due to an altered POU_{HD} DNA binding sequence (TAGT *versus* TAAT). Based on the structure reported here, we would add to this hypothesis that the reduced affinity imposed by the altered DNA binding sequence on the *UTF1* element is compensated by the larger interprotein interface with Sox2 relative to that on the *FGF4* element. In the case of Oct1, however, this compensation is not sufficient, because of the longer interdomain linker length, which, as discussed above, would be predicted to lower the affinity of Oct1 *versus* Oct3 for DNA.

Concluding Remarks—Here we present the solution structure of a ternary complex formed by the DNA binding domains of Oct1 and Sox2 bound to the *Hoxb1* autoregulatory element. Because of the large size of this complex (M_r ~42,000) by NMR standards, a combination of techniques was required to solve the structure. These include the extensive use of dipolar couplings to accurately orient the domains, translational restraints from intermolecular NOE data, various isotope-labeling schemes to simplify the NMR spectra, and the application of conjoined rigid body/torsion angle dynamics based on the known structures of the binary complexes. To our knowledge, this structure represents the first ternary macromolecular complex solved by NMR to date. The techniques employed provide unique insights into how direct interaction of Oct1 and Sox2 leads to synergistic transcription regulation. In particular, alternative binding orientations of POU_S evident in the Oct1·DNA binary complex are quelled by direct interaction with the DNA binding domain of Sox2.

The above model helps explain how other binding partners for Oct1 can achieve the same synergism in transcriptional activation. Crystal structures of small peptides from the SNAP190 subunit of the $SNAP_C$ transcription factor (46) and the OCA-B B-cell specific coactivator (47, 48) show that both peptides make contact with DNA and the POU domains of Oct1. Although these peptides recognize different regions of Oct1 than those contacted by Sox2, these interactions are still likely to lead to a similar cooperative increase in DNA binding affinity for Oct1. Combinatorial control of transcription depends on bringing together transcription factors with different binding partners in a cell-specific, temporally regulated manner. Using the same transcription factors in different contexts requires adaptability of both the protein and DNA binding surfaces. Decreased dissociation of the POU_S domain brought about by a relatively small hydrophobic protein:protein interface provides the mechanism by which the Oct1·Sox2·*Hoxb1*-DNA ternary transcription factor complex meets these requirements.

Acknowledgments—We thank J. Iwahara, A. Bax, C. Bewley, J. Boisbouvier, J. Chou, L. Dailey, and J. Wu for helpful discussions and advice; F. Delaglio, D. Garrett, and J. Baber for technical support.

REFERENCES

1. Levine, M., and Tjian, R. (2003) *(2003) Nature* **424,** 147–151
2. Ptashne, M., and Gann, A. (2002) *Genes & Signals*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
3. Wolberger, C. (1999) *Annu. Rev. Biophys. Biomol. Struct.* **28,** 29–36
4. Kamachi, Y., Uchikawa, M., and Kondoh, H. (2000) *Trends Genet.* **16,** 182–187
5. Dailey, L., and Basilico, C. (2001) *J. Cell. Physiol.* **186,** 315–328
6. Murphy, E. C., Zhurkin, V. B., Louis, J. M., Cornilescu, G., and Clore, G. M. (2001) *J. Mol. Biol.* **312,** 481–499
7. Klemm, J. D., Rould, M. A., Aurora, R., Herr, W., and Pabo, C. O. (1994) *Cell* **77,** 21–32
8. Nishimoto, M., Fukushima, A., Okuda, A., and Muramatsu, M. (1999) *Mol. Cell. Biol.* **19,** 5453–5465
9. Ambrosetti, D., Schöler, H. R., Dailey, L., and Basilico, C. (2000) *J. Biol. Chem.* **275,** 23387–23397
10. Di Rocco, G., Gavala, A., Popperl, H., Krumlauf, R., Mavilio, F., and Zappavigna, V. (2001) *J. Biol. Chem.* **276,** 20506–20515
11. Gaufo, G. O., Flodby, P., and Capecchi, M. R. (2000) *Development* **127,** 5343–5354
12. Goto, N. K., Gardner, K. H., Mueller, G. A., Willis, R. C., and Kay, L. E. (1999) *J. Biomol. NMR* **13,** 368–374
13. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6,** 27–293
14. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) *J. Magn. Reson.* **94,** 214–220
15. Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 12366–12371
16. Tugarinov, V., Muhandiram, R., Ayed, A., and Kay, L. E. (2002) *J. Am. Chem. Soc.* **124,** 10025–10035
17. Clore, G. M., and Gronenborn, A. M. (1998) *Trends Biotechnol.* **16,** 22–34
18. Bax, A, Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R., and and Zhu, G. (1994) *Methods Enzymol.* **239,** 79–106
19. Clore, G. M., Starich, M. R., and Gronenborn, A. M. (1998) *J. Am. Chem. Soc.* **120,** 10571–10572
20. Hansen, M. R., Hanson, P., and Pardi, A. (2000) *Methods Enzymol.* **317,** 220–240

21. Yang, D., Venters, R. A., Mueller, G. A., Choy, W. Y., and Kay, L. E. (1999) *J. Biomol. NMR* **14,** 333–343
22. Schwieters, C. D., and Clore, G. M. (2001) *J. Magn. Reson.* **152,** 288–302
23. Schwieters, C. D., Kuszewski, J., Tjandra, N., and Clore, G. M. (2003) *J. Magn. Reson.* **160,** 66–74
24. Schwieters, C. D., and Clore, G. M. (2001) *J. Magn. Reson.* **149,** 239–244
25. Carson, M. (1991) *J. Appl. Crystallogr.* **24,** 958–961
26. Reményi, A., Linz, K., Nissen, L. J., Reinbold, R., Scholer, H. R., and Wilmanns, M. (2001) *Mol. Cell* **8,** 569–580
27. Kuszewski, J., Schwieters, C. D., and Clore, G. M. (2001) *J. Am. Chem. Soc.* **123,** 3903–3918
28. Clore, G. M., and Kuszewski, J. (2003) *J. Am. Chem. Soc.* **125,** 1518–1525
29. Huang, K., Louis, J. M., Donaldson, L., Lim, F. -L., Sharrocks, A. D., and Clore, G. M. (2000) *EMBO J.* **19,** 2615–2628
30. Bax, A., Kontaxis, G., and Tjandra, N. (2001) *Methods Enzymol.* **339,** 127–174
31. Clore, G. M., and Garrett, D. S. (1999) *J. Am. Chem. Soc.* **121,** 9008–9012
32. Ottiger, M., and Bax, A. (1998) *J. Am. Chem. Soc.* **120,** 12334–12341
33. Zweckstetter, M., and Bax, A. (2002) *J. Biomol. NMR* **23,** 127–137
34. Wang, G., Louis, J. M., Sondej, M., Seok, Y.-J., Peterkofsky, A., and Clore, G. M. (2000) *EMBO J.* **19,** 5635–5649
35. Cornilescu, G., Lee, B. R., Cornilescu, C. C., Wang, G., Peterkofsky, A., and Clore, G. M. (2002) *J. Biol. Chem.* **277,** 42289–42298
36. Zweckstetter, M., and Bax, A. (2000) *J. Am. Chem. Soc.* **122,** 3791–3792
37. Scaffidi, P., and Bianchi, M. E. (2001) *J. Biol. Chem.* **276,** 47296–47302
38. Braddock, D. T., Cai, M., Baber, J. L., Huang, Y., and Clore, G. M. (2001) *J. Am. Chem. Soc.* **123,** 8634–8635
39. Ruckert, M., and Otting, G. (2000) *J. Am. Chem. Soc.* **122,** 7793–7797
40. Braddock, D. T., Louis, J. M., Baber, J. L., Levens, D., and Clore, G. M. (2002) *Nature* **415,** 1051–1056
41. Klemm, J. D., and Pabo, C. O. (1996) *Genes Dev.* **10,** 27–36
42. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry Part III: The Behavior of Biological Macromolecules*, pp. 979–1018, W. H. Freeman & Co, San Francisco
43. Clore, G. M. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 9021–9025
44. Gorenstein, D. G. (1994) *Chem. Rev.* **94,** 1315–1338
45. Reméniy, A., Lins, K., Nissen, L. J., Reinbold, R., Schöler, H. R., and Wilmanns, M. (2003) *Genes Dev.* **17,** 2048–2059
46. Hovde, S., Hinkley, C. S., Strong, K., Brooks, A., Gu., L., Henry, R. W., and Geiger, J. (2002) *Genes Dev.* **16,** 2772–2777
47. Chasman, D., Cepek, K., Sharp, P. A., and Pabo, C. O. (1999) *Genes Dev.* **13,** 2650–2657
48. Sauter, P., and Matthias, P. (1998) *Mol. Cell. Biol.* **18,** 7397–7409
49. Schwieters, C. D., and Clore, G. M. (2002) *J. Biomol. NMR* **23,** 221–225
50. Laskowski, R. A., MacArhur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291