Disease-causing Mutations in Cartilage Oligomeric Matrix Protein Cause an Unstructured Ca\(^{2+}\) Binding Domain*

Quinn Kleerekoper‡‡, Jacqueline T. Hecht††, and John A. Putkey‡‡‖

From the ‡Departments of Biochemistry and Molecular Biology and ††Pediatrics, ‡‡Structural Biology Research Center, University of Texas, Houston Medical School, Houston, Texas 77030

Received for publication, October 15, 2001, and in revised form, December 11, 2001
Published, JBC Papers in Press, January 8, 2002, DOI 10.1074/jbc.M109944200

Chondrocytes from pseudoachondroplasia (PSACH) and multiple epiphyseal dysplasia (EDM1) patients display an enlarged rough endoplasmic reticulum that accumulates extracellular matrix proteins, including cartilage oligomeric matrix protein (COMP). Mutations that cause PSACH and EDM1 are restricted to a 27-kDa Ca\(^{2+}\) binding domain (type 3 repeat). This domain has 13 \(\text{Ca}^{2+}\)-binding loops with a consensus sequence that conforms to \(\text{Ca}^{2+}\)-binding loops found in EF hands. Most disease-causing mutations are found in the 11-kDa C-terminal region of this domain. We expressed recombinant native and mutant forms of the type 3 repeat domain (T3) and its 11-kDa C-terminal region (T3-Cterm). T3 and T3-Cterm bind \(13\) and \(8\) mol of \(\text{Ca}^{2+}\)/mol of protein, respectively. CD, one-dimensional proton, and two-dimensional \(^1\text{H}-^1\text{H}\) NMR HSQC spectra of \(\text{Ca}^{2+}\)-bound T3-Cterm indicate a distinct conformation that has little helical secondary structure, despite the presence of 13 EF hand \(\text{Ca}^{2+}\)-binding loops. This conformation is also formed within the context of the intact T3. 19 cross-peak positions found between \(8.0\) and \(11.4\) ppm are consistent with the presence of strong hydrogen bonding patterns, such as those in \(\beta\)-sheets. Removal of \(\text{Ca}^{2+}\) leads to an apparent loss of structure as evidenced by decreased dispersion and loss of all down field resonances. Deletion of Asp-470 (a mutation found in 22% of all PSACH and EDM1 patients) decreased the \(\text{Ca}^{2+}\)-binding capacity of both T3 and T3-Cterm by about 3 mol of \(\text{Ca}^{2+}\)/mol of protein. Two-dimensional \(^1\text{H}-^1\text{H}\) NMR HSQC spectra of mutated T3-Cterm showed little evidence of defined structure in the presence or absence of \(\text{Ca}^{2+}\). The data demonstrate that \(\text{Ca}^{2+}\) is required to nucleate folding and to maintain defined structure. Mutation results in a partial loss of \(\text{Ca}^{2+}\)-binding capacity and prevents \(\text{Ca}^{2+}\)-dependent folding. Persistence of an unstructured state of the mutated \(\text{Ca}^{2+}\) binding domain in COMP is the structural basis for retention of COMP in the rough endoplasmic reticulum of differentiated PSACH and EDM1 chondrocytes.

Pseudoachondroplasia (PSACH) and multiple epiphyseal dysplasia (EDM1) are related autosomal dominant skeletal dysplasias that are characterized by disproportionate short stature, lax joints, and early osteoarthritis (1, 2). These diseases are caused by mutations in cartilage oligomeric matrix protein (COMP). COMP, a member of the thrombospondin family of proteins (3), is found in bone growth plates, tendon bundles, ligaments, and smooth muscles (4–7). High concentrations of the protein are seen in the territorial matrix of chondrocytes (4) of mature cartilage from normal individuals but is depleted in the matrix of cartilage from PSACH and EDM1 patients (7). In these patients, COMP is sequestered in the rough endoplasmic reticulum (rER) (8), which becomes massively enlarged with alternating electron dense and electron lucent material (9, 10). Other extracellular matrix proteins, such as type IX collagen (8), are also retained in the enlarged rER of PSACH chondrocytes. These observations suggest that expression of the PSACH and EDM1 cellular phenotype results, at least in part, from improper protein processing/trafficking in chondrocytes.

COMP is a homopentamer composed of multidomain subunits. Each subunit consists of four domains as follows: an N-terminal association domain, followed by an epidermal growth factor-like type 2 repeat domain, a \(\text{Ca}^{2+}\)-binding type 3 repeat domain, and a C-terminal globular domain. Interestingly, 72 of 76 known disease-causing mutations in COMP are localized to the type 3 repeat \(\text{Ca}^{2+}\)-binding domain, and the majority of these are found in the C-terminal portion of this domain (8, 11, 12). Several general mechanisms could contribute to the observed cellular phenotype of PSACH and EDM1 chondrocytes. First, mutations may compromise the function of COMP. Second, mutations may cause gain of aberrant intracellular function or undesirable protein-protein interactions. A third possibility is that mutations may induce improper folding that leads to aggregation of partially folded intermediates or undesirable nonspecific interactions between mutant COMP and other proteins in the rER. Such a mechanism would be dominant negative and would be independent of the normal function of COMP.

To investigate these potential mechanisms, we have initiated a study to provide detailed information on the structural consequence of mutations in the \(\text{Ca}^{2+}\)-binding type 3 repeat domain of COMP. This domain has 13 \(\text{Ca}^{2+}\)-binding loops that conform to the consensus sequence of an EF hand \(\text{Ca}^{2+}\)-binding loop such as those found in calmodulin and troponin C (13). A typical EF hand, or helix-loop-helix \(\text{Ca}^{2+}\)-binding motif, consists of a calcium-binding loop flanked by helices of 10–12 amino acids. In contrast, 5 pairs of putative \(\text{Ca}^{2+}\)-binding loops

* This work was supported in part by National Institutes of Health Grant RO1HL45724 (to J. P.) and Robert A. Welch Foundation Grant AU-1144 (to J. P.). 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: Dept. of Biochemistry and Molecular Biology, Structural Biology Research Center, University of Texas, Houston Medical School, Houston, TX 77030. Tel.: 713-500-6061; Fax: 713-500-0652; E-mail: John.Putkey@uth.tmc.edu.
‡ The abbreviations used are: PSACH, pseudoachondroplasia; EDM1, multiple epiphyseal dysplasia; COMP, cartilage oligomeric matrix protein; T3, type 3 repeat domain; T3-Cterm, T3 and its 11-kDa C-terminal region; rER, rough endoplasmic reticulum; BME, 2-mercaptoethanol; DTT, dithiothreitol; MOPS, 4-morpholinepropanesulfonic acid; LDL, low density lipoprotein.
in COMP are separated by only 1 or 3 amino acids. Thus, the Ca\(^{2+}\) binding domain in COMP and other thrombospondins represent a novel utilization of the EF hand Ca\(^{2+}\)-binding loop. We report here comparative analysis of the wild type and mutant COMP Ca\(^{2+}\) binding domains, with a focus on the C-terminal 11-kDa subdomain that incurs the majority of disease-causing mutations. The recombinant mutant proteins have Asp-470\(^{-}\) deleted; a mutation that is found in 22% of all PSACH and EDM1 patients. The data are consistent with a model in which removal of Ca\(^{2+}\) or mutation leads to a loss of defined structure in the Ca\(^{2+}\) binding domain of COMP.

**EXPERIMENTAL PROCEDURES**

**Protein Expression and Isolation—**Selected regions of the COMP cDNA were amplified by PCR and inserted into the expression plasmid pET28d. All recombinant proteins encoded Ala as the second amino acid to enhance overall expression in *Escherichia coli* BL21 (DE3). The COMP derivatives could be easily isolated by standard procedures developed for cardiac troponin C (14). The cells were grown and induced as described previously (15) in either L-broth or M9 media prepared using \(^{15}N\)-ammonium chloride. Cells from 4 liters of culture were collected by centrifugation and resuspended in 120 ml of 50 mM Tris, pH 7.5, 0.2 mM EDTA, 0.1% BME. Resuspended cells were frozen at \(-70°C\), thawed, and then sonicated using a Branson sonicator with medium tip 3 times for 5 min at 50% power and 80% duty cycle while kept on ice. The lysed cells are centrifuged at 45,000 rpm in a Beckman Ti45 rotor. All COMP derivatives are found in the soluble fraction.

Soluble proteins were applied to a 2.5 x 30-cm Macro-Q (Bio-Rad) anion exchange column at a flow rate of 5 ml/min. The column was washed with 50 mM Tris, pH 7.5, 100 mM KCl, 0.2 mM EDTA, 0.1% BME until a base-line \(A_{\text{nm}}\) was achieved. Bound proteins were then eluted by a linear 100--500 mM KCl gradient (5 ml/min for 120 min). Fractions containing COMP proteins were pooled and made 1 M in ammonium sulfate, which precipitates most proteins but leave COMP soluble. After centrifugation, the soluble fraction was dialyzed against 50 mM Tris, pH 7.5, 100 mM KCl, 1 mM DTT, 0.1 mM oxidized DTT, and then applied to a Waters semi-preparative DEAE Protein Pak high performance liquid chromatography column. After washing, the bound proteins were eluted with a 100--500 mM KCl gradient. Fractions containing COMP proteins were pooled, dialyzed against 20 mM, ammonium bicarbonate, pH 8.0, 10 mM DTT, and lyophilized. The lyophilized proteins were then dialyzed in vitro by first resuspending in 50 mM Tris, pH 7.5, 8 M guanidine HCl, 10 mM DTT, 2 mM CaCl\(_2\). This was then diluted 50--100-fold into 50 mM Tris, pH 7.5, 100 mM KCl, 20 mM CaCl\(_2\), 1 mM DTT, 0.1 mM oxidized DTT at room temperature. The refolded protein was concentrated and dialyzed against the buffer of choice for a given experiment. Proteins were analyzed by electrospray mass spectroscopy to confirm sequence and by amino acid analysis to determine protein concentration. The N-terminal Met residue was removed by the bacterium.

**Calcium Binding—**Refolded proteins were transferred to a buffer of 50 mM MOPS, pH 7.0, 100 mM KCl, 2 mM CaCl\(_2\), 1 mM DTT, 0.1 mM oxidized DTT and concentrated to a protein concentration of about 4 mg/ml. A 1-ml volume of protein was placed in a Sildylider (Pierce) dialysis cassette and dialyzed at room temperature against the same buffer containing trace amounts of \(^{45}\)CaCl\(_2\). The time necessary to reach a linear 100--500 mM KCl gradient (5 ml/min for 120 min). Fractions containing COMP proteins were pooled and made 1 M in ammonium sulfate, which precipitates most proteins but leave COMP soluble. After centrifugation, the soluble fraction was dialyzed against 50 mM Tris, pH 7.5, 100 mM KCl, 1 mM DTT, 0.1 mM oxidized DTT, and then applied to a Waters semi-preparative DEAE Protein Pak high performance liquid chromatography column. After washing, the bound proteins were eluted with a 100--500 mM KCl gradient. Fractions containing COMP proteins were pooled, dialyzed against 20 mM, ammonium bicarbonate, pH 8.0, 10 mM DTT, and lyophilized. The lyophilized proteins were then dialyzed in vitro by first resuspending in 50 mM Tris, pH 7.5, 8 M guanidine HCl, 10 mM DTT, 2 mM CaCl\(_2\). This was then diluted 50--100-fold into 50 mM Tris, pH 7.5, 100 mM KCl, 20 mM CaCl\(_2\), 1 mM DTT, 0.1 mM oxidized DTT at room temperature. The refolded protein was concentrated and dialyzed against the buffer of choice for a given experiment. Proteins were analyzed by electrospray mass spectroscopy to confirm sequence and by amino acid analysis to determine protein concentration. The N-terminal Met residue was removed by the bacterium.

**NMR Spectroscopy—**Refolded proteins were transferred to a buffer of 50 mM MOPS, pH 7.0, 100 mM KCl, 2 mM CaCl\(_2\), 1 mM DTT, 0.1 mM oxidized DTT and concentrated to a protein concentration of about 4 mg/ml. A 1-ml volume of protein was placed in a Sildylider (Pierce) dialysis cassette and dialyzed at room temperature against the same buffer containing trace amounts of \(^{45}\)CaCl\(_2\). The time necessary to reach a linear 100--500 mM KCl gradient (5 ml/min for 120 min). Fractions containing COMP proteins were pooled and made 1 M in ammonium sulfate, which precipitates most proteins but leave COMP soluble. After centrifugation, the soluble fraction was dialyzed against 50 mM Tris, pH 7.5, 100 mM KCl, 1 mM DTT, 0.1 mM oxidized DTT, and then applied to a Waters semi-preparative DEAE Protein Pak high performance liquid chromatography column. After washing, the bound proteins were eluted with a 100--500 mM KCl gradient. Fractions containing COMP proteins were pooled, dialyzed against 20 mM, ammonium bicarbonate, pH 8.0, 10 mM DTT, and lyophilized. The lyophilized proteins were then dialyzed in vitro by first resuspending in 50 mM Tris, pH 7.5, 8 M guanidine HCl, 10 mM DTT, 2 mM CaCl\(_2\). This was then diluted 50--100-fold into 50 mM Tris, pH 7.5, 100 mM KCl, 20 mM CaCl\(_2\), 1 mM DTT, 0.1 mM oxidized DTT at room temperature. The refolded protein was concentrated and dialyzed against the buffer of choice for a given experiment. Proteins were analyzed by electrospray mass spectroscopy to confirm sequence and by amino acid analysis to determine protein concentration. The N-terminal Met residue was removed by the bacterium.

**Circular Dichroism—**CD spectra were collected on a JASCO J-710 spectrophotometer using a cylindrical cuvette with a 0.5-mm path length. Proteins that were refolded in vitro were dialyzed against 10 mM MOPS, pH 7.0, 100 mM KCl, 1 mM reduced DTT, 0.1 mM oxidized DTT, 2 mM CaCl\(_2\), and then diluted to a concentration of 5--20 \(\mu\)M in the same buffer. CD spectra were collected before and after the addition of 3 mM EDTA.

**NMR—**NMR spectroscopy was carried out on a Bruker Avance DRX 600, or Varian Unity Plus 600 spectrometers equipped with an inverse triple resonance gradient probes. Proteins used for NMR samples varied in protein concentration from 1 to 2 \(\mu\)M and were exchanged into a NMR buffer consisting of 20 mM Tris, pH 7.0, 100 mM KCl, 20 mM CaCl\(_2\),...
Typically, the Ca\(_2\)RER is rich in Ca\(^{2+}\) repeat (17). The native folding environment of COMP in the NEB pre-activated at residues 101, 124, and 721, which are outside the type \(\alpha\) helices. Tissue-derived human COMP is glycosylated in which Asp-470 was deleted to mimic the most common disease mutations. The molecular weight of the recombinant proteins was described previously (16). The N- and C-terminal regions of the type 3 repeat have distinguishing characteristics. This includes a different distribution of short and long interloop spacers, different patterns of Cys residues, and the presence of His residues only in the C-terminal half. These features, together with the fact that 66% of disease-causing mutations are located in the C-terminal half of the type three repeat, led us to construct an expression plasmid for an 11-kDa protein called T3-Cterm which encodes amino acids 420–524 (loops 2–13) and \(\alpha\) points (10). A pair of mutated derivatives, T3\(\Delta\)D and T3-Cterm\(\Delta\)D, was generated in which Asp-470 was deleted to mimic the most common disease-causing mutation. The net effect of this mutation was to remove the single amino acid between loops 10 and 11.

We used a bacterial system to express COMP proteins because it provides high levels of expression and efficient labeling of proteins with stable isotopes required for multidimensional heteronuclear NMR. Tissue-derived human COMP is glycosylated at residues 101, 124, and 721, which are outside the type 3 repeat (17). The native folding environment of COMP in the rER is rich in Ca\(^{2+}\) and contains chaperones such as protein-disulfide isomerase, which may aid in the proper folding of the Cys-rich type 3 repeat. In contrast, the intracellular environment of bacteria is low in Ca\(^{2+}\) and high in reducing potential. The proteins used in this study have all been refolded in vitro as described under "Experimental Procedures." Fig. 2 shows an SDS gel of the purified native and mutant proteins. Both T3 and T3-Cterm proteins migrate with apparent relative molecular masses that are greater than predicted from primary sequence. The molecular weight of the recombinant proteins was confirmed by electrospray mass spectroscopy.

Ca\(^{2+}\) Binding to Type 3 Repeat Domain and Subdomain—Typically, the Ca\(^{2+}\)-binding properties of proteins are determined by sequential addition of Ca\(^{2+}\) to the apoprotein. Unlike regulatory EF hand proteins that are designed to cycle between apo- and Ca\(^{2+}\)-bound forms, COMP may require Ca\(^{2+}\) to maintain a proper structure that is not readily regained if Ca\(^{2+}\) is removed from the folded protein and then replenished. Thus, we generated Ca\(^{2+}\)-binding isotherms for recombinant COMP proteins by sequential dilution of Ca\(^{2+}\) after the proteins had been refolded in the presence of Ca\(^{2+}\). Fig. 3 compares the Ca\(^{2+}\)-binding curves of T3 and T3-Cterm with T3\(\Delta\)D and T3-Cterm\(\Delta\)D. T3 binds 12–13 mol of Ca\(^{2+}\) per mol of protein with a \(K_{D,10}\), of about 0.25 mM. T3-Cterm binds 8–9 mol of Ca\(^{2+}\) per mol of protein with a \(K_{D,10}\), of about 0.5 mM. The mutant T3\(\Delta\)D and T3-Cterm\(\Delta\)D bind less Ca\(^{2+}\) than the wild type proteins at all concentrations of free Ca\(^{2+}\), and with slightly lower affinity. At 2 mM Ca\(^{2+}\) both T3 and T3-Cterm bind about 3 more mol of Ca\(^{2+}\)/mol of protein than the mutated derivatives. This suggests that the effect of deletion of Asp-470 in the C-terminal half of the type 3 repeat domain is restricted to Ca\(^{2+}\)-binding sites in this subdomain.

Circular Dichroism—The predicted secondary structure of T3-Cterm (PSIPRED V2.0) indicates a high content of coiled structures. To confirm this, we performed circular dichroism on all proteins in the presence and absence of Ca\(^{2+}\) as an initial
Mutations in COMP Cause Loss of Structure

**Mutations in COMP Cause Loss of Structure**

**FIG. 4.** CD spectra for wild type and mutant T3 and T3-Cterm in the presence and absence of Ca$^{2+}$.

**screen for differences in protein structure.** Preliminary NMR studies indicated that low ionic strength has a significant effect on the structure of T3-Cterm. Because chloride ions interfere with CD measurements in the far UV range, and because fluoride precipitates with Ca$^{2+}$, we performed all measurements in 100 mM KCl and show CD spectra from 195 to 240 nm. Fig. 4 compares the CD spectra of the wild type and mutant T3 proteins (Fig. 4A) and T3-Cterm (Fig. 4B). The large negative ellipticity at 200 nm and small negative values at 220 nm indicate high content of coiled secondary structure and relatively little helical structure for all proteins. The CD data shows several overall trends for T3 and T3-Cterm proteins. The wild type proteins displayed an increase in negative ellipticity at ~200 nm and a shift to higher wavelength upon addition of Ca$^{2+}$ to the apoproteins. Unlike the wild type proteins, mutant ΔD proteins behaved differently. Addition of Ca$^{2+}$ to apoT3ΔD resulted in a decrease in negative ellipticity at ~200 nm, a similar shift to higher wavelength as seen for wild type T3, and less change than T3 in [θ] at all higher wavelengths. T3-CtermΔD exhibited similar characteristics as T3ΔD at ~200 nm upon Ca$^{2+}$ addition but to a greater degree. Unlike T3-Cterm, there is a substantial decrease in ellipticity at all wavelengths for T3-CtermΔD. The data suggest the T3-Cterm proteins undergo different structural transitions in response to Ca$^{2+}$. The changes observed for T3-Cterm could represent an unfolded to folded transition, but for T3-CtermΔD the results are most consistent with an unfolded to aggregated transition.

**NMR—CD** measures the sum of secondary structures in a protein. It is sensitive to structural changes, but it does not provide great insight into the nature of these changes. We employed heteronuclear NMR to determine the effect of Ca$^{2+}$ and mutation on the structure of the type 3 repeat. The focus of the present study is T3-Cterm because it incurs the majority of disease-causing mutations. All spectra were collected using proteins that had been refolded in vitro in the presence of Ca$^{2+}$ and a 10:1 ratio of reduced and oxidized DTT (1/0.1 mM). Subsequent addition of excess DTT to the refolded protein did not greatly affect the NMR spectra (data not shown). Calcium was removed from the refolded protein to generate the apoform by sequential dilution and concentration with a buffer containing EDTA.

Fig. 5 compares the amide region of the one-dimensional proton spectra for wild type T3-Cterm (Fig. 5A) and T3-CtermΔD (Fig. 5B) in the presence and absence of Ca$^{2+}$. The spectrum of Ca$^{2+}$-bound T3-Cterm is well dispersed with a number of distinct down field-shifted resonances. Removal of Ca$^{2+}$ from wild type T3-Cterm dramatically reduces or eliminates the down field shifted resonances and generally decreases resonance dispersion. Spectra for the mutant T3-CtermΔD show little evidence of the down field shifted resonances, and the resonances are much less dispersed than for wild type T3-Cterm in the presence or absence of Ca$^{2+}$. The two groups of chemical shifts at about 6.8 and 7.6 ppm correspond to the side chain amides of Gln and Asn residues. Thus, chemical shifts of the backbone H$^N$ for apoT3-Cterm and T3-ΔD in the presence or absence of Ca$^{2+}$ are found in a narrow spectral window between 7.8 and 8.8 ppm. This lack of dispersion is characteristic of unstructured protein (18).

Fig. 6, A and B, shows the $^1$H-$^1$5N HSQC spectrum of T3-Cterm in the presence and absence of Ca$^{2+}$, respectively. The spectrum of Ca$^{2+}$-bound T3-Cterm is well dispersed and characteristic of a protein with defined structure. 93 cross-peaks (excluding side chain amides) resolved in this contour level, which agrees well with the 96 backbone amide protons predicted from the primary sequence. The spectrum shows 19 cross-peaks down field of 9.0 ppm in the proton dimension. These down field resonances likely correspond to protons that participate in strong hydrogen-bonded structures, such as β-sheets. Four of the down field cross-peaks are identified as the H$^N$ protons of the four glycine residues at the 6th position of Ca$^{2+}$-binding loops 9–12.3 The H$^N$ chemical shifts of Gly residues in the 6th position of EF hand Ca$^{2+}$-binding loops of calmodulin and troponin C shift down field in the presence of Ca$^{2+}$ (19–22). Removal of Ca$^{2+}$ causes a significant spectral change that is dominated by a substantial decrease of dispersion in the proton dimension (Fig. 6B). This is consistent with the one-dimensional spectra in Fig. 5. Most notably, all of the down field resonances seen in the presence of Ca$^{2+}$ are absent upon removal of Ca$^{2+}$. There is little overlap of resonances between the spectra for T3-Cterm in the presence and absence of Ca$^{2+}$.

Fig. 6, C and D, shows the $^1$H-$^1$5N HSQC spectra for T3-CtermΔD in the presence and absence of Ca$^{2+}$, respectively. In contrast to wild type T3-Cterm, the mutant protein shows little resonance dispersion in the presence of Ca$^{2+}$. The 19 down field

---

3 Identification of the Gly residues was made by assigning the neighboring spin systems of each Gly in Ca$^{2+}$-bound $^{13}$N,$^{15}$C-enriched T3 C-term using triple resonance experiments HNCACB, CBCA(CO)NH, (C(CO)NH), and H(CCO)NH.
not overlap with corresponding resonances in T3 (data not shown). Thus, we conclude that Ca\(^{2+}\)-bound T3-Cterm assumes a single dominant conformation in solution and that T3-Cterm is an independently folding subdomain of the larger type 3 repeat Ca\(^{2+}\) binding domain in COMP.

We wished to collect \(^{1}H,^{15}N\) HSQC spectra for T3-Cterm at sub-saturating levels of Ca\(^{2+}\). Rather than add Ca\(^{2+}\) to the apoprotein, which may not readily refold under conditions used to collect NMR spectra, we elected to dilute Ca\(^{2+}\) in a sample that had been refolded in the presence of calcium. The total Ca\(^{2+}\) in a 2 mM sample of T3-Cterm was diluted to a final predicted total Ca\(^{2+}\) concentration of 0.1 mM by sequential rounds of dilution and concentration. Fig. 8A compares the spectra of the same protein at high (black) or low (red) concentrations of Ca\(^{2+}\). All resonances observed at saturating Ca\(^{2+}\) are also present at the lower Ca\(^{2+}\) concentration, but the relative intensities were about 30% those observed for the Ca\(^{2+}\)-saturated protein. At the lower concentration of Ca\(^{2+}\), additional resonances are observed between 7.8 and 8.8 ppm that are characteristic of the apoprotein. Fig. 8B expands the boxed region in A. The cross-peaks in B are the sum of resonances observed in spectra of the Ca\(^{2+}\)-saturated (C) and apoproteins (D). The presence of distinct structured and unstructured populations at sub-saturating levels of Ca\(^{2+}\) is typical of a slow exchange process associated with high affinity ligand binding. The equilibrium Ca\(^{2+}\) binding constant of COMP is on the order of 10\(^{-4}\) M\(^{-1}\). Thus we were surprised to see evidence of slow exchange for COMP. Additional studies will be necessary to characterize this phenomenon.

The results in Fig. 8 suggest that very low affinity binding of Ca\(^{2+}\) to mutated T3-Cterm\(\Delta D\) could be detected as a minor population of protein with a dispersed HSQC spectrum. Thus, \(^{1}H,^{15}N\) HSQC spectra of T3-Cterm\(\Delta D\) were analyzed at a very low contour level. Fig. 9 shows this analysis in red with the corresponding spectra of Ca\(^{2+}\)-saturated T3-Cterm shown in black. Only the down field region of the spectral window is shown because the very intense signals at this low contour level obscure comparisons in other spectral regions. At a low contour level, dispersed cross-peaks are observed for T3-Cterm\(\Delta D\). The relative intensities of the cross-peaks seen for T3-Cterm\(\Delta D\) are less than 2% those of wild type protein, and the chemical shifts are not identical between the two proteins. These data suggest that a small population of the mutant protein exists in a more structured state in the presence of 20 mM free Ca\(^{2+}\). This likely reflects a dramatic decrease in Ca\(^{2+}\) affinity, rather than complete inhibition of Ca\(^{2+}\) binding due to deletion of Asp-470. Nevertheless, deletion of Asp-470 causes an unstructured or scrambled state in the bulk of the population of protein in the presence or absence of biologically relevant concentrations of Ca\(^{2+}\).

**DISCUSSION**

Dominant negative mutations in COMP cause the rER in chondrocytes of PSACH and EDM1 patients to become grossly enlarged and to retain COMP and other extracellular matrix proteins. 72 of 76 known disease-causing mutations are located in the type 3 repeat Ca\(^{2+}\)-binding domain of COMP. 70% of these mutations are in the C-terminal portion of the type 3 repeat, and 22% involve deletion of 1 of 5 Asp codons (Asp-469—73). This essentially deletes the single Asp-470 spacer between Ca\(^{2+}\)-binding loops 10 and 11 (see Fig. 3). The goal of the current study was to apply NMR to gain insights into the structural consequence of this deletion. We report the first high resolution NMR data for a member of the thrombospondin family of proteins. The results have general implications with respect to the effect of Ca\(^{2+}\)-binding proteins in normal and disease states.
Primary sequence analysis suggests that the N- and C-terminal portions of this domain have distinct structures. Our data show that T3-Cterm is indeed a structurally distinct subdomain that can fold independent of the intact type 3 repeat domain. We hypothesize that the N-terminal portion of the type 3 repeat is also a structurally distinct subdomain. Independence of the N and C subdomains of the type 3 repeat may insulate them from global effects of mutations. For example, deletion of Asp-470 in the C-terminal subdomain may minimally affect the structure of the N-terminal subdomain. This would be analogous to what is observed for a dimer of LDL-A Ca²⁺/H11001 ligand binding domains from the LDL receptor, which are structurally independent, and are insulated from the effects of mutation in the paired domain (23). Structurally distinct subdomains in the type 3 repeat may confer important functional characteristics. For example, the adjacent globular domain is known to bind collagen (24), and mutations in the type 3 repeat affect collagen binding (25). Although the biological significance of these observations is unknown, the structural integrity of both type 3 repeat subdomains may be necessary to orient properly the globular domain for optimal ligand binding.

The data presented here argue that the Ca²⁺-binding loops in COMP provide a coordination sphere that is similar if not identical to that of a canonical EF hand. Four of the Ca²⁺-binding loops in T3-Cterm (loops 9–12) have Gly at the 6th position (Gly-440, -463, -476, and -499). Glycine residues at this relative position in the canonical EF hands of CaM and tropinin C allow a tight −90° turn of the Ca²⁺-binding loops (13). In the presence of Ca²⁺, the amide proton of this Gly residue forms a strong hydrogen bond with the carboxyl oxygen of the acidic residue at loop position 1. This confers an unusually large down field chemical shift (8.9–11.0 ppm) on the HN of this Gly in the EF hands of cardiac troponin C (19, 20), skeletal troponin C (21), and calmodulin (22) that can be diagnostic for Ca²⁺ binding. The circled cross-peaks centered on 10 ppm in Fig. 6 correspond to the HN of Gly-440, -463, -476, and -499 in Ca²⁺-binding loops 9–12, which we attribute to a similar structural feature associated with Ca²⁺ binding. Thus we propose that loops 9–12 provide a coordination geometry and pentagonal bipyramidal distribution of ligands that is similar to prototypical EF hands. These four Ca²⁺-binding loops may constitute a stable core structure for T3-Cterm that is hypersensitive to mutations.

Although Ca²⁺-binding loops in T3-Cterm conform to those found in a canonical EF hand, it is unlikely that the Ca²⁺-binding domain of COMP consists of multiple copies of true helix-loop-helix EF hand motifs. These observations are consistent with the primary sequences of T3 and T3-Cterm, which are dominated by Ca²⁺-binding loops. The first 9 amino acids in an EF hand Ca²⁺-binding loop have a coiled or β-strand secondary structure, whereas the last three amino acids form the first portion of the F helix. Four pairs of Ca²⁺-binding loops in COMP are separated by only a single amino acid, and most of the longer (11 amino acids) interloop spacers have two Pro
residues. It is likely that these features inhibit extensive helix formation between Ca\textsuperscript{2+}-binding loops. In addition, the 19 amide protons with chemical shifts between 9.0 and 11.4 ppm in the spectra for Ca\textsuperscript{2+}-bound T3-Cterm are consistent with the presence of strong hydrogen bonding patterns such as those found in β-sheets (26), but the structural context remains in question.

There is precedent for the use of EF hand-like Ca\textsuperscript{2+}-binding loops outside the context of α-helices. Vyas et al. (27) reported a variant EF hand in bacterial galactose-binding protein that has a partial EF loop bordered by helix and β-strand secondary structures. The Ca\textsuperscript{2+}-binding motif found in bacterial alkaline protease (28, 29) and the β-propeller domain of integrin α\textsubscript{v}β\textsubscript{3} (30) have a short Ca\textsuperscript{2+}-binding loop flanked by β-strands. However, the Ca\textsuperscript{2+}-binding loops in these proteins differ significantly from the 12-amino acid loop of an EF hand in that the −Z Ca\textsuperscript{2+} coordination site is not provided by position 12 of the loop. We predict that the type 3 repeat domain in COMP has multiple copies of a motif that uses canonical EF hand Ca\textsuperscript{2+}-binding loops in a novel context.

The NMR spectra reported here show that removal of Ca\textsuperscript{2+} from T3-Cterm results in a general loss of proton resonance dispersion. All of the down field amide resonances seen for Ca\textsuperscript{2+} bound T3-Cterm are absent from the one-dimensional \textsuperscript{1}H and two-dimensional \textsuperscript{1}H-\textsuperscript{15}N HSQC spectra of the apoprotein. This is consistent with a fully unstructured state or an ensemble of rapidly converting structures (18) for T3-Cterm in the absence of Ca\textsuperscript{2+}. We hypothesize that Ca\textsuperscript{2+} binding nucleates folding and is required to maintain a folded state with defined tertiary topology stabilized by hydrogen bonding patterns. This defined topology is lost in the absence of Ca\textsuperscript{2+}. Under appropriate oxidizing conditions, removal of Ca\textsuperscript{2+} may induce scrambling of disulfide bonds. However, we do not feel that potential disulfide scrambling is the primary cause of loss of defined structure, because high concentrations of DTT fail to promote resonance dispersion in the apoprotein (data not shown).

Whereas a number of EF hand containing proteins (e.g. CaM and skeletal troponin C, calbindin D\textsubscript{9k}, calcyclin, S100B) are well structured in the absence of Ca\textsuperscript{2+}. The removal of Ca\textsuperscript{2+} can have severe global structural consequences for proteins with Ca\textsuperscript{2+}-binding motifs. The cysteine-rich LDL-A modules of the LDL receptor and the Tva cellular receptor both require Ca\textsuperscript{2+} for proper folding (31, 32). Removal of Ca\textsuperscript{2+} from these motifs results in an unstructured or scrambled ensemble. The overall effect of removal of Ca\textsuperscript{2+} on the NMR spectra of the

**FIG. 7.** \textsuperscript{2+}\textsuperscript{1}H-\textsuperscript{15}N HSQC of T3-Cterm and T3. The down field amide region of Ca\textsuperscript{2+}-bound T3 (red) and T3-Cterm (black) are overlaid to highlight similarities and differences. Buffer conditions are described in the Fig. 5.

**FIG. 8.** \textsuperscript{1}H-\textsuperscript{15}N HSQC spectra of T3-Cterm at different levels of Ca\textsuperscript{2+}. A compares HSQC spectra of T3-Cterm at high (black) and low (red) levels of Ca\textsuperscript{2+}. Protein was refolded in the presence of 20 mM Ca\textsuperscript{2+} and concentrated to about 2 mM in a buffer that contained 20 mM Ca\textsuperscript{2+} as described under “Experimental Procedures.” The total Ca\textsuperscript{2+} was about 35 mM. After collection of an HSQC spectrum, total Ca\textsuperscript{2+} in the sample was effectively diluted to about 0.1 mM by successive rounds of dilution of the protein into a Ca\textsuperscript{2+}-free buffer followed by concentration to the original volume. The relative intensities of the down field cross-peaks at low Ca\textsuperscript{2+} levels were about 30% of the Ca\textsuperscript{2+}-saturated protein. Contour levels were adjusted optimally to observe cross-peaks in both spectra. B shows and an expanded view of the indicated region from A. The black contours correspond to resonances found in the Ca\textsuperscript{2+}-saturated protein (C), and the green contours correspond to those found in the fully apoprotein (D).
LDL-A motif is very similar to that reported here for T3-Cterm. Dramatic structural effects of removal of Ca\(^{2+}\) have also been observed in EF hand proteins from invertebrates. The EF hand protein calerythin from *Saccharopolyspora erythraea* is in equilibrium between a structured and less structured form in the absence of Ca\(^{2+}\) (33), and the sarcoplasmic EF hand protein NSCP from the invertebrate *Nereis diversicolor* appears unstructured in the absence of Ca\(^{2+}\) (34).

Deletion of Asp-470 causes a loss of about 3 mol of Ca\(^{2+}\)/mol of protein for the intact type 3 repeat domain as well as T3-Cterm. This is consistent with the results of Chen et al. (35) who observed that intact pentameric \(\Delta 470\) COMP bound 2–4 mol less Ca\(^{2+}\) per subunit than the wild type protein. The NMR spectra show that T3-Cterm\(\Delta D\) is largely unstructured in the presence or absence of 20 mM free Ca\(^{2+}\). Because the wild type protein is unstructured in the absence of Ca\(^{2+}\), we conclude that inhibition of Ca\(^{2+}\) binding is the primary reason the mutated protein fails to adopt defined structure. These data are consistent with electron micrographs of native and mutant COMP pentamers (36). In the presence of Ca\(^{2+}\) the stalk region of native COMP homopentamers appears compact, and in the absence of Ca\(^{2+}\) they appear more extended. The stalk region in mutant COMP homopentamers appears more extended in the presence or absence of Ca\(^{2+}\). The data presented here suggest that the more extended conformation is not the result of a misfolded yet defined structure but rather to an unfolded or unstructured Ca\(^{2+}\) binding domain that may have a high content of extended and flexible random coil.

It is not presently clear by what mechanisms the unstructured mutated type 3 repeat domain in COMP promotes intracellular protein retention resulting in a grossly enlarged rER in PSACH and EDM1 chondrocytes. Retention of COMP is gradual and requires or is at least enhanced by expression of other extracellular matrix proteins in the rER of differentiated chondrocytes (36, 37). This suggests the possibility of specific interactions between mutated COMP and other extracellular proteins that are being processed in the rER. But the unstructured mutated Ca\(^{2+}\) binding domain of COMP raises the potential for nonspecific interactions as well. Unstructured or partially folded proteins can associate in a concentration-dependent manner with other partially folded proteins, primarily via hydrophobic interactions. Intracellular protein crowding would enhance such interactions (38) and possibly promote intracellular deposition of long lived partially folded intermediates that result from pathological conditions (39). Differentiated chondrocytes may provide a crowded protein environment that favors heterologous interactions between mutated COMP and a variety of other large extracellular proteins that are being processed in the rER.

The potential role of chaperone proteins must also be considered as suggested previously (16, 35). Creemers et al. (40) demonstrated that binding of BiP to lymphoma proprotein convertase prevents aggregation but slows protein maturation. Thrombospondin transiently associates with BiP, Grp94, and Erp72 (41, 42), and native and mutant COMP are in close proximity to these proteins as well (43). Prolonged association of chaperonins with an unstructured COMP Ca\(^{2+}\) binding domain may inhibit protein secretion and/or disrupt the equilibrium distribution of chaperonins between other proteins in the crowded chondrocyte rER, thus leading to global trafficking problems in the active rER of differentiated chondrocytes.

In summary, we have shown that the type three repeat Ca\(^{2+}\) binding domain of COMP is composed of distinct subdomains. Removal of Ca\(^{2+}\), and a disease-causing mutation, both led to a dramatic loss of structure in the C-term subdomain. This is likely central to the pathogenesis of PSACH and MED and is a striking example of a disease-causing mutations that cause a loss of structure in a human Ca\(^{2+}\)-binding protein.

**Acknowledgments**—We thank Dr. Ed Nikonowics at Rice University and Dr. David Gorenstein and the Sealy Center for Structure Biology at the University of Texas Medical Branch, Galveston, TX, for the use of their respective NMR facilities. We also thank Wen Liu and Natasha Bogdanov for their expert technical assistance.

**REFERENCES**

1. Cooper, R. R., Ponseti, I. V. & Maynard, J. A. (1973) *J. Bone Jr. Surg. Am.* 55, 475–484
2. McKeand, J., Rotta, J. & Hecht, J. T. (1996) *Am. J. Med. Genet.* 63, 406–410
3. Sage, E. H. & Bornstein, P. (1991) *J. Biol. Chem.* 266, 14831–14834
4. Hedblom, E., Antonsson, P., Hjerpe, A., Aeschlimann, D., Paulsson, M., Rosa-Pimentel, E., Sommarin, Y., Wendel, M., Oldberg, A. & Heinegard, D. (1992) *J. Biol. Chem.* 267, 6132–6136
5. Shen, Z., Heinegard, D. & Sommerin, Y. (1995) *Matrix Biol.* 14, 733–781
6. DiCesare, P., Hauser, N., Lehman, D., Pasumarti, S. & Paulsson, M. (1994) *FEBS Lett.* 354, 237–240
7. Hecht, J. T., Deere, M., Putsam, E., Cole, W., Vertel, B., Chen, H. & Lawler, J. (1998) *Matrix Biol.* 17, 269–278
8. Maddox, B. K., Deere, D. R., Sakai, L. Y., Charbonneau, N. L., Morris, N. P., Ridgway, C. C., Boswell, B. A., Sussman, M. D., Horton, W. A., Bachinger, H. P. & Hecht, J. T. (1997) *J. Biol. Chem.* 272, 30993–30997
9. Maynard, J. A., Cooper, R. R. & Poindessiat, I. V. (1972) *Lab. Invest.* 26, 40–44
10. Stanescu, V., Stanescu, R. & Maroteaux, P. (1984) *J. Bone Jr. Surg. Am.* 66, 817–836
11. Hecht, J. T., Francomano, C. A., Briggs, M. D., Deere, M., Conner, B., Horton, W. A., Warman, M., Cohn, D. H. & Blanton, S. H. (1993) *Genomics* 18, 661–666
12. Hecht, J. T., Nelson, L. D., Crowder, E., Wang, Y., Elder, F. F., Harrison, W. R., Francomano, C. A., Prange, C. K., Lennon, G. G. & Deere, M. (1995) *Nat. Genet.* 10, 325–329
13. Strynadka, N. C. J. & James, M. N. G. (1989) *Annu. Rev. Biochem.* 58, 951–998
14. Putkey, J. A., Sweeney, H. L. & Campbell, S. T. (1989) *J. Biol. Chem.* 264, 30997–31002
15. Lin, X., Krudy, G. A., Howarth, J., Brito, R. M. M., Tsuda, S., Kay, C. M., Smillie, L. B. & Sykes, B. D. (1997) *Matrix Biol.* 16, 325–329
16. Hou, J., Putkey, J. A. & Deere, M. (2000) *Cell Calcium* 27, 309–314
17. Sogn, J., Boynton, R. E., McInerney, A., Marshak, D. R., Olsom, H. H., Heinegard, D. & Barry, F. P. (1997) *J. Biol. Chem.* 272, 14120–14126
18. Dyonon, H. J. & Wright, P. E. (1998) *Nat. Struct. Biol.* 5, 499–503
19. Krudy, G. A., Brito, R. M. M., Putkey, J. A. & Rosevear, P. R. (1992) *Biochemistry* 31, 1595–1602
20. Sia, S., Li, M. X., Spyracopoulos, L., Gagne, S. M., Liu, W., Putkey, J. A. & Sykes, B. D. (1997) *J. Biol. Chem.* 272, 18216–18221
21. Li, M. X., Gagne, S. M., Teuda, S., Ray, C. M., Smillie, L. B. & Sykes, B. D. (1995) *Biochemistry* 34, 8330–8340
22. Ikura, M., Kay, L. E. & Bax, A. (1990) *Biochemistry* 29, 4659–4667
23. Bieri, S., Atkins, A. R., Lee, H. T., Winnor, D. J., Smith, R. & Kroon, P. A.
Mutations in COMP Cause Loss of Structure

(1998) Biochemistry 37, 10994–11002
24. Rosenberg, K., Olsson, H., Morgelin, M. & Heinegard, D. (1998) J. Biol. Chem. 273, 20397–20403
25. Thur, J., Rosenberg, K., Nitche, D. P., Pihlajamaa, T., Ala-Kokko, L., Heinegard, D., Paulsson, M. & Maurer, P. (2001) J. Biol. Chem. 276, 6083–6092
26. Pardi, A., Wagner, G. & Wuthrich, K. (1983) Eur. J. Biochem. 137, 445–454
27. Vyas, N. K., Vyas, M. N. & Quiocho, F. A. (1987) Nature 327, 635–638
28. Springer, T. A., Hun, J. & Takagi, J. (2000) Cell 102, 275–277
29. Baumann, U., Wu, S., Flaherty, K. M. & McKay, D. D. (1993) EMBO J. 12, 3357–3364
30. Xiong, J.-P., Stehle, T., Diefenbach, B., Zhang, R., Dunker, R., Scott, D. L., Joachimiak, A., Goodman, S. L. & Arnaout, M. A. (2001) Science 294, 339–345
31. Wang, Q.-Y., Dolmer, K., Huang, W., Gettins, P. G. W. & Rong, L. (2001) J. Virol. 75, 2051–2058
32. Blacklow, S. C. & Kim, P. S. (1996) Nat. Struct. Biol. 3, 758–761
33. Aitio, H., Laakso, T., Pihlajamaa, T., Torkkeli, M., Kilpelainen, I., Drakenberg, T., Seimaa, R. & Annila, A. (2001) Protein Sci. 10, 74–82
34. Precheur, B., Cox, J. A., Petrova, T., Misplenter, J. & Cruescu, C. T. (1996) FEBS Lett. 395, 89–94
35. Chen, H., Deere, M., Hecht, J. T. & Lawler, J. (2000) J. Biol. Chem. 275, 26538–26544
36. Delot, E., Brodie, S. G., King, L. M., Wilcox, W. R. & Cohn, D. H. (1998) J. Biol. Chem. 273, 26692–26697
37. Hecht, J. T., Montufar-Solis, D., Decker, G., Lawler, J., Daniels, K. & Duke, P. J. (1998) Matrix Biol. 17, 625–633
38. Ellis, J. R. (2001) Trends Biochem. Sci. 26, 597–604
39. Minton, A. P. (2000) Curr. Opin. Struct. Biol. 10, 34–39
40. Creemers, J. W. M., Van de Loo, J.-W. H. P., Plets, E., Hendershot, L. M. & Van de Ven, W. H. M. (2000) J. Biol. Chem. 275, 38842–38847
41. Prabakaran, D., Kim, P. S., Dixit, V. M. & Arvan, P. (1996) Eur. J. Cell Biol. 70, 134–141
42. Kuznetsov, G., Chen, L. B. & Nigam, S. N. (1997) J. Biol. Chem. 272, 3057–3063
43. Hecht, J. T., Hayes, E., Snuggs, M., Decker, G., Montufar-Solis, D., Mwalle, F., Stevens, J. & Duke, P. J. (2001) Matrix Biol. 20, 251–261
44. DiCesare, P. E., Morgelin, M., Mann, K. & Paulsson, M. (1994) Eur. J. Biochem. 223, 927–937