A Multidomain TIGR/Olfactomedin Protein Family with Conserved Structural Similarity in the N-terminal Region and Conserved Motifs in the C-terminal Region*

Michelle L. Green‡ and Teri E. Klein

Based on the similarity between the TIGR (trabecular-meshwork inducible glucocorticoid response) (also known as myocilin) and olfactomedin protein families identified throughout the length of the TIGR protein, we have identified more distantly related proteins to determine the elements essential to the function/structure of the TIGR and olfactomedin proteins. Using a sequence walk method and the Shotgun program, we have identified a family including 31 olfactomedin domain-containing sequences. Multiple sequence alignments and secondary structure analyses were used to identify conserved sequence elements. Pairwise identity in the olfactomedin domain ranges from 8 to 64%, with an average pairwise identity of 24%. The N-terminal regions of the proteins fall into two subgroups, one including the TIGR and olfactomedin families and another group of apparently unrelated domains. The TIGR and olfactomedin sequences display conserved motifs including a residual leucine zipper region and maintain a similar secondary structure throughout the N-terminal region. The correlation between conserved elements and disease-associated mutations and apparent polymorphisms in human TIGR was also examined to evaluate the apparent importance of conserved residues to the function/structure of TIGR. Several residues have been identified as essential to the function and/or structure of the human TIGR protein based on their degree of conservation across the family and their implication in the pathogenesis of primary open-angle glaucoma. Additionally, we have identified a group of chitinase sequences containing several of the highly conserved motifs present in the C-terminal region of the olfactomedin domain-containing sequences. Molecular & Cellular Proteomics 1:394–403, 2002.

Our previous work has established that the relationship between the TIGR (trabecular-meshwork inducible glucocorticoid response) (also known as myocilin) and olfactomedin proteins is evident not only in the C-terminal olfactomedin domain but also in the N-terminal domain of both proteins (1). Although the N-terminal domain of olfactomedin is divergent (2), the C-terminal domain is highly conserved among the TIGR and olfactomedin proteins. To identify sequence elements important to the structure and function of this family in a broader context, a more divergent set of sequences must be evaluated (3).

Karavanich and Anholt (2) first compared the olfactomedin-related proteins from rat, mouse, human, and frog and identified several conserved motifs. The pairwise identities among these sequences ranged from 22 to 98% identical. Upon identification of the human TIGR protein’s homology with olfactomedin, Nguyen et al. (4) confirmed that TIGR possessed several of the motifs identified by Karavanich and Anholt (2). Rozsa et al. (5) extended their evaluations to a slightly more divergent set of proteins, including a rat latrophilin sequence and a human olfactomedin-related sequence.

Subsequently, Kulkarni et al. (6) have identified a human olfactomedin gene family of which TIGR/myocilin is one member. The members of this family, designated HOLFA, HOLFB, HOLFC, HOLFD, and HTIGR, are distributed in various tissue types. For example, HOLFB is expressed in pancreatic and prostate tissues, and HOLFC is expressed in cerebellum. In their study, the researchers concentrated their analysis on the olfactomedin domain of the sequences because of extreme differences in sequence lengths among the olfactomedin-related proteins. When the evolutionary relationships between these domains are evaluated, HOLFA and HOLFC appear to be related phylogenetically, as do HOLFB and HOLFD. TIGR/myocilin appears to belong to a separate subgroup. In addition to the human gene family, Kulkarni et al. (6) examined 13 other human and non-human sequences, including latrophilins and several ESTs.1

In this study, we have extended the analysis of the human olfactomedin-related gene family. Despite continued interest in the olfactomedin domain-containing (ODC) proteins, the native function has not been determined for any of these proteins. Presumably, sequence elements conserved across long evolutionary times are likely to be essential to the struc-

1 The abbreviations used are: EST, expresses sequence tag; ODC, olfactomedin domain-containing; LR, leucine-rich; LZ, leucine-zipper.

From Stanford Medical Informatics, Stanford University School of Medicine, Stanford, California 94305-5479
Received, April 1, 2002, and in revised form, May 8, 2002
Published, MCP Papers in Press, May 21, 2002, DOI 10.1074/mcp.M200023-MCP200
tural or functional success of a protein family (7). To elucidate potential sequence epitopes important to the native functions of the TIGR sequence, we have attempted to identify a larger, more diverse family of related sequences (<30% identical). Based on this larger set, we have identified conserved sequence motifs that may be key to the function or structure of the family. Using these motifs, we have also identified a family of proteins that may be distantly related to the TIGR and olfactomedin proteins.

**EXPERIMENTAL PROCEDURES**

**Identification of New Sequences**—Beginning with the set of 18 olfactomedin-related sequences of Kulkarni et al. (6) and several TIGR and olfactomedin orthologs (a total of 22 sequences after sequences more than ~75% identical were removed from the olfactomedin-related sequences), we used each sequence2 to query the GenBankTM databases to identify additional ODC sequences. From this diverse set, we searched the databases using the individual region of each protein that aligned with each of the three exon ajonctions of TIGR. We then evaluated the output from these searches using Shotgun (8) to identify hits that are found by multiple query sequences. Shotgun hits were evaluated based on their sequence similarity with our initial set (using both pairwise and multiple alignments) to determine whether they could be related sequences. The newly accepted sequences were used to query GenBankTM, these data were added to our set of database output, and the Shotgun search was repeated.

**Secondary Structure Prediction**—Several programs are available currently for predicting secondary structure from sequence. Because these algorithms employ different methods for prediction, they can result in contradictory results. To maximize the accuracy of our predictions, we used Jpred (9) jura.ebi.ac.uk:8888/, a server that incorporates information from several programs including PHD, PREDATOR, and 2PRED. Jpred can predict secondary structure for single sequences or multiple alignments, thus incorporating evolutionary data, in addition to the primary sequence. The server calculates a consensus secondary structure from the individual predictions thereby increasing the accuracy.

The Jpred server can perform a multiple sequence alignment on a set of unaligned sequences and provide predictions based on that alignment. Alternately, a set of aligned sequences may be submitted for secondary structure prediction. Multiple alignments are sensitive to the composition of the sequences included in the alignment (7). To evaluate our hypothesis that the ODC sequences are related in the N-terminal region and, hence, have similar structures, several separate predictions were made. All of the sequences containing a domain.

---

2 The following sequences were used to search for distant homologs: HTI8GR, NCBI accession number NP_000252; BTI8GR, NCBI accession number BAA82152; MITI8GR, NCBI accession number NP_034995; RTI8GR, NCBI accession number BAA34190; COLFA, NCBI accession number AAF40413; FOLFA, NCBI accession number Q07081; MOLF, NCBI accession number AAB84058; HOLFA, NCBI accession number A07984; ROLFA, NCBI accession number A0003420; HOLF, NCBI accession number AA47264; HOLF, NCBI accession number AAD20056; Holden, NCBI accession number W53028; MPA3, NCBI accession number BAA28767; HKIAA, NCBI accession number BAA47484; LHAT2, NCBI accession number CAA10458; BLAT1, NCBI accession number AD09191; BLAT3, NCBI accession number AD09324; PLAT1, NCBI accession number A09324; PLAT2, NCBI accession number A09324; WOLF1, NCBI accession number A0852933; WOLF2, NCBI accession number C0B04088.

---

3 The following sequences were identified by our searches: DOLFA, NCBI accession number AAF48788; JOLFN, NCBI accession number AV670132 and NCBI accession number AV670161; MIDN, NCBI accession number AVID87584; HOLF, NCBI accession number XP_00131313; HOLF, NCBI accession number CAC17635; HEST4, NCBI accession number A5L62289; HEST3, NCBI accession number BF530912; HL-P4, NCBI accession number T23140; ZEST1, NCBI accession number AW15690; ZEST2, NCBI accession number AW154600. Because these searches were completed, the HOLF sequence has been deleted from GenPept at the request of the submitter; however, a mouse sequence 86% identical to HOLF is still present with NCBI accession number AAH05485.
Conserved Elements in the Olfactomedin Protein Family

ing sequences are equally divergent from both the TIGR and olfactomedin sequences but do appear to be evolutionarily related to the group.

Consistent with the conclusions of Karavanich and Anholt (2) and Kulkarni et al. (6), the C-terminal region is well conserved and appears to have evolved slowly, predominantly via point mutations; the N-terminal region, on the other hand, seems to have evolved more quickly. Because of the different apparent rates of evolution between the ends of the molecule, we have analyzed these two regions separately.

**N-terminal Region**

Fig. 2 shows a ClustalW alignment of a representative set of the N-terminal regions of the sequences of Kulkarni et al. (6), olfactomedins, TIGR orthologs, and all of the sequences we have identified in this work containing an N-terminal region. In this region, the similarity among the JOLHN, MKIDN, HOLFE, HOLFF, and HOLFC sequences and the remaining olfactomedin, TIGR sequences is evident. The GXCX(T) consensus motif is clear in these sequences, as well as the TIGR and olfactomedin sequences. Also, several of the sequences possess a residual leucine-zipper (LZ) or a leucine-rich (LR) region in their N terminus. Fig. 2 also shows the obvious disparity among the latrophilins (represented by BLAT1), DOLFA, and WOLF2 and the remaining sequences. These proteins lack the GXCX(T) motif, as well as a residual LR/LZ region.

In our pursuit of distantly related sequences that might be part of a TIGR/olfactomedin superfamily, we searched for homology to the N-terminal region of the sequences shown in Fig. 2. Our congruence analysis using Shotgun did not identify any distantly related sequences with significant similarity to the set other than their LR/LZ regions. In fact, the list of Shotgun hits consisted primarily of sequences with LR or LZ regions. The hits included proteins such as myosin, plectin, kinesin, tropomyosin, and several other coiled-coil sequences. Previous work has shown that this similarity does not appear to be a sufficient basis to infer an evolutionary relationship (1).

In distantly related proteins, structure is often more highly conserved than sequence (11). Although proteins sharing 30% sequence identity are expected to exhibit the same fold structure, many proteins with statistically insignificant sequence similarity may also have similar folds (12, 13). Because the coiled-coil structure and LR/LZ domain appeared to be the basis for the similarity between our query sequences and their Shotgun hits, we decided to evaluate our sequences based on their structures. Three-dimensional structural information is not available for any of these sequences; hence we compared the predicted secondary structures for the proteins.
Because the composition of a set of sequences can significantly affect a multiple alignment and, hence, the resultant secondary structure prediction, several separate predictions were made for the N-terminal region. The TIGR-like sequences, the olfactomedin-like sequences, and the latrophilin sequences were each submitted separately for secondary structure predictions. Fig. 2 shows a representative sequence from each prediction with the secondary structure elements, helices and β-strands, indicated for each group. The sequences that did not clearly fall into one of these groups were also submitted individually to confirm their predictions were consistent with the larger group.

The predicted secondary structures confirm the results of the sequence alignments; DOLFA, WOLF2, and the latrophilins appear to be part of groups separate from that of the olfactomedin- and TIGR-related sequences. All three sequences (DOLFA, WOLF2, and BLAT1) are primarily sheet-like in the N-terminal region. The TIGR- and olfactomedin-related sequences are all primarily helical. Although there are differences in the initiation and termination of the helices and sheets in the separate predictions, the predictions are quite similar for the TIGR- and olfactomedin-related sequences. JOLHN, HOLFE, and HOLFF are also predicted to be predominantly helical in the N-terminal region.

### C-terminal Region

**Secondary Structure of the C-terminal Region**—We also used Jpred to predict the secondary structure of the C-terminal region. The entire family of ODC sequences aligned by ClustalW was submitted to the Jpred server. Each program predicted that the majority of this domain is composed of β-sheets, with only a few very short helical regions as shown in Fig. 3 for the HTIGR sequence. The figure also includes the N-terminal region of TIGR to highlight the explicit difference in the predicted structures of the two domains.

**Conserved Sequence Elements**—In the C-terminal olfactomedin domain, our analysis included 32 sequences described previously and in the current work. Karavanich and Anholt (2) identified five conserved motifs in this domain; however, these motifs were assessed using only four sequences, rat, mouse, and frog olfactomedin and an EST from *Caenorhabditis elegans*. Our analysis of a larger evolutionarily divergent collection of sequences has enabled us to specify a more stringent

---

**TABLE I**

| Name     | Accession No. | Species/description | Region of homology |
|----------|---------------|---------------------|--------------------|
| DOLFA    | AAF48788      | CG6867 gene product (*Drosophila melanogaster*) | C-terminal         |
| JOLHN    | AV670132      | OLNH cell line cDNA library (OLb) *Oryzias latipes* cDNA clones | N- and C-terminal |
| MKIDN    | AIA987584     | *Mus musculus* cDNA clone | N- and C-terminal |
| HOLFF    | XP_001313     | Sugano mouse kidney mkia Mus musculus cDNA clone | N- and C-terminal |
| HEST3    | BF530912      | NCI_CGAP_Bm67 Homo sapiens cDNA clone | N- and C-terminal |
| HEST4    | AL562289      | LTI_NFL003_NBC3 Homo sapiens cDNA clone | C-terminal        |
| ZEST1    | AW115690      | Sugano Kawakami zebrafish DRA Danio rerio cDNA clone 2597775 | C-terminal        |
| ZEST2    | AW154600      | Sugano Kawakami zebrafish DRA Danio rerio cDNA clone 2639120 | C-terminal        |
| HL-P4    | T23140        | Halibut pituitary cDNA T23140 | C-terminal        |
set of conserved residues that might be used to recognize more distant members of this group.

As discussed previously, we divided the olfactomedin domain into nine segments for our analysis and then removed redundant sequences to $<75\%$ identity. The conserved motifs within each of the nine segments are shown in Fig. 4. Segments 2 and 4 are not shown, because they did not contain any residues strongly conserved by identity or by type. Some of the motifs shown in Fig. 4 may appear to include sequences with higher than $75\%$ identity. The $<75\%$ sequence identity comparison and the identification of conserved residues were performed on the entire segment, rather than just the conserved portions shown in the figure. For example, in region 1, BLAT1 and BLAT3 are 100% identical (HQSGAWCKDPL); however, the surrounding sequences in segment 1 are more divergent ($<75\%$ identical). The exception to this rule is the HTIGR sequence. To maintain a consistent reference sequence throughout the comparison, and because we were interested primarily in the relationship of the other family members to the HTIGR sequence, it was added back to the alignments after the conservation analyses were completed.

Five residues within the nine segments of the olfactomedin domain are conserved completely across all of the sequences (excluding those not present in a particular segment). Three of these conservations occur in segment 1 (Gly-268, Trp-270, and Asp-273 in the HTIGR sequence), one occurs in segment 5 (Glu-385), and one occurs in segment 9 (Asn-480). Several of the highlighted motifs were identified by Karavanich and Anholt (2) and Nguyen et al. (4). The additional evolutionary information included in our divergent alignment helps to clarify the residues that have been conserved across evolutionary time. For instance, the DDXDXDEGLW motif is not as well conserved across our whole set of ODC sequences as it is among the olfactomedin and TIGR sequences alone.

One of the most extensive conserved motifs does not contain any fully conserved residues. The area containing the cysteine shown to be involved in the formation of oligomers in olfactomedin and believed to be involved in dimerization in TIGR (Cys-433; region 7) has no fully conserved residues. This surprising result is primarily because of the presence of two sequences, HEST4 and ZEST2. The ZEST2 sequence is consistent with the conserved motif, with the exception of an arginine in place of the conserved cysteine. Although the HEST4 sequence possesses a cysteine in the conserved position, the surrounding residues do not match the motif. Both HEST4 and ZEST2 are EST sequences and contain nonsense codons upstream of this motif. Future studies should investigate the accuracy of these sequences in this region.

Comparison of HTIGR Mutations and Conserved Residues—Presumably, sequence elements essential to the function or structure of a family of evolutionarily related proteins should be more highly conserved than less essential elements. When functional or structural information is available for a family of proteins, differences among the proteins can be useful for determining which conserved residues are key elements in its function or structure (3). Partially conserved or divergent residues, which impart differences in function or structure, may also be identified.

Because the normal function and three-dimensional structure of the TIGR and olfactomedin protein families have not been elucidated, another source of information must be found. Rozsa et al. (5) evaluated the effects of mutations on the predicted secondary structure in several TIGR/olfactomedin-related proteins (HTIGR, MTIGR, FOLFA, ROLFA, HOLFA, and RCL2B). The predicted effects of these mutations on the secondary structure of human TIGR provide some insight into the pathogenesis of glaucoma. Likewise, our previous work examined mutations in light of the evolutionary relationship...
between the TIGR (HTIGR, MTIGR, RTIGR, and BTIGR) and oolfactomedin (COLFA, ROLFA, FOLFA, HOLFC, and MOLFA) sequences. Evaluating HTIGR mutations in terms of the conserved sequence elements across a more diverse set of sequences should provide additional understanding of the structural and functional characteristics of the ODC family. The comparison of disease-associated mutations versus non-disease-associated polymorphisms in terms of their degree of conservation as the family has evolved may help to elucidate those residues/motifs that are functionally or structurally essential.

We compared 11 polymorphisms and 27 mutations catalogued by several researchers (5, 14–23) to the fully and partially conserved residues in our alignment of the ODC C-terminal domains. Table II provides a summary of each segment of the alignment, including the fully conserved resi-
dues within the region, each variant in the region, whether the variant has been shown to be associated with glaucoma, and whether the wild-type residue is conserved (partially or fully) within our family.

Overall, mutations and polymorphisms are evenly distributed across the alignment; however, as shown in Fig. 5, they do tend to occur in clumps of variants concentrated in a particular group of residues leaving several variant-free regions within the sequence. Interestingly, the first 30% of the alignment contains only about 1% of the disease-related variants known currently. More than one-third (37%) of the mutations occur in residues that are not conserved across the family of sequences examined here. Another 37% occur in partially conserved residues, whereas the remaining 26% occur in highly or completely conserved residues.

**Identification of a Potential Distant Relative**

Our search for distant TIGR/olfactomedin relatives also identified a chitinase A enzyme from *Pseudoalteromonas* sp. Fig. 6 shows the alignment of the chitinase sequence with the three query sequences that identified the chitinase in their search output. Although the similarity is low, we decided to investigate further, because even a distant relationship to a protein with known function and structure might provide useful information about the ODC family.

Homologs were identified by querying GenBank with the chitinase sequence. Fig. 7 shows segments of the alignment of the ODC sequences with the chitinase homologs at 40% identity. Several conserved motifs are highlighted, including motifs completely conserved across the collection of ODC sequences. The partially conserved DXDXXDEGLW signature is evident in Fig. 7C. The conserved Glu-385 is not conserved within the chitinase sequences; however, the remainder of the motif is present, with the addition of a four-residue insert in the XXX region. The region in Fig. 7D includes the highly conserved Cys-433, Gly-434, and Tyr-437 residues. Cys-433 is replaced by a conserved aliphatic residue in the chitinases, whereas Gly-434 is completely conserved across both groups. Tyr-437 is partially conserved within the chitinases. Other motifs include the conserved Y(N/D) motif in Fig. 7E and the (Y/F/W)(Y/F/W)(D/N) motif in Fig. 7F.

**DISCUSSION**

In this work we have identified several new members of the olfactomedin/TIGR family. Analysis of these new sequences, which are more divergent than the previous set of olfactomedin or TIGR proteins evaluated, has provided valuable information regarding the evolutionary relationships within the family.

**Analysis of the N-terminal Region**—The addition of the new sequences to our analysis has provided additional support that the olfactomedin and TIGR proteins are related throughout the length of the molecule rather than only in the olfactomedin domain in the C terminus. In Fig. 2, the GXXC motif and the leucine-rich/leucine-zipper regions of the sequences are highlighted. All of the olfactomedin- and TIGR-related sequences possess a residual leucine zipper aligned with the leucine zipper of TIGR. The additional sequences have also helped to adjust our previous alignments (1). When we align only the TIGR and olfactomedin sequences the N-terminal region includes a (A/V)LEE(E/Y)K motif spanning residues 151 through 156 in HTIGR and 135 through 140 in MOLFA. With the addition of the more divergent sequences, that region appears to be only partially conserved among the HTIGR, COLFA, and HOLFC proteins and is significantly more divergent among the other sequences.

In addition to the support from our sequence analysis, structural analysis of the N-terminal regions provides additional support that these sequences are related. As discussed above, our Shotgun analysis revealed mostly leucine-rich/leucine-zipper sequence hits such as myosin, kinesin, and other coiled-coil proteins. Coupled with the identification of a residual leucine zipper in the N-terminal region of the olfactomedins, the analysis suggested that conserved elements might be evident in the secondary structure rather than the primary structure. Because the evolutionary relationship be-

---

4 Chitinase A, NCBI accession number AAC79665.
tween the TIGR and olfactomedin-like sequences appears to be distant, we completed predictions for each set separately, as well as in combination with the other sequences that don’t clearly fit into either category. Our secondary structure predictions confirm that the helical nature of the N-terminal region is conserved in the olfactomedin-like and TIGR-like sequences. As shown in Fig. 2, the predicted structure for both groups is predominantly helical with a limited strand region at the very beginning of the sequence.

The latrophilins, on the other hand, are mostly strand-like in the N-terminal region. These proteins are transmembrane G protein-coupled receptors with large extracellular N-terminal regions containing the olfactomedin domain. The family is also characterized by several splice variants. It is likely that the olfactomedin domain in these proteins is the only domain related to the olfactomedins. The drastically different strand nature of the region preceding the olfactomedin domain would appear to support this conclusion.

Rozsa et al. (5) used the Chou-Fasman (24) and Garnier-Osguthorpe-Rosman (25) algorithms to predict the secondary structure of the wild-type and allelic variants of TIGR with glaucoma-associated mutations. Although both Chou-Fasman and Garnier-Osguthorpe-Rosman algorithms predicted the N-terminal region of TIGR to be a mixture of helical and strand regions, it is predominantly helical. Likewise, although their C-terminal prediction contains both strands and helices, the composition is primarily strands. Several of the programs used by the Jpred server for the prediction of secondary structure have improved significantly on the accuracy of predictions over those obtained with Chou-Fasman or Garnier-Osguthorpe-Rosman algorithms. The Jpred prediction, a consensus of multiple secondary structure predictions, improves the average prediction accuracy to 72.9% (9). Based on these improvements, the differences in the prediction of the secondary structure of TIGR between the work of Rozsa (5) and our present work are not surprising.

In our previous work (1), we hypothesized that the exon 2 region of TIGR was the result of an insertion event into a distant olfactomedin relative. Our current analysis has revealed two additional pieces of evidence supporting this hypothesis. The JOLHN sequence is the first non-olfactomedin sequence we have identified that is clearly related, yet significantly divergent from TIGR in the N-terminal region (32% identity). JOLHN aligns with the exon 1 region of TIGR through the end of the leucine zipper. After a large gap (27 amino acids), the alignment resumes in the middle of exon 2. Additionally, alignment with the divergent set of olfactomedins (COLFA, HEST3, HOLF, FOLFA, and HOLF) also results in a large gap (30 amino acids) after the end of the leucine zipper. The alignment resumes for a short region in the middle of exon 2 suggesting the presence of large insertions in the exon 2 region.

Analysis of the C-terminal Domain—Our analysis of the C-terminal region of the sequences shown in Fig. 1 reveals several significant findings. The predicted secondary structure of the C-terminal regions of the family appears to be completely different from that of the N-terminal region, evidence of the multidomain structure of this family. Multiple alignments of these sequences revealed several conserved motifs that might be used to confirm additional members of this family. In particular, the motifs in segments 1, 5, and 7 and the short (Y/N)N(P/A/S) motif in segment 9 are highly conserved within the family.

Because neither functional nor structural information is available for any of the sequences included in our analysis, we used HTIGR mutation data to evaluate the significance of conserved and non-conserved residues within our collected family of sequences. From our comparison, it is possible to identify residues that may be important to some functional or structural characteristic specific to the TIGR family of proteins, as well as residues that might be essential in the overall structural or functional characteristics of the entire family of sequences. For example, as shown in Table II, Cys-433 is almost completely conserved in segment 7. The C433R mutation has been shown to be associated with glaucoma in seven of 25 unrelated Brazilian patients with juvenile-onset open-angle glaucoma (22) suggesting that some key functional or structural role has been interrupted by this substitu-
tion. Interestingly, one of the zebrafish EST sequences (ZEST2) contains an arginine aligning with Cys-433 of HTIGR. The W286R mutation, on the other hand, occurs in a residue conserved only within the TIGR and latrophilin sequences, perhaps indicating an important role for this residue in these two families. Similarly, T293K occurs in the region of an apparent insertion in the TIGR proteins that is absent from the remaining sequences. Table III provides a summary of other conserved and non-conserved HTIGR mutations that may highlight regions that are important to the overall function of the protein.

Polymorphisms in the HTIGR sequence are also summarized in Table III. Not surprisingly, over 60% of the observed polymorphisms actually occur in non-conserved residues. Most polymorphisms in conserved residues appear to be relatively conservative substitutions, e.g. V329M and V402I. The non-disease associated phenotype of these changes is consistent with the assumption that a moderate change in a single amino acid will not drastically affect the local structure or function of the molecule. Unlike the V495I, V329M, and V402I variants, S425P is a non-conservative substitution in a partially conserved motif and might be expected to impact the success of the TIGR molecule. The fact that this variant is not disease-associated may warrant further investigation to confirm its status as a polymorphism.

Another interesting result of our analysis is the identification of two large regions where disease-causing mutations in conserved residues appear to be concentrated. The clustering of mutations in residues 360 to 380 and 477 to 481 may indicate regions in which to focus efforts to further elucidate the structure/function of the ODC proteins.

The Chitinase Family—In addition to the family of ODC sequences, we have identified another family that may be functionally related, but it is unlikely that all of these sequences have the same functional characteristics. Evidence to specify the significance of the relationship between these families has not been obtained; however, given the limited availability of structural or functional information available for the ODC family, further investigation is warranted.

Although several questions regarding the ODC family remain unanswered, our study has provided important insights into the diversity within this group of related sequences. Although several of the sequences appear to contain only an olfactomedin domain, as is the case for the latrophilins and several of the EST sequences, the relationship extends outside this domain in sequences such as TIGR, the neuronal olfactomedin-like sequences, HOLFE, and HOLFF. Within the olfactomedin domain, we have identified several conserved motifs that will be useful for identifying and confirming future additions to this group of sequences and for investigating the elements necessary to maintain the structural and/or functional success of these molecules. Two regions with an increased frequency of mutations in conserved residues point to potential areas of investigation into the structural/functional operation of the ODC proteins. All of these elements should lead to a clearer understanding of the involvement of human TIGR in the pathogenesis of glaucoma and the normal function(s) of this family of sequences.

Acknowledgment—We thank Dr. Thai Nguyen for helpful comments and suggestions in the preparation of this manuscript.

* This work was supported in part by NLM, National Institutes of Health Grant NLM 07033 (to M. L. G.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Stanford Medical Informatics, Stanford University School of Medicine, 251 Campus

---

**Table III**

Summary of the significance of conserved and non-conserved mutations and polymorphisms in HTIGR

| Disease-associated? | Conserved? | Variants | Predicted significance |
|---------------------|------------|----------|------------------------|
| Yes                 | Yes        | D380G, V426F, C433R, Y437H, N480K, P481T/L, G246R, Q337R, P361S, G364V, P370L, T377M, I465M, I477N/S, I499F | Potential residues important to the overall structure or protein-protein interactions of the family may be functionally related, but it is unlikely that all of these sequences have the same functional characteristics |
| Yes                 | No         | W286R, T293K, E323K, E352K, T353I, GQ367-8V, S393R, K423E, A445V, R470C | Conservative substitutions might not affect the structure/function of TIGR |
| No                  | Yes        | V329M, V402I, S425P, V495I | Non-conserved residues not likely essential to the success of the molecule |
| No                  | No         | G252R, G367R, K398R, R422C/H, Y473C, K500R | |

---

"advertisement in accord-"
REFERENCES

1. Green, M. L., Do, H., Polansky, J. R., Nguyen, T. D., and Klein, T. E. (2001) Similarities and differences between the TIGR and olfactomedin proteins. Invest. Ophthalmol. Vis. Sci. 42, 656 (Abstr. 3531)

2. Karavanich, C. A., and Anholt, R. R. (1998) Molecular evolution of olfactomedin. Mol. Biol. Evol. 15, 718–726

3. Babbitt, P., and Gerlt, J. (1997) Understanding enzyme superfamilies. Chemistry as the fundamental determinant in the evolution of new catalytic activities. J. Biol. Chem. 272, 30591–30594

4. Nguyen, T. D., Chen, P., Huang, W. D., Chen, H., Johnson, D., and Polansky, J. R. (1998) Gene structure and properties of TIGR, an olfactomedin-related glycoprotein cloned from glucocorticoid-induced trabecular meshwork cells. J. Biol. Chem. 273, 6341–6350

5. Rozsa, F. W., Shimizu, S., Lichter, P. R., Johnson, A. T., Othman, M. I., Scott, K., Downs, C. A., Nguyen, T. D., Polansky, J. R., and Richards, J. E. (1998) GLC1A mutations point to regions of potential functional importance on the TIGR/MYOC protein. Mol. Vis. 4, 20

6. Kulkarni, N., Karavanich, C., Atchley, W., and Anholt, R. (2000) Characterization and differential expression of a human gene family of olfactomedin-related proteins. Genet. Res. 76, 41–50

7. McClure, M., Vasi, T., and Fitch, W. (1994) Comparative analysis of multiple protein-sequence alignment methods. Mol. Biol. Evol. 11, 571–592

8. Pegg, S. C.-H., and Babbitt, P. C. (1999) Shotgun: getting more from sequence similarity searches. Bioinformatics 15, 729–740

9. Cuff, J., Clamp, M., Siddiqui, A., Finlay, M., and Barton, G. (1998) JPred: a consensus secondary structure prediction server. Bioinformatics 14, 892–893

10. Thompson, J., Higgins, D., and Gibson, T. (1994) ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680

11. Murzin, A. (1998) How far divergent evolution goes in proteins. Curr. Opin. Struct. Biol. 8, 380–387

12. Orengo, C., Jones, D., and Thornton, J. (1994) Protein superfamilies and domain superfolds. Nature 372, 631–634

13. Abagyan, R., and Batalov, S. (1997) Do aligned sequences share the same fold? J. Mol. Biol. 273, 355–368

14. Adam, M. F., Belmouden, A., Binisti, P., Brézin, A. P., Valtot, F., Béchetoille, A., Dascotte, J. C., Copin, B., Gomez, L., Chaventré, A., Bach, J. F., and Garchon, H. J. (1997) Recurrent mutations in a single exon encoding the evolutionarily conserved olfactomedin-homology domain of TIGR in familial open-angle glaucoma. Hum. Mol. Genet. 6, 2091–2097

15. Alward, W. L., Fingert, J. H., Coote, M. A., Johnson, A. T., Lerner, S. F., Junqua, D., Durcan, F. J., McCartney, P. J., Mackey, D. A., Sheffield, V. C., and Stone, E. M. (1998) Clinical features associated with mutations in the chromosome 1 open-angle glaucoma gene (GLC1A). N. Engl. J. Med. 338, 1022–1027

16. Angius, A., De Gioia, E., Loli, A., Fossarello, M., Sole, G., Orzalesi, N., Grignolo, F., Cao, A., and Pirastu, M. (1998) A novel mutation in the GLC1A gene causes juvenile open-angle glaucoma in 4 families from the Italian region of Puglia. Arch. Ophthalmol. 116, 793–797

17. Fingert, J. H., Ying, L., Swiderski, R. E., Nystuen, A. M., Arbour, N. C., Alward, W. L., Sheffield, V. C., and Stone, E. M. (1998) Characterization and comparison of the human and mouse GLC1A glaucoma genes. Genome Res. 8, 377–384

18. Raymond, V., Clepet, C., Dubois, S., Winstall, E., Falardeau, P., Cote, G., Anctil, J. L., Amyot, M., Plante, M., and Morissett, J. (1997) Normal homozygotes for primary open-angle glaucoma caused by the TIGR/GLC1A gene: Evidence for a new form of dominance in man. Am. J. Hum. Genet. 61, 21 (Abstr. 101)

19. Stollova, D., Child, A., Fleck, B. W., Barsoum-Homsy, M., Ozdemir, N., Brice, G., Pitts Crick, R., and Sarfarazi, M. (1997) Identification of new TIGR mutations in juvenile-onset primary open angle glaucoma (JOAG) families from different populations. Am. J. Hum. Genet. 61, 412 (Abstr. 2411)

20. Stone, E. M., Fingert, J. H., Alward, W. L. M., Nguyen, T. D., Polansky, J. R., Sunden, S. L. F., Nishimura, D., Clark, A. F., Nystuen, A., Nichols, B. E., Mackey, D. A., Ritch, R., Kainenak, J. W., Craven, E. R., and Sheffield, V. C. (1997) Identification of a gene that causes primary open angle glaucoma. Science 275, 668–670

21. Suzuki, Y., Shirato, S., Taniguchi, F., Ohara, K., Nishimaki, K., and Ohta, S. (1997) Mutations in the TIGR gene in familial primary open-angle glaucoma in Japan. Am. J. Hum. Genet. 61, 1202–1204

22. Vasconcellos, J. P. C., Melo, M. B., Costa, V. P., Tsukumo, D. M. L., Basseres, D. S., Bordin, S., Saad, S. T. O., and Costa, F. F. (2000) Novel mutation in the MYOC gene in primary open angle glaucoma patients. J. Med. Genet. 37, 301–303

23. Zhou, Z., and Vollrath, D. (1999) A cellular assay distinguishes normal and mutant TIGR/myocilin protein. Hum. Mol. Genet. 8, 2221–2228

24. Chou, P. Y., and Fasman, G. D. (1974) Prediction of protein secondary structure from amino acid sequence. Methods Enzymol. 266, 540–553