Metaphyseal Chondrodysplasia Type Schmid Mutations Are Predicted to Occur in Two Distinct Three-dimensional Clusters within Type X Collagen NC1 Domains That Retain the Ability to Trimerize*

(Received for publication, July 30, 1998, and in revised form, October 30, 1998)

Debora S. Marks‡, Carl A. Gregory‡, Gillian A. Wallis, Andy Brass, Karl E. Kadler, and Raymond P. Boot-Handford§

From the Wellcome Trust Centre for Cell-Matrix Research, School of Biological Sciences, University of Manchester, Oxford Road, Manchester, M13 9PT United Kingdom

Metaphyseal chondrodysplasia type Schmid (MCDS) is caused by mutations in COL10A1 that are clustered in the carboxyl-terminal non-collagenous (NC1) encoding domain. This domain is responsible for initiating trimerization of type X collagen during biosynthesis. We have built a molecular model of the NC1 domain trimer based on the crystal structure coordinates of the highly homologous trimeric domain of ACRP30 (adipocyte complement-related protein of 30 kDa or AdipoQ). Mapping of the MCDS mutations onto the structure reveals two specific clusters of residues as follows: one on the surface of the monomer which forms a tunnel through the center of the assembled trimer and the other on a patch exposed to solvent on the exterior surface of each monomeric unit within the assembled trimer. Biochemical studies on recombinant trimeric NC1 domain show that the trimer has an unusually high stability not exhibited by the closely related ACRP30. The high thermal stability of the trimeric NC1 domain, in comparison with ACRP30, appears to be the result of a number of factors including the 17% greater total buried solvent-accessible surface and the increased numbers of hydrophobic contacts formed upon trimerization. The 27 amino acid sequence present at the amino terminus of the NC1 domain, which has no counterpart in ACRP30, also contributes to the stability of the trimer. We have also shown that NC1 domains containing the MCDS mutations Y598D and S600P retain the ability to homotrimerize and heterotrimerize with wild type NC1 domain, although the trimeric complexes formed are less stable than those of the wild type molecule. These studies suggest strongly that the predominant mechanism causing MCDS involves a dominant interference of mutant chains on wild type chain assembly.

Type X collagen is a short chain, homotrimeric collagen (α1,X)3 expressed specifically by hypertrophic chondrocytes in the endochondral growth plate (1). The expression of type X collagen is also re-activated during fracture repair and in osteoarthritis (2–4). The human α1(X) collagen chain consists of a short amino-terminal non-collagenous domain 2 of 37 amino acids followed by a triple helix-forming collagenous domain with 154 Gly-X-Y repeats and a carboxyl-terminal non-collagenous domain (NC1)1 of 161 amino acids (5). The α1(X)3 molecule is thought to assemble into a hexagonal lattice within the extracellular matrix in a fashion similar to that of type VIII collagen (6, 7). Type X collagen is part of a family of collagen-like proteins sharing a condensed gene structure, a collagen triple helical domain, and in particular, a highly conserved carboxy-terminal non-collagenous (NC1-like) domain (8, 9). This family includes collagen types X and VIII, C1q component of complement, hibernation proteins (10), cerebellin (11, 12), and ACRP30, an abundant serum protein implicated in energy homeostasis and obesity (13).

The precise function of type X collagen remains to be determined (14–16), but mutations in the COL10A1 gene cause metaphyseal chondrodysplasia type Schmid (MCDS), an autosomal dominant form of human skeletal dysplasia (17, 18). An intriguing finding is that virtually all of the mutations causing MCDS occur within the carboxy-terminal non-collagenous NC1 domain of type X collagen (Ref. 19 and references therein). The only two MCDS mutations not found in the NC1 domain affect the putative signal peptide cleavage site upstream of the non-collagenous domain 2 in the molecule (20). The role of the carboxy-terminal non-collagenous domain of most collagens (including type X) in initiating intracellular α-chain selection, assembly, and helix formation is well established (Refs. 8 and 21 and references therein). Assembly studies based on cell-free translation of recombinant RNA encoding wild type and MCDS transcripts of type X collagen suggest that mutant chains do not interfere with the trimerization of the wild type protein based on SDS-PAGE assays (22). These findings have led to the suggestion that the phenotype of MCDS is best explained by haplo-insufficiency. In support of this hypothesis, in one individual with MCDS, only wild type mRNA for type X collagen could be detected in the growth plate, suggesting that the mutant transcript is unstable and rapidly degraded (23). However, haplo-insufficiency is in discord with the clustering of both point and frameshift mutations in the NC1 domain of type X collagen; in haplo-insufficiency one would expect to find MCDS-causing frameshift mutations to be randomly distributed through the gene (19). The non-random clustering of

* This work was supported by Grants B0547 and G0520 from the Arthritis Endochondral Research Campaign and Grant 019512 from The Wellcome Trust. 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.

‡ These authors contributed equally to the data presented in this paper.

§ To whom correspondence should be addressed: University of Manchester, 2.205 Stopford Bldg., Oxford Rd., Manchester, M13 9PT United Kingdom. Tel.: 44 161 275 5097; Fax: 44 161 275 5082; E-mail: ray.boot-handford@man.ac.uk.

The abbreviations used are: NC1, non-collagenous carboxyl-terminal domain of type X collagen; MCDS, metaphyseal chondrodysplasia type Schmid; PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; TNF, tumor necrosis factor; nt, nucleotide(s).
MCDS mutations in the NC1 domain are more consistent with a mechanism involving dominant interference in which the mutant chains retain the ability to trimerize. It is possible that previous investigations have failed to detect trimerization of chains containing MCDS mutations due to the harsh assay conditions, such as SDS-PAGE, that have been employed.

The crystal structure of the NC1-like domain of ACRP30, a protein closely related to type X collagen, has been recently solved at 2.1-A resolution (24). We have therefore performed molecular modeling of the NC1 domain, based on the ACRP30 crystal structure coordinates, to gain greater insight into the structural basis of MCDS. The model reveals that the MCDS mutations are localized to two specific regions of the folded monomeric NC1 domain and that many of the MCDS mutations may not totally abolish the ability of affected chains to trimerize. Mutagenesis experiments, based on information obtained from the model, demonstrate that NC1 domains containing specific MCDS mutations retain the ability to form trimers and provide an explanation for the unusually high thermal stability exhibited by the assembled NC1 trimer.

EXPERIMENTAL PROCEDURES

Alignments—The amino acid sequence of the NC1 domain of human type X collagen was used to probe EMBL, GenBank, and PIR as data base using the advanced BLAST 2.0 search (25). We used a cut-off expectancy value (E) of 1 (such that no more than 1 match is expected to be found merely by chance, according to the stochastic model of Karlin and Altschul (26)) and a gapped alignment using the Blosum 62 matrix (27). Multalin (28) was used to re-align the resulting 25 sequences using the Blosum 62 matrix using a penalty of 12 for gap opening and 2 for gap extending. None of the resulting aligned sequences have three-dimensional structures in the Brookhaven data base.

Molecular Modeling—All molecular modeling was performed on an R5000 02 Silicon graphics workstation using QUANTA® and CHARMM 23.1B programs. The three-dimensional model of the NC1 domain of type X collagen was built based on the coordinates of human ACRP30 (29). The file containing the coordinates was rewritten manually to make a Protein Data Base file format that could be imported by QUANTA®. A homology model was built by copying the coordinates of the backbone of the ACRP30 trimer and the coordinates of identical residues in the type X collagen NC1 domain. The remaining side chains were built in the Protein Design module using the Ponder and Richards' rotamer library. Disordered loops in the crystal structure of ACRP30 (loop G-H in M1 and loops G-H and A-A' in M2) were modeled wherever possible by overlaying the intact loop of a corresponding part of one of the other two monomers. This left a gap between Glu590 to Glu661 (corresponding to Asp224 to Gly 228) in ACRP30 in each monomer.

The resulting trimer of the NC1 domain of type X collagen was energy-minimized (using the steepest descent followed by Newton-Raphson algorithm) to gradient convergence (<0.01 root mean square) removing bad steric and electrostatic contacts. The water molecules from the crystal structure were overlaid, and those internal to the trimer were incorporated into the model. The Protein Health module in QUANTA was used to check the integrity of the model using a Ramachandran map and to identify buried hydrophilic or exposed hydrophobic residues.

The solvent-accessible surfaces of monomers and assembled trimers were calculated to estimate the surface area buried upon trimerization. The local environment of each residue on the interfaces was examined to locate those residues important for the trimer structure. The interfaces are not symmetric, because of a slight twist and stagger to the arrangement of the β-sandwich monomers (24). Particular attention was given to the position of those residues in the NC1 model trimer that differed in ACRP30 and related proteins (see Fig. 1).

Preparation of the Constructs—DNA constructs were generated by PCR from the lambda genomic COL10A1 clone, HX3 (30). DNA encoding the entire human type X collagen NC1 domain was amplified using the primers 5'-CCAGCTCAACGACATATGCTGAGGT3' (primer A, sense nucleotides 1550–1563), which incorporates an Ndel site coding for an in-frame Met codon at the amino terminus of the extended NC1 coding region, was used together with the antisense primer B. To produce the truncated NC1 translation construct, DNA encoding a portion of the NC1 domain from the Met codon was generated by PCR using the primer 5'-GGGGTAACATATGCTGCTGT3' (nucleotides 1632–1656) together with the antisense primer B. To create constructs containing MCDS and related point mutations, site-directed mutagenesis was carried out using the PCR-based single overlap extension procedure described elsewhere (29) using oligonucleotides A and B (described above) and mutagenic oligonucleotides (nucleotides 1575–1899) sense, 5'-CCAGGAATTACTCTTTTCATAC-3' and antisense, 5'-GTATGAAAAAGTATATTCCTGG-3' (for Y598F; sense, 5'-CCAGGAATTACTCTTTTCATAC-3' and antisense, 5'-GTATGAAAAAGTATATTCCTGG-3' for Y598A; 5'-CCAGGAATTACTCTTTTCATAC-3' and antisense, 5'-GTATatatatatatatatatatatat-3' and antisense, 5'-GTATGAAAAATGTATATTTCTTG-G-3' for Y598D; and nucleotides 1881–2004) sense, 5'-ATATACCGTTTCCGACCTGCGTGTT-3' and antisense, 5'-CACCTGTTGAAATGAACCTGTT-3' for Y598S). The resulting PCR fragments were cloned into the T/A cloning vector pCR 2.1 (Invitrogen). Constructs generating sense RNA upon transcription from the T7 promoter were identified by restriction mapping and sequenced.

Coupled In Vitro Transcription and Translation—In vitro expression was carried out using the TNT T7 polymerase-coupled rabbit reticulocyte lysate system (Promega). 1 μg of each purified construct was expressed in 40 μl of T7 premix containing 2 μl of [35S]methionine (1000 Ci/mmol, NEN Life Science Products) made up to a final volume of 50 μl with sterile distilled water. Reactions were incubated at 30 °C for 90 min. 5-μl aliquots of in vitro expression reaction were added to 10 μl of sample buffer (50 mM Tris/HCl, pH 6.8 (room temperature), 50% (v/v) glycerol, 0.025% (w/v) bromphenol blue; the concentration of SDS in the sample buffer was either 10% (w/v) or 0.75% (w/v) as indicated under "Results"). Where heating is indicated, samples were overlaid with 20 μl of mineral oil and incubated at the appropriate temperature for 5 min using Perkin-Elmer 480 thermal cycler prior to analysis by SDS-PAGE. All samples were resolved on standard 12% or 16% PAGE gels containing 0.1% SDS according to Laemmli (30). Gels were fixed for 1 h in 10% (w/v) acetic acid containing 10% (w/v) methanol for 30 min and washed twice in 30% (w/v) methanol containing 3% (v/v) glycerol. Gels were dried and exposed to Kodak Biomax Film, or, if quantitation was required, phosphorimaged (Fuji-Bas).

Immunoprecipitation of NC1 Domains with the X34 Monoclonal Antibody—In vitro translations containing labeled NC1 domains were subjected to centrifugation at 109,000 x g for 1 h to remove particulate 50 μl of the supernatant was then dialyzed against NET buffer overnight (100 mM Tris/HCl, pH 7.4, 100 mM NaCl, 1% (v/v) Tween 20, 1 mM EDTA). 5 μl of hybridoma medium containing the X34 monoclonal antibody, which recognizes a conformation-dependent epitope in the NC1 domain of native type X collagen (31), was added to the dialyzed translation followed by incubation for 16 h at 4 °C with gentle inversion. Following the incubation, 10 μl of protein A-Sepharose (Amersham Pharmacia Biotech) was added to the mixture and the incubation was allowed to proceed for a further 4 h, after which the protein A-Sepharose was recovered by centrifugation at 12,000 × g (control incubations were carried out in the absence of X34). The pelleted protein A-Sepharose was washed in 500 μl of NET buffer three times followed by resuspension in sample buffer containing 1% (v/v) β-mercaptoethanol. Prior to SDS-PAGE analysis, the samples were heated for 2 min at 68 °C.

RESULTS

Alignments—The complete NC1 domain of human type X collagen was aligned with the carboxyl-terminal non-collagenous domain of mouse ACRP30 (Fig. 2). The type X collagen NC1 domain contains an extra 27 residues at the amino terminus that are not present in ACRP30 and a deletion of ACRP30-302. The globular domain of ACRP30 exhibits 40% identity and 65% similarity with the equivalent region of the type X collagen NC1 domain. Overlaying the secondary structural elements from the crystal structure of ACRP30 on the alignment
FIG. 1. Sequence alignment of C1q-like domains with a BLAST 2.0 E value of less than 0.015. BLAST 2.0 was used to probe the data bases with the NC1 domain of type X collagen sequence and retrieved 58 sequences with an E value less than 10. After discarding redundancy we aligned 25 sequences in Multalin using the Blosum 62 matrix and default penalty values. Those residues in red are conserved throughout the alignment; those in blue are conserved in at least 70% of sequences, and those highlighted in yellow show the MCDS point mutations.
revealed that those residues forming β-strands of ACRP30 were mostly conserved in the NC1 domain, whereas differences were predominantly localized to the loop regions (Fig. 2). Therefore, the secondary structural elements are highly conserved. Within the β-strands, those residues forming the hydrophobic core of the monomer of ACRP30 (see below) are particularly conserved. This conservation strongly indicates that the tertiary fold of the NC1 domain is identical to the β-sandwich of ACRP30. In addition, the conservation of key interface residues between monomers (see below and Fig. 2)

**Fig. 1—continued**

**Fig. 2.** Alignment of mouse ACRP30 (Swiss-Prot accession number Q60994, acr3 mouse) and human type X collagen (Swiss-Prot accession number Q03692, cala human) from the end of the triple helical domains to their carboxyl termini. Conserved residues are shown in white font with black shading; * indicates the position of amino acid substitutions causing MCDS; β-strands are indicated by shaded boxes and labeled according to ACRP30 and TNF-α nomenclature (24). The alignment was calculated in Multalin (28) with default parameters then adjusted by hand in the triple helical region.
indicates that the quaternary structures of these two proteins are similar. This pattern of conservation, based on alignments, also holds across the complete family of related proteins including collagen types X and VIII, C1q, ACRP30, hibernation protein and cerebellin (see Fig. 1).

**Molecular Model of Trimeric NC1 Domain**—A model of the trimeric NC1 domain (Fig. 3) was built using the three-dimensional coordinates of trimeric ACRP30 based on the alignment shown in Fig. 2. The amino-terminal 27 residues of the NC1 domain of type X collagen were not included in the model since...
there is no equivalent sequence in ACRP30 (Fig. 2). The 8 loop regions in the NC1 domain were modeled on the ACRP30 structure since their lengths are conserved, and they are anchored spatially by the β-strands.

Upon trimerization (Fig. 3a), the mode of packing between the NC1 monomer subunits produces a tight association contributed to by the loss of solvent-accessible surface and hence providing a free energy gain. The total buried solvent-accessible surface of the NC1 domain on trimerization is 6222 Å², compared with 5324 Å² for ACRP30. The three monomers pack together forming a twisted tunnel that extends through the core of the structure. The hydrophobic contacts that close the packing between the NC1 monomers are identical or show only conservative changes with respect to ACRP30. These include Ala595, Ile641, Leu673, Ile696, Val651, Ile641, Phe675, and Ala678 in the NC1 domain. However, in the NC1 domain, Ala595, Ile641, Val651, and Ala678 are replaced by Gly595, Leu598, Val651, and Ala678, respectively.

TABLE I

| Point mutation | Position of residue and direction of side chain | Comments on wild type residue orientation |
|----------------|-----------------------------------------------|-------------------------------------------|
| Y597H          | On β-strand C, interior face of trimer, side chain pointing to interior of monomer | OH group hydrogen bonds to Ser602 via water |
| Y598D          | On β-strand C, interior face of trimer, side chain pointing into trimer tunnel | OH group forms pyramidal geometry with 2 other Tyr608 and a water molecule (see Fig. 3) |
| S600P          | On β-strand C, interior face of trimer, side chain pointing into trimer tunnel | OH groups from 3 monomers form a network of H bonds both with water and directly connecting the three monomers via Ser571 and Ser600. Proline could break the β-strand |
| S671P          | On terminal β-strand H, on interior face of trimer with side chain pointing into monomer interface | One monomer forms a direct contact with neighboring monomer Ser600. Ser600 from M2 and M3 form a network of H bonds with water and both the side chain and backbone of neighboring Ser600. Proline could break the strand |
| C591R          | On β-strand B on outside face of trimer with sulfydryl pointing to the interior of the monomer | Arginine could not be easily accommodated in the hydrophobic interior |
| G595E          | On bottom surface of trimer at the start of β-strand C | This glycine forms a sharp turn which would not be favorable for glutamic acid |
| L614P          | On β-strand D, exterior face of trimer, side chain contributing to hydrophobic core of monomer | This strand immediately precedes loop DE (see Fig. 1). Proline could disrupt the β-strand conformation and dislocate the loop |
| N617K          | On loop DE, exterior face of trimer, side chain points out into solution | This loop maps onto corresponding loops in C1q and TNF-α which are involved in molecular recognition |
| G618V          | On loop DE, exterior face of trimer, this glycine forms the bend in the loop | This glycine has a ϕ torsion angle of 170° which would be unfavorable for valine and proline |
| L644R          | Toward end of β-strand F, bottom face of trimer, side chain completely buried | Contributing to hydrophobic core of monomer. Tightly packed region so arginine would cause steric clash as well as unfavorable electrostatic contacts |
| D648G          | At the beginning of β-strand G, bottom face of trimer side chain points towards Aα567 and Lys616 | Possible salt bridge between COO− and NH₃⁺ of Lys616 |
| W651R          | On β-strand G, on outside face of trimer, side chain partially exposed and points towards loop II | Aromatic ring almost all exposed; could be involved in hydrophobic contact with other molecule or another trimer |

* Indicates residues located in a cluster on an interior patch.  
* Indicates residues located on a surface patch on the exterior of trimer.

Location of MCDS Point Mutations in the NC1 Trimer—All NC1 domain residues that are affected by single amino acid substitutions causing MCDS were displayed simultaneously on the monomer (Fig. 3d) and trimer (Fig. 3e) models revealing a striking clustering of position. The substituted residues are localized to two distinct regions. The side chains of residues that are substituted in MCDS either form the polar and hydrophobic-bonded region within the trimer tunnel (e.g. Tyr608 and Ser609, see Fig. 3c) or form a patch on the external surface of the monomer/trimer assembled from part of loop D–E and part of β-strand G (Figs. 1 and 2). A detailed summary of the location and orientation of amino acid residues substituted within the NC1 domain in MCDS is presented in Table I.

Trimerization and Stability of Recombinant Human Type X Collagen NC1 Domain—The model was first used in an attempt to predict motifs or residues responsible for the unusual stability of the NC1 domain of type X collagen (see below and “Discussion”). Expression of the full-length NC1 domain in a reticulocyte lysate system resulted in detection of the monomeric protein (Mᵣ = 20,000) and a complex of Mᵣ = 45,000 that is stable on SDS-PAGE in reducing conditions (Fig. 4a, lane 4). Similarly, expression of an “extended” NC1 construct containing 11 Gly-X-Y repeats of the collagenuous domain of type X collagen (Fig. 4a, lane 9) and a “truncated” construct lacking the amino-terminal 27 residues of the NC1 domain (Fig. 4a, lane 1) produced monomers of Mᵣ = 28,000 (extended) and 17,000 (truncated) as well as complexes of Mᵣ = 60,000 and 35,000 (respectively). Based on Mᵣ, the complexes formed appeared too large to be dimers but too small to be trimers. In order to resolve the molecular composition of the complexes, co-expression studies were performed. Co-expression of the truncated and extended constructs (Fig. 4a, lane 2) and full-length and extended constructs (Fig. 4a, lane 5) produced the expected NC1 monomer and complex bands as well as two extra complex bands of intermediate Mᵣ. The fact that 4 complex
bands were formed in each co-expression reaction indicates clearly that the complexes are trimeric associations of the monomers running slightly faster on SDS-PAGE than would be predicted. It should be noted that the data in Fig. 4a were produced adding the translation mixes to sample buffer containing 10% SDS (final SDS concentration 6.7%) and running on the SDS-PAGE gel without prior heating.

To check that the trimeric NC1 domains had adopted a native conformation, translation reactions containing co-expressed full-length and extended NC1 domains were immunoprecipitated with the X34 monoclonal antibody which recognizes a conformation-dependent epitope in the native trimeric NC1 domain but not monomeric NC1 (31). SDS-PAGE analysis of the immunoprecipitated co-translation reactions revealed that the X34 monoclonal antibody recognized the trimer but not monomeric NC1 domains present in the co-translation. The control immunoprecipitation carried out in the absence of X34 revealed that low levels of monomeric and trimeric NC1 domain bound nonspecifically to the protein A-Sepharose (Fig. 4b, lane 3). Comparison of the control immunoprecipitation with that carried out in the presence of the X34 antibody demonstrated that in addition to adopting a native structure, all of the trimeric NC1 domain was stable in the conditions used for SDS-PAGE analyses.

In order to investigate the stability of the trimers formed in the system, translations of the full-length NC1 domain (Fig. 4c) were heated for 5 min in 6.7% SDS at the temperatures indicated prior to loading on the gel. The trimer formed is stable at 60 °C and, while some denatures to monomer at 80 °C, a significant proportion is still trimeric at 100 °C (see also Fig. 5a). The full-length monomer appears unaffected by the exposure to increasing temperatures (Fig. 4d). ACRP30 lacks the amino-terminal 27 residues found in the NC1 domain of type X collagen (Fig. 2) and also does not form a trimer capable of withstanding SDS-PAGE (13). In order to see whether the amino-terminal NC1 sequence plays a role in determining stability, it was deleted. In comparison with the full-length NC1 domain (Fig. 4c), the domain lacking the amino-terminal 27 residues responded in a different fashion in two critical ways. First, as the temperature increases above 40 °C, the monomer is seen to aggregate into a high M, complex that does not enter the resolving gel (Fig. 4d). Second, the trimer, which is stable at 60 °C and unaffected by the aggregation process exhibited by the monomer, is completely denatured at 80 °C (Fig. 4d).

**Assembly and Stability of NC1 Domains Containing MCDS and Related Mutations**—In order to examine the role in determining trimer stability of Tyr^608^, which forms a network of hydrogen bonds in the trimer tunnel (Fig. 3e) and is mutated to Asp in one case of MCDS (Table I), the stabilities of NC1 domains containing Y598F, Y598A, and Y598D were examined (Fig. 5). Removal of the OH group (Y598F) produced an approximately 10 °C drop in the melting temperature of the trimer, and a further 10 °C decrease was found for the trimer containing the Y598A mutation. Under these conditions, in which the sample is placed in 6.7% SDS to prevent “geling” of the translation mix during heating prior to loading on the gel, a trimer

**FIG. 4.** Recombinant NC1 domain of human type X collagen forms a stable SDS-PAGE-resistant trimeric complex in vitro. DNA constructs encoding either the full-length NC1 domain (F), or an extended NC1 domain possessing the first 11 Gly-X-Y repeats at the amino terminus (E), or a NC1 domain truncated by 27 residues from the amino terminus (T) were transcribed and translated in the presence of [35S]methionine by an in vitro expression system. a, autoradiographs of translated samples exposed to 6.7% SDS and then resolved by standard SDS-PAGE analysis under reducing conditions without prior heating. Lane 1, expression of truncated (T) NC1 construct; lane 2, co-expression of truncated and extended (E) NC1 constructs; lane 3, expression of the extended NC1 construct; lane 4, expression of the full-length (F) NC1 construct; lane 5, co-expression of full-length and extended NC1 constructs. b, 12% SDS-PAGE analysis of immunoprecipitations of translations containing normal length and extended NC1 domains. Lane 1, whole co-translation reaction; lane 2, material immunoprecipitated by monoclonal antibody X34; lane 3, negative control, material bound nonspecifically by the protein A-Sepharose only. c and d, full-length (F in c) and truncated (T in d) NC1 constructs were analyzed exactly as described in a except that the samples were heated at the temperatures indicated above the gels for 5 min prior to loading. Top indicates the boundary between the stacking and resolving gels.
could not be detected when expressing the construct containing the MCDS mutation Y598D (see Fig. 5a). Indeed, trimerization could not be detected for a second MCDS mutation, S600P, under the same conditions (Fig. 5b).

To investigate whether the high level of SDS (6.7% final concentration) to which samples were exposed prior to gel
electrophoresis was denaturing potential trimer complexes formed by MCDS constructs, the concentrations of SDS in the sample buffer were reduced. Translations separated under native conditions failed to resolve in the gel, as did samples taken up in a final concentration of 0.1% SDS and run on standard SDS-PAGE gels (data not shown). However, when samples were exposed to 0.5% SDS and (without heating) run on standard 0.1% SDS-PAGE gels, resolution of wild type monomers and trimers was apparent (Fig. 5c, lanes 1 and 2). Analysis under the same conditions of NC1 domains containing the Y598D and S600P MCDS substitutions (Fig. 5c, lanes 3 and 5) demonstrated that each formed a trimeric complex. When the same MCDS translations were exposed to 1% SDS prior to running on the gel, the levels of detectable trimer were reduced by more than 50% (results not shown). Co-translation of each MCDS NC1 mutant construct with the extended wild type construct led to the formation of heterotrimers (Fig. 5c, lanes 4 and 6).

DISCUSSION

Structure and Assembly of the Wild Type Human NC1 Domain—The NC1 domain of type X collagen is a putative asymmetric trimer composed of three 10 β-strand jelly roll motifs based on modeling of the ACRP30 crystal structure, which is remarkably similar to the structure of the TNFs (24). In both TNF and C1q, the sequences involved in molecular recognition (e.g. receptor and IgG binding, respectively) have been shown to include loop D–E and loop A–A (see Ref. 24 and references therein). These sequences map to two specific exposed patches on each monomer within the assembled trimer. Interestingly, the residues forming these loops differ in composition and in length across the whole family of type X collagen-like proteins as shown in Figs. 1 and 2. These data suggest strongly that the equivalent regions in type X collagen are involved in molecular recognition events such as supramolecular assembly (6) or interaction with cells.

Previous studies on type X collagen assembly have demonstrated that trimers generated by in vitro translation possess an unusually high thermal stability (Ref. 22 and references therein). We demonstrate here that this thermal stability is a consequence of the stability of the NC1 domain (see Fig. 4). It is of interest to note that the intact wild type NC1 domain trimerizes spontaneously forming a native structure (Fig. 4, a and b), whereas when translating the full-length type X collagen polypeptide in the same system, either Ca2+ or microsomal membranes are required to achieve trimerization (22). We assume that the presence of Ca2+ or microsomal membranes (and associated chaperone proteins), prevents the unfolded collagenous domain from interfering with the folding and trimerization of the nascent NC1 domain. Furthermore, it is noteworthy that recombinant human NC1 domain expressed in bacteria also spontaneously forms a thermally stable native trimere that is specifically recognized by the conformation-dependent X34 monoclonal antibody (32). ACRP30 and other members of this family of related proteins, with the exception of type VIII collagen (34), do not appear to exhibit this level of thermal stability. The modeling of the NC1 domain has revealed a number of features that may help explain the differences in stability. First, the model revealed that the packing of the aromatic side chains within the hydrophobic core of the monomers of both NC1 and ACRP30 is organized in a “herring bone” pattern similar to that associated with hydrophobic cores of other stable proteins (34). However, the NC1 domain contains a more densely packed hydrophobic core than ACRP30 because of the presence of three additional hydrophobic side chains that may increase stability of the monomer and trimer (35). Second, on trimerization, the total buried solvent-accessible surface of the NC1 domain is 17% greater than ACRP30, and the hydrophobic plug at the base of the NC1 domain has additional contributions from Pro150, Val151, and Ile688 (corresponding residues in ACRP30 are Tyr, Arg, and Gly, respectively). Third, sequence alignments show that type X collagen has an additional 27 residues at the amino terminus of the NC1 domain that are not present in ACRP30 or other members of the family of proteins with the exception of the two chains of type VIII collagen (Figs. 1 and 2).

To examine whether the additional 27-residue sequence stabilizes the NC1 domain, we compared full-length and truncated NC1 domains. In the absence of the 27 residues, monomeric NC1 forms high molecular weight aggregates at temperatures above 40 °C (Fig. 4d) implying that the monomer has a defined structure in SDS at temperatures below 40 °C that becomes denatured at higher temperatures leading to aggregation. However, no aggregation is exhibited by the full-length NC1 monomer, even at 100 °C (Fig. 4c), strongly suggesting that it has a folded structure that is resistant to denaturation. The trimer is still formed by the shortened NC1 protein although its stability is reduced (compare Fig. 4, c and d) presumably because of the decreased stability of the monomer.

Structure and Assembly of NC1 Domains Containing MCDS and Related Mutations—Mapping MCDS point mutations onto the NC1 model revealed a striking three-dimensional clustering of affected residues into two distinct domains (Fig. 3, d and e). The first cluster of residues substituted in MCDS is located on an external patch that includes loop D–E and is known to play a role in molecular recognition in structurally related molecules such as TNF and C1q as described above. Since these residues are located on the external surface of the trimer, it seems unlikely that the associated MCDS mutations would prevent the trimerization process. The second cluster of mutations affect the polar region of the trimer tunnel. It is significant that not one of the MCDS point mutations (see Table I) affect the hydrophobic plug at the base of the trimer tunnel (Fig. 3b) or any of the monomer–monomer hydrophobic contacts that are likely to drive the trimerization process.

On the basis of the findings described above, we predicted that NC1 domains containing MCDS point mutations would trimerize but form less stable complexes that may be difficult to detect on SDS-PAGE gels. In order to test this hypothesis, the role of Tyr598 in determining NC1 trimer stability was examined. Tyr598 not only participates in a hydrogen bond network as shown in Fig. 3c but is also mutated to Asp in MCDS (Table I). Replacement of the Tyr598 by either Phe or Ala did not prevent the formation of SDS-resistant trimers although their thermal stabilities were considerably reduced (Fig. 5a). Under equivalent conditions, in which the translation mix was dissolved in sample buffer containing 6.7% SDS, no trimers could be detected when expressing NC1 domains containing MCDS Y598D or S600P mutations (Fig. 5b). However, both Y598D and S600P mutant NC1 domains were shown to form homotrimers, and heterotrimers with wild type NC1 domains, that are detectable in low concentrations of SDS (Fig. 5c) but denature as the level of SDS is increased (see Fig. 5b). These data clearly demonstrate that the stabilities of trimers containing MCDS mutant chains (in terms of sensitivity to the concentration of SDS and temperature) are considerably reduced compared with that of the wild type. Nevertheless, the fact that non-covalently associated MCDS mutant trimers can be detected on SDS-PAGE indicates that the stabilities of these mutant complexes are still greater than equivalent non-covalent trimeric complexes formed during the assembly of collagens with the exception of type VIII (33). We also have preliminary evidence suggesting that NC1 domains containing frameshift mutations
causing MCDS also retain the ability to trimerize. The MCDS mutant NC1 domains examined here not only retain the ability to trimerize but several α1(X) chains containing similar MCDS substitutions have recently been shown to fold stable collagenous triple helices.

In conclusion, the study presented here has produced a three-dimensional model of the NC1 domain of type X collagen that reveals a clustering of MCDS mutations that is consistent with the mutant chains retaining the ability to participate in the trimerization process. In addition, we have demonstrated that the NC1 domain forms an unusually stable trimer and that NC1 domains containing MCDS mutations retain the ability to trimerize. The modeling and protein assembly studies suggest strongly that the predominant molecular mechanism causing MCDS involves a dominant interference of mutant chains on wild type assembly.

Acknowledgments—We thank Larry Shapiro for the ACRP30 coordinates, Klaus Von der Mark for providing the X34 monoclonal antibody, and Mike Grant and Neil Bulleid for constructive discussions and suggestions.

REFERENCES
1. Schmid, T. M., and Linsenmayer, T. F. (1985) J. Cell Biol. 100, 598–605
2. Grant, W. T., Wang, G. J., and Balian, G. (1987) J. Biol. Chem. 262, 9844–9849
3. Girkontaite, I., Frischholz, S., Lammi, P., Wagner, K., Swoboda, B., Aigner, T., and Von der Mark, K. (1996) Matrix Biol. 15, 231–238
4. Thomas, J. T., Creswell, C. J., Rash, B., Heylard, J., Freemont, A. J., Grant, M. E., and Boot-Handford, R. P. (1991) Biochem. Soc. Trans. 19, 804–808
5. Thomas, J. T., Creswell, C. J., Rash, B., Nicolai, H., Jones, T., Solomon, E., Grant, M. E., and Boot-Handford, R. P. (1991) Biochem. J. 280, 617–623
6. Wada, C., and Ohitani, H. (1991) Mol. Brain Res. 9, 71–77
7. Orade, Y., Oberdick, J., Molinar-Rode, R., and Morgan, J. J. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 1069–1073
8. Scherer, P. E., Williams, S., Fogliano, M., Baldini, G., and Lodish, H. F. (1995) J. Biol. Chem. 270, 26746–26749
9. Jacenko, O., LuValle, P. A., and Olsen, B. R. (1993) Nature 365, 56–61
10. Rosati, R., Horan, G. S. B., Pinero, G. J., Gorofalo, S., Keene, D. R., Horton, W. A., De Crombrugghe, B., and Behringer, R. R. (1994) Nat. Genet. 6, 129–135
11. Wada, C., and Ohitani, H. (1991) Mol. Brain Res. 9, 71–77
12. Urade, Y., Oberdick, J., Molinar-Rode, R., and Morgan, J. J. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 1069–1073
13. Scherer, P. E., Williams, S., Fogliano, M., Baldini, G., and Lodish, H. F. (1995) J. Biol. Chem. 270, 26746–26749
14. Jacenko, O., LuValle, P. A., and Olsen, B. R. (1993) Nature 365, 56–61
15. Rosati, R., Horan, G. S. B., Pinero, G. J., Gorofalo, S., Keene, D. R., Horton, W. A., De Crombrugghe, B., and Behringer, R. R. (1994) Nat. Genet. 6, 129–135
16. Wada, C., and Ohitani, H. (1991) Mol. Brain Res. 9, 71–77
17. Urade, Y., Oberdick, J., Molinar-Rode, R., and Morgan, J. J. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 1069–1073
18. Wallis, G. A., Rash, B., Sweetman, W. A., Thomas, J. T., Super, M., Evans, G., Grant, M. E., and Boot-Handford, R. P. (1994) An. J. Hum. Genet. 54, 169–178
19. Wallis, G. A., Rash, B., Sykes, B., Bonaventure, J., Maroteaux, P., Zabel, B., Wynnedavies, R., Grant, M. E., and Boot-Handford, R. P. (1996) J. Med. Genet. 33, 450–457
20. Ikegawara, S., Nakamura, K., Nagano, A., Haga, N., and Nakamura, Y. (1997) Hum. Mutat. 9, 131–135
21. Bulleid, N. J., Dalley, J. A., and Lees, J. F. (1997) EMBO J. 16, 684–6701
22. Chan, D., Weng, Y. M., Hocking, A. M., Golub, S., Mc Quillan, D. J., and Bateman, J. F. (1996) J. Biol. Chem. 271, 15360–15372
23. Chan, D., Weng, Y. M., Graham, H. K., Silience, D. O., and Bateman, J. F. (1998) J. Clin. Invest. 101, 1490–1499
24. Shapiro, L., and Scherer, P. E. (1998) Curr. Biol. 8, 335–338
25. Pearson, W. R. (1990) Methods Enzymol. 183, 63–88
26. Karlin, S., and Altschul, S. F. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 2264–2268
27. Henikoff, S., and Henikoff, J. G. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 10915–10919
28. Corpet, F. (1988) Nucleic Acids Res. 16, 10881–10890
29. Ling, M. M., and Robinson, B. H. (1997) Anal. Biochem. 254, 157–178
30. Laemmli, U. K. (1970) Nature 227, 680–685
31. Girkontaite, I., Frischholz, S., Lammi, P., Wagner, K., Swoboda, B., Aigner, T., and Von der Mark, K. (1996) Matrix Biol. 15, 231–238
32. Gregory, C. A. (1998) Structural and Functional Studies on Reombinant Human Non-Collagenous Carboxy-terminal (NC1) Domain of Human Type X Collagen. Ph.D. thesis, University of Manchester, UK
33. Ilidge, C., Kiely, C., and Shuttleworth, A. (1998) J. Biol. Chem. 273, 22091–22095
34. Burley, S. K., and Petsko, G. A. (1986) Science 229, 23–28
35. Malakauskas, S. M., and Mayo, S. L. (1998) Nat. Struct. Biol. 5, 470–475

D. S. Marks, C. A. Gregory, G. A. Wallis, A. Brass, K. E. Kadler, and R. P. Boot-Handford, unpublished observations.

S. McLaughlin and N. Bulleid, personal communication from School of Biological Sciences, University of Manchester, United Kingdom.