

A Role for Collagen IV Cross-links in Conferring Immune Privilege to the Goodpasture Autoantigen

STRUCTURAL BASIS FOR THE CRYPTICITY OF B CELL EPITOPES^{*[5]}

Received for publication, May 6, 2008 · Published, JBC Papers in Press, May 22, 2008, DOI 10.1074/jbc.M803451200

Roberto M. Vanacore^{†§}, Amy-Joan L. Ham^{¶||}, Jean-Philippe Cartailier^{†§}, Munirathinam Sundaramoorthy^{†§¶1}, Parvin Todd^{†§}, Vadim Pedchenko^{†§}, Yoshikazu Sado^{**}, Dorin-Bogdan Borza^{†§¶¶}, and Billy G. Hudson^{†§¶¶2}

From the [†]Center for Matrix Biology, Departments of [§]Medicine, [¶]Biochemistry, and ^{**}Pathology, ^{||}Mass Spectrometry Research Center, Vanderbilt University Medical Center, Nashville, Tennessee 37232 and ^{**}Division of Immunology, Shigei Medical Research Institute, Okayama 701, Japan

The detailed structural basis for the cryptic nature (crypticity) of a B cell epitope harbored by an autoantigen is unknown. Because the immune system may be ignorant of the existence of such “cryptic” epitopes, their exposure could be an important feature in autoimmunity. Here we investigated the structural basis for the crypticity of the epitopes of the Goodpasture autoantigen, the $\alpha 3\alpha 4\alpha 5$ noncollagenous-1 (NC1) hexamer, a globular domain that connects two triple-helical molecules of the $\alpha 3\alpha 4\alpha 5$ collagen IV network. The NC1 hexamer occurs in two isoforms as follows: the M-isoform composed of monomer subunits in which the epitopes are accessible to autoantibodies, and the D-isoform composed of both monomer and dimer subunits in which the epitopes are cryptic. The D-isoform was characterized with respect to quaternary structure, as revealed by mass spectrometry of dimer subunits, homology modeling, and molecular dynamics simulation. The results revealed that the D-isoform contains two kinds of cross-links as follows: S-hydroxylysyl-methionine and S-lysyl-methionine cross-links, which stabilize the $\alpha 3\alpha 5$ -heterodimers and $\alpha 4\alpha 4$ -homodimers, respectively. Construction and analysis of a three-dimensional model of the D-isoform of the $\alpha 3\alpha 4\alpha 5$ NC1 hexamer revealed that crypticity is a consequence of the following: (a) sequestration of key residues between neighboring subunits that are stabilized by domain-swapping interactions, and (b) by cross-linking of subunits at the trimer-trimer interface, which stabilizes the structural integrity of the NC1 hexamer and protects against binding of autoantibodies. The sequestered epitopes and cross-linked subunits represent a novel structural mechanism for conferring immune privilege at the level of quaternary structure. Perturbation of the quaternary structure may be a key factor in the etiology of Goodpasture disease.

Autoimmune disease results from a failure of the adaptive immune system to differentiate between self- and nonself-antigens. During the early steps of development, the immune system learns to recognize self by inducing “tolerance” against self-antigens. Several mechanisms, including clonal deletion, clonal anergy, and receptor editing, have been described for inducing self-tolerance of potential autoreactive lymphocytes (1). However, certain T cell self-epitopes, termed cryptic epitopes, are unable to induce tolerance during lymphocyte development. Although the mechanisms underpinning this remain elusive, under certain pathophysiological conditions such as inflammation, cryptic T cell self-epitopes may become apparent to the immune system and trigger a response that may lead to autoimmunity (2–4).

There is emerging evidence for the existence of cryptic B cell epitopes that may also play an important role in autoimmunity. In Goodpasture disease (GP),³ the GP autoantibodies target the glomerular and alveolar basement membranes, specialized forms of extracellular matrix, causing rapidly progressive glomerulonephritis and pulmonary hemorrhage, respectively (5, 6). The GP epitopes are cryptic within the native autoantigen, as first described by Wieslander *et al.* (7). The autoantigen is now known as the $\alpha 3\alpha 4\alpha 5$ noncollagenous-1 (NC1) hexamer, which connects two triple-helical protomers within the $\alpha 3\alpha 4\alpha 5$ network of collagen IV of certain basement membranes (6). The GP epitopes, designated E_A and E_B, are cryptic within the hexamer and reside in the $\alpha 3$ NC1 subunit (8).

The crypticity of GP epitopes is a feature of the quaternary structure of the $\alpha 3\alpha 4\alpha 5$ NC1 hexamer, which occurs in M- and D-isoforms. The M-isoform is composed of only NC1 monomer subunits. The epitopes are accessible to the GP autoantibodies that induce NC1 hexamer dissociation into monomer subunits (9). In contrast, the D-isoform is composed of both NC1 monomer and dimer subunits, but the epitopes are inaccessible (cryptic), indicating a role of dimers in the mechanism of crypticity (9). The epitopes can be exposed, however, by experimental dissociative conditions, such as low pH or protein

* This work was supported, in whole or in part, by National Institutes of Health Grants DK18381, DK065123 (to B. G. H.), DK62524 (to M. S.), and AI69874 (to D. B. B.). 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.

[5] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Table 1 and Fig. 9.

¹ To whom queries about NC1 hexamer modeling should be addressed.

² To whom correspondence should be addressed: Dept. of Medicine, Division of Nephrology, Vanderbilt University Medical Center, B-3102 Medical Center North, 11615 Twenty-first Ave. South, Nashville, TN 37232-2372. Tel.: 615-322-7299; Fax: 615-322-7381; E-mail: billy.hudson@vanderbilt.edu.

³ The abbreviations used are: GP, Goodpasture; NC1, noncollagenous domain-1; GBM, glomerular basement membrane; TBM, testis basement membrane; MS, mass spectrometry; MS/MS, tandem MS; LC, liquid chromatography; ESI+, electrospray ionization in positive mode; GdnHCl, guanidine hydrochloride; CID, collision-induced dissociation; ELISA, enzyme-linked immunosorbent assay; Met-Hyl, S-hydroxylysyl-methionine; HPLC, high pressure liquid chromatography.

Collagen IV Cross-links and Immune Privilege

denaturants that perturb the NC1 hexamer quaternary structure (7, 9). Of particular note, immunization with $\alpha 3$ NC1 monomer and dimer subunits, having exposed B cell epitopes, induces antibody-mediated glomerulonephritis in rats, rabbits, and mice (10–13). In contrast, their cryptic equivalents, $\alpha 3$ -containing NC1 hexamers, are nonpathogenic (10–12).

Cryptic B cell epitopes have also been implicated in Graves disease and anti-phospholipid disease. In Graves disease, the most common form of hyperthyroidism, autoantibodies target the thyrotropin receptor, which induces hyperactivity of the thyroid and results in goiter and thyrotoxicosis (14). The epitope is partially sequestered between the extracellular A subunit and membrane-bound B subunit of the thyrotropin receptor, rendering it inaccessible for binding of autoantibodies. The epitope, however, can be exposed as a result of intramolecular proteolytic cleavage of the receptor (15). Immunization with the A subunit, which bears the exposed epitope, induces disease. However, immunization with intact thyrotropin receptor, in which the epitope is cryptic, does not induce an immunogenic response. Another example is anti-phospholipid disease where the autoantigen is plasma $\beta 2$ -glycoprotein I in which the cryptic epitope is sequestered by a carbohydrate moiety (16, 17). Interactions with cell membrane phospholipids induce a conformational change in $\beta 2$ -glycoprotein I structure that exposes the cryptic epitope by relocating the carbohydrate moiety away from the autoantigen, allowing binding of pathogenic autoantibodies.

Collectively, the crypticity and pathogenicity of B cell epitopes of the three different autoantigens described above suggest that the exposure of cryptic epitopes upon perturbation of tertiary or quaternary structure is a critical feature that triggers an autoimmune response. An understanding of the structural basis of the crypticity of various autoantigens can provide new insights into the mechanisms of autoimmunity.

In this study, the GP autoantigen, $\alpha 3\alpha 4\alpha 5$ NC1 hexamer, was investigated to determine the molecular basis of the crypticity of epitopes of the D-isoform. This included the elucidation of features of the quaternary structure of the $\alpha 3\alpha 4\alpha 5$ NC1 hexamer by mass spectrometry analysis of NC1 dimer subunits, homology modeling, and molecular dynamics simulation. The findings reveal a novel structural mechanism for immune privilege involving sequestration of epitopes within the autoantigen quaternary structure and stabilization of the structure by cross-linking of subunits.

MATERIALS AND METHODS

Bovine testes were purchased from Pel-Freez biologicals (Rodgers, AR). Bacterial collagenase (CLSPA) was purchased from Worthington.

Immunoprecipitation and Western Blotting—Analysis of TBM hexamers by immunoprecipitation with Mab1 and Mab3 (Wieslab AB, Lund, Sweden) was performed as described previously (18). Antibody-antigen complexes were precipitated with protein G-agarose beads, solubilized in Laemmli sample buffer, and analyzed by immunoblotting. All samples were resolved into NC1 monomers and dimers by SDS-PAGE in replicate 4–22% gradient gels under nonreducing conditions. After transfer to Immobilon P and blocking with casein, $\alpha 1$ – $\alpha 5$

NC1 domains were identified by immunoblotting with specific monoclonal antibodies, as described previously (18). Affinity-purified GP IgG (10 μ g) was incubated with native TBM NC1 hexamers (5–10 μ g) in 200 μ l of Tris-buffered saline, pH 7.4, overnight at room temperature. NC1 subunits, prepared by incubating hexamers with 6 M GdnHCl for 15 min at 60 °C, were diluted at least 20-fold prior to incubation with GP IgG for 1 h. The precipitated NC1 domains were detected by Western blotting using monoclonal antibodies as described above.

Enzyme-linked Immunosorbent Assay (ELISA)—The assay was performed as described before (9), but with modifications. Briefly, 96-well Maxisorp microtiter plastic plates were coated overnight with an anti- $\alpha 3\alpha 4\alpha 5$ NC1 hexamer monoclonal antibody (200 μ l of a 1 μ g/ml solution in carbonate buffer, pH 9.6) and blocked with 2 mg/ml bovine serum albumin. Samples containing TBM NC1 hexamers were diluted with incubation buffer (TBS, pH 7.4, with 1 mg/ml bovine serum albumin and 0.05% Tween 20) and allowed to bind to immobilized antibody for 1 h at 37 °C. To dissociate TBM NC1 hexamers, some antigen samples were treated with 6 M GdnHCl for 15 min at 56 °C before incubation with the immobilized antibody. The $\alpha 3\alpha 4\alpha 5$ NC1 hexamers captured by the immobilized antibody were incubated with three sera of GP patients. Binding of GP autoantibodies was detected by using an alkaline phosphatase-conjugated anti-human antibody (Sigma) and *p*-nitrophenyl phosphate as chromogenic substrate. Absorbance at 405 nm was measured with a Spectramax 190 plate reader (Molecular Devices, Sunnyvale, CA).

Isolation of $\alpha 3\alpha 5$ -heterodimers and $\alpha 4\alpha 4$ -homodimers—TBM NC1 hexamers were prepared as described previously by Kahsai *et al.* (19). Briefly, NC1 hexamers were solubilized from bovine testis by the collagenase digestion method and purified on DE52-cellulose and Sephacryl S-300 columns. Separation of the subunits was performed by acidification of NC1 hexamers with trifluoroacetic acid (0.1% v/v) and fractionation on a C_{18} reverse-phase column (Vydac, Hesperia, CA). NC1 subunits were eluted with increasing concentrations of acetonitrile, and protein elution was monitored by absorbance at 214 nm. Each chromatographic peak was analyzed by Western blotting using monoclonal antibodies Mab1 (anti- $\alpha 1$ NC1 domain) and Mab3 (anti- $\alpha 3$ NC1 domain) from Wieslab AB, Lund, Sweden, and H42 (anti- $\alpha 4$ NC1 domain) and H52 (anti- $\alpha 5$ NC1 domain) (20). The fractions containing $\alpha 3\alpha 5$ -heterodimers (P1) or $\alpha 4\alpha 4$ -homodimers (P2) were denatured and reduced by mixing equal volumes of protein solution with a 0.4 M Tris-HCl, pH 8.5, buffer containing 8 M GdnHCl, and 50 mM dithiothreitol. After heating samples in a boiling water bath for 20 min, P1 and P2 proteins were alkylated with 50 mM iodoacetamide for 30 min at room temperature.

Reduced/alkylated P1 and P2 fractions were each fractionated in a TSK SW_{xl}3000 ToSo-Hass HPLC column equilibrated with 4 M GdnHCl, 50 mM Tris-HCl, pH 7.5. Fractions containing NC1 dimer subunits were precipitated with ethanol at –20 °C and resuspended in 0.1 M ammonium bicarbonate followed by proteolytic digestion with sequencing grade modified trypsin (Promega, Madison, WI) at an ~1:25 ratio at room temperature for 16 h. Cross-linked peptides of $\alpha 4\alpha 4$ -homodimers were obtained by fractionating the P2 fraction on SDS-PAGE,

which was stained with Bio-Safe colloidal Coomassie Brilliant Blue G-250 (Bio-Rad). The “ $\alpha4\alpha4$ -homodimer” band was “in-gel”-digested with trypsin (Promega, Madison, WI) and Asp-N endoproteinase (Princeton Separations, Aldephia, NJ). Sample “clean up” with ZipTipTM-C₁₈ pipette tips (Millipore, Billerica, MA) was performed prior to mass spectrometry analysis.

Mass Spectrometry Analysis of $\alpha3\alpha5$ -heterodimers and $\alpha4\alpha4$ -homodimers—Mass spectrometry analyses were performed on a ThermoFinnigan LTQ linear ion trap mass spectrometer equipped with a Surveyor LC pump and a microelectrospray source controlled by Xcalibur 1.4 software as described before (21). LC-ESI/MS³ analysis was performed using data-dependent scanning in which one full MS spectrum (full mass range of 400–2000 atomic mass units) was followed by one MS/MS spectrum. The three most intense ions in each MS/MS spectrum were subjected to an additional fragmentation (MS³) analysis.

Software for Sequence and Post-translational Modification Analysis—Peptides derived from collagen IV NC1 domains were identified by searching the bovine subset of the Uniref100 data base using the TurboSEQUENT version 27 (revision 12) algorithm (Thermo Electron, San Jose, CA) on a high speed, multiprocessor Linux cluster in the Advanced Computing Center for Research and Education at Vanderbilt University. The search files were adapted for the identification of peptides that included the amino acid modifications of interest *i.e.* –48 atomic mass units on Met⁹³, +46 atomic mass units on Lys²¹¹, or +62 atomic mass units on Lys²¹¹ (46 + 16 atomic mass units in the case of Hyl²¹¹) as well as carboxyamidomethylation of cysteine (+57 atomic mass units) and oxidation of methionine (+16 atomic mass units). Theoretical protein, peptide, and peptide fragment ion masses were generated using General Protein Mass Analysis for Windows version 6.01.1 (Lighthouse Data, Odense, Denmark).

Homology Modeling of the D- $\alpha3\alpha4\alpha5$ NC1 Hexamer—A model of the $\alpha3\alpha4\alpha5$ NC1 hexamer was constructed by a combination of homology modeling and molecular dynamics simulation. The crystal structure of bovine $\alpha1\alpha2\alpha1$ NC1 hexamer (Protein Data Bank accession code 1M3D) was used as the starting template (22). The sequences of human $\alpha3$, $\alpha4$, and $\alpha5$ NC1 domains were aligned with bovine $\alpha1$, $\alpha2$, and $\alpha1$ NC1 sequences, respectively, using ClustalW (23). Of the available $\alpha1\alpha2\alpha1$ NC1 trimers in the asymmetric unit, the one having the overall lowest *B*-factor average was chosen. Chains A ($\alpha1$), B ($\alpha1$), and C ($\alpha2$) were mutated to human $\alpha3$, $\alpha5$, and $\alpha4$ sequences, respectively. In those cases where there were insertions or deletions in the primary sequence, the sequence variations were reiteratively modeled and stereochemically refined, using the molecular graphics software O, version 8 (24). The initial human $\alpha3\alpha4\alpha5$ NC1 trimer model was duplicated to form a “head-to-head” NC1 hexamer using the symmetry operator obtained by the superposition of the two NC1 trimers of the $\alpha1\alpha2\alpha1$ NC1 hexamer crystal structure. The resulting NC1 hexamer model was subjected to molecular dynamics simulation using VMD (25) and NAMD 2.5 (26).

Molecular Dynamics Simulation of the $\alpha3\alpha4\alpha5$ NC1 Hexamer—The homology model was placed in a sphere of water using the SOLVATE module of VMD (25). The molecu-

lar dynamics simulations were carried out with the program NAMD (26) utilizing the CHARMM27 force field (27). The boundary conditions started with the first boundary potential of 10 kcal/mol·Å² at a distance of 75 Å, with the potential exponent set to 2. Force field parameters were chosen in such a way that both van der Waals and electrostatic interactions between 1 and 4 atoms were taken into account. A switching function was employed to smooth van der Waals and electrostatic interactions to zero between 10 and 15 Å. The initial temperature of the system was set at 310 K, and the simulation was run at a step size of 2 fs/step (10 steps per cycle). The bonds were treated as rigid, including those involving hydrogen atoms, and Langevin dynamics were used with a dampening coefficient of 5/ps, set at 310 K (hydrogen atoms excluded). The resulting simulation provided a minimized model after 400 iterations. General analysis of the final model was performed using Chimera (28).

Model Analysis of the Epitopes on $\alpha3\alpha4\alpha5$ NC1 Hexamer—Model analysis and graphics utilized PyMOL 1.0 and MMTSB toolkit (29). Surface area analysis was performed with GETAREA 1.1 (30) and NACCESS 2.1.1 (31). The accessible surface area of the $\alpha3$ NC1 monomer was analyzed within and outside of the NC1 hexamer. We were primarily interested in the accessible surface area of the epitopes on the $\alpha3$ NC1 domain for GP autoantibodies. This semi-quantitative analysis identified residues that are sequestered by neighboring NC1 monomers.

RESULTS

Strategy for Obtaining the Quaternary Structure of the $\alpha3\alpha4\alpha5$ NC1 Hexamer—The quaternary structure of the $\alpha3\alpha4\alpha5$ NC1 hexamer was further defined as the foundation to delineate the molecular basis of the crypticity of epitopes of the D-isoform. In our previous study, we proposed a quaternary structure of the $\alpha3\alpha4\alpha5$ NC1 hexamer based on a composition of ($\alpha3$)₂($\alpha4$)₂($\alpha5$)₂, which allows for a total of 18 possible combinations of different NC1 subunit arrangements (32). The proposed quaternary structure was based on two-dimensional electrophoresis and immunoblotting as well as immunoprecipitation studies using monoclonal antibodies and GP autoantibodies. These studies clearly demonstrated a counterclockwise arrangement of the $\alpha3$, $\alpha4$, and $\alpha5$ within each NC1 trimer. However, the spatial orientation and connections between NC1 subunits across the trimer-trimer interface were only speculated because the existence of a $\alpha3\alpha5$ -heterodimer had not been unambiguously established.

The identification of an $\alpha3\alpha5$ -heterodimer, with respect to its constituent NC1 monomers and the chemical nature of a cross-link that connects the NC1 monomers, can provide novel information that further defines the quaternary structure of the $\alpha3\alpha4\alpha5$ NC1 hexamer. An example is the $\alpha1\alpha1$ -homodimer, first observed by SDS-PAGE (33, 34), that is derived from the $\alpha1\alpha2\alpha1$ NC1 hexamer. This NC1 dimer is stabilized by a cross-link between Met⁹³ and Lys²¹¹ of two $\alpha1$ NC1 monomers juxtaposed at the trimer-trimer interface of the NC1 hexamer. The cross-link was first proposed by Than *et al.* (35), based on electron density connectivity in the crystal structure of the $\alpha1\alpha2\alpha1$ NC1 hexamer. Subsequently, we (21) used mass spectrometry analysis as an independent chemical approach to identify and

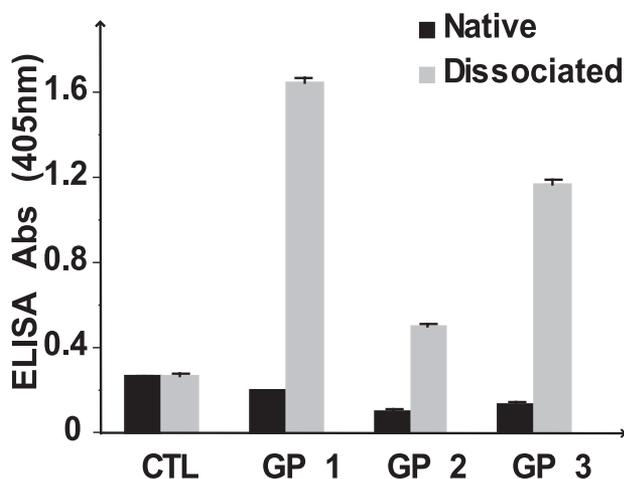


FIGURE 1. Crypticity of GP epitopes in TBM NC1 hexamers. GP autoantibody binding to TBM NC1 hexamers was assayed by ELISA using three GP patient sera (GP1, GP2, and GP3) and a normal serum as control (CTL). GP sera were assayed for their binding to immobilized TBM NC1 hexamers (500 ng/well) under native (black bars) and dissociated (gray bars) conditions (GdnHCl). The assay was carried out as described under "Materials and Methods," and the absorbance at 405 nm of triplicates was averaged. The TBM NC1 hexamers are a mixture of two populations, $\alpha3\alpha4\alpha5$ and $\alpha1\alpha2\alpha1$ NC1 hexamers, as described in Fig. 2.

characterize the putative cross-link and showed it to be a novel covalent bond between Met⁹³ and hydroxylysine 211 (Hyl²¹¹) rather than lysine. The cross-link is characterized by a sulfonium ion linkage and was termed *S*-hydroxylysyl-methionine. Upon NC1 hexamer dissociation, the $\alpha1\alpha1$ -homodimer retains information about the spatial orientation and connectivity between NC1 monomers at the trimer-trimer interface of the parent NC1 hexamer.

In this study, we hypothesized that the *S*-hydroxylysyl-methionine cross-links connect the putative $\alpha3\alpha5$ -heterodimers and the $\alpha4\alpha4$ -homodimers in the D-isomer of the $\alpha3\alpha4\alpha5$ NC1 hexamer. Because large quantities were required for this study, the $\alpha3\alpha4\alpha5$ NC1 hexamer was isolated from TBM, which contains an abundance of the $\alpha3$ chain of collagen IV (19) in comparison with that of GBM. The NC1 dimer subunits were characterized to acquire new information about the quaternary structure, spatial orientation, and connectivity between monomers within the parent $\alpha3\alpha4\alpha5$ NC1 hexamer.

Isolation and Characterization of $\alpha3\alpha4\alpha5$ NC1 Hexamers from Testis Basement Membranes—TBM NC1 hexamers required chemical and immunochemical characterization to establish that they display cryptic epitopes as did those of GBM (9). TBM NC1 hexamers, consisting of a potential mixture of $\alpha3\alpha4\alpha5$ and $\alpha1\alpha2\alpha1$ NC1 hexamers, were isolated as described previously on a gel filtration column (see Fig. 2A in Ref. 19). The crypticity of epitopes in the TBM NC1 hexamers for autoantibodies from three GP patients was evaluated by an ELISA. In each case, the epitopes were cryptic in the native NC1 hexamers, but upon dissociation by GdnHCl the epitopes became exposed and reacted with GP autoantibodies (Fig. 1).

To validate that the GP autoantibody reactivity and crypticity are properties of $\alpha3\alpha4\alpha5$ NC1 hexamers, the TBM NC1 hexamers, presumed to be a mixture of $\alpha3\alpha4\alpha5$ and $\alpha1\alpha2\alpha1$ NC1 hexamers, were characterized for composition of NC1 domains by the same strategy that we previously used for GBM

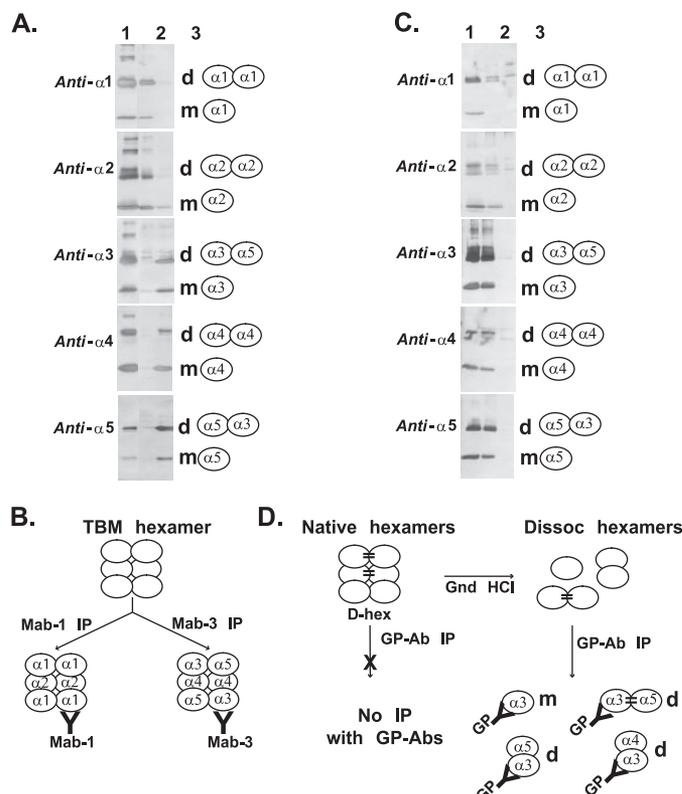


FIGURE 2. Characterization of the TBM NC1 hexamers with respect to NC1 domain composition and crypticity. A, TBM NC1 hexamers (lane 1) were fractionated using Mab1 (lane 2) or Mab3 (lane 3) monoclonal antibodies to NC1 hexamers containing $\alpha1$ and $\alpha3$ subunits, respectively. Mab1- and Mab3-bound fractions were precipitated with protein G beads, resolved by SDS-PAGE into NC1 monomer (m) and NC1 dimer (d) subunits, and then analyzed by Western blot in five replica membranes for the presence of $\alpha1$ – $\alpha5$ NC1 domains, using monoclonal antibodies H1 (*anti- $\alpha1$*), H22 (*anti- $\alpha2$*), H31 (*anti- $\alpha3$*), H43 (*anti- $\alpha4$*), and H53 (*anti- $\alpha5$*). Lane 1 shows the staining of TBM NC1 hexamer (total) which serves as reference for the migration of NC1 domain subunits. B, diagram of the immunoprecipitation (IP) with Mab1 and Mab3 separating the two distinct NC1 hexamer populations. C, immunoprecipitation of NC1 domains from dissociated (lane 2) or native (lane 3) TBM NC1 hexamers using GP antibodies (Ab) (IgG fraction). GP-bound NC1 hexamers were subjected to Western blotting analysis for the presence of $\alpha1$ – $\alpha5$ NC1 domains as described for Fig. 2A. TBM NC1 hexamers were dissociated (Dissoc) with 6 M GdnHCl for 15 min at 60 °C. D, diagram illustrating the immunoprecipitation of the different forms of $\alpha3$ NC1 domains from native and dissociated NC1 hexamers.

NC1 hexamers (32). Monoclonal antibodies to either $\alpha1$ (Mab1) and $\alpha3$ (Mab3) NC1 domains were used to precipitate $\alpha1$ - or $\alpha3$ -containing NC1 hexamers, respectively. The antibody-bound fractions were analyzed by SDS-PAGE and Western blot using antibodies specific for $\alpha1$ – $\alpha5$ NC1 domains (Fig. 2A). Mab1-bound fractions contained monomers and homodimers of $\alpha1$ and $\alpha2$ NC1 domains derived from $\alpha1\alpha2\alpha1$ NC1 hexamers but not $\alpha3$, $\alpha4$, or $\alpha5$ NC1 domains (Fig. 2A, lane 2, all blots). In contrast, Mab3-bound fractions contained monomers and dimers of $\alpha3$, $\alpha4$, and $\alpha5$ NC1 domains derived from $\alpha3\alpha4\alpha5$ NC1 hexamers but not $\alpha1$ or $\alpha2$ NC1 domains (Fig. 2A, lane 3, all blots). Collectively, these results demonstrate that TBM NC1 hexamers contain two populations, $\alpha1\alpha2\alpha1$ and $\alpha3\alpha4\alpha5$ NC1 hexamers, as illustrated in Fig. 2B, which reflect the presence of two separate collagen IV networks.

Crypticity of TBM $\alpha3\alpha4\alpha5$ NC1 hexamers was evaluated by immunoprecipitation as described previously for GBM (9).

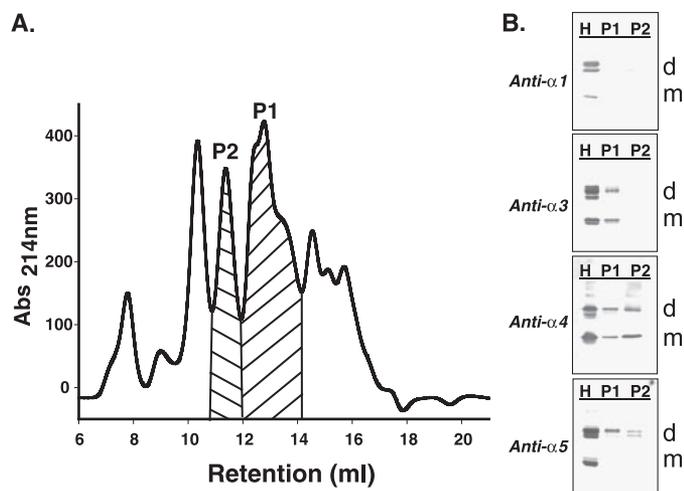


FIGURE 3. Isolation of $\alpha 3\alpha 5$ -heterodimer and $\alpha 4\alpha 4$ -homodimer. *A*, separation of NC1 hexamer subunits by a C_{18} reverse-phase HPLC column. Protein elution was monitored by measuring absorbance at 214 nm. As shown in the chromatographic profile, fractions were grouped into pools P1 and P2. *B*, P1 and P2 fractions were analyzed by Western blots using antibodies Mab1 (*anti- $\alpha 1$*), Mab3 (*anti- $\alpha 3$*), H42 (*anti- $\alpha 4$*), and H52 (*anti- $\alpha 5$*) against NC1 domains. The migration of NC1 monomers and dimers is indicated by *m* and *d*, respectively. Fraction P1, composed mainly of $\alpha 3\alpha 5$ -heterodimer, and fraction P2, enriched in $\alpha 4\alpha 4$ -homodimer, were subsequently used in cross-link analyses.

TBM NC1 hexamers in native or dissociated conditions were incubated with GP autoantibodies. The immunoprecipitated NC1 domains were analyzed by Western blotting using chain-specific antibodies for $\alpha 1$ – $\alpha 5$ chains (Fig. 2*C*). Under dissociating conditions, GP antibodies precipitated monomeric and dimeric forms of $\alpha 3$ NC1 domain (Fig. 2*C*, lane 2, *anti- $\alpha 3$* blot). In addition, the $\alpha 3$ dimer also stained with the anti- $\alpha 5$ NC1 domain antibody (Fig. 2*C*, lane 2, *anti- $\alpha 5$* blot), suggesting the presence of $\alpha 3\alpha 5$ -heterodimers and/or $\alpha 3\alpha 3$ - and $\alpha 5\alpha 5$ -homodimers. The presence of $\alpha 4$ NC1 domain in the anti- $\alpha 4$ blot is because of its strong noncovalent interactions with $\alpha 3$ NC1 domain (32). In addition, there is a monomer-sized band in the anti- $\alpha 5$ blot that corresponds to $\alpha 5$ NC1 domain (Fig. 2*C*, lane 2), reflecting a strong interaction with $\alpha 3$ NC1 monomer. Conversely, under native conditions (Fig. 2*C*, lane 3 of all blots) GP antibodies do not react with NC1 hexamers as no NC1 domains were detected after immunoprecipitation.

Collectively, the immunoprecipitation studies demonstrate that the GP epitopes are cryptic within the $\alpha 3\alpha 4\alpha 5$ NC1 hexamer of testis basement membrane, as illustrated in Fig. 2*D*. The results are consistent with the ELISA results presented in Fig. 1, where reactivity was observed only after NC1 hexamer dissociation. Therefore, the $\alpha 3\alpha 4\alpha 5$ NC1 hexamers in TBM exist in the D-isofrom, harboring cryptic epitopes to GP autoantibodies. In contrast, both M- and D-isofroms occur in GBM NC1 hexamers, where the latter is predominant. Moreover, the D-hexamers from TBMs were isolated in milligram quantities that were required for cross-link analysis.

Isolation of NC1 Dimers from $\alpha 3\alpha 4\alpha 5$ D-hexamer—To facilitate cross-link analysis, involving proteolytic digestion and mass spectrometry, enriched fractions of $\alpha 3\alpha 5$ -heterodimers and $\alpha 4\alpha 4$ -homodimers were obtained by fractionation of TBM NC1 hexamers on a reverse-phase C_{18} column (Fig. 3*A*). The

acidic conditions of reverse phase dissociate NC1 hexamers with concomitant separation of the different NC1 monomers and dimers. Fractions containing the respective NC1 dimers were identified by SDS-PAGE and immunoblotting using chain-specific monoclonal antibodies (Fig. 3*B*). Western blot analysis shows that P1 is mainly composed of $\alpha 3$ NC1 monomers and $\alpha 3\alpha 5$ -containing NC1 dimers⁴ (Fig. 3*B*). Although P1 fraction also contains $\alpha 4$ NC1 monomer, they did not interfere with mass spectrometry analysis. Fraction P2 predominantly contains $\alpha 4$ NC1 monomers and $\alpha 4\alpha 4$ -homodimers. Fractions containing $\alpha 1$ and $\alpha 2$ NC1 domains elute before P2 in the chromatogram (data not shown). Fractions P1 (enriched in $\alpha 3\alpha 5$ -containing NC1 dimers) and P2 (enriched in $\alpha 4\alpha 4$ -homodimers) were further fractionated on a gel filtration column to devoid them of NC1 monomers for cross-link analysis (data not shown).

Strategies for Cross-link Analyses of NC1 Dimers—We utilized a strategy of proteolytic digestion coupled with mass spectrometric analysis, as we described previously to characterize the cross-link of the $\alpha 1\alpha 1$ -homodimers (21). We searched for peptides encompassing the conserved Met⁹³ and Lys²¹¹ residues, which are presumed to harbor a cross-link (Fig. 4). The quaternary structure of $\alpha 3\alpha 4\alpha 5$ D-hexamers is presumed to be composed of up to three NC1 dimers, which result from the covalent interaction of NC1 monomers at the trimer-trimer interface; thus, two $\alpha 3\alpha 5$ -heterodimers and one $\alpha 4\alpha 4$ -homodimer are expected upon dissociation of the $\alpha 3\alpha 4\alpha 5$ NC1 hexamer (Fig. 5*A*) (32). The *in silico* analysis of the sequences of $\alpha 3$, $\alpha 4$, and $\alpha 5$ NC1 domains indicates that trypsin would generate three peptides containing Met⁹³ and three peptides containing Lys²¹¹, all of which have different lengths and amino acid compositions because of local sequence differences between three NC1 domains (Fig. 4 and Fig. 5*A*). A random combination between Met⁹³ with Lys²¹¹ containing peptides from any of the three NC1 domains would generate a total of nine different peptide complexes. However, quaternary structure constraints (21, 32, 35) would only allow the formation of cross-links between $\alpha 3$ and $\alpha 5$ NC1 domains to form $\alpha 3\alpha 5$ -heterodimers and between two $\alpha 4$ NC1 domains to a form $\alpha 4\alpha 4$ -homodimer. As illustrated in Fig. 5*A*, trypsin digestion of $\alpha 4\alpha 4$ -homodimer should generate the T-6498 complex. Likewise, trypsin digestion of $\alpha 3\alpha 5$ -heterodimers would generate the T-4481 and T-3959 complexes, each connected by the *S*-hydroxylysyl-methionine (Met-Hyl) cross-link (Fig. 5*A*).

Each of these peptide complexes (T-6498, T-4481, and T-3959) should display a fragmentation pattern by mass spectrometry that is characteristic of the Met-Hyl cross-link, as described previously for the $\alpha 1\alpha 1$ -homodimers (21). Fragmentation of the Met-Hyl cross-link by collision-induced dissociation (CID) breaks the bond between C- γ and S- δ of Met⁹³ resulting in a mass loss of 48 atomic mass units from Met⁹³ (thiomethyl group) and a concomitant mass gain

⁴ We used the term “ $\alpha 3\alpha 5$ -containing NC1 dimers” here to indicate that the composition of these NC1 dimers is uncertain prior to the findings of the present study. The NC1 dimers could represent $\alpha 3\alpha 5$ -heterodimers as well as $\alpha 3\alpha 3$ - or $\alpha 5\alpha 5$ -homodimers co-migrating in SDS-PAGE.

Collagen IV Cross-links and Immune Privilege

of 46 atomic mass units by Hyl²¹¹ (Fig. 5B). This characteristic fragmentation profile was used to identify tryptic peptide complexes containing the Met-Hyl cross-link by search-

ing for cross-link-derived peptides containing either of these two modified amino acids using the SEQUEST algorithm (36) as indicated under "Materials and Methods."

```

Col14A1  SVDHGLVTRHSQTTDDPQCPPGPKILYHGYSLLYVQGNERAHGQDLGTAGSCLRFKFTM
Col14A2  -ISIGYLLVKHSQTDQEPMPVGMNKLWSGYSLLYFEGQEKAHNQDLGLAGSCLARFSTM
Col14A3  AVMRGFVFTTRHSQTTAIPSCPEGTEPLYSGFSLLEFVQGNQAHGQDLGTAGSCLQRFTTM
Col14A4  GYLSGFLLVLSHSQTDGEPTECPMGMPRLWTGYSLLYLEGQERAHNQDLGLAGSCLPIFSTL
Col14A5  SVAHGFLITRHSQTTDAPQCPQGTLOVYEGFSLLYVQGNKRAHGQDLGTAGSCLRRFSTM
Col14A6  -LRVGYTLVKHSQSEVPLCPSGMSQLWVGYSLLFVEGQEKAHNQDLGFAGSCLPRFSTM
          *: .. ***: * * * * : : *::: : : : : * * * * * * * * * * * * * * *
          * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Col14A1  PFLFCNINNVCFASRNDYSYWLSTPEPMPMSMAPITGENIRPFISRCVCEAPAMVMAV
Col14A2  PFLYCNPGDVCYASRNDKSYWLST--TAPLPMPVAAEDIRPYISRCVCEAPAVAIIV
Col14A3  PFLFCNINDVCFASRNDYSYWLSTPAMIPMDMAPITGRALEPYISRCTVCEGPAIAIIV
Col14A4  PFAYCNIHQVCHYARRNDRSYWLAS--TAPLPMTPLSEDIRPYISRCVCEAPAVAIIV
Col14A5  PFMFCNINNVCFASRNDYSYWLSTPEPMPMSMQPLKQSIQFFISRCVCEAPAVVIAV
Col14A6  PFIYCNIQVCHYARRNDRSYWLST--TAPIPMPVQGTQIQYISRCVCEAPSAQIAIV
          ** : * * : * * : * * * * * * * * * : : * * * * : : * * * * * * * * : : * *
          * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Col14A1  HSQTIQIPQCPTGWSSLWIGYSFVMHTSAGAEGSQALASPGSCLEEFRSAPFIECHG-R
Col14A2  HSQDVSIPHCAGWRSWLWIGYSFLMHTAAGDEGGQSLVSPGSCLEDFRATPFIECNGAR
Col14A3  HSQTTDIPPCAGWISLWKGFSFIMFTSAGSEAGQALASPGSCLEEFRSAPFIECHG-R
Col14A4  HSQDQSIQPCPRAWRSWLWIGYSFLMHTGAGDQGGQALMSPGSCLEDFRAAPFLECQGRQ
Col14A5  HSQTIQIPHCQGWDSLWIGYSFMMHTSAGAEGSQALASPGSCLEEFRSAPFIECHG-R
Col14A6  HSQDITIPQCPLGWRSLWIGYSFLMHTAAGAEGGQSLVSPGSCLEDFRATPFIECSGAR
          *** ** * * . * * * * * * * * * * * * * * * * * * * * * * * * * * *
          * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Col14A1  GTCNYYANYSFWLATIERSEMFKK-PTPSTLKA93GELR-THVSRQVCMRRT-
Col14A2  GTCHYYANKYSFWLTTIPE-QSFQGTSPADTLKAGLIR-THISRCQVCMKNL-
Col14A3  GTCNYYNSYSFWLASLDPKRMFRK-PIPSTVKAGELE-NISRCQVCMKMRP
Col14A4  GTCHFFANKYSFWLTTVVRPDLQFSSAPLPD211LKESHARQKISRQCVKHS-
Col14A5  GTCNYYANSYSFWLATVDVSDMFSK-PQSETLKA93GDLR-TRISRCQVCMKRT-
Col14A6  GTCHYFANKYSFWLTTVEERQQFGEEPSETLKA93QLH-TRVSRQVCMKSV-
          ***** : : * * * * * * * * * * * * * * * * * * * * * * * * *
  
```

FIGURE 4. Sequence alignment shows that Met⁹³ and Lys²¹¹ are conserved in the NC1 domains of collagen IV. A multiple sequence alignment between the NC1 domains of bovine collagen IV α 1- α 6 chains shows that amino acids Met⁹³ (red) and Lys²¹¹ (blue) are conserved in all six chains. The red box around Lys²¹¹ denotes the consensus sequence for lysyl hydroxylase (53). The alignment was produced by ClustalW.

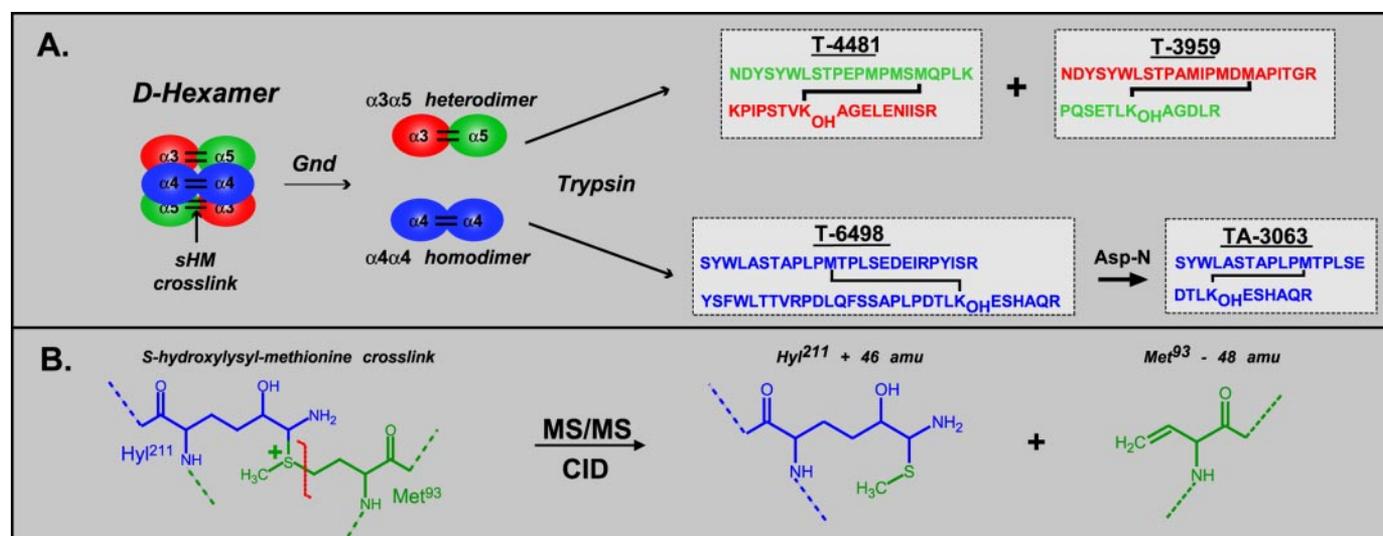


FIGURE 5. Experimental strategy for cross-linking analysis of NC1 dimers of the α 3 α 4 α 5 D-hexamer. A shows that the α 3 α 4 α 5 D-hexamer putatively contains two heterodimers, composed of α 3 (red) and α 5 (green) NC1 domains, and one homodimer, composed of α 4 (blue) NC1 domains. The NC1 dimers are presumed to contain Met-Hyl cross-links (black bars) that were previously found in the α 1 α 1-homodimer (21). The Met-Hyl cross-link is formed between Met⁹³ from an NC1 domain of one trimer and a Hyl²¹¹ from an NC1 domain of the opposite trimer (Lys²¹¹ is post-translationally modified to Hyl²¹¹). A total of six Met-Hyl cross-links (two per NC1 dimer) can be formed at the trimer-trimer interface (21, 35). Dissociation of α 3 α 4 α 5 D-hexamers yields α 3 α 5-heterodimers and α 4 α 4-homodimer. Trypsin digestion of α 3 α 5-heterodimers releases two peptide complexes (T-4481 and T-3959), each containing the Met-Hyl cross-link. Trypsin digestion of α 4 α 4-homodimer yields one complex (T-6498) containing the Met-Hyl cross-link, which is further digested with Asp-N endopeptidase to reduce its size (TA-3063 complex). B, proposed structure of the Met-Hyl cross-link based on mass spectrometry analysis of the α 1 α 1-homodimer of placenta basement membrane (21). The side chains of Met⁹³ and Hyl²¹¹ are connected by a sulfonium ion. Upon fragmentation of the cross-link by CID, the covalent bond between C- γ and S- δ breaks, resulting in the loss of 48 atomic mass units (thiomethyl group) from Met⁹³ with a concomitant gain of 46 atomic mass units by Hyl²¹¹.

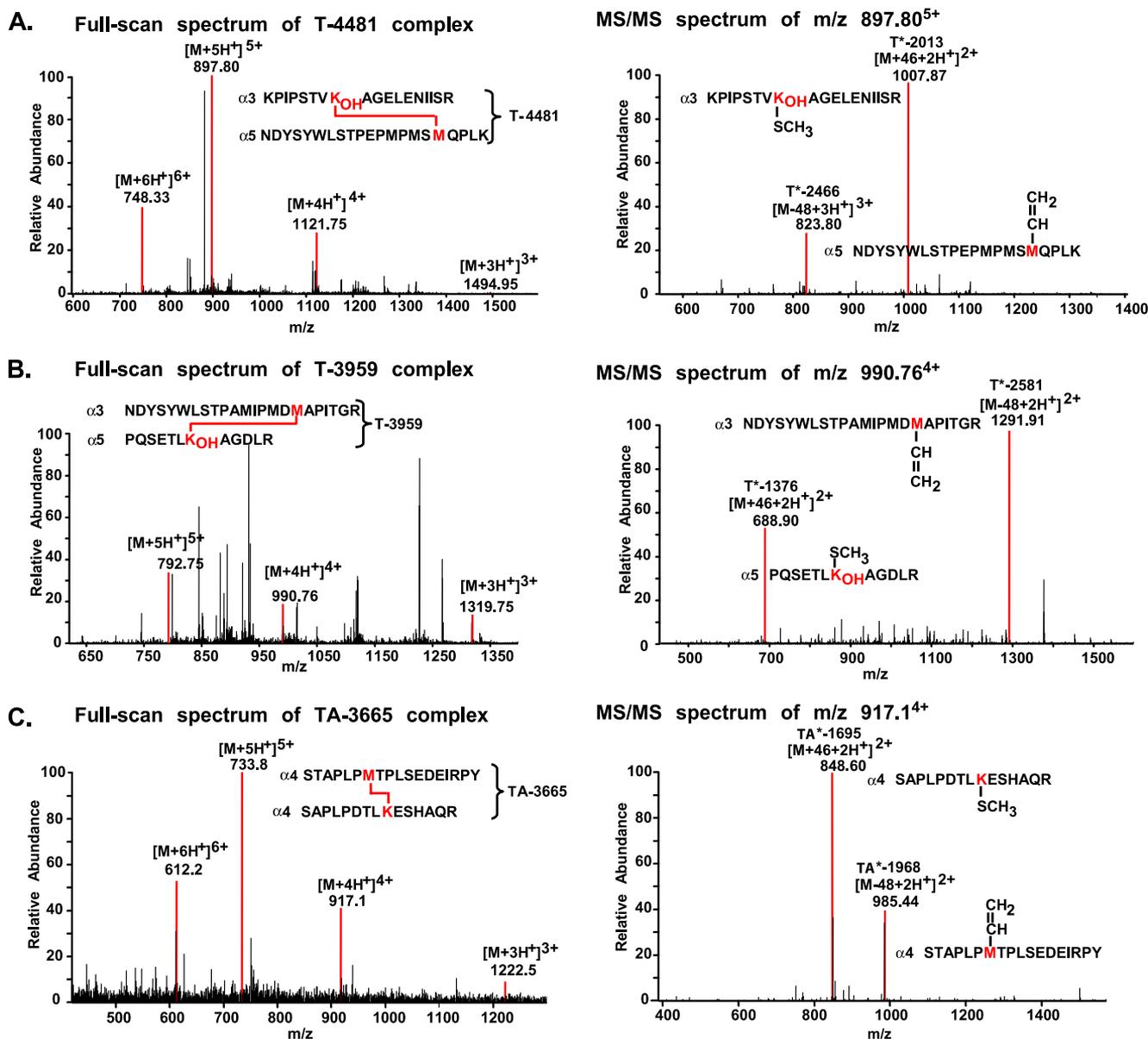


FIGURE 6. Mass spectrometry analysis of tryptic peptide complexes derived from $\alpha 3\alpha 5$ -heterodimer and $\alpha 4\alpha 4$ -homodimers. A, fraction P1 (Fig. 3) containing mainly $\alpha 3\alpha 5$ -heterodimers was digested with trypsin and analyzed by LC-ESI/MS³. *Left panel*, ESI spectrum of the T-4481 complex constituted by NDYSYWLSTPEPMPMSMQPLK (T-2514) and KPIPSTV_{KOH}AGELENIISR (T-1967) peptides. Multiply charged ions of the complex (+3, +4, +5, and +6) are indicated in red for better visualization. *Right panel*, MS/MS spectrum generated by fragmentation of the quintuply charged ion 897.8 m/z produces the 823.8 m/z and 1007.87 m/z ions. The 823.8 m/z is the triply charged ion for T*-2466 peptide, which represents the loss of 48 atomic mass units by the T-2514 peptide. The 1007.87 m/z is the doubly charged ion for T*-2013 peptide, which represents the gain of 46 atomic mass units by the T-1967 peptide. The asterisk indicates that such tryptic peptide contains a modified residue. The sequences of T*-2466 and T*-2013 peptides, including their amino acid modifications, were confirmed by further fragmentation (MS³) of 823.8 m/z and 1007.87 m/z ion, respectively (supplemental Fig. 9A). B, *left panel*, ESI spectrum of the T-3959 complex constituted by NDYSYWLSTPAMIPMDMAPITGR (T-2629) and PQSETL_{KOH}AGDLR (T-1330) peptides. Multiply charged ions of the complex (+3, +4, and +5) are indicated in red for better visualization. *Right panel*, MS/MS spectrum generated by fragmentation of the 990.76 m/z ion produces the 1291.91 m/z and 688.90 m/z ions. The 1291.91 m/z is the doubly charged ion for T*-2581 peptide, which represents the loss of 48 atomic mass units by the T-2629 peptide. The 688.90 m/z is the doubly charged ion for the T*-1376 peptide, which represents the gain of 46 atomic mass units by the T-1330 peptide. The asterisk indicates that such tryptic peptide contains a modified residue. The sequence of T*-2581 and T*-1376 peptides, including their amino acid modifications, were confirmed by further fragmentation (MS³) of 1291.91 m/z and 688.90 m/z ions, respectively (supplemental Fig. 9B). C, LC-ESI/MS³ analysis of fraction P2 (Fig. 3), containing mainly $\alpha 4\alpha 4$ -homodimers, after digestion with trypsin and Asp-N endoproteinase. *Left panel*, ESI spectrum of the TA-3665 complex constituted by STAPLPM_TPLSEDIRPY (TA-2016) and SAPLPDTL_KESHAQR (TA-1649) peptides. Multiply charged ions of the complex (+3, +4, +5, and +6) are indicated in red for better visualization. *Right panel*, MS/MS spectrum generated by fragmentation of 917.1 m/z ion produces the 985.44 m/z and 848.60 m/z ions. The 985.44 m/z is the doubly charged ion for TA*-1968 peptide, which represent the loss of 48 atomic mass units by TA-2016 peptide. The 848.60 m/z is the doubly charged ion for the TA*-1695 peptide, which represents the gain of 46 atomic mass units by the TA-1649 peptide. The asterisk indicates that such peptide contains a modified residue. The sequences of TA*-1968 and TA*-1695 peptides, including their amino acid modifications, were confirmed by further fragmentation (MS³) of 985.44 m/z and 848.60 m/z ions, respectively (supplemental Fig. 9C).

derived from the $\alpha 3$ NC1 domain plus 46 atomic mass units, indicating the characteristic fragmentation of the Met-Hyl cross-link (Fig. 6A, right panel).

To confirm that these ions correspond to the peptide sequences predicted in Fig. 5A, including the chemical modifications of Met⁹³ and Hyl²¹¹, 823.8 m/z and 1007.87 m/z ions

Collagen IV Cross-links and Immune Privilege

were selected for further fragmentation (MS^3) (supplemental Fig. 9A). The *b*- and *y*-series of the MS^3 fragmentation profile for the 828.8 *m/z* ion are consistent with the sequence for the $\alpha 5$ Met⁹³-containing peptide lacking the 48 atomic mass units at Met⁹³ (T*-2466 peptide). Similarly, the MS^3 fragmentation profile of 1007.87 *m/z* ion is consistent with the sequence for the $\alpha 3$ Hyl²¹¹-containing peptide in which Hyl²¹¹ has gained 46 atomic mass units (T*-2013 peptide).

Fig. 6B (left panel) shows a full scan ESI+ mass spectrum containing the multiple charged ions (+3, +4, +5, and +6) that represent a T-3959 complex, as was predicted in Fig. 5A for a second cross-link in the $\alpha 3\alpha 5$ -heterodimer. The average mass obtained by deconvolution of the ESI+ mass spectrum is equivalent to that calculated for the T-3959 complex. Fragmentation of the quadruply charged ion (990.76 *m/z*, Fig. 6B, right panel), produces a doubly charged ion (1291.91 *m/z*) that corresponds to the mass of the Met⁹³-containing peptide derived from $\alpha 3$ NC1 domain minus 48 atomic mass units (T*-2581 peptide) and a doubly charged ion at 688.90 *m/z*, which is equivalent to the mass of the Hyl²¹¹-containing peptide derived from the $\alpha 5$ NC1 domain plus 46 atomic mass units (T*-1376 peptide), indicating the characteristic fragmentation pattern of the Met-Hyl cross-link. To confirm that these ions correspond to the peptide sequences predicted in Fig. 5A, including the chemical modifications of Met⁹³ and Hyl²¹¹, a fragmentation profile (MS^3) of the 1291.91 *m/z* and 688.90 *m/z* ions was acquired as indicated above (supplemental Fig. 9B).

Thus, mass spectrometric analysis demonstrates the presence of both T-4481 and T-3959 complexes, as predicted in Fig. 5A, in which peptides derived from $\alpha 3$ and $\alpha 5$ NC1 monomers are linked by the Met-Hyl cross-link. The presence of the Met-Hyl cross-link in the T-4481 and T-3959 complexes establishes a covalent interaction between $\alpha 3$ and $\alpha 5$ NC1 monomers, which unambiguously establishes the existence of an $\alpha 3\alpha 5$ -heterodimer, and verifies the identity of a $\alpha 3$ -containing-heterodimer first observed by Kleppel *et al.* (37). Moreover, the location of the cross-link reveals that the spatial orientation and connectivity of the $\alpha 3$ and $\alpha 5$ NC1 monomers is homologous to that established previously for the $\alpha 1\alpha 1$ NC1 homodimer. In addition, among the two possible cross-linking sites per dimer (35), the $\alpha 3\alpha 5$ -heterodimer provides the first chemical evidence that Met-Hyl cross-links occur at both sites (Fig. 5A), a structural feature that was not possible to discern in the previous analysis of $\alpha 1\alpha 1$ -homodimers (21).

Mass Spectrometry Analysis of $\alpha 4\alpha 4$ -Homodimers—The $\alpha 4\alpha 4$ -homodimers were digested with trypsin and analyzed by mass spectrometry as described above for the $\alpha 3\alpha 5$ -heterodimers. However, none of the peptide sequences containing Met⁹³ – 48 atomic mass units or Hyl²¹¹ + 46 atomic mass units were found by SEQUEST, possibly because of the large size of the cross-linked peptides. Therefore, the predicted T-6498 complex (Fig. 5A) could not be detected in the ESI- MS^3 analysis.

For further MS analyses, the tryptic complexes were further digested with Asp-N endoproteinase in an attempt to yield the smaller predicted TA-3063 complex (Fig. 5A). The digested products were examined by an ESI- MS^3 experiment in which the peptide sequences containing Met⁹³ – 48 atomic mass units

and Hyl²¹¹ + 46 atomic mass units were searched with SEQUEST. Although a SEQUEST search did not find the predicted TA-3063 complex, it did reveal an unexpected TA-3665 complex composed of two $\alpha 4$ -derived peptides as follows: one peptide encompassing Lys²¹¹ and the other Met⁹³. The mass difference between the predicted TA-3063 complex and the observed TA-3665 complex reflects the replacement of Hyl²¹¹ with Lys, together with additional residues at the amino or carboxyl terminus, indicating that cleavage occurred at sites other than those predicted (see below). Fig. 6C, left panel, shows the full scan mass spectrum containing the multiple charged ions (+3, +4, and +5) of the TA-3665 complex. Fragmentation (MS/MS) of quadruply charged ion (917.41 *m/z*) generated the fragment ions 848.6 *m/z* and 985.44 *m/z* whose masses are consistent with an $\alpha 4$ Lys²¹¹-containing peptide plus 46 atomic mass units (TA*-1695 peptide) and an $\alpha 4$ Met⁹³-containing peptide minus 48 atomic mass units (TA*-1968 peptide), respectively. To confirm the sequences and amino acid modifications of these peptides, 848.6 *m/z* and 985.44 *m/z* ions were subjected to further CID fragmentation (MS^3). The predicted *b*- and *y*-ion fragmentations for the ion 848.6 *m/z* are consistent with sequence for the TA*-1695 peptide and clearly shows the extra 46 atomic mass units attached to Lys²¹¹, rather than hydroxylysine (supplemental Fig. 9C). Furthermore, the *b*- and *y*-ions produced by fragmentation of 985.44 *m/z* are consistent with the sequence for TA*-1968 peptide in which Met⁹³ is missing 48 atomic mass units.

Hydroxylysine was expected at position 211 of $\alpha 4$ NC1 domain because it occurs in the $\alpha 1\alpha 1$ -homodimer (21) and in the $\alpha 3\alpha 5$ -heterodimer (see above). The absence of hydroxylysine, however, is consistent with the amino acid composition of the NC1 domains reported many years ago (7), which demonstrated that all the NC1 domains contain hydroxylysine except for M3, which would later be named the $\alpha 4$ NC1 domain. Collectively, these results demonstrate that the TA-3665 complex is composed of two peptides that are connected by a covalent cross-link between the side chains of Met⁹³ and Lys²¹¹, which unambiguously establishes the existence of an $\alpha 4\alpha 4$ -homodimer. Furthermore, the results establish that the NC1 monomers of the $\alpha 4\alpha 4$ -homodimer are covalently connected by a novel cross-link involving Met⁹³ and unmodified Lys²¹¹, designated herein as *S*-lysyl-methionine (Met-Lys). The amino acid sequence of the $\alpha 4$ NC1 domain may provide an explanation for the absence of hydroxylation of Lys²¹¹ as discussed below.

DISCUSSION

In this study, the Goodpasture autoantigen, the $\alpha 3\alpha 4\alpha 5$ NC1 hexamer, was characterized with respect to its quaternary structure, as revealed by mass spectrometry analysis of cross-links in NC1 dimer subunits, homology modeling, and molecular dynamics simulation.⁵ This information provided the framework for interpreting the structural basis for the crypticity of B cell epitopes of the GP autoantigen. The findings reveal

⁵The $\alpha 3\alpha 4\alpha 5$ NC1 hexamer model obtained via homology modeling and molecular dynamics simulation is available at the Vanderbilt Center for Matrix Biology website.

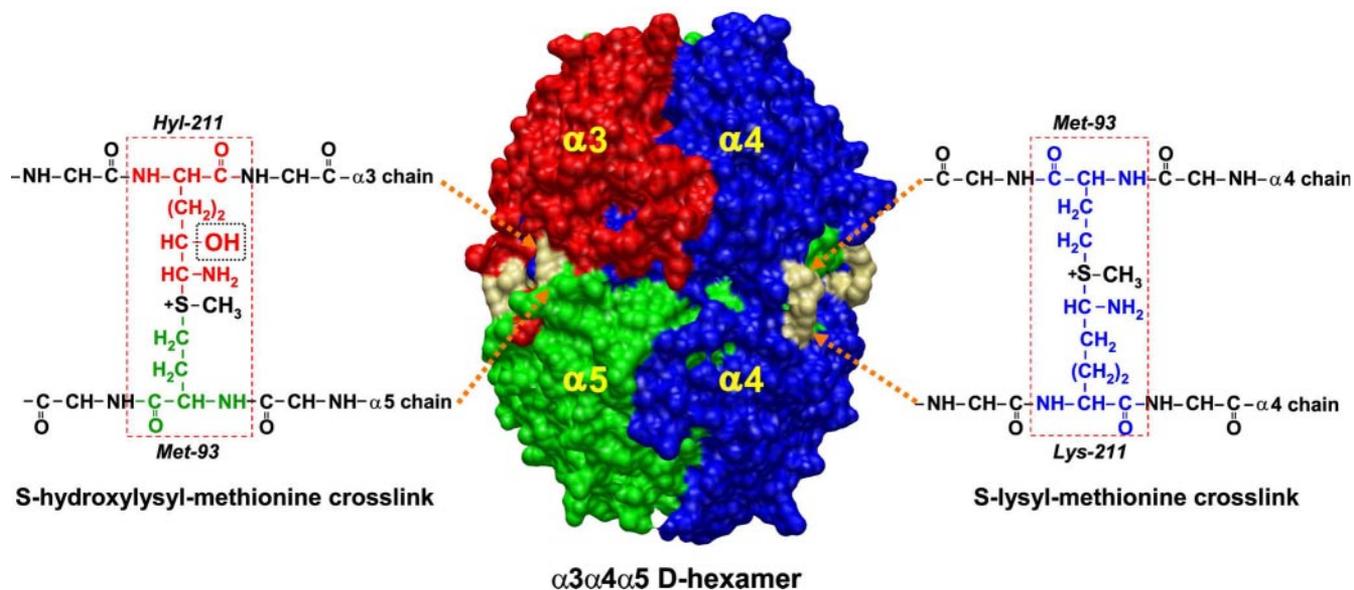


FIGURE 7. Location of Met-Hyl and Met-Lys cross-links and quaternary structure of $\alpha3\alpha4\alpha5$ D-hexamers. This model was generated by homology modeling using the coordinates of the $\alpha1\alpha2\alpha1$ NC1 hexamer (22) and the cross-linking information obtained here. The NC1 hexamer is composed of two trimeric caps each consisting of $\alpha3$ (red), $\alpha4$ (blue), and $\alpha5$ (green) NC1 domain. The two heterotrimers interact through a large planar interface forming NC1 hexamer. Heterodimers ($\alpha3\alpha5$ and $\alpha5\alpha3$) and homodimers ($\alpha4\alpha4$), which are stabilized by Met-Hyl and Met-Lys cross-links (khaki color), respectively, were assembled into a NC1 hexamer. Molecular graphics images were produced using the University of California, San Francisco, Chimera (28).

new molecular features of the quaternary structure of the GP autoantigen that may provide clues about the etiology of GP disease.

Cross-links of the $\alpha3\alpha4\alpha5$ NC1 D-isoform—Using mass spectrometry, we characterized NC1 dimer subunits in search of a putative cross-link analogous to that previously found for the $\alpha1\alpha1$ -homodimer (21). We searched for tryptic peptide complexes that correspond to regions of the $\alpha3$, $\alpha4$, and $\alpha5$ NC1 domains encompassing the conserved Met⁹³ and Lys²¹¹ residues, which were presumed to harbor a cross-link (Fig. 4). The analyses revealed the presence of Met-Hyl cross-links in the $\alpha3\alpha5$ -heterodimers, analogous to that found in the $\alpha1\alpha1$ -homodimers, as described previously (21). Surprisingly, the findings revealed a Met-Lys cross-link in the $\alpha4\alpha4$ -homodimers, indicating that Lys²¹¹ of the $\alpha4$ NC1 monomer does not undergo post-translational modification to hydroxylysine. Typically, hydroxylation of lysine is catalyzed by lysyl hydroxylase, which recognizes the X-Lys-Gly motif in the triple-helical domain of various collagens (38). It also occurs in nonhelical regions of collagen I at the end of the triple-helical domain that contains an X-Lys-(Ala/Ser) sequence (38). At these nonhelical regions, the lysine and hydroxylysine residues undergo enzymatic oxidation, which induces formation of aldehyde-derived cross-links that stabilizes the collagen fibril. This sequence is identical to the X-Lys²¹¹-Ala sequence present in the NC1 domains of $\alpha1$, $\alpha2$, $\alpha3$, $\alpha5$, and $\alpha6$ chains of collagen IV (Fig. 4), and which undergo hydroxylation and form the Met-Hyl cross-links. However, this sequence differs in the NC1 domain of the $\alpha4$ chain, where it is X-Lys²¹¹-Glu (Fig. 4), a distinction that may explain the lack of hydroxylation of the $\alpha4\alpha4$ -homodimer.

The existence of the Met-Lys rather than Met-Hyl cross-link in the $\alpha4\alpha4$ -homodimer indicates that the hydroxylation of Lys²¹¹ is not required for cross-link formation. However, this hydroxyl group in the $\alpha3\alpha5$ -heterodimers and $\alpha1\alpha1$ -ho-

modimers may contribute to the stability of the cross-link, as in the case of collagen I where the hydroxylysine-derived cross-links are more stable than those derived from lysine residues (39). It may also serve as an attachment site for carbohydrate units, galactose, or the disaccharide glucosyl-galactose and play a role in regulating the self-assembly of protomers into collagen IV networks (21). Clearly, hydroxylation of collagen IV molecules is crucial for normal basement membrane function as was recently shown from studies of lysyl hydroxylase null mice (40, 41).

Three-dimensional Model of the D-isoform of the $\alpha3\alpha4\alpha5$ NC1 Hexamer—The identification of the $\alpha3\alpha5$ -heterodimer and the $\alpha4\alpha4$ -homodimer together with the chemical nature and location of cross-links establishes the spatial arrangement and connectivity of the NC1 monomers at the trimer-trimer interface of the $\alpha3\alpha4\alpha5$ NC1 hexamers, as illustrated in Fig. 7. The direct contacts between $\alpha3$ and $\alpha5$ NC1 monomers via the Met-Hyl cross-link, as well as two $\alpha4$ NC1 monomers via the novel Met-Lys cross-link, are illustrated. Notably, the proposed chemical structure for the Met-Lys cross-link is also shown (Fig. 7). The findings verify the quaternary structure and the chain composition of the protomers of the $\alpha3\alpha4\alpha5$ network of collagen IV, as previously proposed (32), and establish the end-to-end (NC1-to-NC1) connections between protomers of this network.

The overall quaternary structure (Fig. 7) provides the framework to construct a three-dimensional model (Fig. 8) of the $\alpha3\alpha4\alpha5$ NC1 hexamer by homology modeling and molecular dynamics simulation, using the atomic coordinates from the crystal structure of the homologous $\alpha1\alpha2\alpha1$ NC1 hexamer (22). The quaternary structure is characterized by two $\alpha3\alpha4\alpha5$ NC1 heterotrimers, where the $\alpha3$, $\alpha4$, and $\alpha5$ NC1 domains are arranged counterclockwise (viewed from the triple-helical pole) in a pseudo-3-fold symmetry axis. The two NC1 hetero-

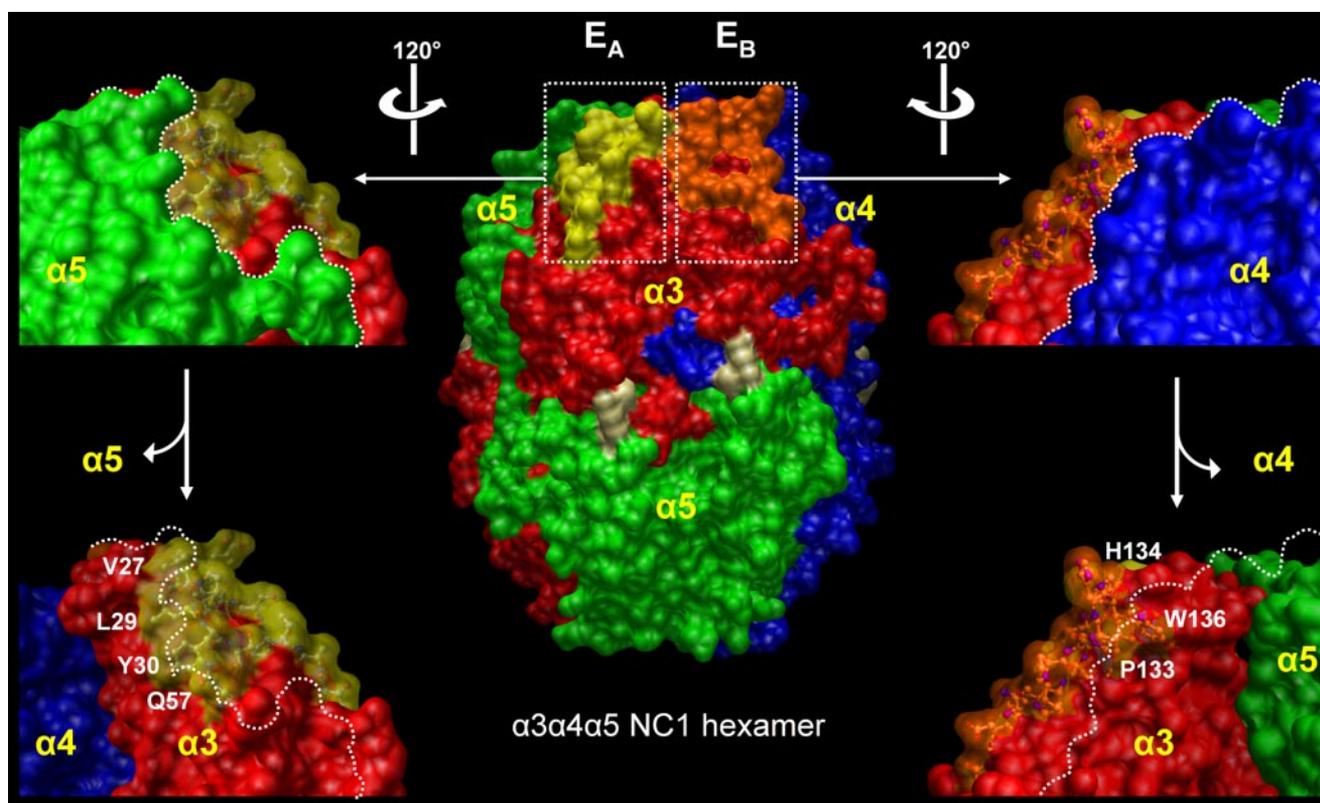


FIGURE 8. A three-dimensional model for the D-isomer of the $\alpha3\alpha4\alpha5$ NC1 hexamer showing the location and quaternary structure of the E_A and E_B epitopes. A solid surface representation (solvent exclusion) is shown for the three-dimensional atomic structure model for the $\alpha3\alpha4\alpha5$ D-hexamer (center). The model depicts the location of the immunodominant epitopes for Goodpasture antibodies, E_A (yellow) and E_B (orange), within the $\alpha3$ NC1 domain as well as the location of the S-hydroxylysyl-methionine and S-lysyl-methionine cross-links (khaki color). The left and right panels show portions of the $\alpha3$ - $\alpha5$ (left) and $\alpha5$ - $\alpha4$ (right) NC1 domain interfaces, which illustrate the structural setting of the E_A and E_B epitopes within the quaternary structure of the $\alpha3\alpha4\alpha5$ D-hexamer. In each case, the NC1 hexamer was rotated about 120° in the direction marked by the arrows for optimal view. Lower insets represent the same view except that the $\alpha5$ (left) or $\alpha4$ (right) NC1 monomers were removed leaving a footprint (dotted line) on $\alpha3$ NC1 monomer to visualize the residues that were sequestered by the neighboring $\alpha5$ or $\alpha4$ NC1 domain. Sequestered residues are labeled in white captions. Molecular surfaces images were produced using University of California, San Francisco, Chimera (28).

trimers interact through an extensive planar trimer-trimer interface to form an NC1 hexamer. Each monomer of the $\alpha3\alpha4\alpha5$ NC1 hexamer contains a large domain swapping β -hairpin element that associates with a complementary binding site on an adjacent NC1 monomer, a feature that is critical in the assembly of NC1 trimers (22, 42). Furthermore, the NC1 hexamer is stabilized by up to four Met-Hyl and two Met-Lys cross-links. These cross-links can be envisioned to function as “molecular fasteners” that reinforce the quaternary structure of NC1 hexamers, which in turn strengthen the connection between two adjoining triple-helical molecules of the $\alpha3\alpha4\alpha5$ collagen IV networks.

Location and Structural Features of E_A and E_B Epitopes—The three-dimensional model (see above) provides the framework for describing the location and tertiary and quaternary structures of the GP epitopes, as presented in Fig. 8. In previous studies, the residues that encompass the E_A and E_B epitopes on the $\alpha3$ NC1 domain were determined (8, 32, 43–46). The Molecular Dynamics-minimized homology model revealed that a number of these residues are located on the surface but that residues Val²⁷, Leu²⁹, Tyr³⁰, and Gln⁵⁷ are sequestered under the neighboring $\alpha5$ NC1 monomer at the bottom of a surface cleft, formed by the lateral interaction between $\alpha3$ and $\alpha5$ NC1 monomers (Fig. 8, left panel). The sequestration is sta-

bilized by domain swapping interactions between neighboring NC1 monomers. Of the 1164 \AA^2 accessible surface area of the E_A epitope, 130 \AA^2 (11%) is buried for these four residues. The supplemental Table 1 provides a summary of the change in accessible surface area for each of these epitope residues.

The E_B epitope, encompassing residues Thr¹²⁷ through Lys¹⁴¹, as defined in previous studies (8, 32, 43, 44), is located on a ridge of the $\alpha3$ NC1 monomer adjacent to the $\alpha4$ NC1 monomer. MD-minimized homology modeling revealed that a number of these residues are located on the surface but that residues Pro¹³³, His¹³⁴, and Trp¹³⁶ are sequestered by the neighboring $\alpha4$ NC1 monomer (Fig. 8, right panel). Of the 1,030 \AA^2 accessible surface area of the E_B epitope, 87 \AA^2 (9%) is buried for these three residues (supplemental Table 1).

Structural Basis for the Crypticity of E_A and E_B Epitopes of the D-isomer—The structural basis for the crypticity of epitopes of the D-isomer was determined by a comparison of the structural and immunochemical features of the M- and D-isomers, together with an analysis of the quaternary structure of the E_A and E_B epitopes within the $\alpha3\alpha4\alpha5$ NC1 hexamer (see below). As we described in a previous study (9), the epitopes of the M-isomer are accessible to GP autoantibodies, which disrupt interactions between neighboring subunits, whereas the epitopes of the D-isomer are inaccessible (cryptic) and the iso-

form is impenetrable by GP autoantibody. The structural feature that distinguished the M- and D-isoforms was the presence of NC1 dimer subunits in the latter, a feature that correlated with the crypticity of epitopes (9). As reported herein, the $\alpha 3\alpha 5$ -heterodimer and the $\alpha 4\alpha 4$ -homodimer are held together by the Met-Hyl and Met-Lys cross-links, respectively. Thus, the cross-links play a key role in the crypticity of epitopes, as envisaged below.

In the absence of cross-links, as in the M-isoform, certain epitope residues are accessible on the surface for interaction with GP autoantibodies. This interaction, in turn, leads to the displacement of neighboring subunits, exposing the sequestered residues (Fig. 8) and allowing binding of GP autoantibodies to all residues that constitute the epitope. However, the cross-links present in the D-isoform render it impenetrable to GP autoantibodies such that sequestered residues do not become exposed for binding of GP autoantibodies. Thus, the cross-links provide a belt-like structural reinforcement that increases the structural integrity of the $\alpha 3\alpha 4\alpha 5$ NC1 hexamer, which protects against displacement of neighboring NC1 subunits (Fig. 8). Furthermore, that the GP autoantibodies do not bind to the D-isoform indicates that there are key residues within the epitope that are inaccessible because they are sequestered by neighboring subunits. These residues are Val²⁷, Leu²⁹, Tyr³⁰, and Gln⁵⁷ for the E_A epitope and Pro¹³³, His¹³⁴, and Trp¹³⁶ for the E_B epitope, as identified in the three-dimensional model (Fig. 8). It follows that these same residues are also sequestered in the M-isoform, but in the absence of cross-links, they are accessible to GP autoantibody.

A Novel Molecular Mechanism for Conferring Immune Privilege to GP Epitopes, a Role for Collagen IV Cross-links—Immune privilege originally implied a location anatomically protected from attacks by the immune system, such as eye, brain, and testis (47). More recently, other cellular and molecular mechanisms have been described that confer immune privilege. These include cytokines that inhibit innate or adaptive immunity, systemic regulatory T cells, and local cytotoxic mechanisms producing apoptosis of effector T cells induced by local FasL-Fas interactions (48, 49). Here we describe a novel mechanism for conferring immune privilege at the level of quaternary structure to an extracellular matrix protein. This structural mechanism involves two key features as follows: (a) the sequestration of key residues of an epitope within quaternary structure, and (b) the further stabilization of the integrity of the quaternary structure by cross-links. These structural features may prevent the immune system from recognizing the otherwise pathogenic epitopes that reside in the $\alpha 3\alpha 4\alpha 5$ collagen IV network of certain basement membranes.

The proposed mechanism also suggests that perturbation of the quaternary structure of the $\alpha 3\alpha 4\alpha 5$ NC1 hexamer exposing the pathogenic epitopes could be a key factor in the etiology of Goodpasture disease. Perturbations could include a disruption of domain-swapping interactions, which leads to the exposure of sequestered residues, eliciting the production of pathogenic antibodies, and/or a disruption of cross-links, which render the NC1 hexamer more susceptible to disruption, exposing immunogenic residues. Several environmental factors, including cigarette smoking, exposure to organic solvents, reactive oxygen

species, viruses and bacteria, have been associated with the onset of Goodpasture disease (50–52). Such factors could play a role in the loss of structural immune privilege of the GP autoantigen.

Acknowledgment—The technical assistance of Mohamed Rafi is greatly appreciated.

REFERENCES

- Goodnow, C. C., Sprent, J., Fazekas de St Groth, B., and Vinuesa, C. G. (2005) *Nature* **435**, 590–597
- Nossal, G. J. (1991) *Eur. J. Biochem.* **202**, 729–737
- Sercarz, E. E. (2002) *Nat. Immunol.* **3**, 110–112
- Moudgil, K. D., and Sercarz, E. E. (2005) *Trends Immunol.* **26**, 355–359
- Hudson, B. G., Reeders, S. T., and Tryggvason, K. (1993) *J. Biol. Chem.* **268**, 26033–26036
- Hudson, B. G., Tryggvason, K., Sundaramoorthy, M., and Neilson, E. G. (2003) *N. Engl. J. Med.* **348**, 2543–2556
- Wieslander, J., Langeveld, J., Butkowsky, R., Jodkowski, M., Noelken, M., and Hudson, B. G. (1985) *J. Biol. Chem.* **260**, 8564–8570
- Borza, D. B., Netzer, K. O., Leinonen, A., Todd, P., Cervera, J., Saus, J., and Hudson, B. G. (2000) *J. Biol. Chem.* **275**, 6030–6037
- Borza, D. B., Bondar, O., Colon, S., Todd, P., Sado, Y., Neilson, E. G., and Hudson, B. G. (2005) *J. Biol. Chem.* **280**, 27147–27154
- Abbate, M., Kalluri, R., Corna, D., Yamaguchi, N., McCluskey, R. T., Hudson, B. G., Andres, G., Zoja, C., and Remuzzi, G. (1998) *Kidney Int.* **54**, 1550–1561
- Kalluri, R., Gattone, V. H., II, Noelken, M. E., and Hudson, B. G. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 6201–6205
- Wang, X. P., Fogo, A. B., Colon, S., Giannico, G., Abul-Ezz, S. R., Miner, J. H., and Borza, D. B. (2005) *J. Am. Soc. Nephrol.* **16**, 3563–3571
- Sado, Y., Boutaud, A., Kagawa, M., Naito, I., Ninomiya, Y., and Hudson, B. G. (1998) *Kidney Int.* **53**, 664–671
- Chazenbalk, G. D., Pichurin, P., Chen, C. R., Latrofa, F., Johnstone, A. P., McLachlan, S. M., and Rapoport, B. (2002) *J. Clin. Investig.* **110**, 209–217
- Chen, C. R., Pichurin, P., Nagayama, Y., Latrofa, F., Rapoport, B., and McLachlan, S. M. (2003) *J. Clin. Investig.* **111**, 1897–1904
- Giannakopoulos, B., Passam, F., Rahgozar, S., and Krilis, S. A. (2007) *Blood* **109**, 422–430
- de Laat, B., Derksen, R. H., van Lummel, M., Pennings, M. T., and de Groot, P. G. (2006) *Blood* **107**, 1916–1924
- Boutaud, A., Borza, D.-B., Bondar, O., Gunwar, S., Netzer, K.-O., Singh, N., Ninomiya, Y., Sado, Y., Noelken, M. E., and Hudson, B. G. (2000) *J. Biol. Chem.* **275**, 30716–30724
- Kahsai, T. Z., Enders, G. C., Gunwar, S., Brunmark, C., Wieslander, J., Kalluri, R., Zhou, J., Noelken, M. E., and Hudson, B. G. (1997) *J. Biol. Chem.* **272**, 17023–17032
- Ninomiya, Y., Kagawa, M., Iyama, K., Naito, I., Kishiro, Y., Seyer, J. M., Sugimoto, M., Oohashi, T., and Sado, Y. (1995) *J. Cell Biol.* **130**, 1219–1229
- Vanacore, R. M., Friedman, D. B., Ham, A. J., Sundaramoorthy, M., and Hudson, B. G. (2005) *J. Biol. Chem.* **280**, 29300–29310
- Sundaramoorthy, M., Meiyappan, M., Todd, P., and Hudson, B. G. (2002) *J. Biol. Chem.* **277**, 31142–31153
- Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., and Thompson, J. D. (2003) *Nucleic Acids Res.* **31**, 3497–3500
- Jones, T. A., and Kjeldgaard, M. (1997) *Methods Enzymol.* **277**, 173–208
- Humphrey, W., Dalke, A., and Schulten, K. (1996) *J. Mol. Graphics* **14**, 27–38
- Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) *J. Comput. Chem.* **26**, 1781–1802
- MacKerell, A. D., Jr., Banavali, N., and Foloppe, N. (2000) *Biopolymers* **56**, 257–265
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) *J. Comput. Chem.* **25**,

Collagen IV Cross-links and Immune Privilege

- 1605–1612
29. Feig, M., Karanicolas, J., and Brooks, C. L., III (2004) *J. Mol. Graphics Model* **22**, 377–395
30. Fraczekiewicz, R., and Braun, W. (1998) *J. Comput. Chem.* **19**, 319–333
31. Hubbard, S. J., and Thornton, J. M. (1993) *NACCESS Computer Program*, University College London, London, UK
32. Borza, D.-B., Bondar, O., Todd, P., Sundaramoorthy, M., Sado, Y., Nishimiyama, Y., and Hudson, B. G. (2002) *J. Biol. Chem.* **277**, 40075–40083
33. Siebold, B., Deutzmann, R., and Kuhn, K. (1988) *Eur. J. Biochem.* **176**, 617–624
34. Weber, S., Engel, J., Wiedemann, H., Glanville, R., and Timpl, R. (1984) *Eur. J. Biochem.* **139**, 401–410
35. Than, M. E., Henrich, S., Huber, R., Ries, A., Mann, K., Kuhn, K., Timpl, R., Bourenkov, G. P., Bartunik, H. D., and Bode, W. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, 6607–6612
36. Yates, J. R., III, Eng, J. K., and McCormack, A. L. (1995) *Anal. Chem.* **67**, 3202–3210
37. Kleppel, M. M., Fan, W., Cheong, H. I., and Michael, A. F. (1992) *J. Biol. Chem.* **267**, 4137–4142
38. Kivirikko, K. I., and Pihlajaniemi, T. (1998) *Adv. Enzymol. Relat. Areas Mol. Biol.* **72**, 325–398
39. Kivirikko, K. I., and Myllylä, R. (1984) in *Extracellular Matrix Biochemistry* (Piez, K. A., and Reddi, A. H., eds) pp. 83–118, Elsevier Science Publishing Co., Inc., New York
40. Ruotsalainen, H., Sipila, L., Vapola, M., Sormunen, R., Salo, A. M., Uitto, L., Mercer, D. K., Robins, S. P., Risteli, M., Aszodi, A., Fassler, R., and Myllylä, R. (2006) *J. Cell Sci.* **119**, 625–635
41. Rautavuoma, K., Takaluoma, K., Sormunen, R., Myllyharju, J., Kivirikko, K. I., and Soininen, R. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101**, 14120–14125
42. Khoshnoodi, J., Cartailier, J. P., Alvares, K., Veis, A., and Hudson, B. G. (2006) *J. Biol. Chem.* **281**, 38117–38121
43. Netzer, K. O., Leinonen, A., Boutaud, A., Borza, D. B., Todd, P., Gunwar, S., Langeveld, J. P., and Hudson, B. G. (1999) *J. Biol. Chem.* **274**, 11267–11274
44. David, M., Borza, D. B., Leinonen, A., Belmont, J. M., and Hudson, B. G. (2001) *J. Biol. Chem.* **276**, 6370–6377
45. Hellmark, T., Burkhardt, H., and Wieslander, J. (1999) *J. Biol. Chem.* **274**, 25862–25868
46. Hellmark, T., Segelmark, M., Unger, C., Burkhardt, H., Saus, J., and Wieslander, J. (1999) *Kidney Int.* **55**, 936–944
47. Medawar, P. B. (1948) *Br. J. Exp. Pathol.* **29**, 58–69
48. Griffith, T. S., Brunner, T., Fletcher, S. M., Green, D. R., and Ferguson, T. A. (1995) *Science* **270**, 1189–1192
49. Ferguson, T. A., and Griffith, T. S. (2006) *Immunol. Rev.* **213**, 228–238
50. Stevenson, A., Yaqoob, M., Mason, H., Pai, P., and Bell, G. M. (1995) *Q. J. Med.* **88**, 23–28
51. Donaghy, M., and Rees, A. J. (1983) *Lancet* **2**, 1390–1393
52. Kalluri, R., Cantley, L. G., Kerjaschki, D., and Neilson, E. G. (2000) *J. Biol. Chem.* **275**, 20027–20032
53. Robins, S. P., and Bailey, A. J. (1975) *Biochem. J.* **149**, 381–385