The Paired Domain of Pax3 Contains a Putative Homeodomain Interaction Pocket Defined by Cysteine Scanning Mutagenesis

Sergio Apuzzo and Philippe Gros*

Department of Biochemistry, McGill University, Montreal, Quebec, Canada H3G 1Y6

Received January 3, 2006; Revised Manuscript Received March 25, 2006

ABSTRACT: Pax3 is a transcription factor that plays an important regulatory role during neurogenesis, myogenesis, and formation of neural crest cell derived structures. Pax3 has two DNA binding domains, a paired domain (PD) and paired-type homeodomain (HD) that show complete interdependence for DNA binding, with mutations in one domain impairing DNA binding by the other domain. Cooperative interactions between the PD and HD of Pax3 suggest that the two domains may physically interact for DNA binding. Site-specific modification with thiol reagents in single cysteine Pax3 mutants was used to determine which segment of the PD may interact with the HD. Twenty-four single cysteine mutants were independently introduced in the second α-helix (α2, positions 59−80) and in the β-hairpin structure (positions 40−41) at the amino terminal portion of the PD. These mutants were tested for their ability to bind to PD (P6CON, P3OPT) and HD-specific DNA targets (P2), and the effect of treatment with *N*-ethylmaleimide on these binding properties was established. In the PD, single cysteine mutants CL/Q40C, CL/I59C, CL/V60C, CL/P69C, CL/S70C, CL/I72C, CL/S73C, CL/L76C, CL/V78C, and CL/S79C displayed NEM sensitive DNA binding toward both PD and HD targets. Three PD mutants (CL/L41C, CL/A63C, and CL/H64C) showed unusual behavior, with DNA binding to PD targets being NEM insensitive while DNA binding by the HD was abrogated by NEM treatment. Three-dimensional modeling of the NEM sensitive PD cysteine mutants reveal that they are not randomly distributed, but rather that they cluster in a hydrophobic pocket. We propose that this hydrophobic pocket may serve as a docking site for the HD during DNA binding by the intact protein.

Pax3 is a member of the Pax family of nine transcription factors that play critical roles in different aspects of mammalian development (*1*). Pax3 expression occurs in the developing somites, neural tube, and neural crest cell derived structures. Heterozygosity for a loss-of-function mutation at *Pax3* in *splotch* (*Sp*) mice causes pigmentary defects (white belly spot) while homozygotes show defects in neurogenesis (spina bifida), myogenesis (absence of limb muscle), and formation of neural crest cell derivatives including melanocytes (*2*−*7*). In humans, mutations in *PAX-3* cause Waardenburg syndrome, a pathology characterized by pigmentary disturbances, craniofacial abnormalities, and sensory deafness (*8, 9*).

Pax proteins are defined by a unique DNA binding domain, the paired domain (PD) first identified in the *Drosophila* paired (Prd) protein (*10*). Several Pax proteins, including Pax3, contain a second DNA binding domain known as the paired-type homeodomain (HD). Pax3 also shows a conserved octapeptide and a proline-serine-threonine rich C-terminal trans-activation domain, both of which are involved in protein−protein interactions (*1, 11*). The crystal structure of DNA-bound Pax6 has been solved and reveals that the PD is a bipartite structure, with each domain consisting of three α-helices with the last two forming a typical helix−turn−helix (HTH) motif (PAI and RED) (*12, 13*); the C-terminal helix of each HTH makes critical contacts in the major groove of DNA (*12, 13*). Located upstream of the N-terminal subdomain (PAI), a β-turn motif makes contacts in the minor groove of DNA (*12, 13*). The crystal structure of the DNA-bound paired-type HD of the Prd protein (*14*) reveals a conserved three helical fold with the two most C-terminal helices forming a HTH motif with α3 making DNA specific major groove contacts (*14*−*16*). The HD of Pax proteins has the unique ability to homodimerize on palindromic sequences of the type TAAT($N_{2/3}$)ATTA (*15, 17*). The identity of the amino acid residue at position 50 in α3 of HD determines both DNA binding specificity and dimerization potential (*15, 17*).

Although the PD and HD can bind to cognate DNA sequences when expressed individually, genetic and biochemical data indicate that the two domains are functionally interdependent in intact Pax3 (*9, 18*−*20*). For instance, the *Splotch-delayed* (*Sp*^{d}) allele of Pax3 bears a mutation (G42R) in the β-hairpin of the PD that not only abrogates DNA binding by the PD but also impairs DNA binding by the HD (*20*). Conversely, studies of the HD mutation R53G found in a Waardenburg patient show that R53G not only impairs DNA binding by the HD but also uncouples DNA binding by the PD (*22*). Additional studies in chimeric PAX3 proteins have shown that the PD can modulate both DNA binding specificity and dimerization potential of heterologous HDs (*21*).

* To whom all correspondence should be addressed. Mailing address: Department of Biochemistry, McGill University, 3655 Drummond, Room 907, Montreal, QC, Canada H3G-1Y6. Tel: 514-398-7291. Fax: 514-398-2603. E-mail: philippe.gros@mcgill.ca.

10.1021/bi060009z CCC: $33.50
Published on Web 05/17/2006

Previously, we have created a Pax3 mutant (Cys-less; CL-Pax3) in which all endogenous cysteine residues had been removed. CL-Pax3 retains both PD and HD DNA binding activity, and we have used this mutant backbone for the reintroduction of single cysteine residues at strategic locations of the PD and HD (*23*). We have monitored the effect of site-specific modification of such single cysteine mutants by *N*-ethyl maleimide on the DNA binding properties of either domain (*23*). Introduction of a cysteine at position 82 (α3 helix of PAI domain; C82) did not affect DNA binding by Pax3, but site-specific modification of C82 impaired DNA binding by both PD and HD (*23*). Conversely, NEM modification of the single cysteine HD mutant V263C abrogated both the HD and PD DNA binding activity, while modification of independent mutations at nearby positions (V265C and S268C) had no effect on PD activity (*23*). These studies further demonstrated the reciprocal regulation of the PD and HD in intact Pax3.

Recently, we modified the Pax3 protein by insertion of single or multiple Factor Xa protease cleavage sites at strategic positions in or near the PD and HD. Protease sensitivity studies in Pax3 mutants modified at position 100 in the PD linker separating the PAI and RED motif, as well as at position 216 immediately upstream of the HD, revealed that DNA binding by Pax3 resulted in a more compact and less solvent exposed protein (*24*). In addition, it was observed that DNA binding by the PD not only caused a structural change in the PD but also caused a conformational change in the HD; similarly, DNA binding by the HD also caused a conformational change in the PD, providing a structural basis for the functional interdependence of the two DNA binding domains of Pax3 (*24*).

The major objective of this study was to try to identify individual residues in the Pax3 PD that may come in close proximity to the HD during DNA binding by the intact protein. We reasoned that such residues may play a key role in the regulation of DNA binding that the PD of Pax3 exerts on its own HD or on heterologous HD fused to it (*21*). We focused our search on two regions of the amino terminal portion of the PD. We have previously shown that the G42R mutation in the PD of the *Sp*^{d} allele of Pax3 impairs both PD and HD DNA binding but that removal of helix 2 of the PD (PAI) restores DNA binding by the HD (*20, 21*). In addition, DNA binding studies with the Prd *Drosophila* protein using hybrid DNA targets optimized for both PD and HD (PH0 probe) show that the PD and HD bind their respective sites only when they are everted and juxtaposed in the same molecule (e.g. PH0 probe) (*28*). Modeling studies using the available structures from the PD of PAX6 and the HD of Prd suggests that when Pax proteins, with a PD and a HD, are bound to the PH0 DNA sequence, the protein segment N-terminal to the HD comes in close proximity to helix 2 of the PAI subdomain of the PD (*28*). Together, these studies suggest that helix 2 may be a candidate for interaction with the HD. Finally, structural studies of the ternary complex formed between the PD of Pax5 and the ETS domains of Ets-1 (*29*) reveal that the β-turn within the β-hairpin structure at the amino terminus of the PD is involved in protein:protein interactions. More specifically, residues in helix 2 of the PAI subdomain as well as residues equivalent to Pax3 Q40 and L41 in the β-hairpin make contact with the ETS domain (*29*). Therefore, we hypothesized that residues in the β-hairpin structure and in helix 2 of the PD may play a role in intramolecular interactions between the PD and the HD of Pax3.

We have used site-specific modification of single cysteine mutants to identify the residues in the amino terminal portion of the PD that may be involved in functional or physical interaction with the HD. Clustering of NEM sensitive single cysteine mutants impairing DNA binding by the HD on the three-dimensional structure of the PD identifies a hydrophobic pocket which may serve as a docking site for the HD during DNA binding by the intact protein.

MATERIALS AND METHODS

Mutagenesis. The construction of the pMT2 expression plasmid containing the entire protein-encoding region of wild-type Pax3 cDNA has been previously described (*20*). This pMT2/Pax3 construct encodes for all 479 amino acids of the murine Q+ isoform of Pax3 (*20*). This cDNA was modified by the in-frame addition of antigenic epitope derived from the human c-Myc protein (c-Myc epitope, EQKLISEEDL) at the N-terminus as well as a poly-histidine tail (His_6), an HA hemagglutinin epitope (YPYDVPDYAS), and a termination codon at the C terminus of the protein. This was accomplished by PCR-mediated mutagenesis with mutagenic primers: P3-Myc (5′-CTCGAATTCATGGAG-CAGAAGTTAATCAGCGAAGAGGATCTCACCA-CGCTGGCCGGCGCTGTGCCCAGGATG-3′) and P3-HA (5′-TTTAGCGGATCCGAATTCTTAGTGATGGTGGT-GATGGTGTCCCGCGGCGTAATCTGGAACGTCA-TATGGATATCCGAACGTCCAAGGCTTACTTTG-3′). Both primers were engineered with *Eco*RI restriction sites at their ends. The resulting 1.5-kb PCR product was digested with *Eco*RI and ligated into the corresponding site of mammalian expression vector pMT2, and the resulting construct was designated pMT2/Myc-Pax3-HA WT (wild type). The seven endogenous cysteines of Pax3 were sequentially mutated to either serine or glycine in order to generate the cysteine-less construct, pMT2/Myc-Pax3-HA CL, as previously described (*23*). The CL construct was then digested at the unique *Cla*I site at position 189 of the Pax3 coding region, and two protease Xa cleavage sites were introduced in-frame with the rest of the Pax3 encoding region using double-stranded oligonucleotides with *Cla*I compatible cohesive ends, as described (*24*). The oligonucleotide used to introduce 2 protease Xa sites, (5′)-CGATATCGAAGGTAGAATA-GACGGCCGAAT-(3′), permitted the generation of the pMT2/Myc-Pax3-HA *2Xa* CL plasmid.

Twenty-four single cysteine mutants were created in the PD, at amino acid positions 40 and 41 in the β-hairpin structure, and at positions 59−80, which includes the entire length of the α2 helix of the PAI subdomain and flanking regions. All mutants were created by PCR-mediated mutagenesis using the mutagenic oligonucleotides listed in Table 1 along with pMT2/Myc-Pax3-HA *2Xa* CL plasmid as a template. The generation of all single cysteine mutants also required the use of the flanking primers: (5′)-ATGACCA-CGCTGGCCGGCGCTGTG-(3′) and (5′)-AGTGAGAGGG-GAGAGAGCATAGTC-(3′). In all cases, a PCR product of 1247bp was produced and digested with *Sma*I. The 330bp *Sma*I fragment was swapped into the endogenous *Pax3 Sma*I sites (pst 342−672) of the pMT2/Myc-Pax3-HA *2Xa* CL

Table 1: Oligonucleotides Used for Pax3 Mutagenesis[a]

Substitution	Mutagenic primer (5'-3')
PD SCM	
Q40C	GGCCGAGTCAAC**TGT**CTCGGAGGAGTA
L41C	CGAGTCAACCAG**TG**CGGAGGAGTATTT
I59C	ATCCGCCACAAG**TGT**GTGGAGATGGCC
V60C	CGCCACAAGATA**TGT**GAGATGGCCCAC
E61C	CACAAGATAGTG**TGT**ATGGCCCACCAT
M62C	AAGATAGTGGAG**TGT**GCCCACCATGGC
A63C	ATAGTGGAGATG**TG**CCACCATGGCATT
H64C	GTGGAGATGGCC**TG**CCATGGCATTCGG
H65C	GAGATGGCCCAC**TG**TGGCATTCGGCCG
G66C	ATGGCCCACCAT**T**GCATTCGGCGGAGC
I67C	GCCCACCATGGC**TG**TCGGCCGAGCGTC
R68C	CACCATGGCATT**TG**TCCGAGCGTCATT
P69C	CATGGCATTCGG**TGT**AGCGTCATTTCT
S70C	GGCATTCGGCCG**T**GCGTCATTTCTCGC
V71C	ATTCGGCCGAGC**TG**CATTTCTCGCCAG
I72C	CGGCCGAGCGTC**TG**TTCTCGCCAGCTT
S73C	CCGAGCGTCATTT**G**TCGCCAGCTTCGC
R74C	AGCGTCATTTCTT**G**CCAGCTTCGCGTG
Q75C	GTCATTTCTCGC**TGT**CTTCGCGTGTCC
L76C	ATTTCTCGCCAG**TG**TCGCGTGTCCCAT
R77C	TCTCGCCAGCTTT**G**CGTGTCCCATGGA
V78C	CGCCAGCTTCGC**TGT**TCCCATGGATCC
S79C	CAGCTTCGCGTGT**G**CCATGGATCCGTC
H80C	CTTCGCGTGTCC**TG**TGGATCCGTCTCT

[a] Nucleotide substitutions leading to amino acid changes are indicated in bold.

plasmid. Each mutation was verified by nucleotide sequencing, and the accessibility of restriction sites used for cloning was verified by endonuclease fragmentation.

Expression and Detection of Pax3 Mutants. The expression plasmids were used to transiently transfect COS-7 monkey cells. One million cells were plated in Dulbecco's modified Eagle medium containing 10% fetal bovine serum and were transfected by the calcium phosphate coprecipitation method using 15 μg of plasmid DNA doubly purified by ultracentrifugation on cesium chloride density gradients. Calcium-DNA precipitates were placed onto the cells for 5 h and then treated with HBS (0.14 M NaCl, 5 mM KCl, 0.75 mM Na_2HPO_4, 6 mM dextrose, 25 mM HEPES, pH 7.05) containing 15% glycerol for 1 min. The cells were then washed and placed in complete Dulbecco's modified Eagle medium. Whole cell extracts were prepared 24 h following glycerol shock by sonication in a buffer containing 20 mM HEPES (pH 7.6), 0.15 M NaCl, 0.5 mM DTT, 0.2 mM EDTA, 0.2 mM EGTA, and a mixture of protease inhibitors: aprotinin, pepstatin, and leupeptin used at 1 mg/mL and phenylmethylsulfonyl fluoride used at 1 mM. These extracts were stored frozen at −70 °C until use. To assess Pax3 mutant protein expression and stability, aliquots of whole cell extracts were analyzed by electrophoresis on acrylamide-containing SDS gels (SDS−PAGE), followed by electrotransfer onto nitrocellulose membranes and immunoblotting. Immunodetection was performed with mouse monoclonal anti-HA antibody (BabCO, Berkeley, CA) at a dilution of 1:1000 and visualized by enhanced chemiluminescence using a sheep anti-mouse horseradish peroxidase conjugated secondary antibody (Amersham Biosciences). Following anti-HA probing the membranes were submerged in stripping buffer (100 mM 2-mercaptoethanool, 2% SDS, 62.5 mM Tris-HCl, pH 6.7) and incubated at 50 °C for 30 min. The membranes were then washed with TBST buffer (10 mM Tris-HCl, pH 8, 150 mM NaCl, 0.1% Tween 20) at room temperature. Following blocking, the membranes were probed with mouse monoclonal anti-Myc antibody (BabCO) at a dilution of 1:1000 and visualized by enhanced chemiluminescence using a sheep anti-mouse horseradish peroxidase conjugated secondary antibody (Amersham Biosciences).

Electrophoretic Mobility Shift Assays. Electrophoretic mobility shift assays (EMSA) were performed as previously described (*20*). Each protein:DNA binding reaction was carried out using approximately 8 μg of total cell extracts from transiently transfected COS-7 monkey cells and 10 fmol (0.06 μCi) of radioactively labeled double stranded oligonucleotides containing either PD or HD recognition sites or both. The final concentration of labeled oligonucleotide in the binding reaction is 0.0005 μM. Whole cell extracts were incubated with ^{32}P-labeled PD or composite PD and HD specific probes in a volume of 20 μL containing 10 mM Tris-HCl (pH 7.5), 50 mM KCl, 1 mM DTT, 2 mM spermidine, 2 mg/mL BSA, and 10% glycerol. To reduce nonspecific binding, 1 μg of poly(dI-dC)poly(dI-dC) was included in binding studies with PD-specific probes and composite probes with both PD and HD sites, while 2 μg of heat-inactivated salmon sperm DNA was added to binding reactions involving HD specific probes. Following a 30 min incubation at room temperature, samples were electrophoresed at 12 V/cm in 6% acrylamide:bis-acrylamide (29:1) gels containing 0.25 or 0.5X TBE (1X TBE is 0.18 M Tris-HCl, 0.18 M boric acid, 4 mM EDTA, pH 8.3). Gels were dried under vacuum and exposed to Kodak BMS film with an intensifying screen. PD-specific sequences P6CON (5′)-TGGAATTCAGGAAAAATTTTCACGCTTGAGTT-CACAGCTCGAGTA-(3′) (*25*) and P3OPT (5′)-TGGTG-GTCACGCCTCATTGAATATTA-(3′) (*26, 27*), HD-specific sequence P2 (5′)-GATCCTGAGTCTAATTGATTACTG-TACAGG-(3′) (*15*), and the composite PD and HD specific sequence PH0 (5′)-GATTTCTTCCAATTAGTCACGCT-TGAGTG-(3′) (*28*) were synthesized as complementary oligonucleotide pairs and were designed in order to have recessed 3′ ends for end labeling with [α-^{32}P] dATP (3000 Ci/mmol; NEN) using the Klenow fragment of DNA polymerase. The thiol-specific reagent *N*-ethylmaleimide (NEM) from Pierce was prepared as a 32 mM stock in water. NEM was added as a 0.5 μL aliquot to a 4 μL volume of whole cell extract (final concentration of NEM was 3.55 mM), followed by a 30 min incubation at room temperature prior to the addition of the [^{32}P]-labeled probe and EMSA.

RESULTS

To study proximity relationships between the PD and the HD of Pax3, we used cysteine scanning mutagenesis to independently mutate Q40 and L41 (β-hairpin), as well as each position of helix 2 and flanking positions within the PAI subdomain of the PD (Pax3 positions 59−80) (see Figure 1). Mutants were constructed by PCR-based mutagenesis, expressed in COS-7 monkey cells, and tested for their DNA binding properties for HD or PD targets with or without prior site-specific modification with NEM.

Protein Expression and DNA Binding Properties of Pax3 Single Cysteine Mutants. To facilitate detection of the wild-

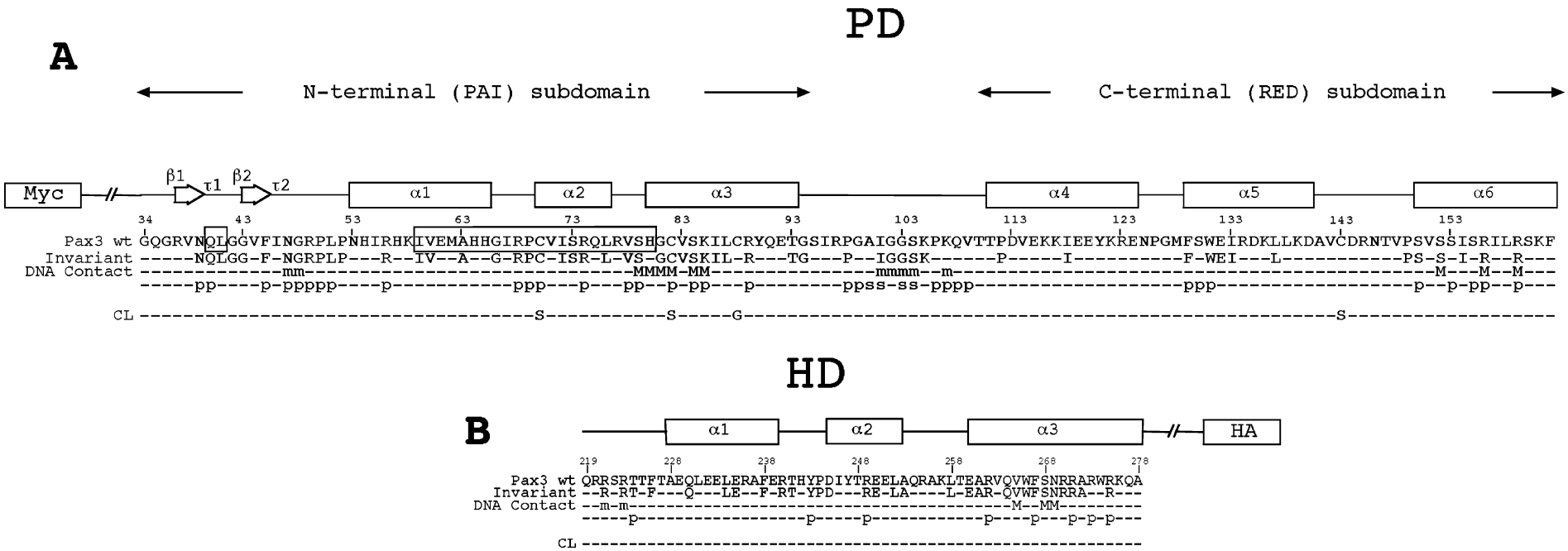

FIGURE 1: Site-directed mutagenesis of the paired domain of Pax3. (A) The amino acid sequence for positions 34−162 of Pax-3 is shown, including a schematic representation of the N-terminal (PAI) and C-terminal (RED) subdomains of the PD, together with structural features based on the three-dimensional structure of the PD of the Pax6 protein (β, β-strand; τ, β-turn; α, α-helix) (*13*). Invariant residues among all known PDs are identified below the schematic representation. The type of predicted DNA contacts made by these residues (p, phosphate; m, minor groove; M, major groove) is shown. The positions and nature of the mutations introduced in Pax-3 to create the Cys-less (CL) mutant are indicated. Boxed residues show areas targeted when generating single cysteine mutants. (B) Schematic representation of the Pax3 homeodomain, including the presence and position of predicted structural features, invariant amino acid residues, number and types of DNA contacts (as for panel A). The position of the c-Myc and hemagglutinin (HA) epitope tags inserted in-frame at the amino and carboxyl termini of Pax3, respectively, is shown.

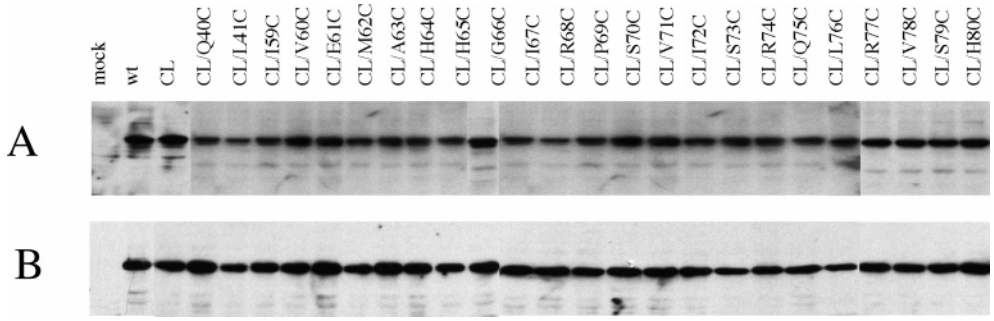

FIGURE 2: Expression of Pax3 mutants in COS-7 monkey cells. COS-7 cells were transiently transfected with wild-type (wt) *Pax3*, or with a *Pax3* mutant devoid of cysteine residues (Cys-less; CL) or with a number of independent *Pax3* mutants. The "mock" labeled lane refers to whole cell extracts from untransfected COS7 cells. *Pax3* cDNAs were cloned into pMT2 expression plasmid, and total cell extracts from transiently transfected COS-7 monkey cells were separated by SDS−PAGE (10% acrylamide) and transferred to nitrocellulose membranes. Immunodetection of Pax3 was carried out with mouse anti-HA (A) and anti-c-Myc (B) monoclonal antibodies and a horseradish peroxidase-conjugated secondary antibody.

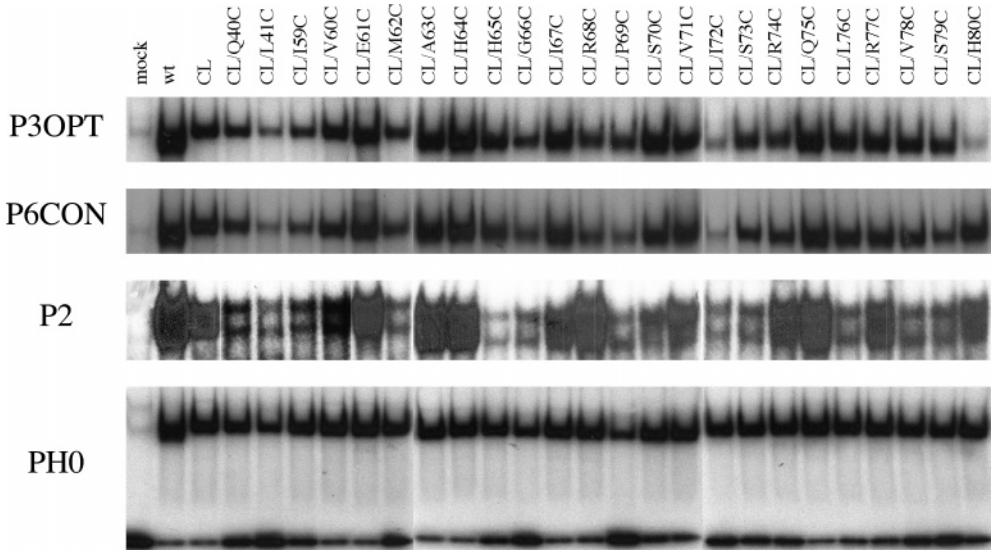

FIGURE 3: DNA binding properties of Pax3 mutants analyzed by electrophoretic mobility shift assays. Whole cell extracts from control nontransfected cells (mock) and from cells expressing individual *Pax3* mutants were used in electrophoretic mobility shift assays to evaluate the DNA binding properties of Pax3 single cysteine mutants against paired domain (P6CON and P3OPT) and homeodomain binding sites (P2), and against a probe with both paired domain and homeodomain binding sites (PH0). Protein−DNA complexes were formed using total COS-7 cell extracts and were resolved on 6% acrylamide nondenaturing gels, as described under Materials and Methods. The free probe is shown for the EMSA performed with the PH0 probe only.

type (WT), Cys-less (CL), and all single cysteine mutants (SCM), c-Myc and hemagglutinin (HA) epitope tags were inserted in-frame at the N-terminus and C-terminus of all constructs, respectively. All constructs were made using the pMT2 expression vector followed by transient transfection into COS-7 monkey cells. Immunoblotting of whole cell extracts with either anti-HA or anti-c-Myc monoclonal antibodies indicates comparable levels of expression for the WT, CL, and of all the SCMs (Figure 2). This suggests that none of the mutations introduced affect protein stability in COS-7 cells. The effect of single cysteine insertions in the β-hairpin and in helix 2 on the DNA binding properties of the PD and HD was examined in each mutant by electrophoretic mobility shift assays (EMSAs). The DNA binding of the PD SCMs was examined using oligonucleotide probes P3OPT and P6CON previously shown to reveal binding determinants present in both the amino (PAI) and carboxyl (RED) subdomains of the PD. Most of the SCMs retained PD DNA binding activity similar to the CL protein used as control, with the possible exception of CL/L41C and CL/I72C which showed reduced binding to P3OPT (Figure 3). All SCMs showed similar binding properties for either PD probe, P3OPT or P6CON, with the notable exception of CL/H80C that shows little P3OPT binding activity yet retains wild-type P6CON binding activity. The molecular basis for the unique differential binding of CL/H80C to P3OPT and P6CON is addressed in the Discussion. The effect of cysteine substitutions on DNA binding properties of the HD was evaluated using a target sequence (P2), which contains the sequence TAAT(N_2)-ATTA previously shown to support cooperative dimerization of Pax3 (*15*, *17*). Most SCM retained near wild-type binding activity toward P2 with the exception of CL/Q40C, CL/L41C, CL/M62C, CL/H65C, CL/G66C, CL/P69C, and CL/I72C that showed decreased binding. These results indicate that some of the mutations in the PD affect DNA binding by the HD, in agreement with previously published results (*20*, *22*). Finally, all mutants bound the PD−HD composite site present in the PH0 probe, suggesting that the decreased binding to either PD or HD targets seen for some of the mutants did not impair their ability to bind to the composite site.

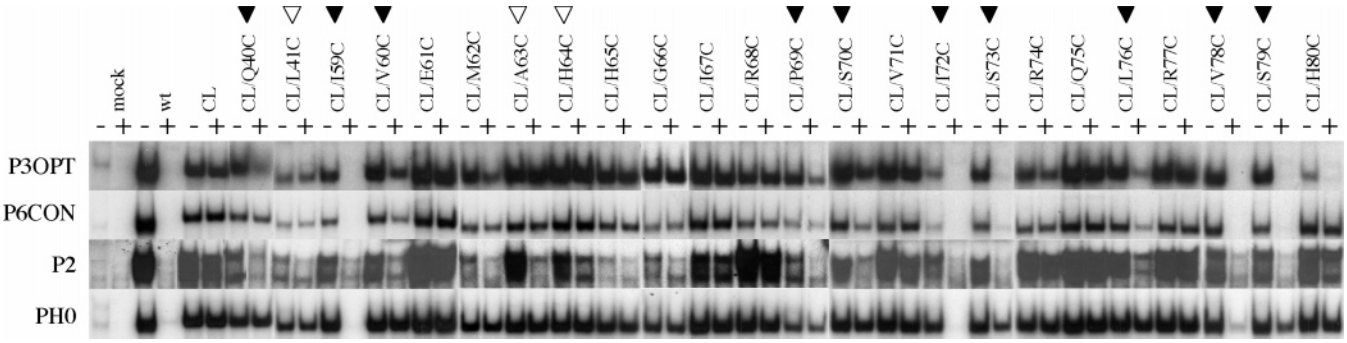

FIGURE 4: Effect of site-specific modification on DNA binding properties of Pax3 mutants. Effects of *N*-ethyl maleimide (NEM) treatment on the paired domain and homeodomain DNA binding properties of wild-type (WT), Cys-less (CL), and single cysteine mutants in the PD. Whole cell extracts from COS7 monkey cells (mock) or from cells expressing different Pax3 proteins were incubated with 0 (−) or 3.55 mM (+) NEM prior to electrophoretic mobility shift assays. The DNA binding properties of the PD were evaluated with target sites P3OPT and P6CON, while the HD was tested using the P2 probe. The ability to bind the composite sequence PH0 was also assessed. Filled arrowheads indicate SCMs with PD and HD NEM sensitive DNA binding domains, and empty arrowheads indicate SCMs with only HD NEM sensitive DNA binding.

Effect of Thiol Specific Reagents on DNA Binding by Single Cysteine Pax3 Mutants. The thiol-reactive compound *N*-ethyl maleimide (NEM) can form covalent adducts with Cys residues and was used for site-specific modification of SCMs. The effect of NEM on DNA binding by the PD and HD of WT, CL, and SCMs was tested in EMSA. Whole cell extracts from COS-7 cells expressing individual mutants were preincubated with 3.55 mM *N*-ethyl maleimide (NEM), and the ability of each mutant to bind via their PD (with P3OPT and P6CON) or HD (with P2) or both simultaneously (with PH0) was monitored by EMSA. The ability of WT Pax3 to bind to either PD or HD probes or the composite PH0 site is completely abrogated by prior incubation with NEM. On the other hand, the CL-Pax3 mutant is insensitive to the effect of NEM, and retains DNA binding to all probes (Figure 4).

NEM-sensitivity studies showed three types of responses in the SCMs. The first was that NEM had no effect for any of the probes analyzed, similar to that seen for the CL control. This group included SCM at positions 61, 62, 65, 66, 67, 68, 71, 74, 75, and 77. The absence of NEM effect in this group can reflect either inaccessibility of the corresponding Cys to NEM or that modification at that site is without consequence on the PD or HD ability to bind the target sequences. The second group was characterized by a loss or reduction of DNA binding by both PD and HD and included SCM at positions 40, 59, 60, 69, 70, 72, 73, 76, 78, and 79 (Figure 4). The third and most intriguing group corresponded to SCM in which NEM had no apparent effect on DNA binding by the PD, but caused a modest (position 41) or severe loss (positions 63 and 64) of DNA binding by the HD. Of note, NEM did not affect the ability of SCMs to bind the composite PH0 site, except for SCMs CL/P69C, CL/S73C, CL/V78C, and CL/S79C, which show partial NEM sensitivity, and CL/I59C and CL/I72C, which display complete NEM sensitivity. Finally, CL/H80C showed a unique behavior in that NEM impaired DNA binding by this mutant to one PD probe (P3OPT) but not the other (P6CON), while having no effect on HD DNA binding (Figure 4).

The positions of PD residues (PAI subdomain) that show NEM sensitivity for DNA binding are shown (Figure 5) in a 3-dimensional DNA-bound PD model which is based on the high-resolution crystal structure of PAX6 (*13*). Residues colored red identify positions at which NEM modification of a Cys reduces DNA binding by both the PD and HD, while those colored yellow show positions at which NEM modification of a cysteine impairs DNA binding only by the HD. This analysis shows that NEM sensitive SCMs are not randomly distributed but that they appear to cluster together in the 3-D structure of the DNA-bound PD. The NEM sensitive regions of the PD include portions of β-turn (pst: 40, 41) within the β-hairpin and segments of α1 (pst: 59, 60, 63, 64) and α2 (pst: 69, 70, 72, 73, 76) as well the α2−α3 loop (pst: 78, 79). These positions appear to be solvent exposed and map to the opposite plane of the DNA binding determinants of helix 3 (Figure 5). It is important to note that 8 of 13 NEM sensitive SCMs are located at positions that hold a hydrophobic side chain (L41, I59, V60, A63, P69, I72, L76, and V78), suggesting that the subdomain identified forms a three-dimensional hydrophobic pocket where the HD may dock during DNA binding.

DISCUSSION

Examining NEM sensitive PD SCMs in the 3-D model of the DNA-bound PAI subdomain structure (Figure 5) indicates that these informative SCMs cluster together. This clustering of portions of the β-turn of the β-pin and portions of α1, α2, and the α2−α3 loop suggests that these regions, that are not adjacent to one another in the primary sequence, may come into close proximity in the tertiary structure of the PD to form a HD binding pocket. The docking of HD segments to this PD hydrophobic pocket may provide the physical contact underlying the functional interdependence of the PD and HD for DNA binding noted in many mutant Pax3 variants (*9, 18*−*20, 22*).

We have chosen to use site-specific modification of single cysteine mutants (*23, 30, 31*) to probe functional interactions between the two DNA binding domains of Pax-3. Although this technique has been extensively used to decipher structure:function relationships in a number of soluble and membrane proteins (*23, 30, 31*), one of the limitations is that positions that do not show NEM sensitivity are not informative in the analysis. Indeed, for such single cysteine insertions one cannot distinguish inaccessibility of the mutant to the alkylator from successful modification being without consequence for protein function. Although we did not take into account mutants that lacked NEM sensitivity to DNA binding, we believe that most of the

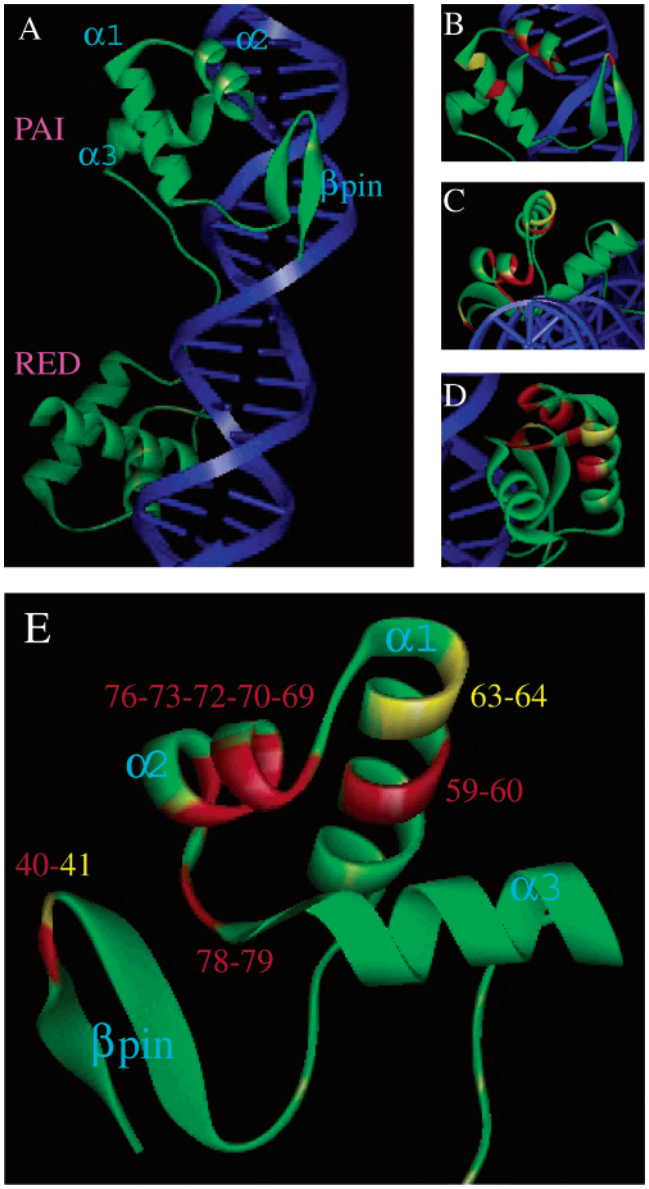

FIGURE 5: Structure of the DNA-bound Paired domain. The paired domain (PD) is shown as a green ribbon drawn through the α carbon backbone. The DNA strands are shown as gray ribbons through the sugar−phosphate backbone, and bases are shown as protrusions from the ribbons. The PD positions targeted for cysteine substitution and at which NEM modification causes a decrease in PD and HD DNA binding activities are indicated in red, while those leading to a reduction in HD DNA binding activity only are indicated in yellow. Panel A shows the DNA-bound PD with the PAI and RED subdomains indicated. Also labeled are the α-helices of the PAI subdomain only along with the *β*-hairpin structure. Panels B, C, and D show the PAI subdomain from three different viewpoints. Panel E shows the DNA-facing side of the PAI subdomain with the DNA strands removed. The three helices and the *β*-hairpin position numbers of residues that are NEM sensitive when substituted with cysteine are shown.

positions studied herein are indeed accessible to NEM at the concentration and temperature conditions used. Circular dichroism (CD) studies of Pax5 (*32*) and Pax8 (*32*) reveal that the PD is largely unstructured in solution, adopting a more structured conformation with a larger α-helical content upon DNA binding. This is in agreement with our own protease sensitivity studies of Pax3 in solution and DNA-bound, that show a more compact less-accessible PD and HD upon DNA binding (*24*).

NEM modification of several SCMs (red regions in Figure 5) affected DNA binding by both the PD and the HD. This behavior is common for many Pax3 mutants bearing independent mutations in different portions of the PD that we have previously analyzed. In general, they may reflect an NEM mediated structural change in the PAI region of the PD that disrupts directly or indirectly conformation of the critical DNA binding α3 helix. Disruption of HD DNA binding in these mutants may be due to a direct structural change in the HD "docking" site identified here, or may suggest that docking of the HD onto the PD can happen only when the PD is in a DNA-bound state. Those SCM positions at which NEM modification only impaired DNA binding by the HD, while leaving intact the DNA binding properties of the PD (yellow regions, Figure 5), were of particular interest. Indeed, these mutants demonstrate directly that several residues in the identified "docking" site can modulate HD function without altering the DNA binding properties of the neighboring DNA binding determinants of the PAI subdomain, including the critical helix 3. Such mutants provide an important internal control for specificity of the site-specific modification strategy we have chosen to use. Overall, we favor an interpretation where NEM modification at all sensitive positions in the PD identified here causes a steric hindrance that reduces the ability of the HD to dock onto the HD binding pocket in the PD. Therefore HD docking onto the PD may be absolutely required for HD DNA binding, including dimerization on P2-type target sequences.

The location of a HD docking pocket at the "head" portion of the PAI subdomain is consistent with the hypothetical model proposed by Desplan (*28*) for the interaction of the PD and HD domains of Prd when bound to the chimeric probe PH0. Our results analyzed in the light of this model suggest that the HD can dock between helices 1 and 2 of the PAI subdomain and the DNA template. This situation is very similar to the interaction reported between the PD of Pax5 and another DNA-binding module, the ETS domain of Ets-1 (*29*). Ets proteins bind poorly to suboptimal Pax5/Ets binding sites, but affinity drastically increases when Pax5 is present through cooperative interactions between the 2 proteins (*29*). Residues of the *β*-turn in the *β*-hairpin as well as a portion of α2 of the DNA-bound Pax5 PD have been shown to be responsible for the recruitment of the ETS domain. The equivalent of the NEM sensitive residue Q40C in the Pax5 PD (Q22) is known to participate in hydrogen bonding and van der Waals contacts with Q336 and Y395 of the ETS domain. The equivalent of the NEM sensitive Pax3 residue L41C (L23 in Pax5) associates with the hydrocarbon portion of the side chain of residue K399 of the ETS domain of Ets1 via hydrophobic coupling. Finally the equivalent of the Pax3 residue R74, located in α2 of the PAI subdomain, in Pax5 (R56) contacts residue D398 and K399 of the ETS domain (*29*). The CL/R74C mutant did not show PD or HD NEM sensitivity, but residues flanking this position in α2 (pst: 72, 73 and 76, 78, 79) were found to be NEM sensitive. Taken together, the analysis of PD:HD interactions in SCM Pax3 mutants reported here, and structural studies of the interface of Pax5-ETS proteins interaction in the DNA-bound state, strongly suggest that the *β*-turn of the *β*-hairpin and parts of α2 of the PD perform critical protein−protein interaction functions.

Cysteine substitution at position 80 created a mutant (CL/H80C) that retains PD binding activity to P6CON but abrogates PD binding activity to P3OPT. Initially this was counterintuitive since one would suspect that a cysteine substitution would more likely disrupt binding to a DNA sequence optimized for the Pax3 PD (P3OPT) than to a DNA sequence corresponding to the consensus sequence for the PD of Pax6 (P6CON). Position 80 is the first residue in the DNA binding recognition helix α3 of the PAI subdomain, and it plays a critical role in DNA binding specificity (*12*, *13*). PD proteins, such as Pax3 and *Drosophila* Prd, have a histidine residue at the first position of α3 and have a preference for the 5′-GTCACGC-3′ DNA sequence (*12*, *13*). Other PDs, such as the Pax6 PD, show an asparagine at the first position of PAI α3 and have a preference for the 5′-TTCACGC-3′ DNA sequence (*12*, *13*). A histidine residue at position 1 of the PAI α3 hydrogen bonds the first guanine of the 5′-GTCACGC-3′. Instead of making hydrogen bonds with adenine in an AT base pair, the Asn at this position interacts with the thymine in the AT base pair via hydrophobic coupling (*12*, *13*). The methylene group (−CH_2−) β carbon of the side chain participates in hydrophobic interactions with the methyl moiety of first thymine of the 5′-TTCACGC-3′ sequence. This contact is stabilized by a water-mediated interaction between the amide group of the Asn side chain and the sugar phosphate backbone of DNA (*12*, *13*). The substitution of a cysteine at position 80 of Pax3 essentially converts the specificity of the PD from Prd-like to Pax6-like. This can be accounted for since the hydrophobic cysteine side chain is more likely able to participate in hydrophobic interaction with a methyl group of a thymine than participate in hydrogen bonding to a guanine of DNA. This would account for the unusual behavior of this mutant toward different PD targets.

Eight out of thirteen NEM sensitive positions in the PAI subdomain consist of residues with hydrophobic side chains. This suggests that interaction of the HD onto this putative PD docking site most likely involves hydrophobic interactions, and corresponding hydrophobic residues in the HD. The identity of these residues remains unknown at present, although several observations in chimeric and mutant proteins (*21*) point to the N-terminal arm immediately upstream of helix 1 of the HD as a good candidate for such interactions. However, the identification of a subset of PD positions at which NEM modification impairs DNA binding by the HD and PD provides anchor points for the systematic search of potential interacting residues in the HD by a similar cysteine scanning mutagenesis approach. In this approach, a bifunctional sulfhydryl cross-linker (e.g. copper phenanthroline) could be used to look for cross-links between individual PD anchor cysteines, and a series of additional single cys mutants in discrete portions of the HD (*30*, *31*). In this scheme, successful cross-links can manifest themselves as species of higher molecular mass on SDS−PAGE and can be further validated using inserted factor Xa cleavage sites with immunoblotting against epitope tags inserted at the NH2 and COOH-terminus of the proteins.

Together, results of site-specific modification experiments agree with the proposition that the PD and HD do not function as independent DNA binding modules in Pax3, but instead the PD and HD may physically and functionally interact, in the full length Pax3 protein, so that one domain can modulate the DNA binding properties of the other domain. These interactions may be important for target site selection by the Pax3 protein in vivo. The results of this study also explain the loss of PD and HD activity seen in the G42R Sp^d mutant variant of Pax3. G42R maps within the boundaries of the proposed hydrophobic HD docking pocket, and independent mutations at that site behave as complete loss-of-function (*33*). The introduction of a bulkier and charged side chain at position 42 may prevent HD docking via steric hindrance, loss of hydrophobic interface, or both.

REFERENCES

1. Stuart, E. T., Kioussi, C., and Gruss, P. (1994) *Annu. Rev. Genet. 28*, 219−36.
2. Bober, E., Franz, T., Arnold, H. H., Gruss, P., and Tremblay, P. (1994) *Development 120*, 603−12.
3. Beechey, C. V., and Searle, A. G. (1986) *Mouse News Lett. 75*, 28.
4. Auerbach, R. (1954) *J. Exp. Zool. 127*, 305−29.
5. Franz, T. (1990) *Acta Anat. 138*, 246−53.
6. Franz, T. (1989) *Anat. Embryol. 180*, 457−64.
7. Goulding, M., Lumsden, A., and Paquette, A. J. (1994) *Development 120*, 957−71.
8. Baldwin, C. T., Hoth, C. F., Amos, J. A., da-Silva, E. O., and Milunsky, A. (1992) *Nature 355*, 637−8.
9. Baldwin, C. T., Hoth, C. F., Macina, R. A., and Milunsky, A. (1995) *Am. J. Med. Genet. 58*, 115−22.
10. Bopp, D., Burri, M., Baumgartner, S., Frigerio, G., and Noll, M. (1986) *Cell 47*, 1033−40.
11. Mansouri, A., Chowdhury, K., and Gruss, P. (1998) *Nat. Genet. 19*, 87−90.
12. Xu, W., Rould, M. A., Jun, S., Desplan, C., and Pabo, C. O. (1995) *Cell 80*, 639−50.
13. Xu, H. E., Rould, M. A., Xu, W., Epstein, J. A., Maas, R. L., and Pabo, C. O. (1999) *Genes Dev. 13*, 1263−75.
14. Wilson, D. S., Guenther, B., Desplan, C., and Kuriyan, J. (1995) *Cell 82*, 709−19.
15. Wilson, D., Sheng, G., Lecuit, T., Dostatni, N., and Desplan, C. (1993) *Genes Dev. 7*, 2120−34.
16. Treisman, J., Gonczy, P., Vashishtha, M., Harris, E., and Desplan, C. (1989) *Cell 59*, 553−62.
17. Schafer, B. W., Czerny, T., Bernasconi, M., Genini, M., and Busslinger, M. (1994) *Nucleic Acids Res. 22*, 4574−82.
18. Glaser, T., Walton, D. S., and Maas, R. L. (1992) *Nat. Genet. 2*, 232−9.
19. Lalwani, A. K., Brister, J. R., Fex, J., Grundfast, K. M., Ploplis, B., San Agustin, T. B., and Wicox, E. R. (1995) *Am. J. Hum. Genet. 56*, 75−83.
20. Underhill, D. A., Vogan, K. J., and Gros, P. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 3692−6.
21. Fortin, A. S., Underhill, D. A., and Gros, P. (1998) *Nucleic Acids Res. 26*, 4574−81.
22. Fortin, A. S., Underhill, D. A., and Gros, P. (1997) *Hum. Mol. Genet. 6*, 1781−90.
23. Apuzzo, S., and Gros, P. (2002) *Biochemistry 41*, 12076−85.
24. Apuzzo, S. and Gros, P. (2004) *J. Biol. Chem. 279*, 33601−12.
25. Epstein, J., Cai, J., Glaser, T., Jepeal, L., and Maas, R. (1994) *J. Biol. Chem. 269*, 8355−61.
26. Chalepakis, G., and Gruss, P. (1995) *Gene 162*, 267−70.
27. Epstein, J. A., Lam, P., Jepeal, L., Maas, R. L., and Shapiro, D. N. (1995) *J. Biol. Chem. 270*, 11719−22.
28. Jun, S., and Desplan, C. (1996) *Development 122*, 2639−50.
29. Garvie, C. W., Hagman, J., and Wolberger C. (2001) *Mol. Cell 8*, 1267−76.
30. Frillingos, S., Sahin-Toth, M., Wu, J., and Kaback, H. R. (1998) *FASEB J. 12*, 1281−99.
31. Loo, T. W., and Clarke, D. M. (2000) *J. Biol. Chem. 275*, 5253−6.
32. Tell, G., Scaloni, A., Pellizzari, L., Formisano, S., Pucillo, C., and Damante, G. (1998) *J. Biol. Chem.* 273, 25062−72.
33. Underhill, D. A., and Gros, P. (1997) *J. Biol. Chem. 272*, 14175−82.

BI060009Z