Structure, Stability, and Function of hDim1 Investigated by NMR, Circular Dichroism, and Mutational Analysis†

Yu-Zhu Zhang,*,‡,⊥ Hong Cheng,‡ Kathleen L. Gould,§ Erica A. Golemis,*,‡ and Heinrich Roder*,‡

Division of Basic Science, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111 and the Howard Hughes Medical Institute and Department of Cell and Developmental Biology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232

Received March 26, 2003; Revised Manuscript Received June 9, 2003

ABSTRACT: The 142 amino acid Dim1p protein is a component of the U4/U6·U5 tri-snRNP complex required for pre-mRNA splicing and interacts with multiple splicing-associated proteins. To gain further insight into the structural basis of its function, we determined the solution structure of the reduced form of the dominant negative human hDim1 ($hDim1_{1-128}$) using multidimensional NMR spectroscopy. This dominant negative hDim1 assumes a thioredoxin-like fold, confirming previous NMR and crystallographic results. However, in contrast to a recent crystal structure, the NMR solution structure for the reduced form of $hDim1_{1-128}$ presented here, along with thermodynamic data, indicates that the presence of a disulfide bond between Cys38 and Cys79 is structurally and functionally unimportant. Comparison of the truncated $hDim1_{1-128}$ with the full-length protein, using NMR and circular dichroism spectroscopy, indicates that the 14 C-terminal residues can undergo a local unfolding transition and assume alternative conformations, which appear to play a functional role. Other residues essential for hDim1 function are identified by using mutational and genetic approaches. The residues thus identified are not identical with those previously shown to govern Dim1 interaction with defined protein partners.

The Dim1 protein is extraordinarily well conserved throughout the eukaryotic kingdom, with 79% sequence identity over its entire length of 142 amino acids and full complementation function maintained between human (hDim1) and *Schizosaccharomyces pombe* (Dim1p) orthologs (*1*, *2*). Dim1 functions at multiple levels in control of pre-mRNA processing. Dib1p, the *Saccharomyces cerevisiae* Dim1p ortholog, has been identified by mass spectrometry following purification as a component of the pre-mRNA splicing machinery in two independent studies aimed at identifying novel elements of the *S. cerevisiae* U4/U6·U5 snRNP (*3*, *4*). Dib1p was found to be required for the splicing of a non-messenger RNA, the U3 RNA (*4*). Elimination of expression of the Dim1 ortholog *dml-1* in *Caenorhabditis elegans* by RNA interference leads to embryonal lethality during gastrulation with an arrest phenotype consistent with global disruption of zygotic gene expression and compatible with a global defect in pre-mRNA splicing (*5*−*7*). These results suggested an essential, highly conserved function for Dim1 family proteins in eukaryotes.

Analyses of the interaction properties of Dim1 proteins have provided direct connections to the pre-mRNA splicing machinery. It has been shown that hDim1 interacts with hnRNP F and hnRNP H′ (*5*), proteins that function in tissue-specific enhancement of pre-mRNA splicing (*8*, *9*) and contain RNA recognition motif (RRM)-like sequences that confer the ability to directly interact with poly(rG) (*10*). hDim1 also interacts with Npw38/PQBP-1 (*5*, *11*, *12*). This protein possesses a specific RNA binding activity for poly-(rG) (*11*), colocalizes with the SRm160/SRm300 splicing co-activators when coexpressed with Dim1 (*5*, *13*, *14*), and under some circumstances, regulates RNA polymerase II-dependent transcription (*15*). Finally, Dib1p has been found to interact strongly with Prp6p (*16*), a protein required for the accumulation of U4/U6·U5 tri-snRNP (*17*, *18*). Together, these studies suggest that hDim1 functions at multiple control points in the splicing of pre-mRNAs, potentially as part of a large spliceosomal complex involving numerous protein−protein interactions. Which of these contacts is critical for the essential function of Dim1 is not yet clear.

Past studies by our group using NMR[1] coupled with molecular modeling (*2*) and by others using X-ray crystallography (*19*) have indicated that hDim1 adopts a thioredoxin fold with a C-terminal (14 amino acid) extension. Analysis

† This work was supported by NIH Grants RO1 CA63366 (E.G.) and RO1 GM056250 (H.R.), Core Grant CA-06927 (to Fox Chase Cancer Center), an appropriation from the Commonwealth of Pennsylvania (to Fox Chase Cancer Center), NIH Training Grant T32 CA09035, ACS Fellowship PF-98-290-01-CSM (Y.Z.), and the HHMI (K.L.G.). The NMR facility was supported in part by a grant from the Kresge Foundation.

* Corresponding authors. Tel.: 312-567-3484; fax: 312-567-3494; e-mail: zhangy@iit.edu (Y. Z.). Tel.: 215-728-2860; fax: 215-728-3616; e-mail: EA_Golemis@fccc.edu (E. A. G.). Tel.: 215-728-3123; fax: 215-728-3574; e-mail: H_Roder@fccc.edu (H. R.).

‡ Fox Chase Cancer Center.

§ Vanderbilt University School of Medicine.

⊥ Present address: Department of Biological, Chemical, and Physical Sciences, Illinois Institute of Technology, Chicago, IL 60616.

[1] Abbreviations: NMR, nuclear magnetic resonance; TOCSY, total correlation spectroscopy; HSQC, heteronuclear single-quantum correlation; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; rmsd, root-mean-square deviation; DTT, dithiothreitol; GuHCl, guanidine hydrochloride.

10.1021/bi034486i CCC: $25.00
Published on Web 07/26/2003

of amino acid conservation throughout evolution and mutational studies (*2*) indicates that the ability to take part in redox reactions through formation of a disulfide bond is probably not important for in vivo Dim1 activity, arguing against a simple analogy to thioredoxin. However, deletion of the C-terminal extension to make $hDim1_{1-128}$ results in the creation of a dominant negative form of Dim1 (*2*). This suggests that regulation of the activity of the thioredoxin-like core by the C-terminal extension may be important for Dim1 function. The functional importance of the C-terminal extension is further supported by the fact that the *dim1-35* mutant initially isolated in *S. pombe* (*1*), which contains a point mutation at the junction between this extension and the thioredoxin-like core (G126D), manifests a temperature-sensitive phenotype leading to loss of cellular viability. However, at present, the role of this extension for Dim1 function is poorly understood.

Dim1 provides a useful model for analysis of the folding and function of proteins in complexes involving many protein−protein interactions. In this study, we have taken a systematic approach to elucidating the structure and function of hDim1. (1) We report the solution structure for the reduced dominant negative form of Dim1 and compare it to the crystal structure of the oxidized full length Dim1 protein. (2) The consequences of deleting residues 129−142 for Dim1 folding and stability are investigated using CD spectroscopy. (3) The in vivo complementation properties of a systematic series of alanine mutants pinpoint surface residues critical for protein function. (4) For selected Dim1 mutants with reduced in vivo function, we study their impact on folding and stability. (5) Finally, we compare residues shown to be required for complementation to those necessary for interaction with partner proteins, hnRNP F and Npw38/PQBP-1. Together, these studies identify functionally important regions on Dim1 and can be used to guide further studies of Dim1 in pre-mRNA splicing.

MATERIALS AND METHODS

Protein Expression, Purification, Characterization, and Mutagenesis. Protein expression and purification protocols for full-length hDim1 and $hDim1_{1-128}$ have been described (*2*). Mutants hDim1 G126D, hDim1-8 (R86A, K88A), and hDim1-12 (R124A/K125A/R127A) were expressed and purified by similar procedures. The mutagenesis strategy used to create alanine scan derivatives of hDim1 has been reported (*5*). For expression in *S. pombe*, a series of alanine scan mutants discussed in ref *5* were cloned into the vector pMNS21L.

We have previously found that extended storage of hDim1 samples (1−2 months at 4 °C or ∼1 week at room temperature) results in complete cleavage of the polypeptide chain at positions 128 or 129 (*2*, *5*). The mechanism of the cleavage awaits further study because the flanking amino acids do not match the recognition sequence for any known peptidase. Truncated forms (1-128/129) of the mutant proteins were also obtained after prolonged storage of the purified proteins. The short peptides were removed by using a centrifugal filter device (Ultrafree; Millipore, Bedford, MA) with a 5 kDa molecular weight cutoff. The molecular weight of the purified truncated proteins was determined by mass spectroscopy, which showed two peaks corresponding to 1-128 and 1-129 fragments, but no peak corresponding to full-length protein.

Circular Dichroism Spectroscopy. CD spectra were acquired on an Aviv 62A spectrometer (Aviv, Lakewood, NJ). GuHCl-induced unfolding transitions were measured by recording the ellipticity at 222 nm on a series of Dim1 samples (11.5 μM in 50 mM sodium phosphate buffer, pH 7.82) prepared by mixing aliquots of equimolar stock solutions of native and fully unfolded (6.5 M GuHCl) Dim1. The unfolding equilibrium of reduced Dim1 was measured in the presence of 2 mM DTT. CD spectra over the range from 190 to 300 nm were recorded on oxidized and reduced (1 mM DTT) samples of wild-type and mutant Dim1 (4 μM) in 10 mM potassium phosphate, pH 7.82. Protein concentration was determined by measuring the absorption at 280 nm in 6.5 M GuCl. All CD measurements were performed in 2-mm quartz cuvettes thermostated at 20 °C.

NMR Data Collection. NMR data were collected on a Bruker DMX-600 spectrometer equipped with a 5-mm *x*,*y*,*z*-shielded pulsed-field gradient triple-resonance probe. Stable $hDim1_{1-128}$ samples at concentrations suitable for NMR studies can be prepared only at pH higher than 7.5. Better spectra are obtained at physiological temperature (37 °C) than at room temperature (20 °C). Hence, all NMR experiments were carried out at pH 7.65 and 37 °C. NMR data processing and sequential assignments have been described (*2*). Side-chain assignments were obtained from 3D ^{15}N TOCSY-HSQC (*20*) and 3D HCCH-TOCSY (*21*, *22*).

Distance restraints were obtained from cross-peak volumes in 3D ^{15}N NOESY-HSQC and 3D ^{13}C-edited aliphatic NOESY (*20*). Cross-peaks were categorized as weak (1.8−5 Å), medium (1.8−3.5 Å), or strong (1.4−2.7 Å) in each data set based on the ratio of the peak volume of cross-peaks to that of the relevant diagonal peak. The ratio of weak, medium, and strong peaks was approximately 57:23:13.

$^3J_{HNH\alpha}$ coupling constants were obtained from a 3D HNHA experiment (*23*, *24*). Backbone ϕ angles were constrained to $-60° \pm 30°$ for $^3J_{HNH\alpha} < 5.5$ Hz, $-120° \pm 40°$ for $^3J_{HNH\alpha} > 9$ Hz, $-120° \pm 50°$ for $^3J_{HNH\alpha}$ between 8−9 Hz, $-120° \pm 60°$ for $^3J_{HNH\alpha}$ between 7−8 Hz. For $^3J_{HNH\alpha}$ between 5.5−7 Hz, the backbone ϕ angles were constrained to $-80° \pm 30°$ and omitted if such restraints were violated.

Hydrogen bonds were assigned to slowly exchanging amide protons from a series of 46 min HSQC spectra, with the first spectrum initiated about 2 h after the start of buffer exchange, using a Millipore 5K filter. In the first HSQC spectrum of the hydrogen exchange experiment, 42 cross-peaks were observed. Among these, 14 were confirmed to be involved in hydrogen bonds by the measurement of $^{3h}J_{NC'}$ coupling across hydrogen bond from a 3D HNCO experiment optimized for magnetization transfer across the weak coupling between H-bonded amide nitrogen and proton-acceptor carbonyl carbon (*25*). Each hydrogen bond was specified as two distance restraints (HN−O distance and N−O distance). In the first round of structure calculation, only the hydrogen bonds identified by both experiments were used.

Structure Calculation. Structure calculations were performed on a Silicon Graphics O2 workstation, using the simulated annealing protocol within CNS1.0 (see http://cns.csb.yale.edu). In each round of structure calculation, random structures were generated from the initial structure and subjected to 1000 steps of high-temperature dynamics

at 50 000 K for a total of 15 ps. Each structure was then cooled to 0 K over a period of 15 ps (1000 steps at 15 fs/step). A second, slow-cool annealing was carried out over a total of 15 ps (3000 steps at 5 fs/step) from 2000 to 0 K. Restrained energy minimization was then performed for 2000 steps. The initial structure for the first round of calculation was an extended peptide generated from the sequence. The initial structure for subsequent calculations was one of the structures with low CNS energy from the previous round. In the final round, a family of 120 structures was obtained with root mean square deviation (rmsd) for bond lengths <0.01 Å, for bond angles $<1.0°$, and for improper $<1.0°$. There were no NOE violations >0.5 Å and no dihedral angle violations $>5°$. Among these, the 20 structures with lowest overall energy were selected for further analysis. The coordinates have been deposited in the Protein Data Bank as entry 1PQN. Graphic displays of structures were generated by Rasmol (*26*), Molscript (*27*), and Raster3D (*28*).

Complementation Analyses in S. pombe. 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 as described (*29*). 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 (*31*). Analyses were performed as described by Zhang et al. (*2*). Briefly, the *nmt* promoter used to express hDim1 from the pMNS21L vector is regulated by thiamine, producing high levels of hDim1 and derivatives in the absence of thiamine and low levels in the presence of thiamine (*32*). Wild-type *dim1*$^+$ parental strains are viable from 25 to 37°C and grow optimally at 30−32 °C; *dim1-35* mutant strains are viable at 25 °C and inviable at 36.5 °C (*1*). *Dim1* null strains are inviable at all temperatures and maintained as *dim1::his3/dim1*$^+$ heterozygote diploids. For analysis of complementation in a null background, plasmids containing relevant mutants are transformed into the diploid, the diploid sporulated in the presence or absence of thiamine, and His$^+$ Leu$^+$ colonies selected at 32 °C. In cases where colonies were not obtained under these conditions, sporulation was performed at 25 °C to determine if the defect was due to temperature sensitivity. For analysis of rescue of the *dim1-35* hypomorph, plasmids containing relevant mutants in hDim1 were transformed into the haploid strain, and shift in permissive temperature was scored in the presence or absence of thiamine.

RESULTS AND DISCUSSION

Solution Structure of Reduced hDim1$_{1-128}$. Since the truncated form of human Dim1, hDim1$_{1-128}$, acts as a dominant negative inhibitor of the full-length protein in yeast, the highly conserved C-terminal segment clearly plays an important functional role. Another question raised by the recent crystal structure of hDim1 concerns the relevance of disulfide bond formation between the two cysteine residues of hDim1. To address these issues, we used heteronuclear multidimensional NMR methods to determine the solution structure of hDim1$_{1-128}$ in its disulfide-reduced form at pH 7.65 and 37 °C. The NMR-based structure calculation used 1378 distance restraints (including 84 for hydrogen bonds) and 67 dihedral angle restraints (Table 1). Figure 1 shows backbone traces for the 20 structures with lowest energy among a total of 120 structures calculated. The rmsd of this set of 20 structures is 0.8 ± 0.1 Å for the backbone heavy atoms of the residues involved in secondary structure (Table 1). The rmsd for the whole backbone is substantially higher because of the presence of two extended loops for which few experimental constraints were obtained. Figure 2 depicts a ribbon diagram of the energy-minimized average structure. Analyses of the pattern of NOEs and the directly observed hydrogen bonds reveal a mixed parallel and antiparallel triple-stranded β-sheet with an additional short strand (Figure 3).

Table 1: Structural Statistics of the NMR Solution Structure of the Reduced hDim1$_{1-128}$[a]

Experimental Restraints Used in the Final Structure Calculation	
inter-residue NOEs	852
sequential($\|i-j\| = 1$)	467
medium range ($\|i-j\| < 5$)	204
long range ($\|i-j\| \geq 5$)	181
intra-residue NOEs	442
dihedral restraints ($^3J_{HNH\alpha}$)	67
hydrogen bonds restraints	84
rmsd from Ideal Geometry	
bonds (Å)	0.0012 ± 0.0001
angles (deg)	0.26 ± 0.004
impropers (deg)	0.099 ± 0.006
rmsd from Experiment Restraints	
distance	0.0060 ± 0.0004
dihedral angle (deg)	0.05 ± 0.02
rmsd from Mean Structure	
11−19, 25−31, 36−48, 56−62, 80−85, 88−90, 91−93, 109−124 backbone	0.8 ± 0.1

[a] The statistical parameters and rmsd values were calculated for the 20 structures with lowest overall energy.

Previously, we predicted that hDim$_{1-128}$ has a thioredoxin-like structure, based on preliminary NMR results and molecular modeling (*2*). The structure calculated from NMR data confirms that hDim1$_{1-128}$ assumes a thioredoxin-like fold with a 10-amino acid insertion in the last loop of thioredoxin (*2*). Overall, the structure of hDim1$_{1-128}$ in its reduced state is very similar to a recently reported crystal structure of the oxidized form of hDim1 (*19*). The rmsd between the backbone residues involved in the secondary structure of the 20 solution structures and the oxidized crystal structure is 1.4 ± 0.2 Å. The structure consists of a mixed (parallel and antiparallel) four-stranded β-sheet (β1: 25−31; β2: 56−62; β3: 80−85; β4: 88−90) flanked by three α-helices (α1: 11−21; α2: 36−47; α3: 109−122). Residues 96−101 in the loop connecting β4 and α3 are undefined in the X-ray structure. In the NMR spectra, most of these residues were also poorly defined, and unambiguous constraints involving these residues could not be obtained (*2*). The main difference between the crystal structure of hDim1 and the NMR structure of hDim1$_{1-128}$ is that in the crystal structure, residues 91−93 and 129−131 form a short parallel β sheet, whereas in the final 20 NMR structures residues 91−93 assume a range of structures that do not register as regular β-sheet, presumably because of deletion of the 14 C-terminal residues.

In thioredoxin, the cysteines in the canonical Cys-X-X-Cys motif form a functional disulfide bond at the N terminus of helix α2. The corresponding region of the hDim1 sequence

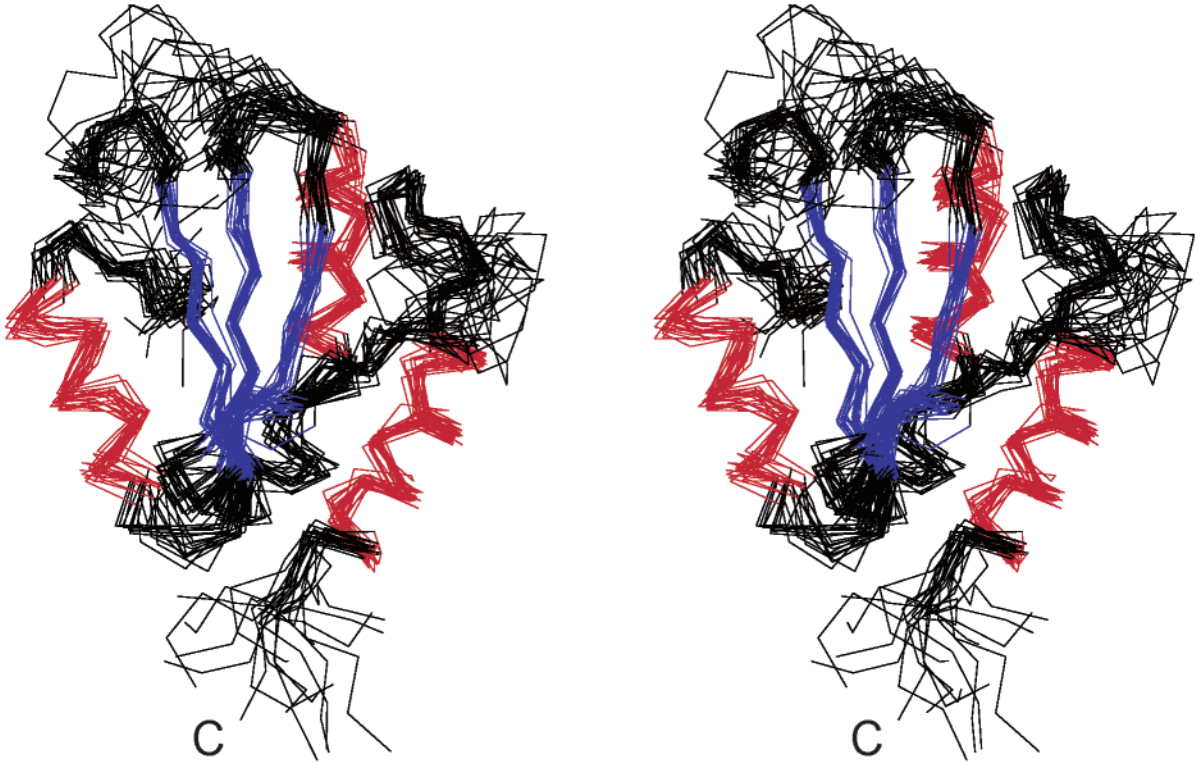

FIGURE 1: Stereoview of the final ensemble of 20 structures of hDim1$_{1-128}$ superimposed using the backbone-heavy atoms of residues in the secondary structure. α-Helices are shown in red, and β-strands are shown in blue.

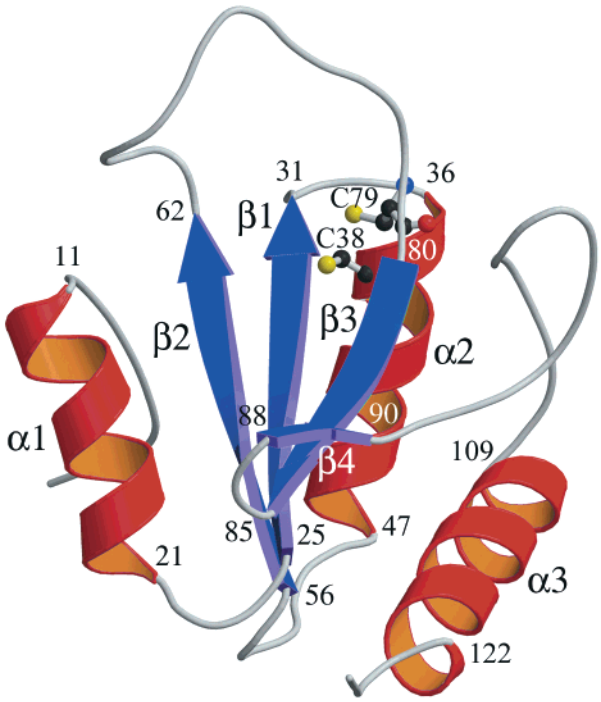

FIGURE 2: Ribbon diagram of the hDim1$_{1-128}$ NMR structure with lowest overall energy. The two cysteine residues are shown in ball-and-stick representation. The first and the last residues of each secondary structure element are numbered.

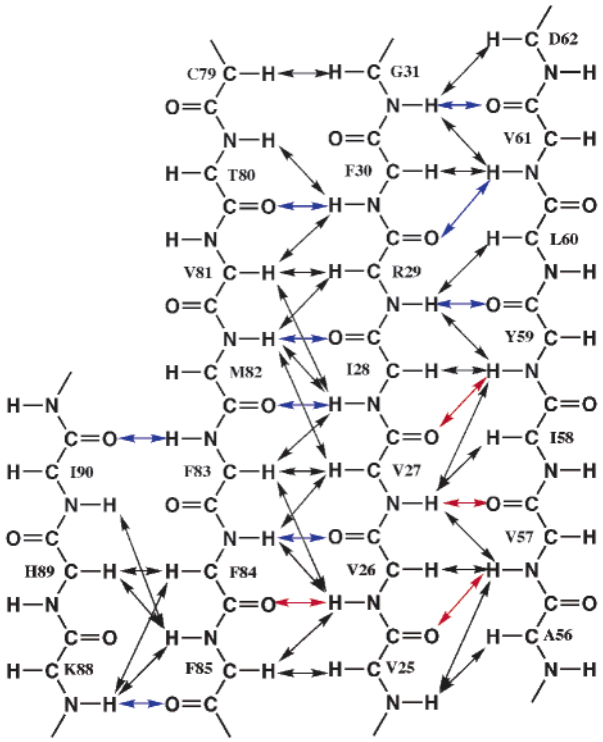

FIGURE 3: Schematic representation of the β-sheet region of hDim1$_{1-128}$. Constraints used in the final NMR structure calculation are indicated as follows: black arrows, NOE connectivities; blue arrows, hydrogen bonds that can be observed by hydrogen exchange but not in the $^{3h}J_{NC'}$-HNCO experiment; and red arrows, hydrogen bonds observed in both experiments.

contains only one cysteine, Cys38, corresponding to the second Cys in the Cys-X-X-Cys motif of thioredoxin. Although disulfide-bond formation with a second cysteine at position 79 on strand β3 of hDim1 is possible, the structural and/or functional significance of this linkage is unclear. A mixture of oxidized and reduced forms was observed in the crystal structure of hDim1 (*19*), despite the presence of DTT during the crystallization. However, Cys79 is not an evolutionarily conserved residue (*2*), suggesting

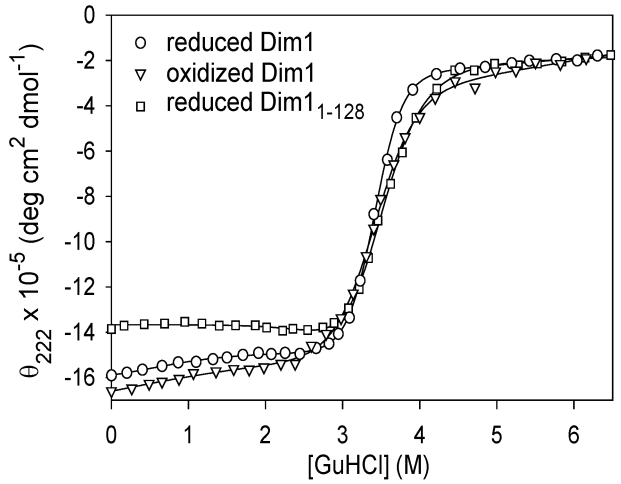

FIGURE 4: Unfolding transitions of hDim1 induced by GuHCl monitored by CD at 222 nm (20 °C, 50 mM sodium phosphate, pH 7.82). Reduced (○) and oxidized (▽) forms of full-length Dim1 are compared with the truncated reduced form, $hDim1_{1-128}$ (□). In the reduced samples, formation of the 38−79 disulfide bond was prevented by addition of 2 mM DTT.

that this disulfide bond is not essential for Dim1 function.

DTT was added to the NMR sample to eliminate the possibility of intermolecular disulfide bridge formation. Comparison of the chemical shifts of the β-carbons of Cys38 and Cys79 (27.8 and 29.2 ppm, respectively) with the "random coil" values reported by Wishart and Case (*33*) (28.0 ppm for the reduced and 41.1 ppm for the oxidized Cys) confirms that the protein is fully reduced under the conditions used in our NMR studies. Under these conditions, the β-carbons of the two cysteines are at a distance of 5.7 ± 0.3 Å. In contrast, the average distance between two β-carbons of disulfide-bonded cysteines is 3.9 ± 0.3 Å in the PDB database (containing a total of 18 660 disulfide bonds). Although the distance seen in the solution structure is clearly inconsistent with a covalent bond, it appears short enough for the disulfide bond to form without major structural rearrangement, which is consistent with the fact that both reduced and oxidized forms can coexist in the same crystal lattice (*19*). This conclusion was further confirmed by our observation that the chemical shifts of amide protons and nitrogens for hDim1 under oxidizing conditions are very similar to those of the DTT-reduced protein (data not shown) and by the observation that the CD spectra of reduced and oxidized hDim1 (full length) are virtually identical (data not shown). To determine the contribution of the disulfide bond to protein stability, we measured the GuHCl-induced unfolding transitions of full-length hDim1 in the presence and absence of DTT (Figure 4). The transitions occur at similar midpoint concentrations (C_m), but the oxidized protein shows a shallower transition (smaller m value), resulting in a lower value for the free energy of unfolding, ΔG_u, in the absence of denaturant (Table 2). The observed decrease in ΔG_u upon formation of the disulfide bond (from 11.0 to 7.3 kcal/mol) indicates that formation of the 38−79 disulfide bond is energetically unfavorable. Given the reducing environment in the cell, the disulfide bond is therefore unlikely to be present under physiological conditions. Taken together, these observations strongly support the biological evidence that Dim1 function is not linked with formation of a disulfide bond (*2*).

Structural and Functional Role of the hDim1 C-Terminal Tail. The ability of the truncated form of hDim1, $hDim1_{1-128}$, to act as a dominant negative inhibitor of the full-length protein in yeast, together with the fact that the residues in the C-terminal extension are highly conserved (*2*), indicates that the C-terminal 14 amino acids play an important functional role. To determine the contribution of the C-terminal tail to protein stability, we used CD measurements in the α-helical region (222 nm) to compare the GuHCl-induced unfolding transitions for the full-length and truncated forms of reduced hDim1. Although both forms unfold over a similar range of GuHCl concentrations (Figure 4), the truncated form shows a shallower transition, resulting in a substantially lower free energy of unfolding in the absence of denaturant (Table 2). In addition to the main transition centered around 3.4 M GuHCl, the wild-type protein shows a small decrease in the magnitude of the CD signal at 222 nm between 0 and 2 M GuHCl consistent with a slight decrease in helical secondary structure. In contrast, the molar ellipticity for the truncated form is less negative (corresponding to a ∼15% decrease in secondary structure) and nearly GuHCl-independent between 0 and 3 M GuHCl. While small changes in the pretransition region are often dismissed as uninteresting baseline effects, the fact that they are observed only for the full-length protein suggests that they may reflect a local unfolding transition involving the 14 C-terminal residues of Dim1. In fact, the unfolding curve of hDim1 can be fitted as a three-state equilibrium involving a shallow minor transition ($C_{mp} \approx 0.5$ M, $m_p \approx 1.1$ kcal mol^{-1} M^{-1}) preceding the main unfolding transition ($C_m = 3.41$ M, $m = 3.2$ kcal mol^{-1} M^{-1}; Table 2). Our conclusion that the pretransition changes reflect local unfolding of the C-terminal extension is supported by the observation that all full-length proteins investigated, including oxidized Dim1 (Figure 4) and several mutant forms (Figure 8), exhibit very similar behavior at low denaturant concentrations.

In the crystal structure, the 14 C-terminal amino acids show extended structure, including a short segment of parallel β-sheet involving residues 129−131 and β-strand 4 (residues 91−93). Because of progressive cleavage of the C-terminal tail, we were unable to carry out a full NMR structure determination of the full-length protein. However, we were able to record an HSQC spectrum of the intact protein, which allowed us to measure changes in amide nitrogen and proton chemical shifts associated with truncation (Figure 5). Not surprisingly, some of the largest chemical shift changes are found near the cleavage site (residues 118−123) and between residues 80 and 95, indicating perturbations due to the loss of contacts with tail residues. However, there are substantial chemical shift changes in other parts of the protein, suggesting more widespread structural perturbations.

Comparison of the CD spectra for full-length hDim1 and $hDim1_{1-128}$ shows that truncation results in significant changes in secondary structure (Figure 6a). Interestingly, the difference spectrum (in terms of molar ellipticity) obtained by subtracting the spectrum of the truncated form from that of the full-length protein exhibits a broad negative band around 220 nm consistent with a significant decrease in α-helical secondary structure on truncation (Figure 6b). Using the program k2D (http://www.embl-heidelberg.de/∼andrade/k2d.html), we estimate that ∼75% of the 14 C-terminal residues of wild-type hDim1 are in an α-helical confor-

Table 2: Equilibrium Parameters for GuHCl-Induced Unfolding of Dim1[a]

Dim1 protein	C_m (M)	m (kcal mol^{-1} M^{-1})	ΔG_0 (kcal mol^{-1})
wt (oxidized)	3.39 ± 0.02	2.16 ± 0.11	7.32 ± 0.002
wt (reduced)	3.41 ± 0.01	3.24 ± 0.07	11.05 ± 0.01
G126D (reduced)	3.36 ± 0.41	2.78 ± 1.17	9.34 ± 0.48
(1-128) (reduced)	3.35 ± 0.33	1.72 ± 0.53	5.76 ± 0.17
D1-8 (reduced)[b] (R86A/K88A)	2.59 ± 0.01	2.92 ± 0.20	7.56 ± 0.002
D1-12 (reduced) (R124A/K125A/R127A)	3.35 ± 2.21	2.63 ± 1.58	8.81 ± 3.49

[a] Measured by CD at 222 nm in 50 mM sodium phosphate, pH 7.82, 20 °C; disulfide-reduced samples were measured in the presence of 2 mM DTT. [b] Major transition; a four-state fit yields two additional transitions with $C_{m1} = 0.87 \pm 0.07$ M, $m_1 = 1.64 \pm 0.40$ kcal mol^{-1} M^{-1}, and $C_{m3} = 4.51 \pm 0.12$ M, $m_3 = 1.15 \pm 0.21$ kcal mol^{-1} M^{-1}.

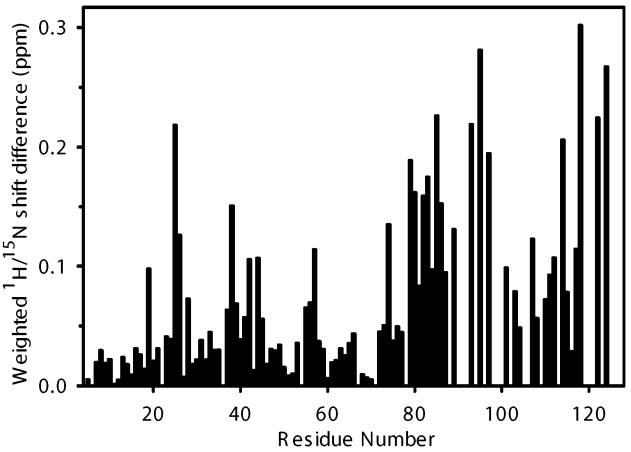

FIGURE 5: Differences in backbone amide ^{1}H and ^{15}N chemical shifts between hDim1$_{1-128}$ and full-length hDim1, expressed in terms of the weighted geometric average $\Delta\delta$ (ppm) = $\sqrt{(\Delta\delta^{1H})^2+(0.17*\Delta\delta^{15N})^2}$ plotted vs residue number.

mation. In addition, the program AGADIR (http://www.embl-heidelberg.de/Services/serrano/agadir/agadir-start.html) predicts a substantial α-helix content (~60%) for a peptide corresponding to the C-terminal segment of hDim1 (LVVSPKDYSTKYRY) under conditions similar to those used here (pH 7.0, 278 K, and 0.1 ionic strength). Thus, in contrast to the crystal structure, the 14 residues of hDim1, which extend beyond the thioredoxin-like core of the protein, appear to assume a partially α-helical conformation in solution.

These apparently contradictory observations may be explained if the C-terminal segment is inherently flexible and can assume alternative conformations, depending on the conditions and structural environment, such as crystal constraints or interactions within the splicing complex. This scenario is supported by the existence of a local unfolding transition involving the C-terminal residues of hDim1 suggested by our CD data. It is also possible that residues 129−131 are involved in β-sheet interactions with strand β4, as seen in the crystal structure, while the rest of sequence adopts a helical conformation.

Identification of Essential Residues for hDim1 Function. As noted above, Dim1 proteins are evolutionarily extremely well conserved. This fact, coupled with the observation that hDim1 is a component of the U4/U6·U5 tri-snRNP, suggests that hDim1 may be embedded in a multisubunit complex and engaged in multiple protein−protein or protein−RNA interactions. To determine the amino acids required for the

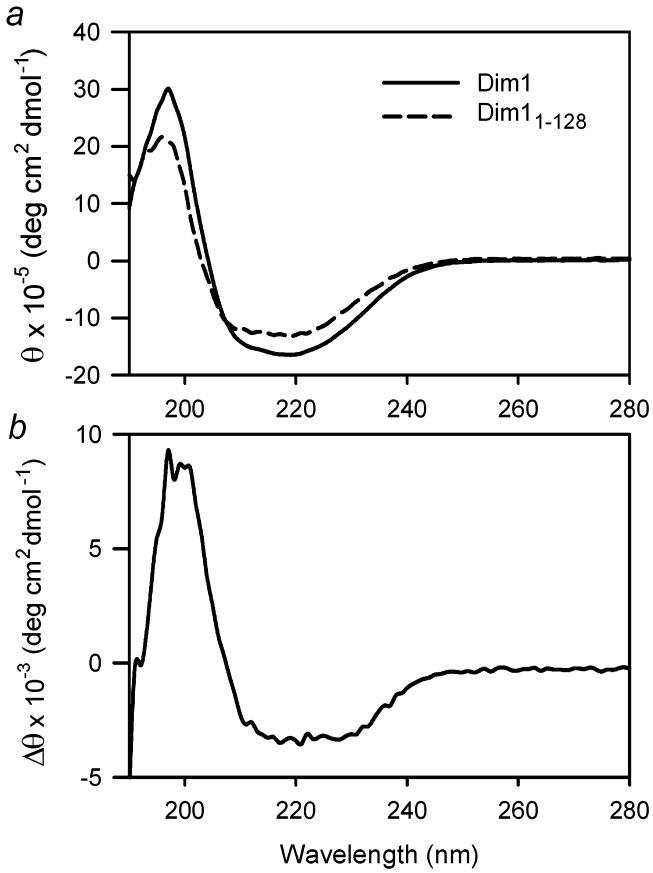

FIGURE 6: Effect of C-terminal truncation of hDim1 on the CD spectra. (a) Spectra (molar ellipticity) of full-length hDim1 and the truncated form, hDim1$_{1-128}$. (b) Difference spectrum (full-length minus truncated).

essential function of hDim1, we generated an alanine scan mutant series that systematically targeted charged residues throughout the hDim1 coding sequence and tested each of these mutants for its ability to complement *dim1* null and *dim1-35* temperature-sensitive mutants in *S. pombe*. Results of these analyses are summarized in Table 3.

As previously described, wild-type mammalian forms of hDim1 can complement both the *dim1-35* and *dim1*$^-$ null mutations in *S pombe* (*1*). It is clear from Table 3 that mutants hDim1-1, -3, -6, -7, -10, -11, -13, and -14 are fully functional. In contrast, mutant hDim1-8 completely lost its ability to rescue either the null or temperature-sensitive alleles under any growth conditions. Mutants hDim1-2, -4, -5, -9, -12, and -15 showed partial rescue. hDim1-4 and -12 were temperature sensitive (ts) for the ability to complement the *dim1::his3* allele. For hDim1-12, such ts complementation

Table 3: Summary of Alanine Scan Mutations in Dim1 and Corresponding Complementation and Phenotypes

			Rescue of					
			Dim1::his				*dim1-35*	
			25 °C		32 °C			
mutant no.	sequence no. of targeted residues	sequence of targeted residues	+B1	−B1	+B1	−B1	+B1	−B1
WT hDim1			Y	Y	Y	Y	Y	Y
1	7−9	**HLH**[a]			Y	Y	Y	Y
2	21−24	EEDR			Y	Y	N	N
3	33−35	**D**W**D**			Y	Y	Y	Y
4	40−43	**KMDE**	Y	Y	N	N	N	Y
5	50−53	**EK**V**K**			N	Y	Y	Y
6	65−68	**E**VP**D**			Y	Y	Y	Y
7	71−74	**KMYE**			Y	Y	Y	Y
8	86−88	**RNK**	N	N	N	N	N	N
9	107−109	EDK			Y	Y	N	Y
10	111−114	**EMVD**			Y	Y	Y	Y
11	122	G			Y	Y	Y	Y
12	124−127	**RKGR**	Y	Y	N	Y	N	N
13	134−135	KD			Y	Y	Y	Y
14	139−141	**KYR**			Y	Y	Y	Y
15[b]	1−8	deleted			N	Y	N	Y
16[c]	126	**G**→D						

[a] Bold face indicates residues changed to alanine. [b] See legend to Figure 7. [c] For mutant 16, the original G126D mutation identified in the *dim1-35* allele was introduced into the human hDim1 background to allow comparative evaluation versus the remaining mutations.

was only observed in the presence of thiamine (low level expression). Strikingly, although hDim1-2 and -12 were able to complement the null mutation, they were unable to complement the originally defined *dim1-35* allele (G126D mutant) even with high-level expression, suggesting that these alleles were targeting an overlapping function. Conversely, hDim1-4 was unable to rescue the null allele of *dim1*, but could rescue *dim1-35*, indicating the possible presence of two distinct essential Dim1 activities. All other mutants evaluated behaved equivalently to wild-type hDim1.

To gain additional insight into the significance of the results described above, we mapped the conserved residues and the position of mutations that affected Dim1 function on the structure of the protein (Figure 7), from which two points are worth noting. First, the hDim1-2, -8, and -12 mutations all cluster in a strongly conserved region surrounding G126, the single residue that is altered (G126D) in the *dim1-35* mutant. All three mutants fail to rescue the *dim1-35* allele, although two (hDim1-2 and -12) are competent to rescue the null allele at 25 °C. It has been noted previously (*19*) that residues R86, K88, R124, K125, and R127 targeted in these mutants form a basic patch likely to interact with a single target containing negatively charged surfaces, such as RNA. This structure−function analysis supports the idea that an essential activity of the protein resides in this region. This may represent a catalytic or binding activity. Second, the hDim1-4 mutant targets a separate face of the Dim1 molecule that is spatially far from the basic pocket. As K40, D42, and E43 are also critical for Dim1 function and hDim1-4 has discrete complementation phenotype, it would appear to identify a second critical site for Dim1.

In a previous study (*5*), we identified Npw38/PQBP-1 and hnRNP F as hDim1-interacting proteins and mapped residues in hDim1 that are important for these interactions by combined mutational and two-hybrid analyses. Comparing the essential regions for Dim1 function with the regions involved in interactions with Npw38/PQBP-1 and hnRNP F, it is clear that residues shown to contribute to hnRNP F and Npw38/PQBP-1 binding overlap only partially with those shown to be required for Dim1's essential function(s). Overlap is seen in both hDim1-2 and -4 mutants. Moreover, a number of the sites shown to impact binding to the interactors, most notably hDim1-6, have no effect on Dim1's essential function(s). In short, separate sets of specific residues on the surface of the hDim1 molecule are required to stabilize interactions with these individual partner molecules (*5*) and for Dim1's essential function.

Effect of Mutations on Structure and Stability. The results from our alanine scan mutant series and yeast complementation experiments show that mutants hDim1-2, -4, -8, and -12 partially or completely lost their ability to rescue the nonfunctional *dim1::his* or *dim1-35* alleles (Table 3). We have shown by Western analysis that the steady-state level of overexpressed Lex A-fused derivatives of mutated Dim1 proteins is not affected (*5*). Therefore, loss of the rescue ability is apparently not caused by gross decrease of levels of hDim1 in the cell. To determine whether the loss in Dim1 function is related to changes in structure and/or stability, we used CD spectroscopy to further characterize some of the variants. We were able to express and purify wild-type hDim1, hDim1-8 (R86A/K88K), hDim1-12 (R124A/K125A/R127A), and the G126D mutant. The solubility of two variants (hDim1-2 and -4) was insufficient for biophysical studies.

The far-UV−CD spectra for all mutants analyzed are nearly identical to that of the wild type (cf. Figure 6), indicating that the secondary structure is not significantly perturbed by the mutations. To determine the effect of mutations on protein stability, we used the CD signal at 222 nm to monitor the unfolding equilibrium as a function of GuHCl. To ensure complete reduction of the disulfide bond, experiments were performed in the presence of 2 mM DTT. Between 3 and 4 M GuHCl, wild-type hDim1, G126D, and hDim1-12 show a sharp change in the CD signal at 222 nm, corresponding to an abrupt loss in the negative band

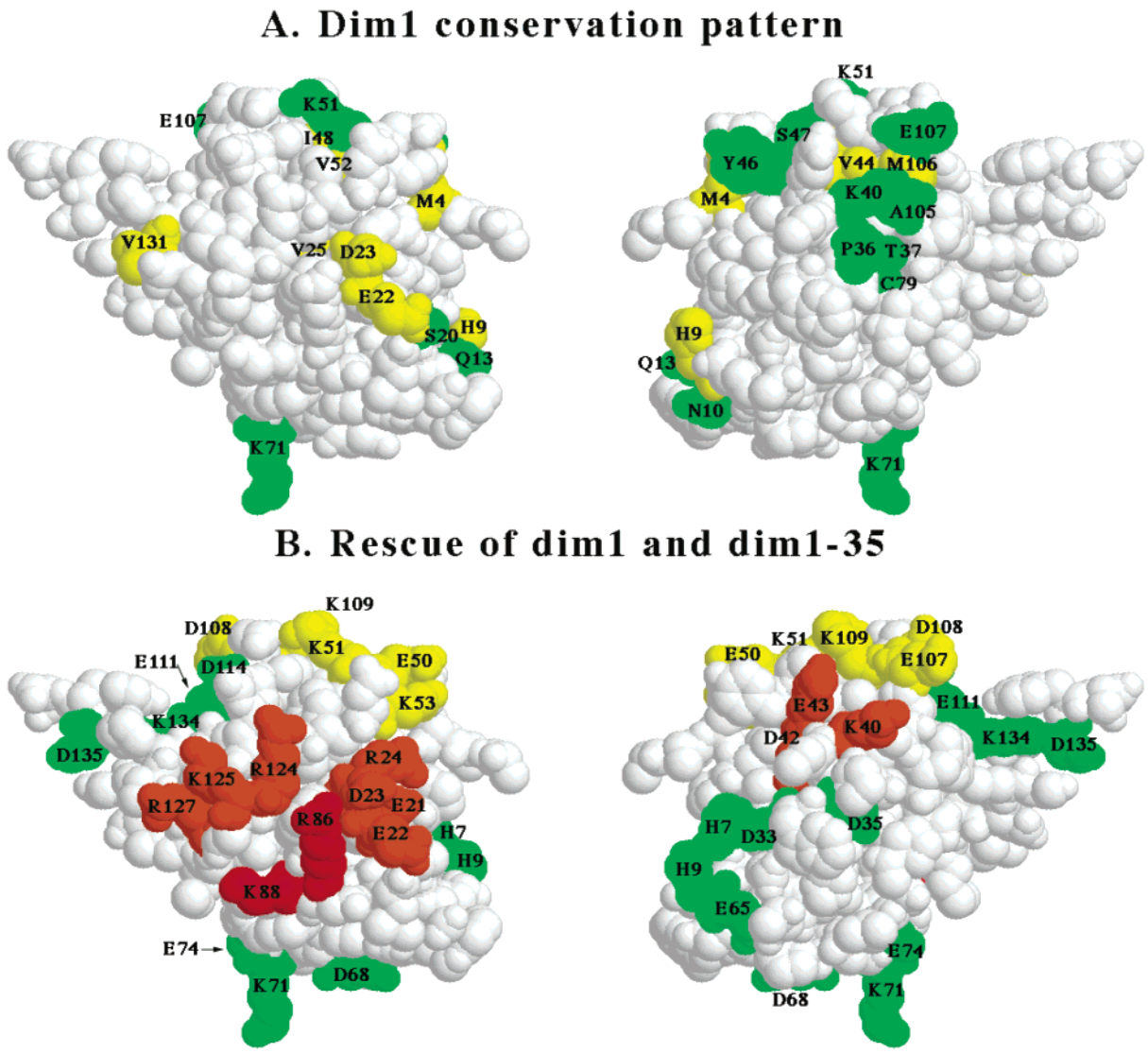

FIGURE 7: Conservation pattern and mutation effect on function for Dim1. Note: mutant phenotypes are the product of two or more amino acid changes for each construct evaluated (see Table 3 for details). For easy comparison with the mutation effect on the interaction between Dim1 and hnRNP F and PQBP-1 (*2*), the coordinate of the oxidized hDim1 1qvc (*5*) was used. (A) The evolutionary conservation of Dim1. White, amino acid completely conserved; yellow, conservative substitutions allowed; green, variance tolerated. (B) Rescue of *dim1::his3* and *dim1-35* in *S. pombe*. Red, no rescue; orange, partial rescue (rescue of either null or hypomorph, but not both); yellow, rescue under conditions of high-level hDim1 mutant expression; green, rescue at all conditions tested (see Results and Discussion sections). Residues shown in white were not mutated.

dominated by an α-helical secondary structure (Figure 8). The free energies of unfolding, determined from the midpoint concentration (C_m) and slope (*m* value) of the main transition (Table 2), are nearly identical, indicating that these mutations have negligible effects on stability. As in the case of wild-type Dim1 (Figure 4), the mutants show small changes in ellipticity at denaturant concentrations below the onset of the main unfolding transition centered around 3.4 M GuHCl, which can be interpreted as a second transition between the native state and a native-like intermediate. Indeed, a three-state equilibrium model (U ↔ I ↔ N), including a partially folded intermediate (I) in addition to the fully unfolded (U) and native (N) states, yields an excellent fit to the data. This interpretation is supported by the fact that the different forms of Dim1 studied show significant variations in their pretransition behavior. The changes are smallest in the case of the truncated form, $Dim1_{1-128}$ (Figure 4), and most pronounced in the case of the least stable variant, R86A/K88A (Figure 8b).

Remarkably, this variant (also termed Dim1-8) shows clear evidence for a third transition at denaturant concentrations above the main transition. This complex unfolding curve, which is highly unusual for a relatively small globular protein, can be fitted to a four-state model (U ↔ I_1 ↔ I_2 ↔ N). The second transition (the major one) has a C_m of 2.6 M, which is significantly lower than that of the wild type and other mutants (Table 2). Thus, the R86A/K88A mutation causes a major decrease in structural stability as well as loss in cooperativity of the unfolding transition. Despite the disruptive nature of these amino acid changes, the protein appears fully folded in the absence of denaturant, as indicated by its native-like far-UV−CD spectrum (data not shown). In addition, Dim1-8 shows a highly dispersed $^{15}N-^{1}H$ HSQC spectrum very similar to that of the wild-type protein (Figure

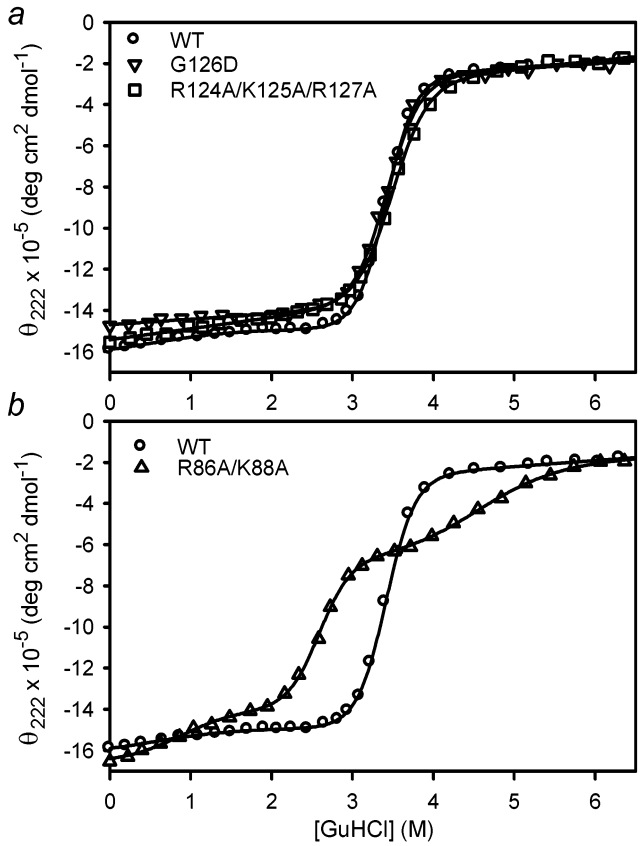

FIGURE 8: Effect of mutations on the folding−unfolding equilibrium. (a) The GuHCl-induced unfolding transition of wild-type hDim1 (○) is compared with that of G126D (▽) and the R124A/K125A/R127A (hDim1-12; □). (b) Unfolding transition of R86A/K88A (hDim1-8; △), along with wild-type hDim1 (○). The solid lines represent fits of 3-state or 4-state unfolding mechanisms (see text).

9a). In Figure 9b, the average changes in amide nitrogen and hydrogen chemical shifts (Dim1-8 minus wild type) are plotted vs residue number. Significant chemical shift changes are found only for residues at or near the site of mutations (residues 85−90) and a few residues involved in tertiary contacts with this region (residues 21−23).

On the basis of their crystal structure, Reuter et al. (*19*) noted that Arg86 and Lys88 join with Arg124, Lys125, and Arg127 to form a basic patch likely to interact with a target molecule containing a negatively charged surface, such as RNA. Our GuHCl-unfolding results indicate that R86 and R88, located in the C terminus of β3 and the N terminus of β4 (Figure 2), are important for maintaining a stable, cooperatively folded structure. Arg86 is surrounded by several acidic residues (Glu21, Glu22, and Asp23), and its replacement with Ala may disrupt important electrostatic interactions. This is consistent with the fact that some of the largest chemical shift perturbations associated with the R86A/R88A mutations are found for residues in the loop region containing residues 21−23 (Figure 9). Thus, the loss of rescue ability in hDim1-8 can result from the fact that mutation of Arg86 and Lys88 disrupts a basic patch also involving residues Arg124, Lys125, and Arg127. However, we cannot rule out the possibility that the decrease in structural stability and disruption in the cooperativity of the folding/unfolding transition contributes to the functional defects.

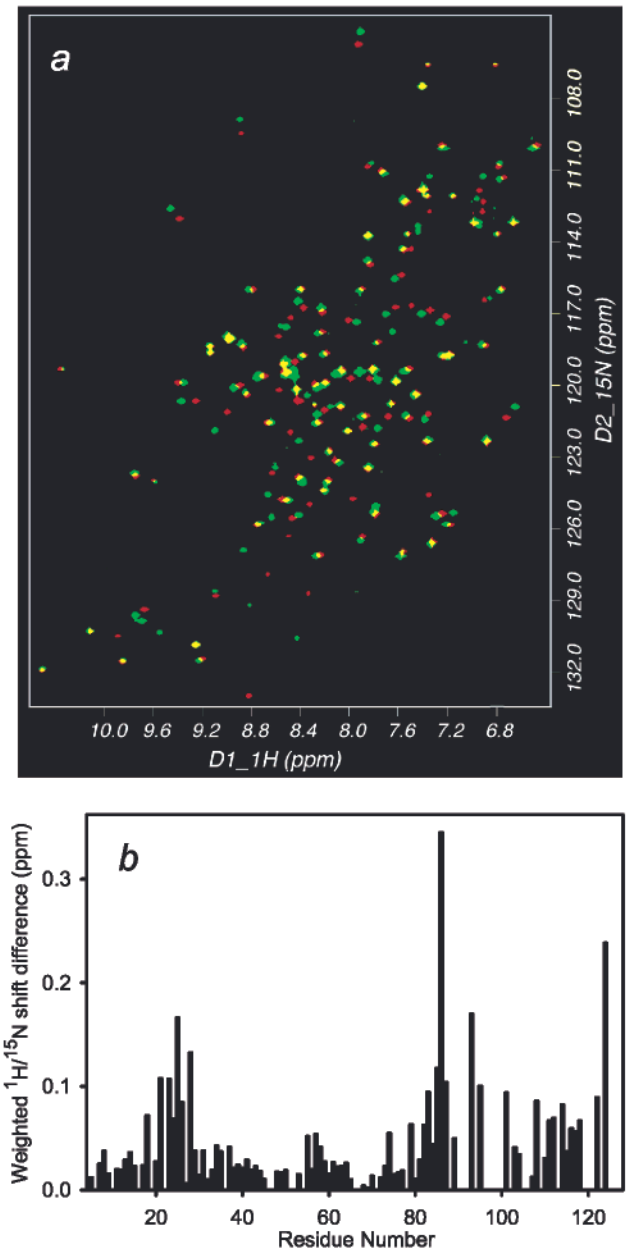

FIGURE 9: Chemical shift changes induced by the R86A/K88/A mutation of reduced hDim1$_{1-128}$. (a) HSQC spectrum of hDim1-8$_{1-128}$ (red) plotted together with that of the wild-type hDim1$_{1-128}$ (green). Yellow peaks indicate residues with little or no changes in chemical shifts for both ^{1}HN and ^{15}N resonances. (b) Weighted geometric average of amide proton and nitrogen chemical shift differences (see Figure 5) between wild-type Dim1$_{1-128}$ and the R86A/K88/A mutant.

In conclusion, we have determined that the dominant negative Dim1 (hDim1$_{1-128}$) forms the thioredoxin-like core of Dim1. The C-terminal extension of Dim1, whose removal largely affects the basic region of the wild-type protein, is required for its function (*2*). Mutagenesis and yeast complementation studies indicate that residues involved in the basic patch are important for Dim1 function. If the basic patch is involved in interactions with other splicing factors (e.g., Prp6 and other components of the tri-snRNP) or with RNA in conjunction with other splicing subunits, our data suggest that Dim1 plays an important role in general pre-mRNA splicing. Our structural and thermodynamic studies further show that (1) the carboxyl-terminal 14 amino acids can

assume multiple conformations in the solution, (2) in addition to the β-strand formed by residues 129−131, as seen in the crystal structure, it is likely that the rest of the sequence adopts a helical conformation, and (3) structural flexibility of the C-terminal segment is likely to be important for the biological function of hDim1.

ACKNOWLEDGMENT

We are grateful to Ying Tong Wang and Anna Feoktistova for excellent technical assistance and to Jenny Glusker and Ramil Latypov for their careful reading of the manuscript.

REFERENCES

1. Berry, L. D., and Gould, K. L. (1997) *J. Cell Biol. 137*, 1337−1354.
2. Zhang, Y. Z., Gould, K. L., Dunbrack, R. J., Cheng, H., Roder, H., and Golemis, E. A. (1999) *Physiol. Genomics 1*, 109−118. (http://physiolgenomics.physiology.org/cgi/reprint/1/3/109)
3. Stevens, S. W., and Abelson, J. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 7226−7231.
4. Gottschalk, A., Neubauer, G., Banroques, J., Mann, M., Luhrmann, R., and Fabrizio, P. (1999) *EMBO J. 18*, 4535−4548.
5. Zhang, Y., Lindblom, T., Chang, A., Sudol, M., Sluder, A. E., and Golemis, E. A. (2000) *Gene 257*, 33−43.
6. Powell-Coffman, J. A., Knight, J., and Wood, W. B. (1996) *Dev. Biol. 178*, 472−483.
7. Ferguson, K. C., Heid, P. J., and Rothman, J. H. (1996) *Genes Dev. 10*, 1543−1556.
8. Min, H., Chan, R. C., and Black, D. L. (1995) *Genes Dev. 9*, 2659−2671.
9. Chou, M.-Y., Rooke, N., Turck, C. W., and Black, D. L. (1999) *Mol. Cell. Biol. 19*, 69−77.
10. Honore, B., Rasmussen, H. H., Vorum, H., Dejgaard, K., Liu, X., Gromov, P., Madsen, P., Gesser, B., Tommerup, N., and Celis, J. E. (1995) *J. Biol. Chem. 270*, 28780−29789.
11. Komuro, A., Saeki, M., and Kato, S. (1999) *Nucleic Acids Res. 27*, 1957−1965.
12. Waragai, M., Lammers, C.-H., Takeuchi, S., Imafuku, I., Udagawa, Y., Kanazawa, I., Kawabata, M., Mouradian, M. M., and Okazawa, H. (1999) *Hum. Mol. Genet. 8*, 977−987.
13. Blencowe, B. J., Bauren, G., Eldridge, A. G., Issner, R., Nickerson, J. A., Rosonina, E., and Sharp, P. A. (2000) *RNA 6*, 111−120.
14. Blencowe, B. J., Issner, R., Nickerson, J. A., and Sharp, P. A. (1998) *Genes Dev. 12*, 996−1009.
15. Okazawa, H., Rich, T., Chang, A., Lin, X., Waragai, M., Kajikawa, M., Enokido, Y., Komuro, A., Kato, S., Shibata, M., Hatanaka, H., Mouradian, M. M., Sudol, M., and Kanazawa, I. (2002) *Neuron 34*, 701−713.
16. Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleish, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J. M. (2000) *Nature 403*, 623−627.
17. Abovich, N., Legrain, P., and Rosbash, M. (1990) *Mol. Cell. Biol. 10*, 6417−6425.
18. Galisson, F., and Legrain, P. (1993) *Nucleic Acids Res. 21*, 1555−1562.
19. Reuter, K., Nottrott, S., Fabrizio, P., Luhrmann, R., and Ficner, R. (1999) *J. Mol. Biol. 294*, 515−525.
20. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) *J. Biomol. NMR 4*, 845−858.
21. Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) *J. Magn. Reson. 88*, 425−431.
22. Bax, A., Clore, G. M., Driscoll, P. C., Gronenborn, A. M., Ikura, M., and Kay, L. E. (1990) *J. Magn. Reson. 87*, 620−627.
23. Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) *J. Biomol. NMR 4*, 871−878.
24. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
25. Wang, Y. X., Jacob, J., Cordier, F., Wingfield, P., Stahl, S. J., Lee-Huang, S., Torchia, D., Grzesiek, S., and Bax, A. (1999) *J. Biomol. NMR 14*, 181−184.
26. Sayle, R. A., and Milner-White, E. J. (1995) *Trends Biochem. Sci. 20*, 374.
27. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
28. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−523.
29. Moreno, S., Klar, A., and Nurse, P. (1991) *Methods Enzymol. 194*, 795−823.
30. Maundrell, K. (1993) *Gene 123*, 127−130.
31. Prentice, H. L. (1992) *Nucleic Acids Res. 20*, 621.
32. Fleig, U. N., and Nurse, P. (1991) *Mol. Gen. Genet. 226*, 432−440.
33. Wishart, D. S., and Case, D. A. (2001) *Methods Enzymol. 338*, 3−34.

BI034486I