Reciprocal Regulation of DNA Binding of the Paired domain and Homeodomain of Pax3.

Sergio Apuzzo

Department of Biochemistry

McGill University, Montreal

2006

A thesis submitted to the Faculty of Graduate Studies and Research in full fulfillment of the requirements of the degree of Doctor of Philosophy

Your file Votre référence
ISBN: 978-0-494-32290-1
Our file Notre référence
ISBN: 978-0-494-32290-1

Copyright permission form journals:

This work contains the following articles:

Apuzzo, S., A. Abdelhakim, A. S. Fortin and P. Gros (2004). "Cross-talk between the paired domain and the homeodomain of Pax3: DNA binding by each domain causes a structural change in the other domain, supporting interdependence for DNA Binding." J Biol Chem 279(32): 33601-12.

Apuzzo, S. and P. Gros (2002). "Site-specific modification of single cysteine Pax3 mutants reveals reciprocal regulation of DNA binding activity of the paired and homeo domain." Biochemistry 41(40): 12076-85.

Apuzzo, S. and P. Gros (2006). "The paired domain of Pax3 contains a putative homeodomain interaction pocket defined by cysteine scanning mutagenesis." Biochemistry 45(23): 7154-61.

These articles are reproduced in full in this work with permission from *Biochemistry* and *The Journal of Biological Chemistry.*

Abstract

Pax3 is a gene that encodes for a transcription factor involved in embryogenesis. Pax3 heterozygote mutants in mice (Splotch) and humans (Waardenburg Syndrome) exhibit pigmentary disturbances. Also, neural tube defects (spina bifida) develop in homozygous *splotch* mice. Pax3 has 2 DNA binding domains, a paired Domain (PD), and a paired-type homeodomain (HD). We have used site-specific modification of single Cys Pax3 mutants (in a Cys-less backbone) to probe structure: function relationships. N-Ethyl maleimide (NEM) modification of a single Cys in the PD (C82) abrogates DNA binding by the PD but also by the HD. Likewise, NEM modification of V263C in the HD abrogates DNA binding by both the HD and the PD. These results demonstrate that the PD and HD do not function as independent DNA binding modules in Pax3, but are functionally interdependent. Pax3 was also modified by insertion of Factor Xa protease sites in either the PD (Xa100) or the HD (Xa216). The effect of DNA binding by the PD and HD on conformational changes in the other domain was monitored by accessibility of the Factor Xa site. Binding by the PD causes a structural change not only in the PD but also in the HD; similarly, DNA binding by the HD causes a conformational change in the PD. These results provide a structural basis for the functional interdependence of the PD and HD in Pax3. Site-specific modification of single Cys mutants was further used to identify a cluster of solvent exposed residues in the β-hairpin structure and in the first 2 helices of the PD, which when modified (with NEM) impair DNA binding by the HD. Three-dimensional modeling shows that these residues form a hydrophobic pocket possibly defining the site of interaction of the HD. Pax3 double cysteine mutants were also generated and used with bifunctional cysteine specific reagents to monitor proximity relationships in Pax3. We have observed that cysteines placed in the N-terminal portion of the HD can be crosslinked with cysteines placed in the putative HD interaction pocket in the PD, suggesting proximity of the two segments. Taken together, these results

suggest that residues in or close to the PD helix 2 and those in or close to the N-terminal arm of the HD undergo direct physical interaction, which we postulate is responsible of the functional interdependence we observe between the PD and HD of Pax3.

Résumé

Pax3 est un facteur transcriptionel impliqué dans le développement embryonnaire. De manière hétérozygote, la mutation de Pax3 provoque une désorganisation pigmentaire et est reliée à l'appellation "Splotch" chez la souris et "Syndrôme de Waardenburg" chez l'humain. Dans sa forme homozygote, la mutation de Pax3 cause des anomalies du tube neural (Spina bifida) chez la souris. Pax3 possède deux domaines de liaisons à l'ADN: un domaine paired (PD) et un homéodomaine (HD). Nous avons utilisé une technique de modification chimique ciblant l'unique cysteine d'une série de formes mutantes de Pax3 construites à partir d'un Pax3 modifié, à l'intérieur duquel tous les résidues de cysteine indigènes avaient préalablement été remplacées. Permettant l'analyse de sites spécifiques, cette technique fut utilisée afin d'étudier la relation entre la structure et le fonctionnement du facteur Pax3. La modification au N-Ethyl Maleimide (NEM) d'une cysteine unique à l'intérieur du PD (C82) provoque l'inhibition de l'affinité à lier l'ADN par le PD aussi bien que par le HD. De manière similaire, la modification au NEM de V263C à l'intérieur du HD se transcrit aussi par la perte de l'habilité à lier l'ADN par le HD aussi bien que par le PD. Ces résultats nous démontrent que le PD et le HD ne fonctionnent pas simplement de façon indépendante, mais bien par interdépendance pour permettre à Pax3 de se lier à l'ADN (Apuzzo et al., 2002). Pax3 fut aussi modifié par l'insertion de sites de facteur de protéase Xa en son PD (Xa100) ou en son HD (Xa216). Les effets de la liaison à l'ADN par le PD et par le HD sur les changements de conformation dans l'un et l'autre domaine furent examiné par accessibilité au site du facteur Xa. La liaison du PD produit un changement structurel non seulement dans le PD, mais aussi dans le HD, et la liaison du HD produit également un effet similaire dans le PD ainsi que dans le HD. Ces résultats amènent une explication moléculaire à l'interdépendance fonctionnelle entre le PD et le HD à l'intérieur de Pax3 (Apuzzo et al., 2004). La modification de mutants à cysteine unique positionnée à des sites spécifiques fut utilisée de nouveau pour identifier un cluster de

résidus exposés au solvant dans la structure de l'épingle à cheveux beta et dans les deux premières hélices du domain PD. La modification au NEM de ces résidus correspond à une perte de la capacité à lier l'ADN par le HD et la modélisation 3-D de ces mêmes résidus du PD les démontre comme faisant partie d'une portion hydrophobe du site d'interaction avec le HD (Apuzzo et al., 2006). Enfin, des résidus de cysteines furent insérées aux mêmes sites que pour les mutants de Pax3 à cysteine unique, et ce, en differentes combinaisons de deux cysteines par mutants (mutants double-Cys), afin d'y combiner des reactifs bifonctionnels specifiques aux cysteines pour y détailler les différentes relations de proximité à l'intérieur de Pax3 par pontage. Nous avons observé que des cysteines placées dans la portion amino-terminale du HD pouvaient être pontées à des cysteines placées dans une portion du PD putative comme site d'interaction avec le HD, ce qui suggère une proximité entre ces deux segments (Apuzzo et al., 2006). En conclusion, ces résultats semblent démontrer que des résidus près de ou à l'intérieur même de la seconde hélice du PD ainsi que certains résidus près de ou à l'intérieur même du segment amino-terminal du HD sont impliqués dans une intéraction physique directe, ce que nous proposons être responsable de l'interdépendance fonctionnelle observée entre le PD et le HD du facteur Pax3.

Preface

The studies in Chapters 2,3,4 and 5 of the thesis have been published in the following journals:

Chapter 2: Apuzzo, S., and Gros, P. (2002). Site-Specific Modification of Single Cysteine Pax3 Mutants Reveals Reciprocal Regulation of DNA Binding Activity of the Paired and Homeodomain. Biochemistry 41, 12076 – 12085.

Chapter 3: Apuzzo, S., Abdelhakim A., Fortin, A. S., and Gros, P. (2004). Cross-talk between the Paired Domain and the Homeodomain of Pax3. Journal of Biological Chemistry, Vol. 279, No. 32, 336 01 – 33612.

Chapter 4: Apuzzo, S., and Gros, P. (2006). The Paired Domain of Pax3 Contains a Putative Homeodomain Interaction Pocket Defined by Cysteine Scanning Mutagenesis. Biochemistry 45, 7154 – 7161.

Chapter 5: Apuzzo, S., and Gros, P. (2006). Cooperative interactions between the two DNA binding domains of Pax3: Helix 2 of the Paired domain is in close proximity of the amino terminus of the Homeodomain (submitted to JBC).

Contribution of Authors

Dr. A. S. Fortin and Aliaa Abdelhakim:

For studies shown in Chapter 3, both Dr. Fortin and Ms. Abdelhakim generated a subset of the expression plasmids.

Dr. Kalle Gehring (McGill University) and Dr. H. R. Kaback (UCLA):

For studies presented in Chapter 2, both Dr. Gehring and Dr. Kaback have participated in constructive discussions and provided helpful suggestions.

Dr. Philippe Gros:

Throughout all four projects undertaken Dr. Gros provided essential supervision, and guidance. Together we have collaborated on the experimental strategies used as well as discussed all data obtained. He has also played an active role in the writing and editing of all the manuscripts that are included in this thesis.

Table of contents

List of figures

Chapter 3

Chapter 4

List of Tables

Acknowledgement

I want to express my appreciation for all the knowledge and freedom that my thesis supervisor, Dr. Philippe Gros, has provided me during my doctoral studies. As a result of the experiences I had under his supervision and guidance I feel properly prepared for any scientific challenge I may face in the future. As a result of my stay in his lab I am not only able to confidently conduct research, design projects and write scientific papers, but I am also able to teach various scientific topics (chemistry, biology, biochemistry) and present scientific findings (my own and those of others) in a systematic, thought provoking manner. I enjoyed the many teaching opportunities he has allowed me to undertake which includes training several undergraduates assigned to his lab. I look forward to implementing my research and teaching skills in the future when I acquire my own lab. I would also like to show my gratitude for Dr. A. S. Fortin who has offered me guidance during the first few years of graduate studies. I also want to thank all my students who have contributed to the progression of my research, especially to Aliaa Abdelhakim, who has made a contribution to my third chapter. I am also grateful to the funding agencies that have provided studentships to me during my doctoral studies. They include: Fonds da la Recherche en Santé du Québec, Faculty of Medicine, McGill Graduate Studies Fellowship and the Canadian-Italian Business and Professional Association (CIBPA). I would like to thank past and present members of the lab that have contributed to my joyful experience in the lab, they include: Zoha Kibar, Lucy Mitsos, Jie Cai, Isabelle Carrier, Elena Torban, John Forbes, Karine Turcotte, Anne Fortier, Patrick Fortin, Melissa Mathieu, Normand Groulx, Hui Jun Wang, and Susan Gauthier. I would like to thank Charles Meunier, Luisa De Marte, Geneviève Thibault, Anne-Marie Patenaude, Gundula Min-oo and Joanne Berghout for helping me revise my thesis. Last but not least, I want to send my heart-felt gratitude to my family and friends who have all

supported me emotionally during my studies, they include my mother Maria Apuzzo,my sister Anna Maria Apuzzo, my brothers Enrico and John Apuzzo, and my good friends Tania Pietrantonio and Jon Paterson.

Objectives of the presented work

Pax3 is a member of a family of transcription factors all of which are all involved in development (Dahl et al., 1997). Members of the Pax family are defined by the presence of a 128 amino acid DNA binding domain known as the paired domain (PD) (Walther et al., 1991). Some members of the Pax family, which includes Pax3, also contain a second DNA binding domain, the paired-type homeodomain (HD) (Stuart et al., 1994). Many studies have shown that for members of the Pax family that contain both domains the PD and the HD may be structurally independent and show a functional interdependence of DNA binding; that is the two domains each consist of their own primary and secondary structures but their ability to bind DNA involves some sort of communication between the domains. The objective of the work presented is to determine the mechanism underlying this observed interdependence. These studies will contribute to understanding the different modes in which Pax proteins bind DNA and regulate transcriptional expression of target genes during embryogenesis.

Chapter 1

Literature Review

1.1 **Historical perspective**

Many genes involved in the development of *Drosophila melanogaster* were discovered via large scale genetic screening (Nusslein-Volhard and Wieschaus, 1980). Many genes responsible for mammalian development have been discovered by sequence homology with known developmental genes of the fly. These genes include the mammalian family of Pax genes (Burri et al., 1989; Deutsch et al., 1988; Dressler et al., 1990; Stapleton et al., 1993; Walther et al., 1991; Wallin et al., 1993; Walther and Gruss, 1991). Pax family members are defined by the presence of the paired box, which encodes a DNA binding domain known as the PD. The paired box sequence was identified in flies as a region of homology between the *Drosophila* Prd gene and two linked genes present at the *gooseberry* locus (Bopp et al., 1986). The Prd gene of *Drosophila* was discovered during large genetic screens and is classified as a pair-rule segmentation gene (Nusslein- Vollard and Wiechaus, 1980). Mammalian Pax genes were first identified by homology with the paired box sequence present in the *Paired* (Prd) *Drosophila* gene. This led to the discovery of the Pax1 protein followed by the other members of the Pax family.

Sequence analysis of the *Prd* gene identified a homeobox, which codes for the DNA binding domain known as the HD (Frigerio et al., 1986). The possible function of the PD as a DNA binding domain was suggested by the discovery of a DNA binding activity of the Prd protein that was not mediated by the HD (Treisman et al., 1989). Extensive studies of both the *Drosophila* Prd and the mammalian Pax-1 proteins revealed that the PD is a sequence specific DNA binding domain (Chalepakis et al, 1991; Treisman et al., 1991). Analysis of naturally occurring and engineered mutations in Pax genes confirmed their important roles in various stages of embryogenesis. The first part of this literature review will provide an overview of the mammalian Pax gene family, starting with a brief outline of the classification of Pax genes

into subfamilies, followed by an exploration of the genetic disorders that are caused by loss-of-function mutations in individual Pax genes. Subsequently, the biochemical properties as well as known physiological targets of Pax proteins will be discussed.

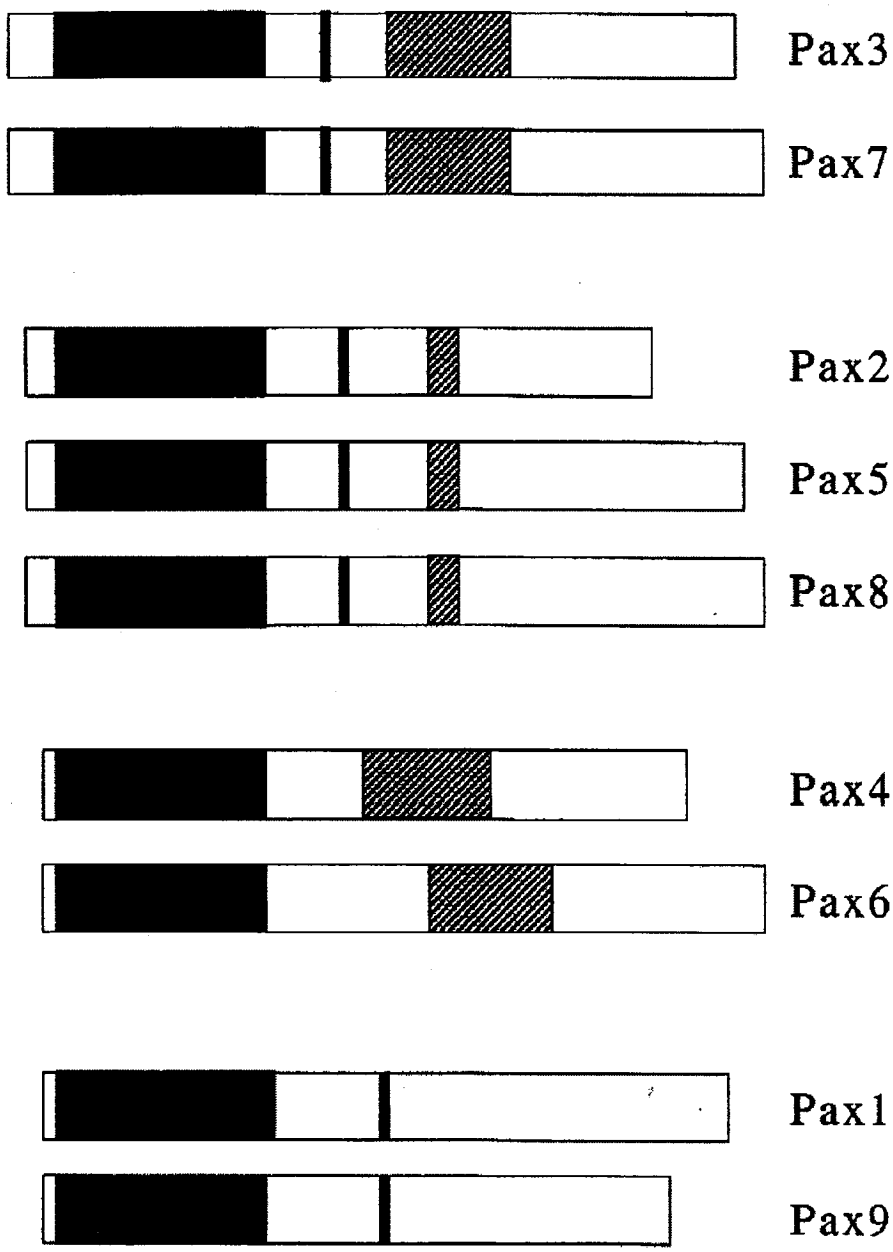
1.2 Classification of Pax genes into subfamilies

The Pax family consists of nine members, Pax1 to Pax9. The family is defined by the presence a paired domain (PD) at the N-terminus (Figure 1). Pax3 Pax4, Pax6 and Pax7, like Prd, also encode for a second DNA binding domain, the paired-type homeodomain (HD) (Frigerio et al., 1986). Pax2, Pax5 and Pax8 proteins only contain a portion of the HD and Pax1 and Pax9 do not have any HD (reviewed in Dahl et al., 1997; Mansouri et al., 1996; Stuart and Gruss, 1995; Stuart et al., 1994). With the exception of Pax4 and Pax6, all Pax members also contain a conserved octapeptide motif positioned between the PD and the HD. All Pax proteins contain a transactivation domain at their C-termini.

This classification scheme is supported by the observation that within the same subfamily Pax genes share a high degree of sequence identity within their and outside of their PD, show similar genomic organizations and have similar temporal and spatial expression patterns (Dahl et al., 1997; Mansouri et al., 1996; Stuart and Gruss, 1995; Walther et al., 1991).

Figure 1

Schematic representations of the structural features of members of the Pax family. All members contain the 128 amino acid PD (large black box) at the N-terminus. A subset of Pax proteins also contain a conserved octapeptide motif (small black box) and/or a complete or partial HD (hatched box). Pax proteins are classified into subfamilies based on the presence or absence of these 3 structural features.


Pax3
Pax7
Pax2
Pax5
Pax8
Pax4
Pax6
Pax1
Pax9

1.2.1 The Pax1/Pax9 subfamily

Pax1 and Pax9 are members of the same subfamily and contain PDs with 98% sequence identity as well as the conserved octapeptide, but do not have a HD (Chalepakis et al., 1991; Deutsch et al., 1988; Neubuser et al., 1995; Stapleton et al., 1993; Wallin et al., 1993). Both members share very similar, but not identical, temporal and spatial expression patterns in the limb buds, the developing vertebral column, and in adult and embryonic thymus (Deutsch et al., 1988; Neubuser et al., 1995). The developing vertebral column arises from Pax1/Pax9 expressing sclerotome cells located in the ventral somites. Interestingly, only Pax9 is expressed in mesenchyme involved in tooth development (Neubuser et al., 1997). In agreement with their expression patterns, Pax1 and Pax9 mutants develop abnormalities of the thymus as well as certain skeletal elements. The *Un (Undulated)* mouse Pax1 mutant allele contains a glycine-to-serine mutation in the PD and results in vertebral (Balling et al., 1988; Wright, 1947) and thymus dysgenesis (Dietrich and Gruss, 1995; Timmons et al., 1994; Wallin et al., 1996). A less severe phenotype is observed in Pax1 heterozygotes thereby demonstrating that *Un* Pax1 mutation is not recessive but semi-dominant, a characteristic common to several Pax genes (Dietrich and Gruss, 1995; Wilm et al. 1998). Pax9 mutants, generated by homologous recombination, have no phenotype in the heterozygote but homozygotes are devoid of teeth. No defects are seen in the skeletal system of homozygotes suggesting that Pax1 compensates for the lack of Pax9 activity during skeletogenesis (Peters et al., 1998). Human PAX9 heterozygotes suffer from autosomal dominant oligodentia and experience the agenesis of six or more permanent teeth (Stockton et al., 2000). Pax1 and Pax9 double knock out mice have more severe vertebral dysgenesis than Pax1-/- or Pax9-/- knock out mice and reveal a gene dosage effect in vertebral column development (Peters et al.,

1999). Therefore Pax1 and Pax 9 have unique roles during embryogenesis but cooperate or have redundant functions during the development of the skeletal system.

1.2.2 The Pax2/Pax5/Pax8 subfamily

Members of this subfamily contain a PD as well as an octapeptide and a partial HD, which contains only the first 23 amino acids of the 60 amino acid domain (Adams et al., 1992; Asano and Gruss, 1992; Dressler et al., 1990; Plachov et al., 1990). All members of this subfamily are expressed in the developing central nervous system, more specifically, in the midbrain-hindbrain boundary (Adams et al., 1992; Asano and Gruss, 1992; Nornes et al., 1990; Plachov et al., 1990). Pax2 and Pax8 are both expressed in the developing kidney and are both downregulated once terminal differentiation of kidney epithelial structures occurs (Dressler and Douglas, 1992; Plachov et al., 1990; Poleev et al., 1992). Pax2 and Pax8 expression is observed in kidney epithelium in the Wilm's tumor renal malignancy where embryonically-derived tissues are unable to complete the mesenchymal-epithelial transition (Dressler and Douglas, 1992; Eccles et al., 1995; Poleev et al., 1992; Tagge et al., 1994). Pax2, Pax5 and Pax8 also have distinct expression domains. Pax2 is expressed in the developing inner ear and in the developing eye (optic cup and optic stalk), while Pax5 expression is found in B-lymphoid tissues and lineages involved in early B-cell differentiation. Pax8 is found expressed in the developing and adult thyroid gland (Nornes et al., 1990; Torres et al., 1996; Adams et al., 1992; Plachov et al ., 1990; Zannini et al., 1992).

Autosomal dominant human PAX2 mutations give rise to renal-coloboma syndrome, which is characterized by optic nerve colobomas, renal hypoplasia (e.g. fewer nephrons) as well as auditory and central nervous system defects (Eccles and Schimmenti, 1999; Schimmenti et al., 1997). Optic nerve coloboma results from lack of closure of the optic fissure (Favor et al., 1996; Torres et al., 1996). The *Krd* (kidney and retinal defects) mouse is a heterozygous mutant of Pax2 and displays retinal and renal defects. In accordance with the

semi-dominant nature of Pax2 mutations, both the mouse and human, Pax2 mutant homozygotes experience kidney agenesis and embryonic death (Eccles and Schimmenti, 1999; Keller et al., 1994). Induced overexpression of Pax2 is also detrimental to proper development and results in severe kidney abnormalities. The nature of these abnormalities suggests that persistent expression of Pax2 inhibits terminal differentiation of glomerular and tubular epithelia (Torres et al., 1995). Pax5 mutants display aberrant midbrain/hindbrain formation as well as a complete lack of functional B cells since B cell differentiation is halted at an early pre-B-cell stage (Urbanek et al., 1994). Mouse Pax8 homozygous mutants experience thyroid dysgenesis, more specifically, they develop smaller thyroid glands (Mansouri et al., 1998). Also, PAX8 mutations in humans results in hypothyroidism (Macchia et al., 1998).

1.2.3 The Pax3/Pax7 subfamily

Members of this subfamily contain all three structural features: a PD, octapeptide and the HD. Both Pax3 and Pax7 have an additional isoform that results from alternative splicing and that differs from the original sequence by the addition of a glutamine residue in the linker between the two subdomains within the PD (Vogan et al., 1996).

Both genes are expressed during neurogenesis with overlapping but not identical spatial and temporal expression domains. During early embryogenesis both genes are expressed in the dorsal neural tube (Goulding et al., 1991; Jostes et al., 1990; Stoykova and Gruss, 1994). Pax3 and Pax7 expression requirement in neurogenesis is supported by the observation that diffusible factors Wnt1, 3a, 4 and 6 are required to maintain Pax3 and Pax7 expression in the neural tube (Otto et al., 2006). One difference of note is that Pax3 expression precedes the closure of the neural tube (Pax3 expression is first detected in the tips of the neural folds) whereas Pax7 expression begins soon after closure. Pax3 and Pax7

expression is also evident in neural crest cells, a population of cells derived from the dorsal neural tube following closure that have the capacity to migrate to various areas of the embryo to give rise to several unrelated structures (Conway et al., 1997; Goulding et al., 1991; Mansouri et al., 1996). Neural crest cells contribute to the formation of melanocytes, the peripheral nervous system (e.g. dorsal root ganglia) and the heart, among other structures.

Pax3 and expression can also be found in somitic mesoderm that gives rise to myoblasts and then muscle tissues (Goulding et al., 1994; Jostes et al., 1990). A portion of Pax3 expressing cells in the somite have the ability to migrate into the limb to form limb musculature, but this expression is suppressed once in the limb and the activation of myogenic markers occurs. Pax7 expression is seen in proliferating myoblasts, in satellite cells as well as adult skeletal muscles (Goulding et al., 1994; Jostes et al., 1990; Seale et al., 2000; Williams; Ordahl, 1994 and Oustanina et al., 2004). Satellite cells are quiescent myogenic precursor cells located between the basal membrane and the sarcolemma of myofibers and they are characterized by Pax7 expression (Kawiak et al., 2006). Overall, it is obvious that Pax3 and Pax7 play a role in myogenesis, neurogenesis and formation of neural crest cell derived structures.

The Pax3 heterozygous mutant mouse *Splotch* has hypopigmentation defects which consist of white spotting of the belly, feet and tail. This pigmentation disturbance is due to the inability of a sufficient number of neural crest cell derived melanocytes to migrate to these areas (Auerbach, 1954). Many *Splotch* alleles exist and are either naturally occurring or radiation induced. Characterization of these alleles has shown that they are the result of deletions, missense, nonsense or frameshift mutations in Pax3 (Epstein et al., 1991; Epstein et al., 1993; Goulding et al., 1993; Vogan et al., 1993). The mutation in the Sp allele is complex and lies within intron 3 of the Pax3 gene resulting in the production of four distinct, aberrantly spliced mRNA transcripts. Three of these transcripts result in premature termination with the

absence of an intact PD, octapeptide motif, and HD. The fourth lacks exon 4, resulting in the loss of the C-terminal part of the PD as well as loss of the octapeptide motif (Epstein et al., 1993). The mutation in the Sp^{2H} mutant allele is a 32 basepair deletion in the HD (Epstein et al., 1991).

All Sp mutations are lethal in the homozygous state but the longevity of the homozygous mutant Pax3 mouse is different between alleles. Mutant alleles Sp and Sp^{2H} cause embryonic death during mid-gestation due to the absence of a septum in the outflow tract of the heart, which is formed only when a sufficient number of cardiac neural crest cells migrate from the neural tube to the heart. This septum is required to convert the outflow tract of the primordial heart into the aorta and pulmonary artery. The absence of the septum results in constant mixing of oxygenated and non-oxygenated blood and leads to heart failure in mid-gestation (Conway et al., 1997; Franz, 1989). The Sp^{r} (Sp-retarded) allele causes pre-implantation death of the embryo in the homozygous mutant. This drastic effect is due to the fact that this allele is mutated by a cytogenetically detectable deletion of the Pax3 locus as well as surrounding loci. In fact, the initiation of a positional cloning effort to define the molecular basis of the mouse Splotch phenotype began with an analysis of Sp^{r} mice (Evans et al., 1988). The hypomorph Sp^{d} (Sp-delayed) allele consists of a missense mutation where glycine 9 of the PD (glycine 42 in the Pax3 protein) is changed to an arginine. Sp-delayed homozygous mutant embryos survive to term but die soon after birth due to the lack of the diaphragm muscle required for breathing (Auerbach, 1954; Beechey and Searle, 1986; Dickie, 1964). Sp homozygous mutants experience neural tube and neural crest defects as well as an absence of limb musculature (Auerbach, 1954; Franz, 1990; Franz, 1989; Franz, 1993; Franz et al., 1993; Bober et al., 1994; Goulding et al., 1994). The lack of closure in the neural tube of Sp mutants results in spina bifida, exencephaly or anencephaly (Auerbach, 1954). Sp homozygous mutants display defects in the formation of the peripheral nervous system as

well. More specifically, the absence of Pax3 causes dysgenesis of dorsal root ganglia (Kioussi et al., 1995; Serbedzija and McMahon, 1997; Tremblay et al., 1995). Recently, it has been found that decreased Pax3 expression induced by maternal hyperglycemia (diabetic pregnancy) leads to an increase in incidence of neural tube defects (Loeken et al., 2006). This observation is in line with the belief that Pax3 is involved in neural tube formation.

Homozygous Sp mice also display a disruption in myogenesis. These mice lack structural organization of the somites that results in malformations or reduction of trunk musculature. Also, these mice lack any migrating myoblasts that are responsible for populating the limb bud and giving rise to limb muscles, consequently no limb musculature results (Bober et al., 1994; Franz et al., 1993; Goulding et al., 1994).

Humans with only one functional copy of PAX3 develop Waardenburg syndrome (WS). This semi-dominant disorder is characterized by pigmentary disturbances, sensorineuronal deafness and craniofacial abnormalities (Waardenburg, 1951). Although there are some differences in the phenotypes between the Sp mouse and human WS patients, the similarities in the structures affected implies that the role of Pax3 is conserved. Although, severe limb musculature defects are observed in some WS homozygotes, none show the neural tube defects that are prevalent in Sp mouse homozygous mutants (Ayme and Philip, 1995; Zlotogora et al., 1995).

Since no naturally occurring Pax7 mutant allele exists, one was created by homologous recombination (Mansouri et al., 1996). Unlike other Pax proteins, no observable phenotype exists in the heterozygous mutant. Pax7 homozygous mutants survive to term but die soon after birth due to insufficient muscle growth. At birth, Pax7 homozygous mutants are grossly normal but upon closer inspection one can observe a decrease in muscle mass and fiber caliber. These observations seem to be due to the fact that Pax7-/- mice lack satellite

cells required for post-natal muscle growth, maintenance and repair (Mansouri et al., 1996; Mansouri et al., 1996; Seale et al., 2000).

A new Pax3 null allele was made recently by replacing the first exon with a gene encoding Cre recombinase (Engleka et al., 2005). This allele was used with a β-galactosidase reporter gene to generate a fate-map of Pax3 derivatives in the developing mouse. As expected, neural crest and some somatic derivatives have been identified (Engleka et al., 2005). However, this study also revealed Pax3-expressing precursors in the colonic epithelium of the hindgut and within the urogenital system (Engleka et al., 2005). This Pax3 expression has not been characterized to date.

1.2.4 The Pax4/Pax6 subfamily

Members of this subfamily contain both a PD and HD but lack the octapeptide motif (Matsushita et al., 1998). Unlike Pax4, Pax6 expression patterns have been thoroughly studied. Like most members of the Pax family, Pax6 is also expressed in the developing central nervous system. Pax6 is expressed in the ventral neural tube upon neural tube closure and is also expressed in the developing brain, more specifically in the forebrain, telencephalon, diencephalon and myelecephalon (Puschel et al., 1992; Stoykova and Gruss, 1994; Walther and Gruss, 1991). Pax6 expression in the eye has been extensively studied and early on expression occurs in the optic vesicle and the optic stalk, while later in development expression is seen in the retina, cornea and the lens. In accordance with similarities between lens and nasal placode formation, Pax6 is also found expressed in olfactory epithelium and the olfactory bulbs (Grindley et al., 1995; Puschel et al., 1992; Walther and Gruss, 1991). Pax6 expression has also been detected in the developing pancreas but only in certain endocrine cells (Sander et al., 1997; St-Onge et al., 1997; Turque et al., 1994). Pax4 transcripts have also

been detected in endocrine precursor cells of the pancreas (Smith et al., 1999; Sosa-Pineda et al., 1997).

Targeted inactivation of Pax4 led to defects in the pancreas only (Sosa- Pineda et al., 1997). The adult endocrine function of the pancreas is controlled by four islet cell types known as α, β, δ and PP cell which respectively synthesize glucagons, insulin, somatostatin and the pancreatic polypeptide. Development of these cells occurs from pluripotent progenitors that initially commit to the β/δ or α/PP cell fate and terminally differentiate later during development (Dohrmann et al., 2000). Pax4-/- mice die soon after birth and exhibit growth retardation as well as dehydration. These mice are unable to form mature β and γ cells and have an increased number of α cells (Sosa-Pineda et al., 1997). This observation suggests that Pax4 plays a role in endocrine cell differentiation during pancreatic development.

Pax6 mutant alleles are semi-dominant in both mouse and human. In the heterozygous state Pax6 mutations cause the Small eye (Sey) phenotype in mice and aniridia in humans (Glaser et al., 1994; Hanson et al., 1993; Hanson et al ., 1994; Hill et al., 1991; Hogan et al., 1988; Hogan et al., 1986; Jordan et al., 1992; Ton et al., 1991). Variations in phenotype exist for different Pax6 mutant alleles in both mice and humans but all cause ocular malformations in the heterozygous mutants including cataracts and opacification of the cornea, iris hypoplasia and absence or reduction in size of the lens (Callaerts et al., 1997). Pax6-/- mutants die at birth and exhibit malformations in the central nervous system and lack eyes and a nose (agenesis of nasal cavities or olfactory bulbs) (Grindley et al., 1995; Stoykova et al., 1996). Pax6-/- mice also develop defects in the pancreas. Some Pax6 mutant alleles cause agenesis of only glucagon producing α cells while others result in a decrease in number of all 4 of the pancreatic endocrine cell types (Sander et al., 1997; St-Onge et al., 1997). Therefore, a strong correlation exists between the expression patterns of Pax6 and the organs that are malformed when Pax6 expression is downregulated. As for many other Pax genes, the semi-

dominance observed in naturally occurring mutants suggests a dose requirement for Pax6 during eye development. This is further supported by the observation that transgenic mice that overexpress Pax6 also develop ocular defects that are very similar to those that occur when Pax6 is underexpressed (Schedl et al., 1996). There is a strong sensitivity of cells involved in ocular development to changes in endogenous levels of Pax6 expression.

The Pax6 homolog in *Drosophila*, eyeless (ey), results in agenesis of the fly's compound eyes when mutated. Also, when it is ectopically expressed in leg, wing and antennal imaginal discs leads to the formation of ectopic eyes (Hauck et al., 1999; Quiring et al., 1994; Xu et al., 1999; Czerny et al ., 1999; Halder et al., 1995). Also, ectopic lens and retinal tissues arise in *Xenopus* when Pax6 is ectopically expressed (Chow et al., 1999). These studies suggest that the role of Pax6 in eye development is conserved among species and holds a key position in the ocular developmental pathway.

1.3 The Biochemistry of Pax Proteins

The conservation of sequence and function of Pax gene homologs among different subfamilies as well as different species suggests that the biochemistry of Pax proteins is conserved as well. There are nevertheless significant differences in primary structure that likely have significant functional implications.

1.3.1 Paired Domain Biochemistry

1.3.1.1 DNA Binding Activity of the Paired domain

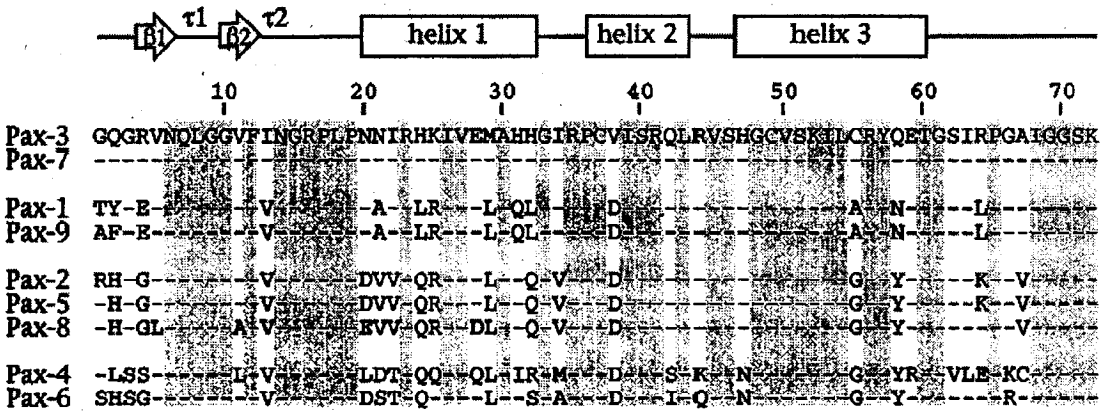
The PD is the structural element that defines the Pax family of proteins and is also the element that is the most highly conserved. This is evident by the fact that the PD of the mouse and human Pax3 homologs share a 100% sequence identity (Hoth et al.,1993; Walther et al., 1991). Also, the sequence identity of the PD of Pax proteins within each subfamily is typically 90% or more (Walther et al., 1991). At least 55 of the 128 residues that constitute

the PD are accepted as invariant among all PDs across different subfamilies as well as different species (Walther et al., 1991) (Figure 2).

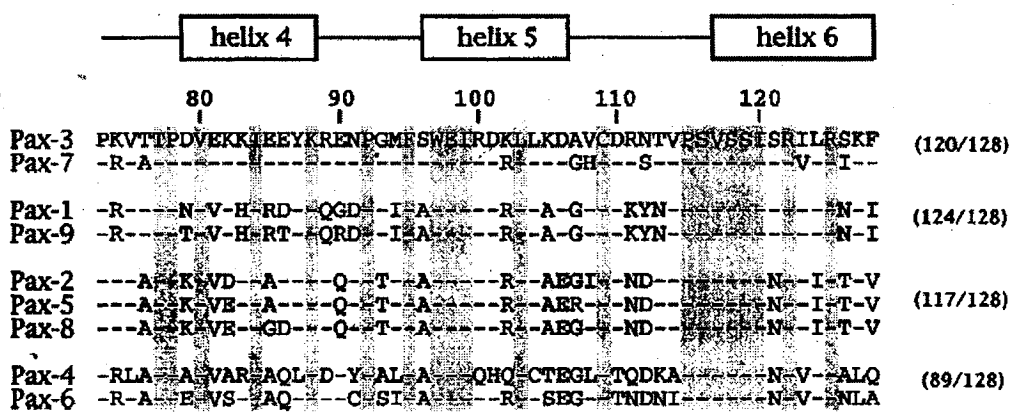
Figure 2

Multiple sequence alignment of the PDs of mouse Pax proteins. Only the entire sequence of Pax3 is shown and is aligned with the PD of the rest of the eight Pax genes grouped into individual subfamilies (Adams et al., 1992; Balling et al., 1988; Dressler et al., 1990; Goulding et al., 1991; Jostes et al., 1990; Matsushita et al., 1998; Plachov et al., 1990; Wallin et al., 1993; Walther and Gruss, 1991). Invariant residues are shaded in gray while identical positions are in dashes. Following the alignment, the number of identical residues among subfamily members is shown. With the use of the Prd and Pax6 crystal structures, a schematic representation of the PD was formulated. This representation shows the secondary structure of the PD and is displayed above the sequence alignment (Xu et al., 1999; Xu et al., 1995). (τ, β-turn, β, β-strand).

N-terminal subdomain



C-terminal subdomain



Sequence analysis of the promoter of the *Drosophila even-skipped* gene (e5) revealed the presence of a HD specific DNA sequence element. Mutations in the Prd protein that inactivated HD DNA binding activity did not abrogate Prd affinity for the e5 sequence thereby suggesting the presence of another DNA binding element within the Prd protein (Treisman et al., 1989). Subsequent deletion studies of the Prd protein permitted the discovery of the PD as a DNA binding domain and revealed that the e5 sequence held both HD and PD specific DNA binding elements that are adjacent to one another (Treisman et al., 1991). The DNA binding capacity of the PD was also supported by studies of the Pax1 protein. This protein contains only a PD with no HD, and was found to bind weakly to the e5 sequence and strongly to e5 sequence derivatives (Chalepakis et al., 1991). Interference studies and analysis of DNA sequences with Pax1 affinity were used to conclude that the PD recognizes a DNA sequence that extends 24 base pairs (Chalepakis et al., 1991). Sequence alignment and footprinting studies with the in vivo and in vitro derived DNA binding sequences that have affinity to Pax2,5,6 or 8, reveals that the PD binds a relatively degenerate consensus sequence (Adams et al., 1992; Barberis et al ., 1989; Epstein et al., 1994; Kozmik et al., 1992; Zannini et al., 1992). Structure-function studies with the PD of Pax5 resulted in the formation of the bipartite structure model of the PD, which suggests that it contains 2 subdomains: the PAI or N-terminal and the RED or C-terminal subdomains. According to the model each subdomain interacts specifically with different DNA sequences that are relatively degenerate. A truncated Pax5 protein that lacks the last 36 amino acids of the PD is only able to recognize DNA via the use of the PAI subdomain. The PD DNA sequences recognized only through the use of the N-terminal subdomain are termed type II sequences. Binding to Type I sequences requires the

use of both the PAI and RED subdomains of the PD and this is supported by the observation that the truncated Pax5 protein is unable to bind Type I DNA sequences (Czerny et al., 1993).

Figure 3

PD recognition sequences.

A) H2A-2.2 (Czerny et al., 1993), CD19-2A (Czerny et al., 1993), Nf3' (Epstein et al., 1995) and P6CON (Epstein et al., 1994), are some type I DNA sequences that contain binding sites for both the PAI and RED subdomains which are underlined. The DNA sequences e5 (Treisman et al., 1991), PRS4 (Chalepakis et al., 1991) and H2A2.1 (Czerny et al., 1993), are some type II DNA sequences that hold only N-terminal subdomain specific binding sites. Comparison of these sequences permitted the derivation of a consensus sequence that is displayed above the multiple sequence alignment along with a schematic representation of the bipartite structure of the PD. Bases in type I or II sequences that have diverged from the consensus sequence are displayed in italics.

B) The consensus binding sequences of the PD of Prd (Jun and Desplan, 1996), Pax3 (Chalepakis and Gruss, 1995; Epstein et al., 1996), Pax6 (Epstein et al., 1994), Pax4 (Fujitani et al., 1999), Pax2 (Epstein et al., 1994) and Pax8 (Jun and Desplan, 1996).

A)

	Consensus	TG/CC GT CA/TC CGC . . CA . TG . . C/T
Type I	H2A-2.2	ACAGTCGTCACCCACCGCTGCGTCACAA
	CD19-2/A	GCGGTGGTCACGCCTCAGTGCCCCATTC
	Nf3'	GTGTGTGTCACGCTTATTTTCCTGTACT
	P6CON	GGAATTTTCACGCTTGAGTTCACAGCTC
Type II	e5	GATTAGCACCGTTCCGCTCAGGCTGTCCCGT
	PRS4	TGGGCTCACCGTTCCGCTCAGGCTGTCCCGT
	H2A-2.1	TTTCTTGACACGCTGAAGAAATAGTTG

B)

Prd	T/C CGTCACG C/G TT GA/CG C
Pax-3	TCGTCACG C/C TT-A
Pax-6	A–TTCACGC A/T T G/C A-T TC/GA -C
Pax-4	G--G/T T C/G A T/A GCGTG-GGG/TCA -T-
Pax-2	T-GTCA C/T GC A/G TGA
Pax-8	A/C -GTCAC G/C C-T/C -A C/G T G/A

In accordance with the model Type II sequences are only half the length (10 to 12 base pairs) of those recognized by both subdomains (24 base pairs) (Figure 3).

DNA binding studies of alternatively spliced isoforms of Pax6 (5a isoform), 8 (S insertion isoform) and 3 (Q insertion isoform), that create insertions within the PD which alter the DNA binding activity of the PAI and RED subdomains, further support the bipartite model of the PD (Kozmik et al., 1997; Epstein et al., 1994; Vogan et al., 1996). The Pax6 isoform contains a 14 amino acid insertion (Pax6-5a) and the Pax8 isoform holds a one amino acid insertion (Pax8-S) in the PD (Glaser et al., 1992; Kozmik et al., 1997; Puschel et al., 1992; Walther et al., 1991). These insertions prevent the PAI subdomain of PAX6 from binding DNA and permits the PD of these insertion isoforms to bind a new sequence termed 5aCON exclusively through the use of the RED subdomain. This is in contrast to the wildtype PD which binds DNA with the use of the PAI subdomain alone, or both the PAI and RED subdomains, but never with the wildtype RED subdomain alone (Czerny et al., 1993; Epstein et al., 1994; Kozmik et al., 1997). These isoforms confirm the DNA binding capacity of the RED subdomain and support the bipartite model of the PD. The Pax3 isoform (Pax3/Q+) shows a glutamine insertion in the linker between the PAI and RED subdomain (Vogan et al., 1996). DNA binding studies of sequences that bind the Pax/Q- and Pax3/Q+ isoforms reveal that both isoforms bind type II sequences through the PAI subdomain with equal affinity but that the Q+ isoform binds type I sequences with 2 to 5 fold reduced affinity (Vogan et al., 1996). Therefore the presence of the extra glutamine residue interferes with DNA binding activity of the RED subdomain. Hence the relative contribution of the PAI and RED subdomain to DNA binding is modulated by these PD splice variants.

1.3.1.2 Paired Domain Structure

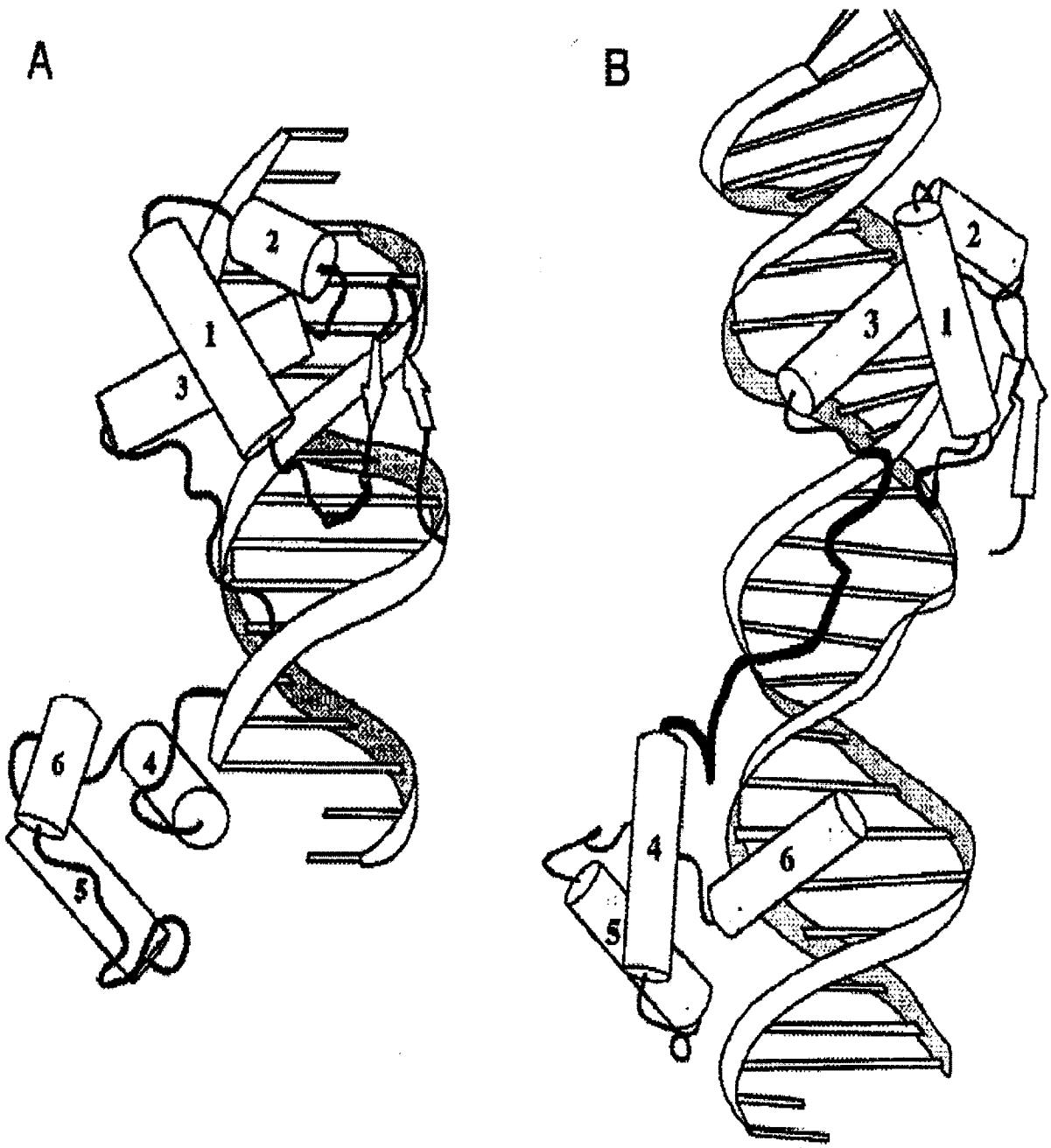
The crystal structures of the DNA bound PD of the Prd and Pax6 proteins confirm the bipartite model (Xu et al., 1999; Xu et al., 1995) (Figure 4).

Figure 4

DNA bound Structures of the PD of Prd and Pax6.

A) The *Drosophila* Prd DNA bound PD structure determined via X-ray crystallography. The DNA sequence used is 15 base pairs long and is a type II sequence (Xu et al., 1995).

B) The mammalian Pax6 DNA bound PD structure resolved via X-ray crystallography. The DNA sequence used is 24 base pairs in length and is a type I sequence (Xu et al., 1999). Each subdomain consists of three helical folds with the two most C-terminal helices forming a helix-turn-helix motif. The six helices are indicated with numbers 1 through 6. For both the Prd and Pax6 PDs the PAI subdomain engages in DNA contacts with the use of the β-hairpin motif as well as its most C-terminal helix. The linker is not simply a tether between the subdomains but adopts an extended conformation on Type I sequences and makes base specific contacts in the minor groove. Unlike the Pax6 PD, the Prd RED subdomain does not contact DNA. The RED subdomain of Pax6 binds DNA and makes base specific contacts in the major groove with its most C-terminal helix.


A
1
2
3
4
5
6
B
1
2
3
4
5
6

The PD consists of two globular units, PAI and RED, connected by a short linker sequence. The PAI subdomain consists of a three helical fold, very similar to the three helical fold of HDs, with the two most C-terminal helices forming a helix-turn-helix motif. Helices 1 and 2 pack together in an anti-parallel manner and are perpendicular to helix 3. Helix 2 makes non-specific phosphate backbone contacts while helix 3 is involved in making base specific contacts with the major groove of DNA. Preceding the three helical fold lies a β-hairpin and a β-turn. The β-hairpin consists of two anti-parallel β-strands joined by a type I β-turn. This structure is involved in making non-specific phosphate-backbone DNA contacts. The β-turn that follows the β-hairpin structure is a type II turn and unlike most other DNA binding proteins participates in base-specific contacts in the minor groove of DNA. The PAI and RED subdomains are structurally independent. The type II β-turn and portions of the helix-turn-helix motif pack against the linker via hydrophobic interactions. The sub-linker amino acid sequence is not simply a tether that joins the PAI and RED subdomain, but adopts an extended conformation on DNA and participates in base specific contacts via the minor groove (Xu et al., 1995; Xu et al., 1999).

The Pax6 PD is bound to a longer 26 base pair type I DNA sequence and has a PAI subdomain that is very similar to that of the Prd PAI subdomain. Unlike the Prd RED subdomain, the Pax6 RED subdomain contacts DNA with the use of its most C-terminal helix, which makes base-specific contacts in the major groove. The lack of DNA contact formation of the RED subdomain in the Prd crystal structure is in agreement with the biochemical data which suggests that Prd binds DNA exclusively through the use of the PAI subdomain (Xu et al., 1995). Another difference between the Prd and Pax6 PD structures is the specificity observed for the base pair at position 4 of DNA. The first residue of helix 3 of the PAI subdomain contacts this base pair. In the Prd protein, as well as most other proteins

including Pax3, a G/C base pair is preferred at this position, while for the Pax6 protein an A/T base pair lies at this position. This difference in specificity is due to the identity of the first residue of helix 3. In Prd and other Pax proteins this residue is a histidine, which is able to participate in hydrogen bonds with the GC base pair. For Pax6, this residue is instead an asparagine that is able to recognize the AT base pair via hydrophobic coupling. The methylene group of aspargine undergoes hydrophobic interaction with the methyl group of thymine which is reinforced by a water mediated hydrogen bond between the amide in the side chain of the asparagine residue and the sugar-phosphate backbone (Xu et al., 1995; Xu et al., 1999).

Despite the resolution of the PD 3D structure, accounting for the degeneration observed in the DNA sequences recognized by the PD is not possible. In addition, these X-ray crystallography studies did not address the possibility of other proteins binding DNA alongside the PD and influencing PD DNA binding specificity.

1.3.2 Paired-type Homeodomain Biochemistry

1.3.2.1 DNA Binding Activity of the Homeodomain

The superfamily of HDs includes the paired-type HD family which, like most HDs consists of 60 residues and displays DNA binding activity. Paired-type HDs can be found in mammalian Pax3, 7, 4 and 6 proteins as well as the *Drosophila* Prd, Gooseberry and Eyeless proteins (Goulding et al., 1991; Jostes et al., 1990; Walther and Gruss, 1991; Walther et al., 1991). Some proteins also contain paired-type HDs without the presence of a PD (e.g. Phox protein). Most members of the HD superfamily of proteins recognize DNA elements that contain a TAAT core motif (Gehring et al., 1994). There is some specificity for the 2 base pairs found immediately 3' to the core motif and the identity of residue 50 of the HD is a major determinant for this specificity (Treisman et al., 1989). The feature unique to the

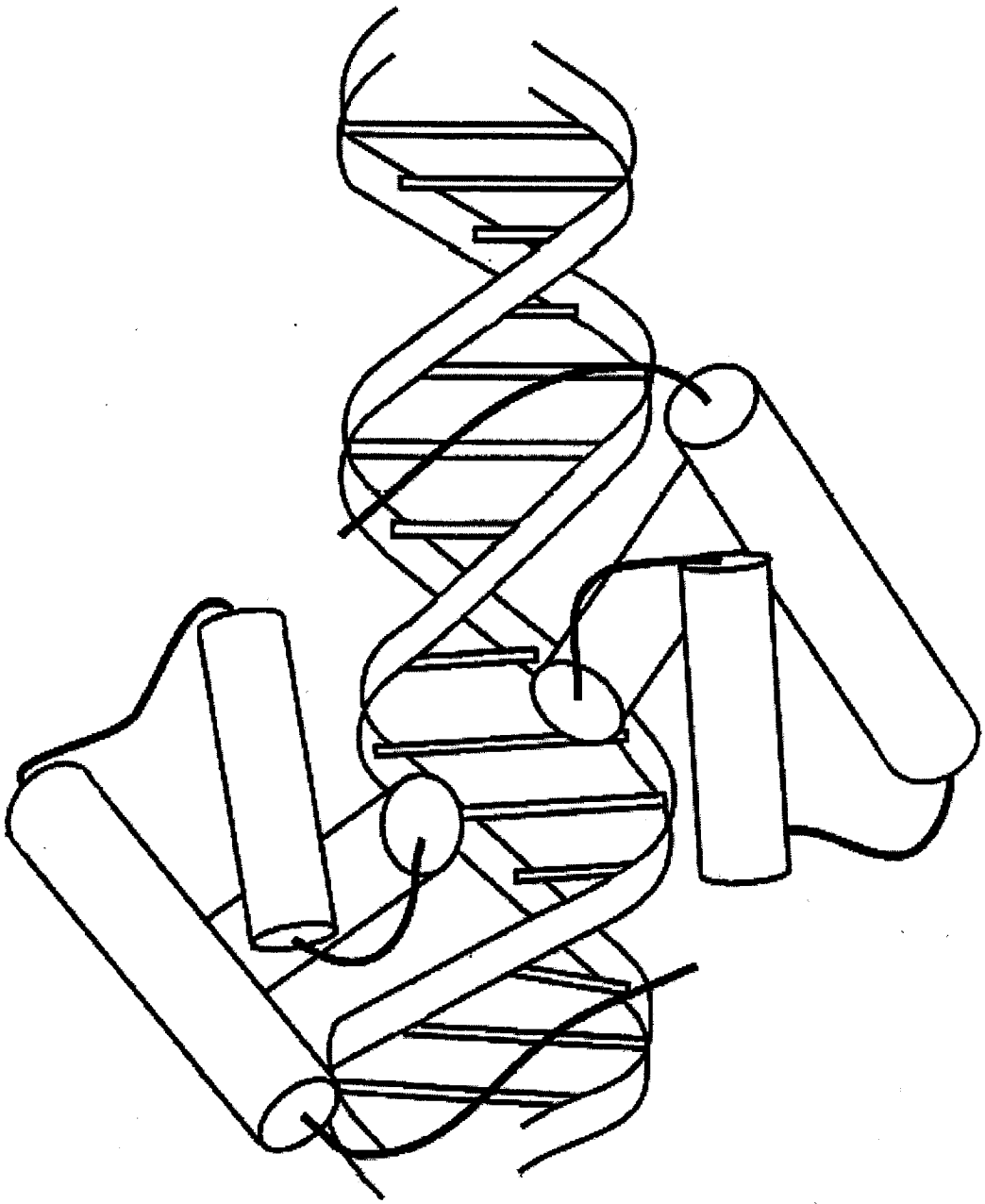
paired-type family of HDs is their ability to form cooperative homo and heterodimers on palindromic sites of the type 5'-TAAT $(N)_{2-3}$ ATTA-3'. Non-paired-type HDs may dimerize on DNA but require domains extrinsic to the HD. Paired-type HDs rely entirely on the 60 residues within the HD for their dimerization potential. The length half site spacer and the extent of cooperation observed during dimerization is determined by the identity of residue 50 (Wilson et al., 1993). Paired-type HDs which hold a glutamine at position 50 (Q50) preferentially dimerize on a palindromic DNA site that contains a 3 base pair spacer (P3 probes), while HDs with serine at this position (S50) dimerize more efficiently on DNA with a 2 base pair spacer (P2 probes). For both S50 and Q50 paired-type HDs, binding of the first HD to DNA increases the affinity of the second HD for the remaining DNA TAAT sequence by 50 fold for S50 HDs and over 100 fold for Q50 HDs (Wilson et al., 1993).

1.3.2.2 Paired-type Homeodomain Structure

Crystal structures of many HDs have been resolved but the one of particular interest for our studies and that serves as a model for cooperative DNA binding of paired-type HDs, is the crystal structure of the paired-type HD of the *Drosophila* Prd protein bound to the P3 site (Wilson et al., 1995) (Figure 5). Of note is that the amino acid identity at position 50 has been altered from the native serine to glutamine. This substitution was done in order to optimize the cooperativity of the two Prd HDs complex to the P3 DNA sequence (Wilson et al., 1993; Wilson et al., 1995).

Figure 5

Schematic representation of the HD dimer bound to DNA. This structure was obtained by X-ray crystallography of the *Drosophila* Prd paired-type HD bound to the P3 DNA sequence (Wilson et al., 1995). The base specific contacts made by the HD involve the third helix (recognition helix), which binds DNA via the major groove, and the N-terminal arm, which contacts the minor groove of DNA. Dimerization on DNA requires the formation of reciprocal contacts between the N-terminus of helix 2 of one HD with the N-terminal arm of the other domain. Dimerization is also supported by water-mediated hydrogen bonds and hydrophobic contacts made between N-termini of the two recognition helices (Wilson et al., 1995).



When bound to the P3 DNA sequence the Prd paired-type HD adopts a three helical fold structure very similar to the structure of the HDs of engrailed and antennapedia (Billeter et al., 1993; Kissinger et al., 1990; Qian et al., 1989; Wilson et al., 1995). The two most C-terminal helices form a helix-turn-helix motif with helix 2 anchoring the motif over DNA by making phosphate backbone contacts and helix3 participating in specific DNA contacts by contacting bases in the major groove of DNA. The first and second helices are anti-parallel and perpendicular to helix3. The third helix is able to make base specific contacts with the use of three critical invariant residues: V47, N51 and Q/S50. Residue V47 participates in hydrophobic coupling to the methyl group of a thymine base while N51 makes both direct and water mediated hydrogen bonds with both adenine and thymine bases of DNA. Position 50 is involved in water-mediated hydrogen bonds with a thymine base of DNA as well as a base immediately 3' to the TAAT core motif (Wilson et al., 1995). Preceding the three helical fold lies the N-terminal arm. The N-terminal arm adopts an extended conformation on DNA and is able to form specific contacts with DNA by contacting bases in the minor groove of DNA (Wilson et al., 1995).

The two HDs adopt a head-to-head conformation on DNA. Dimerization on DNA requires the formation of reciprocal contacts between the N-terminus of helix 2 of one HD with the N-terminal arm of the other domain. Dimerization is also supported by water-mediated hydrogen bonds and hydrophobic contacts made between N-termini of the two recognition helices (Wilson et al., 1995). These reciprocal contacts occur on the interaction surfaces between the HDs and are possible only if the DNA helix is distorted from the ideal B form by bending approximately 21 degrees (Wilson et al., 1995). The 3D structure of the HD on P2 is postulated to be similar to the dimer formed on P3 although slight differences would be required in order to accommodate for the missing base pair in the spacer that may bring the two HDs closer together (Wilson et al., 1993; Wilson et al., 1995). Even though the Prd HD

crystal structure explains the dimerization potential of paired-type HDs, it is not able to account for the different half-site preferences of S50 and Q50 HDs.

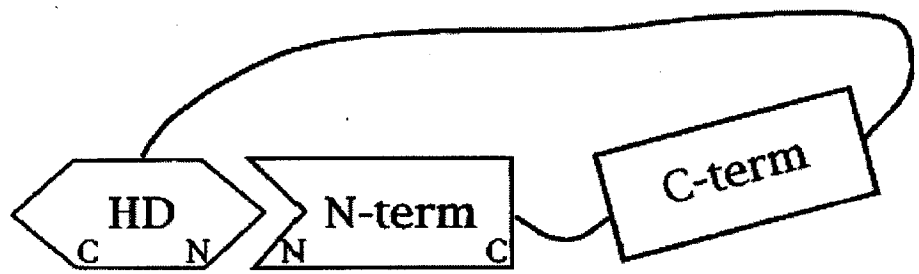
1.3.3 Functional Dependence between the Paired and Homeodomains

Many observations and studies indicate that the paired and HD are structurally and functionally independent. For instance, the DNA bound complexes of both the PD and paired-type HD indicate the domains have different DNA binding specificity. Also, many naturally occurring proteins contain a PD and no HD (e.g. Pax1) while there are other proteins that contain a paired-type HD and no PD (e.g. Phox protein) suggesting that these domains can function as autonomous DNA binding domains. Despite this, many studies indicate that when present in the same polypeptide chain the paired and paired-type HD can cooperate for DNA binding and function interdependently.

Lack of functional independence between the PD and the HD was first observed with studies of the Prd protein and segments of the even-skipped promoter sequence (Treisman et al., 1991). The *even-skipped* promoter sequence contains two elements that can bind the Prd protein. The e5 element consists of both a PD and HD recognition sites that overlap and that can be bound independently by the PD or the HD. Another element, e4, has paired and HD specific DNA sites that lie adjacent to one another. Recognition of e4 by the Prd protein requires binding of both domains simultaneously (Treisman et al., 1989). Other studies have also indicated that binding of the domains is cooperative. For instance, in vitro selection for DNA oligos with high affinity for the Prd protein from a pool of random oligonucleotides has led to the discovery of the PH0 probe. The PH0 probe has a very high affinity for the Prd protein and contains both a PD and a HD recognition sequence and they are placed in juxtaposed fashion and have an inverted relative orientation (Figure 6) (Jun and Desplan, 1996).

Figure 6

Schematic representation of the Prd protein bound to the PH0 and PTE sites that contain adjacent PD and HD recognition sites placed in an inverted relative orientation. The PTE site (Fujioka et al., 1996) from the *even-skipped* promoter and the in vitro selected PH0 (Jun and Desplan, 1996) sequences are aligned.



PHO GATCTCTTCA*TTA*GTCACGCTTGAGTG

PTE TTGATGGGCA*TTA*GCCTCGGTGAGTGG

Studies in which the PD and HD are singly mutated or the spacing or orientation of the two sites are changed in the PH0 sequence have shown that all these alterations decrease the affinity of the Prd protein for the probe by almost a hundred fold. This would indicate that the PD and HD cooperate when binding the PH0 sequence (Jun and Desplan, 1996). These observations are also made when untethered PD and HDs of Prd were used during the DNA binding assays, although tethering was observed to contribute slightly to the cooperativity between the PD and HD (Jun and Desplan, 1996). Importantly a PH0-like sequence was identified in the promoter region of the even-skipped gene, called PTE, that also contains PD and HD juxtaposed sites placed in an inverted fashion, which has been shown to regulate even-skipped gene expression by Prd (Fujioka et al, 1996). Hence, these studies suggest that the PD and HD can bind DNA simultaneously and with cooperativity and that this mode of DNA binding has biological relevance.

Lack of phenotypic rescue by transgenic expression of either the PD or the HD of Prd individually in Prd-/- mutants indicates that both domains are required for proper Prd function (Bertuccioli et al., 1996; Miskiewicz et al., 1996). Also, rescue attempts of the Prd-/- fly with Prd proteins with mutations only in the PD or only in the HD both failed (Bertuccioli et al., 1996; Miskiewicz et al., 1996).

Similar observations have been made with the Pax3 protein; mutations in the PD disrupt PD as well as HD function and mutations in the HD abrogate HD and PD DNA binding activity. The Sp-delayed mouse mutant allele contains a G42R mutation that abrogates PD DNA binding activity but also decreases the Pax3 HD affinity for the P2 probe significantly (Underhill et al., 1995). The DNA binding abilities of many PAX3 missense mutations that are associated with the human Waadenburg syndrome have also supported the functional dependence hypothesis of the PD and HD. Many of the missense mutations map to the PAI (not RED) subdomain and abrogate the PD and affect the HD DNA binding activity

as well. WS PD mutations studied thus far compromise both PD and HD DNA binding activity (Fortin et al., 1997). Also, it was determined that the R53G Waardenburg Syndrome mutation in the HD abrogates both the HD as well as the PD DNA binding activity (Fortin et al., 1997).

Two specific naturally occurring mutations in the Pax6 PD have been shown to severely affect the DNA binding activity of the HD. Interestingly the I87R mutation in the RED subdomain abrogates PD and HD DNA binding activity while the R26G PAI subdomain mutant has no PD DNA binding activity but has an elevated HD DNA binding and transactivation ability in comparison to PAX6 (Singh et al., 2000). This PAX6 study reveals that the two subdomains of the PD influence the function of the HD differently.

In addition to influencing the ability of the HD to bind DNA, the PD has been shown to influence HD DNA sequence specificity. Full-length constructs of Pax3 with a PD and HD are able to dimerize only on P2 sequences but when the Pax3 HD is expressed alone, dimerization is observed on both P2 as well as the P3 sequences (Underhill and Gros, 1997). Hence, the presence of the PD in Pax3 prevents dimerization of the HD on the P3 probe and this effect can be transferred onto a heterologous HD. The paired-type HD of the Phox protein (Q50 HD) is able to dimerize on the P3 probe and monomerize on the P2 probe. Once the Phox HD is covalently linked to the PD of Pax3 it loses both P2 and P3 binding at both the monomer and dimer level (Underhill and Gros, 1997; Fortin et al., 1998). Overall, these findings indicate that the paired and HD functionally interact and influence each other's recognition of DNA.

1.3.4 Biochemistry of the Octapeptide

All members of the Pax family except Pax4 and Pax6 (as well as the *Drosophila* Prd) contain the octapeptide structural motif that is located in the linker sequence between the PD

and the HD. The exact position of the octapeptide within the linker is variable among different members of the Pax family. As with the PD, the octapeptide sequence identity is higher for members within the same subfamilies and the generally accepted consensus sequence for the octapeptide is (H/Y)S(I/V)(N/S)G(I/L)LG (Noll et al., 1993). Little is known about the biochemistry of the octapeptide but studies with the Pax2/5/8 subfamily suggest that it may have a role in transcription regulation by the Pax protein. More specifically, the deletion of the octapeptide in this subfamily resulted in an increase in transcriptional activation of Pax target genes by 2 to 3 fold as compared to the unaltered Pax protein (Eberhard et al., 2000; Lechner and Dressler, 1996). This possible transcription repression effect of the octapeptide may involve Pax association with co-repressor molecules of the Groucho family of proteins (Eberhard et al., 2000, Koop et al., 1996; Stifani et al., 1992). Studies with Pax5 and the Groucho family member protein Grg4 show that the proteins interact and that this interaction results in transcription inhibition by the Pax5 C-terminal transactivation domain. This repression requires interaction of Pax5 and Grg4 at two distinct locations in Pax5, one of them being the octapeptide (Eberhard et al., 2000). Therefore the octapeptide may serve a role in protein-protein interaction. Further studies are required to elucidate the exact role and mechanism of the octapeptide motif.

1.3.5 Transcription Regulation by Pax Proteins

The expression patterns of Pax proteins in the embryo as well as their homology to other known regulators of development suggest that Pax genes play a role in development (Noll, 1993; Chalepakis et al., 1993). The discovery of nuclear localization signals in the protein sequence as well as the presence of DNA binding domains supported the role of Pax proteins as transcription factors (Bopp et al., 1989; Chalepakis et al., 1991; Dressler and Douglas, 1992; Krauss et al., 1992; Walther et al., 1991). Transactivation studies

of constructs with Pax binding sites upstream of reporter genes have shown that Pax proteins activate transcription (Chalepakis et al., 1991; Chalepakis et al., 1994; Fujitani et al., 1999; Glaser et al., 1994; Kozmik et al., 1992; Lechner and Dressler, 1996; Schafer et al., 1994; Kozmik et al., 1993; Nornes et al., 1996). Studies done with Pax3, 6 and 9 proteins show a dose-dependent bell-shape transactivation response. At low concentrations these transcription factors activate transcription but as the concentration increases transcriptional activation does not plateau but actually begins to decrease (Chalepakis et al., 1994; Czerny and Busslinger, 1995; Nornes et al., 1996; Vogan et al., 1996). Also, it has been shown that the transcription of certain target genes of Pax2, 4 and 5 is actually repressed following exposure to these Pax proteins (Fujitani et al., 1999; Nutt et al., 1998; Schwarcz et al., 2000; Smith et al., 1999). These observations suggest that Pax proteins may have both transcription activation as well as repression capabilities.

1.3.5.1 Repression and Activation Domains of Pax Proteins

Studies have localized the transactivation domain to the C-termini of Pax proteins following both the PD and the HD. The homology of the DNA binding modules among Pax proteins is much higher than the homology found for the C-terminal transactivation domain (Walther et al., 1991). Despite this, the homology of the C-terminal of Pax proteins is sufficiently high to suggest that it holds a conserved function (Adams et al., 1992; Czerny and Busslinger, 1995; Dorfler and Busslinger, 1996; Glaser et al., 1992; Kozmik et al., 1993; Vorobyov et al., 1997). The C-terminal portion of Pax proteins does not share any homology with transactivation domains of other transcription factors but sequence analysis reveals that it is rich in proline, serine and threonine residues, a characteristic that is typical of transactivation domains (Chalepakis et al., 1991; Chalepakis et al., 1993). Deletion studies with Pax3/Pax7 (Schafer et al., 1994; Chalepakis et al., 1994; Sheng et al., 1997), Pax6/Pax4

(Fujitani et al., 1999; Glaser et al., 1994; Smith et al., 1999), Pax2/Pax5/Pax8 (Dorfler and Busslinger, 1996; Lechner and Dressler, 1996) and Pax9 (Nornes et al., 1996) have shown that Pax dependent transactivation requires the C-termini of the proteins, supporting the role of the C-terminus as a transcription activation domain. Many of these same studies have also demonstrated the existence of inhibitory domains in Pax proteins that repress transcription of target genes. For instance, Pax2, 4, 5, 8 and 9 have C-terminal domains that contain adjacent inhibition and activation sequences (Dorfler and Busslinger, 1996; Fujitani et al., 1999; Nornes et al., 1996). As mentioned earlier, the octapeptide motifs of Pax2, 5 and 8 possess the ability to repress transcription (Eberhard et al., 2000; Lechner and Dressler, 1996), and the first 90 amino acids of Pax3 contain transcription inhibitory sequences (Chalepakis et al., 1994). It was observed that deletion of these inhibitory sequences leads to a noticeable rise in transactivation of Pax target genes. As was the case with the Pax5/Grg4 study, the inhibitory domains of Pax proteins generally do not function independently thereby suggesting that repression of transcription by Pax proteins requires the use of Pax-specific negative regulators (Dorfler and Busslinger, 1996; Eberhard et al., 2000). This discovery implies that the transcriptional effect of Pax proteins may be regulated by variations in the combinatorial and relative activity of transcription activating and transcription inhibiting domains.

1.3.5.2 Regulation of Pax Protein Transactivation Activity

The exact molecular mechanism of the regulation of transcription by Pax proteins is still largely unknown but many factors have been identified that play a role in modulating the ability of Pax proteins to activate transcription. For instance, Pax 3, 5, 6 and 9 transcription activation abilities have been shown to be affected by the concentration of the protein (Chalepakis et al., 1994; Czerny and Busslinger, 1995; Vogan et al., 1996; Wallin et al., 1998). As mentioned earlier the transcriptional response of the Pax3, 6 and 9 proteins is dose

dependent (Chalepakis et al., 1994; Czerny and Busslinger, 1995; Nornes et al., 1996; Vogan et al., 1996). In the case of Pax5, activating and inhibiting DNA sequence elements of certain promoters of target genes compete for Pax5 binding and consequently modulate the Pax5 induced expression of the downstream gene. At low concentrations of Pax5 only high affinity activating DNA elements are bound by Pax5 and lead to optimal transcription of the downstream gene. At high Pax5 levels both high affinity activating sequences as well as low affinity inhibiting sequences are bound by Pax5 leading to less transcription of the target gene by the Pax5 protein (Wallin et al., 1998).

Another factor that influences the transcription activation abilities of Pax proteins is alternative splicing. For example, the Q+ and Q- Pax3 isoforms have different DNA binding affinity for type I sequences and transcription activation of genes downstream of type I sequences differs between the two isoforms (Vogan et al., 1996). Similarly, different Pax6 isoforms have different DNA binding specificity and transactivation activities (Epstein et al., 1994). Pax6 isoforms induce the transcription of reporter genes to similar extents when the gene is fused upstream to a promoter sequence that contains either RED domain specific sequences (5aCON) or PAI and RED specific Type I sequences (P6CON). The Pax(5a) isoform has DNA binding affinity for the 5aCON sequence only and is actually more efficient than Pax6 at inducing transcription from 5aCON DNA sequence containing promoters (Epstein et al., 1994). It has been observed that the Pax6(5a) isoform can squelch the transactivation ability of Pax6, implying that the relative concentration of these isoforms is another factor that modulates transcription activation of Pax6 specific target genes (Czerny and Busslinger, 1995; Epstein et al., 1994). This possibility is validated by the fact that mutations in Pax6 that change the ratio of these two isoforms result in a unique phenotype (Epstein et al., 1994). Unlike Pax3 and 6, isoforms of Pax8 and 9 have been shown to have similar DNA binding specificity, but like Pax3 and Pax6 isoforms they differ in their ability to

activate downstream genes (Kozmik et al., 1993; Nornes et al., 1996). This differential transactivation ability of Pax8 and 9 isoforms is due to the alternative splicing (shifts in reading frame and exon skipping) that results in different C-terminal sequences (Kozmik et al., 1993; Nornes et al., 1996). Altogether, these studies suggest that alternative splicing can create functional diversity and can also lead to the modulation of target gene transcription.

1.3.6 Interactions of Pax Proteins with other Transcription Factors

Many Pax interacting proteins have been identified with the use of yeast-two-hybrid screening and have been characterized as transcriptional co-repressors of Pax target genes (Eberhard et al., 2000; Hollenbach et al., 1999; Magnaghi et al., 1998). As was mentioned earlier, a known interacting protein of Pax5 that acts as a co-repressor of transcription is the Groucho Grg4 protein. Other members of the Groucho family can also repress Pax5 mediated transactivation (Eberhard et al., 2000). Groucho proteins are known co-repressors of transcription that are able to bind to the N-terminal tails of histones and interact with histone deacetylases, presumably to activate deacetylation of histones and encourage DNA packing and condensation which makes DNA less available for transcription (Chen et al., 1999; Choi et al., 1999; Fischer and Caudy, 1998; Parkhurst, 1998). The Pax5/Grg4 complex is a result of the interaction of the transactivation domain and octapeptide motif of Pax5 with the glutamine rich and serine-proline rich domains of Grg4 (Eberhard et al., 2000). Interaction of Grg4 and Groucho with Pax2 and Pax8 also leads to transcription inhibition (Eberhard et al., 2000).

Pax3 is also known to interact with co-repressors. In vitro studies have shown that Pax3 interacts with HIRA, a homolog of the Saccharomyces cerevisiae co-repressor proteins Hir1p and Hir2p, that are involved in regulating the cell-cycle dependent transcription of histone genes (Magnaghi et al., 1998; Spector et al., 1997). This interaction involves the HD of Pax3 and the 2 contiguous domains in the C-terminus of HIRA (Magnaghi et al., 1998).

The suspected role of HIRA as a co-repressor of Pax3 target genes remains to be shown but evidence does exist that these proteins may interact in vivo. HIRA and Pax3 are both expressed in neural crest cell explants and have overlapping expression in the developing neuroepithelium and its derived neural crest (Magnaghi et al., 1998). Another protein detected via yeast-two-hybrid screening with Pax3 as bait is hDaxx (Hollenbach et al., 1999). The interaction requires the proline-serine-threonine rich C-terminal sequence of hDaxx with the inter-domain linker (including the octapeptide) and HD of Pax3 (Hollenbach et al., 1999). Many reasons suggest that hDaxx may behave as a co-repressor of Pax3 target genes. The hDaxx protein is a nuclear protein, like Pax3, and has 2 short regions that share homology with the co-repressor Sin3 protein (Hollenbach et al., 1999; Kiriakidou et al., 1997). The fusion protein that consists of the hDaxx protein linked to the DNA binding domain of GAL4, represses transcription of a constitutively active gene when the protein is DNA bound. This supports a role of hDaxx as a repressor of transcription (Hollenbach et al., 1999). Also, incubation of Pax3 protein with increasing amounts of hDaxx leads to progressive inhibition of Pax3 mediated transcriptional activation (Hollenbach et al., 1999).

Pax proteins may also interact with co-activators to enhance transcription. Pro-glucagon gene expression in pancreatic α cells can be induced by promoter binding of Pax6 or cdx2 (a HD protein) individually (Hussain and Habener, 1999). When both Pax6 and cdx2 are expressed in this cell type, a heterodimer forms that synergistically activates the pro-glucagon gene (Anderson et al., 1999; Hussain and Habener, 1999; Ritz-Laser et al., 1999). This synergism is further enhanced when the dimer forms a complex on DNA with the co-activator p300 protein (Hussain and Habener, 1999). Pax6 and cdx2 can interact individually with p300, but Pax6/p300 complex formation is enhanced when cdx2 is present (Hussain and Habener, 1999). Recently, the PDZ-binding motif containing TAZ protein was identified in a yeast two-hybrid screen with Pax3 as bait (Murakami et al., 2006). TAZ and Pax3 expression

partially overlap in the embryo and in vitro assays have indicated that TAZ can enhance the transcriptional activity of Pax3 on the Mitf promoter (Murakami et al., 2006). These results suggest that TAZ is a bona fide co-activator that promotes expression of Pax3 target genes.

Interaction of transcription factors with components of the basal transcription machinery is yet another mechanism used by transcription factors to modulate transcription efficiently. For example, Pax3, 5 and 6 interact with the TBP (TATA-binding protein) component of the basal transcription machinery (Cvekl et al., 1999; Eberhard and Busslinger, 1999). This interaction has been shown with GST pull-down assays and with co-immunoprecipitation of endogenous proteins in certain cell-lines and tissues (Cvekl et al., 1999; Eberhard and Busslinger, 1999). This interaction is mediated by the N-terminal portion of the HD suggesting that the partial HD in Pax5 is a protein-protein interaction motif (Eberhard and Busslinger, 1999). Other studies have indicated that many other HD containing proteins interact with TBP (Zhang et al., 1996; Cvelk et al., 1999; Eberhard and Busslinger, 1999; Um et al., 1995; Zwilling et al., 1994). The TBP shares sequence homology with the protein-protein interaction pocket domain of the pRB (retinoblastoma protein), that is required for pRB interaction with transcription factors. Consequently TBP and pRB frequently bind to the same transcription factors (Hagemeier et al., 1993; Hagemeier et al., 1993; Hateboer et al., 1993). Pax3 5 and 6 can bind the pocket domain of the activated underphosphorylated form of pRB (Cvekl et al., 1999; Eberhard and Busslinger, 1999; Wiggin et al., 1998). Reporter assays with Pax3 have shown that this interaction inhibits transcriptional activation (Wiggin et al., 1998). No clear role for these interactions in vivo has been determined as yet, we can only speculate that the interaction of Pax proteins with TBP and pRB may modulate the transactivation abilities of Pax proteins. Pax3 downregulation is a pre-requisite to terminal differentiation of neuronal progenitor cells (Reeves et al., 1998) and pRB is known to also promote terminal differentiation by inhibiting progression through the cell cycle (Weinberg,

1995). One may speculate that the pRB may promote neuronal differentiation by repressing the activity of Pax3, alternatively, Pax3 may inhibit pRB to prevent neural differentiation. It is interesting to note that the CD19 promoter contains no TATA box and therefore cannot recruit TBP on its own but instead has a Pax5 binding site in the -30 region (Kozmik et al., 1992). Also note that the transcription of the CD19 gene has been shown to have a strict requirement for Pax5 (Kozmik et al., 1992). Perhaps the role of the interaction between TBP and Pax5 is to recruit TBP to the promoter of CD19 in order to establish the basal transcriptional machinery.

Pax3 transactivation activity can also be regulated via functional antagonism. Pax3 and the HD containing Msx1 proteins are co-expressed in migrating myoblasts that have delaminated from the somite and are headed toward the limb bud in order to develop the limb musculature (Bendall et al., 1999). Expression of Pax3 in these cells commits them to the myogenic lineage and prolonged expression will eventually lead to terminal differentiation into myotubes by activating the expression of myogenic regulatory factors such as MyoD (Bendall et al., 1999). Msx1 expression is necessary for maintaining the committed myoblasts in an undifferentiated state during migration in order for the cells to reach and populate the limb bud and subsequently give rise to limb musculature (Bendall et al., 1999). When Pax3 binds the PD DNA binding site in the MyoD promoter sequence a mild activation of MyoD occurs (Bendall et al., 1999). Msx1 expression results in strong repression of MyoD expression. In vitro studies have shown that the PD of Pax3 and the N-terminal arm of the HD of Msx1 interact to form a Pax3:Msx1 complex and inhibit Pax3 DNA binding activity on the MyoD promoter sequence (Bendall et al., 1999). Therefore the Pax3 and Msx1 physical interaction mediates the functional antagonism of these proteins on MyoD expression and on terminal differentiation of myoblasts. Also, Pax5 and PU.1 can reciprocally inhibit their transactivation abilities when they physically interact on enhancer sequences of Pax5 or PU.1

target genes such as Igk (Maitra and Atchison, 2000). Altogether, these studies confirm that Pax proteins can use functional antagonism as a means to modulate their activity and that this antagonism is possible via intermolecular interactions of Pax proteins with other proteins. Knockdown experiments in Xenopus have shown that Msx1 and Pax3 are also both required for proper neural crest formation (Monsoro-Burq et al., 2005). Perhaps a similar functional antagonism mechanism is at play between Msx1 and Pax3 during the development of neural crest cell derivatives.

In contrast to functional antagonism there are examples in the literature of the formation of complexes between Pax and other proteins that encourages the function of Pax as an activator of transcription. For instance, Pax5 is known to cooperate with several different members of the Ets family to activate the promoter of the mb-1 gene (Fitzsimmons et al., 1996; Wheat et al., 1999). Examples of Ets proteins that can be recruited by Pax5 to the mb-1 promoter include the GABPα/β heterodimer, Fli-1 and c-Ets-1. The mb-1 gene is expressed in maturing B cells and encodes the Igα chain which is involved in intracellular signaling of the B cell receptor complex (Hagman et al., 2000). The mb-1 promoter holds a Pax5 recognition site adjacent to a sub-optimal Ets binding site. Pax5 is able to bind the promoter on its own but Ets proteins are unable to bind the sub-optimal DNA site unless they are recruited by Pax5 to form a ternary complex. Both the Pax5 and Ets DNA binding sites are required for ternary complex formation since mutation in either DNA binding site reduces mb-1 expression (Fitzsimmons et al., 1996). The Ets factors contain a winged helix-turn-helix DNA binding domain that consists of four α helices and a β-sheet formed by four anti-parallel β-strands. The third helix behaves as a recognition helix and makes base specific contacts in the major groove of DNA to recognize the core GGAA/T motif (Sharrocks et al., 1997). Recruitment by Pax5 requires a critical aspartic acid residue immediately C-terminal to the recognition helix of Ets (Fitzsimmons et al., 1996). Modeling studies and X-ray

crystallography studies have confirmed that the Pax5/Ets ternary complex on the composite site of mb-1 is mediated by a salt bridge formed between the glutamine residue in the β-hairpin with the critical aspartic acid of the Ets domain (Wheat et al., 1999; Garvie et al., 2001). Both the β-hairpin and helix 2 of the PD of Pax5 make contacts with Ets when recruitment occurs. This suggests a possible role for these structures in Pax proteins as protein-protein interaction segments. Thus studies of the Pax3/Msx1 and the Pax5/Ets complexes indicate that physical interactions with other proteins may lead to either functional antagonism or functional cooperation.

1.4 Target Genes of Pax Proteins

Both biochemical and genetic experiments have revealed that Pax proteins largely function as regulators of transcription of downstream targets at various times during development. The identification of target genes of Pax proteins has proven difficult but several have been discovered through phenotypic analysis of Pax mutants. The discovery of Pax target genes was facilitated once DNA recognition sequences were defined for both the PD and HD. Proteins with temporal and spatial expression patterns as well as mutant phenotypes similar to those seen with Pax proteins have been identified as possible target genes of the Pax family. In some cases, a direct regulatory relationship was discovered between Pax and candidate target genes by the identification of Pax responsive elements in the promoter or enhancer sequences. Further confirmation of Pax regulation of candidate target genes was possible with the observation that the expression of such genes was altered in mice deficient in corresponding Pax genes. This approach was used to determine many targets of Pax proteins, and includes genes that encode for cell-adhesion molecules, signal transduction molecules, transcription factors and protein products specific to certain terminally differentiated cells. Only an overview of Pax3 target genes will follow.

1.4.1 Target Genes and Biological Pathways of Pax3 in Neurogenesis

Few Pax3 target genes involved in neural tube formation have been discovered to date. The genes that are suspected to be transcriptionally regulated by Pax3 in the neural tube are not activated but inhibited by Pax3. One possible gene that may be regulated by Pax3 in the developing neural tube is Engrailed-1 (En-1), a homeobox containing transcription factor that shares homology with the *Drosophila* engrailed protein (Joyner et al., 1985; Davis et al., 1988). Like members of the Pax2/5/8 subfamily, En-1 plays a role in defining the mid-hindbrain boundary and is expressed in cells bordering this junction (Song et al., 1996). It has been observed that in the Pax3/Pax7 double knock out, En-1 expression is expanded dorsally in the neural tube thereby suggesting that Pax3 may inhibit En-1 transcription in the dorsal neural tube where Pax3 is expressed (Mansouri and Gruss, 1998). One can speculate that the mechanism of this inhibition may involve the binding of Pax3 to the promoter of En-1 which may disrupt the assembly and/or function of the transcriptional machinery. This possibility is supported by the fact that two putative Pax binding sites have been identified in the enhancer sequence of the En-1 gene (Song et al., 1996). Pax3 may regulate En-1 in order to promote proper dorsal-ventral patterning of the neural tube.

Cross-regulation between Pax3 and Pax7 may be possible in the developing neural tube, more specifically Pax7 expression may be negatively regulated by Pax3 (Borycki et al., 1999). Both Pax3 and Pax7 are expressed in the dorsal neural tube throughout the anterior-posterior axis but differ slightly in their exact temporal and spatial expression patterns (Goulding et al., 1991; Jostes et al., 1990; Stoykova and Gruss, 1994). Pax3 is expressed in the dorsal half of the neural tube and expression is initiated prior to closure, while Pax7 expression occurs in a smaller area of the dorsal neural tube and only following neural tube closure (Goulding et al., 1991; Jostes et al., 1990; Stoykova and Gruss, 1994). This hypothesis

is supported by the observation that in Splotch mice Pax7 expression is up regulated in the neural tube and its expression domain is also expanded to include areas in the dorsal tube where only Pax3 is normally expressed. Also, transfection of Pax3 into C2C12 cells decreases the levels of endogenous Pax7 (Borycki et al., 1999). The relevance of Pax7 repression by Pax3 during neural tube formation is unknown but may have to do with proper timing of neural tube closure.

1.4.2 Target Genes and Biological Pathways of Pax3 in Myogenesis

A target gene of Pax3 involved in myogenesis is the Met receptor tyrosine kinase signaling protein (Daston et al., 1996; Epstein et al., 1996; Yang et al, 1996). Late in development, Met expression occurs in a wide variety of embryonic tissues (Sonnenberg et al., 1993) but during early development the protein is expressed in two distinct populations in the developing somite: the ventrolateral and dorsomedially portions of the dermo-myotome (Bladt et al., 1995; Yang et al., 1996). During early development some Met expressing myoblasts in the somite delaminate from the ventrolateral portions of the somites and migrate into the developing limb buds. Upon arrival to the limb bud, terminal differentiation of the myoblasts into myotubes occurs and myogenic differentiation markers are expressed (Epstein et al., 1996; Yang et al., 1996). The expression domain of Met overlaps with that of Pax3. Pax3 is also expressed in the somite-derived myoblasts prior to and during migration to the limb bud to give rise to form limb musculature. Pax3 is actually used as a marker for migrating myoblasts (Bober et al., 1994; Williams and Ordahl, 1994; Goulding et al., 1994). Also, complete loss-of function of Met or Pax3 results in very similar mutant phenotypes. Both Met-/- and Pax3-/- embryos develop myoblasts with an inability to migrate into the limb bud and this leads to the formation of embryos with no limb muscles (Bober et al., 1994; Goulding et al., 1994; Bladt et al., 1995). Another observation that indicates that Met is a

Pax3 target gene is the observation that Sp/Sp mice have no detectable levels of the Met transcript in the ventrolateral somites (Daston et al., 1996; Epstein et al., 1996; Yang et al., 1996). Promoter studies with Met have led to the identification of a Pax binding site that has affinity for Pax3 as shown by electrophoretic mobility assays (Epstein et al., 1996). The ability of this site to mediate Pax3 dependent transcription activation is still to be determined either in cell culture and/or in the developing embryo. Hence, overlapping expression domains, similar loss-of-function phenotypes and the presence of a Pax binding site in the promoter sequence of Met indicate that Pax3 regulates expression of Met in migrating limb muscle precursor cells.

Evidence has surfaced which implicates the Dach2, Six1, Eya2 and Pax3 proteins as members of a molecular pathway that is involved in regulation of myogenesis (Heanue et al., 1999). It has been found that these four proteins have overlapping expression domains in the dorsal somite as well as in hypaxial myoblast precursors (Heanue et al., 1999). Pax3 and Dach2 must positively regulate each other's expression since overexpression of any one of these genes leads to the overexpression of the other. Eya2 is able to synergistically activate transcription of target genes with Six1 or Dach2. Ectopic expression or over expression of Eya2 with Dach2, or Six1 in somites leads to a synergistic increase in expression of Pax3 and the myogenic specific genes MyoD, Myogenin and MHC (Myosin heavy chain) (Heanue et al., 1999). It remains to be determined whether these myogenic specific genes are direct transcriptional targets of the Pax-Eya-Six-Dach pathway, but it has been demonstrated that Eya and Six have binding activity at the Myogenin promoter. In addition, it has been shown that their binding leads to synergistic activation of Myogenin (Spitz et al., 1998; Ohto et al., 1999).

Pax3 is a member of another pathway that is involved in myogenesis and includes the basic helix-loop-helix myogenic regulatory factors Myf5 and MyoD. Pax3 was implicated in

this pathway as a result of misexpression experiments and the analysis of mutant phenotypes that result when there is a loss of either of these genes individually or in pairs. Mice with either the MyoD or Myf5 genes knocked out show normal development of skeletal muscles but the loss of both genes results in no myogenesis of skeletal muscles. These observations suggest that MyoD and Myf5 either cooperate during muscle formation or have redundant functions (Rudnicki et al., 1993). Pax3 Sp/Sp mice lack limb muscles but Pax3/Myf5 double knock out mice develop no body muscles and lack MyoD expression, hence these two genes lie genetically upstream from the MyoD gene (Tajbakhsh et al., 1997). The upstream position of Pax3 relative to MyoD was also suggested with Pax3 ectopic expression studies. Ectopic expression of Pax3 leads to a rise in MyoD expression. Also, Pax3 expression in cell cultures induces the expression of reporter genes with MyoD regulatory elements fused upstream. This implies that MyoD may be a direct target gene of Pax3 (Bendall et al., 1999; Maroto et al., 1997). Although the hierarchical relationship between Pax3, Myf5 and MyoD has been established, further studies need to be carried to assess the molecular basis of the pathway in myogenesis. The different effects of the removal of these three genes on normal development of distinct myogenic lineages (head muscles and hypaxial and epaxial muscles) imply that the regulatory interactions between these genes are complex and may require the involvement of lineage specific proteins.

1.4.3 Target Genes and Biological Pathways of Pax3 in Melanogenesis

Pax3, along with Mitf and Sox10, is implicated in a pathway required for proper melanocyte development. All three genes have overlapping expression in melanoblasts and have similar phenotypes in heterozygote mutants. Human heterozygotes for loss-of-function mutations in any of these three genes have Waardenburg syndrome, albeit with different clinical subtypes depending on which of the genes are mutated (Baldwin et al., 1992; Pingault

et al., 1998; Tassabehji et al., 1994). All subtypes result in auditory and pigmentary defects. Therefore defects in these genes must disturb melanogenesis and therefore pigmentation, but the requirement for melanocytes in inner ear development is unknown. What is known is that melanocytes contribute to the cochlea and their absence is the cause of sensorineural deafness in Waardenburg syndrome (Auerbach, 1954; Hodgkinson et al., 1993; Southard-Smith et al., 1998; Steingrimsson et al., 1994). Mitf homozygote mutants are viable albinos that suffer from microphtalmia (small eye phenotype) (Baldwin et al., 1992; Pingault et al., 1998; Tassabehji et al., 1994). Mitf, a basic helix-loop-helix zipper transcription factor, regulates the expression of dopachrome tautomerase, tyrosinase and tyrosinase-related protein-1 (Trp-1), which encode proteins essential for the synthesis of melanin from tryptophan and tyrosine amino acids. This observation provides a molecular basis for the pigmentation defects observed in Mitf heterozygote mutants (Bertolotto et al., 1998; Yasumoto et al., 1997). Also, ectopic Mitf expression in fibroblasts changes these cells into melanocytes and therefore Mitf seems to be critical for the initiation of melanogenesis program during development (Tachibana et al., 1996). Pax3 involvement in melanogenesis is also supported by the discovery that Pax3 can transactivate both the Mitf and Trp-1 genes (Galibert et al., 1999; Watanabe et al., 1998). Also, Dopachrome tautomerase, an enzyme involved in melanin synthesis, is regulated directly or synergistically by Pax3 and Mitf (Jiao et al., 2006). The Mitf promoter can be bound and regulated by the Sox10 protein, but when Pax3 and Sox10 bind Mitf gene regulatory elements synergistic activation results (Bondurand et al., 2000; Potterf et al., 2000; Verastegui et al., 2000; Watanbe et al., 1998). A recent study in transgenic mice showed that during melanogenesis, Pax3 is essential for the proliferation of committed melanoblasts whereas Mitf ensures melanoblast cell survival within and immediately following emigration from the dorsal neural tube (Hornyak et al, 2001). Pax3 simultaneously functions to initiate a melanogenic cascade in specific neural crest cells while acting

downstream to prevent terminal differentiation by competing with Mitf and Sox10 for occupancy of an enhancer required for expression of the mature melanocyte specific dopachrome tautomerase (Dct), an enzyme required in melanin synthesis (Lang et al, 2005). Pax3 mediated repression of Dct and competition with Mitf for binding to the Dct promoter is carried out with Grg4, a protein known to bind Pax proteins and to function as a co-repressor (Lang et al., 2005). This is presumably to permit the increase melanoblast cell number prior to differentiation into melanocytes. These studies suggest that there is a hierarchy among the Pax3, Mitf and Sox10 transcription factors during melanogenesis and this accounts for the pigmentary disturbances associated with the Waardenburg syndrome. It is interesting to note that Pax3 regulation elements can be found in the Mitf promoter and contain PD and HD binding DNA sequences (Corry et al., 2005). The Trp-1 promoter contains only a PD binding site (Corry et al., 2005). This suggests that Pax3 can regulate target genes through alternate modes of DNA recognition. The role of Pax3 in melanogenesis is also supported by the observation that PAX3 re-expression is consistently observed in cutaneous malignant melanoma (CMM) and appears linked to progression of CMM (Blake et al., 2005).

Hair graying is an evident indication of aging, but not much is known about its causes. Recent papers have revealed evidence that hair graying is a result of incomplete melanocyte stem cell maintenance and inducate that Pax3 and Mitf are key molecules that help regulate the balance between melanocyte stem cell maintenance and differentiation and play a role in hair graying (Steingrimsson et al., 2005).

Interestingly, Pax3 also commits somitic cells to the myogenic lineage, but unlike in melanogenesis, Pax3 is also able to initiate myogenic differentiation by inducing the expression of myogenic regulatory factors such as MyoD (Bendall et al., 1999). As mentioned earlier Pax3-controlled myoblast differentiation is actually inhibited by Msx1 in migrating limb muscle precursors in order to sustain the proliferative and migratory capacity of the

migrating myoblasts until they reach and populate the limb bud (Bendall et al., 1999). In other words, Pax3 commits some neural crest cells to the melanogenesis pathway and simultaneously inhibits their differentiation to maintain their proliferative state, while Pax3 expression in myoblasts commits cells to the myogenic lineage and induces differentiation as well. These findings confirm the role of Pax3 in cell commitment and cell differentiation, although, for differentiation it may serve to activate or inhibit it.

1.4.4 Target Genes and Biological Pathways of Pax3 in Heart Formation

The role of Pax3 in heart formation was first identified when it was determined that Sp/Sp mice die in mid-gestation due to heart failure (Conway et al., 1997; Franz, 1989). Upon closer inspection it was found that septation of the outflow tract of the developing heart did not occur in these mice. This condition is known as persistent truncus arteriosus and is characterized by the lack of septation between the pulmonary artery and the aorta. In wildtype mice Pax3 is expressed in a subset of neural crest cells known as cardiac neural crest cells that migrate to the developing heart. Upon arrival to the heart these cells populate within the outflow tract and lead to the formation of the pulmonary artery/aorta septum. It was found that in Sp/Sp mice there is an elevated level of expression of the Msx2 gene in cardiac neural crest cells (Kwang et al., 2002). When Msx2 was knocked out in Sp/Sp mice cardiac neural crest cells were able to septate the outflow tract. Also, a 560 base pair element was identified upstream of the Msx2 gene that had affinity for Pax3 in vitro (Kwang et al., 2002). These observations indicate that perhaps Pax3 binds to an Msx2 regulatory sequence and represses Msx2 expression which, for still unknown reasons, is required for proper function of cardiac neural crest cells and septation of the outflow tract.

1.5 Pax3 and Tumorigenesis

Typical characteristics of tumor or cancerous cells are uncontrolled proliferation, resistance to terminal differentiation and cell immortality; in contrast, normal cells proliferate in a pre-programmed controlled manner. In order to support the quick processes of organogenesis and growth, many developmental genes, including the Pax genes, serve to maintain the proliferative, undifferentiated and anti-apoptotic potential of embryonic cells. Hence, the de-regulation of developmental control genes is frequently associated with tumorigenesis.

The chromosomal translocation between human chromosome 13 (FKHD gene containing locus, 13q14) and chromosome 2 (PAX3 gene containing locus 2q35) results in a pediatric tumor known as alveolar rhabdomyosarcoma (ARMS) (Barr et al., 1999). This translocation fuses the N-terminus of the PAX3 gene with the C-terminus of the Forkhead related (FKHR) gene (Barr et al., 1993; Davis et al., 1994). The resulting fusion protein consists of the PD and HD of Pax3 fused to the transctivation domain of FKHR (Davis et al., 1994; Galili et al., 1993; Shapiro et al., 1993). The structural nature of the fusion protein suggests that the molecular basis for ARMS is misexpression of PAX3 target genes. The transactivation ability of the PAX3-FKHR fusion protein is greater than that of PAX3 despite the fact that the chimeric protein binds PAX3 specific DNA sequences with lower affinity (Bennicelli et al., 1995; Fredericks et al., 1995; Sublett et al., 1995). The gain in transactivation ability of the fusion protein is due to a weakened susceptibility to the N-terminal repression domain of PAX3. The PAX3 and FKHR proteins have transactivation domains of similar potency but the PAX3 and PAX3-FKHR proteins do not. This implies that the transactivation domains are modulated differently and it may be possible that they are differentially regulated by the N-terminal repression domain of Pax3 (Bennicelli et al., 1996).

Deletion of the PAX3 repression domain in PAX3 and PAX3-FKHR proteins results in the relief of transcriptional repression in both proteins, but the fusion protein only experiences ten to one hundred folds less relief from repression (Bennicelli et al., 1996). This suggests that the repression domain is able to inhibit transcriptional activity of PAX3 more efficiently than that of the fusion protein, thereby possibly accounting for the more potent transcriptional activity observed for the fusion protein. Microarray studies have recently shown that PAX3-FKHR expression in ARMS tumors display an expression profile that is different from PAX3 expressing cells (Davicioni et al., 2006). The Met receptor, a transcriptional target of Pax3, has a role in PAX3-FKHR-mediated transformation (Taulli et al., 2006). Also, PAX3-FKHR induced expression in C2C12 cells promotes proliferation whilst blocking myogenesis (Wang et al., 2005). This is in line with the belief that PAX3-FKHR retains cells in a proliferative state by inhibiting terminal differentiation. Terminal differentiation in these cells has been shown to be a result of repression of MyoD and myogenin expression (Wang et al., 2005).

The chimeric protein has many oncogenic effects including a potent transformation activity. This activity requires the HD of PAX3 and the transactivation domain of FKHR (Lam et al., 1999; Scheidler et al., 1996; Xia et al., 2004). PAX3-FKHR protein is also a more potent inhibitor of myogenic differentiation than PAX3, as shown in C2C12 myoblasts (Epstein et al., 1995). Like PAX3, the fusion protein is able to induce the expression of many muscle related genes, in agreement with the muscle characteristics of ARMS (Khan et al., 1999). Some studies have also shown that, like PAX3, the fusion protein has anti-apoptotic activity which is presumably the cause of tumor cell survival. ARMS cells express high levels of PAX3 protein as well as the fusion protein (Bernasconi et al., 1996; Schafer et al., 1994). RNAi mediated depletion of PAX3 expression in ARMS tumor cells results in reduced cell viability and ectopic expression of PAX3 can prevent apoptosis. This implies that an elevated level of PAX3 expression is required for the survival of tumor cells (Bernasconi et al., 1996;

Margue et al., 2000). Supporting this hypothesis is the finding that both PAX3 and the fusion protein can bind the promoter and activate the expression of the anti-apoptotic BCL-XL gene (Margue et al., 2000). Perhaps the anti-apoptotic activity of PAX3 and PAX3-FKHR is mediated via the activation of this gene. RNAi studies, in which the expression of wildtype PAX3 in the somites of a normal embryo is repressed, have shown that these cells experience an increased frequency of apoptosis, which implies that the cell-survival, anti-apoptotic function of Pax3 is not specific to tumor cells (Borycki et al., 1999).

Tumors have evolved numerous mechanisms for evading the immune system; this is also the case with the PAX3-FKHR oncoprotein. PAX3-FKHR alters expression of genes that are normally regulated by the Janus kinase/signal transducer and activator of transcription (STAT) signaling pathways (Nabarro et al., 2005). This occurs as a result of a specific interaction between PAX3-FKHR and the STAT3 transcription factor, which results in a dramatic reduction in tumor MHC expression and an alteration in cytokine concentrations to inhibit surrounding inflammatory cells and immune detection (Nabarro et al., 2005).

Chapter 2

Site-Specific Modification of Single Cysteine Pax 3 Mutants Reveals Reciprocal Regulation of DNA Binding Activity of the Paired and Homeodomain

Abstract

The mechanism by which the paired domain (PD) and the homeodomain (HD) act together in the intact Pax3 protein to recognize DNA is unclear and was studied in a Pax3 mutant (Pax3-CL) devoid of cysteines. Pax3-CL binds to PD (P6CON-P3OPT sites) and HD (P2, P1/2 sites) DNA site sequences with near wild-type activity but, contrary to Pax3, in an *N*-ethyl maleimide (NEM) insensitive fashion. The Pax3-CL backbone was used for cysteine scanning mutagenesis and for site-specific NEM modification. Five single cysteine replacements were independently introduced in the PD, while eight were inserted in the HD. NEM sensitivity of PD and HD DNA binding was investigated in DNA-binding competent mutants. In the PD mutant C82, NEM abrogated DNA binding by the PD but also abolished DNA binding by the Cys-less HD. Likewise, in the HD mutant V263C, NEM modification abrogated DNA binding not only by the HD, but also by the Cys-less PD. The transfer of NEM sensitivity to the PD seen in V263C was specific and not due to simple loss of HD DNA binding since alkylation of adjacent V265C and S268C, although impairing HD DNA binding did not affect PD DNA binding. Thus, the PD and HD do not function as independent DNA binding modules in Pax3 but seem functionally interdependent.

Introduction

Pax-3 is a member of the mammalian Pax family (Stuart et al., 1994), a group of nine DNA-binding transcription factors structurally defined by a highly conserved DNA binding domain known as the PD, which was initially identified in the paired segmentation gene of *Drosophila* (Bopp et al., 1986). Pax proteins play a key role in directing tissue patterning and development of different organs during embryogenesis (Dahl et al., 1997), and mutations in Pax genes impair normal development of the skeleton (Pax1), kidney (Pax2), eye (Pax6), pancreas (Pax4), and thyroid (Pax8) (Balling et al., 1988; Sanyanusin et al., 1995; Torres et al., 1995; Jordan et al., 1992; Glaser et al., 1992; Hill et al., 1991; Sosa-Pineda et al., 1997; Mansouri et al., 1998; Macchia et al., 1998). Pax3 is expressed in the developing neural tube, in neural crest cell derivatives, and in migrating limb muscle precursor cells (Bober et al., 1994; Goulding et al., 1991), and a naturally occurring mutation in mouse *Pax3* (splotch) causes profound defects in neurogenesis (spina bifida and exencephaly) and myogenesis (no limb musculature) (Bober et al., 1994; Beechey et al., 1986; Auerbach et al., 1954; Franz et al., 1990; Franz et al., 1989; Goulding et al., 1994). Likewise, mutations in human *PAX3* cause Waardenburg syndrome, a pathology associated with pigmentary disturbances, cranio-facial abnormalities, and sensorineuronal deafness (Baldwin et al., 1992; Baldwin et al., 1995). It has been proposed that Pax3 plays a dual role in neurogenesis and myogenesis via transcriptional regulation of specific targets genes and by maintaining replicative potential of specific, Pax3 positive, epithelial, and migratory cell populations (Dahl et al., 1997).

Pax proteins show a modular structure. In addition to the PD, certain Pax proteins possess a second DNA binding domain, the paired-type HD (Stuart et al., 1994; Noll et al., 1993). A highly conserved octapeptide motif is also found in certain Pax proteins in the linker

separating the PD from the HD. Finally, the C-terminal half of Pax proteins shows a P/S/T domain resembling transactivating domains of other transcription factors. A high-resolution three-dimensional structure has been obtained for the PD-DNA complex of Prd and Pax6 and shows this domain to be bipartite (Xu et al., 1995; Xu et al., 1999) with N-terminal (PAI) and C-terminal subdomains (RED) each folding into three α-helices, with the last two forming a helix-turn-helix (HTH) DNA binding motif. In the PAI subdomain, a unique β-hairpin structure, consisting of 2 antiparallel β-strands joined by a type I β-turn, contacts the sugar-phosphate backbone to help anchor the subdomain to DNA. PAI holds a type II β-turn lying immediately C-terminal to the hairpin and makes base specific contacts with the minor groove. Helix 3 of PAI fits directly into the major groove and participates in base specific contacts with DNA (Xu et al., 1995; Xu et al., 1999). Both N- and C-terminal subdomains contribute to DNA binding to certain type of sequences (CD19/2, P6CON) (Czerny et al., 1993; Epstein et al., 1994). Other Pax proteins bind DNA mostly through the PAI domain, while certain isoforms of Pax6 and Pax8 bind DNA exclusively through the RED domain (Kozmik et al., 1997; Epstein et al., 1994). Finally, the linker segment joining the PAI and RED domains also make base-specific contacts in the minor groove (Xu et al., 1999), and alternative splicing of a single glutamine residue in this segment of Pax3 generates proteins with distinct DNA binding properties (Vogan et al., 1996; Vogan et al., 1997).

Paired-type HDs found in Pax proteins define a specific highly conserved subgroup (Kappen et al., 1993). The crystal structure of the DNA bound form of the paired-type HD has been determined and shows three α helical segments with helices 2 and 3 forming a HTH motif (Wilson et al., 1995). Although helix 1 does not make DNA contacts, residues N-terminal to it contact the minor groove. Helix 3 makes extensive DNA contacts in the major groove and is important for sequence specificity (Wilson et al., 1993; Treisman et al., 1989). A unique characteristic of paired-type HDs is their ability to cooperatively dimerize on

palindromic sequences of the type TAAT-(N_{2-3})-ATTA. Position 50 (helix 3) of the HD plays a key role in binding specificity and dimerization potential of HDs; while the HD of Phox (Gln50) can only dimerize on TAAT-(N_3)-ATTA sequences showing a three-nucleotide spacer, the HD of Pax3 (Ser50) can only dimerize (in the context of a PD) on sequences harboring a two-nucleotide spacer (Wilson et al., 1993; Schafer et al., 1994).

Although both the PD and HD of Prd and Pax3 can bind DNA on their own (Underhill et al., 1995), a large body of data suggests that they do not function independently in intact Pax3, perhaps conferring additional sequence binding specificity to the protein. Indeed, both PD and HD are required for function in Pax3 and Pax6 and mutations in either domain show similar loss-of-function phenotypes in vivo (Glaser et al., 1992; Baldwin et al., 1995; Lalwani et al., 1995). Interestingly, a mutation (G42R) in the PD of Pax3 found in the splotch-delayed mouse mutant (Sp^d) not only shows reduced DNA binding to PD oligos, but also causes reduced DNA binding to HD-specific oligos (Underhill et al., 1995). Conversely, a WS mutation in the HD (R53G) also modulates binding of Pax3 to PD-type sequences (Fortin et al., 1997). Additional studies in Pax3-Phox chimeras have shown that the Pax3 PD can modulate the DNA binding specificities and dimerization potential of a heterologous HD (Fortin et al., 1998). Deletion of helix 2 of the PD in the context of the Sp^d mutation restores DNA binding by the HD, identifying this helix as a key structural element in this regulation (Forti et al., 1998). These results have suggested that the PD and HD are functionally interdependent and interact for final target site selection. However, the mechanistic basis and the protein segments involved in this functional interaction remain poorly understood so far.

Site-specific modification of cysteines with thiol-specific reagents is a versatile tool for structure: function studies (Frilligos et al., 1998; Loo et al., 2000). This is best accomplished in the Cys-less molecular backbone in which single cysteines are inserted by

mutagenesis in predetermined strategic locations (Frilligos et al., 1998). For example, important information on the chemical environment of a specific residue or protein segment (e.g., transmembrane domains vs. solvent-exposed loops) can be obtained by reacting the modified protein with thiol reagents of different chemical and physical properties, such as varying size and degree of hydrophobicity. Residues in substrate-binding pockets can be identified by Cys-scanning mutagenesis as positions that confer NEM sensitivity to substrate binding and, conversely, at which prior reaction with substrate protects against NEM alkylation. Dynamic changes in structure following substrate binding or catalytic activity can be further studied in quenching experiments, using purified protein modified by fluorescent sulfhydryl reagents (Frillingos et al., 1998). Finally proximity relationships can be studied in mutants containing Cys pairs, by a variety of methods including cross-linking with bifunctional reagents (Frillingos et al., 1998; Loo et al., 2000), FRET (Qu et al., 2001), excimer fluorescence (Wang et al., 1992; Jung et al., 1993; Sen et al., 1990), and several others. In the present study, we have created and functionally characterized a Pax3 mutant in which all Cys residues have been replaced by Ser or Gly. We have used this mutant backbone to reintroduce Cys residues at strategically located positions of the PD and HD in individual mutants. The effect of site-specific modification of these single Cys mutants by sulfhydryl reagents on the DNA binding properties of the PD and HD was characterized.

Materials and Methods

Mutagenesis. The construction of the pMT2 expression plasmid encoding a portion of wildtype (wt) *Pax3* cDNA (positions 297-1801) has been previously described (Underhill et al., 1995). This pMT2Pax3 construct encodes the full-length 479 amino acid murine Q+ isoform of Pax3. A 1.3 kb *PstI* internal Pax3 fragment from upstream the initiator AUG to nucleotide position 1581 (residues 1-429) that encodes the PD and the HD was inserted into the corresponding site of plasmid vector pAlter-1™ (ProMega) to create pAlterPax3. The fragment was also inserted into the *Pst I* site of a eukaryotic expression plasmid pMT2 (Fortin et al., 1997) to generate pMT2Pax3(B). The pAlterPax3 plasmid encodes the first 429 residues of Pax3 and consequently encodes for six of the seven endogenous cysteines of Pax3. Altered Sites II in vitro mutagenesis systems (Promega) was used with the pAlterPax3 plasmid to mutate all but the most C-terminal cysteine of Pax3, creating pAlterPax3CL/C429. pAlterPax3CL/C429 was used for mutagenesis in the construction of the single cysteine mutants. The mutagenic oligonucleotides used to generate the Cys-less (CL), and the single cysteine mutants are listed in Table 1. The five most N-terminal endogenous cysteines of Pax3 (Cys70, 82, 88, 143, 355) were substituted to serine or glycine. The single Cys mutants CL/C82 andCL/C88 were obtained at intermediate steps of this process.

The pMT2Pax3(B) plasmid was further modified by PCR-mediated mutagenesis in order to insert in-frame at the C-terminus of the protein, both a polyhistidine tail (His_6) and an antigenic hemagglutinin A epitope followed by a new termination codon. This was carried out using sequence-specific oligonucleotide primers (5')-CAGGTGACAACG-CCTGACGTGGAG-(3') and (5')-CCTTTGGAATTCCT-GCAGTCAATGATGATGATGATGATGTCGC-GAAGCGTAGTCTGGCACATCGTATGGGTATACG-

Table 1. Oligonucleotides used for Pax3 mutagenesis

Substitution	Mutagenic primer (5'-3')
C70S	CATTCGGCC**G**A**G**CGTCATTTC
C82S	GTCCCATGG**A**T**C**CGTCTCTAAG
C88S	CTAAGATCCT**A**G**G**GCAGGTAC
C143S	GGACGCTGTCA**G**C**G**ATCGGAACACTGTG
C355S	CAGCTCTGCCTAC**A**G**T**CT**T**CCCAGCACCAG
S152C	GTGCCCTCAGTG**T**GTTCTATC**TCG**CG**A**ATCCTGAGGAG
S153C	CACTGTGCCCTCAGT**C**AG**C**T**G**TATCAGCCGCATCCTG
S155C	CAGTGAGTTCTATC**T**GCCGCAT**T**CTGAGGAGTAAATTTG
R221C	GAGGAAGCAGCGC**T**GTAGCAGAACCACC
S222C	GCAGCGCAGGTGC**C**G**T**AC**G**ACCTTCACGGC
V263C	GCTTACCGAGGC**G**CGC**T**GTCAGGTCTGGTTTAG
V265C	CGAGTGCAG**T**GTTGGTTTAGCAACCGCCGTGCA**C**GATGGAG
S268C	GCAGGTCTGGTTTTGCAACCG**G**CG**C**GC**C**AGATGGAGGAAAC
R270C	GGTTTAGCAA**T**TGCCGTGCAAG
A272C	GCAGGTCTGGTTTT**CG**AACCGCCGTTGCAGATGGAGGAAAC
W274C	CGTGCAAGATGTAGGAAACAAGC**C**GG**C**GCCAATCAACTG

Nucleotide substitutions leading to amino acid changes are indicated in bold and those that introduce silent restriction sites are underlined.

TACCTCGAGCTGGCTGACACCGTGGTC-(3'). The mutagenic oligonucleotide also replaces Cys429 by a serine. The resulting 1049 bp PCR product was digested with restriction enzymes *Apa I* and *Eco RI* (sites embedded in the oligo) and ligated into the corresponding sites of pMT2Pax3(B) to produce pMT2Pax3C429S/HA. The 12 single Cys mutants generated in pAlterPax3CL/C429 plasmid (mutant oligonucleotides listed in Table 1) were reconstructed in pMT2C429S/HA using endogeneous Pax3 restriction sites *Sma I* (Pax3; pst 342-672) or *Kpn I* (Pax3; pst 563-1500). This permitted the production of the following expression plasmids: pMT2Pax3CL, pMT2Pax3CL/C82, pMT2Pax3CL/C88, pMT2Pax3CL/S152C, pMT2Pax3CL/S153C, pMT2Pax3CL/S155C, pMT2Pax3CL/R221C, pMT2Pax3CL/S222C, pMT2Pax3CL/V263C, pMT2Pax3CL/V265C, pMT2Pax3CL/S268C, pMT2Pax3CL/R270C, pMT2Pax3CL/A272C, and pMT2Pax3CL/W274C. 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 pMT2Pax3 expression plasmids were introduced into COS7 cells via transient transfection. One million cells were plated in Dulbecco's modified Eagle medium (DMEM) containing 10% fetal bovine serum and were transfected by the calcium phosphate coprecipitation method using 15 ug of plasmid DNA doubly purified by ultracentrifugation on cesium chloride density gradients. Cells were exposed to calcium-DNA precipitates 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. Following this treatment the cells were washed once and placed in complete DMEM. Whole cell extracts were prepared 24 h following HBS/15% glycerol treatment by sonication in a buffer containing 20 mM HEPES (pH 7.6), 0.15 M NaCl, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP), 0.2 mM EDTA, 0.2 mM EGTA, and a cocktail of protease inhibitors: aprotinin, pepstatin, and leupeptin at 1 μg/mL and phenylmethysulfonyl fluoride 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 antimouse horseradish peroxidase conjugated secondary antibody (Amersham).

Electrophoretic Mobility Shift Assays. Electrophoretic mobility shift assays were performed as previously described (Underhill et al., 1995). Each protein:DNA binding reaction was carried out using approximately 8 μg of total cell extracts from transiently transfected COS-7 monkey cells and 10fmol (0.06 μCi) of radioactively labeled double stranded oligonucleotides containing either PD or HD recognition sites. The final concentration of labeled oligonucleotide in the binding reaction is 0.0005 μM. Whole cell extracts were incubated with ^{32}P-labeled PD 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, 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 12V/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. Some films were used to perform densitometry studies to quantitate the amount of radiolabeled probe that was protein-bound using a Fuji LAS-1000. PD-specific sequences P6CON (5')-TGGAATTCAGGAAAAATTTTCACGCTTGAGTTCACAGCTCGAGTA-(3') (Epstein et

al., 1994) and P3OPT (5')-TGGTGGTCACGCCTCATTGAATATTA-(3') (Chalepakis et al., 1995; Epstein et al., 1995) and HD-specific sequences P2 (5')-'GATCCTGAGTCTAATTGATTACTGTACAGG-(3') (Wilson et al., 1993) and P1/2 (5')-GATCCTGAGTCTAATTGAGCGTCTGTAC-(3') (Wilson et al., 1993) 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.

Thiol-specific reagents *N*-ethylmaleimide (Pierce) and dibromobimane (Molecular Probes) were prepared as 10 mM stocks in water and 100% dimethyl sulfoxide, respectively, and were stored frozen until use. They were added as a 0.5 µL aliquot to a 4 µL volume of whole cell extract, followed by a 30 min incubation at room temperature prior to the addition of the [^{32}P]-labeled probe and EMSA.

Results

Construction and Characterization of a Pax3 Mutant Devoid of Cysteines. The analysis of structure:function relationships by site-specific modification with sulfhydryl reagents is most conveniently carried out in a protein backbone lacking cysteine residues (Cys-less), where single Cys can be introduced at strategic sites by site-directed mutagenesis. Pax3 has 7 Cys residues at positions 70, 82, 88, and 143 in the PD and at positions 355, 429, and 449 downstream of the HD. A Cys-less Pax3 mutant was created in a Pax3 cDNA fragment (residues 1-429), which lacks the last 50 residues (including Cys449 but which retains wild-type DNA binding activity by the PD and by the HD (Underhill et al., 1995). An alignment of human and mouse Pax protein sequences (Table 2) was used to guide the choice of amino acid to be substituted for each Cys residue. Cys70, Cys82, and Cys143 in the PD are invariant among Pax family members and were replaced by Ser. Serine and Cysteine have a very similar structures, differing only by the size of the Oxygen vs Sulfur atom in the side chain, and therefore these substitution are expected to have a minimal structural consequence. Indeed, a C70S mutation in Pax8 has been previously shown to be without consequence on DNA binding (Tell et al., 1998). Five Pax proteins harbor a Gly at position 88 (Table 2); thus, the Cys88 of Pax3 was mutated to Gly. Finally, Cys355 and Cys429 in the P/S/T-rich trans-activation domain are not conserved in other Pax proteins (Table 2) and were substituted to Ser. All mutations were introduced sequentially by site-directed mutagenesis, and the integrity of the final cDNA was verified by nucleotide sequencing.

The Cys-less (CL) Pax3 construct was further modified by the addition of a hemagglutinin (HA) epitope tag at its C-terminus to facilitate protein detection and was introduced in the pMT2 expression vector followed by transient transfection into COS-7

Table 2. Conservation of cysteine residues in mouse and human Pax proteins

	Endogenous Cysteines of Pax3					
	Cys70	**Cys82**	**Cys88**	**Cys143**	**Cys355**	**Cys429**
Pax Proteins						
Pax3/PAX3	Cys	Cys	Cys	Cys	Cys	Cys
Pax4/PAX4	Cys	Cys	Gly	Cys	Cys	Pro
Pax6/PAX6	Cys	Cys	Gly	Cys	Met	Ser
Pax7/PAX7	Cys	Cys	Cys	Cys	Ser	Cys
Pax1/PAX1	Cys	Cys	Ala	Cys	Cys	Leu
Pax9/PAX9	Cys	Cys	Ala	Cys	Ser	Leu
Pax2/PAX2	Cys	Cys	Gly	Cys	Thr	Ser
Pax5/PAX5	Cys	Cys	Gly	Cys	Pro	Ser
Pax8/PAX8	Cys	Cys	Gly	Cys	Gly	Tyr
Substitution Made	**Ser**	**Ser**	**Gly**	**Ser**	**Ser**	**Ser**

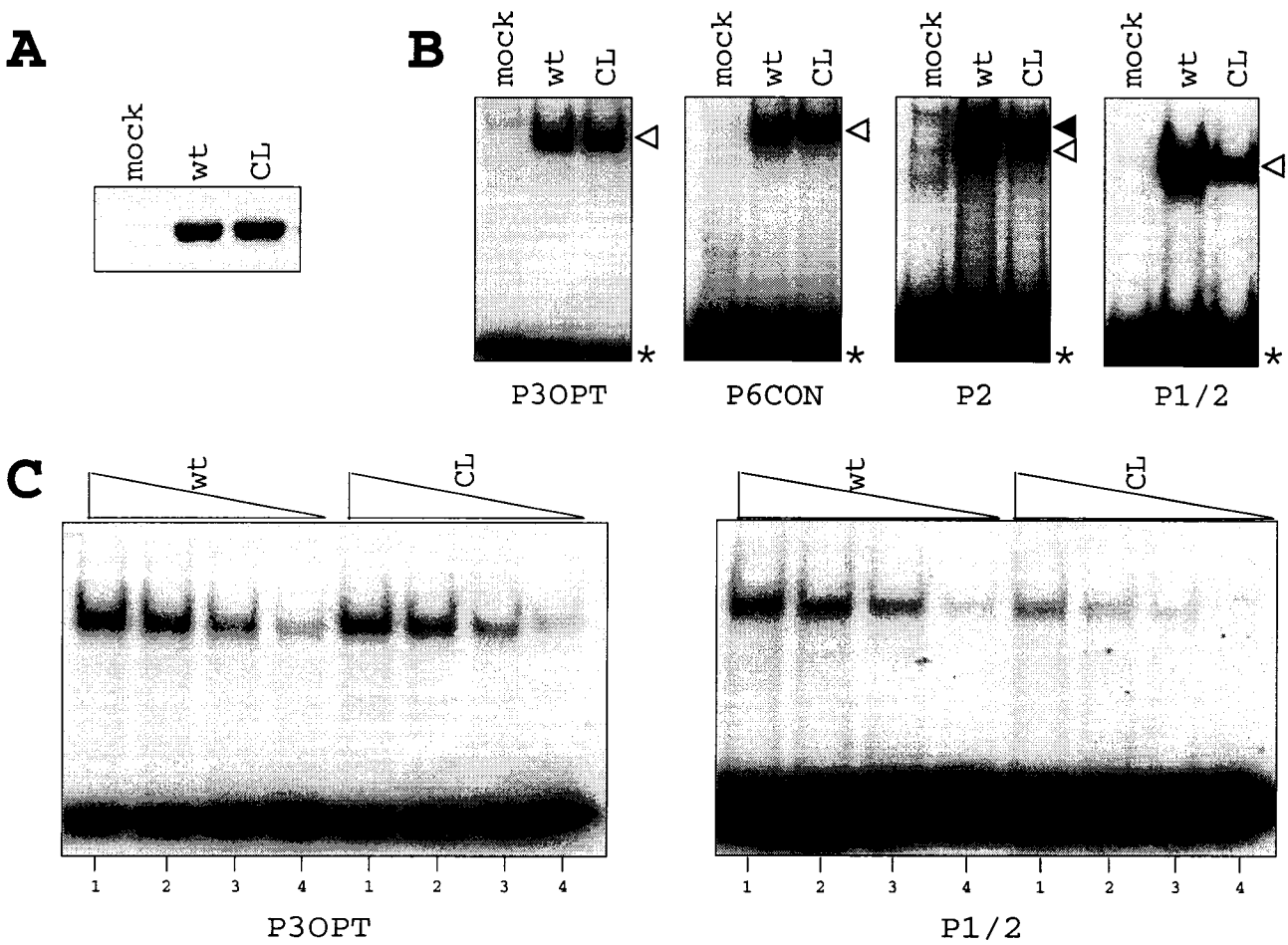
Monkey cells. Immunoblotting of whole cell extracts with an anti-HA monoclonal antibody indicates similar stability and levels of expression of both the wild type (wt) and Cys-less (CL) Pax3 proteins in COS7 cells (Figure 1A). The effect of Cys replacement on DNA-binding properties of the PD and HD of Pax3 was examined by electrophoretic mobility shift assays (EMSA). DNA binding by the PD was examined using oligonucleotide probes P3OPT (Epstein et al., 1994) and P6CON (Chalepakis et al., 1995; Epstein et al., 1995), previously shown to reveal binding determinants present in both the amino (PAI) and carboxy (RED) subdomains of the PD. Results in Figure 1B indicate that the wt and CL Pax3 can both bind these probes, and semiquantitative analysis in the form of dilution series suggests similar affinities of both proteins for P3OPT (Figure 1C) and P6CON (data not shown). The effects of Cys replacement on DNA binding properties of the HD were evaluated using a target sequence (P2) containing the sequence $TAAT(N)_2TAAT$ previously shown to support cooperative dimerization of Pax3 (Wilson et al., 1993). In addition, an oligonucleotide containing half of this sequence (half site, P1/2) and revealing monomeric Pax3 binding by the HD was used. Results in Figure 1B,C indicate that CL Pax3 can bind to P1/2 (monomer, empty arrowhead) and can dimerize on P2 (shaded arrowhead), albeit with somewhat reduced proficiency. Similar levels of wt and CL Pax3 proteins were present in all experiments, as revealed by immunoblotting. Taken together, these experiments indicate that removal of all Cys residues from Pax3 does not have a major effect on DNA binding by either the PD or the HD.

Effects of Thiol-Specific Reagents on DNA Binding by WT and Cys-Less Pax3.

Sulhydryl or thiol-reactive compounds such as *N*-ethyl maleimide (NEM) or dibromobimane (DBB) can form covalent adducts with Cys residues and can be used efficiently for site-specific modification in structure:function studies. The effect of NEM and DBB on DNA binding by the PD and HD of WT and CL-Pax3 proteins was tested in EMSA. In these

Figure 1

PD and HD DNA binding properties of wild-type and of Cys-less Pax3. (A) Immunodetection of wild-type (wt) and Cys-less Pax3 (CL) proteins in whole cell extracts from COS7 monkey cells transfected with corresponding Pax3 cDNAs modified by the in-frame addition of a *Haemophilus influenza* hemagglutinin A (HA) epitope tag at the C-terminus of each protein. Proteins were resolved by SDS-PAGE (10% acrylamide) and transferred to a nitrocellulose membrane. Detection was by enhanced chemiluminescence using a mouse anti-HA monoclonal antibody and a HRP-conjugated secondary antibody. The "mock" labeled lane refers to whole cell extracts from untransfected COS7 cells. (B) Electrophoretic mobility shift assays were used to measure the DNA binding properties of either Pax3 (wt) or CL Pax3 against PD (P3OPT, P6CON) and HD specific binding sites (P2, P1/2). Protein-DNA complexes were formed using total COS7 cell extracts and were resolved on 6% acrylamide nondenaturing gels, as described in Materials and Methods. Open arrowheads identify monomeric Pax3/DNA complexes, while closed arrowheads identify Pax3 dimers bound to the P2 probe. Free oligonucleotide probe is identified by an asterisk. (C) Dilution series of Pax3 (wt) and CL Pax3 proteins to compare relative affinities for P3OPT (PD) and for P1/2 (HD) probes. For each panel, lanes 2, 3 and 4 correspond to 1.5-, 2-, and 4-fold dilutions of total cell extract in lane 1.


A
mock
wt
CL
B
mock
wt
CL
P3OPT
P6CON
P2
P1/2
C
wt
CL
1
2
3
4
P3OPT
P1/2

experiments, extracts from COS cells expressing each protein were incubated with various amounts of NEM prior to EMSA with PD specific (P3OPT, P6CON) and HD specific (P2, P1/2) oligonucleotides (Figure 2). In WT Pax3, PD-specific DNA binding to P3OPT was completely abrogated by 1 mM NEM. Studies with a narrower NEM concentration range showed that 0.3 mM reduced DNA binding by 50%, while no binding was detectable at 0.5 mM (Figure 2, lower panel, Figure 6B). Identical results were obtained when P6CON was used as the PD target sequence; in some experiments, a Pax3-independent, NEM-sensitive P6CON complex of slower electrophoretic mobility was detected in all samples including the negative control (Figure 2; asterisk). NEM sensitivity of DNA binding by the PD was also seen when other PD probes (Nf3', CD19-2A) were used in these experiments (data not shown). A similar analysis of the HD revealed that dimerization of the WT Pax3 on P2 sequences is also NEM-sensitive and is abrogated by NEM concentration of ~0.5 mM (Figures 2 and 6B). Likewise, monomeric binding of WT Pax3 to a P1/2 probe is also largely abrogated by NEM concentrations in the range of 0.1-1 mM; however, we note that a small proportion of monomeric binding to P1/2 is NEM-insensitive, and quantitation of the signal indicates that up to 20% of WT Pax3 binding to P1/2 is NEM-insensitive (Figure 2, lower panel and Figure 6B). In contrast, DNA binding by the PD and the HD in the CL Pax3 mutant was completely NEM insensitive in the same assay for all concentrations tested up to 3.6 mM (Figures 2 and 6B). Finally, identical results were obtained when dibromobimane was used in place of NEM in these experiments (data not shown).

Together, results in Figures 1 and 2 indicate that a Pax3 mutant devoid of Cys residues can still bind DNA through its PD and HD. Thus, inactivation of Pax3 DNA binding by NEM (Figure 2) is probably due to introduction of one or more bulky groups into the molecule rather than the removal of a critical sulfydryl group. Importantly, these results show that the

Figure 2

Effects of *N*-ethyl maleimide treatment on the PD and HD DNA binding properties of WT and CL Pax3. Total cell extracts from COS7 monkey cells (mock) or from cells expressing either wild-type Pax3 (WT) or the Cys-less Pax3 mutant (CL) were incubated in increasing concentrations of *N*-ethyl maleimide (NEM) prior to electrophoretic mobility shift assay. The DNA binding properties of the PD were evaluated with target sites P3OPT and P6CON, while the HD was tested using the P2 and P1/2 sites. Two series of NEM concentrations (in mM) were tested for WT and are indicated on top of each panel. Monomeric Pax3/DNA complexes are indicated by an open arrowhead, while dimers formed on P2 are identified by a shaded arrowhead. The presence of a Pax3-independent band detected in some of the EMSA with P3OPT and P6CON is identified by an asterisk. All gels were exposed for similar time periods.

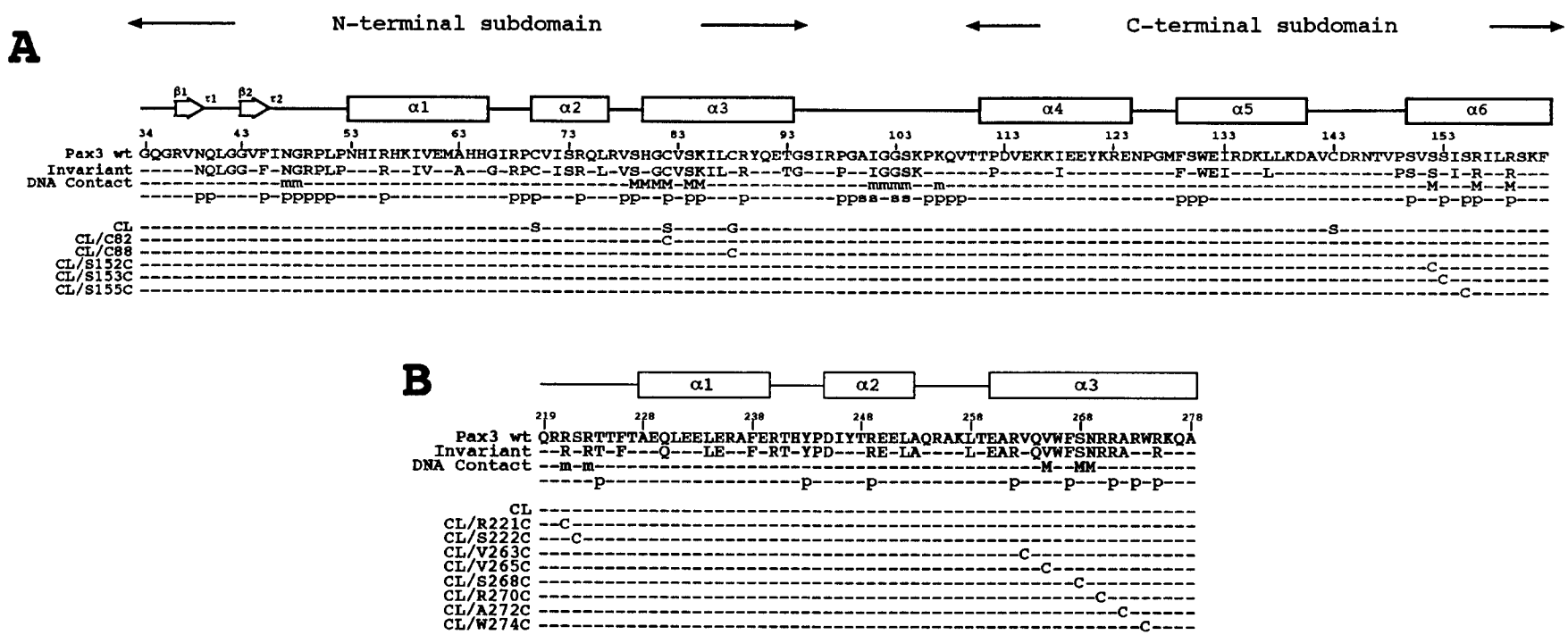
P3OPT
P6CON
P2
P1/2
CL
LM
mock
0
0.001
0.01
0.1
1
3.6
0.28
0.46
0.64
0.82

CL Pax3 mutant can be used as molecular backbone for site-specific modification in Pax3 mutants containing single cysteines inserted by site-directed mutagenesis.

Construction and Characterization of Single Cysteine Pax3 Mutants. Site-specific modification in single cysteine mutants has been used to study, in a dynamic fashion, structural changes in proteins associated with ligand binding or catalysis (Frillingos et al., 1998). Our goal was to use such mutants to study possible interactions between the PD and HD of Pax3, including reciprocal regulation of DNA binding properties. Hence, we aimed to study the effect of site-specific modification of single Cys residues in one domain on DNA binding properties of the other domain. Ideal substitutions for this type of analysis would be Cys insertions in one domain that do not affect DNA binding by this domain but that are strategically positioned close to the DNA molecule so that formation of a bulky adduct at that position upon sulfhydryl modification would hinder DNA binding by this domain. In addition, these insertions should be accessible to NEM, a parameter that must be determined experimentally. The high-resolution crystal structures of the PD of Pax6 (Xu et al., 1999) and of the HD of Prd (Wilson et al., 1995) proteins bound to DNA were used to guide the choice of residues to be mutated to Cys (Figure 3). Five mutations were independently created in the C-terminal helices of each HTH motif of the PD (helices α3 and α6). These two helices make key contacts in the major groove and are essential for DNA binding (Xu et al., 1999). C82 and C88 (α3 helix) are endogenous cysteines highly conserved in Pax proteins; these were re-created as single cys mutants. Likewise serines 152, 153, and 155 in α6 are well conserved in Pax proteins, contact DNA (Figure 3A), and were individually mutated to S152C, S153C, and S155C. In the HD, the N-terminal arm (pst 219-227) is important for regulation of DNA binding (Wilson et al., 1995), through DNA contact with the minor groove (Figure 3B); R221 and S222 in this segment were mutated to Cys. Residues of helix 3 of the HTH motif, including four highly conserved arginines, make extensive base-specific and phosphate

Figure 3

Figure 3 Site-directed mutagenesis of the PD and HD of Pax3. (A) Schematic representation of the N-terminal and C-terminal subdomains of the PD, together with structural features based on the three-dimensional structure of the PD of the *Drosophila* Prd protein (β,β-strand; τ, β-turn; α, α-helix) (Xu et al., 1995; Xu et al., 1999). The amino acid sequence for positions 34-162 of Pax-3 is shown, and invariant residues among all known PDs are identified below. The type of predicted DNA contacts made by these residues (p, phosphate; m, minor groove; M, major groove) is shown. The position and nature of the mutations introduced in Pax-3 to create the Cys-less (CL) mutant, as well as the positions of single cysteine mutants corresponding either to endogenous Cys positions (CL/C82; CL/C88) or novel insertions, are shown below. (B) Schematic representation of the Pax3 HD, including the presence and position of predicted structural features, invariant amino acid residues, number and types of DNA contacts, and positions of single cysteine insertions (as for panel A).

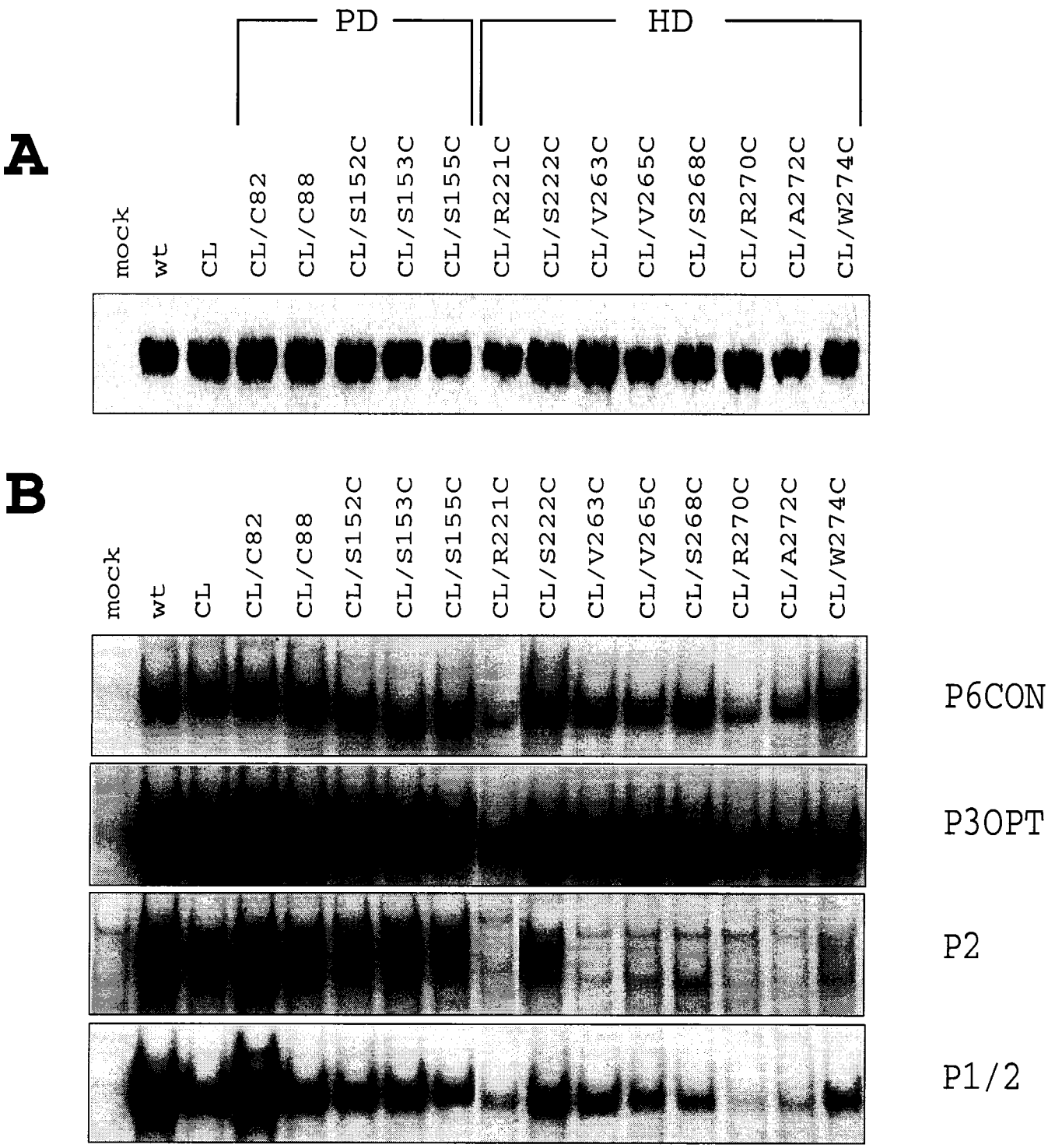

A
N-terminal subdomain
C-terminal subdomain
β1
τ1
β2
τ2
α1
α2
α3
α4
α5
α6
34
43
53
63
73
83
93
103
113
123
133
143
153
Pax3 wt GQGRVNQLGGVFINGRPLPNHIRHKIVEMAHHGIRPCVISRQLRVSHGCVSKILCRYQETGSIRPGAIGGSKPKQVTTPDVEKKIEEYKRENPGMFSWEIRDKLLKDAVCDRNTVPSVSSISRILRSKF
Invariant -----NQLGG-F-NGRPLP---R--IV--A--G-RPC-ISR-L-VS-GCVSKIL-R---TG---P--IGGSK------P------I----------F-WEI---L----------PS-S-I-R--R---
DNA Contact ----------mm----------MMMM-MM----------mmmm--m----------M---M--M---
-----pp----p-ppppp----p----------ppp---p---pp--p-pp---p--------ppss-ss-ppppp----------ppp----------p--p-pp--p---
CL ----------S----------S-----G----------S----------
CL/C82 ----------C----------
CL/C88 ----------C----------
CL/S152C ----------C----------
CL/S153C ----------C----------
CL/S155C ----------C----------
B
α1
α2
α3
219
228
238
248
258
268
278
Pax3 wt QRRSRTTFTAEQLEELERAFERTHYPDIYTREELAQRAKLTEARVQVWFSNRRARWRKQA
Invariant --R-RT-F---Q---LE--F-RT-YPD---RE-LA-----L-EAR-QVWFSNRRA--R---
DNA Contact --m-m----------M--MM----------
-----p----------p-----p----------p----p---p-p-p---
CL ----------
CL/R221C --C----------
CL/S222C ---C----------
CL/V263C ----------C----------
CL/V265C ----------C----------
CL/S268C ----------C----------
CL/R270C ----------C----------
CL/A272C ----------C------
CL/W274C ----------C-----

contacts in the major groove; thus, R221C, S222C, V263C, V265C, S268C, R270C, A272C, and W274C were created in this segment. All mutants were created by site-directed mutagenesis, and the corresponding HA-tagged proteins were expressed by transient transfection in COS-7 cells.

Immunoblotting experiments show that all mutant variants could be expressed to similar levels in COS-7 cells (Figure 4A), indicating that none of the mutations affect protein expression or stability. The effect of single Cys mutations on DNA binding by the PD and HD was analyzed by EMSA, using P6CON-P3OPT and P2-P1/2, respectively (Figure 4B). All the PD SCMs, except C82, have PD and HD DNA binding activities similar to CL Pax3. C82 is the only PD SCM that behaves similarly to WT with respect to DNA binding properties and NEM sensitivity (Figure 4B). Therefore, none of the single Cys mutants inserted in the PD abolished the DNA binding properties of either domain. Most HD SCMs showed near-wild-type activity for PD sequences (P6CON, P3OPT). HD SCMs R221C, R270C, and A272C show a reduction in PD DNA binding activity, though this reduction is small. However, HD mutations had a marked effect on DNA binding by the HD. Although several mutants retained some DNA binding for P1/2, mutants at highly conserved positions R221, R270, and A272 showed severely reduced binding activities. In addition, none except S222C was capable of cooperative dimerization on P2 (Figure 4B). These results confirm the key role of α3 helix in DNA binding by the HD. Reduced P1/2 binding by the mutants is concomitant to an inability to dimerize on P2, showing that integrity of α3 is also essential for cooperative homodimerization of the HD. Of the HD SCMs with binding activity for P1/2, all but S222C have activities less than that of CL Pax3. Due to their severe reduction of HD DNA binding activity, mutants R221C, R270C, and A272C could not be studied further. Nevertheless, several single Cys mutants in the PD and HD did retain robust DNA binding by both domains and were thus suitable for site-specific modification studies.

Figure 4

PD and HD DNA binding properties of Pax3 mutants bearing single cysteine replacements. (A) *Pax3* mutants bearing single cysteine (single Cys) insertions onto the backbone of a Cys-less (CL) *Pax3* mutant were created by site-directed mutagenesis, cloned into pMT2 expression plasmid, and total cell extracts from transiently transfected COS7 Monkey cells were separated by SDS-PAGE and analyzed by immunoblotting with a mouse anti-HA monoclonal antibody, as described in the legend to Figure 1. (B) EMSA analysis of the DNA binding properties of either Pax3 (wt), CL Pax3 (CL), or single Cys mutants against PD (P3OPT, P6CON) and HD specific binding sites (P2, P1/2) was as described in Figure 1.

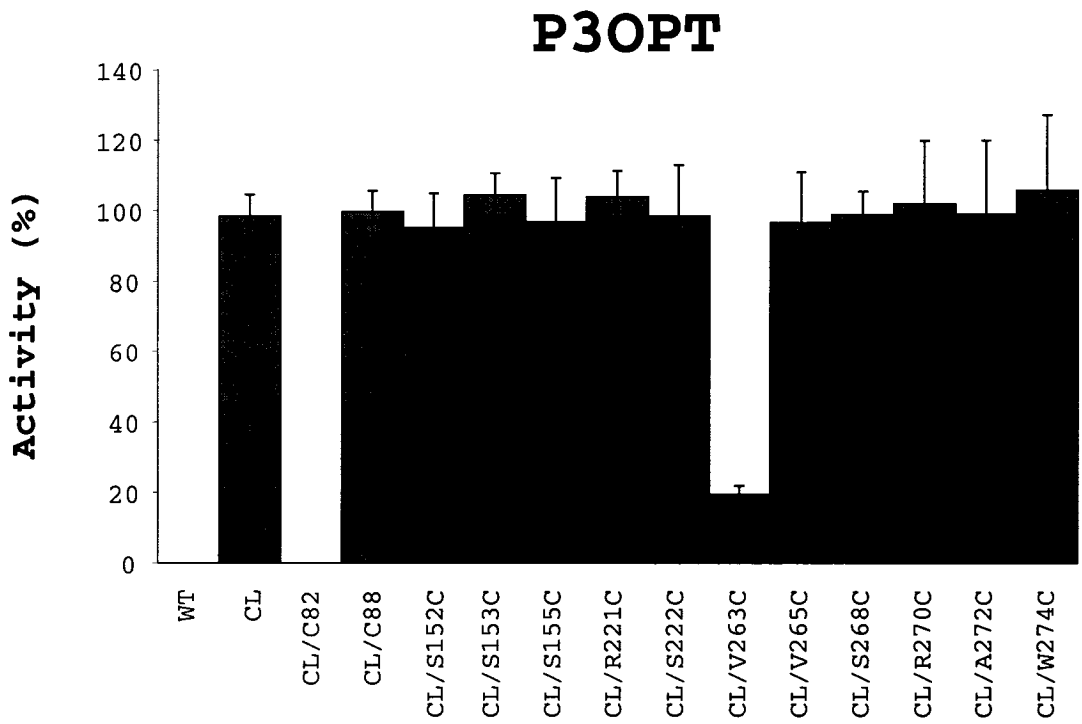

A
PD
HD
mock
wt
CL
CL/C82
CL/C88
CL/S152C
CL/S153C
CL/S155C
CL/R221C
CL/S222C
CL/V263C
CL/V265C
CL/S268C
CL/R270C
CL/A272C
CL/W274C
B
mock
wt
CL
CL/C82
CL/C88
CL/S152C
CL/S153C
CL/S155C
CL/R221C
CL/S222C
CL/V263C
CL/V265C
CL/S268C
CL/R270C
CL/A272C
CL/W274C
P6CON
P3OPT
P2
P1/2

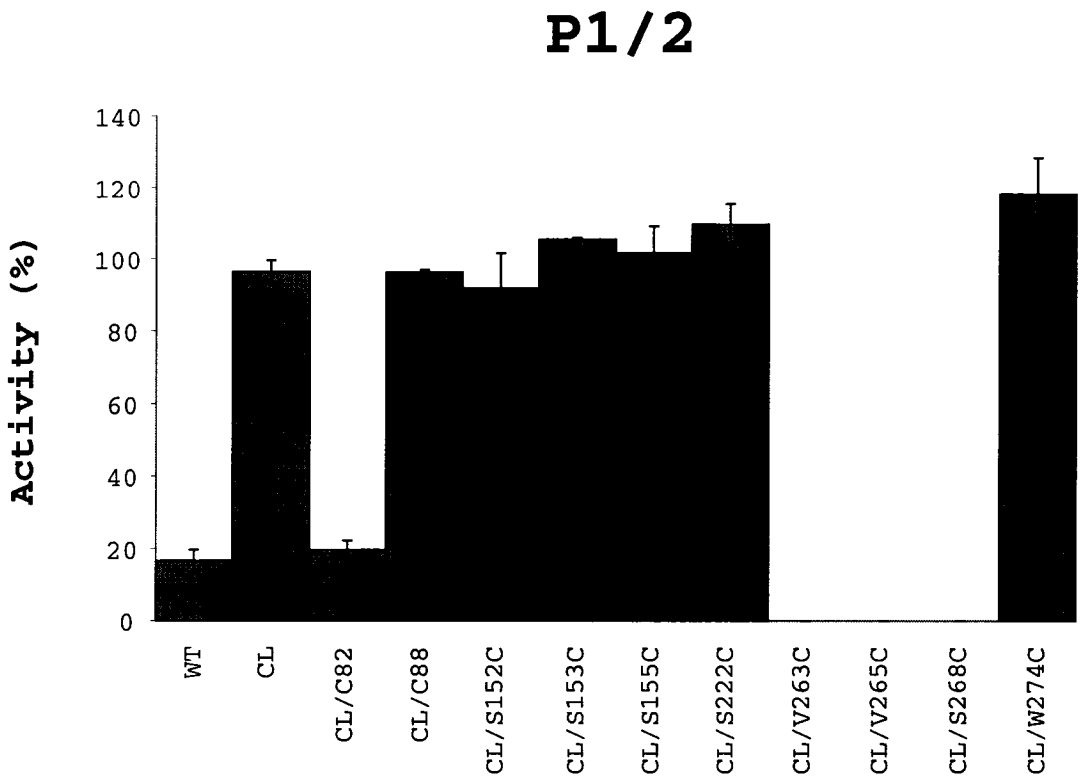
Effect of Thiol Specific Reagents on DNA Binding by Single Cysteine Pax3 Mutants. Initially, whole cell extracts from COS-7 cells expressing individual mutants were incubated with 1 mM *N*-ethyl maleimide (NEM), and the ability of each mutant to bind to P3OPT and P1/2 was monitored by EMSA. Results were quantitated by densitometry and are expressed as a fraction of protein bound to DNA, expressed as a percentage (Figure 5). For the PD single Cys mutants, only the C82 mutant demonstrated NEM sensitivity to PD binding sites (P3OPT). For the HD, mutants V263C, V265C, and S268C showed NEM sensitivity for binding to the P1/2 site, while mutants S222C and W274C were NEM insensitive. Importantly, NEM treatment of PD mutant C82 not only abrogated DNA binding to PD sites P3OPT, but also impaired binding of this mutant to the HD sequence P1/2. Conversely, NEM treatment of the HD mutant V263C not only impaired DNA binding to HD site P1/2, but also strongly reduced binding of this mutant to a PD site (P3OPT) (Figure 5). The NEM sensitivity of PD and HD binding seen in V263C was specific and distinct from that of adjacent mutants V265C and S268C, which displayed NEM sensitive binding to HD site but NEM-insensitive binding to PD site. Finally, although the severe reduction of P1/2 binding displayed by mutants R221C, R270C, and A272C (Figure 4B; see above) precluded the testing of NEM sensitivity toward this sequence (data not shown), NEM treatment did not affect DNA binding of these mutants to the PD site P3OPT. For all mutants studied, identical results were obtained when another sulfhydryl reagent, dibromobimane, was used (data not shown).

Additional dose response studies with NEM concentrations ranging from 0 to 3.6 mM were conducted on single Cys mutants showing NEM sensitivity for DNA binding to PD and HD target sites. Primary EMSA results are shown in Figure 6A and are plotted in Figure 6B after quantitation of the radioactivity in the gel retarded complex. In these studies, the effect of NEM on possible dimerization of C82 on a P2 site was also investigated. For the C82 mutant, NEM concentrations >0.1 mM were sufficient to impair binding to P3OPT and

Figure 5

Figure 5 Effects of *N*-ethyl maleimide treatment on the PD and HD DNA binding properties of Pax3 mutants bearing single cysteine replacements. The effect of NEM treatment of whole cell extracts from COS7 transfected cells on the DNA binding properties of individual mutants was tested for the PD sequence P3OPT and for the HD sequence P1/2, as described in the legend to Figure 2. The signals were quantitated by densitometry from a minimum of three independent experiments, and results are expressed as a ratio (% activity) of the amount of radioactivity present in the DNA/protein complex for each mutant in the absence and presence of prior incubation of the extract with 1 mM NEM.


P3OPT
Activity (%)
140
120
100
80
60
40
20
0
WT
CL
CL/C82
CL/C88
CL/S152C
CL/S153C
CL/S155C
CL/R221C
CL/S222C
CL/V263C
CL/V265C
CL/S268C
CL/R270C
CL/A272C
CL/W274C


P1/2
Activity (%)
140
120
100
80
60
40
20
0
WT
CL
CL/C82
CL/C88
CL/S152C
CL/S153C
CL/S155C
CL/S222C
CL/V263C
CL/V265C
CL/S268C
CL/W274C

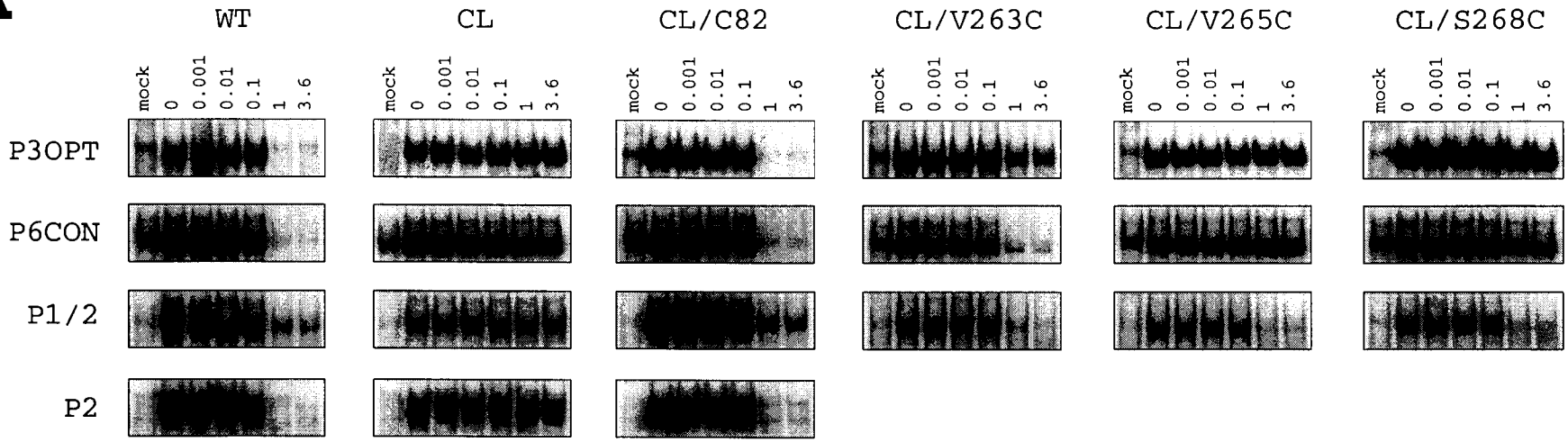
P6CON sequences (IC_{50} ~ 0.35 mM; Figure 6B), but also to P1/2 (IC_{50} ~ 0.4 mM; Figure 6B), with a residual amount of NEM insensitive binding seen for that site (Figure 6A). As expected, NEM also impaired homodimerization of C82 on a P2 HD site (Figure 6A). For the V263C mutant, half-maximal inhibition for binding to PD sites (P3OPT, P6CON) and the HD site (P1/2) were obtained and in both cases were ~0.4 mM NEM (Figure 6A,B). Finally, in both the HD mutants V265C and S268C, although NEM concentrations >0.1 mM were sufficient to abrogate binding to P1/2, binding to PD sites in these mutants was insensitive to [NEM] of up to 3.6 mM (Figure 6A), confirming uniqueness and specificity of NEM sensitivity seen in V263C and C82.

Together, these results demonstrate a reciprocal effect of site-specific modifications of single Cys mutants in the PD and HD of Pax3 on DNA binding by the modified and unmodified site. They support a model in which the PD and HD do not function independently in the mature protein, but rather functionally interact to regulate DNA binding.

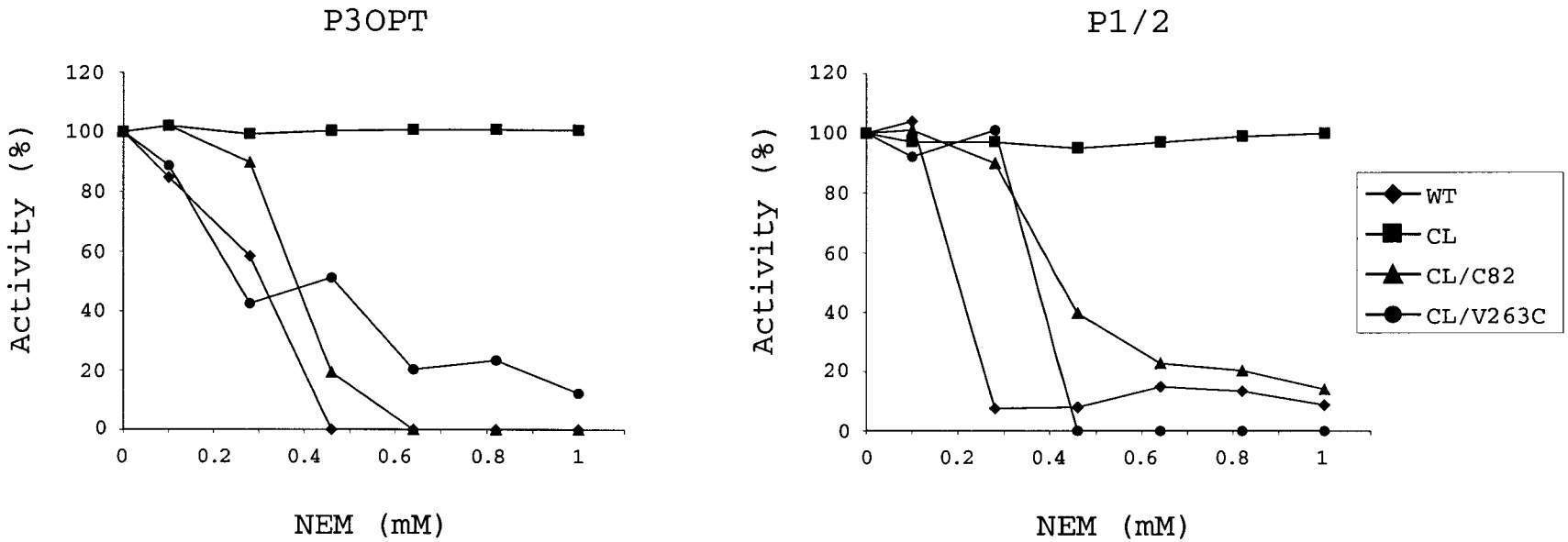
Figure 6

Figure 6 Dose-dependent effect of *N*-ethyl maleimide on DNA binding properties of certain single Cys mutants. (A) The NEM-sensitive DNA-binding properties of single Cys mutants in the PD (C82) and in the HD (V263C, V265C, S268C) were investigated in dose-response experiments as described in the legend to Figure 2 and for NEM concentrations ranging between 0.001 and 3.6 mM. NEM effects on DNA binding by the PD were tested using P3OPT and P6CON, while HD DNA binding was measured with P2 (for WT, CL, and C82) and with P1/2 (for all mutants). Gels were exposed for similar periods of time for mutants and WT protein extracts. (B) The signals were quantitated, and results are expressed as a ratio (% activity) of the amount of radioactivity present in the DNA/protein complex for each mutant in the absence and presence of prior incubation of the extract with the indicated concentration of NEM.

A


WT
CL
CL/C82
CL/V263C
CL/V265C
CL/S268C
mock
0
0.001
0.01
0.1
1
3.6
P3OPT
P6CON
P1/2
P2

B


P3OPT
P1/2
Activity (%)
NEM (mM)
0
20
40
60
80
100
120
0.2
0.4
0.6
0.8
1
WT
CL
CL/C82
CL/V263C

Discussion

We have created and functionally characterized a Pax3 mutant devoid of cysteine residues. One of the challenges of using a Cys-less mutant is that ideally, removal of the Cys residues should not affect function, so that reinsertion of single cysteines can be done in the context of an active protein. At the time of initiating these studies, little information was available on the role of cysteines in Pax proteins. In a study of the isolated PD from Pax5 and Pax8, it was observed that the 3 Cys of the PD (also conserved in Pax3) must be reduced for PD binding to DNA (Tell et al., 1998). Trypsin digestion and mass spectroscopy analysis were used to demonstrate that in the oxidized state a reversible disulfide bridge is formed between Cys37 and Cys49 (Cys70, Cys82 in this study) which inhibits DNA binding by the PD (Tell et al., 1998). In normal nuclear extracts, it was shown that the PD of Pax5 and Pax8 are kept in a DNA-binding competent state by the reducing activity of the Ref-1 protein (Tell et al., 1998). These results established that at the very least, eliminating intramolecular disulfide bonds formation in Pax3 should not be detrimental to DNA binding. The CL-Pax3 mutant created herein could bind to PD and HD target site sequences with characteristics similar to that of wild type Pax3, and this binding was resistant to NEM concentrations of up to 3.6 mM. In the wild type Pax3, NEM blocks DNA binding by the PD and HD at concentrations >0.3 mM. These results show that none of the Cys residues in Pax3 are essential for DNA binding and that it is the introduction of a bulky maleimide group as opposed to removal of a functional thiol moiety that accounts for NEM inhibition of wild type Pax3 activity. The differential effect of NEM on WT and CL-Pax3 indicates that over the concentration range tested, NEM appears to specifically modify cysteines and does not seem to alter other residues important for DNA binding. Second, the observation that in wild-type Pax3, HD DNA binding is NEM-sensitive despite the absence of Cys residues in this domain suggests a possible role of the Cys-containing PD on HD activity. Finally, the observations

that NEM inhibition of Pax3 DNA binding occurs both over a narrow range and at the same concentration for PD and HD targets suggest that alkylation of one or few critical Cys may be responsible for the effect. In purified PD from Pax5/Pax8, Cys82 appears to be most sensitive to reduction (Tell et al., 1998) and is a good candidate for the NEM effect seen in WT Pax3. The unique sensitivity to NEM of the single cysteine mutant C82 also agrees with results of this study.

Functional studies of Pax3 proteins carrying PD mutations found in certain naturally occurring mutant alleles of the *Splotch* mouse *(Sp^d)* (Underhill et al., 1995) or in certain Waardenburg Syndrome patients (Fortin et al., 1997) initially suggested possible regulation of HD DNA binding by the PD. Here we used the CL-Pax3 mutant to further investigate possible functional interdependence of the 2 DNA-binding domains of Pax3. In particular, we wished to determine whether NEM sensitivity of one of the two domains caused by alkylation of a single Cys inserted in this domain would also result in NEM sensitivity in the other, Cys-free domain, and vice-versa. In these studies, we wanted to avoid introducing mutations that would grossly affect the integrity of the DNA binding domains, but rather, we wanted to insert single Cys at neutral positions in close proximity to the DNA molecule in the DNA-bound complex. Alkylation at such positions could impair DNA binding, thus providing a convenient method of ascertaining accessibility of the inserted Cys to thiol reagents. We focused our Cys insertion sites on the α3 helix of each of the HTH motifs of the PD and HD, a helix known to play a key role in DNA binding. In the PD, independent reinsertion of Cys 82 and 88 of PAI was as expected without consequence on PD and HD activity. Likewise, three independent insertions in the RED domain had no effect on DNA binding. In the HD, two insertions were in the N-terminal arm and six were in the α3 helix. Interestingly, 3/8 mutations impaired partly or completely monomeric binding to the P1/2 half site, while 7/8 mutations abolished cooperative dimerization on P2, but none had an effect on PD binding

properties. Thus, dimerization on P2 seemed more mutation sensitive than monomeric binding to P1/2 for the mutant set, possibly reflecting either a threshold effect for dimerization, reduced affinity for P2 compared to P1/2 or both. Therefore, subsequent studies were limited to the five mutants having retained significant monomeric binding to P1/2.

The effect of alkylation on the DNA binding properties of each mutant was tested in dose response experiments for both NEM and dibromobimane with identical results. In the experimental scheme used, we did not ascertain the NEM accessibility of each single Cys insertion. NEM accessibility can be inferred only for insertions that imparted NEM sensitivity to DNA binding by the targeted domain, and we will limit our discussion to these positions. NEM-insensitive insertions could reflect lack of effect of alkylation on DNA binding, or inaccessibility of the residue to NEM and will not be discussed further. As described above, WT Pax3 binding by the PD (P3OPT, P6CON) and by the HD (P2) was NEM sensitive. Of note was a small amount of NEM-insensitive binding to P1/2 (10-20%), detected even at the highest NEM concentration tested. Interestingly, this residual binding was not seen in single Cys HD mutants V263C, V265C, or S268C modified by alkylation. This suggests a small amount of NEM-independent monomeric binding to P1/2 by the HD of Pax3 (as expected from its lack of cysteine), which can be abrogated by sulfhydryl modification of a single Cys insertion in helix 3 of this domain.

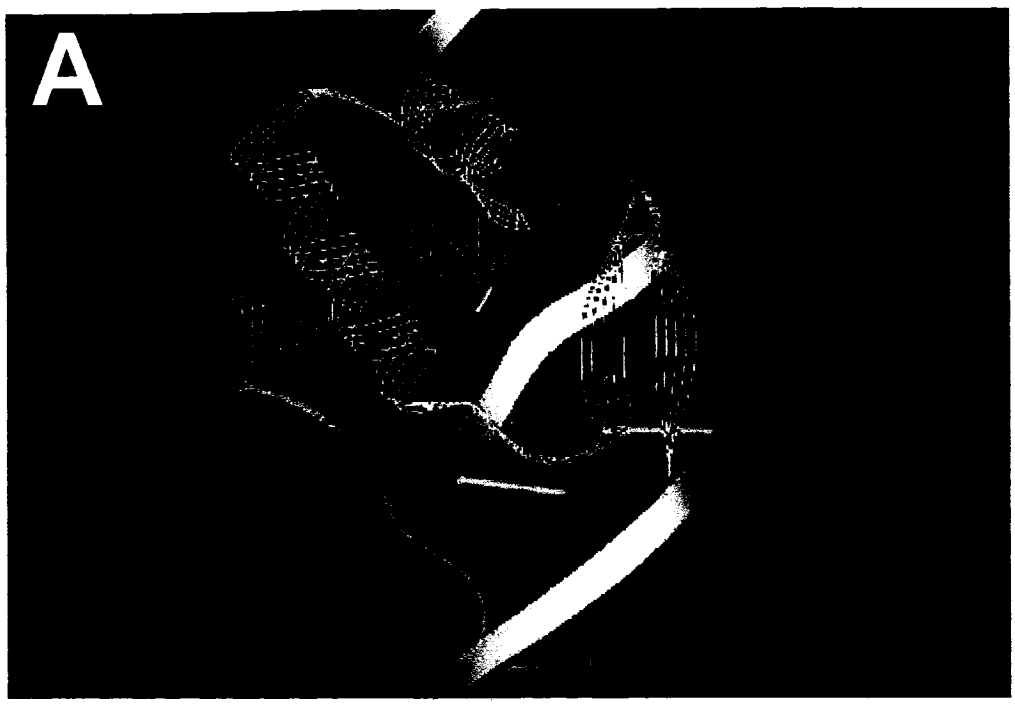
In the PD, NEM alkylation of Cys82 impaired DNA binding by the PD. As shown in Figure 7, examination of Cys82 in the published structures of DNA-bound Pax6 and in Prd indicates that this residue makes phosphate and base specific contacts in the major groove. Thus, it is not surprising that modification by a bulky maleimide group would affect DNA binding. Surprisingly, the NEM sensitivity of PD DNA binding imparted by Cys82 also caused NEM sensitivity for DNA binding by the otherwise unmodified HD. Thus, a small

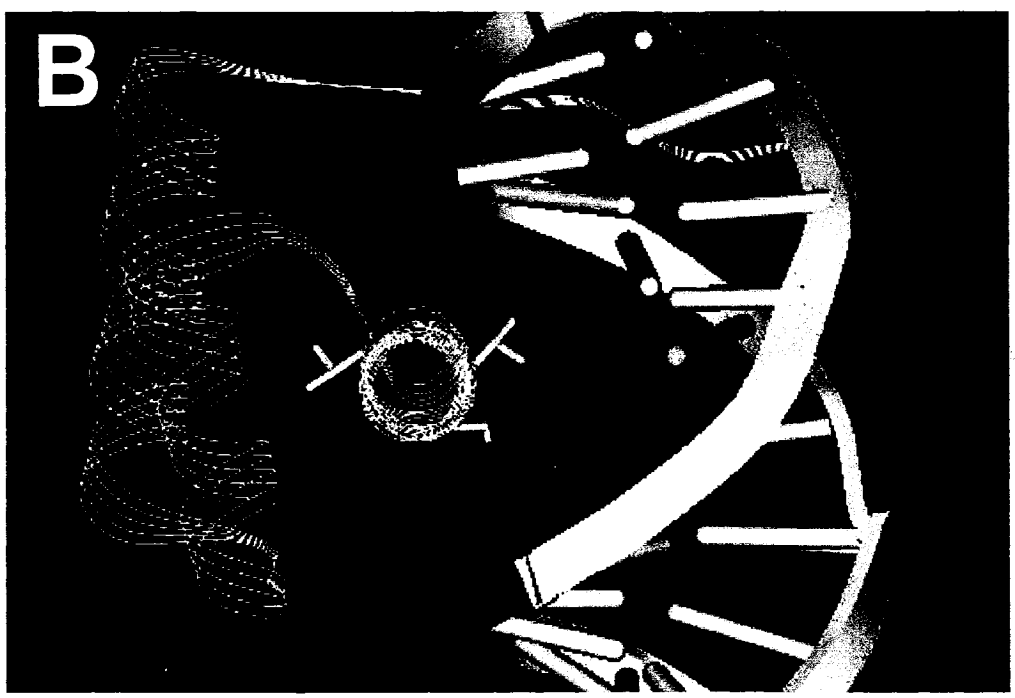
structural change in the PD abrogates the ability of the HD to bind DNA, again suggesting functional interdependence in the form of regulatory interactions between the two domains in the wild type protein, and possibly interrupted by NEM modification of Cys82. Interestingly, WSI mutations have been identified at the flanking Gly81 (G81A) and Ser84 (S84F) (Baldwin et al., 1995) positions. These mutations do not affect DNA binding by either the PD or the HD (Fortin et al., 1997), suggesting that another molecular aspect of Pax3 function may be impaired in these mutants.

In the HD, alkylation of V263C had similar effects to those seen for Cys82 in the PD, in that it not only impaired DNA binding by the HD but also abrogated DNA-binding by the Cys-less PD present in this mutant. Although Val263 does not make DNA contacts per se, it is tucked between R262 and V265 residues in α3 helix that make phosphate contacts and base-specific contact in the major groove, respectively. Thereby, sulfhydryl modification at V263C may adversely affect important DNA contacts made by neighboring residues, possibly uncoupling DNA binding by the HD. Surprisingly, NEM modification of V263C also impaired DNA binding by the PD, therefore mirroring in a reciprocal fashion the effect seen for Cys82 in the PD. The transfer of NEM sensitivity in DNA binding by the PD seen for V263C is specific for this residue and cannot be explained by either (a) a nonspecific disruption of the HD or (b) a general effect of loss of HD DNA binding on PD binding. Indeed, alkylation of downstream positions in α3 in mutants V265C and S268C although impairing DNA binding by the HD was without consequences on PD DNA binding, demonstrating that it is possible to impair DNA binding by the HD without affecting PD activity in the same molecule. Interestingly, the loss of HD DNA binding while retaining PD binding seen after NEM modification of V265C is identical to the phenotype of a WS1 mutation (V265F) previously reported at that position (Fortin et al., 1997). Close examination of the structures of the HD bound to DNA suggests a possible mechanistic basis for the

Figure 7

Figure 7 DNA-bound structure of the N-terminal subdomain of the PD and of the HD. The N-terminal subdomain of the PD (A) and the HD (B) are shown as green ribbons drawn through the xcarbon backbone. The DNA strands are shown as gray ribbons through the sugar phosphate backbone, and bases are shown as protrusions from the ribbons. The position of the endogenous PD cysteine C82 is indicated (A) and faces the major groove. (B) The HD is shown bound to DNA as a monomer. The positions of V263, V265, and S268 residues of the HD indicate that V265 and S268 face DNA while V263 is on the opposite face of the helix.


A


B

distinct effect of alkylation of V265C/S268C and V263C on DNA binding by the PD and HD (Figure 7). V265 and S268 make key base-specific contacts in the major groove separated by approximately one turn of the α3 helix (Wilson et al., 1995). On the other hand, V263 is solvent-exposed and on the other side of the helix in close proximity (tightly packed) to helices α1 and α2 (Figure 7). Therefore, it is tempting to speculate that introduction of a large maleimide group at V263C may not only affect the binding of α3 to DNA but may also alter the packing of the two other helices (α1, α2) of the HD; this latter disruption may have an allosteric effect on the PD activity, an effect not seen after modification of either V265C or S268C. Studies of the Phox-1 protein have previously shown that its HD is responsible for the recruitment and physical interaction with the nuclear factor SRF (Grueneberg et al., 1992). This property is specific to HDs of the paired class (Pax), is independent of the DNA binding specificity of the HD, and is mediated by pairs of charged residues on the solvent exposed face of helices 1 and 2 of the HD (Grueneberg et al., 1995). Such residues are conserved in Pax3 PD, and they may participate in interactions between the HD and PD which may be disturbed by alkylation of V263C. Additional Cys-scanning mutagenesis in helices of the HTH motif will be required to further characterize this possible allosteric effect.

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 the full length Pax3 protein. Rather, it appears that both domains may functionally interact, such that one domain can modulate the DNA binding properties of the other domain. These interactions may be critical for target site selection by the Pax3 protein.

Acknowledgements

The authors are indebted to Dr. K. Gehring (Department Biochemistry, McGill University) and H. R. Kaback (UCLA) for constructive discussions and helpful suggestions during this work.

This study was supported by grants from the Canadian Institutes for Health Research (CIHR) of Canada to P.G. S.A. is supported by a Doctoral studentship from the Fonds de Recherches en Sante du Quebec and P.G. by a Distinguished Scientist Award from the CIHR.

The study discussed in Chapter 2 is a successful "re-enactment" of observations made with naturally occurring Pax3 proteins that abrogate DNA binding activities of the paired and HD. Alkylation of Cys82 with NEM was meant to abrogate PD DNA binding activity in order to determine the effect of an impaired PD on HD DNA binding activity. Like the naturally occurring Splotch-delayed G42R PD mutation, NEM alkylated Cys82 abrogates Pax3 PD and HD DNA binding activities (Underhill et al., 1995). This observation confirms what has been suggested by the characterizations of several naturally occurring PD mutations; that is, the HD DNA binding activity is influenced by the PD. Likewise, alkylation of V263C with NEM was meant to abrogate HD DNA binding activity in order to assess the effect of abrogated HD activity on PD DNA binding activity. Like the Waardenburg syndrome HD mutant R53G, NEM alkylated V263C does not retain HD nor PD DNA binding activities (Fortin et al., 1997). This discovery confirms what has been suggested by the characterizations of other naturally occurring HD mutations: the PD DNA binding activity is influenced by the HD.

To account for the observations made with the alkylated single cysteine mutants we postulate that DNA binding of one domain is sensed by or communicated to the other domain. For example, PD DNA binding may induce a conformational change in the PD as well as the HD and both these conformational changes might be required for proper Pax3 DNA binding by the PD. We propose a similar situation exists for the HD as well.

To study possible conformational changes of both domains when Pax3 is DNA bound by either domain a protease sensitivity approach was implemented. This study required the use of Pax3 mutants with engineered Factor Xa protease cleavage sites and is presented in chapter 3.

Chapter 3

Cross-talk between the Paired Domain and the Homeodomain of Pax3: DNA Binding by each Domain causes a Structural Change in the other Domain, Supporting Interdependence for DNA Binding

Abstract

The Pax3 protein has two DNA binding domains, a paired domain (PD) and a paired-type homeodomain (HD). 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. The mechanistic basis of this functional interdependence is unknown and was studied by protease sensitivity. Pax3 was modified by the creation of Factor Xa cleavage sites at discrete locations in the PD, the HD, and in the linker segment joining the PD and the HD (Xa172, Xa189, and Xa216) in individual Pax3 mutants. The effect of Factor Xa insertions on protein stability and on DNA binding by the PD and the HD was measured using specific target site sequences. Independent insertions at position 100 in the linker separating the first from the second helix-turn-helix motif of the PD and at position 216 immediately upstream of the HD were found to be readily accessible to Factor Xa cleavage. The effect of DNA binding by the PD or the HD on accessibility of Factor Xa sites inserted in the same or in the other domain was monitored and quantitated for multiple mutants bearing different numbers of Xa sites at each position. In general, DNA binding reduced accessibility of all sites, suggesting a more compact and less solvent-exposed structure of DNA-bound *versus* DNA-free Pax3. Results of dose response and time course experiments were consistent and showed that DNA binding by the PD not only caused a local structural change in the PD but also caused a conformational change in the HD (P3OPT binding to Xa216 mutants); similarly, DNA binding by the HD also caused a conformational change in the PD (P2 binding to Xa100 mutants). These results provide a structural basis for the functional interdependence of the two DNA binding domains of Pax3.

Introduction

Pax3 is a member of a family of 9 transcription factors (Stuart et al., 1994), defined by a DNA binding module, the PD, that was first identified in the *Drosophila* protein Paired (Prd) (Bopp et al., 1986). Pax proteins play critical roles during normal embryonic development, and inactivating mutations cause major defects in development of skeleton, muscles, nervous system, eyes, kidneys, and the immune system (Dahl et al., 1997; Balling et al., 1988; Sanyanusin et al., 1995; Torres et al., 1995; Jordan et al., 1992; Glaser et al., 1992; Hill et al., 1991; Sosa-Pineda et al., 1997; Mansouri et al., 1998; Macchia et al., 1998). Pax3 is expressed in developing somites and in neural tube and neural crest cell derivatives and plays a role in the proliferation, migration, and differentiation of cells involved in neurogenesis and myogenesis (Goulding et al., 1991; Bober et al., 1994). The *Pax3* mouse mutant (*splotch, Sp*) displays spina bifida and exencephaly and lacks limb muscles; likewise, mutations in human *PAX3* cause Waardenburg syndrome (WS), a condition characterized by pigmentary disturbances, craniofacial abnormalities and sensorineuronal deafness (Bober et al., 1994; Beechey et al., 1986; Auerbach et al., 1954; Franz et al., 1989; Franz et al., 1990; Goulding et al., 1994; Baldwin et al., 1992; Baldwin et al., 1995). PAX3 also plays a role in cell transformation as shown by the translocation t(2;13) (q35;q14) involving PAX3 (Barr et al., 1993) and forkhead-related transcription factor (Galili et al., 1993) associated with the solid tumor alveolar rhabdomyosarcoma. The translocation leads to the expression of a fusion protein containing the N-terminal DNA binding domain of PAX3 and the C-terminal activation domain of forkhead-related transcription factor (Galili et al., 1993).

Several members of the Pax family, including Pax3, encode a second DNA binding domain, the paired-type HD (Stuart et al., 1994; Noll et al., 1993). Also conserved in the Pax family is the presence of an octapeptide (OP) motif in the segment linking the PD and the HD,

as well as a proline-serine-threonine (PST)-rich C-terminal domain. Both the OP and PST domain are involved in protein-protein interactions to recruit additional transcription factors required for transcription of target genes. The three-dimensional structures of the DNA-bound PD of Prd (Xu et al., 1995) and Pax6 (Xu et al., 1999) reveal that it is bipartite consisting of an N-terminal (PAI) and a C-terminal (RED) subdomains, each formed by a three-helical fold with the two most C-terminal helices forming a helix-turn-helix (HTH) motif that makes DNA-specific contacts in the major groove of DNA. The sublinker, connecting the two subdomains of the PD, adopts an extended conformation on DNA and makes base contacts in the minor groove. Unique to the PAI subdomain is the β-turn and β-hairpin, which precede the three-helical fold and participate in DNA binding (Xu et al., 1995; Xu et al., 1999). The DNA sequences recognized by the PD are of two classes: binding to Class I sequences requires both the PAI and RED subdomains, whereas binding to class II sequences only requires the PAI subdomain (Epstein et al., 1994; Czerny et al., 1993). Some isoforms of Pax6 and Pax8 produce, by alternate splicing, isoforms that bind PD targets exclusively through the RED subdomain (Epstein et al., 1994; Kozmik et al., 1997).

The crystal structure of the DNA-bound form of the paired-type HD (pt-HD) consists of a three-helical fold containing a HTH motif; helix3 makes specific contacts in the major groove of DNA and confers DNA binding specificity of the HD (Kappen et al., 1993; Wilson et al., 1993; Wilson et al., 1995; Treisman et al., 1989). Like other HDs, the extended N-terminal arm motif, preceding the three-helical fold, is also used to make specific DNA contacts but in the minor groove of DNA (Wilson et al., 1993; Wilson et al., 1995; Treisman et al., 1989). The pt-HD class of HDs can bind the TAAT motif but can also uniquely dimerize on palindromic sequences of the type TAAT(N_{2-3})ATTA (Wilson et al., 1993; Schafer et al., 1994). The identity of residue 50 in helix3 determines both DNA specificity and dimerization potential in this class of HDs (Wilson et al., 1993; Schafer et al., 1994). When

covalently linked to a PD the pt-HD of Pax proteins (Ser-50) only permits the dimerization of the HD on palindromic sequences with a two nucleotide spacer (Wilson et al., 1993; Schafer et al., 1994; Underhill and Gros, 1997).

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 the intact Pax3 protein. The *Splotch-delayed* (Sp^d) mouse mutant bears a single G42R mutation in the PD, which abrogates DNA binding by the PD but also impairs DNA binding by the HD. Deletion of helix 2 of the PAI subdomain in the context of the Sp^d mutation has been shown to restore HD DNA binding (Underhill et al., 1995; Fortin et al., 1998). Studies in chimeric PAX3 proteins have shown that the PD can modulate DNA binding specificity and dimerization potential of heterologous HDs (Fortin et al., 1997; Underhill and Gros, 1997). On the other hand, a mutant PAX3 variant from a WS patient bearing a mutation at position 53 of the HD (R53G) shows not only loss of DNA binding by the HD but also by the PD (Fortin et al., 1997). More recent biochemical studies by cysteine scanning mutagenesis and site-specific modification of single cysteine mutants with sulfhydryl reagents have shown that modification of a single cysteine in the PD (Cys-82) disables DNA binding by the PD but also by the HD (Apuzzo et al., 2002). Conversely, modification of a single cysteine at position 263 of the HD (V263C) of the HD abrogates DNA binding by both domains (Apuzzo et al., 2002).

The mechanistic basis of this functional interdependence, including the protein subdomains involved, remains poorly understood but is likely to be relevant for target site selection by Pax3 *in vivo*. One plausible mechanism is that DNA binding by one or both of the DNA binding sites of Pax3 causes conformational changes at or near the other binding site to alter its properties. Thus, we wanted to get insight into the conformations adopted by the Pax3

protein when DNA-free and when bound to PD or HD targets. A number of physicochemical approaches have been used to monitor the effect of substrate binding on protein conformation, including differential immunoreactivity with specific antibodies (Mechetner et al., 1997), tryptophan fluorescence (Weber et al., 1993; Weber et al., 1997; Zhou et al., 1997; Walmsley et al., 1999; Menezes et al., 1990), and site-specific modification of cysteine residues (Liu et al., 1997; Liu et al., 1996). Protease sensitivity has also been used extensively to monitor conformational changes in proteins (Gomes et al., 1996; Stout et al., 1998; Rothman et al., 1997; Morsomme et al., 1998; Zhang et al., 1998). Although partial proteolytic digestion with enzymes such as trypsin, chymotrypsin, and papain has been used in such studies, delineating the cleavage sites is complicated by the necessity to identify proteolytic fragments by epitope mapping with specific antibodies and/or peptide sequencing (Julien et al., 2000). Another implementation of this method involves creating recombinant proteins bearing single heterologous protease cleavage sites (such as Factor Xa) inserted at pre-determined positions in individual mutants. Proteolytic products can be identified using antibodies against antigenic epitope also engineered at convenient positions. In this approach, conformational changes can be studied in a set of recombinant proteins for which structural and functional integrity has been ascertained.

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 (Underhill et al., 1995; Apuzzo et al., 2002). This pMT2/Pax3 construct encodes for all 479 amino acids of the murine Q+ isoform of Pax3 (Vogan et al., 1996). 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 polyhistidine tail (His_6) and 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'-CTCGAATTCATGGAGCAGAAGTTAATCAGCGAAGAGGATCTCACCACGCTGGCCGGCGCTGTGCCCAGGATG-3') and P3-HA (5'-TTTAGCGGATCCGAATTCTTAGTGATGGTGGTGATGGTGTCCCGCGGCGTAATCTGGAACGTCATATGGATATCCGAACGTCCAAGGCTTACTTTG-3'). Both primers were engineered with EcoRI restriction sites at their ends. The resulting 1.5-kb PCR product was digested with EcoRI and ligated into the corresponding site of mammalian expression vector pMT2, and the resulting construct was designated pMT2/Myc-Pax3-HA.

Factor Xa cleavage sites (IE/DGR, Table I) were introduced at different positions in Pax3 by two procedures. The wild type sequence was mutated (without addition of extra amino acid residues) to convert near matches to IE/DGR via PCR-mediated mutagenesis. Individual mutants were created in two independent reactions using complementary oligonucleotide pairs defining each mutation (listed in Table II) as well as P3-Myc and P3-HA; mutated DNA fragments were annealed and repaired, and the full-length cDNAs were

synthesized. Mutants Pax3Xa55, Pax3Xa71, Pax3Xa114, Pax3Xa131, Pax3Xa252, and Pax3Xa259 were created in this fashion and were introduced in pMT2. Mutants Pax3Xa66,

Table 1: Factor Xa cleavage sites introduced in individual Pax3 mutants[a]

Xa Mutant	Wild type sequence	Xa Mutant Sequence
Xa55[b]	IRHK	**IEGR**
Xa71	VISR	**IEGR**
Xa114	VEKK	**IEGR**
Xa131	WEIR	**IEGR**
Xa252	LAQR	**IDGR**
Xa259	TEAR	**IEGR**
Xa66(1)[c,d]	HG▼IR	HG**SIEGRG**IR
Xa66(2)	HGIR	HG**SIEGRASIIEGRG**IR
Xa100(2)	PG▼AI	PGA**GIEGRGAGIEGRG**AI
Xa100(4)	PGAI	PGA**GIEGRGAGIEGRGAGIEGRGAGIEGRG**AI
Xa172(1)	AD▼LE	ADL**DIEGRL**E
Xa172(3)	ADLE	ADL**DIEGRLDIEGRLDIEGRL**E
Xa189(1)	ID▼GI	ID**IEGRIDG**I
Xa189(2)	IDGI	ID**IEGRIDGRIDG**I
Xa216(1)	LK▼RK	LK**IDGRSIKRK**
Xa216(2)	LKRK	LK**IDGRIDGRSIRK**

[a]Introduced amino acids are underlined and engineered Xa protease sites are indicated in bold.

[b]Mutants are designated according to the Pax3 amino acid position at which individual Factor Xa cleavage sites were introduced by site-specific modification of wildtype sequence.

[c]These mutants were created by insertion mutagenesis at the position indicated by an arrow.

[d]The number of Factor Xa sites introduced at that position is indicated in parenthesis.

Table 2: Oligonucleotides used for Site-directed[a] and Insertion Mutagenesis

Xa55	F	GGAGTATTTATCAACGGCCGGCCGCTGCCCAACCATATC**GAAGGCAGAA**TAGTGGAGATGGCCCAC
	R	GTGGGCCATCTCCACTAT**TCTGCCTTC**GATATGGTTGGGCAGCGGCCGGCCGTTGATAAATACTCC
Xa71	F	GGCATTCGGCCTTGC**ATTGAAGG**TCGACAGCTTCGCGTGTCC
	R	GGACACGCGAAGCTGTCGA**CCTTCAAT**GCAAGGCCGAATGCC
Xa114	F	AAACCCAAGCAGGTGACAACTCCGGAC**ATCGAGGGACGT**ATTGAGGAATACAAAAG
	R	CTTTTGTATTCCTCAAT**ACGTCC**CTC**GAT**GTCCGGAGTTGTCACCTGCTTGGGTTT
Xa131	F	AACCCGGGCATGTTTTCAATTGAAGGCAGAGACAAATTGCTCAAGGAC
	R	GTCCTTGAGCAATTTGTCTCTGCCTTCAATTGAAAACATGCCCGGGTT
Xa252	F	ATTTACACCAGGGAGGAGATCGACGGCCGGGCGAAGCTTACCGAGGCC
	R	GGCCTCGGTAAGCTTCGCCCGGCCGTCGATCTCCTCCCTGGTGTAAAT
Xa259	F	GCCCAGAGGGCGAAGCTTATCGAGGGCCGTGTACAGGTCTGGTTTAGCAAC
	R	GTTGCTAAACCAGACCTGTACACGGCCCTCGATAAGCTTCGCCCTCTGGGC
KasI(100)[b]	F	GGCTCCATCCGACCTGGCGCCATCGGCGGCAGC
	R	GCTGCCGCCGATGGCGCCAGGTCGGATGGAGCC
AflII(216)[c]	F	TCTGAACCTGATTTACCGCTTAAGAGGAAGCAGCGCAGG
	R	CCTGCGCTGCTTCCTCTTAAGCGGTAAATCAGGTTCAGA
Xa66(1)[d]		ATCCATAGAAGGTAGAGG
Xa66(2)		ATCCATAGAAGGTAGAGGATCCATAGAAGGTAGAGG
Xa100(2)		GCGCCGGCATAGAAGGTAGAGGCGCCGGCATAGAAGGTAGAG
Xa100(4)		GCGCCGGCATAGAAGGTAGAGGCGCCGGCATAGAAGGTAGAGGCGCCGGCATAGAAGGTAGAG
Xa172(1)		CTAGATATCGAAGGTCGT
Xa172(3)		CTAGATATCGAAGGTCGTCTAGATATCGAAGGTCGTCTAGATATCGAAGGTCGT
Xa189(1)		CGATATCGAAGGTAGAAT
Xa189(2)		CGATATCGAAGGTAGAATAGACGGCCGAAT
Xa216(1)		TTAAGATCGACGGTAGATCTA
Xa216(2)		TTAAGATCGACGGTAGAATAGACGGCCGATCTA

[a]Nucleotide substitutions leading to amino acid changes are indicated in bold, and those that introduce silent restriction sites are underlined.

[b]These oligos were used to introduce a unique KasI restriction site at position 595 of the wildtype Pax3 cDNA sequence.

[c]These oligos were used to introduce a silent and uniqe AflII restriction site at position 940 of the wildtype Pax3 cDNA sequence.

[d]Oligonucleotide sequences used (double stranded) for insertion mutagenesis in the creation of the corresponding Xa mutants as listed in Table 2.

Pax3Xa172, and Pax3Xa189 were created by insertion mutagenesis. For this, pMT2/Myc-Pax3-HA plasmid was digested with BsmI, XbaI, or ClaI, and single or multiple Xa cleavage sites were introduced using double-stranded oligonucleotides with cohesive ends (Tables I and II). These oligonucleotides have a sequence just long enough to encode for one or a few Xa protease sites when placed in-frame with the rest of the Pax3 encoding region. For mutants Pax3Xa100 and Pax3Xa216, the *Pax3* cDNA was modified to introduce unique KasI and AflII sites at nucleotide positions 595 and 940, respectively, using mutagenic primers listed in Table II. These modified *Pax3* cDNAs were subcloned into the EcoRI site of pBluescript (lacks KasI and AflII sites), and one or several Factor Xa sites were independently introduced at the KasI or AflII sites by insertion mutagenesis to create mutants Xa100 (2), Xa100 (4), Xa216 (1), and Xa216 (2) (Tables I and II). In all cases, the presence of the Factor Xa mutations and the integrity of the rest of the Pax3 sequences were verified by nucleotide sequencing. The accessibility of restriction sites used for cloning was verified by restriction enzyme 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 co-precipitation 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 phenylmethysulfonyl 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 Assay—Electrophoretic mobility shift assays were performed as previously described (Apuzzo et al., 2002). Each protein:DNA binding reaction was carried out using ~10 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. The final concentration of labeled oligonucleotide in the binding reaction was 0.5 nM. Whole cell extracts were incubated with ^{32}P-labeled PD-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 bovine serum albumin, and 10% glycerol. Whole cell extracts were also incubated with ^{32}P-labeled HD-specific probes in a volume of

20 µl containing 10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 1 mM DTT, 2 mM $MgCl_2$, 1 mM EDTA, and 5% glycerol. To reduce non-specific binding, 1 µg of poly(dI-dC)poly(dI-dC) was included in binding studies with PD-specific probes, whereas 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'-TGGAATTCAGGAAAAATTTTCACGCTTGAGTTCACAGCTCGAGTA-3') (Xu et al., 1995) and P3OPT (5'-TGGTGGTCACGCCTCATTGAATATTA-3') (Julien et al., 2000; Vogan et al., 1996; Epstein et al., 1996), and HD-specific sequences P2 (5'-GATCCTGAGTCTAATTGATTACTGTACAGG-3') (Wilson et al., 1993; Kozmik et al., 1997) and P1/2 (5'-GATCCTGAGTCTAATTGAGCGTCTGTAC-3') (Wilson et al., 1993; Kozmik et al., 1997) were synthesized as complementary oligonucleotide pairs and were designed to have recessed 3' ends for end labeling with [α-^{32}P]dATP (3000 Ci/mmol; PerkinElmer Life Sciences) using the Klenow fragment of DNA polymerase.

Factor Xa Treatment of Xa Mutants—For time course studies, 10 µg of total cell extracts from transiently transfected COS-7 monkey cells was incubated with or without double-stranded oligonucleotides (final concentration of 2 µM) corresponding to PD and HD binding sites, in a final volume of 10 µl. The whole cell extract was incubated with DNA for 30 min at 20 °C followed by addition of 0.2 µl of 1 µg/µl (0.2 µg) of Factor Xa protease (New England Biolabs). The proteolysis reaction was carried out for pre-determined periods of time (2-180 min) at 20 °C and was stopped by the addition of 5 µl of Laemmli sample buffer (2% (w/v) SDS, 10% (v/v) glycerol, 62.5 mM Tris-HCl, pH 6.8, 100 mM DTT, 0.05% bromophenol blue). For the dose-response assays, 10 µg of total cell extracts from transiently transfected

COS-7 monkey cells was incubated for 30 min at 20 °C with oligonucleotides (final concentration of 2 μM) corresponding to PD and HD binding sites or nonspecific oligonucleotides, in a final volume of 20 μl. Binding reactions done with PD probes (and with a nonspecific oligonucleotide) contained 10 mM Tris-HCl (pH 7.5), 50 mM KCl, 1 mM DTT, 2 mM spermidine, 2 mg/ml bovine serum albumin, 10% glycerol, and 1 μg of poly(dI-dC)poly(dI-dC). Binding reactions done with HD probes (and with a nonspecific oligonucleotide) contained 10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 1 mM DTT, 2 mM $MgCl_2$, 1 mM EDTA, 5% glycerol, and 2 μg of heat-inactivated salmon sperm DNA. Various concentrations of Factor Xa protease, ranging from 0 to 200 ng of Xa protease/microgram of whole cell extract, were then added to the reaction mixture, and proteolytic cleavage was allowed to take place for 15 min at 20 °C. The reaction was stopped by addition of 10 μl of Laemmli sample buffer. For both assays the proteolytic degradation products were separated by SDS-PAGE on 12% polyacrylamide gels, followed by transfer onto nitrocellulose membranes. Immunodetection of Pax3 products was carried out using anti-HA antibody followed by anti-Myc antibody as described above. Films were used to perform densitometry studies to quantify the amount of chemiluminescence using a Fuji LAS-1000.

Results

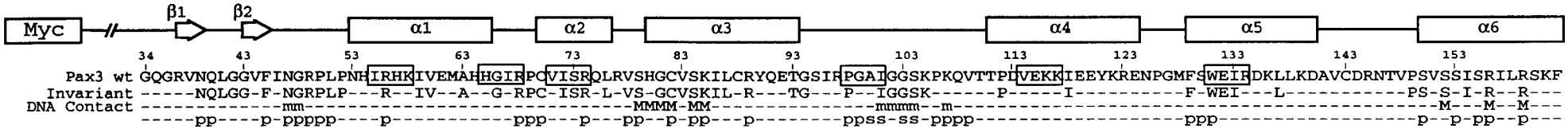
Construction of Pax3 Mutants Bearing Factor Xa Protease Cleavage Sites—A large body of biochemical and genetic data indicate that the PD and the HD of Pax3 are functionally interdependent with strong cooperativity between the two sites. The structural basis for this interdependence remains unclear and was investigated. A major objective of this study was to determine whether or not DNA binding at either the PD or HD of Pax3 causes a conformational change at the other site, thereby providing a possible structural basis for their reported functional interdependence. Possible conformational changes were assessed by protease sensitivity by monitoring accessibility of cleavage sites strategically inserted in the PD and HD of Pax3. A *Pax3* cDNA was modified by the in-frame addition of hemagglutinin (HA) and c-Myc epitope tags at the C and N termini of the protein, respectively (Fig. 1), to facilitate detection of specific proteolytic fragments by immunoblotting. This cDNA was used to insert Factor Xa cleavage sites (IEGR, IDGR; cleavage immediately C-terminal of R) by site-directed mutagenesis. Using the known three-dimensional structures of Pax6 and Paired as molecular templates, Factor Xa sites were introduced either in the PD, in the HD, in the flexible linker separating the PD and the HD, as well as in solvent-exposed segments near the end of individual helices in the HTH motifs of the PD and the HD (see Figs. 1 and 3). Two strategies were used for mutagenesis. First, near-matches in the wild type Pax3 sequence mapping to the highly conserved HTH modules of the PD and HD were independently converted to I(E/D)GR (Table I) to minimize adverse structural changes possibly impairing DNA binding. Targeted in this group of six mutants were helices 1 (α1, Xa55), 2 (α2, Xa71), 4 (α4, Xa114), and 5 (α5, Xa131) of the PD as well as the C terminus of helix 2 (Xa252) and the N terminus of helix 3 (Xa259) of the HD (Table I and Fig. 1). Second, and to maximize

accessibility to proteolytic cleavage, several solvent-exposed, and less conserved linker segments were also targeted for insertion of one or several

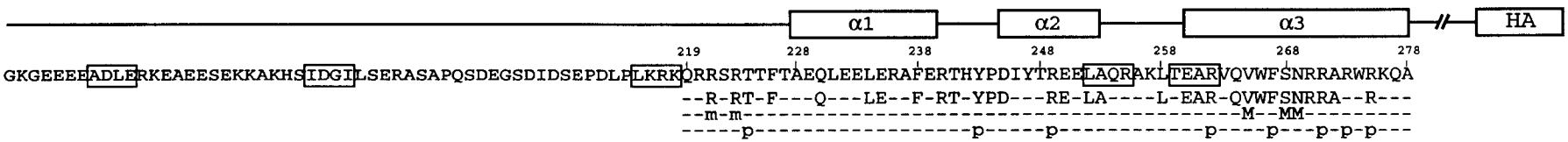
Figure 1

Insertion of Factor Xa protease sites in the PD and in the HD of Pax3. Schematic representation of the PD and HD, including position of α-helices and β-strands (*arrows*) deduced from the known structures of the *Drosophila* Prd, and mammalian Pax6 proteins (Balling et al., 1988; Xu et al., 1995; Xu et al., 1999). The amino acid sequence of Pax3 (positions 34-278) is shown, with invariant residues in members of the Pax family identified below. The type of DNA contacts made by these residues (*p*, phosphate; *m*, minor groove; *M*, major groove) is indicated. Pax3 sequences targeted for modification to create Factor Xa cleavage sites are *boxed*. 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.

PD


Myc
β1
β2
α1
α2
α3
α4
α5
α6
34 43 53 63 73 83 93 103 113 123 133 143 153
Pax3 wt GQGRVNQLGGVFINGRPLPNHIRHKIVEMAHHGIRPCVISRQLRVSHGCVSKILCRYQETGSIRPGAIGGSKPKQVTTPDVEKKIEEYKRENPGMFSWEIRDKLLKDAVCDRNTVPSVSSISRILRSKF
Invariant -----NQLGG-F-NGRPLP---R--IV--A--G-RPC-ISR-L-VS-GCVSKIL-R---TG---P--IGGSK------P-----I----------F-WEI---L----------PS-S-I-R--R---
DNA Contact ----------mm-----------------------------MMMM-MM----------mmmm--m---M---M--M---
-----pp----p-ppppp----p----------ppp---p---pp--p-pp---p-------ppss-ss-pppp-----------------ppp-----------------p--p-pp--p---

HD


α1
α2
α3
HA
219 228 238 248 258 268 278
GKGEEEEADLERKEAEESEKKAKHSIDGILSERASAPQSDEGSDIDSEPDLPLKRKQRRSRTTFTAEQLEELERAFERTHYPDIYTREELAQRAKLTEARVQVWFSNRRARWRKQA
--R-RT-F---Q---LE--F-RT-YPD---RE-LA----L-EAR-QVWFSNRRA--R---
--m-m---M--MM---------
-----p----------p-----p----------p----p---p-p-p---

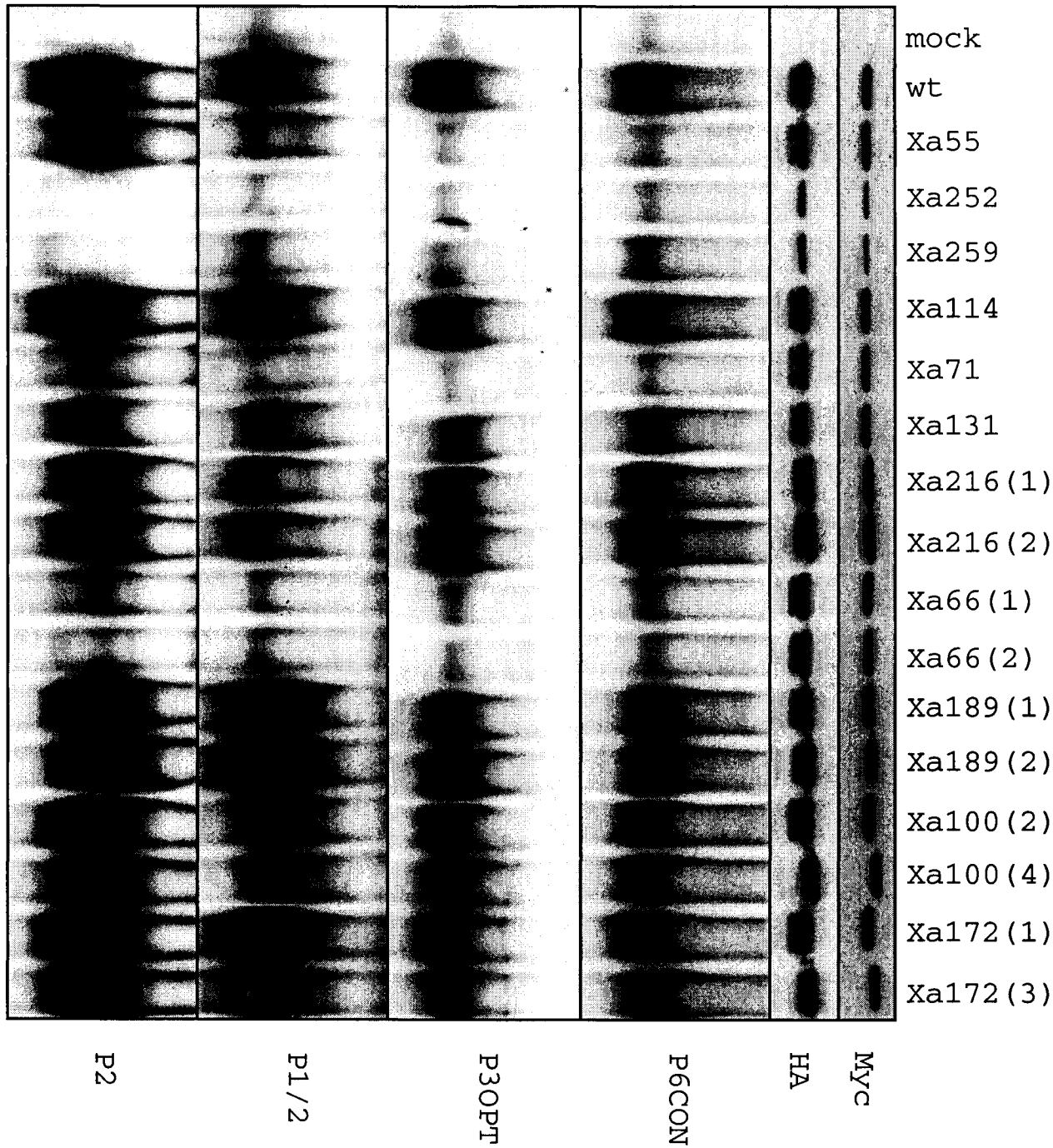
Xa sites (Table I and Fig. 1). Targeted in this group of 10 mutants were the linker separating helix 1 and 2 of the PD (Xa66, 1 and 2 sites), the segment linking the first and second HTH motifs of the PD (Xa100, 2 and 4 sites), and the fragment separating the PD and the HD (Xa172, 1 and 3 sites; Xa189, 1 and 2 sites; Xa216, 1 and 2 sites). Multiple Xa sites were inserted at individual locations not only to maximize accessibility to protease fragmentation but also to provide validation of observed effects.

DNA Binding Properties of Pax3Xa Mutants—Wild type *Pax3* along with the various *Pax3Xa* mutants were introduced in the pMT2 expression plasmid, followed by transient transfection into COS-7 Monkey cells. Immunoblotting of whole cell extracts with either anti-HA or anti-c-Myc monoclonal antibodies indicate similar stability and comparable levels of expression of all mutants in COS-7 cells, with the notable exception of mutants Xa252 and Xa259 (Fig. 2). Reduced levels of expression for Xa252 and Xa259 were noted in multiple transfections and for independent DNA preparations suggesting that mutations at these two positions in the HD may alter protein folding, or processing possibly reducing half-life. The effect of introducing Factor Xa sites on DNA binding properties of the PD and HD of Pax3 in the various mutants was examined by electrophoretic mobility shift assays (EMSAs). DNA binding by the PD was examined using oligonucleotide probes P3OPT (Epstein et al., 1996; Chalepakis et al., 1995) and P6CON (Epstein et al., 1994), previously shown to reveal binding determinants present in both the amino (PAI) and carboxyl (RED) subdomains of the PD (Fig. 2). Mutants at position 55 (Xa55), 66 (Xa66(1); (Xa66(2)), and 71 (Xa71) were found to be severely impaired for DNA binding to P3OPT and P6CON highlighting the critical role of the N-terminal HTH domain (PAI) for DNA binding by the PD. HD Xa mutants Xa252 and

Xa259 also appeared compromised for DNA binding by the PD; however, interpretation of DNA binding results for these mutants was complicated by their low level of

Figure 2

PD and HD DNA binding properties of wild type Pax3 and Pax3 mutants modified by insertion of Factor Xa protease sites. Wild type (*wt*) and mutant Pax3 cDNAs bearing single or multiple Factor Xa protease sites inserted at the position indicated at the *top*, were cloned into pMT2 expression plasmid, and total cell extracts from transiently transfected COS-7 Monkey cells were separated by SDS-PAGE (12% acrylamide) and transferred to nitrocellulose membranes. Immunodetection of Pax3 (*top two panels*) was with mouse anti-c-Myc (*Myc*) and anti-HA (*HA*) monoclonal antibodies and a horseradish peroxidase-conjugated secondary antibody. Extracts from control non-transfected cells (mock) and from cells expressing individual Pax3 mutants were used in electrophoretic mobility shift assays to evaluate the DNA binding properties of Pax3 Xa mutants against PD (*P6CON* and *P3OPT*) and HD-specific binding sites (*P2* and *P1/2*). Protein-DNA complexes were formed using total COS-7 cell extracts and were resolved on 6% acrylamide non-denaturing gels, as described under "Materials and Methods."


mock
wt
Xa55
Xa252
Xa259
Xa114
Xa71
Xa131
Xa216(1)
Xa216(2)
Xa66(1)
Xa66(2)
Xa189(1)
Xa189(2)
Xa100(2)
Xa100(4)
Xa172(1)
Xa172(3)
P2
P1/2
P3OPT
P6CON
HA
Myc

expression in COS-7 cells. The effect of inserting Factor Xa sites on DNA binding properties of the HD was evaluated using a target sequence (P2) containing the sequence $TAAT(N)_2TAAT$ previously shown to support cooperative dimerization of Pax3 (Wilson et al., 1993). In addition, an oligonucleotide containing half of this sequence (half site, P1/2) and revealing monomeric Pax3 binding by the HD was used (Wilson et al., 1993). Results shown in Fig. 2 indicate that DNA binding by the HD was also impaired in the Xa66 mutant, whereas mutants Xa55 and Xa71 show impaired monomeric binding to P1/2 but retain some dimerization potential on P2. Importantly, all other mutants (positions 100, 114, 131, 172, 189, and 216) retained WT binding activity toward the 4 oligonucleotides tested (summarized in Fig. 3), indicating that Factor Xa insertions in the RED subdomain of the PD are not detrimental to function by contrast to mutations at the homologous positions of the other, PAI subdomain. Finally, all insertions in linker segments did not affect DNA binding.

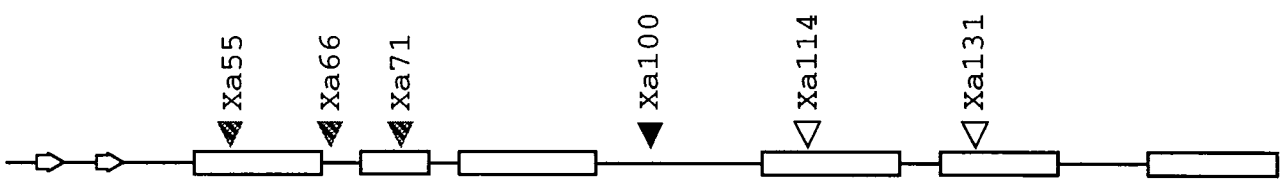
Accessibility of Inserted Xa Sites to Proteolytic Cleavage—The accessibility of the inserted Factor Xa sites to proteolytic cleavage was investigated in mutants showing wild type DNA binding activity. Briefly, whole cell extracts from COS-7 cells expressing either WT or individual Pax3-Xa mutants were incubated with Factor Xa protease (0.2 μg), and at predetermined time points (2-180 min) digestion was stopped and samples were analyzed by SDS-PAGE and immunoblotting with anti-HA (Fig. 4) or anti-c-Myc (Fig. 5) antibodies. The size of the predicted cleavage products immunoreactive with each antibody are shown in Table III. Under these conditions, wild type Pax3 (WT) was almost completely resistant to Factor Xa cleavage, with the full-length 55-kDa immunoreactive species being the prominent band at all time points (*empty arrowhead*; Figs. 4 and 5). Additional minor bands were detected either prior to addition of the protease (35 kDa, 0 min) or at very late time points (30

and 15 kDa, 90-180 min). Although these immunoreactive fragments are derived from the full-length protein, their appearance is likely caused by nonspecific cleavage by Factor Xa at

Figure 3

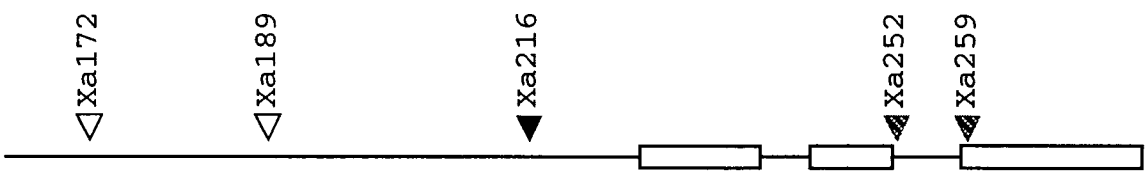
Summary of DNA binding properties and accessibility of protease Xa sites of individual Pax3 mutants. Shown is a schematic representation of the PD and HD of Pax3 according to structural features described in Fig. 1. The sites of modifications to create Factor Xa sites are indicated by *arrowheads*, together with the name of the mutant. The DNA binding properties of each mutant toward PD (*P6CON* and *P3OPT*) and HD targets (*P1/2* and *P2*) is summarized below the schematic diagram (-, absent; +++, wild type). *Gray arrowheads* identify mutants that show important or complete loss of DNA binding; *Filled* (*black*) and *empty arrowheads* identify mutants that retain wild type DNA binding for all sequence targets, and in which the created Factor Xa protease sites are either accessible (*filled arrowheads*) or inaccessible (*empty arrowheads*), under experimental conditions described in Fig. 4.

PD



P6CON	-	-	-	+++	+++	+
P3OPT	-	-	-	+++	+++	++
P1/2	-	-	-	+++	+++	++
P2	++	-	+	+++	+++	+

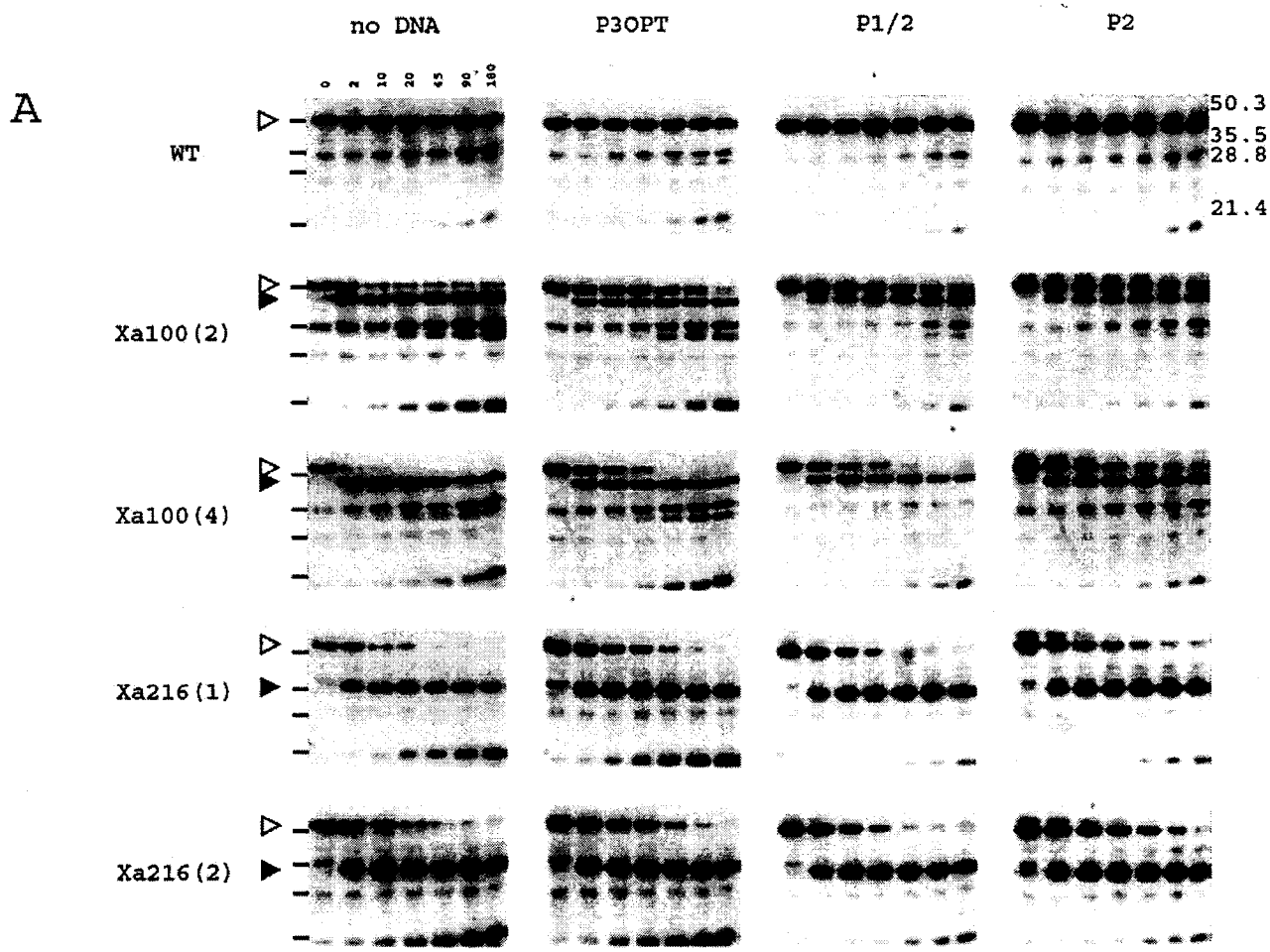
HD



P6CON	+++	+++	+++	-	-
P3OPT	+++	+++	+++	-	-
P1/2	+++	+++	++	-	-
P2	+++	+++	+++	-	-

Figure 4

Accessibility to protease cleavage of Factor Xa sites inserted in selected Pax3 mutants: effect of DNA binding. *A*, ten micrograms of whole cell extracts from COS-7 Monkey cells transfected with either wild type Pax3 (*WT*) or with Pax3 mutants Xa100 (*2*), Xa100 (*4*), Xa216 (*1*), and Xa216 (*2*) were incubated with Factor Xa protease (0.2 μg), either in the absence (no DNA) or after a 30-min incubation with PD (*P3OPT*) and HD target sites (*P1/2* and *P2*) under DNA binding conditions. At pre-determined time points (0-180 min, indicated at the *top*) aliquots were removed and analyzed by SDS-PAGE on 12% acrylamide gels and by immunoblotting with anti-HA antibody. Shown immediately to the *left* of each Pax3 series are *empty arrowheads* that indicate the position of the full-length intact Pax3 products, and *filled* (*black*) *arrowheads* that show the position of major predicted, HA epitope-bearing Factor Xa cleavage products (see Table III). The position and size, in kilodaltons of the molecular mass markers are displayed as *dashes* on the *left* and *numbers* on the *right*, respectively. All immunoblots probed with anti-HA antibody, including those of Fig. S1, were scanned by densitometry (*B*). The intensity of the immunoreactive intact full-length WT protein and of individual Pax3 mutants was determined at both the 0 time point and at the 20-min time points. The amount of intact protein remaining at 20 min (compared with 0 min) was determined for all mutants and for all DNA binding conditions and is expressed as the fraction of intact protein (expressed as a percentage). Several mutants showing inaccessible Factor Xa cleavage sites (Xa114, Xa131, Xa172/3, and Xa189/2) were also included in the analysis as negative controls.


A
no DNA
P3OPT
P1/2
P2
0
2
10
20
45
90'
180
WT
Xa100(2)
Xa100(4)
Xa216(1)
Xa216(2)
50.3
35.5
28.8
21.4

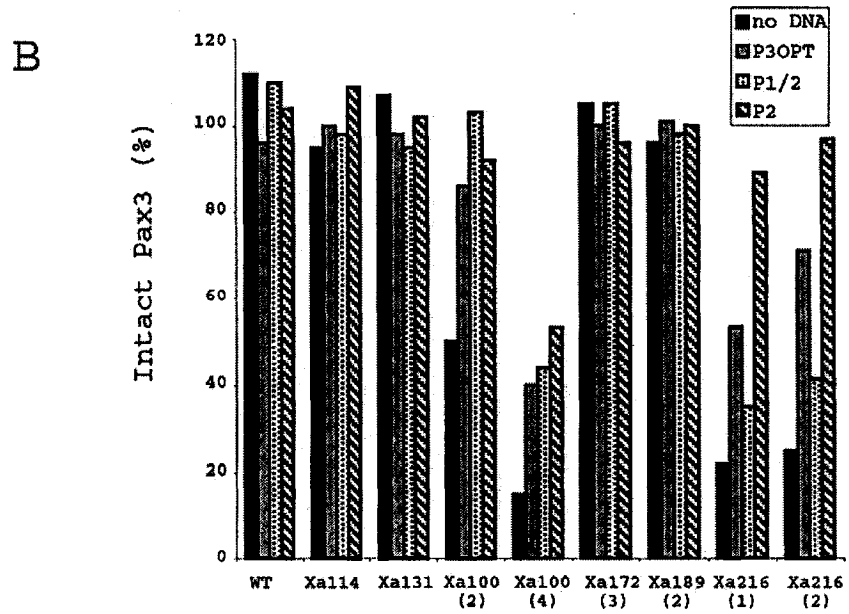
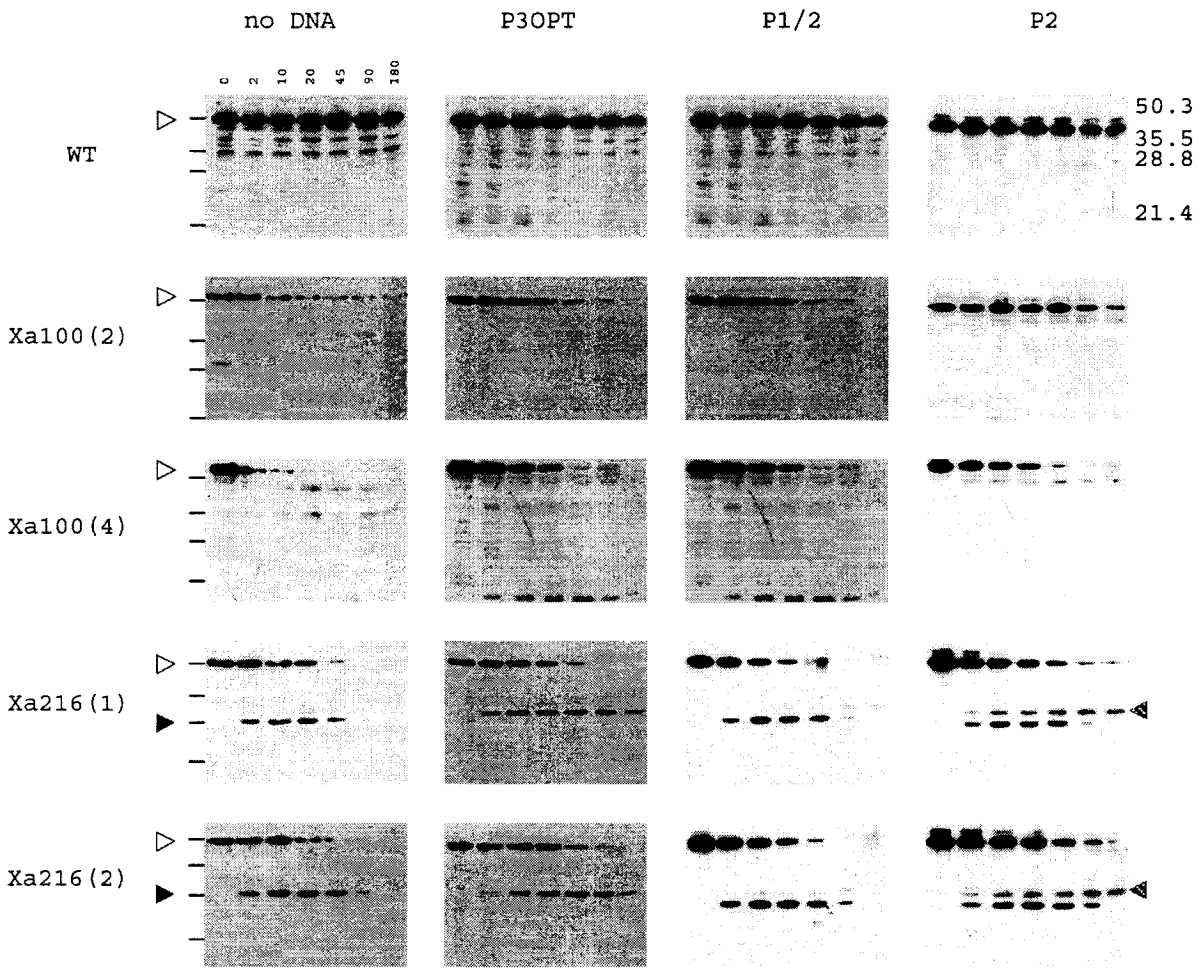

B
Intact Pax3 (%)
120
100
80
60
40
20
0
no DNA
P3OPT
P1/2
P2
WT
Xa114
Xa131
Xa100 (2)
Xa100 (4)
Xa172 (3)
Xa189 (2)
Xa216 (1)
Xa216 (2)

Figure 5

Accessibility to protease cleavage of Factor Xa sites inserted in selected Pax3 mutants: effect of DNA binding. *A*, wild type Pax3 and the different Pax3 mutants were incubated with Factor Xa under different conditions, and digestion products were analyzed by immunoblotting as described in the legend to Fig. 4, except that blots were probed with an anti-c-Myc antibody directed against the N-terminal c-Myc epitope present in all mutants. *B*, all immunoblots probed with anti-Myc antibody, including those of Fig. S2 were scanned by densitometry (Fig. 4*B*). The intensity of the immunoreactive intact full-length WT protein and of individual Pax3 mutants was determined at both the 0 time point and at the 20-min time points. The amount of intact protein remaining at 20 min (compared with 0 min) was determined for all mutants and for all DNA binding conditions and is expressed as the fraction of intact protein (expressed as a percentage). Several mutants showing inaccessible Factor Xa cleavage sites (Xa114, Xa131, Xa172/3, and Xa189/2) were also included in the analysis as negative controls.


no DNA
P3OPT
P1/2
P2
0
2
10
20
45
90
180
WT
Xa100(2)
Xa100(4)
Xa216(1)
Xa216(2)
50.3
35.5
28.8
21.4

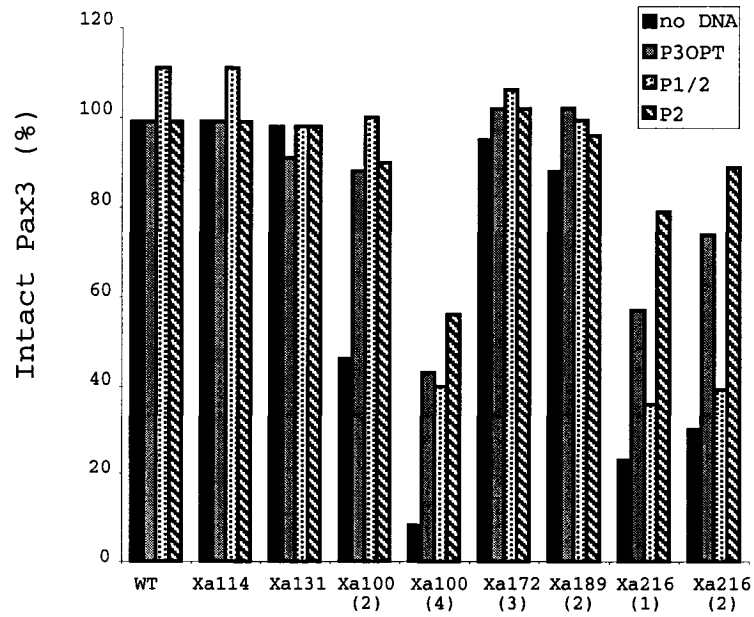

Intact Pax3 (%)
120
100
80
60
40
20
0
no DNA
P3OPT
P1/2
P2
WT
Xa114
Xa131
Xa100 (2)
Xa100 (4)
Xa172 (3)
Xa189 (2)
Xa216 (1)
Xa216 (2)

Table 3: Expected size of Factor Xa proteolytic fragments in individual Pax3 mutants

Pax3 Mutants	[a]Amino Terminal Fragment(1)	[b]Carboxy Terminal Fragment(2)
Xa55	6.7	49.5
Xa71	8.6	47.6
Xa114	13.8	42.4
Xa131	15.9	40.2
Xa252	30.5	25.7
Xa259	31.4	24.8
Xa66(1)	8	48.9
Xa66(2)	8	49.6
Xa100(2)	12.6	44.3
Xa100(4)	12.6	44.3
Xa172(1)	20.8	36.1
Xa172(3)	20.8	36.1
Xa189(1)	22.7	34.3
Xa189(2)	22.7	34.7
Xa216(1)	25.5	31.6
Xa216(2)	25.5	31.6

[a]Expected specific amino terminal proteolytic fragment (kD) detectable using anti-cMyc antibody.
[b]Expected specific carboxy terminal proteolytic fragment (kD) detectable using anti-HA antibody.

suboptimal sites, or cleavage at specific sites by additional proteases contaminating either the cell extract or the commercial Factor Xa preparations. For reasons discussed below (see "Discussion"), these fragments were not considered in our analysis. Stripping the blot and reprobing with anti-c-Myc antibody (Fig. 5) confirmed the resistance of WT Pax3 to digestion by Factor Xa under the conditions tested. Analysis of mutants Xa114, Xa131, Xa172(1/3), and Xa189(1/2) showed results similar to WT Pax3, and indicated almost complete resistance of these proteins to digestion by Factor Xa (supplementary data, Fig. S1), with little if any predicted proteolytic fragments immunoreactive with the HA antibody detected even after 180 min of incubation. Similar results were obtained using the anti-c-Myc antibody to analyze digestion products (supplementary data, Fig. S2). These results indicate that the C-terminal RED subdomain and the linker domain immediately downstream of the PD are probably not solvent-exposed and assume a compact conformation under the conditions tested (Fig. 3). In contrast, mutant Xa100 (2) yielded the expected 44-kDa C-terminal HA-immunoreactive cleavage product upon incubation with Factor Xa (*filled arrowhead*, Fig. 4). The 44-kDa HA fragment was abundant at the earliest time point tested (2 min), and digestion was largely completed by 10 min. Similar rapid cleavage of Xa100 (2), as demonstrated by disappearance of the full-length 55-kDa protein, was verified by immunoblotting with the anti-c-Myc antibody (Fig. 5), although the N-terminal 12-kDa c-Myc-reactive digestion product was not retained on the gel. Increasing the number of Xa sites from two to four in mutant Xa100 (4) produced similar outcome with even more rapid and more complete cleavage at the targeted site. Thus, cleavage at position 100 is specific in these mutants and strongly suggests that the linker separating the PAI and RED subdomains of the PD is solvent-exposed and protease-

accessible. Analysis of mutants Xa216 (1) and Xa216 (2) showed similar results. Both of the specific C-terminal HA-reactive 32-kDa and N-terminal c-Myc-reactive 26-kDa digestion products appeared at 2 min, and digestion of the full-length protein was almost complete by 20-45 min (Figs. 4 and 5). These results indicate that the protein segment immediately upstream of the HD is solvent-exposed and accessible to protease cleavage (Fig. 3).

Effect of DNA Binding on Protease Sensitivity of Pax3 Mutants Xa100 and Xa216— The effect of Pax3 binding to PD (P3OPT) and HD (P1/2, P2) target sequences on the conformation of each domain was analyzed by monitoring the effect of DNA binding on accessibility of Xa cleavage sites present in mutants Xa100 (PD) and Xa216 (HD). Also included in these experiments were WT Pax3 as well as Pax3-Xa mutants previously observed to be resistant to Factor Xa cleavage in time course studies (see supplementary data). Briefly, cell extracts expressing Pax3 proteins were incubated with or without target DNA, followed by addition of Factor Xa and detection of HA (Fig. 4) and c-Myc (Fig. 5) immunoreactive cleavage products appearing over time. The extent of protection from proteolytic fragmentation was further quantitated after densitometry of the immunoblots and is expressed as the fraction of intact full-length Pax3 remaining following 20 min of incubation with Factor Xa (Figs. 4*B* and 5*B*). For the WT Pax3, binding to PD or HD target sequences had no effect on digestion profiles, as expected, with full-length Pax3 remaining the predominant species throughout the incubation period. Likewise, DNA binding by the PD or the HD of mutants Xa114, Xa131, Xa172, and Xa189 did not affect their previously noted resistance to Factor Xa cleavage (Figs. 4*B* and 5*B* and Figs. S1 and S2 in supplementary data). This suggests that DNA binding by either domain in these mutants does not cause a conformational change that increases solvent exposure of the respective Xa bearing segments. By contrast, incubation of mutant Xa100 (2) with P3OPT increased resistance to proteolysis (persistence of 55-kDa protein), suggesting that DNA binding to the PD causes a conformational change reducing

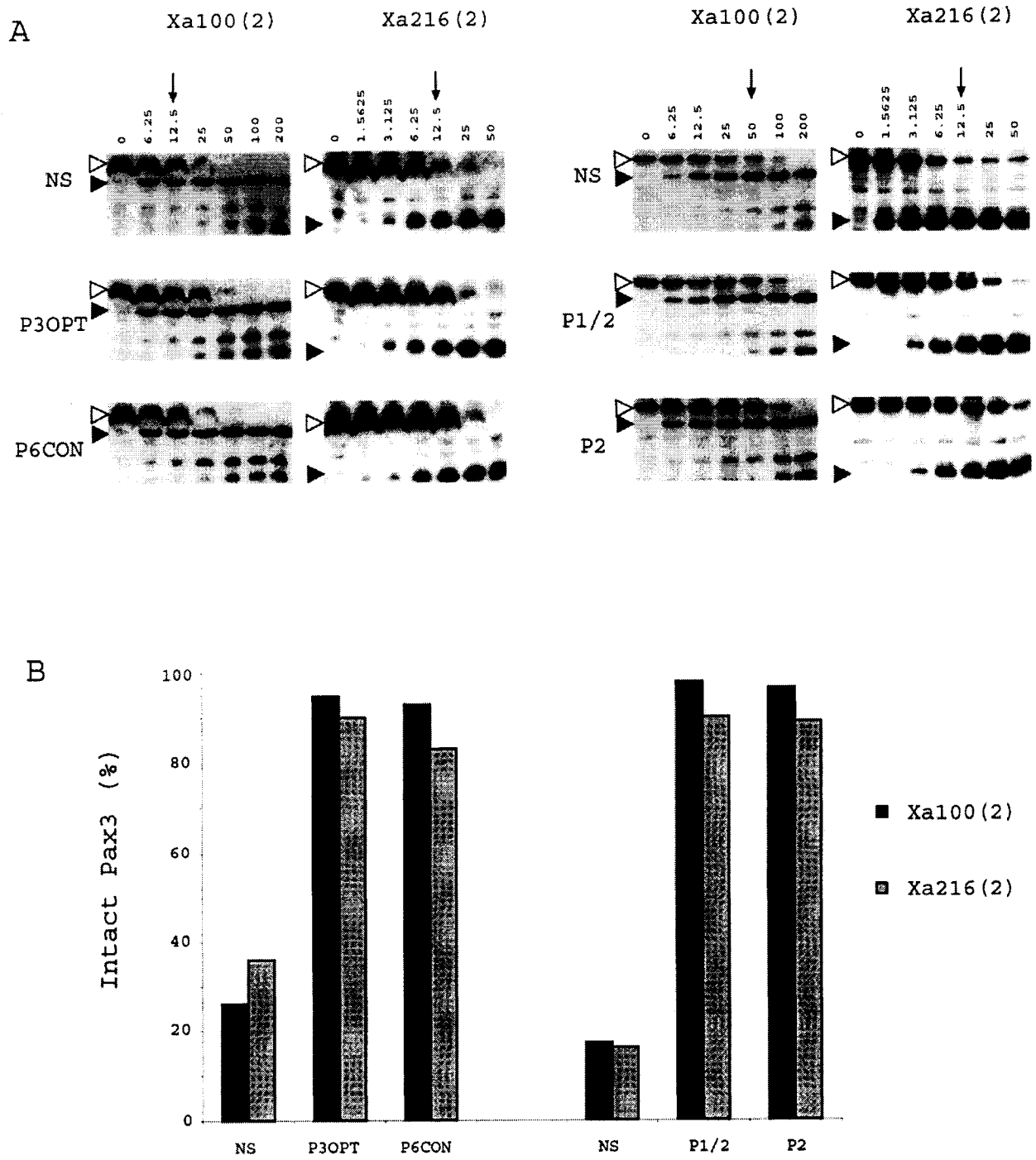
solvent accessibility of the PD. In addition, monomeric binding to P1/2 and in particular dimerization on P2 both also caused a dramatic increased resistance to proteolytic cleavage of Xa100 (2) (from 10% to 40% intact protein), suggesting that DNA binding by the HD in this mutant also causes a conformational change in the PD. Identical results were obtained with Xa100 (4), although the increased susceptibility to cleavage caused by the four consecutive Xa sites at position 100 was maintained in this mutant for all DNA binding conditions tested. In the case of mutants Xa216 (1) and Xa216 (2), mutants that bear Xa sites immediately upstream of the HD, DNA binding by the HD, in particular dimerization on P2, caused a strong increased resistance to proteolysis (from 25% to 85% intact protein) suggestive of a conformational change at that site. In addition, binding of both mutants to the PD target sequence (P3OPT) also increased resistance to proteolysis, suggesting that DNA binding by the PD also causes a conformational change in the HD. In all cases, results of immunoblotting with anti-HA (Fig. 4) and anti-c-Myc antibodies (Fig. 5) were in complete agreement (Figs. 4*B* and 5*B*). Together, results from Xa100 and Xa216 are remarkably similar and suggest that DNA binding at either the PD or HD causes a conformational change at both sites. This change appears to reduce the amount of solvent-exposed area in the protein, suggesting a more compact conformation of the DNA-bound protein.

Specificity of the protective effect of PD and HD target sequences on accessibility of the Xa sites in Xa100 and Xa216 mutants was investigated in dose-response studies. In these experiments, the Pax3 mutants were incubated with PD or HD oligonucleotide probes and DNA-Pax3 complexes were allowed to form; following this, increasing amounts of Factor Xa protease was added to the reaction mixture, which was further incubated for 15 min, and the appearance of specific proteolytic cleavage products was monitored by SDS-PAGE and immunoblotting with the anti-HA antibody (Fig. 6*A*). In these studies, two additional control probes were tested to further validate the specificity of the DNA effect observed in time

course studies: a second, independently derived PD oligonucleotide P6CON, and a PD and HD nonspecific oligonucleotide, which was used as a negative control. Typical immunoblots are shown in Fig. 6*A*, and quantitation of the protective effects of the DNA probes by

Figure 6

Effect of DNA binding on accessibility of Factor Xa sites in Pax3 mutants Xa100 and Xa216: dose-response experiments. *A*, whole cell extracts (10 μg) from COS-7 Monkey cells transfected with Pax3 mutants Xa100 (2), or Xa216 (2) were incubated 30 min with either PD-specific (*P3OPT* and *P6CON*) or HD-specific (*P1/2* and *P2*) oligonucleotides or with a nonspecific (*NS*) oligonucleotide. The buffer conditions for PD *versus* HD DNA binding were different (see "Materials and Methods"), and the negative control was included in each case. Increasing amounts of Factor Xa protease (from 0 to 200 ng/μg of whole cell extract) were added followed by a 15-min digestion period. Proteolytic fragments were separated by SDS-PAGE and detected by immunoblotting using the anti-HA antibody, as described in the legend to Fig. 4. The positions of the full-length intact Xa100 (2) and Xa216 (2) proteins (*empty arrowheads*), and of the major predicted, HA-immunoreactive product (*filled arrowhead*, see Table III) are shown. *B*, for quantifying the effect of DNA binding on accessibility of Factor Xa cleavage sites in dose-response studies, the immunoblots in *A* were scanned by densitometry. The intensity of the immunoreactive intact Xa100 (2) and Xa216 (2) full-length proteins was determined in the absence of Factor Xa and after digestion with a Factor Xa amount indicated by the *arrow* at the *top* of each immunoblot series. The amount of intact protein remaining after Factor Xa digestion was determined for Xa100 (2) (*black*) and Xa216 (2) (*gray*) and for all DNA binding conditions and is expressed as the fraction of intact protein (expressed as a percentage).


A
Xa100(2)
Xa216(2)
Xa100(2)
Xa216(2)
NS
P3OPT
P6CON
P1/2
P2
B
Intact Pax3 (%)
100
80
60
40
20
0
NS
P3OPT
P6CON
NS
P1/2
P2
Xa100(2)
Xa216(2)

densitometry is shown in Fig. 6*B* (percentage of uncleaved Pax3 at concentration indicated by the *arrow* in Fig. 6*A*). These experiments showed that incubation of mutants Xa100 (2) and Xa216 (2) with PD oligonucleotides P6CON and P3OPT reduced sensitivity of both proteins to increasing doses of Factor Xa. Likewise, incubation of both proteins with HD probes P1/2 and P2 similarly increased resistance to Factor Xa fragmentation. The effect of the two PD probes and two HD probes on protease sensitivity was specific and not seen in control DNA-free conditions (data not shown) and upon incubation with a nonspecific target sequence (Fig. 6, *A* and *B*). Therefore, results of dose response experiments are consistent with those obtained in time course experiments and show that DNA binding to the PD and to the HD of Pax3 cause a conformational change both locally as well as in the other DNA binding site.

Discussion

The effect of inserting Factor Xa sites in different domains of Pax3 provides information on the importance of the targeted subdomains in DNA binding by the protein. In the PD, insertion of Factor Xa sites at the three locations tested in the N-terminal PAI subdomain caused loss of PD DNA binding. Because helix 3 makes critical base and phosphate contact in the major groove of DNA and is mutation-sensitive, it was not targeted for insertions. Nevertheless, we observed that mutations in helix 1 (Xa55), in the linker separating helices 1 and 2 (Xa66), and in helix 2 (Xa71) all impaired binding to PD oligonucleotides. Insertion of Xa sites may either disrupt important phosphate contacts made between these helices and DNA (Fig. 1) and/or may destabilize the whole PAI domain. By contrast, two insertions created in helices 1 (Xa114) and 2 (Xa131) of the C-terminal RED domain had no major effect on DNA binding to PD sequences. These results highlight the critical role of the PAI domain in DNA binding by the PD and are in agreement with (*a*) the high degree of primary amino acid sequence conservation of this domain (compared with the RED subdomain) in the Pax/Prd gene family (Xu et al., 1995; Epstein et al., 1994; Czerny et al., 1993; Epstein et al., 1994), (*b*) the clustering of inactivating WS1 mutations to the PAI subdomain of PAX3 (Fortin et al., 1997), and (*c*) the fact that Pax proteins can bind DNA exclusively through their PAI subdomain (Czerny et al., 1993). We note that mutations Xa55 and Xa71 in the PAI domain impaired DNA binding to PD oligonucleotides only, whereas the Xa66 mutant showed impaired DNA binding to both PD and HD sequences. This behavior is similar to WS1 mutants G48R/S and P50L (upstream helix 1 of PAI), respectively, and has been suggested to reflect functional interdependence of the PD and HD in DNA binding (Fortin et al., 1997). Interestingly, insertion of two or even four Xa sites in the linker joining the PAI to the RED (Xa100) had no effect on DNA binding; this linker is well conserved

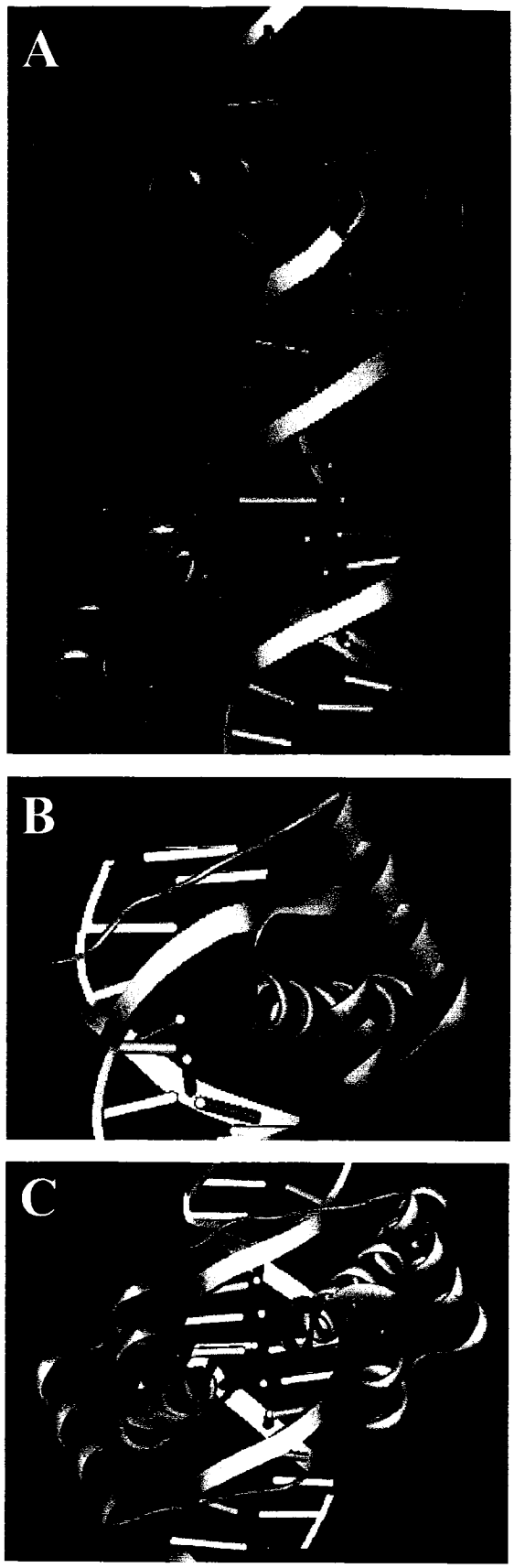
among Pax proteins, and sequences immediately downstream the insertion site make extensive phosphate and base-specific contacts in the minor groove of DNA (Fig. 1). As expected, the three insertions (Xa172, Xa189, and Xa216) in the poorly conserved linker joining the PD and the HD, including one within the octapeptide motif (Xa189) conserved in other Pax proteins (HSIDGILG) (Burri et al., 1989), had no major impact on DNA binding of the corresponding mutants to PD and HD sites. With respect to the HD, two mutations inserted either downstream helix 2 (Xa252) or upstream of the major DNA binding helix 3 (Xa259) appeared to either strongly diminish or abrogate monomeric or dimeric DNA binding to HD target sequences (Fig. 2). Although this conclusion is supported by both the high degree of conservation of the targeted sequences in the paired-type HDs of the Pax family (Fig. 1), and the fact that many WSI mutations map to the HD of Pax3 (Fortin et al., 1997), our inability to express high levels of these mutants precluded a more detailed analysis.

Evaluating the accessibility of inserted Xa sites to protease cleavage can readily provide insight into the solvent exposure of the corresponding Pax3 subdomains. Results of time course studies were very clear and showed that of the seven sites in which insertions preserved DNA binding only two, Xa100 and Xa216, were readily accessible to proteolytic cleavage at the earliest time points of analysis (Fig. 4). In agreement with the proposed structural model of the PD (Fig. 7*A*), these results showed that the PAI to RED subdomain linker (Xa100 mutant) is clearly exposed to solvent. Although this domain is not believed to play a critical role in DNA binding, current structural models suggest that it is in close proximity to DNA (Xu et al., 1995). In addition, alternative splicing of a glutamine residue at position 108 of Pax3/Pax7 is known to alter DNA binding specificity of the PD (Vogan et al., 1996; Vogan et al., 1997). Therefore, mutants Xa100 should be ideally suited to monitor structural changes associated with DNA binding by the PD. By contrast, downstream

insertions into helices 1 (Xa114) and 2 (Xa131) of the RED domain were completely resistant to protease cleavage, possibly suggesting that the HTH motif is either compact or buried in the

Figure 7

Structures of the PD and HD bound to DNA. The structure of the DNA-bound PD (*A*) is adapted from that published for Pax6 (Balling et al., 1988; Xu et al., 1999). The structures of the HD bound to DNA either as a monomer (*B*) or as a dimer (*C*) are taken from that published for the HD of the *Drosophila* Prd protein (Wilson et al., 1995). The DNA strands are shown as *gray ribbons* through the sugar phosphate backbone, and bases are shown as *protrusions from the ribbons*. The Pax3 segments are in *green*, with the positions of discrete sites targeted for modification to Factor Xa sites identified as causing loss-of DNA binding (*red*), as having no effect on DNA binding and being either accessible (*purple*) or not (*blue*) to Factor Xa protease fragmentation.


A
B
C

core of the protein or both (Fig. 7*A*). PD-HD linker insertions mapping far upstream of the HD (Xa172 and Xa189) were resistant to protease cleavage. Considering the observed accessibility of neighboring position 216, this is somewhat surprising; however, it is important to note that Xa sites at positions 172 and 189 were engineered immediately upstream of a hydrophobic residue (Leu and Ile, respectively), a situation known to reduce the efficacy of cleavage by the Factor Xa protease (He et al., 1993).

The effect of binding to PD and HD oligonucleotides on accessibility of all inserted Xa sites was investigated in time course studies (Figs. 4, 5, S1, and S2). In all cases, DNA binding did not increase accessibility of Xa sites, suggesting that DNA binding as a whole does not dramatically increase solvent exposure in the protein. However, this conclusion only applies to the subdomains analyzed and is limited by the fact that the primary amino acid sequence context of individual sites is not identical, possibly influencing protease cleavage irrespective of the three-dimensional structure (DNA-free or DNA-bound) of this subdomain. On the other hand, DNA binding by the PD reduced accessibility of the Xa100 site, reflecting a conformational change to a more compact and less solvent-accessible position of the PAI to RED linker. The effect was specific, was seen in independent mutants bearing different number of Xa sites inserted, was noted upon binding to independent PD site, and was not seen upon binding to an unrelated target site sequence. These results are in accordance with circular dichroism spectroscopy studies with Pax5 and Pax8 PDs that reveal increased helical content in the DNA-bound form of the PD (Tell et al., 1998). Likewise, DNA binding by the HD also resulted in decreased accessibility of the HD N-terminal linker (Xa216), reflecting a more compact structure of the DNA-bound state of this domain. These results are in agreement with x-ray crystallography and NMR studies of HD proteins that reveal an increased order of the

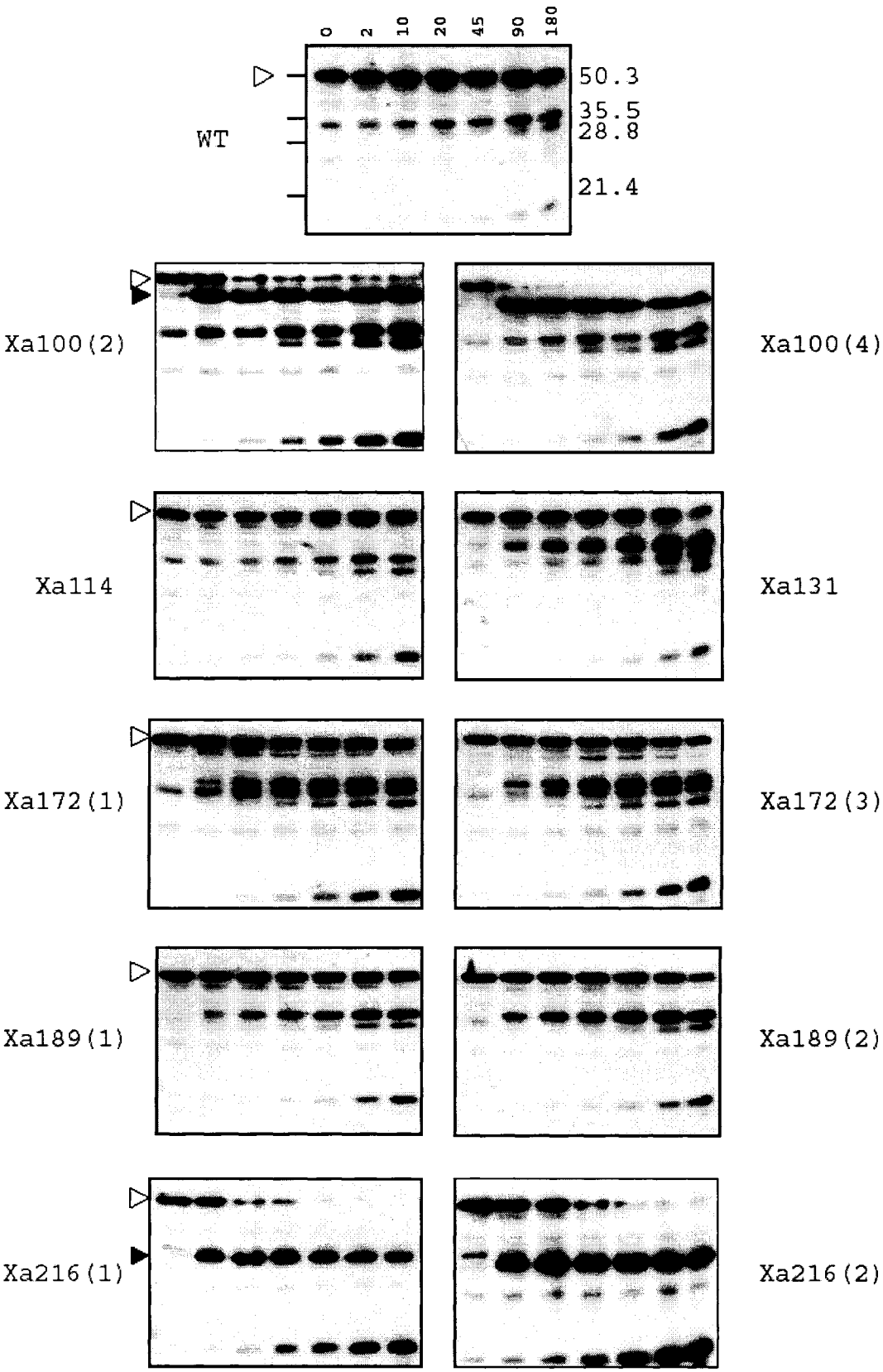
HTH motif, including the N-terminal arm and the recognition helix, of the DNA-bound form of Pbx (Sprules et al., 2003; Sprules et al., 2000; Jabet et al., 1999; Piper et al., 1999), Antennapedia (Qian et al., 1994), engrailed (Clarke et al., 1994), and Oct-1 (Cox et al., 1995) proteins. Increased resistance to protease cleavage of the Xa216 site was more pronounced upon incubation with P2 than with the half site P1/2. This could reflect different stability and steady-state level of conformationally similar HDs in the DNA-bound state (supported by the EMSA data; Fig. 2) or could be indicative of two distinct conformations of DNA-bound HDs in Pax3 monomers and dimers complexed to P1/2 and P2, respectively.

Importantly, DNA binding by the PD also caused a conformational change in the HD (P3OPT binding to Xa216 mutants); likewise, DNA binding by the HD also caused a conformational change in the PD (P2 binding to Xa100 mutants). These results provide for the first time a structural basis for the functional interdependence of the two DNA binding domains previously noted in the study of (*a*) PD and HD DNA binding properties of PAX3 mutants from WS1 patients (Fortin et al., 1997), and (*b*) the effect of PD and HD DNA binding on site-specific modifications of Pax3 mutants bearing single cysteine residues in each domain (Apuzzo et al., 2002). These results further suggest that the PD and HD of Pax3 can functionally interact for the final selection of target site sequences *in vivo*. Finally, results in Fig. 5 indicate that dimerization of Pax3 on P2 results in a conformation that appears distinct from that created in the same mutant upon binding to the PD oligonucleotide. Indeed, comparison of the digestion profile of P2-bound *versus* P3OPT-bound Xa216 mutants identifies a novel nonspecific but c-Myc-immunoreactive cleavage product (*shaded arrow*, Fig. 5), which is present in the former but absent in the latter set of digest. This nonspecific cleavage product is detected only upon binding of Xa216 mutants to P2 and not to P1/2, strongly suggesting that it is caused by dimerization of Pax3 on P2, as opposed to monomeric binding to the half site.

The results of protease sensitivity studies presented here are in agreement with the structural model for combined PD and HD binding to chimeric target sequences proposed for the *Drosophila* Prd protein by Jun and Desplan (Xu et al., 1995). This model, based in part on the sequence arrangement of binding sites selected *in vitro* by sequential amplification from random DNA oligomers (Systematic Evolution of Ligands by EXponential (SELEX) enrichment procedure), suggests that the N terminus of the PAI domain is closely apposed to the N terminus of the HD when bound to the combined PH0 site (Jun et al., 1996). Their model predicts that HD and PD bind to opposite sides of the DNA helix, with helix 2 of the PAI domain in very close proximity of the N-terminal extension of the HD, which was targeted for Xa site insertion in the present study. This model may be used to explain the interaction of the PD of Pax3 with the N-terminal arm of the HD of Msx1 (Bendall et al., 1999). Therefore, the critical position of the N-terminal extension of the HD identified in these studies agrees with our observation that DNA binding by either the PD or HD causes a structural change at that site. Finally, the importance of the N-terminal extension of the HD and of helix 2 of the PAI domain in Pax3 function is highlighted by previous studies from our group showing that (*a*) alteration of the N-terminal HD extension abrogates the ability of the PD to modulate DNA binding specificity of the HD (Fortin et al., 1998), and (*b*) that deletion of helix 2 of the PAI domain restores DNA binding by the HD in the context of an otherwise inactivating mutation (Sp^d) in the PD (Fortin et al., 1998).

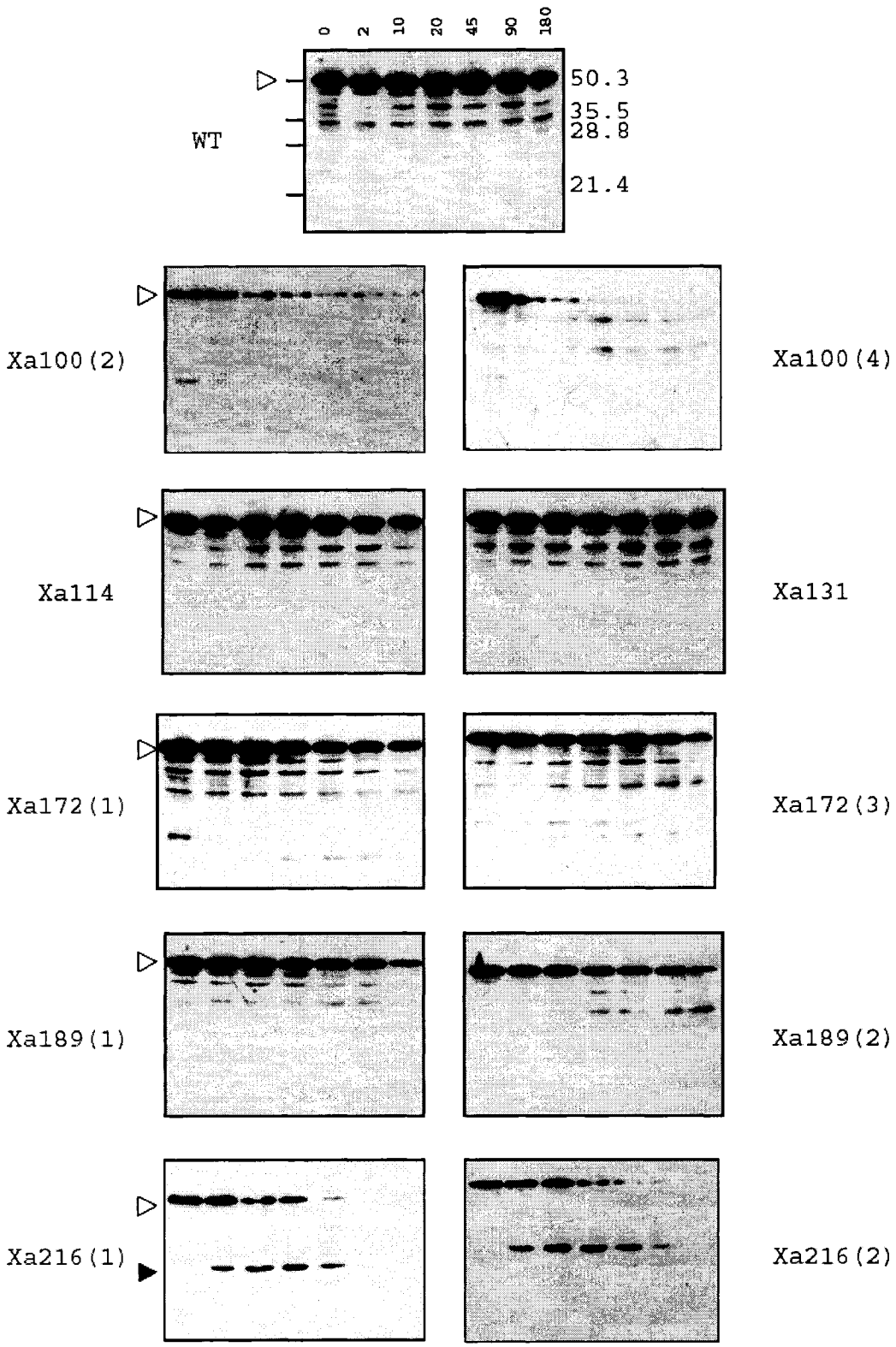
Supplementary Figure S1

Accessibility to Protease Cleavage of Factor Xa sites inserted in Additional Pax3 Mutants: Effect of DNA Binding. Results are shown as described in legend to Figure 4 (anti-HA probed).


0
2
10
20
45
90
180
WT
50.3
35.5
28.8
21.4
Xa100(2)
Xa100(4)
Xa114
Xa131
Xa172(1)
Xa172(3)
Xa189(1)
Xa189(2)
Xa216(1)
Xa216(2)

Supplementary Figure S2

Accessibility to Protease Cleavage of Factor Xa sites inserted in Additional Pax3 Mutants: Effect of DNA Binding. Results are shown as described in legend to Figure 5 (anti-Myc probed).


0
2
10
20
45
90
180
WT
50.3
35.5
28.8
21.4
Xa100(2)
Xa100(4)
Xa114
Xa131
Xa172(1)
Xa172(3)
Xa189(1)
Xa189(2)
Xa216(1)
Xa216(2)

Protease sensitivity studies with Pax3 mutants with engineered Xa cleavage sites have shown that DNA binding of one domain induces a conformational change in the other DNA binding domain. As chapter 3 has indicated, this conformational change was monitored by differential protease sensitivity, more specifically; a reduction of protease sensitivity was observed when Pax3 is DNA bound. We postulate that these conformational changes might be induced by intramolecular physical interactions between the PD and the HD.

The objective of the following study was to identify residues or segments in the PD that may be involved in HD physical interaction in order to shed more light on the functional interdependence of the two domains. The approach used was the same for the study described in chapter 1: cysteine scanning mutagenesis. As mentioned in the literature review, the Pax3 Sp-delayed mutant holds a G42R mutation in the β-hairpin of the PAI subdomain that abrogates paired and HD DNA binding activity (Underhill et al., 1995). Deletion of helix 2 and flanking regions, in the PAI subdomain restores HD DNA binding activity in this mutant protein (Fortin et al., 1998). Based on these observations we postulate that helix 2 and flanking sequences, may form a site where HD physical interaction with the PD occurs. Consequently, this region of the PAI subdomain was targeted for study using cysteine scanning mutagenesis since we believe it holds a HD interaction site.

Chapter 4

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

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.

Introduction

Pax3 is a member of the Pax family of nine transcription factors that play critical roles in different aspects of mammalian development (Stuart et al., 1994). 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 (Bober et al., 1994; Beechey et al., 1986; Auerbach et al., 1954; Franz et al., 1990; Franz et al., 1989; Goulding et al., 1994). In humans, mutations in *PAX-3* cause Waardenburg syndrome, a pathology characterized by pigmentary disturbances, craniofacial abnormalities, and sensory deafness (Baldwin et al., 1992; Baldwin et al., 1995).

Pax proteins are defined by a unique DNA binding domain, the PD first identified in the *Drosophila* paired (Prd) protein (Bopp et al., 1986). Several Pax proteins, including Pax3, contain a second DNA binding domain known as the paired-type 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 (Stuart et al., 1994; Mansouri et al., 1998). 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) (Xu et al., 1995; Xu et al., 1999); the C-terminal helix of each HTH makes critical contacts in the major groove of DNA (Xu et al., 1995; Xu et al., 1999). Located upstream of the N-terminal subdomain (PAI), a β-turn motif makes contacts in the minor groove of DNA (Xu et al., 1995; Xu et al., 1999). The crystal structure of the DNA-bound paired-type HD of the Prd protein (Wilson et al., 1995) 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 (Wilson et al., 1995; Wilson et al., 1993; Treisman et al., 1989). The HD of Pax proteins has the unique ability to homodimerize on palindromic sequences of the type TAAT($N_{2/3}$)ATTA (Wilson et al., 1993; Schafer et al., 1994). The identity of the amino acid residue at position 50 in α3 of HD determines both DNA binding specificity and dimerization potential (Wilson et al., 1993; Schafer et al., 1994).

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 (Baldwin et al., 1995; Glaser et al.,1993; Lalwani et al., 1995; Underhill et al., 1995). 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 (Underhill et al., 1995). 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 (Fortin et al., 1997). Additional studies in chimeric PAX3 proteins have shown that the PD can modulate both DNA binding specificity and dimerization potential of heterologous HDs (Fortin et al., 1998).

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 (Apuzzo et al., 2002). 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 (Apuzzo et al., 2002). 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 (Apuzzo et al., 2002). 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 (Apuzzo et al., 2002). 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 (Apuzzo et al., 2004). 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 (Apuzzo et al., 2004).

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 (Fortin et al., 1998). 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 (Underhill et al., 1995; Fortin et al., 1998). 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) (Jun et al., 1996). 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 (Jun et al., 1996). 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 (Garvie et al., 2001) 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 (Garvie et al., 2001). 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 (Underhill et al., 1995). This pMT2/Pax3 construct encodes for all 479 amino acids of the murine Q+ isoform of Pax3 (Underhill et al., 1995). 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 (Apuzzo et al., 2002). 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 (Apuzzo et al., 2004). The oligonucleotide used to introduce 2 protease Xa sites, (5')-CGATATCGAAGGTAGAATAGACGGCCGAAT-(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')-AGTGAGAGGGGAGAGAGCATAGTC-(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 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),

Table 1. Oligonucleotides used for Pax3 mutagenesis

Substitution	Mutagenic primer (5'-3')
PD SCM	
Q40C	GGCCGAGTCAAC**TGT**CTCGGAGGAGTA
L41C	CGAGTCAACCAG**TGC**GGAGGAGTATTT
I59C	ATCCGCCACAAG**TGT**GTGGAGATGGCC
V60C	CGCCACAAGATA**TGT**GAGATGGCCCAC
E61C	CACAAGATAGTG**TGT**ATGGCCCACCAT
M62C	AAGATAGTGGAG**TGT**GCCCACCATGGC
A63C	ATAGTGGAGATG**TGC**CACCATGGCATT
H64C	GTGGAGATGGCC**TGC**CATGGCATTCGG
H65C	GAGATGGCCCAC**TGT**GGCATTCGGCCG
G66C	ATGGCCCACCAT**T**GCATTCGGCGGAGC
I67C	GCCCACCATGGC**TGT**CGGCCGAGCGTC
R68C	CACCATGGCATT**TGT**CCGAGCGTCATT
P69C	CATGGCATTCGG**TGT**AGCGTCATTTCT
S70C	GGCATTCGGCCG**T**GCGTCATTTCTCGC
V71C	ATTCGGCCGAGC**TGC**ATTTCTCGCCAG
I72C	CGGCCGAGCGTC**TGT**TCTCGCCAGCTT
S73C	CCGAGCGTCATTT**GT**CGCCAGCTTCGC
R74C	AGCGTCATTTCTTGCCAGCTTCGCGTG
Q75C	GTCATTTCTCGC**TGT**CTTCGCGTGTCC
L76C	ATTTCTCGCCAG**TGT**CGCGTGTCCCAT
R77C	TCTCGCCAGCTTTGCGTGTCCCATGGA
V78C	CGCCAGCTTCGC**TGT**TCCCATGGATCC
S79C	CAGCTTCGCGTGT**G**CCATGGATCCGTC
H80C	CTTCGCGTGTCC**TG**TGGATCCGTCTCT

Nucleotide substitutions leading to amino acid changes are indicated in bold

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-mercaptoethanol, 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 (Underhill et al., 1995). 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') (Epstein et al., 1994) and P3OPT (5')-TGGTGGTCACGCCTCATTGAATATTA-(3') (Chalepakis et al., 1995), HD-specific sequence P2 (5')-GATCCTGAGTCTAATTGATTACTGTACAGG-(3') (Xu et al., 1999), and the composite PD and HD specific sequence PH0 (5')-GATTTCTTCCAATTAGTCACGCTTGAGTG-(3') (Jun et al., 1996) 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.

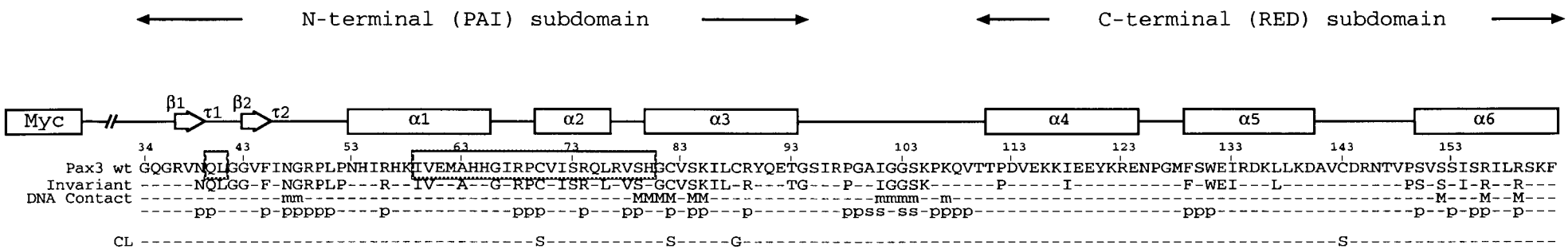
Protein Expression and DNA Binding Properties of Pax3 Single Cysteine Mutants. To facilitate detection of the wild-type (WT), Cys-less (CL), and all single cysteine mutants (SCM), c-Myc and hemagglutinin (HA) epitope tags were inserted in-frame at the N-terminus and C-terminus of all constructs, respectively. 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

Figure 1

Site-directed mutagenesis of the PD 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) (Xu et al., 1999). 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 HD, 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.

PD

A


N-terminal (PAI) subdomain
C-terminal (RED) subdomain
Myc
β1
τ1
β2
τ2
α1
α2
α3
α4
α5
α6
34 43 53 63 73 83 93 103 113 123 133 143 153
Pax3 wt GQGRVNQLGGVFINGRPLPNHIRHKIVEMAHHGIRPCVISRQLRVSHGCVSKILCRYQETGSIRPGAIGGSKPKQVTTPDVEKKIEEYKRENPGMFSWEIRDKLLKDAVCDRNTVPSVSSISRILRSKF
Invariant -----NQLGG-F-NGRPLP---R--IV--A--G-RPC-ISR-L-VS-GCVSKIL-R---TG---P--IGGSK------P-----I----------F-WEI---L----------PS-S-I-R--R---
DNA Contact ----------mm----------MMMM-MM----------mmmm--m----------M---M--M---
-----pp----p-ppppp----p----------ppp---p---pp--p-pp---p--------ppss-ss-pppp----------ppp----------p--p-pp--p---
CL ----------S----------S-----G----------S----------

HD

B

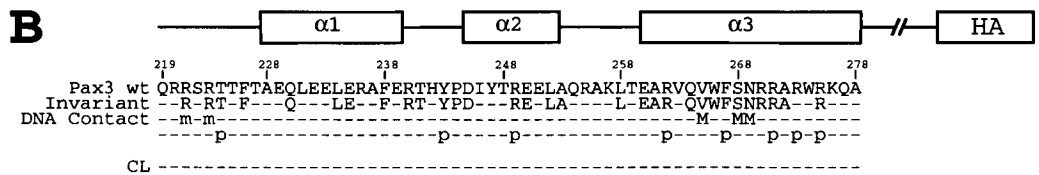

α1
α2
α3
HA
219 228 238 248 258 268 278
Pax3 wt QRRSRTTFTAEQLEELERAFERTHYPDIYTREELAQRAKLTEARVQVWFSNRRARWRKQA
Invariant --R-RT-F---Q---LE--F-RT-YPD---RE-LA----L-EAR-QVWFSNRRA--R---
DNA Contact --m-m----------M--MM---------
-----p----------p-----p----------p----p---p-p-p---
CL ----------

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.

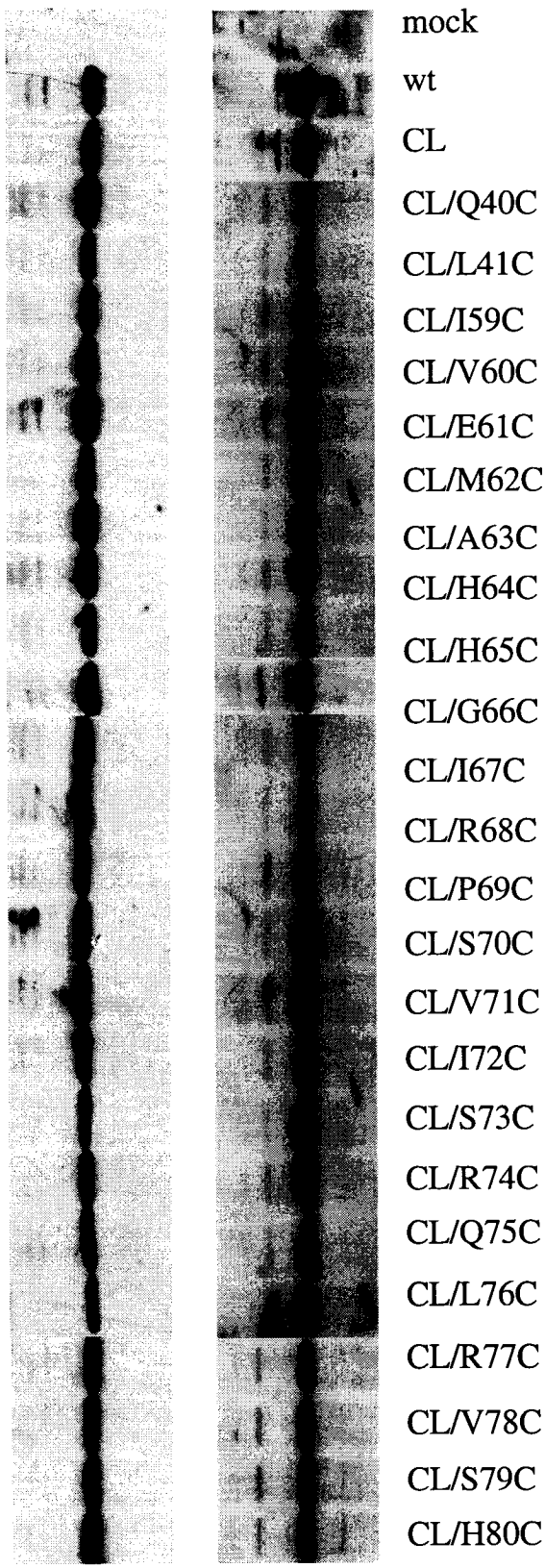
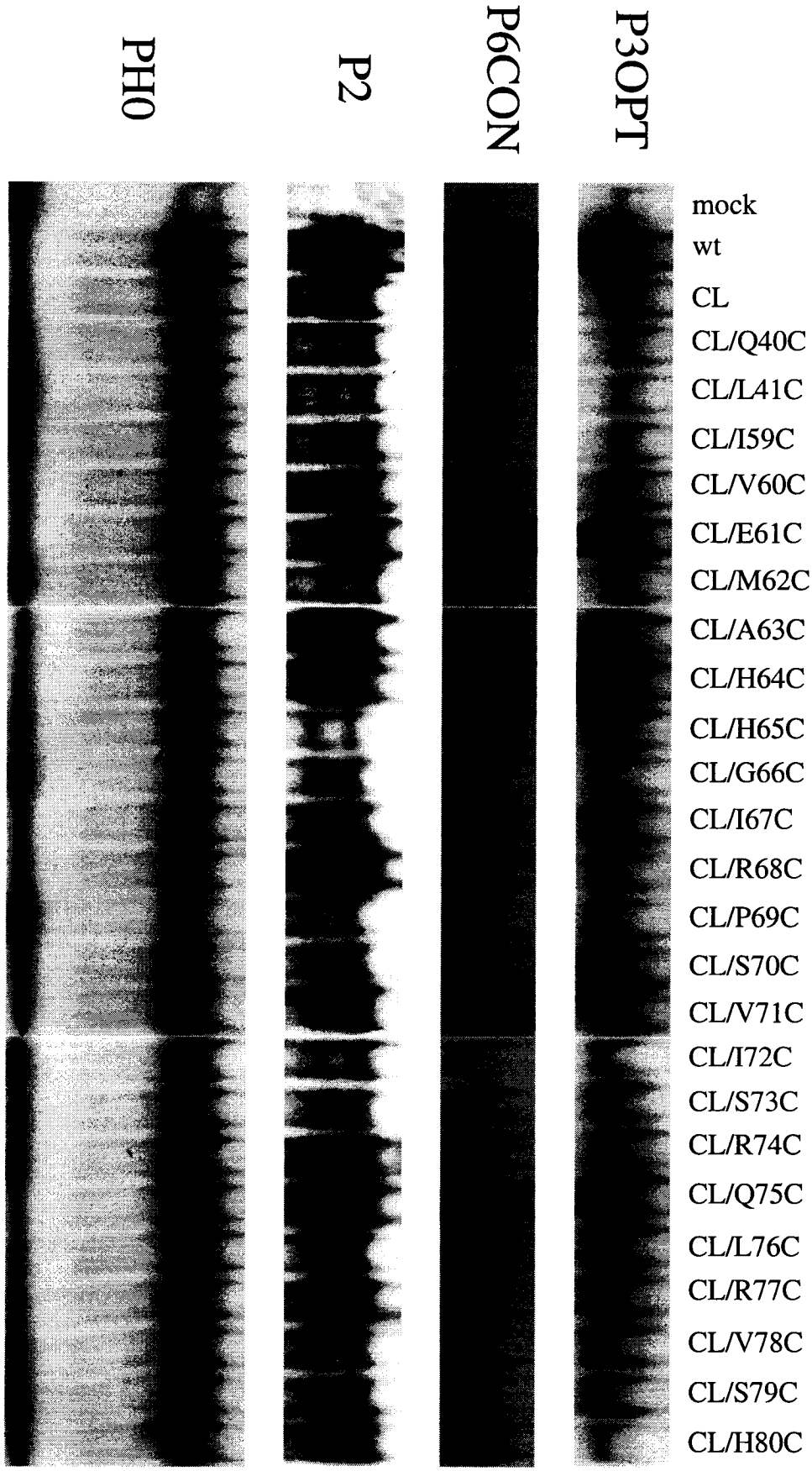

B
A
mock
wt
CL
CL/Q40C
CL/L41C
CL/I59C
CL/V60C
CL/E61C
CL/M62C
CL/A63C
CL/H64C
CL/H65C
CL/G66C
CL/I67C
CL/R68C
CL/P69C
CL/S70C
CL/V71C
CL/I72C
CL/S73C
CL/R74C
CL/Q75C
CL/L76C
CL/R77C
CL/V78C
CL/S79C
CL/H80C

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 PD (P6CON and P3OPT) and HD binding sites (P2), and against a probe with both PD and HD 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.

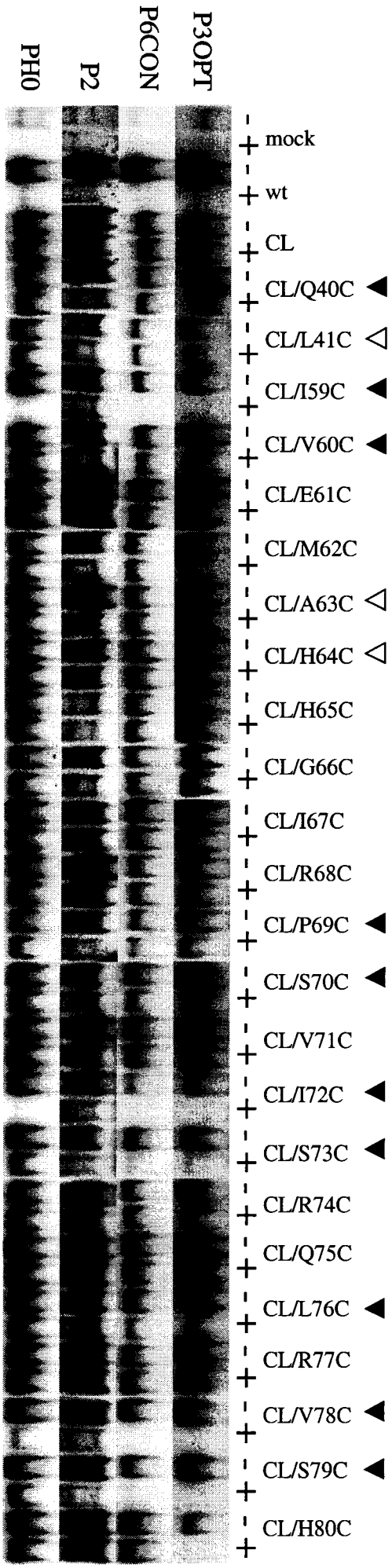

PH0
P2
P6CON
P3OPT
mock
wt
CL
CL/Q40C
CL/L41C
CL/I59C
CL/V60C
CL/E61C
CL/M62C
CL/A63C
CL/H64C
CL/H65C
CL/G66C
CL/I67C
CL/R68C
CL/P69C
CL/S70C
CL/V71C
CL/I72C
CL/S73C
CL/R74C
CL/Q75C
CL/L76C
CL/R77C
CL/V78C
CL/S79C
CL/H80C

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 (Wilson et al., 1993; Schafer et al., 1994). 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 (Underhill et al., 1995; Fortin et al., 1997). 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.

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).

Figure 4

Effect of site-specific modification on DNA binding properties of Pax3 mutants. Effects of *N*-ethyl maleimide (NEM) treatment on the PD and HD 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.

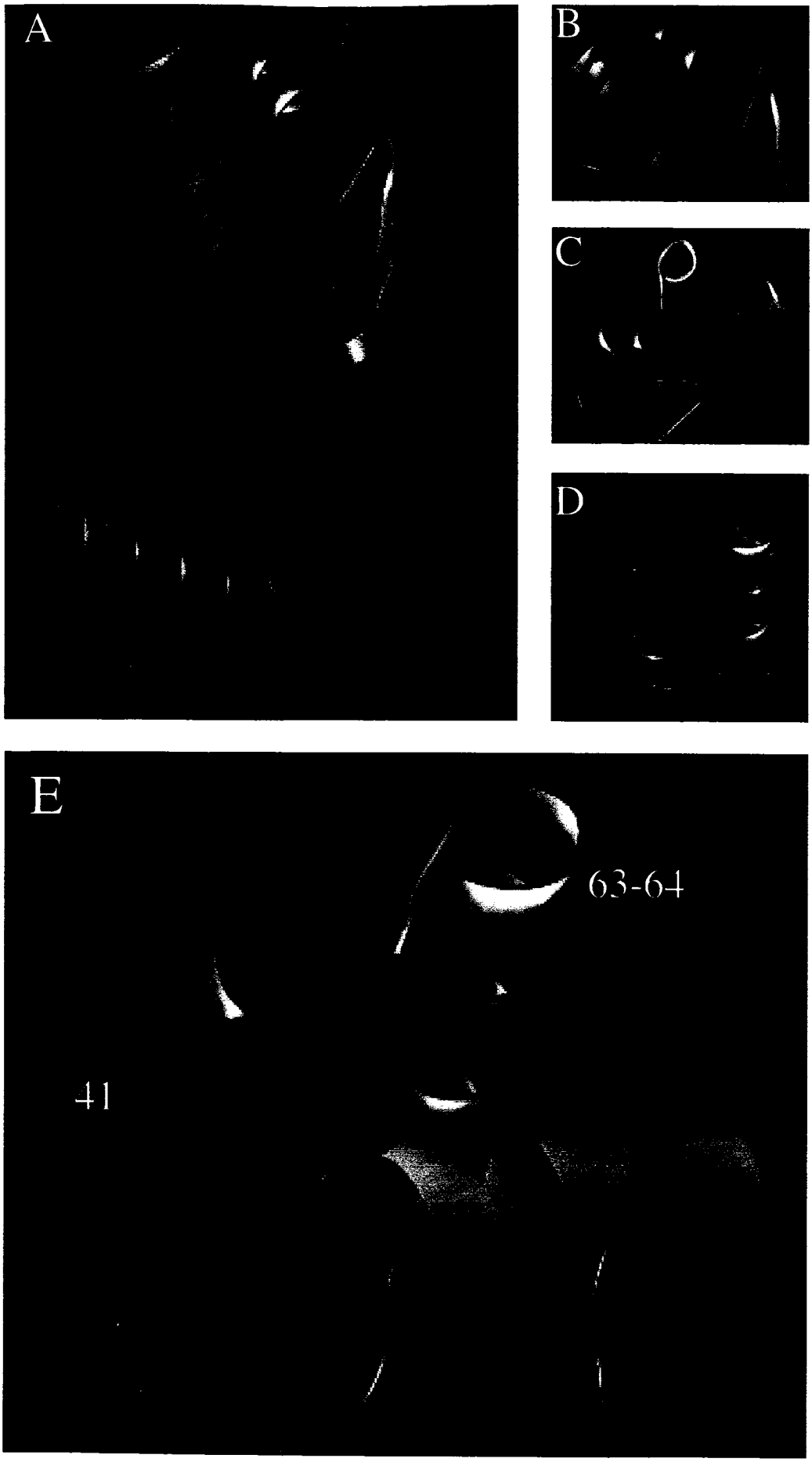

PH0
P2
P6CON
P3OPT
- +
mock
wt
CL
CL/Q40C ◀
CL/L41C ◁
CL/I59C ◀
CL/V60C ◀
CL/E61C
CL/M62C
CL/A63C ◁
CL/H64C ◁
CL/H65C
CL/G66C
CL/I67C
CL/R68C
CL/P69C ◀
CL/S70C ◀
CL/V71C
CL/I72C ◀
CL/S73C ◀
CL/R74C
CL/Q75C
CL/L76C ◀
CL/R77C
CL/V78C ◀
CL/S79C ◀
CL/H80C

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 (Xu et al., 1999). 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

Figure 5

Structure of the DNA-bound PD. The 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.


A
B
C
D
E
63-64
41

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-hairpin 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 (Baldwin et al., 1995; Glaser et al., 1992; Lalwani et al., 1995; Underhill et al., 1995; Fortin et al., 1997).

We have chosen to use site-specific modification of single cysteine mutants (Apuzzo et al., 2002; Frillingos et al., 1998; Loo et al., 2000) 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 (Apuzzo et al., 2002; Frillingos et al., 1998; Loo et al., 2000), 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 positions studied herein are indeed accessible to NEM at the concentration and temperature conditions used. Circular dichroism (CD) studies of Pax5 (Tell et al., 1998) and Pax8 (Tell et al., 1998) 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 (Apuzzo et al., 2004).

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 (Jun et al., 1996) 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 (Garvie et al., 2001). 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 (Garvie et al., 2001). 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 (Garvie et al., 2001). 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 (Xu et al., 1995; Xu et al., 1999). 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 (Xu et al., 1995; Xu et al., 1999). 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 (Xu et al., 1995; Xu et al., 1999). 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 (Xu et al., 1995; Xu et al., 1999). 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 (Xu et al., 1995; Xu et al., 1999). 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 (Fortin et al., 1998) 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 (Frillingos et al., 1998; Loo et al., 2000). 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 (Underhill et al., 1997). 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.

The identification of a putative HD binding site in the PD supports our model that PD: HD intramolecular physical interaction may be responsible for the functional interdependence between the two domains of Pax3. Also, the identification and characterization of the mechanism of this functional interdependence may constitute an essential step towards understanding how proteins holding homologous domains with similar DNA binding activities, such as Pax proteins, achieve functional diversity. This might be especially true for Pax proteins with a PD and a full-length HD, such as Pax3 and Pax6.

The next logical step to be undertaken is to identify residues or portions of the HD that are actually docking onto and physically interacting with residues in the PD. The area of the HD that we suspect may interact with the PD is the N-terminal arm. Our suspicion is partly based on the model of the Pax3 homolog, Prd, bound to the PH0 DNA sequence, which contains both PD and HD recognition sites. This model proposes that when the PD and HD of Prd is bound to the composite PH0 DNA sequence helix 2 of the PD comes into close proximity, and possibly interacts with, the N-terminal arm of the HD (Jun and Desplan., 1996). Also, protease sensitivity studies of Pax3 mutants presented in chapter 3 reveals that a conformational change can be detected in an area of Pax3 that lies immediately upstream of the HD N-terminal arm when the PD is DNA bound (Apuzzo et al., 2004). To assess whether the N-terminal arm of the HD docks onto the putative HD binding pocket in the PD a cysteine scanning mutagenesis approach was used. This involved the use of Pax3 double cysteine mutants as well as bifunctional thiol specific reagents that can crosslink two cysteines that lie in close proximity to one another.

Chapter 5

Cooperative Interactions between the two DNA Binding Domains of Pax3: Helix 2 of the Paired domain is in close proximity of the amino terminus of the Homeodomain

Abstract

Pax3 is a transcription factor that plays an important role during neurogenesis and myogenesis, and Pax3 mutant animals display neural tube defects and lack limb muscles. Pax3 harbors two DNA binding domains, the Paired domain (PD) and a paired-type Homeodomain (HD). Genetic and biochemical data have (i) identified strong cooperative interactions between the PD and HD domains for DNA binding in the intact Pax3 protein, and (ii) suggested an important role for the amino terminal portions of both domains in such cooperativity. We have studied proximity relationships between the PD and HD of Pax3. For this, we have used a cross-linking strategy with the bi-functional thiol reagent bis-maleimidoethane (BMOE) in 21 mutants bearing pairs of cysteine residues (DCM) inserted in strategic locations of a functional Pax3 protein otherwise devoid of endogenous cysteine residues. All 21 DCMs were characterized for protein stability, for DNA binding by the PD and HD, and for the effect of BMOE on protein binding to PD, HD or PD/HD combined DNA targets. BMOE-induced cross-links in DCMs were detected as slower migrating species on immunoblots. Mutants bearing double cysteine insertions I59C/ S222C, S73C/ Q219C and V78C/ K218C showed the most robust cross-linking upon BMOE exposure. These cross-linking studies suggests that portions of helix1 (I59), helix 2 (S73) and the loop between helix 2 and 3 (V78) of the PD are in close proximity to the N-terminal segment of the HD (K218, Q219 and S222) in the tertiary structure of Pax3. These results are compatible with a model in which the PD and HD are organized in an everted arrangement, with the N-terminal portion of the PD being in close proximity to the N-terminus of the HD. This arrangement may be important for the noted PD/HD cooperativity in DNA binding.

Introduction

Pax3 is a member of the Pax family, a group of transcription factors that play critical roles during mammalian development (Bopp et al., 1986; Stuart et al., 1994; Balling et al., 1988; Dahl et al., 1997; Glaser et al., 1992; Hill et al., 1991; Jordan et al., 1992; Macchia., 1998; Mansouri et al., 1998; Sanyanusin et al., 1995; Sosa-Pineda et al., 1997; Torres et al., 1995). Pax3 is expressed in a number of embryonic structures (somites, the neural tube, several neural crest cell derived lineages) and is essential for the normal process of myogenesis and neurogenesis (Bober et al., 1994; Goulding et al., 1991). A mutation in the mouse *Pax3* gene causes severe neural tube defects (spina bifida, exencephaly) and absence of limb musculature; likewise, in humans, mutations in *PAX*3 cause Waardenburg syndrome (WS), a condition associated with sensorineural deafness, cranio-facial abnormalities and pigmentary disturbances (Goulding et al., 1991; Auerbach et al., 1954; Baldwin et al., 1992; Baldwin et al., 1995; Beechey et al., 1986; Franz et al., 1989; Franz et al., 1990; Goulding et al., 1994). Pax3 appears to orchestrate the expression of a number of target genes during normal development, but aberrant *PAX*3 activity causes alveolar rhabdomyosarcoma (Barr et al., 1993; Galili et al., 1993). Structurally, Pax proteins are defined by the presence of a unique DNA binding domain called the PD. In addition, Pax3 and several other Pax proteins contain a second DNA binding domain, a paired-type HD (Stuart et al., 1994; Noll et al., 1993). Other structural domains of Pax proteins include a conserved octapeptide (OP) in the segment linking the PD and HD and proline/serine/threonine-rich (PST) transactivation domain at the carboxyl terminus.

High resolution three-dimensional structures (Xu et al., 1995; Xu et al., 1999) show that the PD is formed by N-terminal (PAI) and C-terminal (RED) subdomains, each consisting of a three helical fold that includes a helix-turn-helix (HTH) motif, with the C-

terminal helix of each HTH making base-specific contacts in the major groove of DNA (Xu et al., 1995; Xu et al., 1999). The linker joining the PAI and RED subdomains also makes base-specific contacts in the minor groove of DNA. In addition to the HTH motif, the PAI subdomain harbors a β-turn structure at its amino terminus that also makes critical contacts in the minor groove of DNA (Xu et al., 1995; Xu et al., 1999). The DNA sequences recognized by the PD can be bound by both subdomains (Class I sequences) or only by the PAI subdomain (Class II sequences) (Czerny et al., 1993; Epstein et al., 1994). The structure of the DNA-bound Prd paired-type HD shows 3 α helices, the last 2 forming a HTH motif (Kappen et al., 1993; Treisman et al., 1989; Wilson et al., 1993; Wilson et al., 1995). The C-terminal helix of the HTH makes base specific contacts in the major groove of DNA and is responsible for DNA binding specificity. The N-terminal segment preceding the HTH motif makes additional contacts in the minor groove of DNA. A unique feature of pt-HDs is their ability to dimerize on palindromic sequences of the type TAAT(N_{2-3})ATTA (Treisman et al., 1989; Wilson et al., 1993; Wilson et al., 1995).

Although the PD and HD can bind to cognate DNA sequences when expressed individually, a large body of genetic and biochemical data indicate that the two domains are functionally interdependent in intact Pax3. For instance, the G42R mutation in the PD of Pax3 from the *Splotch-delayed* (Sp^d) mouse mutant abrogates DNA binding by the PD but also impairs DNA binding by the HD. Interestingly, deleting helix 2 of the PAI subdomain in the context of the G42R mutation (Sp^d) restores DNA binding by the HD suggesting that this helix is involved in cooperative interaction between the PD and HD of Pax3 (Underhill et al., 1995). Studies in chimeric Pax3 proteins have shown that the PD can modulate both DNA binding specificity and dimerization potential of heterologous HDs (Fortin et al., 1998). Conversely, studies of the HD mutation R53G found in a Waardenburg patient shows that it impairs DNA binding by both the HD and the PD (Fortin et al.,1997). Site-specific

modification of single cysteine Pax3 mutants with thiol reagents have shown that independent modification of position 82 in the PD and position 263 in the HD (V263C) abrogate DNA binding by both the PD and HD (Apuzzo et al., 2002). Finally, protease sensitivity studies in Pax3 mutants bearing engineered Factor Xa sites either in the linker separating the PAI and RED motif (position 100), or upstream the HD (position 216), revealed that DNA binding by either the PD or the HD causes a structural change in the other DNA binding domain (Apuzzo et al., 2004), providing a structural basis for the observed interdependence in DNA binding. Additional studies with single cysteine mutants independently introduced in the second α helix (α2, positions 59-80) and in the β-hairpin (positions 40-41) at the amino terminus of the PD identified a number of positions (Q40, I59, V60, P69, S70, I72, S73, L76, V78, S79C) at which NEM modification abrogated DNA binding by the PD and HD (Apuzzo et al., 2006). Three dimensional modeling revealed that these residues are not randomly distributed, but rather that they cluster in a hydrophobic pocket, representing a possible docking site for the HD during DNA binding by the intact protein (Apuzzo et al., 2006).

The objective of the present study was to characterize further the structural basis of the HD/PD functional interaction in Pax3, and possibly identify individual residues of the HD that may come in close proximity to the PD during DNA binding. Proximity relationships were investigated in Pax3 mutants bearing pairs of single Cysteines inserted at different positions of the amino terminus of the PD and of the HD respectively (double cysteine mutants, DCMs), followed by cross-linking with the thiol-specific bi-functional reagent BMOE. For this, we created 21 Pax3 mutants bearing pairs of cysteine between positions 59, 73, and 78 in the amino terminus of the PD on the one hand, and 7 positions (215-218, 219, 222, 225) in the amino terminus of the HD on the other hand. The ability of BMOE to induce cross-linking of cys pairs in Pax3 mutants bound to different DNA targets recognized by either the PD (P3OPT), the HD (P2) or both (PH0) was evaluated. These cross-linking studies

suggest that portions of the PD α1 (I59), α2 (S73) and the loop joining α2 and α3 (V78) are in close proximity to the N-terminus of the HD (K218, Q219 and S222) in the tertiary structure of Pax3.

Material and Methods

Mutagenesis. The construction of the pMT2 expression plasmid containing the entire coding region of wild type (WT) *Pax3* cDNA has been described (Underhill et al., 1995; Apuzzo et al., 2002). *Pax3* was modified by the in-frame addition of c-Myc and HA epitopes at the N and C-termini respectively, thereby generating the pMT2/Myc-Pax3-HA WT construct (Apuzzo et al., 2004). The construction and functional characterization of a Pax3 mutant lacking cysteines (cys-less, CL) has been described (Apuzzo et al., 2002), and this CL Pax3 was similarly modified by insertion of epitopes to create pMT2/ Myc-Pax3-HA CL. The generation of the 3 PD (CL/I59C, CL/S73C, CL/V78C) single cysteine mutants used in this study was previously described (Apuzzo et al., 2006). The 7 other single cysteine mutants have cysteines introduced in N-terminal arm and upstream linker of the HD (CL/L215C, CL/K216C, CL/R217C, CL/K218C, CL/Q219C, CL/S222C and CL/T225C) by PCR mutagenesis using the pMT2/ Myc-Pax3-HA CL as a template and with primers listed in Table I. Pax3 contains a *KpnI* cassette defined by two endogenous *KpnI* restriction sites at positions 563 and 1500. Swapping of *KpnI* cassettes of linker/HD SCM constructs into the 3 PD SCM constructs permitted the generation of the following 21 double cysteine mutants: I59C/L215C, I59/K216C, I59C/R217C, I59C/K218C, I59C/Q219C, I59C/S222C, I59C/T225C, S73C/L215C, S73C/K216C, S73C/R217C, S73C/K218C, S73C/Q219C, S73C/S222C, S73C/T225C, V78C/L215C, V78C/K216C, V78C/R217C, V78C/K218C, V78C/Q219C, V78C/S222C and V78C/T225C. Each mutation was verified by nucleotide

Table 1.Oligonucleotides used for Pax3 mutagenesis

Substitution	Mutagenic primer (5'-3')
SCM	
I59C	ATCCGCCACAAG***TGT***GTGGAGATGGCC
S73C	CCGAGCGTCATTT***G***TCGCCAGCTTCGC
V78C	CGCCAGCTTCGC***TGT***TCCCATGGATCC
L215C	CCTGATTTACCG***TGT***AAGAGGAAGCAGCGCCGATCGAGAACCACCTTC
K216C	GATTTACCGCTG***TGT***AGGAAGCAGCGCAGGTCTAGAACCACCTTC
R217C	GACTCTGAACCAGATCTACCGCTGAAG***T***G***T***AAGCAGCGCAGGTCGCGAACCACCTTC
K218C	CCGCTGAAGAGG***TGT***CAGCGCAGGTCGCGAACCACCTTC
Q219C	GATGAAGGATCCGATATTGACTCTGAACCTGATTTACCGCTGAAGAGGAAG***TGT***CGCAGGAGC
S222C	CAGCGCAGG***T***GCCGTACGACCTTCACG
T225C	GAGCAGAACC***TG***CTTCACGGCAGAGCAGCTGGAGGAACTGGAGCGCGCGTTCGAGAG

Nucleotide substitutions leading to amino acid changes are in bold and italicized and those that introduce silent restriction sites are underlined.

sequencing, and accessibility of restriction sites for cloning was verified by endonuclease fragmentation.

Expression and Detection of Pax3Mutants. 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 co-precipitation 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 phenylmethysulfonyl 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). Subsequently, the immunoblots were submerged in stripping buffer (100 mM 2-mercaptoethanool, 2% SDS, 62.5 mM Tris-HCl, pH 6.7, 30 min at 50 °C), and washed with TBST buffer (10 mM Tris-HCl, pH 8, 150 mM NaCl, 0.1% Tween 20, 20 °C) at room temperature. The membranes were then 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 (Apuzzo et al., 2002). Each protein:DNA binding reaction was carried out using approximately 8 μg of total cell extracts from transiently transfected COS-7 monkey cells and 10fmol (0.06 μCi) of radioactively labeled double stranded oligonucleotides corresponding to different Pax3 target sequences. The final concentration of labeled oligonucleotide in the binding reaction is 0.5nM. Whole cell extracts were incubated with ^{32}P-labeled 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 12V/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')-TGGAATTCAGGAAAAATTTTCACGCTTGAGTTCACAGCTCGAGTA-(3') (Epstein et al., 1994) and P3OPT (5')-TGGTGGTCACGCCTCATTGAATATTA-(3') (Chalepakis et al., 1995; Epstein et al., 1995), HD-specific sequence P2 (5')-GATCCTGAGTCTAATTGATTACTGTACAGG-(3') (Wilson et al., 1993) and the composite PD/HD binding site PH0 (5')-GATTTCTTCCAATTAGTCACGCTTGAGTG-(3') (Jun et al., 1996) were synthesized as complementary oligonucleotide pairs with 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, Pierce) was

prepared as a 32 mM stock in water. Bis-Maleimidoethane (BMOE), a sulfhydryl-to-sulfhydryl cross-linking reagent (Molecular Probes) was prepared as a 32mM stock in 100% dimethyl sulfoxide. To assess the effect of NEM or BMOE on DNA binding they were added as a 0.5 μL aliquot to a 4 μL volume of whole cell extract (final concentration of 3.55mM), followed by a 30 min incubation at room temperature prior to the addition of the [^{32}P]-labeled probe and EMSA. Western blotting was carried out with whole cell extracts treated with non-radiolabeled DNA prior to BMOE treatment to assess the extent of cross-linking in double cysteine mutants. When BMOE-treated whole cell extracts were destined for Western blotting, they were treated with NEM (3.55 mM, 30 minutes at 20 °C) prior to SDS-PAGE to block any remaining un-modified cysteines and to prevent non-specific cross-linking. Films generated from Western blotting were used to perform densitometry studies to quantify the amount of chemiluminescence using a Fuji LAS-1000, as we have previously described (Apuzzo et al., 2002; Apuzzo et al., 2004).

Results

Construction of Pax3 Mutants Bearing Single or Double Cysteine Insertions.

A large body of biochemical and genetic data indicates that DNA binding by Pax3 involves cooperative interactions between the PD and the HD. Indeed, point mutations (Fortin et al., 1997) or site-specific modifications of single cysteines in either PD or HD abrogate DNA binding by both domains (Apuzzo et al., 2002), and DNA binding by either PD or HD causes a structural change in the other domain (Apuzzo et al., 2004). In addition, a number of studies have suggested that this functional cooperation may involve physical interactions between the two domains. The goal of the present study was to investigate proximity relationships between the PD and HD, including the identification of specific residues mapping at the interface of such interactions. We targeted the amino terminal portions of both the PD and the HD as two sites that are most likely to physically interact in intact Pax3. This was based on the following published data. DNA binding studies with the *Drosophila* Prd protein using hybrid DNA sequences (PH0) optimized for combined PD and HD binding show that the two domains are everted (Jun et al., 1996). Additional modeling studies suggest that the N-terminus of the PAI subdomain is closely apposed to the N terminus of the HD when bound to the combined PH0 site (Jun et al., 1996), with helix 2 of the PAI domain in very close proximity of the N-terminal extension of the HD. A critical role of the PAI subdomain helix 2 in interaction with the HD is supported by functional studies showing that deletion of helix 2 in the context of the Sp^d mutant allele of Pax3 (G42R) restores DNA binding by the HD in the context of the mutant protein (Fortin et al.,1998). Finally, studies in Pax3/Phox chimeras suggest that the C-terminal portion of the linker domain separating the PD and HD contributes to functional interactions between the 2 domains in DNA binding (Fortin et al., 1998).

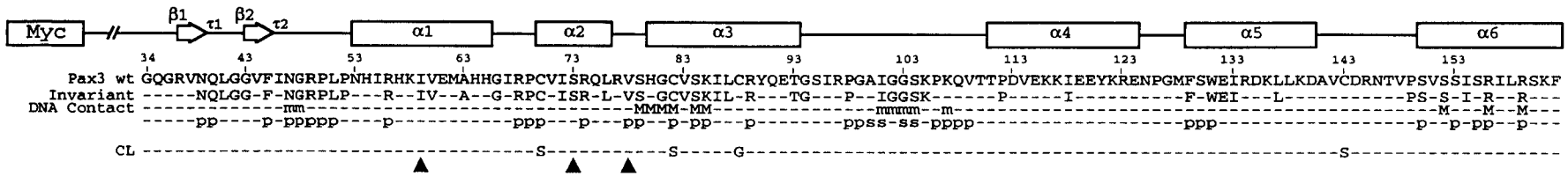
Here, we investigated proximity relationships in the Pax3 protein by cross-linking. For this, we used a Pax3 mutant that is devoid of cysteine (Cys-less, CL) residues but that retains wild type DNA binding properties towards PD or HD target sequences (Apuzzo et al., 2002; Apuzzo et al., 2006). We created double cysteine mutants (DCM) bearing 1 cysteine in individual positions of each the PD and the HD; proximity of the 2 cysteines was assessed by the ability of a bi-functional, thiol-specific cross-linking agent (BMOE) to make a covalent adduct between the two targeted positions. The effect of DNA binding by the PD and the HD on the formation of such adducts was also investigated. The implementation of this strategy requires that both inserted Cys residues be at exposed positions accessible to the sulfhydryl reagent. For this, we selected positions at which Cys insertion do not affect DNA binding, but yet confer N-ethyl maleimide (NEM) modification sensitivity of DNA binding in the context of a single Cys mutant. Such single Cys mutants were then systematically reconstructed as DCM and functionally analyzed.

Previous cysteine scanning mutagenesis identified a hydrophobic pocket in the N-terminal segment of the PAI subdomain at which insertion of single Cys confer NEM sensitivity to DNA binding by the PD and HD in the intact protein (Apuzzo et al., 2006). In this proposed docking site for the HD, positions I59 (α helix 1), S73 (α helix 2) and V78 (α2-α3 linker) of the PAI subdomain (Figure 1) were most sensitive to NEM treatment and were selected for potential PD candidates for the current cross-linking studies. In the HD the N-terminal arm (first 9 residues) was targeted for study. In preliminary experiments, substitution of each of these 9 residues to cysteine in SCM showed that only 3 positions (Q219, S222, T225) could tolerate this replacement without loss of DNA binding by the HD (data not shown). These together with 4 adjacent residues from the linker domain (L215, K216, R217, K218) were selected for potential HD candidates for cross-linking studies (Figure 1).

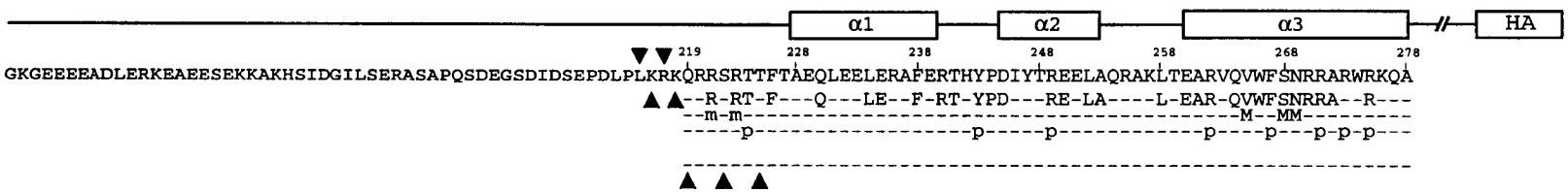
Figure 1

Site-directed Mutagenesis of the Paired Domain of Pax3. The amino acid sequence for positions 34-278 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) (Xu et al., 1999). Invariant residues amongst 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. Residues targeted when generating single cysteine mutants are indicated with filled arrowheads. The figure also includes a schematic representation of the Pax3 linker and homeodomain (HD), including the presence and position of predicted structural features, invariant amino acid residues, number and types of DNA contacts (as for the PD). 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.

PD


Myc
β1
τ1
β2
τ2
α1
α2
α3
α4
α5
α6
34 43 53 63 73 83 93 103 113 123 133 143 153
Pax3 wt GQGRVNQLGGVFINGRPLPNHIRHKIVEMAHHGIRPCVISRQLRVSHGCVSKILCRYQETGSIRPGAIGGSKPKQVTTPDVEKKIEEYKRENPGMFSWEIRDKLLKDAVCDRNTVPSVSSISRILRSKF
Invariant -----NQLGG-F-NGRPLP---R--IV--A--G-RPC-ISR-L-VS-GCVSKIL-R---TG---P--IGGSK------P-----I----------F-WEI---L----------PS-S-I-R--R---
DNA Contact ----------mm----------MMMM-MM----------mmmm--m----------M---M--M---
-----pp----p-ppppp----p----------ppp---p---pp--p-pp---p--------ppss-ss-pppp----------ppp----------p--p-pp--p---
CL ----------S----------S-----G----------S----------

HD


α1
α2
α3
HA
219 228 238 248 258 268 278
GKGEEEEADLERKEAEESEKKAKHSIDGILSERASAPQSDEGSDIDSEPDLPLKRKQRRSRTTFTAEQLEELERAFERTHYPDIYTREELAQRAKLTEARVQVWFSNRRARWRKQA
--R-RT-F---Q---LE--F-RT-YPD---RE-LA----L-EAR-QVWFSNRRA--R---
--m-m----------M--MM---------
-----p----------p-----p----------p----p---p-p-p---

The following single cysteine mutants (SCM) were created in the PD (CL/I59C, CL/S73C and CL/V78C), and in the N-terminal part of the HD (CL/L215C, CL/K216C, CL/R217C, CL/K218C, CL/Q219C, CL/S222C and CL/T225C) using the Pax3 CL backbone as a template. In addition, double cysteine mutants (DCM) were created by combining each of the 3 PD mutations with each of the 7 HD mutations in a group of 21 DCMs.

Protein Expression and DNA Binding Properties of Pax3 Single and Double Cysteine Mutants.

To facilitate detection of wild type (WT), Cys-less (CL) and all single (SCM) and double (DCM) cysteine mutants, cMyc 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 plasmid and were used to transiently transfect COS-7 monkey cells. Immunoblotting of whole cell extracts with either anti-HA or anti-c-Myc monoclonal antibodies show that all SCMs and DCMs can be expressed at similar levels in COS-7 cells (Figure 2). This indicates that none of the mutations had a major effect on protein expression or stability in COS-7 cells. The effects of single and double cysteine substitutions on DNA binding properties of the PD and the HD of Pax3 were assessed in each mutant by electrophoretic mobility shift assays (EMSAs). The DNA binding activity of the PD was measured using the PD specific probes P3OPT (Chalepakis et al., 1995; Epstein et al.,1995) and P6CON (Epstein et al., 1994) which have been shown to contain binding determinants for both the PAI and RED subdomains of the PD. The DNA binding activity of the HD was determined using the P2 target sequence, which contains the TAAT(N_2)-ATTA sequence previously shown to support cooperative dimerization of Pax3. Finally, all mutants were assessed for DNA binding activity to the PD-HD composite site present in the PH0 probe. All SCMs, and DCMs retained PD DNA binding activity for both the P3OPT and P6CON probes

which was similar to that displayed by the WT or CL Pax3 protein (Figure 3). Most mutants retained DNA binding by the HD and showed dimerization on the P2 probe, with the possible exception of SCM CL/L215C and the DCMs S73C/ R217C and V78C/ S222C that showed decreased binding to this target. However, all mutants showed wild type binding activity for the composite probe PH0 (Figure 3). These results indicate that most of the mutations either alone or in combination have no effect on DNA binding by Pax3, and none appear to alter the ability of the protein to bind a PD/HD composite site.

Effect of Thiol specific reagents on DNA Binding by Single and Double Cysteine Mutants.

Both *N*-ethyl maleimide (NEM) and bis-maleimidoethane (BMOE) can covalently modify and alkylate cysteine residues and were used for site-specific modification of SCMs and DCMs. BMOE contains two maleimide functional groups, as opposed to one with NEM, and thus can create covalent adducts (cross-links) between cysteines showing spatial proximity in an otherwise intact protein. The effect of mono-functional NEM (Supplementary figure S1) and bi-functional BMOE (Figure 4) on DNA binding properties of WT, CL and of single and double Cys mutants was evaluated by EMSA, using PD (P3OPT), HD (P2) and combined PD-HD (PH0) target sequences. Whole cell extracts from COS-7 cells expressing, CL or SCMs or DCMs were incubated with either 3.55mM NEM or 3.55mM BMOE prior to testing DNA binding properties by EMSA. For control WT Pax3, treatment with either NEM or BMOE abrogates DNA binding by the PD and HD for both individual and combined target sequence. On the other hand, DNA binding by the Pax3-CL control protein to all probes tested was completely insensitive to NEM and BMOE treatment (Supplementary Figure S1; Figure 4).

The effect of NEM (Supplementary Figure S1) or BMOE treatment (Figure 4) on the DNA binding properties of SCMs and DCMs was identical for each mutant, and only results

Figure 2

Expression of Pax3 Mutants in COS-7 Monkey cells. Wildtype, Cys-less, SCMs and DCMs were cloned into pMT2 expression plasmid, and total cell extracts from transiently transfected COS-7 monkey cells were separated by SDS-PAGE (12% acrylamide) and transferred to nitrocellulose membranes. Immunodetection of Pax3 was done with mouse anti-c-Myc (MYC) and anti-HA (HA) monoclonal antibodies and a horseradish peroxidase-conjugated secondary antibody. The "mock" labeled lane refers to whole cell extracts from untransfected COS7 cells.

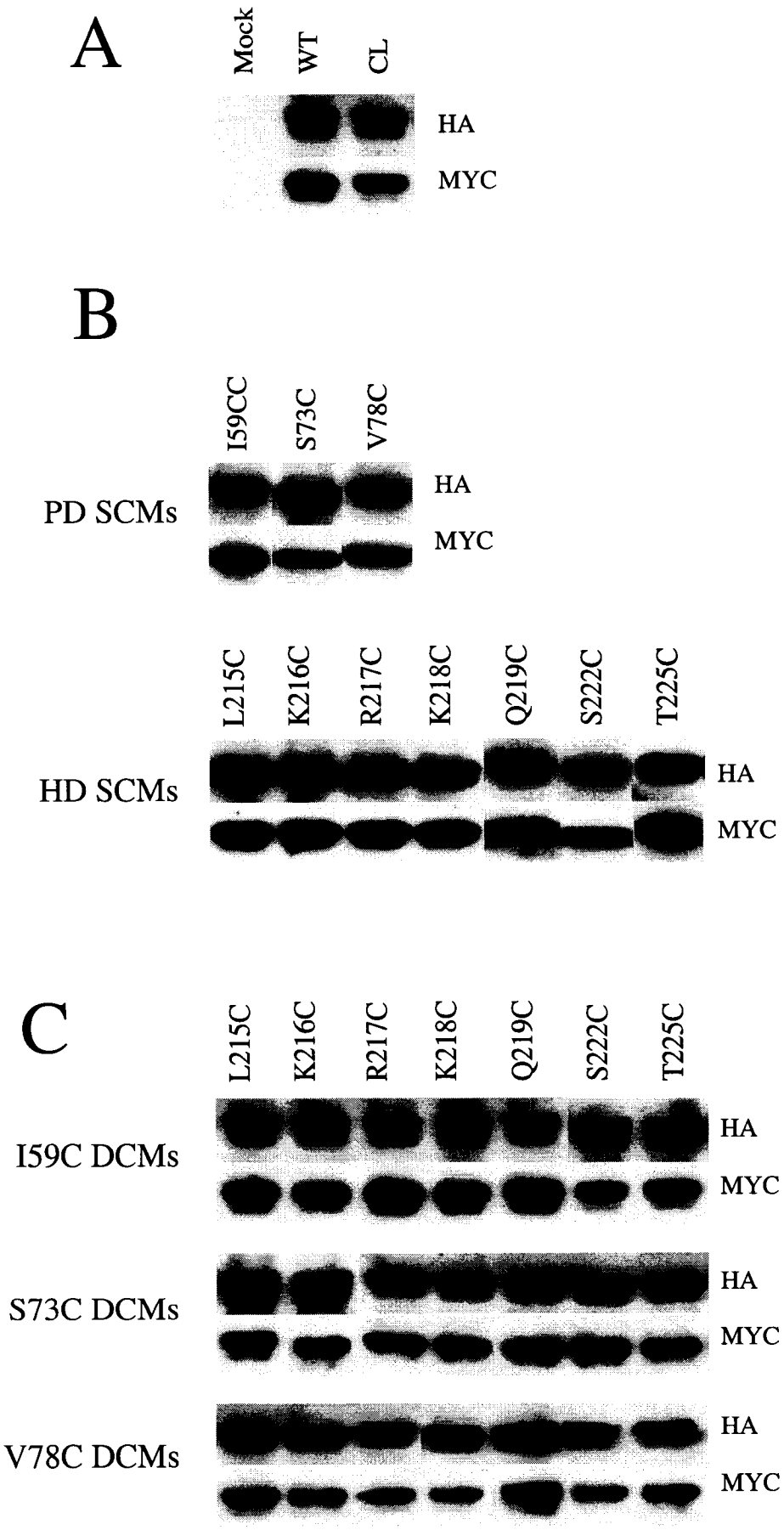

A
Mock
WT
CL
HA
MYC
B
I59CC
S73C
V78C
PD SCMs
HA
MYC
L215C
K216C
R217C
K218C
Q219C
S222C
T225C
HD SCMs
HA
MYC
C
L215C
K216C
R217C
K218C
Q219C
S222C
T225C
I59C DCMs
HA
MYC
S73C DCMs
HA
MYC
V78C DCMs
HA
MYC

Figure 3

Paired domain and homeodomain DNA binding properties of wild-type, Pax3CL, SCMs and DCMs. Electrophoretic mobility shift assays were used to measure the DNA binding properties of Pax3 (WT), Cys-less (CL), single (SCM) and double cysteine Pax3 mutants (DCMs) against paired domain (P3OPT, P6CON), homeodomain (P2) and a paired domain/homeodomain composite site (PH0). Protein-DNA complexes were formed using total cell extracts from transiently transfected COS-7 cells, and were resolved on 6% acrylamide non-denaturing gels, as described in Materials and Methods.

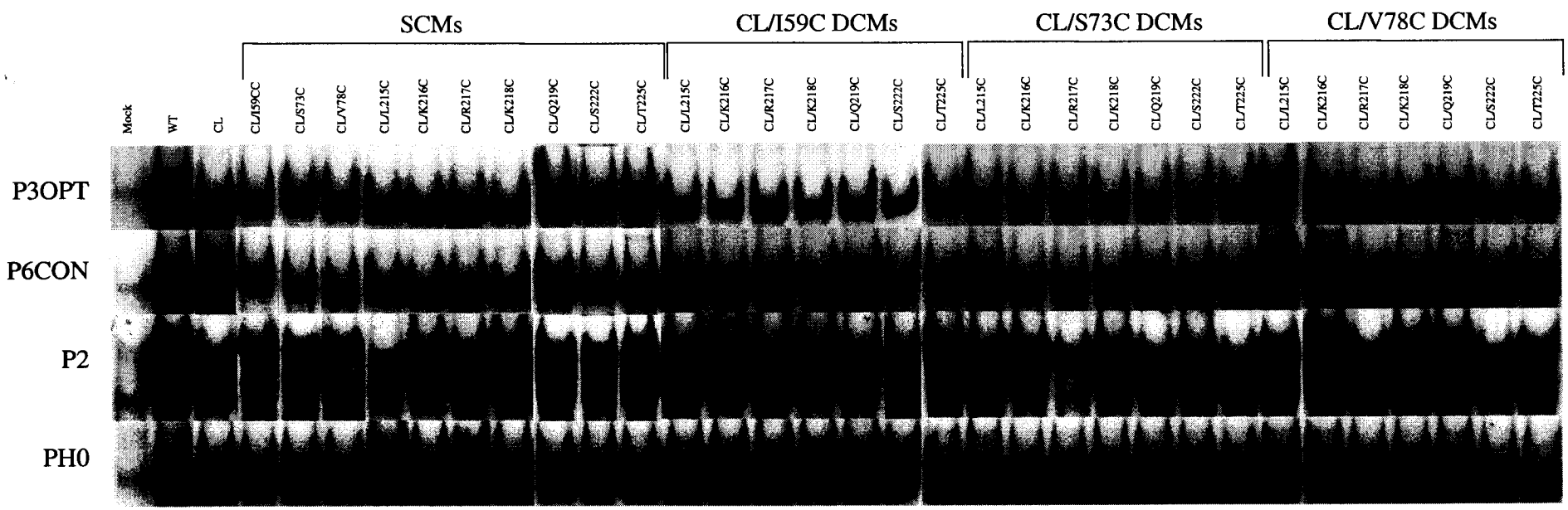
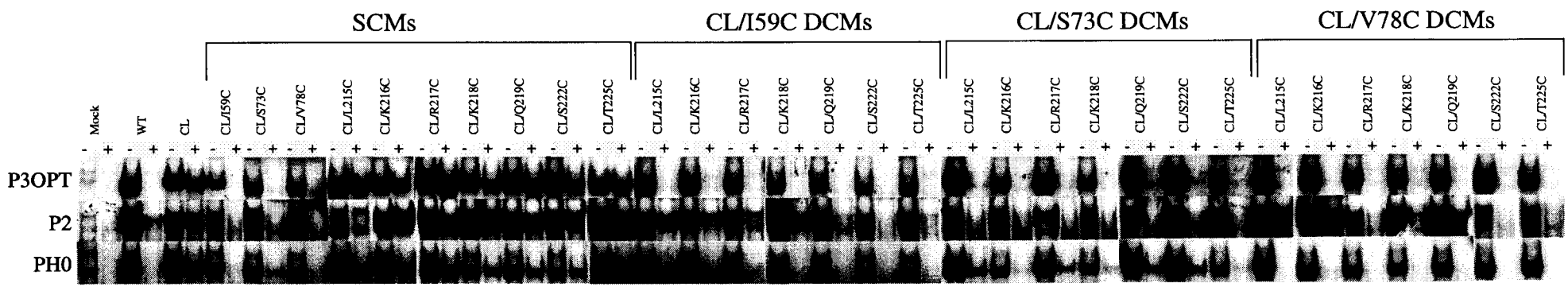

SCMs
CL/I59C DCMs
CL/S73C DCMs
CL/V78C DCMs
Mock
WT
CL
CL/I59CC
CL/S73C
CL/V78C
CL/L215C
CL/K216C
CL/R217C
CL/K218C
CL/Q219C
CL/S222C
CL/T225C
P3OPT
P6CON
P2
PH0

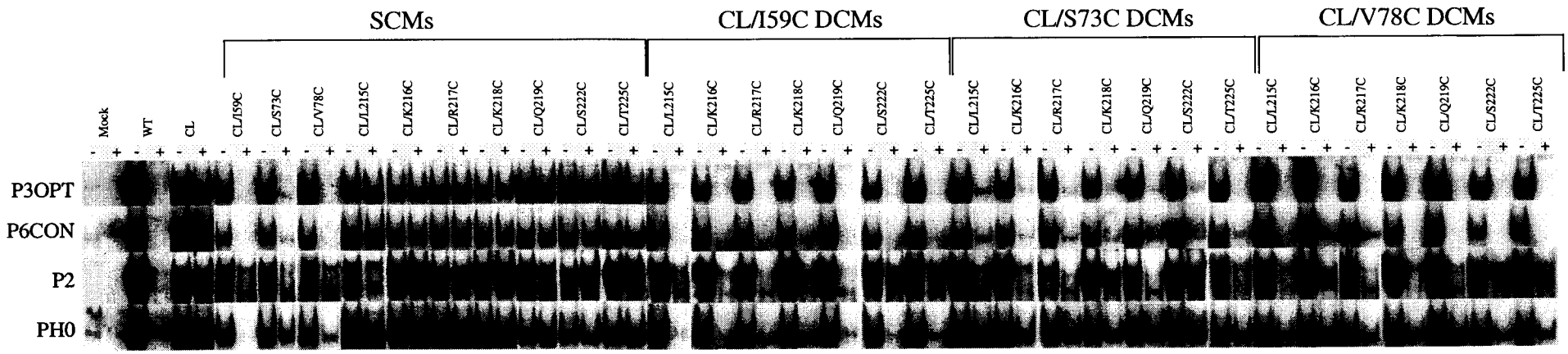
Figure 4

Effect of site-specific Modification on DNA binding properties of Pax3 Mutants. Effects of bis-maleimidoethane (BMOE) treatment on the paired domain and homeodomain DNA binding properties of wild type (WT), Cys-less (CL), single (SCM) and double (DCM) cysteine mutants. Whole cell extracts from COS7 monkey cells (mock) or from cells expressing different Pax3 mutants were incubated with 0 (-) or 3.55mM (+) BMOE prior to electrophoretic mobility shift assays. The DNA binding properties of the different Pax3 proteins were tested against paired domain (P3OPT, P6CON), homeodomain (P2) and a paired domain/homeodomain composite site (PH0).


SCMs
CL/I59C DCMs
CL/S73C DCMs
CL/V78C DCMs
Mock
WT
CL
CL/I59C
CL/S73C
CL/V78C
CL/L215C
CL/K216C
CL/R217C
CL/K218C
CL/Q219C
CL/S222C
CL/T225C
P3OPT
P2
PH0

Supplementary Figure 1

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), single cysteine mutants and double cysteine mutants was determined. Whole cell extracts from COS7 monkey cells (mock) or from cells expressing different Pax3 proteins were incubated with 0 (-) or 3.55mM (+) 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.


SCMs
CL/I59C DCMs
CL/S73C DCMs
CL/V78C DCMs
Mock
WT
CL
CL/I59C
CL/S73C
CL/V78C
CL/L215C
CL/K216C
CL/R217C
CL/K218C
CL/Q219C
CL/S222C
CL/T225C
P3OPT
P6CON
P2
PH0

obtained with BMOE will be reviewed. As we have previously reported for NEM (Apuzzo et al., 2006), alkylation of PD positions 59 (I59C), 73 (S73C), and 78 (V78C) by the BMOE reagent abrogated DNA binding by the PD and HD in the corresponding SCM (Figure 4). This formally verifies that sulfhydryl groups introduced at these positions are accessible to the bi-functional BMOE reagent. Modification of single cysteines inserted at either of the seven positions of the N-terminal portion of the HD was without consequences on DNA binding of corresponding SCMs to the P3OPT probe. On the other hand, BMOE modification of the L215C, R217C, Q219C and S222C SCM mutants had some effect on DNA binding by the HD to the P2 probe. Nevertheless, BMOE treatment had no effect on binding to the PD-HD composite site in any of the 7 HD SCMs tested (Figure 4).

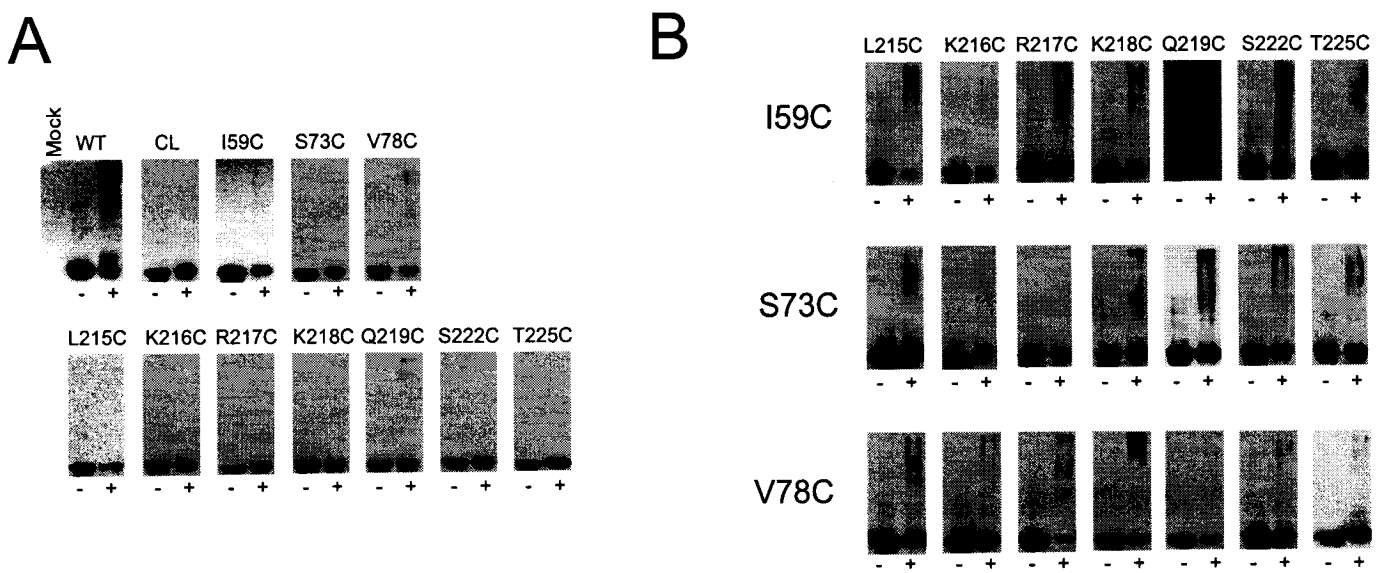
As expected from the observed BMOE-sensitivity of DNA binding by the single mutants I59C, S73C, and V78C, all DCMs showed BMOE sensitivity for all probes tested. DCMs V78C/L215C and V78C/K216C were notable exceptions: although both showed BMOE sensitive DNA binding to PD and PD-HD probes, they showed partial (V78C/L215C) or complete (V78C/K216C) insensitivity for binding to the P2 probe.

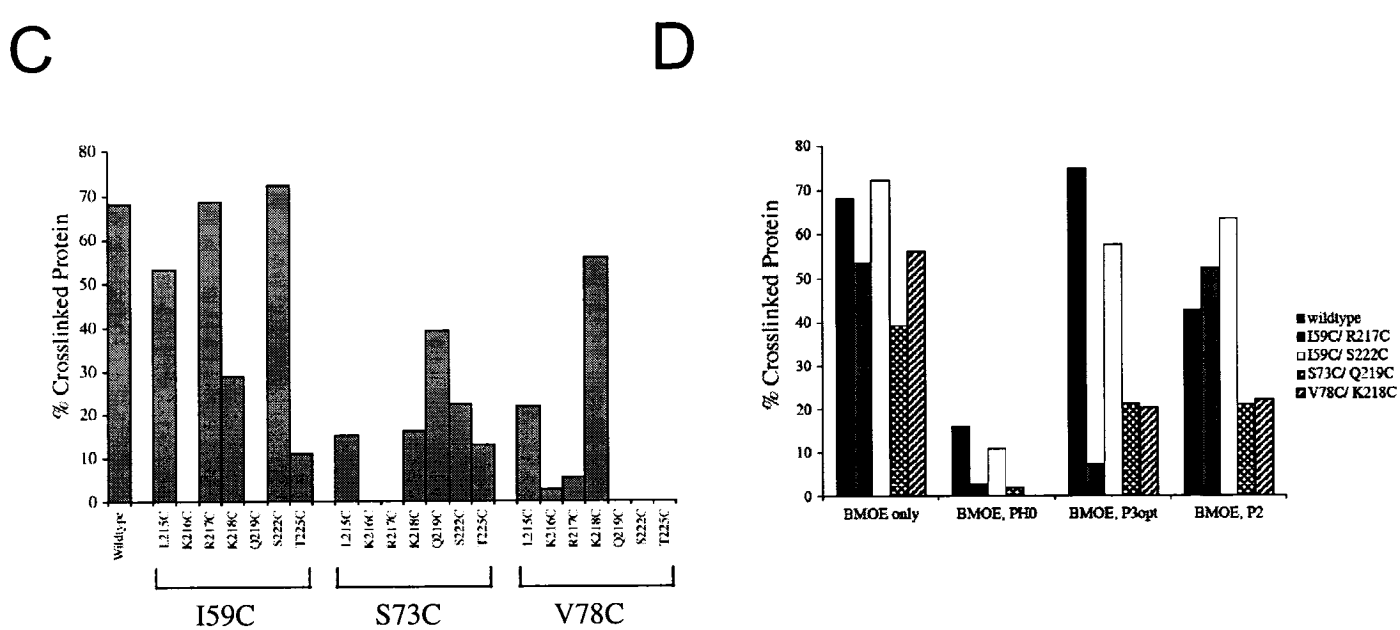
Cross-Linking Studies in Single and Double Cysteine Mutants.

Proximity relationships between pairs of cysteines were studied by cross-linking in DCMs. In these experiments, WT, CL as well as all SCMs and DCMs were treated with BMOE and cross-linked species were detected by SDS-PAGE and immunoblotting with monoclonal anti-HA (Figure 5) or anti-cMyc antibody (data not shown). In these experiments, cross-linked Pax3 proteins are detected as higher molecular mass species of slower electrophoretic mobility on denaturing, detergent-containing (SDS) polyacrylamide gels (Loo et al., 2001). Exposure of WT Pax3 to BMOE followed by SDS-PAGE and immunoblotting results in the appearance of cross-linked Pax3 species of slower mobility that are absent from

Figure 5

Detection and Quantification of Cross-linking Species of Pax3 mutants. Fifteen micrograms of whole cell extracts from COS-7 cells transfected with either wild type (WT), Cys-less (CL), or with the different single cysteine mutants (SCMs) (Panel A) or with double cysteine mutants (panel B) were incubated with 3.55mM BMOE and were analyzed by SDS-PAGE on 12% acrylamide gels and by immunoblotting with anti-HA antibody. Cross-linked species are visible as higher molecular mass bands. (C) The immunoblots of DNA unbound, BMOE exposed double cysteine mutants probed with anti-HA antibody were scanned by densitometry and the % cross-linked protein was determined. (D) The immunoblots of BMOE exposed wildtype Pax3 and selected DCMs probed with anti-HA antibody were scanned by densitometry and the % cross-linked protein was determined with and without prior incubation of a PD specific probe (P3OPT), HD specific probe (P2) and a composite PD/HD probe (PH0).


A
Mock
WT
CL
I59C
S73C
V78C
L215C
K216C
R217C
K218C
Q219C
S222C
T225C
- +
B
I59C
S73C
V78C


C
% Crosslinked Protein
80
70
60
50
40
30
20
10
0
I59C
S73C
V78C
D
% Crosslinked Protein
wildtype
I59C/ R217C
I59C/ S222C
S73C/ Q219C
V78C/ K218C
BMOE only
BMOE, PH0
BMOE, P3opt
BMOE, P2

the control untreated sample (Figure 5A). Such cross-links probably arise from the reaction of BMOE with 2 or more of the 7 endogenous cysteines present in the WT protein. In agreement with this proposal, no such cross-links are detected when the CL mutant is similarly treated with BMOE. These results additionally show that under our experimental conditions, the BMOE reagent can induce the formation of cysteine-specific cross-links. Treatment of all 10 SCMs with BMOE does not result in the formation of slow migrating cross-links. This absence of cross-links in BMOE-treated SCMs together with the demonstrated accessibility of at least single Cys insertions at positions 59, 73 and 78 of the PD (Figure 4) strongly suggest that BMOE does not induce the formation of intermolecular cross-links in these SCMs. Exposure of DCMs to BMOE resulted in the appearance of cross-links for some of the combinations tested (Figure 5B and 5C). This, together with the absence of such cross-links in all 10 similarly treated SCMs (Figure 5B and C) strongly suggest that these BMOE-induced cross-links in DCMs are specific for the individual cysteine pairs and occur in an intra-molecular fashion.

Studies in DCMs with a PD cysteine at position 59, indicated that this residue could react with all N-terminal HD positions with the exception of K216C and Q219C, with the most robust cross-link formed between I59C and S222C (Figure 5B and 5C). To confirm that the detected cross-links were indeed intramolecular (as opposed to intermolecular) SCMs I59C and S222C were mixed together, treated with BMOE and analyzed by SDS-PAGE and immunoblotting. No cross-links were observed under such conditions, confirming that the noted reactivity between I59C and S222C was indeed intramolecular (data not shown). A cysteine at position 73 of the PD (S73C) could react with 5 of the 7 N-terminal HD cysteine insertion, with the strongest reactivity detected for the Q219C. Finally, V78C cross-linked primarily with L215C and particularly K218C, but showed little if any reactivity with any of the other positions.

In summary, of the 7 cysteines inserted in the N-terminal position of the HD, only L215C and K218C form cross-linked species with all three targeted positions in the PD. Also, only K216C fails to significantly cross-link with any of the PD cysteines tested. R217C appears to react only with PD position 59 (I59C) and 78 (V78C), while Q219C reacts only with position 73 (S73C). Finally, HD cysteines S222C and T225C both only cross-link PD cysteines I59C and S73C (Figure 5B and 5C).

Effect of DNA Binding on BMOE-Induced Cross-Linking in Double Cysteine Mutants

The effect of DNA binding by the PD and HD on proximity relationships in the Pax3 protein was further investigated by allowing DCMs to bind PD (P3OPT), HD (P2) or combined PD-HD targets (PH0) prior to treatment with BMOE and resolution of cross-linked products by SDS-PAGE and immunoblotting (Figure 5D and Supplementary figure S2). For these studies, the amount of cross-linked species following BMOE treatment of WT and DCMs (in the absence of DNA) was first quantified by densitometry analysis of the immunoblots and is expressed as a percentage of the total detectable protein on the blot. The effect of DNA binding on the formation of cross-links in WT and DCMs was similarly quantified for each probe. We focused these studies on the 4 DCMs that show the strongest BMOE-induced cross-linking between cysteines inserted in the PD and HD, namely I59C/R217C, I59C/S222C, S73C/Q219C, V78C/K218C (Figure 5D). Incubation of WT and DCM Pax3 variants with the combined PH0 sequence considerably reduced the degree of cross-linking induced by BMOE. Because the effect was seen with all proteins, we suspect that it may reflect a change in overall accessibility of cysteine residues to the BMOE agent in the PH0 bound form of the protein. On the other hand, binding to PD or HD-specific probes had both similar and distinct effects on BMOE-induced cross-linking in individual mutants. For example, in the I59C/S222C mutant, there was no effect of binding to P3OPT or P2

probes on BMOE-induced cross-links, while in the S73C/Q219C and V78C/K218C DCMs, binding to both probes caused a reduction in cross-link formation. On the other hand, the I59C/R217C mutant had a distinct behavior, with significant inhibition of cross-linking when the PD engages in DNA binding and no effect when the protein binds DNA through the HD. These results demonstrate that DNA binding can induce local conformational changes in Pax3 that can be detected by quantitative changes in BMOE-induced cross-linking between individual cysteine pairs of unique DCM variants. In this case, such alterations most likely reflect DNA-induced changes in the distance or orientation of the 2 cysteines, as opposed to overall change of accessibility to the BMOE reagent. Differences in BMOE cross-linking upon DNA-binding would suggest distinct conformations present in the bound and free protein, which differ further depending on the nature of the oligonucleotide bound.

Discussion

We studied proximity relationships in the Pax3 protein by a cross-linking approach based on the use of a bi-functional sufhydryl cross-linker (BMOE) in Pax3 mutants bearing pairs of engineered cysteines. For this, discrete portions of the PD and HD previously shown to play a key role in cooperative interactions of the 2 domains for DNA binding were selected for cysteine insertion. In this approach, and to ascertain that cysteine mutagenesis did not affect overall structure or function of the Pax3 protein, we purposely limited our study to positions at which replacement of the endogenous residue by cysteine was without consequences for DNA binding by the protein. Accessibility of individual Cys substitution to alkylating agents was not determined directly but was inferred from the ability of NEM and BMOE to interfere with DNA binding by the targeted domain. In such an experimental scheme, BMOE-insensitive Cys mutations may reflect a) lack of effect of alkylation on DNA binding, or b) inaccessibility of mutated position to BMOE. The ability of BMOE to induce cross-links in a specific pair of double cysteine mutants was also used as evidence that the 2 positions were accessible to reagent.

The three positions in the PD targeted for Cys mutagenesis (I59, S73, V78) all showed BMOE sensitivity of DNA binding by the PD and the HD, confirming the accessibility of these three positions. The solvent accessibility of these 3 residues is in agreement with previous NEM sensitivity studies (Apuzzo et al., 2006) and the noted capacity of cysteines at all 3 positions to form cross-links with a sub-set of Cysteines introduced in the HD (Figure 5). On the other hand, none of the Cys mutations introduced in the N-terminal part of the HD

showed BMOE sensitivity for DNA binding, except for small effects on HD DNA binding seen in mutants R217C, Q219C, S222C, and L215C. Therefore, it is likely BMOE sensitivity of DNA binding seen in all DCMs is caused in large part by the BMOE reactivity of Cys at the PD positions in these double mutants. On the other hand, the ability of all N-terminal HD Cys substitutions to form cross-links with one or several of the PD Cys insertions indicate that the HD cysteines were indeed accessible to the BMOE cross-linker. One exception is K216C, which failed to generate, upon BMOE exposure, robust cross-links with any of the 3 PD cysteines. In summary, DNA binding and cross-linking studies suggest that all PD and HD cysteines introduced in Pax3, with the possible exception of K216C, were accessible to BMOE and therefore may be solvent-exposed.

The BMOE-induced cross-links detected in WT Pax3 must arise from reaction of any two endogenous cysteines with the bi-functional cross-linker. The endogenous cysteines most likely to be BMOE-reactive are cysteines C70 in the PAI helix 2, C82 in the PAI helix 3 and C143 in the RED subdomain in the loop joining helix 5 and 6 (Tell et al., 1998). Indeed, mass spectroscopy experiments as well as mutagenesis studies have shown that under oxidative conditions a disulfide bridge is formed between Cys70 and 82 as well as Cys 82 and Cys 143 thereby abrogating DNA binding via the PD. Lack of cross-links in Pax3-CL is indicative of the specificity of the BMOE reagent for cysteines and the absence of cross-links in SCMs shows that any cross-links formed in DCMs are more likely the result of intramolecular cysteine cross-linking and not intermolecular cross-linking. In the DCMs analyzed, the most robust cross-links appeared to be formed between I59C and R217C, I59C and S222C, S73C and Q219C and between V78C and K218C. An obvious interpretation of these results is that these pairs of cysteines, and corresponding positions in the wild type protein, are in closest proximity amongst all pairs tested.

The amount of BMOE-induced protein cross-links in WT and mutants was estimated by densitometry and the effect of DNA binding by the PD and HD on the production of such cross-links was measured. Binding to the composite site sequence PH0 caused a dramatic decrease in the overall amount of BMOE-induced cross-linking in all proteins tested. One likely explanation of this observation is that DNA-binding to PH0 increases the compactness and inflexibility of the protein and reduces accessibility of cysteines to BMOE, hence decreasing the number of cross-links formed. This interpretation is in agreement with circular dichroism spectroscopy studies with the PD of Pax5 and 8 showing that the PD adopts a more ordered structure with more helical content upon DNA exposure (Tell et al., 1998). Also, X-ray crystallography and NMR studies of HD proteins have revealed an increased order of the HTH motif, including the N-terminal arm and the recognition helix, of the DNA-bound form of Pbx (Sprules et al., 2003; Sprules et al., 2000; Jabet et al., 1999; Piper et al., 1999), Antennapedia (Qian et al., 1994), engrailed (Clarke et al., 1994) and Oct-1 (Cox et al., 1995) proteins. Also, previous studies have shown that engineered protease sites in the PAI subdomain and N-terminal arm of the HD in Pax3 become less accessible to cleavage upon DNA exposure indicating a more compact conformation of the DNA-bound form (Apuzzo et al., 2004).

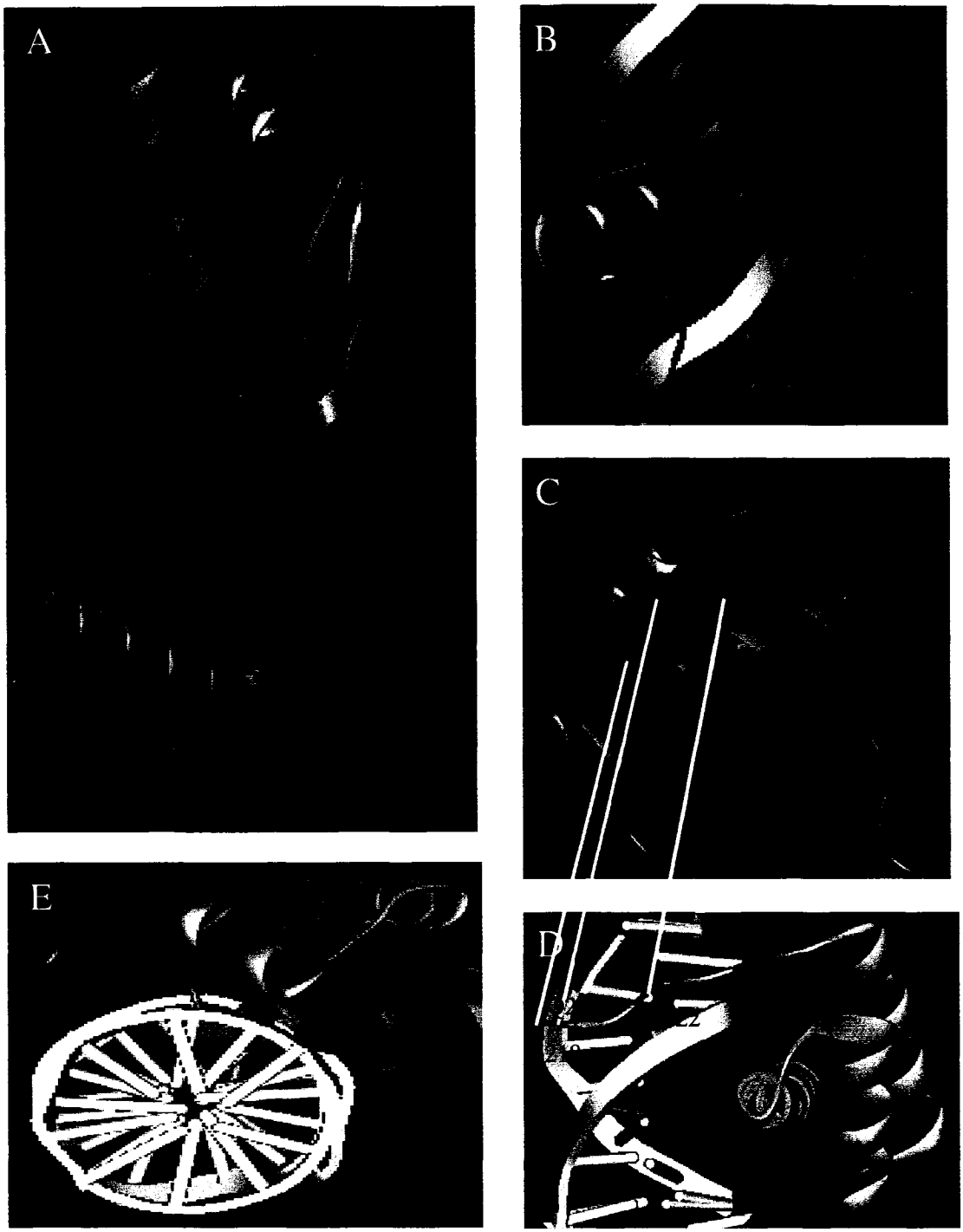
It has been established that for BMOE to induce cross-links, the two sulfhydryl side chains must be located 6-10.5Å apart (average 8Å) (Green et al., 2001). Therefore, one can conclude that the maximum intermolecular distances separating the above-mentioned pairs of residues must be within the range of 6-10.5Å. The pairs of cysteines showing the most robust BMOE-induced cross-links (I59C/R217C, I59C/S222C, S73C/Q219C and V78C/K218C) have been positioned on the three-dimensional structure of the DNA bound models of the PD from Pax6 (Xu et al., 1999) and of the HD from *Drosophila* Prd (Wilson et al., 1995) proteins. From this modeling, an everted positioning of the PD and HD for DNA binding to a

combined site appears required to account for cross-linking results obtained here in DCMs variants. Such a model is in agreement with that previously proposed by Jun and Desplan for combined PD and HD binding of the *Drosophila* Prd protein to chimeric target sequences (Jun et al., 1996).The model suggests that the N-terminus of the PAI domain is closely apposed to the N-terminus of the HD when bound to the combined site. Their model suggests that PD and HD bind opposite sides of DNA helix with the N-terminal arm of the HD in very close proximity to helix 2 of the PAI subdomain. The PD/HD physical interaction suggested in this model may be responsible for the interdependence of DNA binding observed in Pax3 and may also be used to explain the interaction of the PD of Pax3 with the N-terminal arm of Msx1 (Bendall et al., 1999). The model shown in Figure 6 indicates only the most extensive cross-linking patterns observed. The most N-terminal I59 residue is closest to the most C-terminal residue of the N-terminal arm of the HD, S222C. While the most C-terminal PD residue V78 forms the most cross-links with and is closest to the most N-terminal HD residue K218. This is in accordance with the head-to-head orientation of the PD and HD when they are physically interacting. The cross-linking studies and the model proposed in figure 6 support our previous study that implicates the role of residues I59, S73 and V78 in forming a hydrophobic pocket in the PD whereby the HD can dock and participate in the physical interaction responsible for DNA binding functional interdependence (Apuzzo et al., 2006).

These cross-linking studies provide, for the first time, an indication that residues in the PAI subdomain are in close proximity to residues in the N-terminal portion of the HD. The PD and HD residues in close proximity may be the residues responsible for the PD/HD physical interaction that accounts for the functional interdependence of the two DNA binding domains.

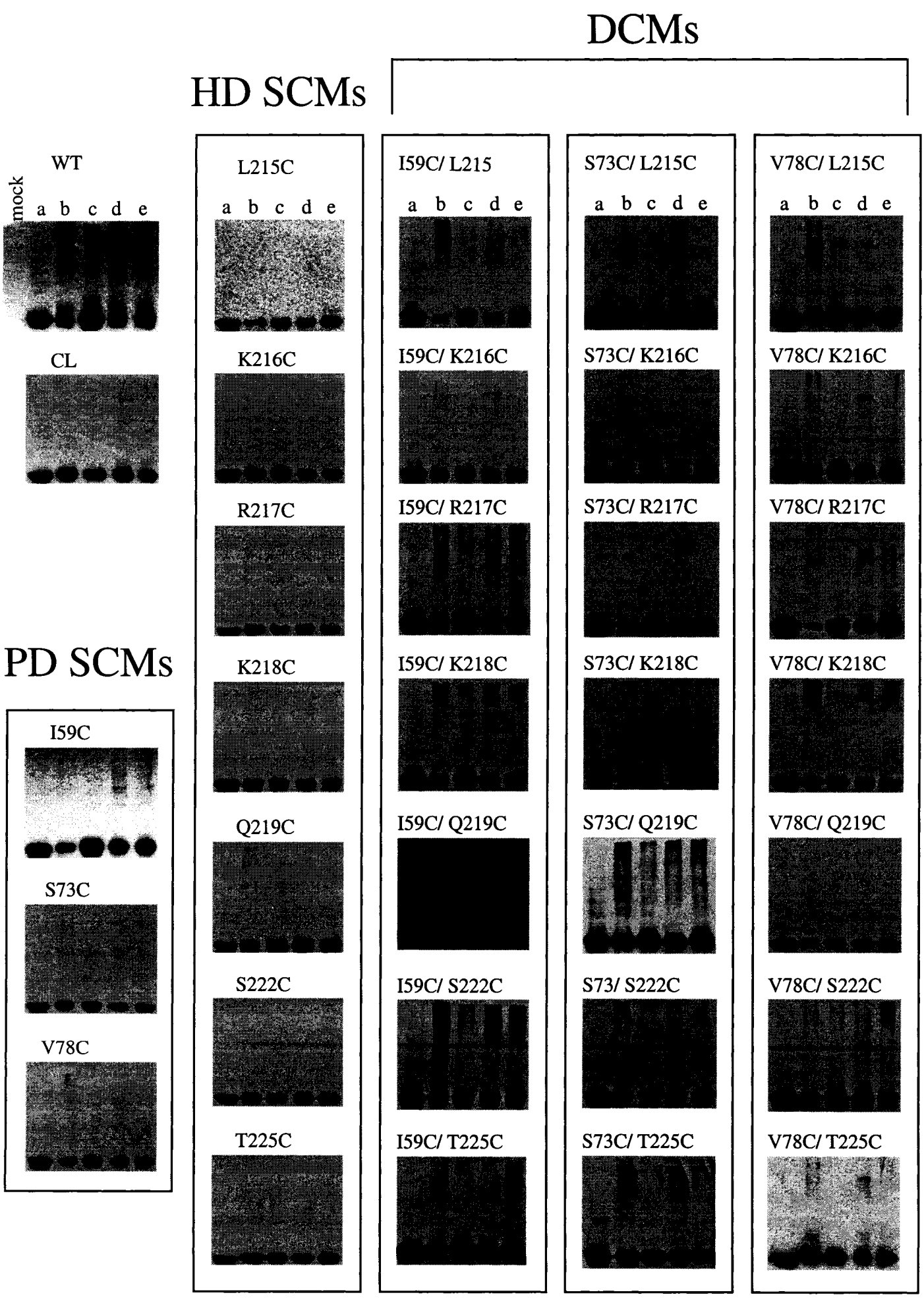
Figure 6

Structure of the DNA-bound Paired domain and Homeodomain. (A) The PAI and RED subdomains of the PD are indicated on a 3D DNA-bound structure. The 3 α-helices and the β-hairpin of the PAI subdomain are indicated. (B) and (E) show different close-up views of the PAI subdomain. (C) View of the DNA-facing PAI subdomain with the DNA structure removed for clarity. (D) The DNA-bound 3D HD structure. The PD and HD are shown as a green ribbon drawn through the α carbon backbone. The DNA strands are shown as blue (A) or gray (B-E) ribbons through the sugar-phosphate backbone, and bases are shown as protrusions from the ribbons. The positions targeted for cysteine substitution and at which BMOE modification causes cross-linking are indicated in red. Turquoise lines that link cysteine substitutions in figure C and D indicate cysteines that undergo extensive cross-links when paired in a DCM and exposed to BMOE. (F) A model of the "head-to-head" orientation of the PD and HD that facilitates PD/HD physical interaction responsible for the interdependence of DNA binding. The black line indicates the linker that joins the PD and HD.


A
B
C
E
D
F
RED
K218
V78
Q219
S73
HD
PAI
S222
I59
C
N
N
C

Supplementary Figure 2

Detection of Cross-linking Species of Pax3 mutants that were incubated with DNA prior to BMOE exposure. Fifteen micrograms of whole cell extracts from COS-7 monkey cells transfected with either wild type (*WT*), Cys-less (CL), a SCM or a DCM Pax3 mutant were incubated with 3.55mM BMOE and were analyzed by SDS-PAGE on 12% acrylamide gels and by immunoblotting with anti-HA antibody. Cross-linked species are visible as higher molecular mass bands. Cross-linked species were detected in mutants exposed to BMOE with or without prior incubation of a PD specific probe (P3OPT), HD specific probe (P2) and a composite PD/HD probe (PH0). The proteins were incubated with no DNA and no BMOE (a), no DNA and BMOE (b) or BMOE with DNA probes specific for the PD (P3OPT) (d), the HD (P2) (e) or both the PD and HD (PH0) (c).


DCMs
HD SCMs
WT
mock
a b c d e
CL
L215C
a b c d e
K216C
R217C
K218C
Q219C
S222C
T225C
I59C/ L215
a b c d e
I59C/ K216C
I59C/ R217C
I59C/ K218C
I59C/ Q219C
I59C/ S222C
I59C/ T225C
S73C/ L215C
a b c d e
S73C/ K216C
S73C/ R217C
S73C/ K218C
S73C/ Q219C
S73/ S222C
S73C/ T225C
V78C/ L215C
a b c d e
V78C/ K216C
V78C/ R217C
V78C/ K218C
V78C/ Q219C
V78C/ S222C
V78C/ T225C
PD SCMs
I59C
S73C
V78C

Chapter 6

Conclusion and Summary

Like many other Pax proteins, Pax3 contains conserved DNA binding domains yet can play different regulatory roles and functions in several diverse processes in embyrogenesis. How can Pax3 be involved in such different diverse processes as neurogenesis, myogenesis and melanogenesis and yet carry only a few conserved DNA binding motifs? Why doesn't redundancy in recognition result in redundancy in function? One possible solution to this question is that Pax proteins use mechanisms involving intra and intermolecular interactions between DNA binding domains to attain functional diversity. The studies described in the previous chapters shed light on what we consider an important set of interactions in Pax3 that mediate DNA recognition: intramolecular PD/HD cooperative interactions. This chapter is to position the work of this thesis in the context of what is already known about Pax protein biochemistry and to assess its impact on Pax research.

Intramolecular Mechanisms for creating Functional Diversity in Pax Proteins

Functional diversity is typical of multiple modular domain transcription factors such as Zinc finger domain proteins. Zinc finger domains interact with DNA as well as one another (intramolecular interaction) to modify specific DNA recognition (Rebar and Pabo., 1994). Like multiple modular domains in Zinc finger transcription factors, proteins with multiple helix-turn-helix (HTH) motifs may interact intramolecularly and in different combinations to permit binding to DNA with different modes thereby allowing for the recognition of a series of different DNA sequences and generating different specificities.

Many Pax proteins contain two DNA binding domains (PD and HD) and consequently hold three distinct DNA binding HTH motifs. Pax proteins are able to recognize different

types of binding sites with the use of different combinations of their HTH motifs. To reinforce this point, what follows are a series of DNA binding modes that Pax proteins and related proteins have been shown to adopt when they bind DNA. Pax3 recognizes the class I DNA sequence P3OPT with the use of both the PAI and RED subdomains (Vogan et al., 1997). Similarly, Pax6 is able to bind the P6CON class I probe with both subdomains in the PD as well (Epstein et al., 1994). Pax3 and other Pax related proteins are also able to bind class II DNA sequences, which requires only the use of the PAI subdomain (Epstein et al., 1994; Czerny et al., 1993). Alternatively spliced isoforms of Pax6 and Pax8 introduce insertions into the PAI subdomain rendering the subdomain nonfunctional (Epstein et al., 1994; Kozmik et al., 1997). These isoforms are able to bind a different set of DNA sequences solely through the use of the RED subdomain (Epstein et al., 1994; Kozmik et al., 1997). It has been shown that the Pax *Drosophila* homolog Prd interacts with the PH0 DNA probe with the cooperative use of the PAI subdomain and the HD (Jun and Desplan., 1996). Along with other Pax and Pax related proteins, Pax3 can bind to HD specific DNA sites as a monomer (on the P1/2 probe) (Apuzzo et al., 2002) or dimer (on the P2 probe) (Apuzzo et al., 2002). The most recently discovered DNA binding mode that may also be used by Pax proteins requires the interaction of the RED subdomain and the HD on DNA. This mode is the major mode of DNA binding used by the product of the *Drosophila* eye gone gene (eyg) known as the Lune Pax-like protein which lacks a PAI subdomain (Jun and Desplan., 1998). Also, recent studies with Pax6 have shown that basic residues in the N-terminal arm and recognition helix of the paired-type HD form an interaction surface. This surface can associate with acidic side chains of residues in helices 1 and 2 of the RED subdomain in the presence and absence of DNA (Bruun et al., 2005).

Therefore, each HTH motif of Pax3, and other related proteins, can bind DNA alone or may associate with another motif as a strategy to generate different specificities. The mode

that we chose to study extensively involves cooperative interactions between the PAI and HD HTH motifs. We believe these interactions exists in Pax3 when the protein is in the DNA unbound state as well as PD DNA bound (P3OPT and P6CON probes), HD DNA (P1/2 and P2 probes) bound and PD and HD DNA bound (PH0 probe) states.

Intermolecular Mechanisms for creating Functional Diversity in Pax Proteins

Another innovative solution implemented by Pax proteins to remedy the requirement of generating functional diversity using only conserved DNA binding domains is through interactions with other proteins on DNA. Pax3 and related proteins can participate in intermolecular interactions with other DNA binding proteins and this may lead to modulation of DNA specificity of Pax proteins. Many proteins are known or suspected to bind Pax proteins and most were mentioned in the previous chapters. Of particular interest are the cooperative complexes that Pax5 can form with members of the Ets family of transcription factors. One extensively studied example of these complexes is formed between Pax5 and c-Ets-1 on the promoter of the mb-1 gene (Fitzsimmons et al., 1996; Wheat et al., 1999; Hagman et al., 2000). A Pax5 recognition site lies adjacent to a sub-optimal Ets binding site in the promoter of mb-1. Pax5 is able to bind the promoter on its own but Ets proteins associate with the sub-optimal recognition site only when they may form a complex with Pax5 to form a ternary complex. Both the Pax5 and Ets DNA binding sites are required on DNA. The formation of this ternary complex is required for optimal expression of the mb-1 gene (Fitzsimmons et al., 1996). Pax5/Ets ternary complex formation on the composite site of mb-1 requires portions of the β-hairpin and helix 2 of the PAI subdomain of Pax5 (Wheat et al., 1999; Garvie et al., 2001). This study and others suggests that these structures in the PD of Pax proteins (β-hairpin and helix 2) can participate in protein-protein intra and intermolecular interaction that permit modulation of DNA specificities.

Studies of Pax3 and Msx1 proteins, on the activity of the MyoD promoter, is yet another example of how Pax proteins can use intermolecular interactions as a means to generate functional diversity. Pax3 can bind the MyoD promoter with the PD to induce MyoD expression. Yet, when complexed with the Msx1 protein, Pax3 no longer retains the same affinity for the MyoD promoter and reduced MyoD expression results (Bendall et al., 1999). In vitro assays indicate that the N-terminal arm of the HD of Msx1 is the element involved in intermolecular interactions with the PD of Pax3 and modulation of Pax3 DNA binding activity (Bendall et al., 1999).

These as well as many other studies of Pax proteins indicate that protein: protein interactions that modulate Pax DNA affinity and specificity may involve the protein segments in or close to the β-hairpin and helix 2 of the PAI subdomain as well as the N-terminal arm of a HD. Studies show that these segments are required for cooperative binding of the PD and HD. The importance of the PD/HD cooperative mode of DNA binding is accentuated by the observation that the rescue of the Prd *Drosophila* mutant only requires a functional PAI domain and HD in the same molecule (Bertuccioli et al., 1999). These observations suggest that, at least for the Prd protein, PD and HD containing proteins may principally act by binding through a combination of its PAI domain and HD. We propose that the cooperative interactions between the PAI subdomain of the PD and the HD of Pax3 are strong enough to influence PD DNA binding activity, HD DNA binding activity as well as binding to composite PD/HD DNA sequences. The methods we chose to explore these PD/HD cooperative interactions are cysteine scanning mutagenesis and protease sensitivity assays, which are well known for their use in protein structure: function studies.

Waardenburg syndrome Missense Mutations and the Cooperative Intramolecular interactions between Pax3 PD and HD

To date over fifty Waardenburg syndrome mutations have been reported. A subset of these mutations are truncating mutations that are located almost anywhere in the PAX3 gene and result in the deletion of either the proline-serine-threonine rich transactivation domain, the HD and the rest of the residues C-terminal to the HD, or the whole gene (Baldwin et al., 1995). Missense mutations cluster to the PAI subdomain and the HD (Baldwin et al., 1995) thereby suggesting a critical role of these protein segments for PAX3 function. These missense mutations cause very similar phenotypes to the truncating mutations previously described and are therefore considered loss of function PAX3 mutant alleles. Therefore, studying the biochemistry of these mutants can give information about the possible roles these protein segments play during DNA binding and transcriptional activity by Pax3.

Most Waardenburg syndrome PAX3 mutant alleles that encode for proteins with missense mutations in the PD change the amino acid identity at conserved positions (Fortin et al., 1997; Xu et al., 1995). The determination of the structural basis of these missense mutations is possible with the use of the PD/DNA co-crystal structures of Pax6 and Prd (Xu et al., 1999; Xu et al., 1995). More than a third of the residues affected in Waardenburg syndrome missense mutations map to invariant areas of the PD that are involved in DNA contacts. Only a few of these missense mutations affect base-specific DNA contacts, most actually map to residues that non-specifically associate with the sugar-phosphate backbone. Mutations at conserved positions required for specific and non-specific DNA contacts compromise DNA binding by the PAX3 protein (Fortin et al., 1997). According to the published Pax6 and Prd crystal structures, some Waardenburg syndrome deleterious missense mutations map to positions that don't associate with DNA. These missense mutations include I59F, V60M and S73L and were previously only thought to disrupt the stability of the helices in which they lie (Fortin et al., 1997; Xu et al., 1999; Xu et al., 1995). Invariant residues I59, and V60 are in the C-terminal half of helix 1 and the well conserved residue S73 lies in the

middle of the second helix of the PAI subdomain. The study described in the chapters 4 and 5 suggests that these three residues play critical roles in the formation of a hydrophobic binding pocket in the PD that is formed from segments of the PAI subdomain helix 1 and 2, the loop between helix 2 and 3 as well as portions of the β-turn in the β-hairpin (Apuzzo et al., 2006). When these three positions are substituted with cysteine and alkylated by a thiol specific reagent, such as N-ethylmaleimide, both PD and HD DNA binding activities are abrogated (Apuzzo et al., 2006). In light of this new discovery, we propose that the Waardenburg syndrome missense mutations I59F, V60M and S73L, which replace the naturally occurring side chain with longer and larger side chains, results in the inability for the PD to properly form the HD hydrophobic binding pocket. Consequently, the PD/HD physical interaction and functional cooperativity is severely affected. This PD/HD physical interaction is required for proper DNA binding of both the PD and HD and therefore its disruption results in impaired PD and HD DNA binding activities (Apuzzo et al., 2006). Another possible interpretation is that the modification of these three positions, via cysteine substitution and site-specific modification or substitution by bulkier amino acids, does not disrupt the formation of the HD binding hydrophobic pocket but prevents the docking of the HD onto the PD via steric hindrance. In either case, PD/HD physical interaction and cooperativity required by both the PD and HD to bind DNA is disrupted.

In conclusion, the structural and functional characterization of naturally occurring Pax3 mutations have played a role towards our improved understanding of the mechanisms that account for the interdependence of DNA binding of the PD and HD of Pax3. Hence, ongoing and future investigations of other Pax3 biochemical activities or functions could greatly benefit from the molecular characterization of naturally occurring mutations.

The Possible Role of the Type I β-turn, α2 and the α2 – α3 loop of the PAI subdomain in the Cooperative Intramolecular interactions between Pax3 PD and HD

Several findings are consistent with a model in which the PD and HD of Pax3 physically interact. For example, the functional interaction between the PD and HD is uncoupled by deletion of the second helix of the PAI subdomain (Fortin et al., 1998). The Prd PD DNA bound structure reveals that this helix is solvent exposed and a therefore a good candidate for being involved in protein: protein interaction (Xu et al., 1995). Also, modeling of the HD and PD crystal structures on the cooperatively bound composite PD/HD PH0 site shows that helix 2 is in a favorable location to physically interact with the N-terminus of the HD (Jun and Desplan, 1996). However, until now, a direct physical interaction between the HD and PD remained to be supported by experimental data.

The Prd DNA bound PD structure reveals that PAI subdomain residues that form the Type I β-turn in the β-hairpin (positions 39 to 43 in Pax3) interacts with residues in α2 (position 73 in Pax3) and the α2 – α3 loop (positions 77 and 78 in Pax3) (Xu et al., 1995). The schematic representation of the PD in figure 1 in chapter 4 reveals that the identity of the Type I β-turn (τ1) residues in Pax3 are N39, Q40, L41, G42 and G43 and that the α2 position 73 residue is a serine and the α2 – α3 loop residues at position 77 and 78 are arginine and valine, respectively. Figure 4 in chapter one clearly shows how the Type I β-turn and portions of α2 and the α2 – α3 loop come in close proximity when the PD is DNA bound. We propose that the Type I β-turn, α2 and α2 – α3 loop together form a protein-protein interaction surface that is required for intra and intermolecular interactions. More specifically, we postulate that these secondary structures come together in the tertiary structure of the PD of Pax3 to form the hydrophobic HD binding pocket described in chapter 4. Also, we believe that disruption of the ability of these segments in the PD to associate will prevent the formation of this HD

binding pocket and consequently preventing PD:HD physical interaction and cooperativity. This would result in the weakening of both PD and HD mediated DNA binding activities, in accordance with the interdependent model of DNA binding of the PD and HD of Pax3.

The mutation G42R in the Type I β-turn in the PD of Pax3, found in the splotch-delayed mouse mutant (Sp^d), not only shows reduced DNA binding to PD oligos, but also causes reduced DNA binding to HD-specific oligos (Underhill et al., 1995; Fortin et al., 1998). We believe that the replacement of the glycine side chain with the larger and bulkier side chain of arginine must sterically prevent the Type I β-turn structure from interacting with residues in α2 and the α2 – α3 loop. Consequently, the hydrophobic HD binding pocket in the PD cannot form properly thereby preventing the PD:HD physical interaction required for PD/HD interdependence of DNA binding. The inability of the Pax3 HD and PD to interact in the G42R mutant compromises both PD and HD mediated DNA binding activities. Deletion of helix 2 of the PAI subdomain in the context of the Sp^d mutation (Pax3 G42RΔα2) has been shown to restore HD DNA binding (Underhill et al., 1995). We propose that for Pax3, and other Pax proteins with a PD and a HD covalently linked within the same polypeptide chain, the HD requires α2 to pack tightly against the Type I β-turn for the PD to modulate HD DNA binding activity. Removal of α2 in the Pax3 G42R mutant uncouples PD dependent HD DNA binding activity.

Another indication that the Type I β-turn structure may play an important role in PD:HD physical interaction and functional interdependence is the fact that each and every residue in this structure is completely conserved among all members of the Pax family as well as among Pax homologs from various other organisms. Another important fact to note is that the multiple sequence alignment of the PD from various genes also indicate that the α2 S73 and the α2 – α3 loop V78 residues are also invariant (see figure 1 of chapter 4).

The involvement of the PAI subdomain Type I β-turn and α2 structures in forming a protein: protein interaction surface is further supported by structural studies of the ternary complex formed between the PD of Pax5 and the Ets domains of Ets-1 on DNA. Pax5 Type I β-turn residues Q22 (Pax3 equivalent residue is Q40) and L23 (Pax3 equivalent residue L41), as well as residue R56 in helix 2 (Pax3 equivalent residue is R74), make contacts with the Ets domain when both proteins are DNA bound (Garvie et al., 2001). The Pax5-Ets structure has led us to believe that NEM modification of NEM sensitive single cysteine mutants Q40C and L41C may mimic the splotch G42R deleterious effect on PD:HD physical interaction. More specifically, we postulate that the side chain replacement of any of the residues in the Type I β-turn with larger or bulkier groups sterically prevents the approach of helix 2 and consequently the formation of the protein: protein interaction site (the hydrophobic HD binding pocket) in the PAI subdomain.

The helix 2 R74C Pax3 single cysteine mutant did not show PD or HD NEM sensitivity, but residues flanking this position (pst: 72, 73 and 76, 78, 79) were found to be NEM sensitive (Apuzzo., 2006). Alkylations at these flanking positions with bulky groups must sterically prevent the approach of the Type I β-turn and, like alkylation of residues in the Type I β-turn, lead to prevention of the protein: protein interaction surface on the PAI subdomain.

Therefore, the analysis of PD:HD interactions in SCM Pax3 mutants, structural studies of the interface of Pax5-Ets proteins interaction in the DNA-bound state, the Prd PD DNA bound structure, and the characterization of the Pax3 G42R splotch mutant protein strongly suggest that the Type I β-turn, α2 and the α2 - α3 loop of the PD interact to form a protein: protein interaction surface. We believe that this surface is used by the PD to associate with the HD (intramolecular interaction) as well as with other transcriptions factors like members of the Ets family of proteins (intermolecular interaction) .

As mentioned previously, the Prd PD DNA bound structure indicates that the Pax3 residues that associate with the Type I β-turn structure include S73 in α2 and V78 in the α2 - α3 loop (Xu et al., 1995). The role of these residues in Type I β-turn interaction and formation of the hydrophobic HD binding pocket is further accentuated by the fact that the alkylation by NEM of single cysteine mutants S73C and V78C abrogate both PD and HD mediated DNA binding activities (Apuzzo et al., 2006). We propose that when unmodified with thiol specific reagents single cysteine mutants S73C and V78C are still able to form similar van der Waals to residues in the Type I β-turn. We believe this since the side chains of valine and cysteine are known to both be able to participate in hydrophobic coupling interactions with the side chains of other hydrophobic amino acids. Also, serine and cysteine share many physico-chemical properties and therefore it is likely that they participate in the very similar van der Waal forces with the side chains of other amino acids. We believe that alkylation of single cysteine mutants S73C and V78C with larger and bulkier groups not only destroy interactions with the Type I β-turn but also sterically prevent the approach of the β-turn to both α2 and the α2 - α3 loop. This would prevent HD binding pocket formation and PD:HD physical and functional interaction. One way to further confirm the association of S73 and V78 with residues in Type I β-turn is to create double cysteine mutants with a cysteine in α2 or in the α2 - α3 loop and another in the Type I β-turn and attempt to cross-link them with a bifunctional thiol specific reagent.

Cooperative Intramolecular interactions between the PD and HD: a comparison of the Pax6 and Pax3 models.

Functional interaction between the PD and HD has also been demonstrated with the Pax6 protein (Sheng et al., 1997; Singh et al., 2000), but the mechanism employed seems to be different from that of Pax3. For Pax3 the RED subdomain seems to play no significant role in the PD/HD physical and functional interaction, this is not the case with Pax6. A naturally

occurring missense mutation in the RED subdomain, I87F, abrogates both PD and HD DNA binding activities, while no missense mutation in the RED subdomain of Pax3 has ever been detected that abrogates HD DNA binding activity. This indicates that a functional (and possibly physical) association may exist between the HD and the RED subdomain of the PD (Singh et al., 2000).

Also, the naturally occurring missense mutation in the PAI subdomain R26G and the alternative splicing isoform Pax6(5a), with a 14 amino acid insertion in the PAI subdomain, both abrogate the DNA binding activity of the PAI subdomain (Singh et al., 2000). PAI subdomain inactivating Waardenburg missense mutations in Pax3 abrogate HD DNA binding activity yet the Pax6 PAI subdomain inactivation mutations indicated above actually increase the DNA binding activity of the HD at both the monomer and dimer level on both P2 and P3 probes (Singh et al., 2000). Two non-exclusive interpretations may account for these observations. PAI subdomain inactivation may uncover a positive interaction between the RED subdomain and the HD or the PAI subdomain inactivation relieves any PAI-mediated negative regulation of the HD DNA binding activity. One must not rule out the possibility that the other potential explanation for these differences between Pax3 and Pax6 may involve sequences outside of the paired domain and homeodomain.

Recent findings confirm that the HD and the RED subdomain of Pax6 physically interact and that this interaction is independent of DNA (Bruun et al., 2005). No such interaction has been observed in Pax3 to date. Mutagenesis/EMSA studies, GST pull-down assays, Yeast two-hybrid assays and FRET were used to study and generate a model of the RED/HD physical interaction in Pax6 (Bruun et al., 2005). This physical interaction is mediated by basic residues in the recognition helix (α3) in the HD and acidic residues in the RED subdomain. More specifically, the basic residues that lie entirely on one side of the HD recognition helix (R57 and R58) and basic residues in the HD N-terminal arm (R3 and R5)

form salt bridges with acidic residues in the RED subdomain helix 4 (E112) and helix 5 (E120 and E128) (Bruun et al., 2005). Most helix 3 HD residues that interact with the RED subdomain are conserved and are not involved in making DNA contacts, therefore they presumably solely serve the purpose of mediating RED subdomain interaction which presumably serves to increase the DNA specificity of the Pax6 protein. Contrastingly, the HD N-terminal arm residues R3 and R5 that make contacts with the RED subdomain in the Pax6 protein both make base specific contacts in the minor groove of DNA (Wilson et al., 1995).

There are two striking similarities and two differences between the models of the PAI/HD physical interaction of Pax3 with that of the RED/HD interaction of Pax6. In both models the solvent exposed portions of the two most N-terminal helices of the three helical folds of the PAI and RED subdomains act as surfaces in the PD that interact with the HD. Also, in both models the N-terminal arm of the HD plays a role in the physical interaction. The Pax6 HD N-terminal arm residue R3 is equivalent to residue R221 in Pax3 and the R5 residue is equivalent to residue R223 in Pax3. The involvement of R221 and R223 in the PAI/HD interaction model of Pax3 was not assessed by the cysteine scanning mutagenesis/site-modification approach used in some of our studies. This was unfortunately the case since the single cysteine mutants R221C and R223C do not retain PD and HD DNA binding activities comparable to the wild type protein and therefore any site-specific modification with thiol specific reagents would have been uninformative (Apuzzo et al., unpublished, Chapter 5). This is most probably due to the loss of the base specific contacts these arginines make with the minor groove of DNA when they are substituted to cysteines. An important observation to note though is that when flanking residues Q219, S222 and T225 were substituted to cysteine they were able to form cross-links with cysteines placed in or near helix 2 of the PAI subdomain (Apuzzo et al., unpublished, Chapter 5). The most obvious difference between the two models is that one involves the use of the HD recognition helix

while the other does not. Also, the RED/HD contacts in Pax6 consists primarily of salt bridges while those in the PAI/HD model of Pax3 is assumed to be mostly mediated by hydrophobic coupling (Bruun et at., 2005; Apuzzo et al., 2006).

Also, the differences in the mechanisms that Pax3 and Pax6 use to mediate PD and HD interactions is further suggested by the observation that Pax6 cooperatively dimerizes on P3 sequences (Singh et al., 2000) regardless of the presence or absence of the PD, yet the Pax3 HD dimerizes on P2 only when present with the Pax3 PD in the same polypeptide but in vitro generated Pax3 HD dimerizes on P3 only (Fortin et al., 1998).

In conclusion, although Pax proteins have highly homologous DNA binding domains that can bind similar DNA sites, these proteins can obtain higher levels of specificity and functional diversity via the distinctive use or combination of DNA binding modules. As of now, functional interactions have been shown for each possible type of pairing of these modules, suggesting that their combined use has more than additive consequences on the DNA binding specificity and activity of Pax proteins. Characterization of these interactions is therefore essential for the proper understanding of specific target gene recognition and regulation by individual Pax proteins. Finally, it is essential to realize that the identification of functional associations between each of the PD subdomains and the HD emphasizes the necessity to use full-length proteins in functional studies to obtain results applicable to the native protein. While the effects of intramolecular interactions on Pax protein DNA binding function are starting to be characterized, the biological relevance, regulation and mechanism of action is still largely unexplored. However, prior to these higher levels of functional complexity being assessed, the biochemical mechanisms responsible for these interactions must be the primary focus for ongoing research. Surely, the complete understanding of the mechanisms that underlie these functional interactions shall promote the elucidation of the roles Pax genes play in molecular pathways as well as in embryogenesis.

References

Adams, B., P. Dorfler, A. Aguzzi, Z. Kozmik, P. Urbanek, I. Maurer-Fogy and M. Busslinger (1992). "Pax-5 encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis." Genes Dev 6(9): 1589-607.

Andersen, F. G., R. S. Heller, H. V. Petersen, J. Jensen, O. D. Madsen and P. Serup (1999). "Pax6 and Cdx2/3 form a functional complex on the rat glucagon gene promoter G1-element." FEBS Lett 445(2-3): 306-10.

Apuzzo, S., A. Abdelhakim, A. S. Fortin and P. Gros (2004). "Cross-talk between the paired domain and the homeodomain of Pax3: DNA binding by each domain causes a structural change in the other domain, supporting interdependence for DNA Binding." J Biol Chem 279(32): 33601-12.

Apuzzo, S. and P. Gros (2002). "Site-specific modification of single cysteine Pax3 mutants reveals reciprocal regulation of DNA binding activity of the paired and homeo domain." Biochemistry 41(40): 12076-85.

Apuzzo, S. and P. Gros (2006). "The paired domain of Pax3 contains a putative homeodomain interaction pocket defined by cysteine scanning mutagenesis." Biochemistry 45(23): 7154-61.

Asano, M. and P. Gruss (1992). "Pax-5 is expressed at the midbrain-hindbrain boundary during mouse development." Mech Dev 39(1-2): 29-39.

Auerbach, R. (1954). "Analysis of the developmental effects of a lethal mutation in the house mouse." J. Exp. Zool. 127: 305-29.

Ayme, S. and N. Philip (1995). "Possible homozygous Waardenburg syndrome in a fetus with exencephaly." Am J Med Genet 59(2): 263-5.

Baldwin, C. T., C. F. Hoth, J. A. Amos, E. O. da-Silva and A. Milunsky (1992). "An exonic mutation in the HuP2 paired domain gene causes Waardenburg's syndrome." Nature 355(6361): 637-8.

Baldwin, C. T., C. F. Hoth, R. A. Macina and A. Milunsky (1995). "Mutations in PAX3 that cause Waardenburg syndrome type I: ten new mutations and review of the literature." Am J Med Genet 58(2): 115-22.

Balling, R., U. Deutsch and P. Gruss (1988). "undulated, a mutation affecting the development of the mouse skeleton, has a point mutation in the paired box of Pax 1." Cell 55(3): 531-5.

Barberis, A., G. Superti-Furga, L. Vitelli, I. Kemler and M. Busslinger (1989). "Developmental and tissue-specific regulation of a novel transcription factor of the sea urchin." Genes Dev 3(5): 663-75.

Barr, F. G. (1999). "The role of chimeric paired box transcription factors in the pathogenesis of pediatric rhabdomysarcoma." Cancer Res 59(7 Suppl): 1711s-1715s.

Barr, F. G., N. Galili, J. Holick, J. A. Biegel, G. Rovera and B. S. Emanuel (1993). "Rearrangement of the PAX3 paired box gene in the paediatric solid tumour alveolar rhabdomyosarcoma." Nat Genet 3(2): 113-7.

Beechey, C. V., Searle, A. G. (1986). "Mutations at the Sp Locus." Mouse News Letters 75: 28.

Bendall, A. J., J. Ding, G. Hu, M. M. Shen and C. Abate-Shen (1999). "Msx1 antagonizes the myogenic activity of Pax3 in migrating limb muscle precursors." Development 126(22): 4965-76.

Bennicelli, J. L., R. H. Edwards and F. G. Barr (1996). "Mechanism for transcriptional gain of function resulting from chromosomal translocation in alveolar rhabdomyosarcoma." Proc Natl Acad Sci U S A 93(11): 5455-9.

Bennicelli, J. L., W. J. Fredericks, R. B. Wilson, F. J. Rauscher, 3rd and F. G. Barr (1995). "Wild type PAX3 protein and the PAX3-FKHR fusion protein of alveolar rhabdomyosarcoma contain potent, structurally distinct transcriptional activation domains." Oncogene 11(1): 119-30.

Bernasconi, M., A. Remppis, W. J. Fredericks, F. J. Rauscher, 3rd and B. W. Schafer (1996). "Induction of apoptosis in rhabdomyosarcoma cells through down-regulation of PAX proteins." Proc Natl Acad Sci U S A 93(23): 13164-9.

Bertolotto, C., P. Abbe, T. J. Hemesath, K. Bille, D. E. Fisher, J. P. Ortonne and R. Ballotti (1998). "Microphthalmia gene product as a signal transducer in cAMP-induced differentiation of melanocytes." J Cell Biol 142(3): 827-35.

Bertuccioli, C., L. Fasano, S. Jun, S. Wang, G. Sheng and C. Desplan (1996). "In vivo requirement for the paired domain and homeodomain of the paired segmentation gene product." Development 122(9): 2673-85.

Billeter, M., Y. Q. Qian, G. Otting, M. Muller, W. Gehring and K. Wuthrich (1993). "Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex." J Mol Biol 234(4): 1084-93.

Bladt, F., D. Riethmacher, S. Isenmann, A. Aguzzi and C. Birchmeier (1995). "Essential role for the c-met receptor in the migration of myogenic precursor cells into the limb bud." Nature 376(6543): 768-71.

Blake, J.A. and M.R. Ziman (2005). "Pax3 transcripts in melanoblast development." Dev Growth Differ 47(9): 627-35.

Bober, E., T. Franz, H. H. Arnold, P. Gruss and P. Tremblay (1994). "Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells." Development 120(3): 603-12.

Bondurand, N., V. Pingault, D. E. Goerich, N. Lemort, E. Sock, C. L. Caignec, M. Wegner and M. Goossens (2000). "Interaction among SOX10, PAX3 and MITF, three genes altered in Waardenburg syndrome." Hum Mol Genet 9(13): 1907-17.

Bopp, D., M. Burri, S. Baumgartner, G. Frigerio and M. Noll (1986). "Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of *Drosophila*." Cell 47(6): 1033-40.

Bopp, D., E. Jamet, S. Baumgartner, M. Burri and M. Noll (1989). "Isolation of two tissue-specific *Drosophila* paired box genes, Pox meso and Pox neuro." Embo J 8(11): 3447-57.

Borycki, A. G., J. Li, F. Jin, C. P. Emerson and J. A. Epstein (1999). "Pax3 functions in cell survival and in pax7 regulation." Development 126(8): 1665-74.

Bruun, J. A., E. I. S. Thomassen, K. Kristiansen, G. Tylden, T. Holm, I. Mikkola, G. Bjorkoy, and T. Johansen (2005). "The third helix of the homeodomain of the paired class homeodomain proteins acts as a recognition helix for both DNA and protein interactions." Nucleic Acids Research 33(8): 2661-75.

Burri, M., Y. Tromvoukis, D. Bopp, G. Frigerio and M. Noll (1989). "Conservation of the paired domain in metazoans and its structure in three isolated human genes." Embo J 8(4): 1183-90.

Callaerts, P., G. Halder and W. J. Gehring (1997). "PAX-6 in development and evolution." Annu Rev Neurosci 20: 483-532.

Chalepakis, G., R. Fritsch, H. Fickenscher, U. Deutsch, M. Goulding and P. Gruss (1991). "The molecular basis of the undulated/Pax-1 mutation." Cell 66(5): 873-84.

Chalepakis, G., M. Goulding, A. Read, T. Strachan and P. Gruss (1994). "Molecular basis of splotch and Waardenburg Pax-3 mutations." Proc Natl Acad Sci U S A 91(9): 3685-9.

Chalepakis, G. and P. Gruss (1995). "Identification of DNA recognition sequences for the Pax3 paired domain." Gene 162(2): 267-70.

Chalepakis, G., F. S. Jones, G. M. Edelman and P. Gruss (1994). "Pax-3 contains domains for transcription activation and transcription inhibition." Proc Natl Acad Sci U S A 91(26): 12745-9.

Chalepakis, G., A. Stoykova, J. Wijnholds, P. Tremblay and P. Gruss (1993). "Pax: gene regulators in the developing nervous system." J Neurobiol 24(10): 1367-84.

Chalepakis, G., J. Wijnholds, P. Giese, M. Schachner and P. Gruss (1994). "Characterization of Pax-6 and Hoxa-1 binding to the promoter region of the neural cell adhesion molecule L1." DNA Cell Biol 13(9): 891-900.

Chen, G., J. Fernandez, S. Mische and A. J. Courey (1999). "A functional interaction between the histone deacetylase Rpd3 and the corepressor groucho in *Drosophila* development." Genes Dev 13(17): 2218-30.

Choi, C. Y., Y. H. Kim, H. J. Kwon and Y. Kim (1999). "The homeodomain protein NK-3 recruits Groucho and a histone deacetylase complex to repress transcription." J Biol Chem 274(47): 33194-7.

Chow, R. L., C. R. Altmann, R. A. Lang and A. Hemmati-Brivanlou (1999). "Pax6 induces ectopic eyes in a vertebrate." Development 126(19): 4213-22.

Clarke, N. D., C. R. Kissinger, J. Desjarlais, G. L. Gilliland and C. O. Pabo (1994). "Structural studies of the engrailed homeodomain." Protein Sci 3(10): 1779-87.

Conway, S. J., D. J. Henderson and A. J. Copp (1997). "Pax3 is required for cardiac neural crest migration in the mouse: evidence from the splotch (Sp2H) mutant." Development 124(2): 505-14.

Conway, S. J., D. J. Henderson, M. L. Kirby, R. H. Anderson and A. J. Copp (1997). "Development of a lethal congenital heart defect in the splotch (Pax3) mutant mouse." Cardiovasc Res 36(2): 163-73.

Corry, G.N. and D.A. Underhill (2005). " Pax3 target gene recognition occurs through distinct modes that are differentially affected by disease-associated mutations." Pigment Cell Res 18(6): 427-38.

Cox, M., P. J. van Tilborg, W. de Laat, R. Boelens, H. C. van Leeuwen, P. C. van der Vliet and R. Kaptein (1995). "Solution structure of the Oct-1 POU homeodomain determined by NMR and restrained molecular dynamics." J Biomol NMR 6(1): 23-32.

Cvekl, A., F. Kashanchi, J. N. Brady and J. Piatigorsky (1999). "Pax-6 interactions with TATA-box-binding protein and retinoblastoma protein." Invest Ophthalmol Vis Sci 40(7): 1343-50.

Czerny, T. and M. Busslinger (1995). "DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5)." Mol Cell Biol 15(5): 2858-71.

Czerny, T., G. Halder, U. Kloter, A. Souabni, W. J. Gehring and M. Busslinger (1999). "twin of eyeless, a second Pax-6 gene of *Drosophila*, acts upstream of eyeless in the control of eye development." Mol Cell 3(3): 297-307.

Czerny, T., G. Schaffner and M. Busslinger (1993). "DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site." Genes Dev 7(10): 2048-61.

Dahl, E., H. Koseki and R. Balling (1997). "Pax genes and organogenesis." Bioessays 19(9): 755-65.

Daston, G., E. Lamar, M. Olivier and M. Goulding (1996). "Pax-3 is necessary for migration but not differentiation of limb muscle precursors in the mouse." Development 122(3): 1017-27.

Davicioni, E., F. G. Finckenstein, V. Shahbazian, J. D. Buckley, T. J. Triche and M. J. Anderson (2006). "Identification of a PAX-FKHR gene expression signature that defines molecular classes and determines the prognosis of alveolar rhabdomyosarcomas." Cancer Res 66(14): 6936-46.

Davis, C. A. and A. L. Joyner (1988). "Expression patterns of the homeo box-containing genes En-1 and En-2 and the proto-oncogene int-1 diverge during mouse development." Genes Dev 2(12B): 1736-44.

Davis, R. J., C. M. D'Cruz, M. A. Lovell, J. A. Biegel and F. G. Barr (1994). "Fusion of PAX7 to FKHR by the variant t(1;13)(p36;q14) translocation in alveolar rhabdomyosarcoma." Cancer Res 54(11): 2869-72.

Deutsch, U., G. R. Dressler and P. Gruss (1988). "Pax 1, a member of a paired box homologous murine gene family, is expressed in segmented structures during development." Cell 53(4): 617-25.

Dickie, M. M. (1964). "New Splotch Alleles in the Mouse." J Hered 55: 97-101.

Dietrich, S. and P. Gruss (1995). "undulated phenotypes suggest a role of Pax-1 for the development of vertebral and extravertebral structures." Dev Biol 167(2): 529-48.

Dohrmann, C., P. Gruss and L. Lemaire (2000). "Pax genes and the differentiation of hormone-producing endocrine cells in the pancreas." Mech Dev 92(1): 47-54.

Dorfler, P. and M. Busslinger (1996). "C-terminal activating and inhibitory domains determine the transactivation potential of BSAP (Pax-5), Pax-2 and Pax-8." Embo J 15(8): 1971-82.

Dressler, G. R., U. Deutsch, K. Chowdhury, H. O. Nornes and P. Gruss (1990). "Pax2, a new murine paired-box-containing gene and its expression in the developing excretory system." Development 109(4): 787-95.

Dressler, G. R. and E. C. Douglass (1992). "Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor." Proc Natl Acad Sci U S A 89(4): 1179-83.

Eberhard, D. and M. Busslinger (1999). "The partial homeodomain of the transcription factor Pax-5 (BSAP) is an interaction motif for the retinoblastoma and TATA-binding proteins." Cancer Res 59(7 Suppl): 1716s-1724s; discussion 1724s-1725s.

Eberhard, D., G. Jimenez, B. Heavey and M. Busslinger (2000). "Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family." Embo J 19(10): 2292-303.

Eccles, M. R. and L. A. Schimmenti (1999). "Renal-coloboma syndrome: a multi-system developmental disorder caused by PAX2 mutations." Clin Genet 56(1): 1-9.

Eccles, M. R., K. Yun, A. E. Reeve and A. E. Fidler (1995). "Comparative in situ hybridization analysis of PAX2, PAX8, and WT1 gene transcription in human fetal kidney and Wilms' tumors." Am J Pathol 146(1): 40-5.

Engleka, K.A., A. D. Gitler, M. Zhang, D. D. Zhou, F. A. High and J. A. Epstein (2005). "Insertion of Cre into the Pax3 locus creates a new allele of Splotch and identifies unexpected Pax3 derivatives." Dev Biol 280(2): 396-406.

Epstein, D. J., M. Vekemans and P. Gros (1991). "Splotch (Sp2H), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3." Cell 67(4): 767-74.

Epstein, D. J., K. J. Vogan, D. G. Trasler and P. Gros (1993). "A mutation within intron 3 of the Pax-3 gene produces aberrantly spliced mRNA transcripts in the splotch (Sp) mouse mutant." Proc Natl Acad Sci U S A 90(2): 532-6.

Epstein, J., J. Cai, T. Glaser, L. Jepeal and R. Maas (1994). "Identification of a Pax paired domain recognition sequence and evidence for DNA-dependent conformational changes." J Biol Chem 269(11): 8355-61.

Epstein, J. A., T. Glaser, J. Cai, L. Jepeal, D. S. Walton and R. L. Maas (1994). "Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing." Genes Dev 8(17): 2022-34.

Epstein, J. A., P. Lam, L. Jepeal, R. L. Maas and D. N. Shapiro (1995). "Pax3 inhibits myogenic differentiation of cultured myoblast cells." J Biol Chem 270(20): 11719-22.

Epstein, J. A., D. N. Shapiro, J. Cheng, P. Y. Lam and R. L. Maas (1996). "Pax3 modulates expression of the c-Met receptor during limb muscle development." Proc Natl Acad Sci U S A 93(9): 4213-8.

Evans, E. P., Burtenshaw, M. D., Beechey, C. V., and Searle, A. G. (1988). "A splotch locus deletion visable by Giemsa banding." Mouse News Letter 81: 66.

Favor, J., R. Sandulache, A. Neuhauser-Klaus, W. Pretsch, B. Chatterjee, E. Senft, W. Wurst, V. Blanquet, P. Grimes, R. Sporle and K. Schughart (1996). "The mouse Pax2(1Neu) mutation is identical to a human PAX2 mutation in a family with renal-coloboma syndrome and results in developmental defects of the brain, ear, eye, and kidney." Proc Natl Acad Sci U S A 93(24): 13870-5.

Fisher, A. L. and M. Caudy (1998). "Groucho proteins: transcriptional corepressors for specific subsets of DNA-binding transcription factors in vertebrates and invertebrates." Genes Dev 12(13): 1931-40.

Fitzsimmons, D., W. Hodsdon, W. Wheat, S. M. Maira, B. Wasylyk and J. Hagman (1996). "Pax-5 (BSAP) recruits Ets proto-oncogene family proteins to form functional ternary complexes on a B-cell-specific promoter." Genes Dev 10(17): 2198-211.

Fortin, A. S., D. A. Underhill and P. Gros (1997). "Reciprocal effect of Waardenburg syndrome mutations on DNA binding by the Pax-3 paired domain and homeodomain." Hum Mol Genet 6(11): 1781-90.

Fortin, A. S., D. A. Underhill and P. Gros (1998). "Helix 2 of the paired domain plays a key role in the regulation of DNA-binding by the Pax-3 homeodomain." Nucleic Acids Res 26(20): 4574-81.

Franz, T. (1989). "Persistent truncus arteriosus in the Splotch mutant mouse." Anat Embryol (Berl) 180(5): 457-64.

Franz, T. (1990). "Defective ensheathment of motoric nerves in the Splotch mutant mouse." Acta Anat (Basel) 138(3): 246-53.

Franz, T. (1993). "The Splotch (Sp1H) and Splotch-delayed (Spd) alleles: differential phenotypic effects on neural crest and limb musculature." Anat Embryol (Berl) 187(4): 371-7.

Franz, T., R. Kothary, M. A. Surani, Z. Halata and M. Grim (1993). "The Splotch mutation interferes with muscle development in the limbs." Anat Embryol (Berl) 187(2): 153-60.

Fredericks, W. J., N. Galili, S. Mukhopadhyay, G. Rovera, J. Bennicelli, F. G. Barr and F. J. Rauscher, 3rd (1995). "The PAX3-FKHR fusion protein created by the t(2;13) translocation in alveolar rhabdomyosarcomas is a more potent transcriptional activator than PAX3." Mol Cell Biol 15(3): 1522-35.

Frigerio, G., M. Burri, D. Bopp, S. Baumgartner and M. Noll (1986). "Structure of the segmentation gene paired and the *Drosophila* PRD gene set as part of a gene network." Cell 47(5): 735-46.

Frillingos, S., M. Sahin-Toth, J. Wu and H. R. Kaback (1998). "Cys-scanning mutagenesis: a novel approach to structure function relationships in polytopic membrane proteins." Faseb J 12(13): 1281-99.

Fujioka, M., P. Miskiewicz, L. Raj, A. A. Gulledge, M. Weir and T. Goto (1996). "*Drosophila* Paired regulates late even-skipped expression through a composite binding site for the paired domain and the homeodomain." Development 122(9): 2697-707.

Fujitani, Y., Y. Kajimoto, T. Yasuda, T. A. Matsuoka, H. Kaneto, Y. Umayahara, N. Fujita, H. Watada, J. I. Miyazaki, Y. Yamasaki and M. Hori (1999). "Identification of a portable repression domain and an E1A-responsive activation domain in Pax4: a possible role of Pax4 as a transcriptional repressor in the pancreas." Mol Cell Biol 19(12): 8281-91.

Galibert, M. D., U. Yavuzer, T. J. Dexter and C. R. Goding (1999). "Pax3 and regulation of the melanocyte-specific tyrosinase-related protein-1 promoter." J Biol Chem 274(38): 26894-900.

Galili, N., R. J. Davis, W. J. Fredericks, S. Mukhopadhyay, F. J. Rauscher, 3rd, B. S. Emanuel, G. Rovera and F. G. Barr (1993). "Fusion of a fork head domain gene to PAX3 in the solid tumour alveolar rhabdomyosarcoma." Nat Genet 5(3): 230-5.

Garvie, C. W., J. Hagman and C. Wolberger (2001). "Structural studies of Ets-1/Pax5 complex formation on DNA." Mol Cell 8(6): 1267-76.

Gehring, W. J., M. Affolter and T. Burglin (1994). "Homeodomain proteins." Annu Rev Biochem 63: 487-526.

Gehring, W. J., Y. Q. Qian, M. Billeter, K. Furukubo-Tokunaga, A. F. Schier, D. Resendez-Perez, M. Affolter, G. Otting and K. Wuthrich (1994). "Homeodomain-DNA recognition." Cell 78(2): 211-23.

Glaser, T., L. Jepeal, J. G. Edwards, S. R. Young, J. Favor and R. L. Maas (1994). "PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects." Nat Genet 7(4): 463-71.

Glaser, T., D. S. Walton and R. L. Maas (1992). "Genomic structure, evolutionary conservation and aniridia mutations in the human PAX6 gene." Nat Genet 2(3): 232-9.

Gomes, X. V., L. A. Henricksen and M. S. Wold (1996). "Proteolytic mapping of human replication protein A: evidence for multiple structural domains and a conformational change upon interaction with single-stranded DNA." Biochemistry 35(17): 5586-95.

Goulding, M., A. Lumsden and A. J. Paquette (1994). "Regulation of Pax-3 expression in the dermomyotome and its role in muscle development." Development 120(4): 957-71.

Goulding, M., S. Sterrer, J. Fleming, R. Balling, J. Nadeau, K. J. Moore, S. D. Brown, K. P. Steel and P. Gruss (1993). "Analysis of the Pax-3 gene in the mouse mutant splotch." Genomics 17(2): 355-63.

Goulding, M. D., G. Chalepakis, U. Deutsch, J. R. Erselius and P. Gruss (1991). "Pax-3, a novel murine DNA binding protein expressed during early neurogenesis." Embo J 10(5): 1135-47.

Green, N. S., E. Reisler, and K. N. Houk (2001). "Quantitative evaluation of the lengths of homobifunctional protein cross-linking reagents used as molecular rulers." Protein Sci 10(7): 1293-1304.

Grindley, J. C., D. R. Davidson and R. E. Hill (1995). "The role of Pax-6 in eye and nasal development." Development 121(5): 1433-42.

Grueneberg, D. A., S. Natesan, C. Alexandre and M. Z. Gilman (1992). "Human and *Drosophila* homeodomain proteins that enhance the DNA-binding activity of serum response factor." Science 257(5073): 1089-95.

Grueneberg, D. A., K. J. Simon, K. Brennan and M. Gilman (1995). "Sequence-specific targeting of nuclear signal transduction pathways by homeodomain proteins." Mol Cell Biol 15(6): 3318-26.

Hagemeier, C., A. J. Bannister, A. Cook and T. Kouzarides (1993). "The activation domain of transcription factor PU.1 binds the retinoblastoma (RB) protein and the transcription factor

TFIID in vitro: RB shows sequence similarity to TFIID and TFIIB." Proc Natl Acad Sci U S A 90(4): 1580-4.

Hagemeier, C., A. Cook and T. Kouzarides (1993). "The retinoblastoma protein binds E2F residues required for activation in vivo and TBP binding in vitro." Nucleic Acids Res 21(22): 4998-5004.

Hagman, J., W. Wheat, D. Fitzsimmons, W. Hodsdon, J. Negri and F. Dizon (2000). "Pax-5/BSAP: regulator of specific gene expression and differentiation in B lymphocytes." Curr Top Microbiol Immunol 245(1): 169-94.

Halder, G., P. Callaerts and W. J. Gehring (1995). "Induction of ectopic eyes by targeted expression of the eyeless gene in Drosophila." Science 267(5205): 1788-92.

Hanson, I. M., J. M. Fletcher, T. Jordan, A. Brown, D. Taylor, R. J. Adams, H. H. Punnett and V. van Heyningen (1994). "Mutations at the PAX6 locus are found in heterogeneous anterior segment malformations including Peters' anomaly." Nat Genet 6(2): 168-73.

Hanson, I. M., A. Seawright, K. Hardman, S. Hodgson, D. Zaletayev, G. Fekete and V. van Heyningen (1993). "PAX6 mutations in aniridia." Hum Mol Genet 2(7): 915-20.

Hateboer, G., H. T. Timmers, A. K. Rustgi, M. Billaud, L. J. van 't Veer and R. Bernards (1993). "TATA-binding protein and the retinoblastoma gene product bind to overlapping epitopes on c-Myc and adenovirus E1A protein." Proc Natl Acad Sci U S A 90(18): 8489-93.

Hauck, B., W. J. Gehring and U. Walldorf (1999). "Functional analysis of an eye specific enhancer of the eyeless gene in *Drosophila.*" Proc Natl Acad Sci U S A 96(2): 564-9.

He, M., L. Jin and B. Austen (1993). "Specificity of factor Xa in the cleavage of fusion proteins." J Protein Chem 12(1): 1-5.

Heanue, T. A., R. Reshef, R. J. Davis, G. Mardon, G. Oliver, S. Tomarev, A. B. Lassar and C. J. Tabin (1999). "Synergistic regulation of vertebrate muscle development by Dach2, Eya2, and Six1, homologs of genes required for *Drosophila* eye formation." Genes Dev 13(24): 3231-43.

Hill, R. E., J. Favor, B. L. Hogan, C. C. Ton, G. F. Saunders, I. M. Hanson, J. Prosser, T. Jordan, N. D. Hastie and V. van Heyningen (1991). "Mouse small eye results from mutations in a paired-like homeobox-containing gene." Nature 354(6354): 522-5.

Hodgkinson, C. A., K. J. Moore, A. Nakayama, E. Steingrimsson, N. G. Copeland, N. A. Jenkins and H. Arnheiter (1993). "Mutations at the mouse microphthalmia locus are associated with defects in a gene encoding a novel basic-helix-loop-helix-zipper protein." Cell 74(2): 395-404.

Hogan, B. L., E. M. Hirst, G. Horsburgh and C. M. Hetherington (1988). "Small eye (Sey): a mouse model for the genetic analysis of craniofacial abnormalities." Development 103 Suppl: 115-9.

Hogan, B. L., G. Horsburgh, J. Cohen, C. M. Hetherington, G. Fisher and M. F. Lyon (1986). "Small eyes (Sey): a homozygous lethal mutation on chromosome 2 which affects the differentiation of both lens and nasal placodes in the mouse." J Embryol Exp Morphol 97: 95-110.

Hollenbach, A. D., J. E. Sublett, C. J. McPherson and G. Grosveld (1999). "The Pax3-FKHR oncoprotein is unresponsive to the Pax3-associated repressor hDaxx." Embo J 18(13): 3702-11.

Hornyak, T. J., D. J. Hayes, L. Y. Chiu and E. B. Ziff (2001). "Transcription factors in melanocyte development: distinct roles for Pax-3 and Mitf." Mech Dev 101(1-2): 47-59.

Hoth, C. F., A. Milunsky, N. Lipsky, R. Sheffer, S. K. Clarren and C. T. Baldwin (1993). "Mutations in the paired domain of the human PAX3 gene cause Klein-Waardenburg syndrome (WS-III) as well as Waardenburg syndrome type I (WS-I)." Am J Hum Genet 52(3): 455-62.

Hussain, M. A. and J. F. Habener (1999). "Glucagon gene transcription activation mediated by synergistic interactions of pax-6 and cdx-2 with the p300 co-activator." J Biol Chem 274(41): 28950-7.

Jabet, C., R. Gitti, M. F. Summers and C. Wolberger (1999). "NMR studies of the pbx1 TALE homeodomain protein free in solution and bound to DNA: proposal for a mechanism of HoxB1-Pbx1-DNA complex assembly." J Mol Biol 291(3): 521-30.

Jiao, Z., Z. G. Zhang, T. J. Hornyak, A. Hozeska, R.L. Zhang, Y. Wang, L. Wang, C. Roberts, F. M. Strickland and M. Chopp (2006). "Dopachrome tautomerase (Dct) regulates neural progenitor cell proliferation." Dev Biol [Epub ahead of print]

Jordan, T., I. Hanson, D. Zaletayev, S. Hodgson, J. Prosser, A. Seawright, N. Hastie and V. van Heyningen (1992). "The human PAX6 gene is mutated in two patients with aniridia." Nat Genet 1(5): 328-32.

Jostes, B., C. Walther and P. Gruss (1990). "The murine paired box gene, Pax7, is expressed specifically during the development of the nervous and muscular system." Mech Dev 33(1): 27-37.

Joyner, A. L., T. Kornberg, K. G. Coleman, D. R. Cox and G. R. Martin (1985). "Expression during embryogenesis of a mouse gene with sequence homology to the *Drosophila* engrailed gene." Cell 43(1): 29-37.

Julien, M. and P. Gros (2000). "Nucleotide-induced conformational changes in P-glycoprotein and in nucleotide binding site mutants monitored by trypsin sensitivity." Biochemistry 39(15): 4559-68.

Jun, S. and C. Desplan (1996). "Cooperative interactions between paired domain and homeodomain." Development 122(9): 2639-50.

Jun, S., R. V. Wallen, A. Goriely, B. Kalionis, and C. Desplan (1998). "Lune/eye gone, a Pax like protein, uses a partial paired domain and a homeodomain for DNA recognition." Proc Natl Acad Sci U S A 10;9(23): 13720-5.

Jung, K., H. Jung, J. Wu, G. G. Prive and H. R. Kaback (1993). "Use of site-directed fluorescence labeling to study proximity relationships in the lactose permease of Escherichia coli." Biochemistry 32(46): 12273-8.

Kappen, C., K. Schughart and F. H. Ruddle (1993). "Early evolutionary origin of major homeodomain sequence classes." Genomics 18(1): 54-70.

Kawiak, J., E. Brzoska, I. Grabowska, G. Hoser, W. Streminska, D. Wasilewska, E. K.Machaj, Z. Pojda and J. Moraczewski (2006). "Contribution of stem cells to skeletal muscle regeneration." Folia Histochem Cytobiol 44(2): 75-9.

Keller, S. A., J. M. Jones, A. Boyle, L. L. Barrow, P. D. Killen, D. G. Green, N. V. Kapousta, P. F. Hitchcock, R. T. Swank and M. H. Meisler (1994). "Kidney and retinal defects (Krd), a transgene-induced mutation with a deletion of mouse chromosome 19 that includes the Pax2 locus." Genomics 23(2): 309-20.

Kessel, M. and P. Gruss (1990). "Murine developmental control genes." Science 249(4967): 374-9.

Khan, J., M. L. Bittner, L. H. Saal, U. Teichmann, D. O. Azorsa, G. C. Gooden, W. J. Pavan, J. M. Trent and P. S. Meltzer (1999). "cDNA microarrays detect activation of a myogenic transcription program by the PAX3-FKHR fusion oncogene." Proc Natl Acad Sci U S A 96(23): 13264-9.

Kioussi, C., M. K. Gross and P. Gruss (1995). "Pax3: a paired domain gene as a regulator in PNS myelination." Neuron 15(3): 553-62.

Kiriakidou, M., D. A. Driscoll, J. M. Lopez-Guisa and J. F. Strauss, 3rd (1997). "Cloning and expression of primate Daxx cDNAs and mapping of the human gene to chromosome 6p21.3 in the MHC region." DNA Cell Biol 16(11): 1289-98.

Kissinger, C. R., B. S. Liu, E. Martin-Blanco, T. B. Kornberg and C. O. Pabo (1990). "Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: a framework for understanding homeodomain-DNA interactions." Cell 63(3): 579-90.

Koop, K. E., L. M. MacDonald and C. G. Lobe (1996). "Transcripts of Grg4, a murine groucho-related gene, are detected in adjacent tissues to other murine neurogenic gene homologues during embryonic development." Mech Dev 59(1): 73-87.

Kozmik, Z., T. Czerny and M. Busslinger (1997). "Alternatively spliced insertions in the paired domain restrict the DNA sequence specificity of Pax6 and Pax8." Embo J 16(22): 6793-803.

Kozmik, Z., R. Kurzbauer, P. Dorfler and M. Busslinger (1993). "Alternative splicing of Pax-8 gene transcripts is developmentally regulated and generates isoforms with different transactivation properties." Mol Cell Biol 13(10): 6024-35.

Kozmik, Z., S. Wang, P. Dorfler, B. Adams and M. Busslinger (1992). "The promoter of the CD19 gene is a target for the B-cell-specific transcription factor BSAP." Mol Cell Biol 12(6): 2662-72.

Krauss, S., M. Maden, N. Holder and S. W. Wilson (1992). "Zebrafish pax[b] is involved in the formation of the midbrain-hindbrain boundary." Nature 360(6399): 87-9.

Kwang, S. J., S. M. Brugger, A. Lazik, A. E. Merrill, L. Y. Wu, Y. H. Liu, M. Ishii, F. O. Sangiorgi, M. Rauchman, H. M. Sucov, R. L. Maas and R. E. Maxson, Jr. (2002). "Msx2 is an immediate downstream effector of Pax3 in the development of the murine cardiac neural crest." Development 129(2): 527-38.

Lalwani, A. K., J. R. Brister, J. Fex, K. M. Grundfast, B. Ploplis, T. B. San Agustin and E. R. Wilcox (1995). "Further elucidation of the genomic structure of PAX3, and identification of two different point mutations within the PAX3 homeobox that cause Waardenburg syndrome type 1 in two families." Am J Hum Genet 56(1): 75-83.

Lam, P. Y., J. E. Sublett, A. D. Hollenbach and M. F. Roussel (1999). "The oncogenic potential of the Pax3-FKHR fusion protein requires the Pax3 homeodomain recognition helix but not the Pax3 paired-box DNA binding domain." Mol Cell Biol 19(1): 594-601.

Lang, D., M. M. Lu, L. Huang, K. A. Engleka, M. Zhang, E. Y. Chu, S. Lipner, A. Skoultchi, S. E. Millar and J. A. Epstein (2005). "Pax3 functions at a nodal point in melanocyte stem cell differentiation." Nature 433(7028): 884-7.

Lechner, M. S. and G. R. Dressler (1996). "Mapping of Pax-2 transcription activation domains." J Biol Chem 271(35): 21088-93.

Liu, R. and F. J. Sharom (1996). "Site-directed fluorescence labeling of P-glycoprotein on cysteine residues in the nucleotide binding domains." Biochemistry 35(36): 11865-73.

Liu, R. and F. J. Sharom (1997). "Fluorescence studies on the nucleotide binding domains of the P-glycoprotein multidrug transporter." Biochemistry 36(10): 2836-43.

Loeken, M.R. (2006) "Advances in understanding the molecular causes of diabetes-induced birth defects." J Soc Gynecol Investig 13(1): 2-10.

Loo, T. W. and D. M. Clarke (2000). "The packing of the transmembrane segments of human multidrug resistance P-glycoprotein is revealed by disulfide cross-linking analysis." J Biol Chem 275(8): 5253-6.

Macchia, P. E., P. Lapi, H. Krude, M. T. Pirro, C. Missero, L. Chiovato, A. Souabni, M. Baserga, V. Tassi, A. Pinchera, G. Fenzi, A. Gruters, M. Busslinger and R. Di Lauro (1998). "PAX8 mutations associated with congenital hypothyroidism caused by thyroid dysgenesis." Nat Genet 19(1): 83-6.

Magnaghi, P., C. Roberts, S. Lorain, M. Lipinski and P. J. Scambler (1998). "HIRA, a mammalian homologue of Saccharomyces cerevisiae transcriptional co-repressors, interacts with Pax3." Nat Genet 20(1): 74-7.

Maitra, S. and M. Atchison (2000). "BSAP can repress enhancer activity by targeting PU.1 function." Mol Cell Biol 20(6): 1911-22.

Mansouri, A., K. Chowdhury and P. Gruss (1998). "Follicular cells of the thyroid gland require Pax8 gene function." Nat Genet 19(1): 87-90.

Mansouri, A. and P. Gruss (1998). "Pax3 and Pax7 are expressed in commissural neurons and restrict ventral neuronal identity in the spinal cord." Mech Dev 78(1-2): 171-8.

Mansouri, A., M. Hallonet and P. Gruss (1996). "Pax genes and their roles in cell differentiation and development." Curr Opin Cell Biol 8(6): 851-7.

Mansouri, A., A. Stoykova, M. Torres and P. Gruss (1996). "Dysgenesis of cephalic neural crest derivatives in Pax7-/- mutant mice." Development 122(3): 831-8.

Margue, C. M., M. Bernasconi, F. G. Barr and B. W. Schafer (2000). "Transcriptional modulation of the anti-apoptotic protein BCL-XL by the paired box transcription factors PAX3 and PAX3/FKHR." Oncogene 19(25): 2921-9.

Maroto, M., R. Reshef, A. E. Munsterberg, S. Koester, M. Goulding and A. B. Lassar (1997). "Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue." Cell 89(1): 139-48.

Matsushita, T., T. Yamaoka, S. Otsuka, M. Moritani, T. Matsumoto and M. Itakura (1998). "Molecular cloning of mouse paired-box-containing gene (Pax)-4 from an islet beta cell line and deduced sequence of human Pax-4." Biochem Biophys Res Commun 242(1): 176-80.

Mechetner, E. B., B. Schott, B. S. Morse, W. D. Stein, T. Druley, K. A. Davis, T. Tsuruo and I. B. Roninson (1997). "P-glycoprotein function involves conformational transitions detectable by differential immunoreactivity." Proc Natl Acad Sci U S A 94(24): 12908-13.

Menezes, M. E., P. D. Roepe and H. R. Kaback (1990). "Design of a membrane transport protein for fluorescence spectroscopy." Proc Natl Acad Sci U S A 87(5): 1638-42.

Miskiewicz, P., D. Morrissey, Y. Lan, L. Raj, S. Kessler, M. Fujioka, T. Goto and M. Weir (1996). "Both the paired domain and homeodomain are required for in vivo function of *Drosophila* Paired." Development 122(9): 2709-18.

Monsoro-Burq, A.H., E. Wang, and R. Harland (2005). "Msx1 and Pax3 cooperate to mediate FGF8 and WNT signals during Xenopus neural crest induction." Dev Cell 8(2): 167-78.

Morsomme, P., S. Dambly, O. Maudoux and M. Boutry (1998). "Single point mutations distributed in 10 soluble and membrane regions of the Nicotiana plumbaginifolia plasma membrane PMA2 H+-ATPase activate the enzyme and modify the structure of the C-terminal region." J Biol Chem 273(52): 34837-42.

Murakami, M., J. Tominaga, R. Makita, Y. Uchijima, Y. Kurihara, O. Nakagawa, T. Asano and H. Kurihara (2006). "Transcriptional activity of Pax3 is co-activated by TAZ." Biochem Biophys Res Commun 339(2): 533-9.

Nabarro, S., N. Himoudi, A. Papanastasiou, K. Gilmour, S. Gibson, N. Sbire, A. Thrasher, M. P. Blundell, M. Hubank, G. Canderan and J. Anderson (2005). "Coordinated oncogenic transformation and inhibition of host immune responses by the PAX3-FKHR fusion oncoprotein." J Exp Med 202(10): 1399-410.

Neubuser, A., H. Koseki and R. Balling (1995). "Characterization and developmental expression of Pax9, a paired-box-containing gene related to Pax1." Dev Biol 170(2): 701-16.

Neubuser, A., H. Peters, R. Balling and G. R. Martin (1997). "Antagonistic interactions between FGF and BMP signaling pathways: a mechanism for positioning the sites of tooth formation." Cell 90(2): 247-55.

Noll, M. (1993). "Evolution and role of Pax genes." Curr Opin Genet Dev 3(4): 595-605.

Nornes, H. O., G. R. Dressler, E. W. Knapik, U. Deutsch and P. Gruss (1990). "Spatially and temporally restricted expression of Pax2 during murine neurogenesis." Development 109(4): 797-809.

Nornes, S., I. Mikkola, S. Krauss, M. Delghandi, M. Perander and T. Johansen (1996). "Zebrafish Pax9 encodes two proteins with distinct C-terminal transactivating domains of different potency negatively regulated by adjacent N-terminal sequences." J Biol Chem 271(43): 26914-23.

Nusslein-Volhard, C. and E. Wieschaus (1980). "Mutations affecting segment number and polarity in *Drosophila*." Nature 287(5785): 795-801.

Nutt, S. L., A. M. Morrison, P. Dorfler, A. Rolink and M. Busslinger (1998). "Identification of BSAP (Pax-5) target genes in early B-cell development by loss- and gain-of-function experiments." Embo J 17(8): 2319-33.

Ohto, H., S. Kamada, K. Tago, S. I. Tominaga, H. Ozaki, S. Sato and K. Kawakami (1999). "Cooperation of six and eya in activation of their target genes through nuclear translocation of Eya." Mol Cell Biol 19(10): 6815-24.

Otto, A., C. Schmidt, and K. Patel (2006). "Pax3 and Pax7 expression and regulation in the avian embryo." Anat Embryol (Berl) 211(4): 293-310.

Oustanina, S., G. Hause, and T. Braun (2004). "Pax7 directs postnatal renewal and propagation of myogenic satellite cells but not their specification." Embo J 23(16): 3430-9.

Parkhurst, S. M. (1998). "Groucho: making its Marx as a transcriptional co-repressor." Trends Genet 14(4): 130-2.

Peters, H., A. Neubuser, K. Kratochwil and R. Balling (1998). "Pax9-deficient mice lack pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities." Genes Dev 12(17): 2735-47.

Peters, H., B. Wilm, N. Sakai, K. Imai, R. Maas and R. Balling (1999). "Pax1 and Pax9 synergistically regulate vertebral column development." Development 126(23): 5399-408.

Pingault, V., N. Bondurand, K. Kuhlbrodt, D. E. Goerich, M. O. Prehu, A. Puliti, B. Herbarth, I. Hermans-Borgmeyer, E. Legius, G. Matthijs, J. Amiel, S. Lyonnet, I. Ceccherini, G. Romeo, J. C. Smith, A. P. Read, M. Wegner and M. Goossens (1998). "SOX10 mutations in patients with Waardenburg-Hirschsprung disease." Nat Genet 18(2): 171-3.

Piper, D. E., A. H. Batchelor, C. P. Chang, M. L. Cleary and C. Wolberger (1999). "Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation." Cell 96(4): 587-97.

Plachov, D., K. Chowdhury, C. Walther, D. Simon, J. L. Guenet and P. Gruss (1990). "Pax8, a murine paired box gene expressed in the developing excretory system and thyroid gland." Development 110(2): 643-51.

Poleev, A., H. Fickenscher, S. Mundlos, A. Winterpacht, B. Zabel, A. Fidler, P. Gruss and D. Plachov (1992). "PAX8, a human paired box gene: isolation and expression in developing thyroid, kidney and Wilms' tumors." Development 116(3): 611-23.

Potterf, S. B., M. Furumura, K. J. Dunn, H. Arnheiter and W. J. Pavan (2000). "Transcription factor hierarchy in Waardenburg syndrome: regulation of MITF expression by SOX10 and PAX3." Hum Genet 107(1): 1-6.

Puschel, A. W., P. Gruss and M. Westerfield (1992). "Sequence and expression pattern of pax-6 are highly conserved between zebrafish and mice." Development 114(3): 643-51.

Qian, Y. Q., M. Billeter, G. Otting, M. Muller, W. J. Gehring and K. Wuthrich (1989). "The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors." Cell 59(3): 573-80.

Qian, Y. Q., D. Resendez-Perez, W. J. Gehring and K. Wuthrich (1994). "The des(1-6)antennapedia homeodomain: comparison of the NMR solution structure and the DNA-binding affinity with the intact Antennapedia homeodomain." Proc Natl Acad Sci U S A 91(9): 4091-5.

Qu, Q. and F. J. Sharom (2001). "FRET analysis indicates that the two ATPase active sites of the P-glycoprotein multidrug transporter are closely associated." Biochemistry 40(5): 1413-22.

Quiring, R., U. Walldorf, U. Kloter and W. J. Gehring (1994). "Homology of the eyeless gene of *Drosophila* to the Small eye gene in mice and Aniridia in humans." Science 265(5173): 785-9.

Rebar, E. J. and C. O. Pabo (1994). "Zinc finger phage: affinity selection of fingers with new DNA binding specificities." Science 263: 671-73.

Reeves, F. C., W. J. Fredericks, F. J. Rauscher, 3rd and K. A. Lillycrop (1998). "The DNA binding activity of the paired box transcription factor Pax-3 is rapidly downregulated during neuronal cell differentiation." FEBS Lett 422(1): 118-22.

Ritz-Laser, B., A. Estreicher, N. Klages, S. Saule and J. Philippe (1999). "Pax-6 and Cdx-2/3 interact to activate glucagon gene expression on the G1 control element." J Biol Chem 274(7): 4124-32.

Rothman, A., Y. Gerchman, E. Padan and S. Schuldiner (1997). "Probing the conformation of NhaA, a Na+/H+ antiporter from Escherichia coli, with trypsin." Biochemistry 36(47): 14572-6.

Rudnicki, M. A., P. N. Schnegelsberg, R. H. Stead, T. Braun, H. H. Arnold and R. Jaenisch (1993). "MyoD or Myf-5 is required for the formation of skeletal muscle." Cell 75(7): 1351-9.

Sander, M., A. Neubuser, J. Kalamaras, H. C. Ee, G. R. Martin and M. S. German (1997). "Genetic analysis reveals that PAX6 is required for normal transcription of pancreatic hormone genes and islet development." Genes Dev 11(13): 1662-73.

Sanyanusin, P., L. A. Schimmenti, L. A. McNoe, T. A. Ward, M. E. Pierpont, M. J. Sullivan, W. B. Dobyns and M. R. Eccles (1995). "Mutation of the PAX2 gene in a family with optic nerve colobomas, renal anomalies and vesicoureteral reflux." Nat Genet 9(4): 358-64.

Schafer, B. W., T. Czerny, M. Bernasconi, M. Genini and M. Busslinger (1994). "Molecular cloning and characterization of a human PAX-7 cDNA expressed in normal and neoplastic myocytes." Nucleic Acids Res 22(22): 4574-82.

Schedl, A., A. Ross, M. Lee, D. Engelkamp, P. Rashbass, V. van Heyningen and N. D. Hastie (1996). "Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities." Cell 86(1): 71-82.

Scheidler, S., W. J. Fredericks, F. J. Rauscher, 3rd, F. G. Barr and P. K. Vogt (1996). "The hybrid PAX3-FKHR fusion protein of alveolar rhabdomyosarcoma transforms fibroblasts in culture." Proc Natl Acad Sci U S A 93(18): 9805-9.

Schimmenti, L. A., H. E. Cunliffe, L. A. McNoe, T. A. Ward, M. C. French, H. H. Shim, Y. H. Zhang, W. Proesmans, A. Leys, K. A. Byerly, S. R. Braddock, M. Masuno, K. Imaizumi, K. Devriendt and M. R. Eccles (1997). "Further delineation of renal-coloboma syndrome in patients with extreme variability of phenotype and identical PAX2 mutations." Am J Hum Genet 60(4): 869-78.

Schwarz, M., F. Cecconi, G. Bernier, N. Andrejewski, B. Kammandel, M. Wagner and P. Gruss (2000). "Spatial specification of mammalian eye territories by reciprocal transcriptional repression of Pax2 and Pax6." Development 127(20): 4325-34.

Seale, P., L. A. Sabourin, A. Girgis-Gabardo, A. Mansouri, P. Gruss and M. A. Rudnicki (2000). "Pax7 is required for the specification of myogenic satellite cells." Cell 102(6): 777-86.

Sen, A. C. and B. Chakrabarti (1990). "Proximity of sulfhydryl groups in lens proteins. Excimer fluorescence of pyrene-labeled crystallins." J Biol Chem 265(24): 14277-84.

Serbedzija, G. N. and A. P. McMahon (1997). "Analysis of neural crest cell migration in Splotch mice using a neural crest-specific LacZ reporter." Dev Biol 185(2): 139-47.

Shapiro, D. N., J. E. Sublett, B. Li, J. R. Downing and C. W. Naeve (1993). "Fusion of PAX3 to a member of the forkhead family of transcription factors in human alveolar rhabdomyosarcoma." Cancer Res 53(21): 5108-12.

Sharrocks, A. D., A. L. Brown, Y. Ling and P. R. Yates (1997). "The ETS-domain transcription factor family." Int J Biochem Cell Biol 29(12): 1371-87.

Sheng, G., E. Harris, C. Bertuccioli and C. Desplan (1997). "Modular organization of Pax/homeodomain proteins in transcriptional regulation." Biol Chem 378(8): 863-72.

Sheng, G., E. Thouvenot, D. Schmucker, D. S. Wilson and C. Desplan (1997). "Direct regulation of rhodopsin 1 by Pax-6/eyeless in *Drosophila*: evidence for a conserved function in photoreceptors." Genes Dev 11(9): 1122-31.

Singh, S., C. M. Stellrecht, H. K. Tang and G. F. Saunders (2000). "Modulation of PAX6 homeodomain function by the paired domain." J Biol Chem 275(23): 17306-13.

Smith, S. B., H. C. Ee, J. R. Conners and M. S. German (1999). "Paired-homeodomain transcription factor PAX4 acts as a transcriptional repressor in early pancreatic development." Mol Cell Biol 19(12): 8272-80.

Song, D. L., G. Chalepakis, P. Gruss and A. L. Joyner (1996). "Two Pax-binding sites are required for early embryonic brain expression of an Engrailed-2 transgene." Development 122(2): 627-35.

Sonnenberg, E., D. Meyer, K. M. Weidner and C. Birchmeier (1993). "Scatter factor/hepatocyte growth factor and its receptor, the c-met tyrosine kinase, can mediate a signal exchange between mesenchyme and epithelia during mouse development." J Cell Biol 123(1): 223-35.

Sosa-Pineda, B., K. Chowdhury, M. Torres, G. Oliver and P. Gruss (1997). "The Pax4 gene is essential for differentiation of insulin-producing beta cells in the mammalian pancreas." Nature 386(6623): 399-402.

Southard-Smith, E. M., L. Kos and W. J. Pavan (1998). "Sox10 mutation disrupts neural crest development in Dom Hirschsprung mouse model." Nat Genet 18(1): 60-4.

Spector, M. S., A. Raff, H. DeSilva, K. Lee and M. A. Osley (1997). "Hir1p and Hir2p function as transcriptional corepressors to regulate histone gene transcription in the Saccharomyces cerevisiae cell cycle." Mol Cell Biol 17(2): 545-52.

Spitz, F., J. Demignon, A. Porteu, A. Kahn, J. P. Concordet, D. Daegelen and P. Maire (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(24): 14220-5.

Sprules, T., N. Green, M. Featherstone and K. Gehring (2000). "Conformational changes in the PBX homeodomain and C-terminal extension upon binding DNA and HOX-derived YPWM peptides." Biochemistry 39(32): 9943-50.

Sprules, T., N. Green, M. Featherstone and K. Gehring (2003). "Lock and key binding of the HOX YPWM peptide to the PBX homeodomain." J Biol Chem 278(2): 1053-8.

St-Onge, L., B. Sosa-Pineda, K. Chowdhury, A. Mansouri and P. Gruss (1997). "Pax6 is required for differentiation of glucagon-producing alpha-cells in mouse pancreas." Nature 387(6631): 406-9.

Stapleton, P., A. Weith, P. Urbanek, Z. Kozmik and M. Busslinger (1993). "Chromosomal localization of seven PAX genes and cloning of a novel family member, PAX-9." Nat Genet 3(4): 292-8.

Steingrimsson, E., N.G. Copeland, and N.A. Jenkins (2005). "Melanocyte stem cell maintenance and hair graying." Cell 121(1): 9-12.

Steingrimsson, E., K. J. Moore, M. L. Lamoreux, A. R. Ferre-D'Amare, S. K. Burley, D. C. Zimring, L. C. Skow, C. A. Hodgkinson, H. Arnheiter, N. G. Copeland and et al. (1994). "Molecular basis of mouse microphthalmia (mi) mutations helps explain their developmental and phenotypic consequences." Nat Genet 8(3): 256-63.

Stifani, S., C. M. Blaumueller, N. J. Redhead, R. E. Hill and S. Artavanis-Tsakonas (1992). "Human homologs of a *Drosophila* Enhancer of split gene product define a novel family of nuclear proteins." Nat Genet 2(4): 343.

Stockton, D. W., P. Das, M. Goldenberg, R. N. D'Souza and P. I. Patel (2000). "Mutation of PAX9 is associated with oligodontia." Nat Genet 24(1): 18-9.

Stout, J. G., Q. Zhou, T. Wiedmer and P. J. Sims (1998). "Change in conformation of plasma membrane phospholipid scramblase induced by occupancy of its Ca2+ binding site." Biochemistry 37(42): 14860-6.

Stoykova, A., R. Fritsch, C. Walther and P. Gruss (1996). "Forebrain patterning defects in Small eye mutant mice." Development 122(11): 3453-65.

Stoykova, A. and P. Gruss (1994). "Roles of Pax-genes in developing and adult brain as suggested by expression patterns." J Neurosci 14(3 Pt 2): 1395-412.

Stuart, E. T. and P. Gruss (1995). "PAX genes: what's new in developmental biology and cancer?" Hum Mol Genet 4 Spec No: 1717-20.

Stuart, E. T., R. Haffner, M. Oren and P. Gruss (1995). "Loss of p53 function through PAX-mediated transcriptional repression." Embo J 14(22): 5638-45.

Stuart, E. T., C. Kioussi, A. Aguzzi and P. Gruss (1995). "PAX5 expression correlates with increasing malignancy in human astrocytomas." Clin Cancer Res 1(2): 207-14.

Stuart, E. T., C. Kioussi and P. Gruss (1994). "Mammalian Pax genes." Annu Rev Genet 28: 219-36.

Sublett, J. E., I. S. Jeon and D. N. Shapiro (1995). "The alveolar rhabdomyosarcoma PAX3/FKHR fusion protein is a transcriptional activator." Oncogene 11(3): 545-52.

Tachibana, M., K. Takeda, Y. Nobukuni, K. Urabe, J. E. Long, K. A. Meyers, S. A. Aaronson and T. Miki (1996). "Ectopic expression of MITF, a gene for Waardenburg syndrome type 2, converts fibroblasts to cells with melanocyte characteristics." Nat Genet 14(1): 50-4.

Tagge, E. P., P. Hanson, G. G. Re, H. B. Othersen, Jr., C. D. Smith and A. J. Garvin (1994). "Paired box gene expression in Wilms' tumor." J Pediatr Surg 29(2): 134-41.

Tajbakhsh, S., D. Rocancourt, G. Cossu and M. Buckingham (1997). "Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD." Cell 89(1): 127-38.

Tassabehji, M., V. E. Newton, K. Leverton, K. Turnbull, E. Seemanova, J. Kunze, K. Sperling, T. Strachan and A. P. Read (1994). "PAX3 gene structure and mutations: close analogies between Waardenburg syndrome and the Splotch mouse." Hum Mol Genet 3(7): 1069-74.

Tassabehji, M., V. E. Newton and A. P. Read (1994). "Waardenburg syndrome type 2 caused by mutations in the human microphthalmia (MITF) gene." Nat Genet 8(3): 251-5.
Taulli, R., C. Scuoppo, F. Bersani, P. Accornero, P. E. Forni, S. Miretti, A. Grinza, P. Allegra, M. Schmitt-Ney, T. Crepaldi and C. Ponzetto (2006). "Validation of met as a therapeutic target in alveolar and embryonal rhabdomyosarcoma." Cancer Res 66(9): 4742-9.

Tell, G., A. Scaloni, L. Pellizzari, S. Formisano, C. Pucillo and G. Damante (1998). "Redox potential controls the structure and DNA binding activity of the paired domain." J Biol Chem 273(39): 25062-72.

Timmons, P. M., J. Wallin, P. W. Rigby and R. Balling (1994). "Expression and function of Pax 1 during development of the pectoral girdle." Development 120(10): 2773-85.

Ton, C. C., H. Hirvonen, H. Miwa, M. M. Weil, P. Monaghan, T. Jordan, V. van Heyningen, N. D. Hastie, H. Meijers-Heijboer, M. Drechsler and et al. (1991). "Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region." Cell 67(6): 1059-74.

Torres, M., E. Gomez-Pardo, G. R. Dressler and P. Gruss (1995). "Pax-2 controls multiple steps of urogenital development." Development 121(12): 4057-65.

Torres, M., E. Gomez-Pardo and P. Gruss (1996). "Pax2 contributes to inner ear patterning and optic nerve trajectory." Development 122(11): 3381-91.

Treisman, J., P. Gonczy, M. Vashishtha, E. Harris and C. Desplan (1989). "A single amino acid can determine the DNA binding specificity of homeodomain proteins." Cell 59(3): 553-62.

Treisman, J., E. Harris and C. Desplan (1991). "The paired box encodes a second DNA-binding domain in the paired homeo domain protein." Genes Dev 5(4): 594-604.

Tremblay, P., M. Kessel and P. Gruss (1995). "A transgenic neuroanatomical marker identifies cranial neural crest deficiencies associated with the Pax3 mutant Splotch." Dev Biol 171(2): 317-29.

Turque, N., S. Plaza, F. Radvanyi, C. Carriere and S. Saule (1994). "Pax-QNR/Pax-6, a paired box- and homeobox-containing gene expressed in neurons, is also expressed in pancreatic endocrine cells." Mol Endocrinol 8(7): 929-38.

Um, M., C. Li and J. L. Manley (1995). "The transcriptional repressor even-skipped interacts directly with TATA-binding protein." Mol Cell Biol 15(9): 5007-16.

Underhill, D. A. and P. Gros (1997). "The paired-domain regulates DNA binding by the homeodomain within the intact Pax-3 protein." J Biol Chem 272(22): 14175-82.

Underhill, D. A., K. J. Vogan and P. Gros (1995). "Analysis of the mouse Splotch-delayed mutation indicates that the Pax-3 paired domain can influence homeodomain DNA-binding activity." Proc Natl Acad Sci U S A 92(9): 3692-6.

Urbanek, P., Z. Q. Wang, I. Fetka, E. F. Wagner and M. Busslinger (1994). "Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP." Cell 79(5): 901-12.

Verastegui, C., K. Bille, J. P. Ortonne and R. Ballotti (2000). "Regulation of the microphthalmia-associated transcription factor gene by the Waardenburg syndrome type 4 gene, SOX10." J Biol Chem 275(40): 30757-60.

Vogan, K. J., D. J. Epstein, D. G. Trasler and P. Gros (1993). "The splotch-delayed (Spd) mouse mutant carries a point mutation within the paired box of the Pax-3 gene." Genomics 17(2): 364-9.

Vogan, K. J. and P. Gros (1997). "The C-terminal subdomain makes an important contribution to the DNA binding activity of the Pax-3 paired domain." J Biol Chem 272(45): 28289-95.

Vogan, K. J., D. A. Underhill and P. Gros (1996). "An alternative splicing event in the Pax-3 paired domain identifies the linker region as a key determinant of paired domain DNA-binding activity." Mol Cell Biol 16(12): 6677-86.

Waardenburg, P. J. (1951). "A new syndrome combining developmental anomalies of the eyelids, eyebrows and nose root with pigmentary defects of the iris and head hair and with congenital deafness." Am J Hum Genet 3(3): 195-253.

Wallin, J., H. Eibel, A. Neubuser, J. Wilting, H. Koseki and R. Balling (1996). "Pax1 is expressed during development of the thymus epithelium and is required for normal T-cell maturation." Development 122(1): 23-30.

Wallin, J., Y. Mizutani, K. Imai, N. Miyashita, K. Moriwaki, M. Taniguchi, H. Koseki and R. Balling (1993). "A new Pax gene, Pax-9, maps to mouse chromosome 12." Mamm Genome 4(7): 354-8.

Wallin, J. J., E. R. Gackstetter and M. E. Koshland (1998). "Dependence of BSAP repressor and activator functions on BSAP concentration." Science 279(5358): 1961-4.

Walmsley, A. R., T. Zhou, M. I. Borges-Walmsley and B. P. Rosen (1999). "The ATPase mechanism of ArsA, the catalytic subunit of the arsenite pump." J Biol Chem 274(23): 16153-61.

Walther, C. and P. Gruss (1991). "Pax-6, a murine paired box gene, is expressed in the developing CNS." Development 113(4): 1435-49.

Walther, C., J. L. Guenet, D. Simon, U. Deutsch, B. Jostes, M. D. Goulding, D. Plachov, R. Balling and P. Gruss (1991). "Pax: a murine multigene family of paired box-containing genes." Genomics 11(2): 424-34.

Wang, W., M. Slevin, S. Kumar and P. Kumar (2005). "The cooperative transforming effects of PAX3-FKHR and IGF-II on mouse myoblasts." Int J Oncol 27(4): 1087-96.

Wang, Q., J. Voss, W. L. Hubbell and H. R. Kaback (1998). "Proximity of helices VIII (Ala273) and IX (Met299) in the lactose permease of Escherichia coli." Biochemistry 37(14): 4910-5.

Watanabe, A., K. Takeda, B. Ploplis and M. Tachibana (1998). "Epistatic relationship between Waardenburg syndrome genes MITF and PAX3." Nat Genet 18(3): 283-6.

Weber, J. and A. E. Senior (1997). "Catalytic mechanism of F1-ATPase." Biochim Biophys Acta 1319(1): 19-58.

Weber, J., S. Wilke-Mounts, R. S. Lee, E. Grell and A. E. Senior (1993). "Specific placement of tryptophan in the catalytic sites of Escherichia coli F1-ATPase provides a direct probe of nucleotide binding: maximal ATP hydrolysis occurs with three sites occupied." J Biol Chem 268(27): 20126-33.

Weinberg, R. A. (1995). "The retinoblastoma protein and cell cycle control." Cell 81(3): 323-30.

Wheat, W., D. Fitzsimmons, H. Lennox, S. R. Krautkramer, L. N. Gentile, L. P. McIntosh and J. Hagman (1999). "The highly conserved beta-hairpin of the paired DNA-binding domain is required for assembly of Pax-Ets ternary complexes." Mol Cell Biol 19(3): 2231-41.

Wiggan, O., A. Taniguchi-Sidle and P. A. Hamel (1998). "Interaction of the pRB-family proteins with factors containing paired-like homeodomains." Oncogene 16(2): 227-36.

Williams, B. A. and C. P. Ordahl (1994). "Pax-3 expression in segmental mesoderm marks early stages in myogenic cell specification." Development 120(4): 785-96.

Wilm, B., E. Dahl, H. Peters, R. Balling and K. Imai (1998). "Targeted disruption of Pax1 defines its null phenotype and proves haploinsufficiency." Proc Natl Acad Sci U S A 95(15): 8692-7.

Wilson, D., G. Sheng, T. Lecuit, N. Dostatni and C. Desplan (1993). "Cooperative dimerization of paired class homeo domains on DNA." Genes Dev 7(11): 2120-34.

Wilson, D. S., B. Guenther, C. Desplan and J. Kuriyan (1995). "High resolution crystal structure of a paired (Pax) class cooperative homeodomain dimer on DNA." Cell 82(5): 709-19.

Wright, M. E. (1947). "Undulated: a new genetic factor in Mus musculus affecting the spine and tail." Heredity 1: 137-41.

Xia, S.J. and F.G. Barr (2004). "Analysis of the transforming and growth suppressive activities of the PAX3-FKHR oncoprotein." Oncogene 23(41): 6864-71.

Xu, H. E., M. A. Rould, W. Xu, J. A. Epstein, R. L. Maas and C. O. Pabo (1999). "Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding." Genes Dev 13(10): 1263-75.

Xu, W., M. A. Rould, S. Jun, C. Desplan and C. O. Pabo (1995). "Crystal structure of a paired domain-DNA complex at 2.5 A resolution reveals structural basis for Pax developmental mutations." Cell 80(4): 639-50.

Yang, X. M., K. Vogan, P. Gros and M. Park (1996). "Expression of the met receptor tyrosine kinase in muscle progenitor cells in somites and limbs is absent in Splotch mice." Development 122(7): 2163-71.

Yasumoto, K., K. Yokoyama, K. Takahashi, Y. Tomita and S. Shibahara (1997). "Functional analysis of microphthalmia-associated transcription factor in pigment cell-specific transcription of the human tyrosinase family genes." J Biol Chem 272(1): 503-9.

Zannini, M., H. Francis-Lang, D. Plachov and R. Di Lauro (1992). "Pax-8, a paired domain-containing protein, binds to a sequence overlapping the recognition site of a homeodomain and activates transcription from two thyroid-specific promoters." Mol Cell Biol 12(9): 4230-41.

Zhang, F., N. Kartner and G. L. Lukacs (1998). "Limited proteolysis as a probe for arrested conformational maturation of delta F508 CFTR." Nat Struct Biol 5(3): 180-3.

Zhang, H., K. M. Catron and C. Abate-Shen (1996). "A role for the Msx-1 homeodomain in transcriptional regulation: residues in the N-terminal arm mediate TATA binding protein interaction and transcriptional repression." Proc Natl Acad Sci U S A 93(5): 1764-9.

Zhou, T. and B. P. Rosen (1997). "Tryptophan fluorescence reports nucleotide-induced conformational changes in a domain of the ArsA ATPase." J Biol Chem 272(32): 19731-7.

Zlotogora, J., I. Lerer, S. Bar-David, Z. Ergaz and D. Abeliovich (1995). "Homozygosity for Waardenburg syndrome." Am J Hum Genet 56(5): 1173-8.

Zwilling, S., A. Annweiler and T. Wirth (1994). "The POU domains of the Oct1 and Oct2 transcription factors mediate specific interaction with TBP." Nucleic Acids Res 22(9): 1655-62.

Appendix