The Orchestration of Mammalian Tissue Morphogenesis through a Series of Coherent Feed-forward Loops

Received for publication, May 24, 2011, and in revised form, October 6, 2011. Published, JBC Papers in Press, October 13, 2011. DOI 10.1074/jbc.M111.264580

Qing Xie and Ales Cvekl

From the Departments of Genetics and Ophthalmology and Visual Sciences, Albert Einstein College of Medicine, Bronx, New York 10461

**Background:** Gene regulatory networks control tissue morphogenesis.

**Results:** Pax6 directly regulates expression of c-Maf in embryonic lens.

**Conclusion:** A novel distal enhancer of c-Maf is required for its expression in the lens through a pair of Pax6-binding sites.

**Significance:** The molecular foundation of lens morphogenesis is on the basis of multiple coherent feed-forward loops comprised of Pax6, c-Maf, and individual crystallins.

Tissue morphogenesis requires intricate temporal and spatial control of gene expression that is executed through specific gene regulatory networks (GRNs). GRNs are comprised from individual subcircuits of different levels of complexity. An important question is to elucidate the mutual relationship between those genes encoding DNA-binding factors that trigger the subcircuit with those that play major “later” roles during terminal differentiation via expression of specific genes that constitute the phenotype of individual tissues. The ocular lens is a classical model system to study tissue morphogenesis. Pax6 is essential for both lens placode formation and subsequent stages of lens morphogenesis, whereas c-Maf controls terminal differentiation of lens fibers, including regulation of crystallins, key lens structural proteins required for its transparency and refraction. Here, we show that Pax6 directly regulates c-Maf expression during lens development. A 1.3-kb c-Maf promoter with a 1.6-kb upstream enhancer (CR1) recapitulated the endogenous c-Maf expression pattern in lens and retinal pigmented epithelium. ChIP assays revealed binding of Pax6 and c-Maf to multiple regions of the c-Maf locus in lens chromatin. To predict functional Pax6-binding sites, nine novel variants of Pax6 DNA-binding motifs were identified and characterized. Two of these motifs predicted a pair of Pax6-binding sites in the CR1. Mutagenesis of these Pax6-binding sites inactivated transgenic expression in the lens but not in retinal pigmented epithelium. These data establish a novel regulatory role for Pax6 during lens development, link together the Pax6/c-Maf/crystallin regulatory network, and suggest a novel type of GRN subcircuit that controls a major part of embryonic lens development.

Generation of complex tissues and organs is the major question of developmental biology (1). It holds the promise to tissue replacement therapy and to understand and correct specific developmental abnormalities. Until recently it was thought that a large number of regulatory factors are required to coordinate embryonic development of a specific tissue through a myriad of complex pathways. It has been shown recently that individual tissues that form the mammalian eye, i.e. the optic cup and ocular lens, can be generated using self-organizing cultures of differentiating embryonic stem cells using three-dimensional matrices and can be governed by a defined set of growth factors (2, 3). However, the molecular mechanisms underlying the formation of distinct tissues are still poorly understood. Tissue formation and organogenesis require intricate temporal and spatial control of gene expression. Embryonic development unfolds information stored in discrete regions of the genome. Specific gene regulatory networks (GRNs) evolved during more than 600 million years of animal evolution (4, 5). These GRNs are comprised of a repertoire of subcircuits or modules in which specific DNA-binding transcription factors control expression of lineage- and differentiation-specific genes (6–9). An important question is to elucidate the mutual relationship between those genes encoding DNA-binding factors that control early stages of cell type specification with those that play major “later” roles during terminal differentiation (4).

Because of its unique cell morphology, the lens is an advantageous tissue to study the molecular mechanisms of temporal and spatial control of gene expression and GRNs (10–12). Although many transcription factors have been identified to be important for lens development, sequential events that control lens development in vivo have not been elucidated. It has been shown earlier that lens progenitor cell formation requires the...
function of at least two essential genes, Pax6 (13) and Six3 (14), both of which are expressed in the head surface ectoderm, the preplacodal region (15), prior to the formation of the lens placode. Genetic studies have shown that Pax6 is an essential gene for multiple stages of lens formation (13, 16–20). In contrast, lens fiber cell differentiation is controlled by a larger group of transcription factors, including c-Maf, Foxe3, Gata-3, Hsf4, Pax6, Pitx3, Prox1, pRb, and Sox1 (21). Loss of function of c-Maf causes the most dramatic lens differentiation defects, including block of cellular elongation and lack of induction of crystallin gene expression in differentiating lens fiber cells (22–24). In the lens, the onset of c-Maf expression is in cells forming the lens placode just prior to its invagination. Transcriptional regulation of c-Maf in vivo during embryonic development is poorly understood (25), although two recent studies identified the direct regulation of c-Maf by BATF in follicular helper T cells (26) and by IL-2/Stat5 in CD4 T cells (27). In addition to the lens, c-Maf is also expressed in bones, hematopoietic cells, the kidney, the pancreas, and the skin, among other tissues. The major roles of c-Maf in these tissues include T helper 2 cell (28) and chondrocyte (29) differentiation. Deficiency of c-Maf causes abnormal macrophage development and is critical for erythroblast island function during fetal erythropoiesis (30). In multiple myeloma, up-regulation of MAF influences tumor metastasis (31). Thus, studies of the transcriptional control of c-Maf are important for a better understanding of the multiple specific roles of this gene in normal development and in cancer.

The paired box (Pax) gene family, comprised of nine genes in mammals, encodes proteins that act as key regulators of organogenesis (32). Mutations in Pax genes cause a plethora of human diseases (33) and mouse developmental abnormalities (32, 34). The paired box encodes an evolutionary conserved 128-amino acid DNA-binding domain, the paired domain (PD), comprised of two interactive subdomains, PAI and RED (35). Alternate splicing can generate multiple splice variants of the PD (e.g. PD5a, (36)) and disrupt its DNA-binding properties (37). Pax6 is a lineage-specific DNA-binding transcription factor that regulates the development of the central nervous system, endocrine pancreas, eye, and olfactory system. Pax6, together with Pax3, Pax4, and Pax7, contains a third DNA-binding domain, a paired-type homeodomain (HD). Previous studies have defined “consensus” binding sites for Pax6 PD (P6CON) (38), PD5a (5aCON) (39), and HD (22). Subsequent crystallographic studies have shown that both PAI and RED are comprised of helix-turn-helix (HTH) folds. However, only the PAI subdomain showed direct contacts with DNA (40). In addition, protein-DNA contacts were also established between the N-terminal β-sheets, and a linker region between PAI and RED subdomains (40). Thus, in contrast to the majority of sequence-specific DNA binding proteins that use a single DNA-binding domain that typically docks with 6–10 bp of DNA, Pax6 can use β-sheets, PAI, linker, RED, and HD, and their various combinations to bind DNA. Although some important direct targets of Pax6 genes have been identified (41), a large body of work remains to be accomplished (42–44). Some transcription factors, especially those with multiple DNA-binding subdomains (e.g. the Pax family, see above), take advantage to bind DNA by a range of combinatorial mechanisms with an elusive consensus binding site as shown for BSAP/Pax5 (45–46). In addition, many transcription factors recognize suboptimal binding sites (47–48) and/or preferentially partner with other factors (49). Thus, better understanding of Pax6 DNA-binding mechanisms is necessary for the identification and functional validation of the GRN downstream of Pax6.

To understand the molecular basis of tissue morphogenesis and its underlying GRNs, we hypothesized that Pax6 could orchestrate the lens differentiation through direct control of c-Maf and the sequential expression of crystallins that make up the bulk structure of the eye lens. To test this hypothesis, we first identified the lens/RPE-specific enhancer of c-Maf and showed an interaction of Pax6 proteins with the c-Maf gene in lens chromatin. To predict functional Pax6-binding sites, we identified and characterized novel Pax6- and Pax6(5a)-binding site variants that aid identification of Pax6-regulated genes in other developmental systems. These findings are summarized in a model comprised of multiple coherent feed-forward loops that regulate crystallin gene expression combined with two Pax6 and c-Maf autoregulatory loops as the molecular foundation of tissue morphogenesis.

**EXPERIMENTAL PROCEDURES**

**Antibodies**—For immunofluorescence staining, primary antibodies used were rabbit anti-GFP (1:1000, Invitrogen) and rabbit anti-c-Maf (1:2000, Bethyl). Secondary antibodies were goat anti-rabbit IgG conjugated with Alexa Fluor 488 or 568 (1:500, Invitrogen). For ChIP, rabbit anti-Pax6 (Millipore) or anti-c-Maf (Bethyl) antibodies were used.

**Transgenic Mouse Production and Analysis of EGFP Expression**—Four reporter plasmids, including c-Maf promoter (-493/+866), CR1 (-4591/-2967), CR2 (-2966/-494), and CR3 (+2582/+4400), were generated in peGFP-1 (Clontech). The DNA sequences of CR1 with mutations (CR1-Mut) and deletion (CR1-De) of the Pax6 binding sites were synthesized (GenScript) and subcloned into peGFP-1 with the c-Maf promoter. The transgenic mice were generated by pronuclear injection of fertilized eggs at the Albert Einstein College of Medicine transgenic core facility. Pax6 heterozygous mice (Pax6+/-knock2) were described earlier (50). EGFP expression was detected by immunofluorescence using a Leica SP2 confocal microscope.

**Immunofluorescence Staining**—Four percent paraformaldehyde-fixed embryos were cryoprotected with 30% sucrose in PBS and embedded in optimal cutting temperature tissue-freezing medium (Triangle Biomedical Sciences) for cryosectioning. Eight-micrometer transverse sections were collected, washed with PBS, and incubated for 30 min with Image iFX™ FX signal enhancer (Invitrogen). Slides were then washed with PBS and incubated overnight at 4 °C with the primary antibody, rabbit anti-GFP (1:1000, Invitrogen) and rabbit anti-c-Maf (1:2000, Bethyl) and diluted in PBS containing 1% BSA and 0.05% Triton X-100. After washing with PBS, slides were incubated for 45 min with secondary antibodies, goat anti-rabbit IgG conjugated with Alexa Fluor 488 or 568 (1:500, Invitrogen), and with DAPI (1:50,000, Invitrogen). Slides were then washed with PBS and mounted with Vectashield (Vector Labs). Immunofluorescence was visualized by using a Leica SP2 confocal microscope in the
AECOM core facility. The percentage of c-Maf-expressing cells = number of cells with c-Maf staining/number of cells with nuclear DAPI staining (only counting the cells posterior of the transitional zone) ¥ 100%. As EGFP protein was located in the elongated cytoplasmic space, the areas of the EGFP-expressing cells were counted with ImageJ. The percentage of EGFP-expressing cells = area of EGFP-expressing cells/area of the lens (posterior of the transitional zone).

Quantitative ChIP (qChIP)—Formaldehyde cross-linked chromatin was obtained from groups of 400 pooled newborn lenses. The sheared chromatin (average size, 600 bp) was generated by sonication. Aliquots of 40 chromatin-prepared lenses were incubated with 5 μg of anti-Pax6 or 5 μg of anti-c-Maf antibodies bound to 20 μl protein G-coated magnetic beads were also used (Invitrogen). The immunoprecipitates were washed three times and resuspended in a buffer containing 10 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 25 mM EDTA supplemented with 0.1 mg/ml RNaseA and 0.2 mg/ml protease K. After the 2-h incubation at 55°C, the cross-links were reversed by overnight incubation at 65°C. Genomic DNA was eluted into 250 μl of water using the QIAquick spin gel purification kit (Qiagen).

The amounts of each specific DNA fragment in immunoprecipitates were determined by quantitative PCR reactions using a standard curve generated for each primer set with 0.04, 0.2, and 1% input DNA samples. Using a standard curve, we transformed cycle threshold (Ct) values into DNA copy numbers. The copy number of a specific DNA fragment in each assay was compared with the copy number of that fragment before immunoprecipitation (“input DNA”). The control antibody (rabbit normal IgG from Calbiochem) was included for each set of the quantitative PCR experiments. The enrichments obtained with IgG were subtracted from the corresponding values obtained with anti-Pax6 and c-Maf antibodies. As we found different enrichments at specific regions of the c-Maf locus, it was essential to determine a critical value to distinguish real binding signals from background noise. Statistical analysis was conducted using the R Software (Version 2.13.1). Analysis of variance was performed using a Student’s t test for the background group (seven sites). Finally, we performed a Student’s t test for the background group (seven sites). The results showed that any site having a signal higher than 0.112 reflects a specific c-Maf binding region.

Quantitative RT-PCR—Relative expression levels of transcripts encoding Pax6, c-Maf, EGFP, and 14 crystallin genes in wild-type and Pax6 heterozygous lenses and Pax6 homozygous lens placodes were determined using quantitative RT-PCR. Total RNA was isolated from microdissected E15.5 and 1-day-old (P1) lenses (WT, Pax6+/−, CR1, and CR1xPax6+−) using the RNeasy micro kit (Invitrogen). Subsequently, cDNA was generated with oligo(dT)20 primers and Superscript III reverse transcriptase (Invitrogen) according to the manufacturer’s protocol. cDNAs from laser capture-microdissected E9.5 lens placodes (WT and Pax6−−) were generously provided by Dr. David C. Beebe (Washington University in St. Louis, MO). Primer design and quantitative PCR reactions are described above. For data normalization, expression of three reference genes, B2M, CCNI, and HPRT, was probed. As no significant changes of expression of B2M, CCNI, and HPRT were found, the final data were expressed relative to the expression level of B2M. For statistical evaluation of the results, p values were calculated from paired Student’s t tests.

Reporter Plasmids and Transient Transfections—The reporter plasmids contained c-Maf promoter and CR1 regions in pGL3 (Promega) as described above. The reporter plasmids driven by six copies of variant Pax6-binding sites were synthesized by GenScript and cloned into a modified vector pGL3 containing the E4 TATA-minimal promoter as described earlier (51). Transient transfections were conducted by Lipofectamine 2000 (Invitrogen) in P19 embryonic carcinoma cells that do not express endogenous Pax6 proteins (51). For studies of the c-Maf promoter, P19 cells were cotransfected with 0.6 μg of firefly luciferase reporter plasmids, 50 ng of each expression plasmid encoding empty vector (pKW10), Pax6 and c-Maf and 20 ng of Renilla-TK using 24-well microplates. The dual luciferase reporter assay system (Promega) was used according to the manufacturer’s instructions. For statistical evaluation of transfection results, p values were calculated as described above.

SELEX—A 95-nt oligonucleotide containing 32 random nucleotides in the middle and fixed flanking sequences was synthesized (5′-GTCAACGTCAGACGGAAATTCGCGGCCGC (N)12CTCGAGGGATCCGCTGCTGCTGGAGG (N)12TCGAGGGATCCGCTGCTGCTGGAGG) using the method described above. The length (32 nt) of the randomer was based on the following rationale: typical binding sites for a transcription factor with only one DNA-binding domain are 6–10 bp, but Pax6 contains multiple functional DNA-binding domains/subdomains. The combined activities of these DNA-binding units may recognize longer sites. A previous study showed that the Pax6 PD domain alone recognizes 20 bp of DNA (38, 40). The Pax6 HD (22) could further extend the Pax6-binding (35). Finally, longer randomers could provide a more flexible plat-
form to promote different types of docking of Pax6 proteins on DNA.

A double-strand DNA random pool was generated by annealing primer A (5′- CGATAAGCTGACGCACTGACG- GATCCCT-3′) and extending it with the Klenow enzyme (10 μl each of 10 mM dATP, dTTP, dCTP, dGTP; 25 μl of 10X buffer 2, 10 μl of 500 μM primer A, 100 μl of 5 μM oligo 95, 5 μl of New England Biolabs Klenow enzyme, 100 μl of H2O) at 25 °C for 3 h. The mixture was then loaded on a 4.5% UltraPure™ agarose 1000 gel (Invitrogen). The specific 96-bp band was cut and gel-purified by QIAquick extraction kit (Qia- gen). One microliter of recombinant proteins with a GST tag was used as a template for PCR amplification in a 50 μl mixture containing 10 pmols of primer A (described above) and B (5′- GTCAACGTCGAGACGGAATTCGCGG-3′) and 1 μl of Taq polymerase. Amplifications were performed for 25 cycles (95 °C for 30 s, 65 °C for 30 s, 72 °C for 30 s) followed by a single cycle of 95 °C for 1 min, 65 °C for 1 min, and 72 °C for 10 min. The PCR products were purified by QIAquick PCR purification kit. One hundred nanograms of purified PCR products were used for the next round of selection. A total of eight rounds of incubation/amplification were performed. PCR products from each round were radiolabeled by [α-32P]dCTP with Klenow enzyme and used as probes for an EMSA to validate the enrichment of Pax6 binding sequences during the selection (supplemental Fig. S2).

Protein Expression and EMSA—The GST-Pax6 (PD/HD) and PD(5a)/HD expression vector were described earlier (52). The proteins were expressed in BL21-Codon Plus DE3 Escherichia coli (Stratagene) and purified by glutathione-Sepharose beads (GE Healthcare). SELEX was performed as described in detail in the supplemental materials. EMSAs were performed in the presence of 1 μg of poly(dl-dC) as a nonspecific DNA competitor (52). A full description of these procedures is provided in the supplemental materials.

RESULTS

A 1.6-kb 5′ Evolutionary Conserved Region CR1 Serves as a lens/RPE-specific Enhancer of the Mouse c-Maf Gene—To identify GRNs that control tissue morphogenesis, we used the lens as a model tissue (11), as its structure and function depends on a coordinated expression of 14 crystallin genes that is under the control of a small group of transcription factors, including Pax6 and c-Maf. To study transcriptional regulation of c-Maf in vivo, we performed a series of transgenic mouse studies. We initially tested whether a 1.3-kb c-Maf promoter fragment could drive expression of the EGFP reporter in transgenic mouse lenses. We generated and analyzed six independent transgenic lines. However, none of them supported EGFP expression in mouse embryos. Next, we tested three adjacent 5′ and 3′ evolutionarily conserved non-coding regions (CR1, CR2, and CR3, Fig. 1, A and B) as potential distal enhancers of c-Maf. Although neither CR2 nor CR3 tested in multiple transgenic lines supported transgenic lens expression, we found that an upstream region CR1 (approximately 1.6 kb) in combination with a 1.3-kb c-Maf promoter strongly activated expression of EGFP in lens and in RPE in three independent transgenic lines (Fig. 1D). The expression pattern of EGFP was similar to the expression of endogenous c-Maf in the lens (compare Fig. 1, C and D, see summary of c-Maf and EGFP positive cells in Fig. 1E and supplemental Fig. S1), as shown earlier (23, 24). These data indicate that CR1 functions as a lens- and RPE-preferred c-Maf enhancer.

Expression of c-Maf Is Regulated by Pax6 During Lens Development—The expression of Pax6 in the prospective lens surface ectoderm (53, 54) commences prior to the onset of c-Maf expression in the lens placode (23, 24), raising the possibility that Pax6 is genetically upstream of c-Maf in the lens lineage. To test whether c-Maf expression is affected by Pax6 haploinsufficiency, we determined the levels of c-Maf transcripts in wild-type and Pax6+/− E15.5 and 1-day-old mouse lenses (P1) by quantitative RT-PCR. The results showed that c-Maf was down-regulated in the Pax6+/− lens compared with that of the wild type (Fig. 2A). In addition, we examined c-Maf expression in Pax6−/− lens placodes. Pax6 inactivation was performed using the le-cre line (13), as this approach allowed initial formation of the lens placode. Expression of both Pax6 and c-Maf was significantly reduced in the mutant placodes (Fig. 2A). Next, we crossed the CR1/c-Maf promoter/EGFP transgenic with a Pax6+/− mouse. We found that CR1/c-Maf promoter-driven EGFP expression was both delayed and reduced in Pax6+/− background (Fig. 2, B–D) confirming that c-Maf is directly or indirectly regulated by Pax6.

To determine whether Pax6 directly interacts with c-Maf locus, we performed qChIP studies within a 12-kb region of the c-Maf locus in lens chromatin. We also tested whether c-Maf binds its own locus, as this would indicate its autoregulation. We found multiple regions occupied by Pax6 and c-Maf, including the promoter and CR1 (Fig. 3A). Transient transfection studies using the c-Maf promoter alone and in combination with CR1 showed that both Pax6 and c-Maf activated these constructs and that their combinations showed synergistic effects (Fig. 3B). A high abundance of Pax6 and c-Maf proteins in the 3′-part of the c-Maf coding region prompted us to test whether an approximately 400-bp exon region could serve as a potential transcriptional enhancer. We found that this fragment in combination with the c-Maf promoter alone did not elicit any EGFP expression in two transgenic lines derived (data not shown). Similarly, a combination of this exon fragment with the CR1/c-Maf promoter, examined in two transgenic lines, did not change the expression profile of the CR1/c-Maf lines shown in Fig. 1 and supplemental Fig. S1. Taken together, these data
led us to hypothesize that CR1 contains at least one functional Pax6-binding site. In contrast, Pax6-binding to the c-Maf pro-
moter fragment was not sufficient to activate EGFP expression
in vivo, although the transient transfection assays showed Pax6
as an activator of this 1.3-kb promoter fragment.

Identification and Functional Characterization of Nine Novel
Pax6 DNA-binding Motifs—Efforts to predict and confirm can-
didate Pax6-binding sites in CR1 of c-Maf using P6CON
sequence (38) failed, as reported earlier for many other Pax6
directly regulated genes (55–57). We hypothesized that natural
Pax6-binding sites are either recognized by mechanisms that
diverge from the DNA-binding mode established for isolated
Pax6 PD because of the additional presence of HD or that addi-
tional auxiliary DNA-binding transcription factors, partners of
Pax6, are involved. To test whether there are novel types of
Pax6-binding sites, we performed systematic evolution of
ligands by exponential enrichment (SELEX (58)) using recom-
binant Pax6 proteins that contained both PD/HD and
PD5a/HD DNA-binding domains as described in the supple-
mental data. The SELEX results (supplemental Fig. S2) pr
ompted us to sequence both “earlier” (n/H11005 4 and 5) and “late” (n/H11005 6 and 8) enriched sequences (a total of 130 sequences for Pax6 and 133 sequences for Pax6(5a) supplemental Table S1).

We reasoned that this approach could identify a wider range of
sequences than those identified in the standard “final” enriched
pool (58). Starting with the “original” Pax6 consensus sequence,
P6CON (38), reiterative computational analyses yielded four
novel sequence motifs: 1-1 (likely recognized by the Pax6

FIGURE 1. Identification and functional characterization of a novel 5’ lens-preferred enhancer in the c-Maf locus. A, evolutionarily conserved regions (CR1, 2, and 3) in the 12-kb region of the mouse c-Maf locus. Mammal Cons displays the conservation of DNA in multiple mammalian species shown in the University of California at Santa Cruz genome browser. B, a diagrammatic summary of four EGFP reporter constructs tested in transgenic mice. The individual conserved regions CR1 (1.6 kb), CR2 (2.5 kb), and CR3 (1.8 kb) were tested in combination with a 1.3 kb c-Maf promoter (+500 to +800). C, expression of c-Maf in the mouse embryonic lens detected by immunofluorescence. D, expression of EGFP driven by CR1/1.3 kb c-Maf promoter in the transgenic mouse eye (line 10). Expression of c-Maf in RPE by in situ hybridizations and using the knocked-in lacZ marker were described elsewhere (23, 24). LP, lens pit; LV, lens vesicle; 1°LF, primary lens fibers; RE, retina; RPE, retinal pigmented epithelium. Scale bars = 100 μm (C and D). Blue, DAPI; red, anti-c-Maf; green, anti-GFP. E, quantitative analysis of c-Maf- and EGFP-expressing cells in three independent transgenic lines (lines 10, 14, and 17; see C and D and supplemental Fig. S1). The percentages of c-Maf- and EGFP-positive cells were calculated as described under “Experimental Procedures.”
FIGURE 2. Temporal and spatial patterns of c-Maf expression in the lens are changed as a result of Pax6 haploinsufficiency. A, quantitative RT-PCR analysis of Pax6 and c-Maf expression in E15.5 and P1 wild-type Pax6+/−/− lenses and wild-type and E9.5 Pax6−/− lens placodes. B, EGFP expression driven by CR1/1.3kb c-Maf promoter in Pax6−/− eye. LP, lens pit; LV, lens vesicle; P1oLF, primary lens fibers; RE, retina; RPE, retinal pigmented epithelium. Scale bars = 100 μm (C and D). Blue, DAPI; green, anti-GFP. C, quantitative analysis of EGFP-positive cells generated in wild-type and Pax6−/− transgenic lenses by ImageJ. Data from three different transgenic lines were combined. D, quantitative analysis of EGFP mRNA levels of three individual transgenic lines (10, 14, and 17) in the wild type and in Pax6−/− transgenic lenses by quantitative RT-PCR.
β-turn, N-terminal HTH fold, and a linker), 2-1 (likely recognized by the Pax6 homeodomain, β-turn, and N-terminal HTH fold), 3-1 (likely recognized by the Pax6-linker and C-terminal HTH fold), and 4-1 (likely recognized by Pax6(5a) homeodomain, linker and C-terminal HTH fold) as shown in Fig. 4B and supplemental Fig. S3. The sequence of motif 1-1 is similar to P6CON (38). However, it excludes the 3′ portion of its site, which, according to crystallographic data (40), does not bind DNA. The novel motif 2-1 is comparable with an optimal binding site of Drosophila Prd, a protein structurally similar to Pax6 (35). The novel motifs 3-1 and 4-1 are most divergent from previously established “optimal” Pax6- and Pax6(5a)-binding sites, respectively. Next, we compiled 37 Pax6-binding sites described in the literature and seven novel sites that we recently identified in parallel studies of genes regulated by Pax6 in the lens and retina. In this group, we found 20 natural sites similar to motif 1-1 (Fig. 4B and supplemental Fig. S4, A and B). Six of these sites had a conserved GC-dinucleotide, the most conserved region in motif 1-1, and their alignment yielded motif 1-2. The remaining 14 sequences gave a consensus motif 1-3 (Fig. 4B and supplemental Fig. S4B). We then identified eight natural Pax6-binding sites similar to a HD-containing motif 2-1 (Fig. 4B and supplemental Fig. S4C). Using motif 3-1, we found 16 natural Pax6-binding sites that were grouped into two variants, motif 3-2 and 3-3 (Fig. 4B), depending on the presence (motif 3-2) or absence (motif 3-3) of a frequent TT-dinucleotide (positions 18 and 19, Fig. 4B, and supplemental Fig. S4, D and E).

EMSAs of both Pax6 and Pax6(5a) proteins showed a range of binding affinities with all 10 probes tested (supplemental Fig. S5) and is summarized in Fig. 4. For Pax6, the affinities were in
the descending order P6CON, approximately motif 2-2 > motif 1-3 > motifs 1-1, 2-1, and 3-3 > motifs 1-2 and 4-1 > motif 3-1. For Pax6(5a), the affinities were in the descending order motif 4-1 > motif 3-1 > motifs 1-3 and 2-1 > motifs 1-2 and 3-2 > motif 1-1 and 2-2. As expected, P6CON did not bind Pax6(5a) (39). Motifs 2-2, 3-3, and P6CON are most specific for binding of Pax6 (on the basis of the ranking of the sites described above), whereas motifs 3-1 and 4-1 are most selective for Pax6(5a) binding.

To functionally characterize these novel Pax6-binding sites, we performed a series of transient transfections using synthetic reporters (Fig. 4A) containing six copies of individual motifs in P19 embryonic carcinoma cells (39, 59). The results demonstrate a dynamic range of transcriptional activation and repression elicited by Pax6 and Pax6(5a) (Fig. 4B). Three motifs, 4-1, 2-1 and 2-2 showed higher activation than the P6CON “reference” sequence. In contrast, motif 1-2 showed transcriptional repression with both Pax6 and Pax6(5a) proteins. In addition, fold changes elicited by Pax6(5a) were smaller compared with those induced by Pax6. The highest activations in this group were found for two motifs, 3-1 and 4-1. From these data we conclude that the majority of novel Pax6-binding motifs support transcriptional activation by Pax6. Pax6(5a) functioned as...
A Pair of Adjacent Pax6-binding Sites in CR1 Is Essential for Its Enhancer Activity in the Lens—Examination of the CR1 region of the c-Maf locus showed a pair of putative Pax6-binding sites related to novel motifs 2-1 (site 2) and 2-2 (site 1) in tail-to-tail orientation separated by 8 bp (Fig. 5A). The in vitro EMSA confirmatory experiments with these individual sites and Pax6 recombinant proteins are shown in Fig. 5B. To design functionally inactive sites, we tested five specific point mutations for each site (Fig. 5B). Mutation 2 (mutation 9) of site 1 (site 2) caused significant loss of Pax6-binding activities on each site, respectively. We then generated multiple transgenic lines containing a deletion (CR1−Del) and a combination of point mutations (mut 2 and mut 9, CR1−Mut) in these two Pax6-binding sites. Analysis of EGFP expression demonstrated a significant loss of expression of the reporter gene in transgenic lenses (compare Figs. 5D and 1D and supplemental Fig. S6) in all independent lines tested. In contrast, expression of EGFP was still detected in RPE cells (Fig. 5C). We conclude that Pax6 directly regulates c-Maf expression via tandem novel binding sites in CR1. Our data also show that relative abundance of Pax6, which is higher in the promoter region compared with CR1, measured via qChIP, does not correlate with functional significance of Pax6-binding to these regions.

Pax6 Is Essential for the Entire Crystallin Gene Expression Profile in the Mouse Embryonic Lens—Pax6 has been reported to directly regulate specific individual crystallin genes (50, 51, 60). To determine the global impact of Pax6 haploinsufficiency on crystallin gene expression, we performed a series of quantitative RT-PCR experiments to detect mRNA levels of all 14 α-, β- and γ-crystallin genes in the wild-type and Pax6+/− E15.5 lens. The expression of all 14 crystallin genes was significantly down-regulated in each Pax6+/− lens tested (Fig. 6). These data further support the central roles of Pax6-dependent gene regulation on the crystallin lens transcriptome.

DISCUSSION

"Core" Lens-specific Gene Regulatory Network—GRNs are composed of multiple subcircuits and provide system-level explanations of cellular differentiation, development, and evolution of multicellular organisms. Individual subcircuits have evolved to serve multitudes of integrated functions (4, 5, 61, 62). The current data showing Pax6 regulation of c-Maf, coupled with earlier studies of Pax6 autoregulation (63) and regulation of crystallin gene expression by c-Maf and Pax6 (50, 51, 60, 64), are summarized in the core lens-specific GRN subcircuit (Fig. 7A).

Pax6 expression is initiated in a broad area of head surface ectoderm prior to the lens progenitor cell formation (53, 54). Expression of Pax6 is maintained in the lens placode (E10), lens vesicle (E11), as well as lens epithelial cells at the anterior and near the lens equator (E12.5–14.5). Autoregulation of Pax6 expression has been shown to be mediated by a single Pax6-binding site in the ectodermal enhancer (63) in concert with other DNA-binding transcription factors, including Pou2f1/Oct-1, Prep1, Six3, and Sox1 (14, 48, 65). However, expression of Pax6 is reduced in terminally differentiated lens fiber cells prior to their denucleation (karyolysis) (52). Expression of c-Maf overlaps with Pax6 in the lens placode and lens vesicle. However, its expression is strongly up-regulated in the posterior compartment of the lens vesicle and in differentiating primary lens fiber cells. Our data show that Pax6 is required to initiate c-Maf expression in the lens placode stage (E10) as well as for c-Maf expression in the differentiating primary lens fiber cells, even though levels of Pax6 expression are reduced in this part of the forming lens (52). Autoregulation of c-Maf likely contributes to high levels of expression of this gene in primary lens fibers. However, the present transgenic data (Fig. 6) suggest that this process is completely Pax6-dependent. Autoregulation of many cell type-specific transcription factors is thought to create a form of cellular memory that is required for
FIGURE 6. Analysis of crystallin gene expression in Pax6 heterozygous lenses. Significant decrease of transcripts encoding two α-, six β-, and six γ-crystallin genes was found in E15.5 Pax6<sup>+/−</sup> lenses by quantitative RT-PCR experiments. The relative mRNA abundance is calculated using the reference gene β2 microglobulin (B2M) as described under “Experimental Procedures.” Samples from one wild–type and three Pax6<sup>+/−</sup> embryos (#1, #2, and #3) were tested in each independent biological replicate. Each reaction was performed in triplicate, and experiments were replicated three times. Mean ± S.D. of individual data and corresponding p values are indicated.
Pax6/c-Maf Circuitry in Lens Differentiation

FIGURE 7. A potent GRN subcircuit comprised of Pax6 and c-Maf is used to control crystallin gene expression in the lens. A, the core regulatory subcircuit. The feed-forward loop is shown in boldface. B, GRN for αA-crystallin gene expression. The core module is expanded by utilization of the cAMP response element-binding protein that regulates Pax6 expression (81) and binds both the αA-crystallin promoter and its 3′-distal enhancer, DCR3 (50). C, GRN for αB-crystallin gene expression. The core module is expanded by utilization of AhR that binds the 5′ distal enhancer (71). D, regulation of γB- and γD-crystallins by the Pax6/c-Maf core module in combination with Hsf4, Prox1, and Sox1 (72–74). E, regulation of γE- and γF-crystallins by the Pax6/c-Maf core module in combination with Hsf4 and Sox1 (72, 73). The Pax6/c-Maf core module is shown in boldface. αA-crystallin, αA and αB; β-crystallins, βA3/ A1, βA2, βA4, βB1, βB2, and βB3; γ-crystallins, γA, γB, γC, γD, γE, and γF; Cryaa, αA-crystallin; Cryab, αA-crystallin; Crygb, β-crystallin; Crygd, γO-crystallin; Cryye, γE-crystallin; Crygf, γF-crystallin.

It has been suggested earlier by in vitro studies that Pax6 could regulate expression of c-Maf through its multiple sites found in its promoter region (-320/+656), although this fragment of DNA was highly active in a fibroblast cell line C3HT101/2 (67). These studies suggested that the rat c-Maf promoter region resides in a 0.9-kb fragment (-320/+656). The present data in transgenic mice showed that even a longer mouse c-Maf promoter fragment (-494/+866) was not active in the eye and in several other tissues tested. Although a -320/-656 promoter fragment was activated by Pax6 in a series of cotransfection tests, and its multiple binding sites were determined using recombinant Pax6 PD protein, these data are difficult to interpret because of the lack of site-directed mutagenesis of those Pax6-PD-binding sites. In our study, we identified a distal enhancer (CR1) that contains two novel functional Pax6-binding sites that are critical for the function of the CR1 enhancer. To show that Pax6 activates the endogenous c-Maf gene, a Pax6-overexpressing fibroblast cell line was established that showed up-regulation of the c-Maf gene (67). In this study, a series of genetic loss of function experiments in mice showed that temporal/spatial patterns of c-Maf expression were disrupted in Pax6 heterozygous lenses and Pax6<sup>-/-</sup> lens placodes. More recently, analysis of a conditional knockout of Pax6 in lens fiber cells revealed no change of c-Maf expression at E15.5 (19). The combination of the present data obtained in Pax6<sup>-/-</sup> lens placodes with the late conditional Pax6 knockout suggests that Pax6 is required to initiate c-Maf expression. However, it appears dispensable for its expression in maturing lens fibers. To explain this further, the chromatin structure of the c-Maf locus in postmitotic lens fiber cells could be established earlier by Pax6-binding in combination with the deposition of modified core histones and chromatin remodeling enzymes that no longer require the presence of primary DNA-binding factors, as proposed for similar developmental models (19, 68, 69).

On the basis of these studies and present data, we propose that Pax6 and c-Maf form a series of branching coherent feed-forward loops to regulate expression of individual α, β- and γ-crystallin target genes (Fig. 7). The production of diverse crystallins, encoded by 14 separate mammalian genes spread on six distinct chromosomes, is central to lens organogenesis. The lens crystallin expression program is completely disrupted in c-Maf-null lenses (24), and c-Maf has been shown to regulate expression of individual crystallin promoters (60, 64, 70). These data also show a disrupted crystallin gene transcriptome in Pax6 heterozygous lenses (Fig. 6). The other transcription factors shown earlier to directly regulate individual crystallin genes, including AhR, Hsf4, Prox1, and Sox1, provide additional levels of temporal and spatial control to specific crystallin genes. The αA- and αB-crystallin genes are regulated by this Pax6/c-Maf core module in combination with CREB (Fig. 7B) (50) or AhR (C) (71), respectively. Hsf4 regulates the expression of six γ-crystallin genes, γA–F (Fig. 7, D and E) (72). Within the γ-crystallin family, Prox1 is essential for only γB- and γD-crystallin expression (Fig. 7D), whereas Sox1 is essential for γB- and modulatory for γD-, γE-, and γF-crystallins (E) (73, 74). The functional advantage of this regulatory subcircuit is that it can control the foundation of the entire crystallin-producing machinery. This machinery is under a number of unique, lens-specific, temporal constraints (75). The primary lens fibers, tightly packed by crystallins reaching concentrations over 1000 mg/ml (76), are formed from E11.5. However, subcellular...
organelles, including nuclei, undergo destruction from E16 to E18. Degradation of these organelles terminates both transcription and translation and is essential for lens transparency (77). We conclude that Pax6 and c-Maf orchestrate the expression of α- and β/γ-crystallin genes through a series of parallel coherent feed-forward loops. This robust expression system generates major lens structural proteins, the crystallins, and, thus, serves as the foundation to form the unique three-dimensional structure, transparency, and refractive power of the ocular lens.

Novel Pax6-DNA Binding Motifs and Improved Identification of Functional Pax6-binding Sites—Newly identified Pax6-binding sites such as that in CR1 of the c-Maf gene suggest that Pax6 proteins employ a large number of combinations of five highly interactive DNA-binding subdomains: HD, β-turn, N-terminal HTH fold (PAI), linker, and C-terminal HTH fold (RED). From comparisons of natural Pax6-binding sites with SELEX-derived motifs, it is apparent that a continuum of many variants of Pax6-binding sequences exists, including those that repress transcription. As Pax6 proteins function as dual activators and repressors (57), it is possible that some of their binding sites induce structural changes in Pax6 proteins that enable them to function as transcriptional repressors. Most importantly, these data will aid analyses of Pax6 ChIP-chip and ChIP-seq data, as it is crucial to identify Pax6-binding sites within the “peak” regions of 200–600 bp. The previous Pax6 consensus sites generated a large number of false positives and false negatives (56, 57). In addition, current data could provide a reliable initial filter to distinguish the direct and indirect Pax6-DNA interactions (78) and facilitate the design of specific mutations for functional studies that probe temporal and spatial control of gene expression by Pax6 in specific tissues, as demonstrated here through a direct control of c-Maf by Pax6 in the lens.

Finally, the currently known Pax6-direct target genes explain only a minor part of the embryonic development in a specific tissue (e.g., lens, retina, cornea, iris, olfactory epithelium, forebrain, cerebellum, and pancreas) (41, 43, 44, 79). It has been shown that Pax6 regulates the batteries of common target genes in distinct tissues, including the lens and neurons (57, 80). The data also suggest that Pax6 acts as a “driver” of lens development and “suppressor” of neuronal development in the lens and vice versa (57). Given the key role of Pax6 and its related genes, e.g. eyeless (ey) in Drosophila and Pax-B and Pax-A in hydrozoan jellyfish, in eye development and evolution of visual systems, the novel data showing a spectrum of Pax6-binding sites enables better identification of those Pax6-target genes that are central to understanding of evolutionary conserved functions of the Pax6 family of genes. Thus, future studies of the role of Pax6 in different tissues and visual systems will build on the model established here for the embryonic lens.

Acknowledgments—We thank Dr. Barbara Birshstein for critical reading of the manuscript. We also thank Drs. David Beebe, Ije Huang, and Ying Liu (Washington University at St. Louis) for the RNAs prepared from wild-type and Pax6-mutated lens placodes. We thank Dr. Yan Xu (Ranapo College of New Jersey) for help with statistical analysis of data, and the Albert Einstein College of Medicine Genomics and Transgenic Mouse Facilities.

REFERENCES

1. Ali, R. R., and Sowden, J. C. (2011) *Nature* 472, 42–43
2. Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T., and Sasaki, Y. (2011) *Nature* 472, 51–56
3. Yang, C., Yang, Y., Brennan, L., Bouhassira, E. E., Kantorov, M., and Cvekl, A. (2010) *FASEB J.* 24, 3274–3283
4. Davidson, E. H. (2010) *Nature* 468, 911–920
5. Peter, I. S., and Davidson, E. H. (2011) *Cell* 144, 970–985
6. Sekiguchi, K., Adachi, T., and Sasai, Y. (2008) *Nat. Rev. Mol. Cell Biol.* 9, 770–780
7. Olson, E. N. (2006) *Science* 313, 1922–1927
8. Orkin, S. H., Wang, J., Kim, J., Chu, J., Rao, S., Theunissen, T. W., Shen, X., and Levasseur, D. N. (2008) *Cold Spring Harbor Symp. Quant. Biol.* 73, 195–202
9. Shoval, O., and Alon, U. (2010) *Cell* 143, 326–e1
10. Chow, R. L., and Lang, R. A. (2001) *Annu. Rev. Cell Dev. Biol.* 17, 255–296
11. Gunghara, L. (2011) *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 1193–1203
12. Medina-Martinez, O., and Jamrich, M. (2007) *Development* 134, 1455–1463
13. Ashery-Padan, R., Marquardt, T., Zhou, X., and Gruss, P. (2000) *Genes Dev.* 14, 2701–2711
14. Liu, W., Lagutin, O. V., Mende, M., Streit, A., and Oliver, G. (2006) *EMBO J.* 25, 5383–5395
15. Streit, A. (2004) *Dev. Biol.* 276, 1–15
16. Hill, R. E., Favor, J., Hogan, B. L., Ton, C. C., Saunders, G. F., Hanson, I. M., Prosser, J., Jordan, T., Hastie, N. D., and van Heyningen, V. (1991) *Nature* 354, 522–525
17. Hogan, B. L., Hirst, E. M., Horsburgh, G., and Hetherington, C. M. (1988) *Development* 103, 115–119
18. Quinn, J. C., West, J. D., and Hill, R. E. (1996) *Genes Dev.* 10, 435–446
19. Shaham, O., Smith, A. N., Robinson, M. L., Taketo, M. M., Lang, R. A., and Ashery-Padan, R. (2009) *Development* 136, 2567–2578
20. Collinson, J. M., Quinn, J. C., Buchanan, M. A., Kaufman, M. H., Wedden, S. E., West, J. D., and Hill, R. E. (2001) *Proc. Natl. Acad. Sci. U.S.A.* 98, 9688–9693
21. Cvekl, A., and Mitton, K. P. (2010) *Hereditas* 155, 135–151
22. Czerny, T., and Busslinger, M. (1995) *Mol. Cell. Biol.* 15, 2858–2871
23. Kawauchi, S., Takahashi, S., Nakajima, O., Ogino, H., Morita, M., Nishizawa, M., Yasuda, K., and Yamamoto, M. (1999) *J. Biol. Chem.* 274, 19254–19260
24. Ring, B. Z., Cordes, S. P., Overbeek, P. A., and Barsh, G. S. (2000) *Development* 127, 307–317
25. Eychène, A., Rocques, N., and Pouponnot, C. (2008) *Nat. Rev. Cancer* 8, 683–693
26. Ise, W., Kohyama, M., Schraml, B. U., Zhang, T., Scher, W., Basu, U., Alt, F. W., Tang, J., Oltz, E. M., Murphy, T. L., and Murphy, K. M. (2011) *Nat. Immunol.* 12, 536–543
27. Rani, A., Afzali, B., Kelly, A., Tewolde-Berhan, L., Hackett, M., Kanhere, A. S., Pedroza-Pacheco, I., Bowen, H., Jurcevic, S., Jenner, R. G., Cousins, D. I., Ragheb, J. A., Lavender, P., and John, S. (2011) *J. Immunol.* 187, 3721–3729
28. Kim, J. H., Ho, I. C., Grusby, M. J., and Glimcher, L. H. (1999) *Immunity* 10, 745–751
29. MacLean, H. E., Kim, J. I., Glimcher, M. J., Wang, J., Kronenberg, H. M., and Glimcher, L. H. (2003) *Dev. Biol.* 262, 51–63
30. Kusakabe, M., Hasegawa, K., Hamada, M., Nakamura, M., Ohsumi, T., Suzuki, H., Tran, M. T., Kudo, T., Uchida, K., Ninomiya, H., Chiba, S., and Takahashi, S. (2011) *BLOOD* 118, 1374–1385
31. Suzuki, A., Iida, S., Kato-Uranishi, M., Tajima, E., Zhan, F., Hamanaka, L., Huang, Y., Ogura, T., Takahashi, S., Ueda, R., Barlogie, B., Shaughnessy, J. D., Ragheb, J. A., Lavender, P., and John, S. (2011) *J. Immunol.* 187, 3721–3729
32. Shi, N., and Epstein, J. A. (2002) *Trends Genet.* 18, 41–47
33. van Heyningen, V., and Williamson, K. A. (2002) *Hum. Mol. Genet.* 11, 1161–1167
34. Graw, J. (2009) *J. Genet.* 88, 469–486
35. Jun, S., and Desplan, C. (1996) *Development* 122, 2639–2650
36. Singh, S., Mishra, R., Arango, N. A., Deng, J. M., Behringer, R. R., and Saunders, G. F. (2002) *Proc. Natl. Acad. Sci. U.S.A.* 99, 6812–6815
