The evolutionarily conserved Dim1 protein defines a novel branch of the thioredoxin fold superfamily

YU-ZHU ZHANG,1 KATHLEEN L. GOULD,2 ROLAND L. DUNBRACK, JR.,1 HONG CHENG,1 HEINRICH RODER,1 AND ERICA A. GOLEMIS1
1Division of Basic Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111; and 2Howard Hughes Medical Institute, Department of Cell Biology, Vanderbilt University, Nashville, Tennessee 37212

Zhang, Yu-Zhu, Kathleen L. Gould, Roland L. Dunbrack, Jr., Hong Cheng, Heinrich Roder, and Erica A. Golemis. The evolutionarily conserved Dim1 protein defines a novel branch of the thioredoxin fold superfamily. Physiol. Genomics 1: 109–118, 1999.—Dim1 is a small evolutionarily conserved protein essential for G2/M transition that has recently been implicated as a component of the mRNA splicing machinery. To date, the mechanism of Dim1 function remains poorly defined, in part because of the absence of informative sequence homologies between Dim1 and other functionally defined proteins or protein domains. We have used a combination of molecular modeling and NMR structural analysis to demonstrate that ~125 of the 142 amino acids of human Dim1 (hDim1) define a novel branch of the thioredoxin fold superfamily. Mutational analysis of Dim1 based on the predicted fold indicates that alterations in the region corresponding to the thioredoxin active site do not affect Dim1 activity. However, removal of a very short carboxy-terminal extension generates a dominant negative form of the protein [hDim1-(1–125)] that when overproduced induces cell cycle arrest in G2, via a mechanism likely to involve alteration of Dim1 association with partner molecules. In sum, this study identifies the Dim1 proteins as a novel sixth branch of the thioredoxin superfamily involved in cell cycle.

mRNA splicing; cell cycle; structure modeling

THE DIM1+ GENE WAS ISOLATED IN Schizosaccharomyces pombe based on the ability of the temperature-sensitive (ts) dim1–35 point mutant to enhance the ts phenotype of cdc2-D217N (5). In a wild-type background, dim1–35 results in a defect in sister chromatid separation, sensitization to microtubule destabilizing agents, loss of viability in mitosis, and a cut [cell untimely torn (47)] terminal phenotype. The mechanism of dim1p protein function in this process is likely to derive at least in part from its recently assigned role affecting the abundance of the anaphase-promoting complex or cyclosome (APC/C) (4), which promotes transitions through mitosis by timed proteolysis of substrates such as PDS1 (7) and cut2+ (14), Ase1 (20), and cyclin B [(23); reviewed in Ref. 19]. Deletion of dim1 in S. pombe results in lethality in G2, with cells moderately elongated but with 2C DNA contents, and interphase nuclei and microtubules, suggesting activity of the dim1p protein is required for entry into M phase (5). In a synthetic lethal screen with dim1–35, Berry et al. isolated the lid1+ (lethal in dim1–35) gene, which was demonstrated to be a novel component of the S. pombe 20S APC/C [homologous to human APC4, (48)]. lid1+ function is required for ubiquitination of the APC/C target cut2p; dim1p is required to maintain steady-state levels of lid1p and the APC/C (4).

The exact role of dim1p in regulating the abundance of the APC/C is not yet clear. One intriguing possibility is that Dim1 may control function of the APC/C, and possibly other proteins required for mitotic entry, at the level of mRNA splicing, because the Saccharomyces cerevisiae dim1p homolog, DIB1, was recently found to be stoichiometrically associated with the U4/U6.U5 tri-SNRP (40). Previous studies in S. pombe (e.g., Refs. 28 and 34) have identified genes such as prp8+ that possess dual function in control of mRNA splicing and mitosis; the most simple model for the mechanism of such a connection would be the existence of spliced transcripts required for progression from G2 through mitosis. The dim1p protein may alternatively or additionally be more directly involved in control of mitotic events; we have previously cloned the human dim1p homolog, hDim1, on the basis of direct physical interaction with HEF1, a human protein hypothesized to function in communicating information about cellular attachment status to the apparatus governing progression through mitosis (25, 26). Elucidation of dim1p/hDim1 function in either of these processes promises to yield intriguing biological insights.

Although the Dim1 proteins are evolutionarily highly conserved, based on inspection of amino acid sequence and cross-species functional complementation (5), standard sequence homology screening programs detect only marginal similarity with proteins outside the set of Dim1 homologs from different organisms. To gain insight into Dim1 activity, we therefore undertook a structure-based analysis. Using protein-threading search algorithms to probe the Protein Databank of experimentally determined protein structures, we unexpectedly found that the bulk of the Dim1 protein was strongly predicted to adopt a thioredoxin fold. This was confirmed by a detailed analysis of the secondary structure of a truncated form of hDim1 (residues 1–128) using multidimensional NMR methods. The results showed that the experimentally determined secondary structure, the positions of hydrophobic core residues in the sequence, and the positions of insertions in the Dim1 sequence generally correspond to patterns in the thioredoxin superfamily of proteins. Based on this preliminary structural model and complementary biological experiments, we have identified the carboxy-
terminal ~14 amino acids of Dim1 that extend beyond the thioredoxin homology region as an essential sequence, which when deleted results in a dominant negative form of hDim1 that induces cell cycle arrest in G2. In a series of functional tests, we determined that this region does not affect hDim1 protein localization, steady-state levels, or phosphorylation, suggesting a role for the carboxy-terminal residues in mediating Dim1 protein interactions with key partners or substrates. These results define the Dim1 protein group as a novel branch of the thioredoxin superfamily that contributes to cell cycle control and provide a guide for future functional analyses of the Dim1 proteins.

MATERIALS AND METHODS

Molecular modeling and sequence analysis. Programs used to search for amino acid sequence homologies of Dim1 include BLAST (1) and PSI-BLAST (2). Initial structure predictions of Dim1 were made using the UCLA-DOE Fold Recognition Server (http://www.doe-mbi.ucla.edu/people/frsvr/frsvr.html) (10). Additional protein secondary structure predictions were made using the programs PHD (36), DSC (22), nnpredict (24), and SIMPA (27). For display of the Dim1 model with noted evolutionary variance, the program RasMol (37) (version 2.6 for Macintosh) was used.

Plasmids. The hDim1 cDNA used as the basis of constructs described herein was that isolated in Ref. 26. This cDNA was PCR amplified with EcoR I and Xho I added to the 5’ and 3’ end, respectively. The PCR product was digested with EcoR I and Xho I to release an insert containing the full-length hDim1 coding sequence, which was used for cloning into the EcoR I- and Xho I-digested vector pBlueScript (Stratagene) to create pBlue-hDim1, the yeast two-hybrid LexA-fusion “bait” vector pEG202 to create pEG202-hDim1 (15), the yeast two-hybrid activation domain-fusion vector pJG4-5, to create pJG4-5-hDim1(18), the green fluorescent protein (GFP)-fusion mammalian expression vector pEGFP-c4 to create pEGFPC4-hDim1 (pEGFP-c4 is a derivative of pEGFP-c1 (Clontech) possessing an altered polynucleotide and the mammalian expression vector pcDNA3.0 (Invitrogen) to create pcDNA3-hDim1.

For expression in S. pombe, pMNS21L-hDim1 was constructed by oligonucleotide-directed PCR mutagenesis to destroy an hDim1 internal BamH I site and insert the hDim1 coding sequence into the Nde I-BamH I cut vector. For overexpression of the hDim1 proteins for NMR studies, hDim1 was PCR amplified from pBlue-hDim1 with flanking Nde I and Blp I sites and recloned into Nde I- and Blp I-restricted vectors pET19b (Novagen) to produce pET19b-hDim1 and pET29b (Novagen) to produce pET29b-hDim1. The expression of hDim1p from both vectors is controlled by the T7 promoter and lac operator.

PCR mutagenesis was used to construct truncated and mutated forms of hDim1p or S. pombe Dim1p described in this study; these were expressed from the pET29b, pEGFP-C4, pMNS21L, pEG202, and pJG4-5 vectors as described above for full-length hDim1.

Protein expression, purification, and labeling. pET29b-hDim1 or pET29b-hDim1(1–128) was transformed into BL21(DE3) E. coli (Novagen). Production of the protein was induced by 1 mM isopropyl-β-D-thiogalactopyranoside for 3 h when the bacteria were grown to OD600 = 1.0–1.2. Unlabeled protein was isolated from bacteria grown in LB medium. 15N-labeled and 13C-labeled proteins were purified from bacteria grown in M9 minimal medium supplemented with 0.15 mM thiamine. For selective incorporation of 15N-labeled valine or lysine, M9 minimal medium containing a mixture of L-amino acids (6) was used.

For hDim1p purification, the cell pellet of a 1-liter culture was resuspended and lysed by sonication. The protein was purified to near homogeneity by three-step fast protein liquid chromatography, Ni-chelating, gel filtration, and Mono Q anion exchange. His-tag hDim1p was purified by the same protocols with minor modification (see RESULTS).

NMR data collection and analysis. NMR spectra were recorded at 37°C on a Bruker DMX-600 spectrometer equipped with a 5-mm x, y, z-shielded pulsed-field gradient triple-resonance probe. All samples for the NMR experiments contained 10 mM phosphate buffer (pH 7.65, uncorrected for isotope effect), 0.2% 2,2-dimethyl-2-silapentane-5-sulfonate (DSS), 0.04% NaN3, and either 5% or 100% H2O. The protein concentration was 1.5 mM. NMR data were processed and analyzed using FEILIX97 (Molecular Simulations). 1H chemical shifts was referenced to DSS, and 13C and 15N chemical shifts were indirectly referenced to DSS (8).

A series of three-dimensional experiments [1H-15N-HSQC (33), NHCO (17, 21), HNCA (17, 21), HNCA/CB (32, 44), and CBCA(CO)NH (16, 32)] were recorded for backbone and Cα resonance assignments. 15N-HSQC was also recorded on selectively 15N-labeled samples [15N]valine or [15N]lysine.

Mass spectrometry. Mass spectrometry (MS) was performed using a LCG quadrupole ion trap liquid chromatography (LC)-MS (Finnigan) and matrix-assisted laser desorption/ionization (MALDI)-time-of-flight (TOF) (PerSeptive Biosystems), with the help of Steven Seeholzer at the Fox Chase Biotechnology Core Facility.

Yeast methods, strains, and media. The S. pombe strains used in this study were leu1-32 ura4-D18 h- and dim1-35 leu1-32 ura4-D18 h-. Strains were grown in yeast extract medium or minimal medium with appropriate supplements (31). To regulate expression of hDim1 cDNAs from the thiamine-repressible nmt1 promoter in the vector pMNS21L (30), cells were grown in the presence or absence of 5 μM thiamine. Transformations were performed by electroporation (35).

Flow cytometry and yeast microscopy. For flow cytometric analysis, cells were fixed in ice-cold 70% ethanol, washed in 50 mM sodium citrate, incubated with 0.1 mg/ml RNase A in 50 mM sodium citrate for 2 h at 37°C, and then stained with 1 μM Sytox green in 50 mM sodium citrate, incubated with 0.1 mg/ml RNase A in 50 mM sodium citrate for 2 h at 37°C, and then stained with 1 μM Sytox green in 50 mM sodium citrate for 2 h at 37°C. Cells were sonicated and analyzed by flow cytometry as described previously (38). To visualize microtubules, cells were fixed in 100% methanol at −20°C for 8 min, washed with phosphate-buffered saline, and processed as described previously (3). Fixed cells were incubated with a 1:20 dilution of the monoclonal TAT-1 antibody (45) followed by a 1:100 dilution of Alexa-conjugated goat anti-mouse secondary antibody (Molecular Probes, Eugene, OR). Yeast fluorescence microscopy was performed on a Zeiss Axioskop with appropriate filters.

Antibodies. All antibody production was done in rabbits at the Laboratory Animal Facility at Fox Chase Cancer Center. Antibody NK16 was raised against a keyhole limpet hemocyanin-conjugated synthetic peptide (GNNKINWAMEDKQEC) derived from hDim1p and was purified chromatographically with an affinity column produced by coupling the peptide antigen to Affi-Gel 15 (BioRad). Antibody 3P4 was raised against bacterially expressed hDim1p purified as described above and was affinity purified with purified hDim1p coupled to AffiGel15 (BioRad).

Mammalian cell transfections and immunofluorescence. Transfections were done using Lipofectamine Plus (Life Technology) to introduce mammalian expression constructs encod-
ing hDim1, GFP-hDim1, hDim1-(1–128), or GFP-hDim1-(1–128) into HeLa cells. Expression levels and size of produced proteins were determined by harvest of whole cell lysates, SDS-PAGE separation of protein species, and Western analysis using antibody to hDim1p as primary, and an enhanced chemiluminescence kit (Amersham) for visualization, using standard protocols. Cells were processed for immunofluorescence as described (26), using antibody to hsDim1 as primary antibody and biotinylated anti-rabbit IgG as secondary; or alternatively by direct observation of GFP after fixation. Observation and image capture of cells were performed using a BioRad MRC 600 confocal microscope.

RESULTS

Extension of the Dim1 family, and structure prediction of Dim1 proteins as new members of the thioredoxin fold superfamly. To gain insight into function of the hDim1 protein, we first utilized standard sequence homology and motif-based programs to probe the sequence of members of the evolutionarily conserved Dim1 family. As a first step, we wished to maximize the number of known Dim1 orthologs in the analysis set. The initial study of Dim1 identified orthologs for the gene in seven species (5). Searching the est database with the tblastn version of BLAST to locate other members of the Dim1 family, we have extended this protein group to include putative Dim1 sequences with sequence identity greater than 65% in 18 species, including three mammalian species (human, rat, mouse), seven plant species (rice, apple, aspen, cotton, corn, Arabidopsis, and soybean), three fungi (S. cerevisiae, S. pombe, and Magnaporthe grisea), two nematodes (Caenorhabditis elegans and Onchocerca volvulus), one slime mold (Dictyostelium discoideum), one platyhelmint (Schistosoma mansoni), and one alveolate (Plasmodium falciparum). We also found three additional sequences (human, mouse, and aspen) that were more closely related to Dim1 than to any other sequence group and seem to form a separate Dim1-like family. This family has been designated the “Dim2” group. The human Dim2 sequence is 37% similar to human Dim1, whereas the human and mouse Dim2 sequences are 97% identical. Alignment of these proteins is shown in Fig. 1.

As none of the Dim1 or Dim2 proteins possessed a defined biochemical function, it was of particular interest to identify significant homologies to proteins outside the Dim1/2 sequence group. BLAST searches initiated with S. pombe, H. sapiens, or S. cerevisiae homologs of Dim1p isolated all known family members with extremely high expectation (E) values (1 × 10^-46). For all members of the Dim1 family, the next nearest match in homology was to several members of the thioredoxin protein family, but with much lower statistical significance (E = 0.004–2.0, above the standard window for inclusion in subsequent rounds of PSI-BLAST). Sequence identities observed between Dim1 proteins and their most homologous thioredoxin relatives ranged from 19% (S. pombe) to 27% (S. cerevisiae) sequence identity over 86 amino acids; by these criteria, a relationship between Dim1 and thioredoxin appeared marginal.

In contrast, use of the UCLA-DOE Structure Prediction Server (http://www.doe-mbi.ucla.edu/people/frsvr/frsvr.html) search program incorporating protein structural information (10) to probe the experimentally determined structures deposited in the Protein Databank resulted in extremely strong predictions of a match between each of the Dim1 protein sequences and multiple thioredoxin structures, with the highest prediction in each case for a fold similar to 3trx (12) (Z value 8.6–10.5). Strikingly, optimum threading of the Dim1 proteins against thioredoxin sequence required minimal gapping in either sequence and extended over the complete length of the thioredoxin sequence (Fig. 1, lines 23 and 24 of each alignment group). This strong prediction for Dim1 structure was reiterated by analysis of Dim1 protein sequence with a variety of other programs using different algorithms to predict secondary structure. The locations of predicted helices and β-sheet segments in Dim1 corresponded very closely to the positions of helices and β-sheet segments in thioredoxin, given the alignments provided by BLAST and by the UCLA server.

This predicted structural fold was of interest for two reasons. First, if valid, it provides a structural frame for Dim1 that might serve as a guide for mutational analysis and functional studies in vivo. Second, all of the previously characterized members of the thioredoxin fold superfamily, including thioredoxin, glutaredoxin, DsbA, glutathione-S-transferase, and glutathione peroxidase, are involved in cellular redox control or in glutathione-related detoxification processes via interaction with thiol- or disulfide-containing substrates (reviewed in Ref. 29). The Dim1 proteins would define a sixth branch of this fold group, which is structurally closely related to the thioredoxin family, but which has a different function related to the regulation of cell cycle progression.

NMR determination of secondary structure of hDim1 confirms predicted thioredoxin homology. To investigate the three-dimensional structure of hDim1p and its structural relatedness to thioredoxin, we overexpressed the hDim1 protein in bacteria. Purified hDim1p was monomeric at millimolar concentrations as determined by sedimentation equilibrium measurements (data not shown). Temperature denaturation of the protein did not occur until ~70°C. Mass spectrometric analysis of prepared samples indicated that the carboxy-terminal 14 residues were cleaved slowly after the protein was purified (data not shown). Part of the carboxy-terminal peptide was also identified by MS-MS spectroscopy, suggesting the cleavage represented action of a specific endoprotease rather than nonspecific degradation of termini. The cleaved form of Dim1, confirmed by mass spectrometry to correspond to a mixed population of Dim1(1–128) > Dim1(1–129), was stable to subsequent degradation over extended periods; intriguingly, the cleavage truncated hDim1 immediately after the predicted thioredoxin homology region. These results suggested the independent evaluation of full-length hDim1 and hDim1(1–128) would be of interest. Accordingly, clones expressing full-length
hDim1 and hDim1-(1–128) were expressed in bacteria, purified, and analyzed by multidimensional NMR methods. Although both forms showed essentially the same cross peak pattern in $^1$H-$^{15}$N correlated spectra (with the exception of peaks for the deleted residues), hDim1-(1–128) had narrower lines and much better signal-to-noise ratios, especially in three-dimensional experiments. This truncated form was chosen for further structural studies.

Resonances of Ca, CO, and Ha of the polypeptide backbone are known to depend strongly on the secondary structure (42, 43). In Fig. 2, the chemical shift indexes of these atoms for 107 assigned residues (from a total of 128) determined by the procedures previously described.
described (42, 43) are plotted against residue numbers. A consecutive run of four or more "1" values in the CO/C\textsubscript{a} index or "2" values in the H\textsubscript{a} index indicates helical conformation and a consecutive run of three or more "1" values in CO/C\textsubscript{a} index or "1" values in H\textsubscript{a} index indicates \(\beta\)-strand structure. As shown, the secondary structures determined by C\textsubscript{a}, CO, and H\textsubscript{a} chemical shift indexes are generally in agreement.

The consensus results of secondary structures for hDim1-(1–128) agree with the prediction that hDim1 assumes a thioredoxin fold. Specifically, of the four \(\alpha\)-helices in thioredoxin, three (the 1st, 2nd, and 4th) are clearly confirmed by the NMR results on hDim1-(1–128) (residues 11–20, 40–51, and 110–123). Because of incomplete backbone assignment in the relevant region (residues 66–72), chemical shift indexes were ambiguous as to whether hDim1 possessed an additional helix corresponding to the third helix in thioredoxin. The identification of \(\beta\)-strands at hDim1 residues 5–7, 24–30, 55–62, 75–84, and 88–94 based on chemical shift indexes and nuclear Overhauser effect (NOE) pattern is also in agreement with its alignment with thioredoxin. The observed \(\alpha\)-helices and \(\beta\)-strands of hDim1 aligned to those found in the NMR structure of thioredoxin (3trx) indicated an insertion of 10 amino acids in hDim1 (residues 99–108). Of note, this insertion coincides with a location previously reported to be tolerant for insertions in other thioredoxin family members (29).

Finally, \(\alpha\)-helices and \(\beta\)-strands have distinctive NOE patterns; analysis of these provides a separate measure of hDim1 secondary structure. The short- and medium-range NOE connectivities for hDim1-(1–128) (Fig. 2) are in good agreement with the secondary structures suggested by the chemical shift indexes, providing independent confirmation of these results. In particular, contiguous stretches of strong NOEs between adjacent NH protons \([d_{NN}(i,i+1)]\) and numerous medium-range NOEs between the \(\alpha\)-proton of residue \(i\) and the amide proton of residues \(i+3\) and \(i+4\) indicate the presence of \(\alpha\)-helical structure for residues 12–21, 37–51, and 110–123. These segments correspond closely to the regions assigned to helices 1, 2, and 4 on the basis of the chemical shift analysis. On the other hand, stretches of strong sequential NOEs are observed between \(\alpha\)- and amide-protons \([d_{NN}(i,i+1)]\) but no medium-range NOEs were observed, confirming the assignment of \(\beta\)-strands for residues 27–31, 58–62, 79–82, and 89–94.

Although detailed solution of the hDim1 structure is in progress (Zhang, Cheng, Golemis, and Roder, unpublished results), completing a highly refined structure takes considerable time. In the interim, the combination of initial secondary structure analyses of the chemical shifts and sequential NOE connectivity of hDim1 with molecular modeling results provided considerable confidence in the hypothesis that Dim1 proteins possess a folding topology closely related to that of thioredoxins. Based on this prediction, it was therefore possible to develop biological tests of hDim1 activity in advance of the completed structure.
Genetic and cell biological tests of structure-based mutations in hDim1 identify a carboxy-terminal regulatory region. As these results strongly supported the idea that the bulk of Dim1 protein sequence (residues 1–128) folds into an amino-terminal thioredoxin-like domain, one issue to consider was whether Dim1 proteins possessed any catalytic activity similar to thioredoxins. Although the CGPC motif required for thioredoxin catalytic activity is not found in Dim1, a DxxC (amino acids 35–38 in hDim1) is essentially invariant through the entire Dim1/Dim2 family, leaving open the possibility that residues in the region of the thioredoxin catalytic pocket might be important for Dim1 activity. Intriguingly, the short carboxy-terminal region extending beyond the thioredoxin homology domain was similarly highly conserved, suggesting that residues 129–142 might encompass an important independent activity or regulatory function for the protein. If regulatory, one attractive possibility was that these sequences might control hDim1 phosphorylation: inspection of the Dim1 family of proteins in this region indicates a well-conserved cluster of potential phosphorylation sites, SPxxYST. The SP site was of particular interest, as its flanking sequences are similar to those found optimal for recognition by mitosis-specific kinases and prolyl isomerases (46), which would be relevant to the dim1 phenotypes identified in mitosis (4, 5). To evaluate these possibilities, mutant forms of dim1p and hDim1 were expressed in S. pombe, either in a dim1–35 mutant background to examine capability to rescue or in a wild-type dim1+ background to test for dominant negative function (Figs. 3 and 4; Table 1).

The nmt promoter used for hDim1 expression is thiamine regulatable, producing high levels of hDim1 and derivatives in the absence of thiamine, and low levels in the presence of thiamine (11). dim1–35 mutant strains are viable at 25°C, semiviable at 32°C, and inviable at 36°C; wild-type dim1+ parental strains are viable from 25–37°C and grow optimally at 30–32°C (5). As previously described, both the dim1–35 and dim1–129 null mutations are completely complemented by wild-type mammalian forms of Dim1 (5). At 25°C, in the presence of thiamine, dim1–35 strains containing empty thiamine-regulatable vector, or vector encoding wild-type hDim1, hDim1-C38A (eliminating any homology to a thioredoxin active site), dim1p-D35C, E36G, E37P (introducing a consensus thioredoxin CGPC active site), hDim1–ΔD35C or hDim1–ΔD35C, Dim1–ΔD35C, E36G, E37P derivatives were viable, indicating complementation, whereas those expressing the two truncated forms, hDim1–Δ128 or hDim1–Δ129, were not (Fig. 3A and data not shown). At 36°C, dim1–35 yeast expressing hDim1 or the hDim1-C38A or hDim1-D35C, E36G, E37P derivatives were viable, indicating complementation, whereas those expressing the two truncated forms, hDim1–Δ128 or hDim1–Δ129, were not (Fig. 3B and data not shown). Strikingly, the two truncated forms of hDim1 also demonstrated a dominant negative effect: in the absence of thiamine, either in dim1–35 yeast grown at 25°C or wild-type yeast grown at 32°C (Fig. 3, C and D, respectively), expression of either hDim1–Δ128 or hDim1–Δ129 resulted in lethality.

The lethality induced by truncated hDim1 could be nonspecific, with cells arrested throughout the cell cycle, or could parallel the dim1–35 hypomorph-induced arrest (cells with a failure of chromosome separation, and a cut phenotype), or could reflect a complete block of Dim1p function by a dominant negative mechanism, resulting in a dim1– null phenotype (arrest in G2) (5). To discriminate among these possibilities, wild-type yeast expressing either empty vector or hDim1–Δ128 were maintained for 20 h at 32°C, then analyzed by FACS (Fig. 4A). Yeast with hDim1–Δ128 were arrested with a uniform 2N (G2) DNA content.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>hDim1-C38A</td>
<td>Same as wild type</td>
</tr>
<tr>
<td>Dim1-(D35C,E36G,E37P)</td>
<td>Same as wild type</td>
</tr>
<tr>
<td>hDim1-(1–128)</td>
<td>Dominant negative</td>
</tr>
<tr>
<td>hDim1-(1–129)</td>
<td>Dominant negative</td>
</tr>
</tbody>
</table>

Library of mutants and phenotypes for hDim1 in Schizosaccharomyces pombe
Additional direct immunofluorescence analysis of hDim1–D128 yeast stained with 4′,6-diamidino-2-phenylindole or antibodies to α-tubulin indicated interphase nuclei and microtubule networks, confirming that no yeast had entered M phase, supporting assignment of G2 arrest (Fig. 4B). Thus removal of the carboxy-terminal sequences of hDim1 extending beyond the thioredoxin fold homology results in the creation of a biologically inactive protein, suggesting that this extension may regulate the core activity of Dim1.

The obvious possibilities to explain the role of the carboxy-terminal residues of hDim1 in regulation would include roles in specifying hDim1 protein phosphorylation, localization, stability, and association with partner molecules. A combined mutational and biochemical approach evaluating the S/T/Y sites encompassed in residues 129–142 of hDim1 has so far indicated this region is not subject to phosphorylation (results not shown). To investigate the roles of residues 129–142 in control of hDim1 intracellular localization, hDim1 and hDim1-(1–128) were transfected into cells to be expressed either as untagged native proteins or as GFP fusions. Observation of localization of hDim1 and hDim1-(1–128) revealed that both protein forms were predominantly nuclear, with a minor component of protein in the cytoplasm, in all cells examined (Fig. 5 and data not shown), arguing against the possibility that the carboxy-terminal sequences grossly altered cellular localization. Finally, in transient transfection assays in which equivalent frequencies of cells were transfected with hDim1 and hDim1-(1–128), comparison by Western blot with α-hDim1 of the steady-state levels of the two forms of protein indicated that both were present at identical levels when assessed at 15 and 42 h after transfection, indirectly suggesting that removal of the 129–142 residues did not significantly affect protein stability (data not shown).

DISCUSSION

Through combined application of molecular modeling, structural determination, and genetic and cell biological analyses, we have established that Dim1 proteins are predominantly composed of an amino-terminal thioredoxin fold module whose activity is governed by a short carboxy-terminal extension, via a mechanism that is likely related to regulation of protein-protein interaction. These results identify a new branch of the thioredoxin fold superfamily family whose structure is highly conserved with thioredoxins; however, in contrast to other superfamily members such as peroxiredoxins, which also maintain similar protein function (39), protein function is apparently quite divergent for the Dim1 group. In general, thioredoxins control stress responses through catalysis of redox reactions. Dim1 proteins regulate G2 progression in S. pombe and S. cerevisiae, via a mechanism that may involve mRNA splicing (40); the region of Dim1 corresponding to the thioredoxin active site is a point of allowed variance in the group, unaffected by directed mutation. These points would define the Dim1 proteins as “topological cousins” (9), rather than functional close relatives, of the thioredoxin group. However, several lines of evidence suggest that the maintenance of the core thioredoxin fold, and its interaction with flanking carboxy-terminal sequences, is particularly important for appropriate Dim1 function. These include the following points.

First, the purified hDim1 protein is cleaved in vitro from a full-length to a Dim1-(1–128) form that elimi-
nates carboxy-terminal sequences beyond the fold homology; on the basis of its resistance to further truncation, this truncated form of hDim1 has significantly greater stability than the full-length form. Second, the naturally occurring \textit{dim1–35} mutation, which results in a hypomorphic form of Dim1 (5), is a single glycine-to-aspartic acid substitution at amino acid 126, at the boundary of the fold-homology region and the carboxy-terminal extension. Third, expression of a form of hDim1 truncated to the Dim1-(1–128) form in \textit{S. pombe} not only fails to rescue a \textit{dim1–35} strain but additionally acts as a dominant negative with a phenotype similar to the \textit{dim1} null mutant. Dominant negative mutations are relatively rare, and such a result indicates that an uncontrolled maintained function of the Dim1-(1–128) fragment is lethal or at minimum indicates that this fragment is independently able to interact with high affinity with key Dim1 partners, blocking their essential activity. Work in this study does not allow definite assignment of function to the short 129–142 region of hDim1, although this region does not appear likely to function in control of Dim1 localization, stability, and potentially phosphorylation (Fig. 5, and data not shown). A reasonable supposition is that this region governs the associative properties of hDim1-(1–128), causing it to fail to release from particular molecular partners, sequestering such required partners from dim1p in \textit{S. pombe}. This model can be investigated with reagents developed in this study. One interesting point to pursue is the question of whether the cleavage event observed in vitro, resolving full-length Dim1 to a Dim1-(1–128) form, occurs in vivo, and whether transformation of Dim1 between full-length and truncated forms plays a role in its function.

Because of the considerable colinearity of the thioredoxin and Dim1 family structures, it is possible to develop a three-dimensional model of the thioredoxin-homologous component of Dim1 based on the experimentally derived thioredoxin coordinates deposited in the Protein Databank (entry 3trx (12)), which can in turn be utilized to gain insight into three-dimensional properties of hDim1 required for biological activity. As shown in Fig. 6, the pattern of evolutionary variance for the Dim1 family group clearly reveals that the solvent-exposed residues of Dim1 do not possess equivalent freedom to diverge; rather, particular regions of the surface are essentially invariant over one billion years of evolution, as estimated from the separation of yeast and humans. Notably, the residues in the region corresponding to the thioredoxin active site (amino acids 35–38) are not in a particularly conserved patch, reinforcing the observation that mutations in this region are not important for protein function (Table 1). Given that characterization of Dim1 to date by biochemical and two-hybrid means (results not shown) argue strongly against the idea that Dim1 homodimerizes, the conserved surface patches thus revealed seem likely to interact with evolutionarily conserved partner elements and can be specifically targeted for detailed genetic and functional analysis. The considerable extent of the conserved solvent-exposed regions identified (Fig. 6A) supports the idea that hDim1 may be a central component of a large functional complex; based on the recent identification of Dim1 as a component of the \textit{S. cerevisiae} U4/U6.U5 small nuclear riboprotein particle (40), it will be of particular interest to examine Dim1 mutations targeted to the conserved regions for effects on cell cycle-regulated proteolysis and splicing.

In considering the possible essential function of Dim1 proteins in conjunction with the tri-snRNP, a probe of the PROSITE MOTIFS database indicates that hDim1 possesses a match to a signature sequence...
highly predictive of an FKBP-type peptidyl-prolyl cis-trans isomerase, (LIVMC)x(YF)xGVL|x1,2(LFT)x2Gx3(DEX)STAEK(x)STAN, with match to sequences 1–18 of Dim1. Inspection of the conservation of this motif through the Dim1 family group (Fig. 1) and inspection of the display of the conserved residues overlaid on the Dim1/thioredoxin template fold (Fig. 6) allow determination, first, that the implicated residues are highly conserved throughout the entire group, and, second, that they localize in a single cluster on the molecule surface. Together, these observations support the significance of the motif-based assignment. Activity of the tri-snRNP involves significant conformational changes during the splicing process, and it has been proposed that prolly isomerases might contribute to such processes (41); in this light, the modeling data provided here suggest it would be attractive to evaluate the Dim1 protein group for such a function. In sum, this study has utilized genomic prediction-based techniques to allow identification of physiologically important elements of the Dim1 protein family in advance of the completion of a protein structure for this novel but clearly important protein group. The ability to rapidly test model-based predictions in genetically tractable organisms highlights the potential power of utilizing evolutionary and structural database analysis in support of traditional genetic and cell biological studies.

We are extremely grateful to Ying Tong Wang and Anna Feoktistova for excellent technical assistance. Jonathon Boyd provided expert assistance with microscopy. We thank Steven Seeholzer for performance of mass spectrometry to assign Dim1 cleavage products, and Tom Coleman for the generous gift of synchronized Xenopus oocyte extracts. We are very grateful to Bob Perry and Sarah Fashena for rigorous review of the manuscript.

This work was supported by National Institutes of Health (NIH) Grants RO1-CA-63366 (E. Golemis) and RO1-GM-47728 (K. L. Gould), NIH Core Funds CA-06927, an appropriation from the Commonwealth of Pennsylvania (to Fox Chase Cancer Center), NIH Training Grant T32-CA-09035 (Y.-Z. Zhang), and American Cancer Society Fellowship PF-98–290–01-CSR (Y.-Z. Zhang). The NMR facility is supported in part by a grant from the Kresge Foundation. K. L. Gould is an Associate Investigator of the Howard Hughes Medical Institute.

Address for reprint requests and other correspondence: E. A. Golemis, W422, Fox Chase Cancer Center, 7701 Burholme Ave., Philadelphia, PA 19111 (E-mail: EA_Golemis@fccc.edu).

REFERENCES

27. Lundgren, K., S. Allan, S. Urasihyama, T. Tani, Y. Ohshima, D. Freundeway, and D. Beach. A connection between pre-mRNA splicing and the cell cycle in fission yeast: cdc28+ is allelic...


