DNA-Binding and Regulation Mechanisms of the SIX Family of Retinal Determination Proteins[†]

Shengyong Hu, Aygun Mamedova, and Rashmi S. Hegde*

Division of Developmental Biology, Cincinnati Children's Hospital Medical Center, and Department of Pediatrics, University of Cincinnati School of Medicine, 3333 Burnet Avenue, Cincinnati, Ohio 45229

Received November 1, 2007; Revised Manuscript Received January 16, 2008

ABSTRACT: The Six/sine oculis proteins are homeodomain transcription factors that are part of the Pax/Eya/Six/Dach retinal determination cascade involved in embryonic cell fate determination. There are six mammalian Six homologues, divided into three classes on the basis of sequence homology. In the present study we examined the DNA-binding specificity and mechanisms of Six2 and Six6 toward the Trex/MEF3 consensus sequence and the core tetranucleotide ATTA commonly recognized by homeodomain proteins. The results suggest that the Six homeodomain does not bind DNA owing to the absence of a key structural feature, the basic N-terminal arm, implicated in canonical homeodomain−DNA binding. Furthermore, the DNA-binding mechanisms and DNA sequence specificity differ among these Six proteins despite the complete conservation of predicted DNA-contacting residues in their homeodomains. Inclusion of 14 amino acid residues immediately C-terminal to the homeodomain of Six6 yields a protein construct able to bind both DNA sequences tested with nanomolar affinity. However, an analogous Six2 construct remains unable to bind DNA. Furthermore, we show that the DNA-binding affinity of Six2 is increased nearly 12-fold by complex formation with the Eyes Absent tyrosine phosphatase, while Six6−DNA binding is not similarly enhanced. This phenomenon could contribute to the synergy observed between Six2 and Eyes Absent in transcriptional activation and in eye development.

The Six[1] (Sine oculis homeobox) family of proteins are transcription factors involved in embryonic cell fate determination (reviewed in ref *1*). Through interaction with DNA and other proteins, they play critical roles in fundamental biological phenomena including transcriptional activation and repression, replication licensing, and signal transduction. Mutations in the Six genes are associated with several developmental disorders, and Six gene expression is upregulated in many cancers. The Six proteins were initially identified as homologues of *Drosophila* sine oculis (SO), a component of the retinal determination cascade (reviewed in ref *2*). The other components of this regulatory cascade are the transcription factor and phosphatase Eyes Absent (Eya) (*3*–*6*), the winged-helix family transcription factor Dachshund (Dach) (*7*), and the paired domain transcription factor (Pax6). Prevailing models hold that the Six/SO proteins bind and translocate the Eya proteins to the nucleus where the complex binds DNA, perhaps leading to transcriptional activation (reviewed in refs *2*, *8*, and *9*).

Six proteins can be placed into three subfamilies on the basis of molecular phylogenetic analysis of their amino acid sequences. These are Six1/2 (*Drosophila* SO homologues), Six3/6 (*Drosophila* OPTIX homologues), and Six4/5 (*Drosophila* D-SIX4 homologues) (*10*). Mammalian Six proteins have two regions of high sequence conservation: a 59-residue homeodomain (HD) and a 115–123 residue "Six" domain (SD) N-terminal to the HD. The Six HD belongs to the K50 class in which the key DNA-contacting Asn at position 50 of the HD is replaced by a Lys (*11*). The Six domain is highly conserved among the Six proteins (43% identity) but bears no resemblance at the primary structural level to proteins/domains of known structure or function. C-terminal to the Six and homeodomains is a region of low conservation in terms of both length and sequence. The regulatory targets of some Six proteins are known, and Six DNA-binding sites have been identified using DNA footprinting and related techniques. Six4 (*12*), Six2, and Six5 (*13*) bind to the "ARE" regulatory element of the Na^+/K^+ ATPase α1 subunit gene, GGTGTCAGGTTGC. The closely related MEF3 motif, TCAGGTT (underlined in the ARE motif), present in the myogenin promoter is also recognized by Six1 and Six4 (*14*). Thus the consensus Six1/2/4/5 binding Trex/MEF3 sequence is TC[G/A]GGT[G/T] (referred to here as Myo-MEF3). In PCR-based selection experiments Six3 was shown to bind with high affinity to the classical homeodomain core tetranucleotide ATTA (*15*), suggesting that the DNA-binding specificity of the Six3/6 class of Six proteins differs from that of the Six1/2 and Six4/5 families. The various Six proteins also differ in their biological functions. Studies in *Drosophila* have typically used rescue of the eyeless phenotype in *so*[1] mutants or the ability to induce ectopic eyes as measures of the *in vivo* functionality of the Six proteins. Differences among the *Drosophila* Six homologues

[†] This work was carried out with the support of NIH Grant EY014648 to R.S.H.

* To whom correspondence should be addressed. Telephone: (513) 636-5947. Fax: (513) 636-6772. E-mail: rashmi.hegde@cchmc.org.

[1] Abbreviations: SIX, sine oculis homeobox; EYA, Eyes Absent; HD, homeodomain; SD, Six domain; SO, sine oculis; DACH, dachshund; NTA, N-terminal arm; GST, glutathione *S*-transferase; GT, glutathione.

10.1021/bi702186s CCC: $40.75
Published on Web 02/23/2008

SO and OPTIX in these rescue and gain-of-function experiments have been ascribed to differences in protein partner selection via the Six domains (*16, 17*).

Here we present data suggesting that Six2 and Six6 differ in their DNA-binding mechanisms and in the consequences of interaction with the protein tyrosine phosphatase Eyes Absent on Six2/6−DNA-binding affinity. Modulation of homeodomain−DNA-binding affinity by a non-DNA-binding cofactor represents an unusual example of functional fine-tuning within a homeodomain family.

EXPERIMENTAL PROCEDURES

Generation of Recombinant Proteins. Human SIX6 and mouse Six2 cDNA were obtained from ATCC (American Type Culture Collection) and used as PCR templates to insert the coding sequences for the constructs Six6(HD) (residues 128–186), Six6(NHD) (residues 112–186), Six6(HDC) (residues 128–200), Six6(NHDC) (residues 112–200), Six2(HD) (residues 124–182), Six2(NHD) (residues 108–182), Six2-(HDC) (residues 124–196), and Six2(NHDC) (residues 108–196) into the *Eco*RI and *Xho*I sites of the vector pGEX-4T to generate GST-fusion proteins. The constructs Six6(HDC) (residues 128–200) and Six6(HDC-ΔNTA) (residues 136–200) were generated by Gateway recombinational cloning (Invitrogen, Carlsbad, CA) and inserted into the destination vector pDEST-HisMBP (*18*) to generate dual-tagged His_6-maltose binding protein−fusion proteins, while full-length Six6 was similarly inserted into the destination vector pDEST-527 to obtain a His_6-tagged protein. The construct Six6(HD)Pitx-2(NTA) was generated using the QuikChange mutagenesis kit (Stratagene, La Jolla, CA) and introducing Arg residues at positions 2, 3, and 5 of the Six6 homeodomain as found in Pitx2. Full-length Six2 was PCR cloned into pET28a to obtain a His_6-tagged protein. Eya3(ED) (residues 238–510) was overexpressed and purified as described previously (*4*). Eya3(179–510) was inserted into the destination vector pDEST-527 by recombinational cloning to obtain a His_6-tagged protein.

All of the above overexpression vectors were transformed into the *Escherichia coli* strain BL21(DE3)RIL, grown at 37 °C and induced with 0.3 mM IPTG at 15 °C. For the GST-fusion proteins, cells were lysed in 20 mM Tris, pH 8, 150 mM NaCl, and 1% Triton X-100 by sonication. The lysis supernatant was applied to GT−agarose beads. The proteins were released from the beads by treatment with thrombin. The His_6-MBP-tagged proteins were purified by Ni-nitrilotriacetic acid resin chromatography. The proteins were eluted from the resin by imidazole gradients and then treated with Tobacco vein mottling virus protease overnight. His-tagged Six2 and Six6 proteins were immobilized on Ni-nitrilotriacetic acid resin columns and eluted with imidazole gradients. All of the proteins were further purified by size exclusion chromatography (Superdex-75 or Superdex-200 columns). All proteins used in DNA-binding studies were >95% pure as estimated by SDS−PAGE and Coomassie staining.

Generation of DNA Probes for Electrophoretic Mobility Shift Analyses (EMSA). PAGE-purified DNA oligonucleotides from Integrated DNA Technologies were used in the binding studies. The sequences and structures of the hairpin oligonucleotides used are presented in Figure 1b.

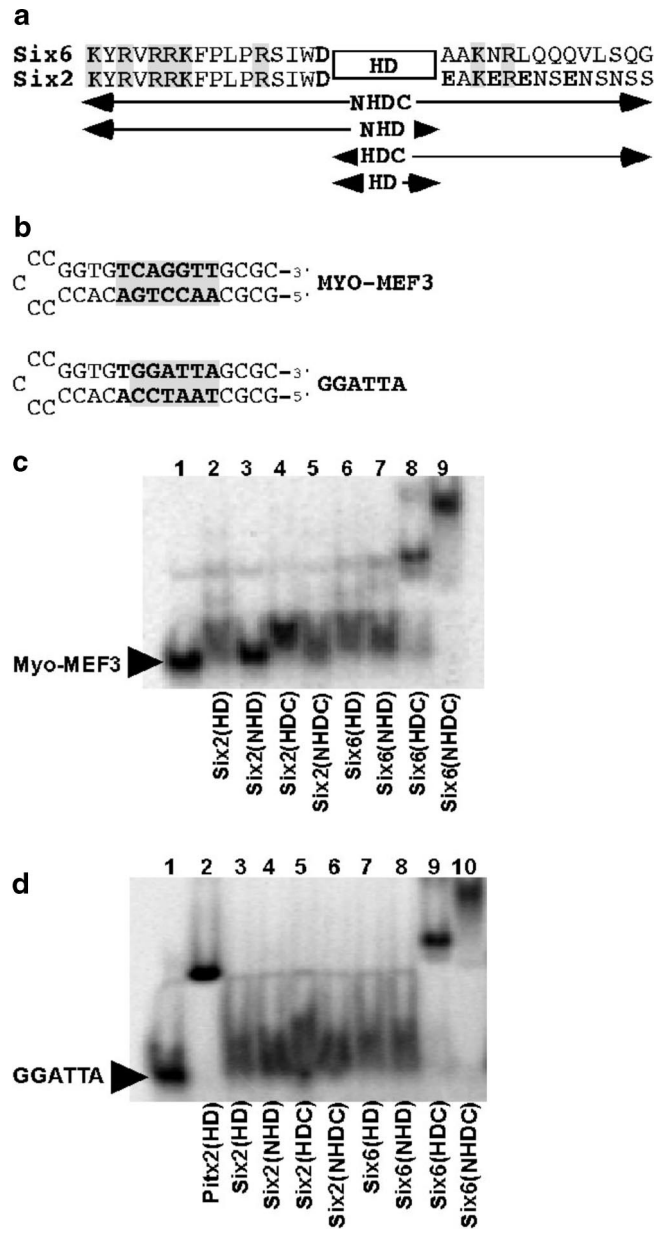

FIGURE 1: (a) Schematic representation of the homeodomain constructs used in experiments. The 59-residue homeodomain (HD) of the Six proteins is represented by an open rectangle. The sequences of the 16 conserved residues N-terminal to the HD and the 14 residues immediately C-terminal to the HD for both Six2 and Six6 are shown. Acidic residues are in bold letters, and basic residues are highlighted. The extent of the HD, NHD, HDC, and NHDC constructs is indicated below the sequences. (b) Oligonucleotide probes used in gel mobility shift experiments. The 5′ and 3′ termini of the probes are shown. The TCAGGTT (Myo-MEF3) and GGATTA sequences being tested for Six binding are highlighted. Flanking base pairs are constant in both probes. (c, d) EMSA of Six2(HD), Six2(NHD), Six2(HDC), Six2(NHDC), Six6(HD), Six6(NHDC), Six6(HDC), Six6(NHDC), and Pitx2(HD) using either the Myo-MEF3 (panel c) or the GGATTA (panel d) probe. In each gel lane 1 is the DNA probe alone, and lanes 2–10 are the various protein constructs indicated. Equivalent amounts of all proteins, in excess over the DNA probe concentration, were used.

Oligos were 5′ end labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase (New England Biolabs). Complementary regions of the oligonucleotides were annealed by heating the labeled probe to 90 °C for 10 min and transferring immediately to ice for 10 min. The only

detectable products were the correctly annealed oligonucleotide hairpins.

DNA-Binding Studies. Quantitative gel mobility shift experiments were conducted following published protocols (*19, 20*). Serial dilutions of the proteins were added to binding buffer, resulting in a final mixture containing 200 pM γ-^{32} P-labeled oligonucleotide, 150 mg/mL bovine serum albumin, 5 μg/mL sonicated salmon sperm DNA, 5 mM DTT, 25 mM HEPES, pH 7.9, 150 mM KCl, 5 mM $MgCl_2$, and 12.5% (v/v) glycerol. The reaction mixtures were incubated in a water bath at 23 °C for 30 min. Mixtures were then loaded directly onto prerun 12% polyacrylamide gels in 0.5× TBE buffer while the gel was running. Electrophoretic separation was run at constant voltage (200 V) for ~90 min, the gels were dried, and the reaction products were visualized by exposure to phosphor storage plates scanned using a PhosphorImager (Molecular Dynamics).

The density of the electrophoretic band representing the protein−DNA complex was quantitated using the ImageQuant software in volume integration mode (Molecular Dynamics). Background density was subtracted from each pixel within the integrated volume to yield the integrated density of the band. Binding isotherms were obtained by monitoring the density of the electrophoretic band representing the protein−DNA complex as a function of protein concentration and analyzed by nonlinear least-squares analysis using ORIGIN version 7.5 (OrginLab Corp., Northampton, MA). The DNA-binding activity of the protein preparations was determined from stoichiometric titrations; the values presented are corrected for this activity. Since the γ-^{32}P-labeled oligonucleotide concentration is much lower than the equilibrium dissociation constants being measured, the approximation that total protein concentration is equal to free protein concentration is made in the analysis. The apparent equilibrium binding constant, K_{eq}, was determined by analysis of the titration curves against the equations

$$p = p_{lower} + (p_{upper} - p_{lower})\bar{Y} \quad (1)$$

$$\bar{Y} = \frac{K_{eq}[X]}{1 + K_{eq}[X]} \quad (2)$$

where $\bar{Y}$ is the fractional saturation of DNA, K_{eq} is the apparent equilibrium binding constant, [X] is the free active protein monomer concentration, p is the apparent saturation calculated from the integrated density of the binding site, and p_{lower} and p_{upper} are the lower and upper limits, respectively, of the titration curve.

Protein Interaction Experiments. His_6-Six2 or His_6-Six6 was incubated with GST-Eya3(238–510), GST-Eya3(179–510), or GST in 20 mM Tris, pH 8, 0.2% NP40, and 500 mM NaCl at 4 °C for 30 min, and the mixture was then loaded on GT−agarose beads. The beads were washed extensively with 20 mM Tris, pH 8, 500 mM NaCl, and 0.2% NP40. Aliquots of the initial protein mixtures used in these experiments, the flow-through from the GT−agarose affinity step, the last wash, and the proteins retained on the beads were run on SDS−PAGE gels and stained with Coomassie blue. As a negative control the interaction of GST-Eya3(238–510) with another transcription factor, the DNA-binding domain of bovine papillomavirus E2 [purified as described previously (*21*)], was tested using similar protocols.

RESULTS

The DNA-Binding Domains of Six2 and Six6 Are Not the Same. In order to define the DNA-binding properties of the Six HD, we tested the minimal 59-residue HD (boundaries defined by alignment with other well-studied HDs; homeodomain definitions reviewed in ref *22*) of Six2 and Six6 in electrophoretic mobility shift assays (EMSA). Two DNA probes were used containing either the Six1/2/4/5 consensus MEF3 sequence TCAGGTT (Myo-MEF3) or the GGATTA core tetranucleotide (underlined) containing the Six3/6 binding site identified by PCR selection (*15*). The GG nucleotides 5′ to the core ATTA in the latter probe are the preferred nucleotides identified in previous studies on the K50 class of HD proteins (*11*), to which the Six proteins belong. Hairpin DNA probes (Figure 1b) were used to ensure proper annealing. The sequence context of the binding sites was kept consistent in both probes. In the absence of protein the labeled DNA probe was visible as a single band. Protein−DNA binding was detected by the presence of bands with reduced mobility on the gels. All of the Six HD constructs used in these experiments (Figure 1a) were purified to greater than 95% homogeneity and contained a Cys to Ser replacement at position 7 of the HD (highlighted in Figure 2a) to avoid oxidation-induced dimerization of the recombinant proteins. The resulting proteins were monomeric in solution as judged by elution profiles on calibrated size exclusion columns and dynamic light scattering analysis (not shown). In these initial experiments designed to identify the DNA-binding domain of the Six proteins, equivalent molar concentrations of the various recombinant purified proteins in excess over the DNA probe concentration were used. The results are summarized in Figure 1c.

The homeodomains of Six2 and Six6 do not yield a clear shifted band, suggesting that they do not form a protein−DNA complex with sufficient stability to be detected by EMSA. In contrast, when a GST-tagged Six2 or Six6 HD was used, we reproducibly saw a loss of free DNA probe and a smeared shifted band (not shown). Thus, to avoid GST tag-induced modification of the intrinsic DNA-binding properties of the Six proteins, we did not use GST-tagged proteins for further experiments. The positive control in these studies was the 60-residue HD of Pitx2, another K50 class homeodomain; the Pitx2 HD bound with high affinity to the GGATTA DNA probe in parallel experiments (Figure 1d, lane 2).

In order to establish the minimal DNA-binding domain of the Six proteins, constructs including various extensions at the N- and C-terminal end of the HD were generated (Figure 1a). The extensions include 16 residues N-terminal to the Six HD that are conserved among the Six proteins and include several basic residues (construct NHD) or 14 residues C-terminal to the Six HD (HDC) that are less well conserved but are predicted to form an extension of the third helix in secondary structure prediction algorithms. The results of EMSA on these constructs suggest that the minimal DNA-binding domains of the Six2 and Six6 proteins are not the same. In the case of Six6 inclusion of the 14-residue stretch immediately C-terminal to the HD is sufficient to allow the formation of a readily detectable protein−DNA complex (Figure 1c,d). The same is not true of Six2; none of the Six2 HD constructs tested here resulted in a clearly shifted DNA band. Substitution of the 14 residues in the C-extension of

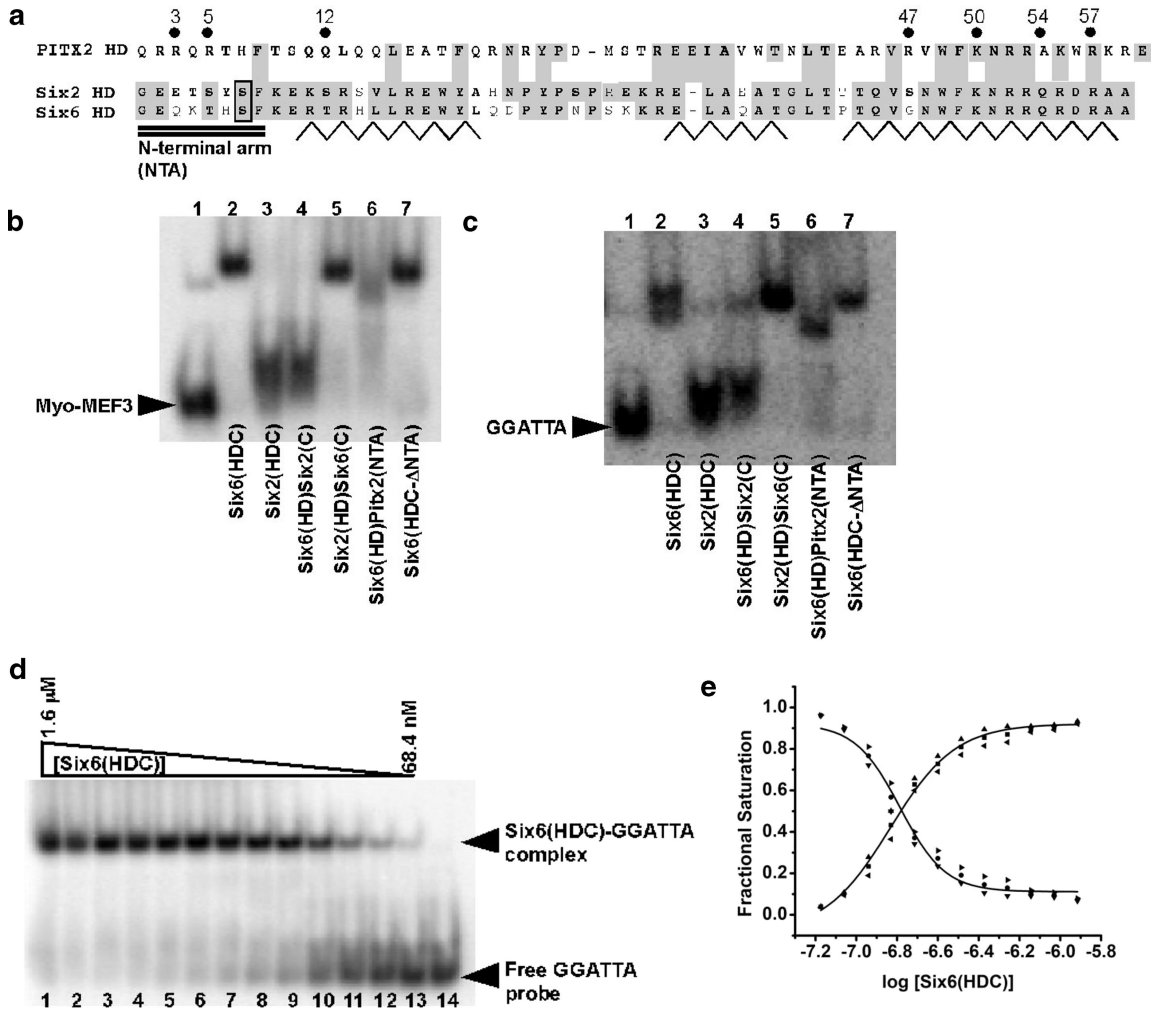

FIGURE 2: (a) Amino acid sequences of the 59-residue homeodomains of Six2 and Six6 are shown aligned with the sequence of the Pitx2 homeodomain used as a positive control in some experiments reported here. Similar residues are highlighted. The closed circles above the sequences indicate residues typically involved in HD−DNA recognition. The numbering shown indicates the position of the residues within the canonical 60-residue homeodomain. The predicted three α-helices are indicated by the zigzag lines. The N-terminal arm (NTA) that is found in the minor groove in most HD−DNA complexes is indicated by the double lines. These NTA residues are deleted in the construct Six6(HDC-ΔNTA). (b, c) EMSA of Six6(HDC-ΔNTA), Six6(HD)Pitx2(NTA), Six2(HD)Six6(C), Six6(HD)Six2(C), Six2(HDC), and Six6(HDC) with Myo-MEF3 (panel b) and GGATTA (panel c) DNA probes. In both gels lane 1 represents DNA alone. Equivalent amounts of all proteins, in excess over the DNA probe concentration, were used. (d) Representative quantitative titration. The binding of Six6(HDC) to the GGATTA probe is shown. Lane 14 is free DNA. Lanes 1−13 represent decreasing concentrations of protein. (e) Representative binding curve. The fraction of DNA probe GGATTA bound by Six6(HDC) is plotted as a function of the log protein concentration.

Six2(HDC) with those found in Six6(HDC) resulted in DNA binding by Six2(HD)Six6(C) (Figure 2b,c). The converse was not true; when the C-extension of Six6 HDC was replaced by that of Six2 HDC, the altered Six6(HD)Six2(C) lost its ability to bind DNA. Thus residues in the C-extension of Six6 specifically enable DNA binding of the Six HD. Of note is that Six2(HD)Six6(C) bound both probes tested (Myo-MEF3 and GGATTA) with comparable affinity.

In order to derive a qualitative assessment of the contributions of the N-terminal arm and the C-terminal extension to Six−DNA binding, we measured apparent equilibrium binding constants (K_{eq}) for Six6 constructs that yielded a clearly defined shifted band in equilibrium titrations, assuming a monomeric protein binding to a single site on the DNA probe. A representative gel and binding curve are shown in panels d and e of Figure 2. The results summarized in Table 1 show that Six6(HDC) binds both DNA probes with comparable affinity and that inclusion of the N-extension does not significantly alter DNA-binding affinity or specificity. Overall, these results suggest that the homeodomain of the Six proteins is not sufficient for DNA-binding and that while the 14 residues C-terminal to the Six6 HD can confer DNA-binding ability to the Six6 HD, the same is not true of the Six2 14-residue C-extension.

Table 1: Apparent Equilibrium Binding Constants (K_{eq}) for Six6 Homeodomain Constructs toward DNA Probes Containing the Consensus Sequence TCAGGTT Found in the Myogenin Promoter (Myo-MEF3) and the Sequence (GGATTA) Typically Recognized by K50 Class Homeodomain Proteins

	K_{eq} (nM)	
protein construct	TCAGGTT (Myo-MEF3)	GGATTA
Six6(HDC)	165 (±12)	158 (±7)
Six6(ΔNTA-HDC)	76 (±2.8)	135 (±3.5)
Six6(NHDC)	44 (±2)	66 (±1.8)

The N-Terminal Arm of the Six HD Lacks Basic Residues Found in Canonical HD Proteins. The homeodomain of the Six proteins lacks the basic "N-terminal arm" (NTA; residues 1–8 of the homeodomain; see Figure 2a) present in most other homeodomain proteins (*23*). Arg residues are generally

present at positions 2, 3, and 5, and they contact both the bases and the backbone of DNA. The number of positive charges on the N-terminal arm has been reported to be important for HD function, with a net charge of +3 or +4 in the seven amino acid region resulting in the highest affinity (*24*). The corresponding region in the Sine oculis, Six1/2/4/5 proteins has a net charge of −2, and in the Optix, Six3/6 proteins there is one positive and one negative charge. Thus the N-terminal arm does not have the electrostatic characteristics that would allow it to wrap into the minor groove, as is the case with most homeodomain proteins examined thus far. In order to test the possibility that the lack of basic residues in the N-terminal arm may contribute to the inability of the Six HD to bind DNA, we introduced arginines at positions 2, 3, and 5 of Six6(HD) [Six6(HD)Pitx2(NTA)]; arginines are present at these positions in the Pitx2 homeodomain (Figure 2a) which was used as a positive control in our experiments. This altered Six6(HD)Pitx2(NTA) protein was able to bind both DNA probes tested under conditions identical to those used for the wild-type Six6(HD) (Figure 2b,c).

To further examine the role of the Six N-terminal arm in DNA interactions, we postulated that the lack of basic residues in the N-terminal arm region of the Six homeodomain would make its contribution to DNA-binding affinity smaller than in the case of classical HD proteins. Indeed, a construct of Six6 HDC with a deletion of residues 1–8 of the HD [Six6(HDC-ΔNTA); Figure 2b,c and Table 1] binds DNA with affinity that is close to that measured for Six6(HDC). In classical homeodomain proteins the deletion of this region results in significantly lowered binding affinity (100-fold) (*25, 26*). Taken together, these results suggest that the DNA-binding mechanism of the Six homeodomain differs from that of well-characterized HD proteins. Specifically, the N-terminal arm is unlikely to make minor groove contacts contributing to either specificity or affinity.

Full-Length Six2 and Six6 Differ in Their DNA Sequence Preferences, and Interaction with Eyes Absent Dramatically Increases Six2−DNA-Binding Affinity. The full-length Six proteins include a conserved protein interaction domain (the Six domain) immediately N-terminal to the homeodomain and a C-terminal domain that houses a transactivation function in Six1/2/4/5. In order to examine whether the differences in the DNA-binding properties of the Six2 and Six6 homeodomains extended to the full-length proteins, we generated and purified recombinant full-length Six2 and Six6 proteins (both with His_6 tags) and used them in EMSA with the Myo-MEF3 and GGATTA DNA probes. Both proteins bound the Myo-MEF3 probe (Figure 3, lanes 5–8). Interestingly, while Six2 was unable to bind to the GGATTA probe (Figure 3, lanes 1–2), Six6 bound the GGATTA sequence (Figure 3, lanes 3 and 4) with affinity comparable to that measured for the Six6−Myo-MEF3 interaction. This difference in DNA-binding specificity among Six2 and Six6 is interesting in light of the complete conservation of predicted DNA contacting residues in the homeodomains of these proteins. Furthermore, Six2(HD)Six6(C) is able to bind both Myo-MEF3 and GGATTA (Figure 2b,c), suggesting that the ability to bind GGATTA is conferred by the C-extension sequence.

Since homeodomain-containing proteins frequently bind DNA cooperatively with other DNA-binding partners, it is likely that the Six−DNA interaction is also regulated in terms of both affinity and specificity by a protein partner. There is genetic and biochemical evidence that Eyes Absent proteins form complexes with Six proteins and that they act synergistically *in vivo* (*27–30*). Thus it is possible that interaction with Eyes Absent could modulate Six−DNA-binding activity. To localize the domain of Eya3 involved in interaction with Six2 and Six6, we performed glutathione *S*-transferase (GST)-pull-down experiments with the bacterially expressed and purified recombinant proteins. Two forms of Eya3 were used: Eya3 (residues 179–510) has been previously reported to be the minimum construct necessary for Six2 interaction (*29*), and Eya3 (residues 238–510) is the conserved Eya domain (ED) that houses the phosphatase activity (*4*). Purified GST-tagged Eya3(238–510) and GST-Eya3(179–510) were independently mixed with either His_6-tagged Six2 or His_6-tagged-Six6. The mixtures were then loaded on glutathione (GT) beads and extensively washed. The final beads along with samples of the loaded mixture, the unbound supernatant, and the final wash were analyzed by SDS−PAGE (Figure 4). His_6-Six2 and His_6-Six6 were retained on GT−agarose beads when premixed with either GST-Eya3(238–510) (Figure 4a,b) or GST-Eya3(179–510) (Figure 4c). In control experiments neither His_6-Six2 (Figure 4a) nor His_6-Six6 (Figure 4b) bound GST. The data thus show that both Six2 and Six6 are able to bind Eya3 and that the ED domain of Eya3 is sufficient for the interaction.

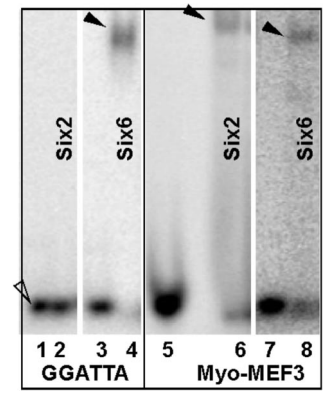

FIGURE 3: EMSA showing that Six2 binds Myo-MEF3, while Six6 binds both GGATTA and Myo-MEF3. Equivalent amounts of His_6-Six2 and His_6-Six6 were used in these experiments. Lanes: (1) free GGATTA probe, (2) GGATTA + His_6-Six2, (3) free GGATTA, (4) GGATTA + His_6-Six6, (5) free Myo-MEF3, (6) Myo-MEF3 + His_6-Six2, (7) free Myo-MEF3, and (8) Myo-MEF3 + His_6-Six6. The open arrow indicates the bands corresponding to the free DNA probes, and filled arrows indicate protein−DNA complexes.

To examine how the interaction of Eya3 with Six2 and Six6 affected Six−DNA interaction, we conducted EMSA with the Myo-MEF3 and GGATTA DNA probes. As noted in Figure 3, both Six2 and Six6 have low intrinsic affinity for DNA, and they showed distinctive preferences regarding DNA sequence. When the Six2 protein was premixed with Eya3(residues 238–510) and used in EMSA with the Myo-MEF3 probe, a supershift as well as an increase in affinity was observed (Figure 5a, lanes 4–6 relative to lanes 7–9). Eya3 (residues 179–510), the previously reported Six2 interaction domain of Eya3 (*28*), also similarly supershifted but had a more dramatic positive impact on the affinity of the Six2−Myo-MEF3 complex (Figure 5a, lanes 1–3). In quantitative EMSA a Six2−Eya3(179–510) complex bound

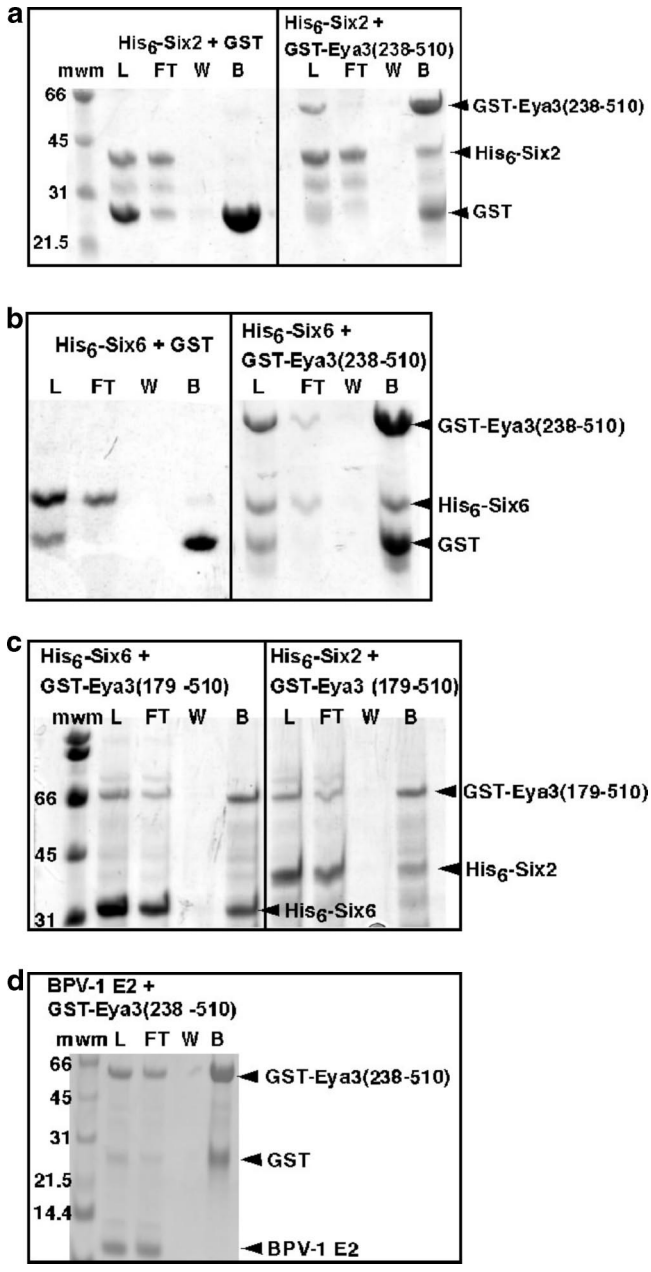

FIGURE 4: Both Six2 and Six6 can form complexes with Eya3(238–510) (a, b) and Eya3(179–510) (c). Key: L, sample loaded on GT−agarose beads; FT, flow-through after loading on GT−agarose; W, last wash of GT−agarose after it was loaded with sample; B, proteins retained on GT−agarose beads after washing; mwm, molecular mass markers (sizes in kDA shown to the left of the gels in panels a and c). (a) GST pulldown using His_6-Six2 and either GST (left panel) or GST-Eya3(238–510) (right panel). (b) GST pulldown using His_6-Six6 and either GST (left panel) or GST-Eya3(238–510) (right panel). (c) GST pulldown using GST-Eya3(179–510) and either His_6-Six6 (left panel) or His_6-Six2 (right panel). (d) Negative control GST pulldown using GST-Eya3(238–510) and BPV-1 E2(325−410) showing that these two proteins do not interact.

Myo-MEF3 with 11.6-fold greater affinity than Six2 by itself. Even in the presence of Eya3(179–510) or Eya3(238–510), Six2 was unable to bind the GGATTA probe. Neither Eya3(238–510) nor Eya3(179–510) was able to bind DNA on its own (data not shown). These results validate the observation that the ED domain of Eya3 is able to bind Six2, as well as suggesting an additional biochemical mechanism

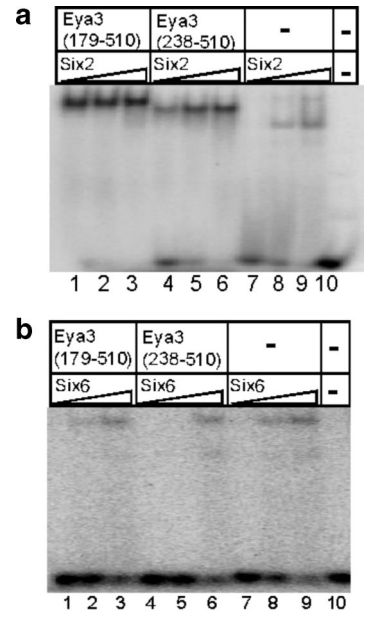

FIGURE 5: The DNA-binding affinity of Six2, but not Six6, is significantly increased in the presence of Eya3. (a) EMSA titrations of Six2 with the Myo-MEF3 probe (free probe, lane 10). Lanes: 1–3, increasing amounts of His_6-Six2 with a constant amount of Eya3(179–510); 4–6, increasing amounts of His_6-Six2 with a constant amount of Eya3(238–510); 7–9, increasing amounts of His_6-Six2. His_6-Six2 concentration ranges used in the three titrations are identical. (b) EMSA titrations of Six6 with the Myo-MEF3 probe (free probe, lane 10). Lanes: 1–3, increasing amounts of His_6-Six6 with a constant amount of Eya3(179–510); 4–6, increasing amounts of His_6-Six6 with a constant amount of Eya3(238–510); 7–9, increasing amounts of His_6-Six6. His_6-Six6 concentration ranges used in the three titrations are identical.

for the previously reported synergistic transactivation of reporter gene transcription in cotransfection experiments with Six and Eya proteins.

Interestingly, even though Six6 can form a complex with Eya3(238–510) and Eya3(179–510) as shown in Figure 4, we were able to detect neither a supershift nor an increase in Six6−DNA-binding affinity in the presence of Eya3-(238–510) or Eya3(179–510) (Figure 5b). Human Six6 and mouse Eya3 were used in these experiments; the amino acid sequences of human and mouse Six6 are 99.2% identical, differing only in two amino acids in the C-terminal region (Pro 201 and Ser 207 in mouse Six6 are respectively Ser and Ala in the human protein).

DISCUSSION

In this report we provide evidence that the Six proteins employ a noncanonical mode of homeodomain−DNA binding and that the Six2 and Six6 proteins differ in their mechanisms of DNA recognition, in their DNA-binding specificities (summarized in Figure 6), and in the effect of Eya3 interaction on Six−DNA binding. Together, these observations provide mechanistic insights into an unusual homeodomain−DNA interaction, as well as suggesting a molecular mechanism for the synergism observed between Six and Eyes Absent proteins in biological processes such as eye development and in biochemical activities such as transactivation.

In a previous report Hazbun et al. (*31*) speculated that residues upstream of the HD of sine oculis (the *Drosophila*

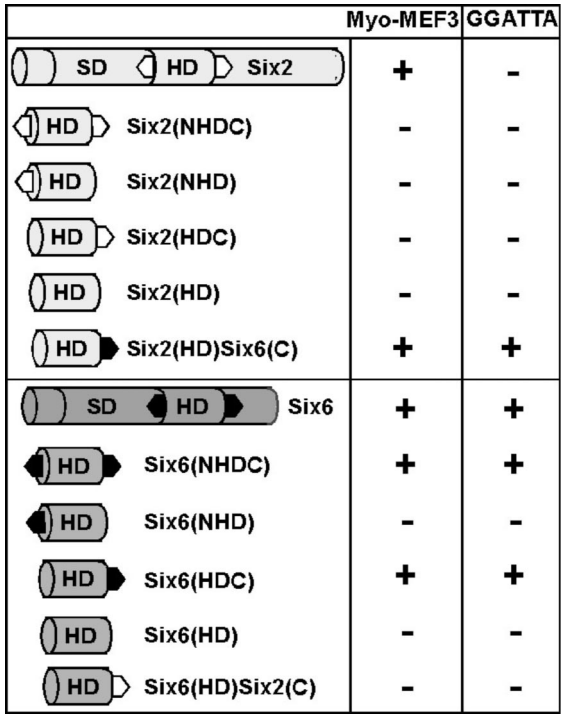

FIGURE 6: Summary of the DNA-binding properties of Six2 and Six6 constructs. A schematic of each construct of Six2 and Six6 is shown along with the results of DNA-binding analyses using either the Myo-MEF3 or the GGATTA probe. (+) indicates DNA binding, (−) indicates no DNA-binding, SD indicates Six domain, HD indicates homeodomain, and the filled and open arrowheads are used to indicate the N- and C-extensions of Six6 and Six2, respectively.

Six1/2 homologue) were similar to the N-terminal arm of the extradenticle HD and may play an analogous minor groove interaction role. However, we find that inclusion of these residues N-terminal to the HD (NHD in Figures 1a and 6) does not enable DNA binding for either Six2 or Six6. In contrast, an extended form of the HD including an additional 14 residues C-terminal to the HD (C-extension) of Six6 (HDC in Figures 1a and 6) is able to bind DNA. Within the context of this DNA-binding Six6(HDC) construct removal of the N-terminal arm region causes little change in DNA-binding affinity, further supporting the model that the N-terminal arm of the Six6 HD does not participate in DNA binding. The relative contribution of the HD N-terminal arm to HD−DNA-binding affinity/specificity may be protein specific. Deletion of the N-terminal arm in the homeodomain leucine zipper protein Hhab-4 dramatically reduced DNA-binding affinity nearly 70-fold (*24*), suggesting a vital DNA interaction role. In contrast, NMR studies of the Pitx2 HD−DNA complex suggest that while the N-terminal arm appears to contact the minor groove of DNA, it retains significant mobility (*32*). The N-terminal arm has also been assigned other roles, as in the case of the murine homeodomain Msx where extensive mutagenesis has shown that it is critical for stabilizing the tertiary structure of the HD (*26*).

Extended HD constructs are known to facilitate DNA binding in the case of other HD proteins. In the Prospero proteins the third helix of the homeodomain extends into the C-terminal prospero domain (*33*). The C-extension of the Prospero HD does not directly participate in DNA binding, rather the homeo and prospero domains together form a composite DNA-binding domain. The primary sequence of the Six proteins C-terminal to the HD is very low, and secondary structure prediction algorithms indicate that the region C-terminal to HDC in all Six proteins is largely unstructured. Thus it is unlikely that the HD- and C-terminal regions of the Six proteins form a composite DNA-binding domain. The PBX proteins represent another instance where the HD extends beyond the classically defined 60 residues (*34*). The PBX C-extension forms a fourth helix upon DNA binding that stabilizes the HD structure. The C-extension is highly conserved among PBX proteins, and it increases PBX−DNA and PBX−HOX binding affinities. As in PBX the C-extension of Six6 increases Six6 DNA-binding affinity. A likely possibility is that inclusion of the C-extension of the Six HD prevents premature truncation of helix 3, thus providing structural stability. The relatively basic residues in the C-extension of Six3/6 may favor DNA interaction more than the acidic amino acids found in Six1/2/4/5, and this increased electrostatic attraction may account for the more relaxed sequence specificity observed in the chimeric construct Six2(HD)Six6(C) relative to Six2 (Figure 6). Precisely how the C-extension of Six6 mediates DNA sequence recognition remains to be determined.

Differences in the biochemical activities of the Six proteins as described here complement several previous reports regarding differences in their biological functions (*35–40*). Recent studies have shown that the C-terminal regions of the SO and OPTIX proteins are not interchangeable; replacement of the SO C-terminal region with that of OPTIX prevents rescue of an eyeless phenotype in the *so*^{1} mutants (*41*). These authors noted two regions of homology in the C-terminal region of OPTIX/Six3/Six6 that are not also present in SO/Six1/Six2, and they speculated that these stretches of amino acids may also play a role in the distinctive protein-partner specificities of OPTIX versus SO. Interestingly, the C-extension that we have identified as playing a role in DNA sequence specificity coincides with one of these two conserved motifs, suggesting that differences in DNA target selection may also underlie the functional divergence among the Six family of proteins.

Homeodomain proteins tend to have relatively low intrinsic DNA sequence specificity and employ different strategies *in vivo* to maintain selectivity in target gene selection. Additional DNA-binding domains and cofactors are commonly used mechanisms to modulate DNA binding site specificity. Typically, these cofactors are also DNA-binding proteins. There is no interaction yet reported between the Six proteins and a sequence-specific DNA-binding protein. However, Six proteins have previously been shown to form functional complexes with the Eyes Absent proteins (*28–30*). Pairs of Six and Eyes Absent proteins when coexpressed synergistically activate transcription in reporter assays (*28*). In *Drosophila* Sine oculis and EYA act synergistically in ectopic eye development assays (*30*). These observations have led to the suggestion that synergism upon cotransfection of Six and Eya may be due to the unmasking of the transactivation domain (TAD) of the Six proteins upon Eya interaction and/or the formation of a composite activation domain composed of the C-terminal TAD of the Six proteins and the N-terminal TAD of the Eya proteins. Our data suggest that Eya3 can also act as a Six2 cofactor by significantly increasing its DNA-binding affinity, without

altering DNA sequence specificity. However, since Eyes Absent by itself is not a DNA-binding protein, Six2−Eyes Absent synergy is likely mediated by a novel mechanism. Several mechanisms could be envisioned: interaction with Eyes Absent could mediate a conformational change in Six2 favoring a DNA-binding competent form, Eyes Absent could structurally stabilize the DNA-binding domain of Six2, or Eyes Absent could favor a DNA-binding competent oligomerization state of Six2. The oligomerization of Six2 in the free and DNA-bound states and the precise mechanism by which Eyes Absent exerts its effects on Six2−DNA binding remain to be experimentally elucidated. Eya3 is coexpressed with Six2 in multiple tissues, underscoring the biological relevance of the biochemical cooperativity. It is curious that Eya3 does not potentiate Six6−DNA binding as it does for Six2, thus representing another point of divergence between Six2 and Six6 with regard to biochemical function. It remains to be determined whether other pairs of Eya and Six proteins also exhibit modulated DNA-binding properties.

Adding another dimension to the biochemical synergy between Six and Eya reported here is the observation that pairs of Six and Eya proteins are associated with cancers. In particular, *SIX1* and *EYA2* are overexpressed in-breast (*42, 43*) and ovarian cancer (*43, 44*), and their overexpression correlates with increased tumor size, metastasis, and poor prognoses. A more detailed understanding of the biochemical mechanisms by which these proteins carry out their cellular roles will be useful in understanding the effects of the reported elevated protein levels in cancers.

ACKNOWLEDGMENT

We thank Chandramohan Kattamuri for cloning some of the overexpression constructs used in these experiments, Heidi Poppendeck for purifying the GST-Eya3(238–510) protein, Michael Spencer for preliminary protein interaction experiments, Dr. Mark Rance for a gift of Pitx2(HD) for use as a control, and Dr. Michael Brenowitz for help in quantitative analysis of the EMSA.

REFERENCES

1. Kawakami, K., Sato, S., Ozaki, H., and Ikeda, K. (2000) Six family genes−structure and function as transcription factors and their roles in development. *BioEssays 22*, 616–626.
2. Wawersik, S., and Maas, R. L. (2000) Vertebrate eye development as modeled in *Drosophila*. *Hum. Mol. Genet. 9*, 917–925.
3. Rayapureddi, J. P., Kattamuri, C., Chan, F. H., and Hegde, R. S. (2005) Characterization of a plant, tyrosine-specific phosphatase of the aspartyl class. *Biochemistry 44*, 751–758.
4. Rayapureddi, J. P., Kattamuri, C., Steinmetz, B. D., Frankfort, B. J., Ostrin, E. J., Mardon, G., and Hegde, R. S. (2003) Eyes absent represents a class of protein tyrosine phosphatases. *Nature 426*, 295–298.
5. Li, X., Oghi, K. A., Zhang, J., Krones, A., Bush, K. T., Glass, C. K., Nigam, S. K., Aggarwal, A. K., Maas, R., Rose, D. W., and Rosenfeld, M. G. (2003) Eya protein phosphatase activity regulates Six1-Dach-Eya transcriptional effects in mammalian organogenesis. *Nature 426*, 247–254.
6. Tootle, T. L., Silver, S. J., Davies, E. L., Newman, V., Latek, R. R., Mills, I. A., Selengut, J. D., Parlikar, B. E., and Rebay, I. (2003) The transcription factor Eyes absent is a protein tyrosine phosphatase. *Nature 426*, 299–302.
7. Kim, S. S., Zhang, R., Braunstein, S. E., Joachimiak, A., Cvekl, A., and Hegde, R. S. (2002) Structure of the retinal determination protein dachshund reveals a DNA binding motif. *Structure (Cambridge) 10*, 787–795.
8. Silver, S. J., and Rebay, I. (2005) Signaling circuitries in development: insights from the retinal determination gene network. *Development 132*, 3–13.
9. Desplan, C. (1997) Eye development: governed by a dictator or a junta? *Cell 91*, 861–864.
10. Seo, H. C., Curtiss, J., Mlodzik, M., and Fjose, A. (1999) Six class homeobox genes in *Drosophila* belong to three distinct families and are involved in head development. *Mech. Dev. 83*, 127–139.
11. Ades, S. E., and Sauer, R. T. (1994) Differential DNA-binding specificity of the engrailed homeodomain: the role of residue 50. *Biochemistry 33*, 9187–9194.
12. Suzuki-Yagawa, Y., Kawakami, K., and Nagano, K. (1992) Housekeeping Na,K-ATPase alpha 1 subunit gene promoter is composed of multiple cis elements to which common and cell type-specific factors bind. *Mol. Cell. Biol. 12*, 4046–4055.
13. Kawakami, K., Ohto, H., Ikeda, K., and Roeder, R. G. (1996) Structure, function and expression of a murine homeobox protein AREC3, a homologue of *Drosophila* sine oculis gene product, and implication in development. *Nucleic Acids Res. 24*, 303–310.
14. Spitz, F., Demignon, J., Porteu, A., Kahn, A., Concordet, J. P., Daegelen, D., and Maire, P. (1998) Expression of myogenin during embryogenesis is controlled by Six/sine oculis homeoproteins through a conserved MEF3 binding site. *Proc. Natl. Acad. Sci. U.S.A. 95*, 14220–14225.
15. Zhu, C. C., Dyer, M. A., Uchikawa, M., Kondoh, H., Lagutin, O. V., and Oliver, G. (2002) Six3-mediated auto repression and eye development requires its interaction with members of the Groucho-related family of co-repressors. *Development 129*, 2835–2849.
16. Kenyon, K. L., Li, D. J., Clouser, C., Tran, S., and Pignoni, F. (2005) Fly SIX-type homeodomain proteins Sine oculis and Optix partner with different cofactors during eye development, Dev. Dyn.
17. Kenyon, K. L., Yang-Zhou, D., Cai, C. Q., Tran, S., Clouser, C., Decene, G., Ranade, S., and Pignoni, F. (2005) Partner specificity is essential for proper function of the SIX-type homeodomain proteins Sine oculis and Optix during fly eye development. *Dev. Biol. 286*, 158–168.
18. Nallamsetty, S., and Waugh, D. S. (2007) A generic protocol for the expression and purification of recombinant proteins in *Escherichia coli* using a combinatorial His6-maltose binding protein fusion tag. *Nat. Protoc. 2*, 383–391.
19. Senear, D. F., and Brenowitz, M. (1991) Determination of binding constants for cooperative site-specific protein-DNA interactions using the gel mobility-shift assay. *J. Biol. Chem. 266*, 13661–13671.
20. Senear, D. F., Dalma-Weiszhausz, D. D., and Brenowitz, M. (1993) Effects of anomalous migration and DNA to protein ratios on resolution of equilibrium constants from gel mobility-shift assays. *Electrophoresis 14*, 704–712.
21. Hegde, R. S., Grossman, S. R., Laimins, L. A., and Sigler, P. B. (1992) Crystal structure at 1.7 Å of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. *Nature 359*, 505–512.
22. Gehring, W. J., Affolter, M., and Burglin, T. (1994) Homeodomain proteins. *Annu. Rev. Biochem. 63*, 487–526.
23. Ades, S. E., and Sauer, R. T. (1995) Specificity of minor-groove and major-groove interactions in a homeodomain-DNA complex. *Biochemistry 34*, 14601–14608.
24. Palena, C. M., Tron, A. E., Bertoncini, C. W., Gonzalez, D. H., and Chan, R. L. (2001) Positively charged residues at the N-terminal arm of the homeodomain are required for efficient DNA binding by homeodomain-leucine zipper proteins. *J. Mol. Biol. 308*, 39–47.
25. Percival-Smith, A., Muller, M., Affolter, M., and Gehring, W. J. (1990) The interaction with DNA of wild-type and mutant fushi tarazu homeodomains. *EMBO J. 9*, 3967–3974.
26. Shang, Z., Isaac, V. E., Li, H., Patel, L., Catron, K. M., Curran, T., Montelione, G. T., and Abate, C. (1994) Design of a "minimAl" homeodomain: the N-terminal arm modulates DNA binding affinity and stabilizes homeodomain structure. *Proc. Natl. Acad. Sci. U.S.A. 91*, 8373–8377.
27. Heanue, T. A., Reshef, R., Davis, R. J., Mardon, G., Oliver, G., Tomarev, S., Lassar, A. B., and Tabin, C. J. (1999) Synergistic regulation of vertebrate muscle development by Dach2, Eya2, and Six1, homologs of genes required for *Drosophila* eye formation. *Genes Dev. 13*, 3231–3243.
28. Ikeda, K., Watanabe, Y., Ohto, H., and Kawakami, K. (2002) Molecular Interaction and Synergistic Activation of a Promoter

by Six, Eya, and Dach Proteins Mediated through CREB Binding Protein. *Mol. Cell. Biol. 22*, 6759–6766.
29. Ohto, H., Kamada, S., Tago, K., Tominaga, S. I., Ozaki, H., Sato, S., and Kawakami, K. (1999) Cooperation of six and eya in activation of their target genes through nuclear translocation of Eya. *Mol. Cell. Biol. 19*, 6815–6824.
30. Pignoni, F., Hu, B., Zavitz, K. H., Xiao, J., Garrity, P. A., and Zipursky, S. L. (1997) The eye-specification proteins So and Eya form a complex and regulate multiple steps in *Drosophila* eye development. *Cell 91*, 881–891.
31. Hazbun, T. R., Stahura, F. L., and Mossing, M. C. (1997) Site-specific recognition by an isolated DNA-binding domain of the sine oculis protein. *Biochemistry 36*, 3680–3686.
32. Chaney, B. A., Clark-Baldwin, K., Dave, V., Ma, J., and Rance, M. (2005) Solution structure of the K50 class homeodomain PITX2 bound to DNA and implications for mutations that cause Rieger syndrome. *Biochemistry 44*, 7497–7511.
33. Yousef, M. S., and Matthews, B. W. (2005) Structural basis of Prospero-DNA interaction: implications for transcription regulation in developing cells. *Structure 13*, 601–607.
34. Piper, D. E., Batchelor, A. H., Chang, C. P., Cleary, M. L., and Wolberger, C. (1999) Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation. *Cell 96*, 587–597.
35. Ando, Z., Sato, S., Ikeda, K., and Kawakami, K. (2005) Slc12a2 is a direct target of two closely related homeobox proteins, Six1 and Six4. *FEBS J. 272*, 3026–3041.
36. Grifone, R., Demignon, J., Houbron, C., Souil, E., Niro, C., Seller, M. J., Hamard, G., and Maire, P. (2005) Six1 and Six4 homeoproteins are required for Pax3 and Mrf expression during myogenesis in the mouse embryo. *Development 132*, 2235–2249.
37. Kobayashi, M., Toyama, R., Takeda, H., Dawid, I. B., and Kawakami, K. (1998) Overexpression of the forebrain-specific homeobox gene six3 induces rostral forebrain enlargement in zebrafish. *Development 125*, 2973–2982.
38. Laclef, C., Souil, E., Demignon, J., and Maire, P. (2003) Thymus, kidney and craniofacial abnormalities in Six 1 deficient mice. *Mech. Dev. 120*, 669–679.
39. Li, X., Perissi, V., Liu, F., Rose, D. W., and Rosenfeld, M. G. (2002) Tissue-specific regulation of retinal and pituitary precursor cell proliferation. *Science 297*, 1180–1183.
40. Zuber, M. E., Perron, M., Philpott, A., Bang, A., and Harris, W. A. (1999) Giant eyes in *Xenopus laevis* by overexpression of XOptx2. *Cell 98*, 341–352.
41. Weasner, B., Salzer, C., and Kumar, J. P. (2007) Sine oculis, a member of the SIX family of transcription factors, directs eye formation. *Dev. Biol. 303*, 756–771.
42. Ford, H. L., Kabingu, E. N., Bump, E. A., Mutter, G. L., and Pardee, A. B. (1998) Abrogation of the G2 cell cycle checkpoint associated with overexpression of HSIX1: a possible mechanism of breast carcinogenesis. *Proc. Natl. Acad. Sci. U.S.A. 95*, 12608–12613.
43. Zhang, L., Yang, N., Huang, J., Buckanovich, R. J., Liang, S., Barchetti, A., Vezzani, C., O'Brien-Jenkins, A., Wang, J., Ward, M. R., Courreges, M. C., Fracchioli, S., Medina, A., Katsaros, D., Weber, B. L., and Coukos, G. (2005) Transcriptional coactivator *Drosophila* eyes absent homologue 2 is up-regulated in epithelial ovarian cancer and promotes tumor growth. *Cancer Res. 65*, 925–932.
44. Behbakht, K., Qamar, L., Aldridge, C. S., Coletta, R. D., Davidson, S. A., Thorburn, A., and Ford, H. L. (2007) Six1 overexpression in ovarian carcinoma causes resistance to TRAIL-mediated apoptosis and is associated with poor survival. *Cancer Res. 67*, 3036–3042.

BI702186S