Structural and functional analyses of disease-causing missense mutations in the forkhead domain of FOXC1

Ramsey A. Saleem1, Sharmila Banerjee-Basu3, Fred B. Berry2, Andreas D. Baxevanis3 and Michael A. Walter1,2,*

1Department of Medical Genetics and 2Department of Ophthalmology, University of Alberta, Edmonton, Alberta, Canada T6G 2H7 and 3Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892-8002, USA

Received July 17, 2003; Revised September 3, 2003; Accepted September 15, 2003

Five missense mutations (P79L, P79T, I91S, I91T and R127H) within the forkhead DNA-binding domain of the FOXC1 transcription factor, identified in patients with Axenfeld–Rieger (AR) malformations, were studied to identify the effects of these mutations on FOXC1 structure and function. Molecular modeling and threading analyses predict that the I91S and T mutations may generate local disruptions to the structure of the forkhead domain while the R127H mutation alters the electrostatic charge of the DNA binding surface of the forkhead domain. The P79L and T mutations are not predicted to grossly perturb the structure of the forkhead domain. Biological analyses indicate that all of these missense mutations cause a range of FOXC1 perturbations, including nuclear localization defects, reduced or abolished DNA binding capacity, and a reduction in the transactivation capacity of FOXC1. These experiments extend our previous hypothesis that reduced transactivation of appropriate target genes by FOXC1, underlie AR malformations mapping to human chromosome 6p25. Importantly, these results can also be applied to predict the consequences of the molecular effects of mutations of other FOX genes that have analogous missense mutations, including FOXP2, FOXE3 and FOXC2.

INTRODUCTION

Transcription factors of the forkhead box (FOX) family have increasingly been shown to be critical for proper execution of the developmental programs of embryogenesis and tissue specific cell differentiation (1,2). FOX proteins are important gene regulators in a wide range of species from yeast to human and have been implicated in diverse biological processes that include tumorigenesis and even language acquisition in humans (3). An approximately 110 amino acid segment, the forkhead domain (FHD), is the common element shared by FOX proteins. The FHD is an evolutionarily conserved variant of the helix–turn–helix DNA binding motif, composed of three α-helices and two large ‘wing-like’ loops (1,2). It is through the FHD that FOX proteins are able to interact with DNA. The FHD also contains nuclear localization signals at the N- and C-termini of the forkhead box that are required to translocate FOX proteins to cell nuclei (4).

Five missense mutations (P79L, P79T, I91S, I91T and R127H) within the forkhead DNA-binding domain of the FOXC1 transcription factor, identified in patients with Axenfeld–Rieger (AR) malformations, were studied to identify the effects of these mutations on FOXC1 structure and function. Molecular modeling and threading analyses predict that the I91S and T mutations may generate local disruptions to the structure of the forkhead domain while the R127H mutation alters the electrostatic charge of the DNA binding surface of the forkhead domain. The P79L and T mutations are not predicted to grossly perturb the structure of the forkhead domain. Biological analyses indicate that all of these missense mutations cause a range of FOXC1 perturbations, including nuclear localization defects, reduced or abolished DNA binding capacity, and a reduction in the transactivation capacity of FOXC1. These experiments extend our previous hypothesis that reduced transactivation of appropriate target genes by FOXC1, underlie AR malformations mapping to human chromosome 6p25. Importantly, these results can also be applied to predict the consequences of the molecular effects of mutations of other FOX genes that have analogous missense mutations, including FOXP2, FOXE3 and FOXC2.

Mutations in FOXC1, a member of the FOX family, underlie Axenfeld–Rieger (AR) malformations mapping to 6p25 (5,6) (Table 1). AR malformations are a genetically heterogeneous group of disorders that map not only to 6p25, but also to 4q25 where mutations in PITX2 cause AR malformations (7), and to loci at 13q14 and possibly 16q24 where the disease genes have yet to be identified (8). While genetically fully penetrant, the clinical presentation of AR malformations is heterogeneous, with variable expressivity even within a particular pedigree. Typically, patients with AR malformations will present with ocular findings including iris hypoplasia, iris adhesions, iridocorneal angle dysgenesis, displacement of Schwalbe’s line and displacement of the pupils. The most severe consequence of these ocular malformations is that approximately 50% of patients with AR malformations will go on to develop glaucoma, a progressively blinding condition. In addition to glaucoma, congenital cardiac defects and dental dysgenesis can also manifest in some individuals with FOXC1 mutations (9–11).

*To whom correspondence should be addressed at: 832 Medical Sciences Building, University of Alberta, Edmonton, AB, Canada, T6G 2H7. Tel: +1 7804929805; Fax: +1 7804926934; Email: mwalter@ualberta.ca
Murine models of Foxc1 null mutations show the spectrum of ocular and systemic defects seen in humans, revealing a broad role for FOXC1 throughout development. Mice with only a single functional copy of Foxc1 show ocular anterior segment defects that include iris hypoplasia, a displacement of Schwalbe’s line, and an aberrant and often underdeveloped iridocorneal angle (12). Foxc1/C0/C0 homozygous null mice die prenatally with hydrocephalus, open eyelids and severe skeletal anomalies (13–15), while studies using compound hetero- and homozygotes of Foxc1 and a highly related gene, Foxc2 (MIM 602402), have shown that both Foxc1 and Foxc2 are required for somitogenesis and development of the cardiovascular system (16). Experiments in zebrafish confirm a role for FOXC1 as an important regulator of somitogenesis; inhibition of synthesis of the zebrafish homolog, Foxc1a, impedes somite formation (17).

Numerous mutations in FOXC1 have been identified, including insertions, deletions, duplications, translocations, and missense mutations (5,9,10,18–23). We have previously characterized five FOXC1 missense mutations (S82T, I87M, F112S, I126M and S131L), finding that these mutations were able to either reduce the stability of the FOXC1 protein (I87M), reduce the ability of FOXC1 to bind DNA (S82T, S131L), or reduce transactivation by FOXC1 even though DNA binding still occurred (F112S, I126M) (24). In the present paper we have utilized molecular modeling in combination with molecular biology studies to investigate five additional disease-causing missense mutations (P79L, P79T, I91S, I91T and R127H) in an effort to extend our understanding of how missense mutations alter the ability of FOXC1 to adopt wild-type conformations, localize properly within the cell, bind DNA and transactivate gene expression.

We have also further characterized an L86F mutation of FOXC1, which has been shown to bind DNA while abolishing transactivation (25). Using the FHD of FOXC1 as a model, these analyses provide insight into the possible mechanisms of general FHD domain function and how these functions are determined by individual amino acids within the FHD.

### RESULTS

#### Localization of FOXC1 and mutant variants of FOXC1 to the cell nucleus

FOXC1 has previously been shown to localize to the nucleus (4,24). This localization is dependent upon a nuclear

### Table 1. Summary of human missense mutations in FOXC1

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Nucleotide position</th>
<th>Clinical features</th>
<th>Diagnosis</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>P79L</td>
<td>236</td>
<td>NL, early-onset glaucoma, posterior embryotoxon, iris hypoplasia, iris strands, scrotum defects, persistence of pupillary membrane, atrial septal defects</td>
<td>Rieger anomaly</td>
<td>(18)</td>
</tr>
<tr>
<td>P79T</td>
<td>235</td>
<td>Early-onset glaucoma, posterior embryotoxon, iris hypoplasia, iris strands, scrotum defects, persistence of pupillary membrane, atrial septal defects</td>
<td>Axenfeld–Rieger syndrome</td>
<td>(21)</td>
</tr>
<tr>
<td>S82T</td>
<td>245</td>
<td>Glaucoma, corectopia, goniodysgenesis, iris hypoplasia, iris strands, posterior embryotoxon, atrial septal defects, hearing loss</td>
<td>Axenfeld–Rieger anomaly</td>
<td>(5)</td>
</tr>
<tr>
<td>L86F</td>
<td>255,256</td>
<td>Glaucoma, posterior embryotoxon, iris hypoplasia, iridocorneal adhesions, corectopia, short stature, obesity, myocardi...</td>
<td>Axenfeld–Rieger malformation</td>
<td>(25)</td>
</tr>
<tr>
<td>I87M</td>
<td>261</td>
<td>Glaucoma, iris strands, goniodysgenesis, posterior embryotoxon</td>
<td>Axenfeld–Rieger anomaly</td>
<td>(5)</td>
</tr>
<tr>
<td>I91S</td>
<td>272</td>
<td>Parents: posterior embryotoxon. Children; iris hypoplasia and severe early-onset glaucoma</td>
<td>Axenfeld–Rieger malformation</td>
<td>(19)</td>
</tr>
<tr>
<td>I91T</td>
<td>272</td>
<td>NL</td>
<td>Axenfeld–Rieger malformation</td>
<td>(19)</td>
</tr>
<tr>
<td>F112S</td>
<td>335</td>
<td>A spectrum of anterior segment defects including corectopia, posterior embryotoxon, iris hypoplasia. Glaucoma from 5 months of age in one patient. Cardiac defects including mitral valve defects, mitral and tricuspid valve defects. Hypodontia, facial anomalies.</td>
<td>Rieger syndrome, Peter’s anomaly</td>
<td>(6,9,40)</td>
</tr>
<tr>
<td>I126M</td>
<td>378</td>
<td>Glaucoma, severe Axenfeld anomalies</td>
<td>Axenfeld anomaly</td>
<td>(6)</td>
</tr>
<tr>
<td>R127H</td>
<td>380</td>
<td>Parents: posterior embryotoxon. Children; iris hypoplasia and severe early-onset glaucoma</td>
<td>Axenfeld–Rieger anomaly</td>
<td>(20)</td>
</tr>
<tr>
<td>S131L*</td>
<td>392</td>
<td>Glaucoma, classic Rieger anomalies, Axenfeld anomalies</td>
<td>Rieger anomaly, Axenfeld anomaly</td>
<td>(6,18)</td>
</tr>
<tr>
<td>M161K*</td>
<td>482</td>
<td>Severe early-onset glaucoma, iris hypoplasia</td>
<td>Axenfeld–Rieger anomaly</td>
<td>(22,23)</td>
</tr>
</tbody>
</table>

NL, not listed. Axenfeld anomaly, iris strands connecting the iridocorneal angle to the trabecular meshwork and posterior embryotoxon; Rieger anomaly, iris hypoplasia, corectopia, or polycoria; Rieger syndrome, Rieger anomaly plus systemic findings; Axenfeld–Rieger anomaly, Axenfeld anomaly plus Rieger anomaly; Axenfeld–Rieger syndrome, Axenfeld–Rieger anomaly plus systemic findings. The asterisk indicates recurring mutations, reported in more than one pedigree.
A putative nuclear localization accessory signal (NLAS) is located at the N-terminus of the forkhead domain that is required but not sufficient to localize FOXC1 to the nucleus (4). Of the missense mutations studied herein, four (P79L, P79T, I91S and I91T) are located within what is thought to be the putative NLAS, therefore the distribution of recombinant FOXC1 within the cell was investigated (Table 2). FOXC1 P79L and FOXC1 P79T show a mild defect affecting nuclear localization. Sixty-three percent of cells expressing FOXC1 P79L and 64% of cells expressing FOXC1 P79T show nuclear localization of FOXC1, compared with 95% for wild-type FOXC1. A reduction in nuclear localization was also seen with FOXC1 I91T with only 52% of the cells showing localization of FOXC1 to the nucleus. Interestingly, disruption to nuclear localization was much more drastic with FOXC1 I91S, in which nuclear localization is reduced to only 15% of the cell population. Nuclear localization of FOXC1 also appears to be severely disrupted when the R127H mutation is present; only 17% of the cells show nuclear localization of FOXC1. The nuclear localization of six previously characterized FOXC1 missense mutant proteins (S82T, L86F, I87M, F112S, I126M and S131L) was also quantified (Table 2). Of these only S131L showed a significant reduction in nuclear localization: 72% compared with 95% for wild-type FOXC1.

EMSAs

Previous work has shown that FOXC1 is able to preferentially bind an in vitro derived binding site and form protein–DNA complexes (24,26). Recombinant FOXC1 mutant proteins were expressed and equalized to wild-type FOXC1 protein levels by western analysis (Fig. 1A). EMSAs used done to assess the effect of these missense mutations on the ability of FOXC1 to form these protein–DNA interactions. FOXC1 carrying the P79T, P79L, I91S or I91T mutation was found to retain DNA binding function, although this function is reduced in comparison to wild-type FOXC1 (Fig. 1B). In contrast the R127H missense mutation appears to disrupt FOXC1 DNA binding entirely.

The FHD of FOXC1 has been shown to bend DNA at an angle of 90° (26). We tested the ability of the FHD in the context of the entire FOXC1 protein to bends DNA, and found that FOXC1 bends the DNA on average approximately 112° (Fig. 2). The effect on DNA bending of the missense mutations that still allowed FOXC1 to form protein–DNA complexes (P79L and T, I91S and T) was then tested. In all cases the mutations did not appear to affect the ability of FOXC1 to bend DNA (data not shown).

A previously characterized FOXC1 missense mutation, I126M, was found to alter the binding site specificity of FOXC1 (6). To determine if these FOXC1 missense mutations were able to alter the DNA binding specificity, oligonucleotides that are variants of the FOXC1 binding site (Table 3) were used in EMSA assays. FOXC1 carrying P79L, P79T, I91S or I91T were of special interest since, though diminished, they still retain DNA binding activity. In all four cases, the DNA binding affinity for the variant oligonucleotides did not appear to deviate from wild-type FOXC1 affinities for these same oligonucleotides (data not shown).

Transactivation assays

Previous work has shown that FOXC1 is able to stimulate transcription, an activity inhibited by some missense mutations (4,24). Recombinant FOXC1 proteins were tested for an ability to drive expression of a luciferase reporter located downstream of a thymidine kinase promoter and FOXC1 binding sites (Fig. 3). FOXC1 carrying the P79L mutation retained 45% of wild-type FOXC1 transactivation activity, while FOXC1 P79T only retained 22% of wild-type FOXC1 activity. The I91 mutations both appear to severely reduce the transactivation of FOXC1 (I91S to 15% and I91T to 14%) when compared with wild-type FOXC1 transactivation levels. Similarly, the level of FOXC1 transactivation was reduced to 16% of wild-type activity by the presence of the R127H mutation.

In order to investigate if the transactivation defects were caused primarily by a reduction or altered nature of the FOXC1–DNA interactions, or if the defect to transactivation was due to overall structural defects to the FOXC1 molecules, a GAL4 DNA binding domain (BD) was positioned in front of the recombinant missense FOXC1 open reading frames. The ability of the exogenous GAL4 BD to rescue the FOXC1 mutant transactivation defect was then tested. It was found that the GAL4 BD was able to rescue transactivation of FOXC1 P79L/T mutants, bringing transactivation to near wild-type FOXC1 levels (Fig. 4). Partial rescue of transactivation was seen with FOXC1 I91S and R127H (27 and 32%, respectively). No transactivation above vector alone was seen with the FOXC1 I91T–GAL4 BD construct. The ability of the GAL4 BD to rescue transactivation of four previously characterized FOXC1 mutants (S82T, L86F, F112S and I126M) that bound the FOXC1 binding site but were unable to transactivate gene expression (24,25) was also tested. When fused to the FOXC1 L86F mutant, the GAL4 BD partially rescued transactivation to 32% of wild-type activity. Interestingly, when the S82T FOXC1 mutant was fused to the GAL4 BD,
transactivation surpassed wild-type activity, giving 190% of wild-type activity. Transactivation of the F112S and I126M FOXC1 mutants is not rescued by the addition of an exogenous GAL4 BD.

Homology modeling of forkhead domain of FOXC1

To date, high-resolution three-dimensional structures of the DNA binding domain for five different forkhead transcription
factors FOXC2 (FREAC-11) (27), FOXO4 (AFX) (28), FOXA3 (HNF3-g) (29), FOXK1a (ILF-1) (30) and Foxd3 (Genesis) (31) have been experimentally determined. Figure 5 shows the sequence alignment of the forkhead domains of FOXC1 and other forkhead family members for which three-dimensional structures are available. The sequence conservation is high within the forkhead family, especially within the secondary structural elements (Fig. 5). The forkhead domain of FOXC1 is very closely related to that of FOXC2 (only two conservative substitutions within the entire DNA-binding region).

Molecular models were generated for the forkhead domains of FOXC1 and disease-causing missense mutations using threading algorithms developed by Bryant and Lawrence (32) as previously described (24). The three-dimensional structure of the forkhead domain of FOXC2 (27) was used as a template for the threading analysis (see Materials and Methods). A summary of the threading results is shown in Table 4. The conformational energies show only a slight difference (∼1 kcal/mol) between the ‘self-thread’ FOXC2 and FOXC1. As expected, the probability values ($E_r/M$) for wild-type FOXC1 meet the statistical criteria of being ≥0.05 (5% $\alpha$-level).

The energy scaffolds for FOXC1 forkhead domain generated in this threading experiment are shown in Figure 6. The energy scaffolds provide a method for visualizing the important intramolecular interactions taking place within a protein. Here, the winged helical bundle is held together by numerous hydrophobic interactions, represented by the thick, magenta-colored cylinders. To further assess the structural effects of disease-causing mutations in the forkhead domain of FOXC1 mutations, threading analyses were performed on all the six missense mutations analyzed in this study. All missense mutations produced threading scores similar to the wild-type FOXC1, suggesting that these missense mutations do not destabilize the structural fold of FOXC1. Although favorable pair-wise interactions between I91 and other hydrophobic residues were eliminated in I91T/S mutants (Fig. 6B), it was not predicted to be sufficient in itself to destabilize the entire structure of the forkhead domain. The location of the disease-causing missense mutations in the molecular model of FOXC1 forkhead domain is shown in Figure 6C and D.

**Structural model for R127H in FOXC1**

Atomic models for the FOXP2 forkhead domain were generated based on the alignments obtained from threading
the R127H mutant protein might interfere with the biological functional properties of a protein reside largely on its surface, a surface of the FOXC1 forkhead domain (Fig. 7). Since the alters the electrostatic charge distribution on the DNA binding core (Fig. 6D). Replacement of Arg 127 by a histidine residue located on the third helix pointing away from the hydrophobic templates. In the molecular model of FOXC1, Arg 127 is

defects associated with the mutations are shown in Figure 8. Five different missense mutations (P79L, P79T, I91S, I91T and R127H) within the FHD upon FOXC1 function. These missense mutations differentially perturb the nuclear localization, DNA binding capacity and transactivation of FOXC1. We also utilized molecular modeling and threading analysis to assist in the understanding of how these missense mutations perturb FOXC1 function. A summary of the mutations, the positions of the mutations within the FHD, and the molecular defects associated with the mutations are shown in Figure 8.

The differences in the molecular consequences of these missense mutations, combined with our previous analyses of five other FOXC1 missense mutations (24) help define the functional composition of the FHD at a single residue level with respect to cellular localization, DNA binding, and transactivation.

**P79 may play a role in the organization of FOXC1 transactivation domains**

FOXC1 P79L and P79T both localize to the nucleus more efficiently than FOXC1 with mutations at positions 91 and 127, but are defective in localization in comparison to the exclusive nuclear localization of wild-type FOXC1. P79 is situated in a segment of the FHD that has been identified as a nuclear localization accessory domain, located at amino acid residues 77–93 of the FHD (4). The nuclear localization accessory signal is required but not sufficient for the exclusive nuclear translocation or retention of FOXC1. It is possible that mutation of the P79 residue is able to disrupt the function of this accessory signal, possibly perturbing, at least to some extent, interaction of FOXC1 with the nuclear translocation machinery. Clearly disruption to nuclear localization is not complete as a substantial proportion of cells still display exclusive nuclear localization. Additionally, this putative nuclear localization accessory domain is tolerant of some amino acid changes as both FOXC1 S82T (24) (Table 2) and L86F (25) localize to the nucleus at wild-type levels.

The fact that P79L and T are both able to bind the FOXC1 binding site supports the idea that the N-terminal portion of the FHD does not play a major role in the formation of the FOXC1 protein–DNA complex. While FOXC1 DNA binding is reduced in the presence of these mutations to levels that are approximately 50% of wild-type FOXC1 levels, the fact that these FOXC1 molecules are still able to bind DNA indicates that the structural integrity of the FOXC1 FHD is maintained. The fact that the P79L and T mutations do not appear to alter either the DNA binding specificity of FOXC1 or the ability of FOXC1 to bind DNA supports the idea that FOXC1 is functionally tolerant of amino acid changes at position P79.

**DISCUSSION**

In this paper we have described the molecular consequences of five different missense mutations (P79L, P79T, I91S, I91T and R127H) within the FHD upon FOXC1 function. These missense mutations differentially perturb the nuclear localization, DNA binding capacity and transactivation of FOXC1. We also utilized molecular modeling and threading analysis to assist in the understanding of how these missense mutations perturb FOXC1 function. A summary of the mutations, the positions of the mutations within the FHD, and the molecular defects associated with the mutations are shown in Figure 8.

The differences in the molecular consequences of these missense mutations, combined with our previous analyses of five other FOXC1 missense mutations (24) help define the functional composition of the FHD at a single residue level with respect to cellular localization, DNA binding, and transactivation.

**P79 may play a role in the organization of FOXC1 transactivation domains**

FOXC1 P79L and P79T both localize to the nucleus more efficiently than FOXC1 with mutations at positions 91 and 127, but are defective in localization in comparison to the exclusive nuclear localization of wild-type FOXC1. P79 is situated in a segment of the FHD that has been identified as a nuclear localization accessory domain, located at amino acid residues 77–93 of the FHD (4). The nuclear localization accessory signal is required but not sufficient for the exclusive nuclear translocation or retention of FOXC1. It is possible that mutation of the P79 residue is able to disrupt the function of this accessory signal, possibly perturbing, at least to some extent, interaction of FOXC1 with the nuclear translocation machinery. Clearly disruption to nuclear localization is not complete as a substantial proportion of cells still display exclusive nuclear localization. Additionally, this putative nuclear localization accessory domain is tolerant of some amino acid changes as both FOXC1 S82T (24) (Table 2) and L86F (25) localize to the nucleus at wild-type levels.

Molecular modeling and threading analyses of the FHD show good agreement with the biochemical analyses of mutations at position P79. Neither FOXC1 P79L nor P79T mutations are predicted to cause large disruptions to the overall structure of the FHD (Table 4). P79 is situated in the N terminal region of the FHD preceding α-helix 1 and is oriented in such a way that the amino acid residue extends away from the DNA (Fig. 6) and thus is unlikely to be involved in formation of the FOXC1-DNA complex. Additionally, P79 is not predicted to be involved in any pairwise interactions, the disruption of which would lead to large structural defects.

The molecular modeling and threading analyses done in this paper are based upon the NMR resolution of the near-identical FOX2 FHD (27). The NMR analysis of FOX2 was performed on the FHD in isolation, thus the threading analyses of FOXC1 done herein are limited to the FHD, independent the context of the entire FOXC1 molecule. Previous work by our laboratory has shown that FOXC1 contains modular N- and C-terminal transactivation domains (4). Transactivation studies show that the FOXC1 P79L and P79T mutations do impair FOXC1 transactivation of a reporter, however there is still a significant level of residual transactivation seen (Fig. 3). Interestingly, the addition of a GAL4 DNA binding domain restores transactivation to near-wild-type FOXC1 levels with respect to cellular localization, DNA binding, and transactivation.

**P79 may play a role in the organization of FOXC1 transactivation domains**

FOXC1 P79L and P79T both localize to the nucleus more efficiently than FOXC1 with mutations at positions 91 and 127, but are defective in localization in comparison to the exclusive nuclear localization of wild-type FOXC1. P79 is situated in a segment of the FHD that has been identified as a nuclear localization accessory domain, located at amino acid residues 77–93 of the FHD (4). The nuclear localization accessory signal is required but not sufficient for the exclusive nuclear translocation or retention of FOXC1. It is possible that mutation of the P79 residue is able to disrupt the function of this accessory signal, possibly perturbing, at least to some extent, interaction of FOXC1 with the nuclear translocation machinery. Clearly disruption to nuclear localization is not complete as a substantial proportion of cells still display exclusive nuclear localization. Additionally, this putative nuclear localization accessory domain is tolerant of some amino acid changes as both FOXC1 S82T (24) (Table 2) and L86F (25) localize to the nucleus at wild-type levels.

Molecular modeling and threading analyses of the FHD show good agreement with the biochemical analyses of mutations at position P79. Neither FOXC1 P79L nor P79T mutations are predicted to cause large disruptions to the overall structure of the FHD (Table 4). P79 is situated in the N terminal region of the FHD preceding α-helix 1 and is oriented in such a way that the amino acid residue extends away from the DNA (Fig. 6) and thus is unlikely to be involved in formation of the FOXC1-DNA complex. Additionally, P79 is not predicted to be involved in any pairwise interactions, the disruption of which would lead to large structural defects.

The molecular modeling and threading analyses done in this paper are based upon the NMR resolution of the near-identical FOX2 FHD (27). The NMR analysis of FOX2 was performed on the FHD in isolation, thus the threading analyses of FOXC1 done herein are limited to the FHD, independent the context of the entire FOXC1 molecule. Previous work by our laboratory has shown that FOXC1 contains modular N- and C-terminal transactivation domains (4). Transactivation studies show that the FOXC1 P79L and P79T mutations do impair FOXC1 transactivation of a reporter, however there is still a significant level of residual transactivation seen (Fig. 3). Interestingly, the addition of a GAL4 DNA binding domain restores transactivation to near-wild-type FOXC1 levels with respect to cellular localization, DNA binding, and transactivation.
The data from the transactivation assays and GAL4 experiments shows that mutations at P79 impair the FOXC1 transactivation domains located outside the FHD; however, when the obligation to bind DNA via the FHD is relieved by the addition of an exogenous DNA binding domain, transactivation levels are restored. While we are unable to model potential interactions between P79 and the transactivation domains located outside the FHD, these data indicate that P79 does not play a large role in maintaining the structure of the FHD, but rather is involved in the proper organization of the FOXC1 transactivation domains when FOXC1 is bound to DNA.

I91 may play a critical role in the proper organization of the FHD with respect to nuclear localization and transactivation

Molecular modeling of the FHD shows that I91 makes important pairwise interactions with other hydrophobic residues within the helical core of the forkhead domain. Disruption of these pairwise interactions is likely to cause localized structural disruptions that may impair FOXC1 function. Mutation of amino acid 91 from isoleucine to threonine reduces the exclusive nuclear localization of FOXC1 to almost half of wild-type levels. Mutation to a serine rather than a threonine, severely reduces localization of the FOXC1 molecule to the nucleus to only 15% of wild-type FOXC1 levels. The basis for the severity of FOXC1 I91S mislocalization is not abundantly clear. Molecular modeling shows that the I91 residue points inwards towards the helical core of the FHD, making the possibility of impaired interaction with the nuclear translocation machinery unlikely. Another possible cause of impaired nuclear localization is potential phosphorylation of the I91S mutation. In silico analysis does not predict either the 91S or 91T residue would be phosphorylated (data not shown) and, again, the orientation of the residue makes this possibility unlikely. It is also possible that localized disruptions caused by the conversion of the isoleucine to a serine residue at position 91 leads to a ripple effect whereby the interactions of other amino acids within the FHD with the nuclear translocation machinery are disrupted. Why the effect of I91S is more severe than I91T remains to be investigated.

FOXC1 I91S and I91T are both able to bind the FOXC1 binding site, although with reduced capacity, again providing evidence that the N-terminal portion of the FHD, including α-helix 1, does not play a major role in the formation of the FOXC1 protein–DNA complex. DNA binding is reduced as FOXC1 I91S and I91T mutant proteins both show 3- to 4-fold reductions in DNA binding compared with wild-type FOXC1, with neither mutation altering the binding specificity of FOXC1 or impairing FOXC1 DNA bending. The fact that the FOXC1 I91 mutant proteins are still able to bind the FOXC1 binding site indicates that the structural integrity of the FOXC1 FHD is maintained, at least to some extent. This is in good agreement with threading analysis, which predicts that.

**Table 4.** Statistics of optimal threads of FOXC1 and its mutations

<table>
<thead>
<tr>
<th></th>
<th>$\Delta G_R/M$</th>
<th>$Z_{(S)}$</th>
<th>$P_{(\text{norm})}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOXC1</td>
<td>1D5V A</td>
<td>126.8</td>
<td>0.0049</td>
</tr>
<tr>
<td>P79L</td>
<td>1D5V A</td>
<td>134.2</td>
<td>0.0059</td>
</tr>
<tr>
<td>P79T</td>
<td>1D5V A</td>
<td>126.6</td>
<td>0.00561</td>
</tr>
<tr>
<td>I91S</td>
<td>1D5V A</td>
<td>121.8</td>
<td>0.00764</td>
</tr>
<tr>
<td>I91T</td>
<td>1D5V A</td>
<td>120.4</td>
<td>0.02114</td>
</tr>
<tr>
<td>R127H</td>
<td>1D5V A</td>
<td>130.0</td>
<td>0.00146</td>
</tr>
<tr>
<td>FOX2</td>
<td>1D5V A</td>
<td>127.9</td>
<td>0.00418</td>
</tr>
</tbody>
</table>

Sums of contact potentials are expressed as a conformational energy $\Delta G_R/M$, the energy associated with non-local, non-bonded interactions. $Z$-scores, in standard deviation units, represent the variance from the mean of $\Delta G_R/M$. (Fig. 4). The data from the transactivation assays and GAL4 experiments shows that mutations at P79 impair the FOXC1 transactivation domains located outside the FHD; however, when the obligation to bind DNA via the FHD is relieved by the addition of an exogenous DNA binding domain, transactivation levels are restored. While we are unable to model potential interactions between P79 and the transactivation domains located outside the FHD, these data indicate that P79 does not play a large role in maintaining the structure of the FHD, but rather is involved in the proper organization of the FOXC1 transactivation domains when FOXC1 is bound to DNA.

**Figure 5.** Multiple sequence alignment of forkhead domains of selected proteins. The sequences shown in single letter amino acid codes are human FOXC1 (Genbank accession no. Q12948), FOXC2 (FREAC-11; Genbank accession no. Q99958), FOXA3 (HNF3a; Genbank accession no. P55318) and FOXO4 (AFX1; Genbank accession no. P98177). Amino acid residues showing absolute identity among these proteins are shown in white against a black background; those positions with conservative substitutions are shown in black against a gray background. The missense mutations of FOXC1 analyzed in this study are shown by solid circles. The missense mutations analyzed in a previous study are shown by filled triangles (24). The positions of the four α-helices defined in the solution structure of human FOXC2 (pdb 1D5V) are schematically represented in the bar below the alignment. The arrows below the alignment indicate the positions of β-strands in the FHD. ALSCRIPT was used to format the alignment (37).
while both I91S and I91T mutations disrupt important pairwise interactions, neither mutation significantly disrupts the overall structure of the FHD.

Both FOXC1 I91S and I91T show severe disruptions to transactivation; this transactivation deficiency cannot be rescued by the addition of the exogenous GAL4–DNA binding domain in either mutant. This data indicates that the localized loss of some of the structural integrity within the FHD produced by the I91S and I91T mutations has intramolecular effects that extend beyond the FHD, silencing the FOXC1 transactivation domains.

The R127H mutation alters the electrostatic charge of the FHD

Of the missense mutations tested herein, the FOXC1 R127H missense mutation caused the largest disruption to FOXC1 function. This mutant protein displayed a severe disruption to nuclear localization, was unable to bind DNA, and showed a severe transactivation defect. While the defect to transactivation is severe at only 16% of wild-type levels, there is still residual transactivation above what is seen with the ΔFHD transactivation levels. It is possible that the EMSAs, which show no
binding of FOXC1 to the FOXC1 binding site, are not sensitive enough to detect small amounts of FOXC1 R127H binding that are sufficient for transactivation above baseline levels. It is also possible that FOXC1 R127H binding occurs when the FOXC1 binding site is presented in context of the reporter plasmid and not in the context of the smaller oligonucleotides used in the EMSAs. Nevertheless, the severe disruption of FOXC1 R127H transactivation and DNA binding demonstrates the importance of this residue with respect to FOXC1 function.

Threading analysis of the effect of the R127H mutation on the structure of the FHD predicts that the overall structure of the FHD is normal, a prediction supported biochemically by the stable protein expression of FOXC1 R127H (Fig. 1A). Molecular modeling of the R127H mutation also predicts that the conversion of arginine 127 to a histidine residue alters the electrostatic charge distribution on the surface of the FHD (Fig. 7). While R127 is not located within the previously defined FOXC1 nuclear localization signal (4), clearly the altered electrostatic charge distribution is able to disrupt the interaction of FOXC1 with the nuclear transport machinery.

R127 is situated in α-helix 3 of the FHD, the recognition helix that interacts within the major groove of DNA. Disruption of the electrostatic charge on the portion of the FHD that makes contact with the DNA would predictably alter the affinity of FOXC1 for DNA. The fact that FOXC1 R127H no longer binds DNA demonstrates that R127 plays a role in the proper organization of the FHD with respect to DNA binding, and also confirms the idea that α-helix 3 plays a substantial role in the formation of FOXC1–DNA complexes.

FOXC1 R127H has a severe transactivation defect that is only weakly rescued by the addition of the exogenous GAL4–DNA binding domain. These data would indicate that the R127H mutation has intramolecular effects that disrupt the transactivation domains located outside the FHD. When the FOXC1 R127H DNA binding deficiency is compensated for by the addition of the exogenous GAL4–DNA binding domain, the modular transactivation domains of FOXC1 still show reduced activity.

The R127H mutation found in AR malformation patients is the positional equivalent to the R553H mutation found in the FHD of FOXP2, a forkhead gene that underlies severe language deficits (3,33). Our data indicate that the R553H mutation identified in FOXP2 would probably also severely perturb the FOXP2 forkhead domain, leading to a similar

Figure 7. Molecular models of the R127H mutation. (A) Surface electrostatic charge distribution of FOXC1. (B) Surface electrostatic charge distribution of R127H. (C, D) Ribbon models showing the positions of Arg 127 and His 127, respectively.
spectrum of molecular defects, disrupting the ability of FOXP2 to transactivate target genes required for proper language development programs.

The addition of an exogenous GAL4 DNA binding domain rescues transactivation defects of FOXC1 S82T and L86F missense mutations

Four previously characterized missense mutations (S82T, L86F, F112S and I126M) (24,25) were tested to determine if the addition of an exogenous GAL4–DNA binding domain (BD) would allow recovery of transactivation. Transactivation was restored with the GAL4–FOXC1 S82T and partially restored with the L86F chimeric proteins; the transactivation ability of FOXC1 F112S and I126M mutant proteins was not restored by the addition of the exogenous GAL4–DNA binding domain. These data would indicate that mutations at position L86 do not exert a large intramolecular influence; thus when the obligation to bind DNA via the FHD is relieved by the presence of the GAL4 BD, the transactivation domains located within FOXC1 are able to exert their influence and activate transcription. These data also indicate that the F112S and I126M missense mutations exert an intramolecular effect that disables the transactivation domains even when DNA binding through the FHD is relieved by the addition of the exogenous DNA binding domain.

Interestingly, the ability of the GAL4–BD–S82T chimera to transactivate at 190% of GAL4–BD–FOXC1 chimera levels may indicate that S82T may exert an influence both within the FHD and within the adjacent inhibitory domain (4). The S82T FOXC1 protein binds the FOXC1 binding site at 25–33% of wild-type FOXC1 levels, yet is able to transactivate at 57% of wild-type FOXC1 levels. Thus, while the S82T mutation impairs DNA binding to some extent, it may have a second effect and act to relieve inhibition of transactivation by the inhibitory domain. A possibility is that when the defect to DNA binding is bypassed by the addition of the exogenous GAL4 DNA binding domain, the effect of the S82T mutation on the inhibitory domain is revealed, and enhancement of transactivation can be detected. Conversely, the I91S, I91T, F112S, I126M and R127H mutations are able to either alter the overall structure of FOXC1 to such an extent that the transactivation domains do not function or these mutations enhance the inhibition of the transactivation inhibitory domain that is adjacent to the FHD.

Genotype–phenotype correlations

FOXC1 mutations and AR malformations do not show strong genotype–phenotype correlations. FOXC1 mutations are highly penetrant, manifesting in AR malformations, but display...
considerable variable expressivity in their phenotypic consequences. The patient with the P79L mutation was described as having Rieger’s anomaly, presumably presenting with ocular defects such as iris hypoplasia, displaced pupils or multiple pupils (18). Patients carrying the P79T mutation are much more seriously affected, presenting with ocular AR malformations and glaucoma, as well as systemic defects including maxillary hypoplasia and cardiac anomalies (21). However biochemical analyses demonstrate that the P79L and P79T mutation disrupt FOXC1 function to similar extents. In contrast to the P79L and T mutations, the FOXC1 R127H retained almost no normal FOXC1 function. The nuclear localization was severely disrupted, DNA binding was undetectable and the transactivation capacity of FOXC1 was drastically reduced and could not be rescued by the addition of an exogenous DNA binding domain. Thus while all the biochemical data would indicate that the R127H mutation is more disruptive to FOXC1 function than the P79T mutation, the phenotype reported for patients with the R127H mutation is less severe than patients with the P79T mutation. Patients with R127H were reported to present with posterior embryotoxon, iris hypoplasia and early-onset glaucoma with no systemic defects (20).

This data, together with our previous analyses of five other FOXC1 missense mutations, indicates no strong genotype–phenotype correlation can be discerned for FOXC1 missense mutations. The largest level of complexity in assessing genotype–phenotype correlations likely comes from the effects of modifier genes within individual patients, and a level of stochastic chance during the development of an individual. Studies in mice also indicate that ocular anomalies occur in Foxc1+/− heterozygotes (12), but the exact nature and severity of the malformation is also highly variably, depending on the individual mouse and the genetic backgrounds of the mice in general, a recapitulation of what is seen in humans and evidence of the importance of modifiers and stochastic developmental events in determining phenotype.

These studies of FOXC1 missense mutations demonstrate that mutations do indeed impair FOXC1 function and also reveal the nature of the functional defect. The combination of biological and computational analysis used in these analyses has established a framework which will now allow for predictive modeling of changes within the FHD for not only members of the C class of FOX proteins but also FHDs of other FOX classes, such as FOXP2 and FOXE3.

MATERIALS AND METHODS

Reports on patients

All patients with FOXC1 missense mutations presented with features of AR malformations. The detailed clinical descriptions of these patients with FOXC1 mutations have been reported elsewhere (18–21).

Homology model building

Threading experiments were performed by the method of Bryant and Lawrence (32), with detailed derivations and methodology provided therein. Each query sequence was threaded through the atomic coordinates of the solution structure of the forkhead domain of FOXC2 (27). Three core segments were defined based on the three-dimensional structure. For each possible alignment, individual pairwise residue interactions were determined based on chemical type and distance intervals, lookup tables for which are present in Bryant and Lawrence (32). Using these values, a conformational energy AGRM, defined as the expected work for substitution of a specific sequence R for a random sequence with the same composition in the context of folding motif M, was then calculated for each alignment. Z-scores (ZR,M) and chance occurrence probabilities (ER,M) were calculated to compare conformational energies for different alignments. Chance occurrence probabilities give the odds that a random sequence of the same length and amino acid composition would yield a threading energy as low as the query sequence R. Calculations of energies and statistical significance were performed using C and S-PLUS subroutines (32). Critical interactions are defined as those having a pairwise interaction energy ≤−1 kcal/mol. All energy scaffold figures were generated using the GRASP software package (34). The MODELER (35) package was also used to build and validate the forkhead domain models, using two experimentally determined structures (FOXC2 and FOXO4).

Plasmid construction and mutagenesis

FOXC1 was cloned into the pcDNA4 His/Max B™ (Invitrogen) as described previously (24). Mutagenesis was done as described previously (24) using the Quickchange mutagenesis kit (Stratagene) and appropriate primers listed below. The primers and resulting protein changes are as follows: P79L forward 5’-cag gaa gag gtc gag GTC gct ggc cat ggt gat gag-3’, reverse 5’-gat gta gta ata ggg cag ctt cac cat gtc ctt g-3’. I91S forward 5’-ctc acc atc acc ctg ggc cgg cag cct gcc cgc gag gac-3’, reverse 5’-gtc cgg ggc ggt ctt ggc gct ggt gct gat ggg gag-3’. R127H forward 5’-aag aac gaa ggg cag ctt cac cac cgg gac cgt ctt g-3’, reverse 5’-gag cca ggt gtt ggt ggt gat gtt gcc cta cga-3’. For P79T and I91T degenerate oligonucleotides were employed. Mutagenized PCR products containing the threonine codon (GCA) were selected: P79T forward 5’-aag gag gtc gga aag ggc ggt gct cta tag cta cat c-3’, reverse 5’-ag tga gat gta atg ggg ggg cag ctt cac cat gtc ctt g-3’. I91S forward 5’-ctc acc atc acc atc ggg cgg cag cct gcc cgc gag gac-3’, reverse 5’-ggg cca ggg ggg ggg cgg ggt ggg gtc ggg ctt ggg ctc cag cgg cag gac cac cac ctc tcg ctc-3’, reverse 5’-ggg cca ggg ggg ggg cgg ggt ggg gtc ggg ctt ggg ctc cag cgg cag gac cac cac ctc tcg ctc-3’. For P79T and I91T degenerate oligonucleotides were employed. Mutagenized PCR products containing the threonine codon (GCA) were selected: P79T forward 5’-aag gag gtc gga aag ggc ggt gct cta tag cta cat c-3’, reverse 5’-ag tga gat gta atg ggg ggg cag ctt cac cat gtc ctt g-3’. I91S forward 5’-ctc acc atc acc atc ggg cgg cag cct gcc cgc gag gac-3’, reverse 5’-ggg cca ggg ggg ggg cgg ggt ggg gtc ggg ctt ggg ctc cag cgg cag gac cac cac ctc tcg ctc-3’. For P79T and I91T degenerate oligonucleotides were employed. Mutagenized PCR products containing the threonine codon (GCA) were selected: P79T forward 5’-aag gag gtc gga aag ggc ggt gct cta tag cta cat c-3’, reverse 5’-ag tga gat gta atg ggg ggg cag ctt cac cat gtc ctt g-3’. I91S forward 5’-ctc acc atc acc atc ggg cgg cag cct gcc cgc gag gac-3’, reverse 5’-ggg cca ggg ggg ggg cgg ggt ggg gtc ggg ctt ggg ctc cag cgg cag gac cac cac ctc tcg ctc-3’. For P79T and I91T degenerate oligonucleotides were employed. Mutagenized PCR products containing the threonine codon (GCA) were selected: P79T forward 5’-aag gag gtc gga aag ggc ggt gct cta tag cta cat c-3’, reverse 5’-ag tga gat gta atg ggg ggg cag ctt cac cat gtc ctt g-3’. I91S forward 5’-ctc acc atc acc atc ggg cgg cag cct gcc cgc gag gac-3’, reverse 5’-ggg cca ggg ggg ggg cgg ggt ggg gtc ggg ctt ggg ctc cag cgg cag gac cac cac ctc tcg ctc-3’.

Potential mutant FOXC1 constructs were sequenced with a Li-COR automated sequencer. Confirmed mutants were subcloned back into the FOXC1 pcDNA4 His/Max vector and these final clones re-sequenced.

GAL4 fusion proteins were constructed by subcloning entire FOXC1 open reading frames from the FOXC1 pcDNA4 His/Max plasmids into EcoRI/Xbal sites in the pGAL DB vector.

Cell transfection

COS-7 cells were grown in Dulbecco’s modified Eagles medium and 10% fetal bovine serum, transfected with 2.5 μg of plasmid for 100 mm plates or 0.75 μg for six-well plates, utilizing FuGene6™ transfection reagent according to the manufacturer’s protocol. Cells were harvested 48 h after transfection and the resulting cell extracts analysed by western
PCR products were [32P]-labeled and EMSA assays done as with FOXC1 BS Reverse–pGEM as a template. The resulting Forward–pGEM as a template, primer pairs 4 and 5 were used (www.cbs.dtu.dk/services/NetPhos/).

Analysis was done using the NETPHOS Prediction Server Phosphorylation prediction.

COS-7 cells using 1 ng pRL, 50 ng G5 reporter plasmid and manufacturers protocol (Promega DLR kit).

HeLa cells were transfected with 1 ng pRL, 50 ng of the FOXC1–BS–TK–luciferase reporter plasmid, and 500 ng of a given FOXC1 expression plasmid. Twenty-four to 48 h post-transfection, cells were harvested and assayed according to the manufacturers protocol (Promega DLR kit).

The GAL4-FOXC1 assays were performed as above but in COS-7 cells using 1 ng pRL, 50 ng G5 reporter plasmid and 500 ng of the GAL4-FOXC1 plasmid.

Phosphorylation prediction

Analysis was done using the NETPHOS Prediction Server (www.cbs.dtu.dk/services/NetPhos/).

ACKNOWLEDGEMENTS

We would like to thank the Ocular Genetics Laboratory for critical reading of this manuscript, and May Yu for technical assistance. This work was supported by the Alberta Heritage Foundation for Medical Research (AHFMR) and the Canadian Institutes for Health Research (CIHR). R.A.S. is supported by a University of Alberta Dissertation Fellowship. F.B.B. is supported by a CIHR post doctoral fellowship. M.A.W. is an AHFMR senior scholar and a CIHR investigator.

REFERENCES


