Protein Structure and Folding:
Identification of S-Hydroxylysyl-methionine as the Covalent Cross-link of the Noncollagenous (NC1) Hexamer of the α1α1α2 Collagen IV Network: A ROLE FOR THE POST-TRANSLATIONAL MODIFICATION OF LYSINE 211 TO HYDROXYLYSINE 211 IN HEXAMER ASSEMBLY

Roberto M. Vanacore, David B. Friedman, Amy-Joan L. Ham, Munirathinam Sundaramoorthy and Billy G. Hudson
doi: 10.1074/jbc.M502752200 originally published online June 10, 2005

Access the most updated version of this article at doi: 10.1074/jbc.M502752200

Find articles, minireviews, Reflections and Classics on similar topics on the JBC Affinity Sites.

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 25 references, 15 of which can be accessed free at http://www.jbc.org/content/280/32/29300.full.html#ref-list-1
Identification of S-Hydroxylysyl-methionine as the Covalent Cross-link of the Noncollagenous (NC1) Hexamer of the α1α1α2 Collagen IV Network

A ROLE FOR THE POST-TRANSLATIONAL MODIFICATION OF LYSINE 211 TO HYDROXYLISINE 211 IN HEXAMER ASSEMBLY*

Received for publication, March 14, 2005, and in revised form, May 16, 2005
Published, JBC Papers in Press, June 10, 2005, DOI 10.1074/jbc.M502752200

Collagen IV networks are present in all metazoa as components of basement membranes that underlie epithelia. They are assembled by the oligomerization of triple-helical protomers, composed of three α-chains. The trimeric noncollagenous domains (NC1) of each protomer interact forming a hexamer structure. Upon exposure to acidic pH or denaturants, the hexamer dissociates into monomers and dimer subunits, the latter reflect distinct interactions that reinforce/cross-link the quaternary structure of hexamer. Recently, the cross-link site of the α1α1α2 network was identified, on the basis of x-ray crystal structures at 1.9-Å resolution, in which the side chains of Met93 and Lys211 were proposed to be connected by a novel thioether bond (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); however, at the higher resolution of 1.5 Å, we found no evidence for this cross-link (Vanacore, R. M., Shanmugasundararaj, S., Friedman, D. B., Bondar, O., Hudson, B. G., and Sundaramoorthy, M. (2004) J. Biol. Chem. 279, 44723–44730). Given this discrepancy in crystallographic findings, we sought chemical evidence for the location and nature of the reinforcement/cross-link site. Trypsin digestion of monomer and dimer subunits excised a ~5,000-Da complex that distinguished dimers from monomers; the complex was characterized by mass spectrometry, Edman degradation, and amino acid composition analyses. The tryptic complex, composed of two peptides of 44 residues derived from two α1 NC1 monomers, contained Met93 and Lys211 post-translationally modified to hydroxylysine (Hyl211). Truncation of the tryptic complex with post-proline endopeptidase reduced its size to 14 residues to facilitate characterization by tandem mass spectrometry, which revealed a covalent linkage between Met93 and Hyl211. The novel cross-link, termed S-hydroxylysyl-methionine, reflects at least two post-translational events in its formation: the hydroxylation of Lys211 to Hyl211 within the NC1 domain during the biosynthesis of α-chains and the connection of Hyl211 to Met93 between the trimeric NC1 domains of two adjoining triple-helical protomers, reinforcing the stability of collagen IV networks.

Collagen IV networks are components of basement membranes that underlie epithelia, compartmentalize tissue, and influence cell behavior. The networks are assembled from a family of six polypeptide chains (α1–α6) that associate forming three subtypes of triple-helical protomers with distinct chain compositions: α1α1α2, α3α4α5, and α5α6α6 (1, 2). The protomers self-assemble by end-to-end associations in which the amino termini of four protomers associate tail-to-tail forming the 7 S domain, and the carboxyl termini of two protomers associate head-to-head through the noncollagenous (NC1) domains, forming dimers. At the interface of the head-to-head connection, the trimeric NC1 domains exist as a hexamer, a stable complex that can be excised by cleavage with collagenase for in vitro studies.

The NC1 domain plays a pivotal role in the assembly of the distinct collagen IV networks. In protomer assembly, the NC1 domains (monomers) of three chains interact, forming a NC1 trimer, to select and register chains for triple-helix formation. In the network assembly, the NC1 trimers of two protomers interact, forming a NC1 hexamer structure, to select and connect protomers. Upon exposure to acidic pH or denaturants, isolated NC1 hexamer dissociates into monomers and dimers, the latter reflecting the presence of cross-links that stabilize the trimer-trimer interface. The cross-links connect α1-like monomers (α1–α1, α1–α5, and α3–α5) and α2-like monomers (α2–α2, α2–α6, and α4–α4) (3, 4). For two decades, the reduc-

* This work was supported by National Institutes of Health Grants R-37DK18381 (to B. G. H.), PO1DK065123 (to B. G. H.), ROI1DK62524 (to M. S.), and RR017806 (to D. B. F). 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.

** Supported by Vanderbilt University through the Academic Venture Capital Fund.

†† To whom correspondence should be addressed: Dept. of Medicine, Division of Nephrology, Vanderbilt University Medical Center, B-3102 Medical Center North, 1161S Twenty-first Ave. South, Nashville, TN 37232-2372. Tel.: 615-322-7299; Fax: 615-322-7381; E-mail: Billy.hudson@vanderbilt.edu.

This paper is available online at http://www.jbc.org
ible dimers were thought to consist of monomers bound by disulfide cross-links (5, 6). However, the recent x-ray crystal structures of the NC1 hexamers of bovine lens capsule base-
membrane and human placenta basement membrane, determined independently by us (7) and Than et al. (8), respec-
tively, have disproved this hypothesis. An alternative ex-
aplanation was proposed by Than et al. (8) in which the cross-link is a thioether bond between Met<sup>93</sup> and Lys<sup>211</sup> that bridges the trimer-trimer interface; the evidence was based on electron density maps at 1.9-Å resolution, suggesting the existence of both cross-linked and noncross-linked residues at this site (8). However, in a subsequent study at the higher resolution of 1.5 Å, we found no evidence for this cross-link (9).

Given this discrepancy in crystallographic findings, we sought chemical evidence in the present study for the location and nature of the reinforcement/cross-link site. We used trypsin digestion in combination with mass spectrometry as a strategy to search for post-translational modifications that may go undetected by crystallography. The results revealed that the site is located at the trimer-trimer interface of the NC1 hexamer, characterized by a novel covalent cross-link: S-hydroxylysyl-
methionine. The cross-link is uniquely labile to conditions typically used for characterization, rendering it a challenge for detection. The findings are the first report of chemical evidence for the location and nature of the reinforcement/cross-link site, and the presence of Hyl within the NC1 domain.

**EXPERIMENTAL PROCEDURES**

**Materials**—Bovine placenta was purchased from Pel-Freeze Biologicals (Rodgers, AR). PBM NC1 hexamer was prepared by collagenase digestion as previously described (10). Bacterial collagenase (CLSPA) was purchased from Worthington (Lakewood, NJ).

**Separation of Monomers and Dimers**—Monomers and dimers of the NC1 domain were isolated as described elsewhere with minor modifications (10). Briefly, PBM hexamers (5 mg) were denatured in 0.2 M Tris-HCl, pH 8.5, buffer containing 4 M GdnHCl and 25 mM DTT and incubated in a boiling water bath for 20 min. Subsequently, the reduced proteins were alkylated with 50 mM iodoacetamide in the dark for 30 min at room temperature. To fractionate dimer and monomer subunits, the denatured hexamer sample was run through a Sephacryl S-300 column (150 × 2.5 cm) that had been equilibrated in 50 mM Tris-HCl, pH 7.8, for 3 h at 7° C. The products of the digestion were immediately analyzed by LC-ESI/MS (see below). To fractionate dimer and monomer subunits, the digested and alkylated samples were separately digested with trypsin, and the trypsin prod-
ucts were fractionated by size-exclusion chromatography on a Superdex peptide column as described under “Experimental Procedures.” The fractionation was monitored by measurement of absorbance at 220 nm for monomers (dashed line) and dimers (solid line). The asterisk indicates the retention time of a tryptic (T) complex(s) of ~5000 molecular weight that distinguishes dimers from monomers.

**LC-ESI/MS/MS and LC-ESI/MS<sup>3</sup>**—The LC-ESI/MS/MS and LC-ESI/MS<sup>3</sup> analyses were performed on a ThermoFinnigan LTQ linear ion trap mass spectrometer equipped with a ThermoFinnigan Surveyor LC pump, microlerespray source, and Xcalibur 1.4 instrument control and data analysis software. HPLC separation of the NC1 trypsin peptides was achieved with a C<sub>18</sub> capillary column at 0.7 ml min<sup>−1</sup> flow rate. Solvent A was H<sub>2</sub>O with 0.1% formic acid and solvent B was acetonitrile containing 0.1% formic acid. The gradient program was: 0–3 min, linear gradient 0–5% B; 3–5 min, 5% B, 50–50 min, linear gradient to 50% B; 50–52 min, linear gradient to 80% B; 52–55 min, linear gradient to 90% B; 55–56 min, 90% B in solvent A. A blank sample (0.1% formic acid) was run between the two analyses. Each sample was subjected to two LC-ESI/MS/MS analyses. In the first analysis, MS/MS spectra of the peptides were obtained using data-dependent scanning in which one full MS spectrum (mass range 400–2000 atomic mass units) was followed by three MS/MS spectra. In the second run, several specific precursor masses were selected for MS/MS analysis in a targeted fashion. For the LC-ESI/MS/MS analyses, the MS<sup>3</sup> spectra were obtained using data-dependent scanning in which one full MS spectrum is followed by one MS/MS spectrum. The three most intense ions in the MS/MS spectrum were selected for a third fragmentation (MS<sup>3</sup>). The MS<sup>3</sup> spectra were obtained using data-dependent scanning in which one full MS spectrum (mass range 400–2000 atomic mass units) was followed by three MS/MS spectra. In the second run, several specific precursor masses were selected for MS/MS analysis in a targeted fashion. For the LC-ESI/MS/MS analyses, the MS<sup>3</sup> spectra were obtained using data-dependent scanning in which one full MS spectrum is followed by one MS/MS spectrum. The three most intense ions in the MS/MS spectrum were selected for a third fragmentation (MS<sup>3</sup>). Software for Sequence and Post-translational Modification Analysis—Samples of trypsin-digested monomers and dimers were analyzed using LC-MS/MS. The data base search algorithm SEQUEST was used to identify peptides from the fragment ions recorded in the tandem mass spectrum. P-Mod (11), a statistics based algorithm, allowed for the successful identification of hydroxylation sites.

**Analytical RP-HPLC of the T<sub>xa</sub>-complex**—The T<sub>xa</sub>-complex peak was analyzed on an AKTA Purifier liquid chromatography system run by UNICORN 4.11 software (Amersham Biosciences). Runs were performed on a Supelcosil LC-318 reversed-phase analytical HPLC column.
The samples were loaded into the column that had been equilibrated with buffer A (95% water with 0.1% trifluoroacetic acid and 5% acetonitrile) with a flow rate of 1 ml/min. The peptides were eluted with a linear gradient up to 40% buffer B (0.085% trifluoroacetic acid, 95% acetonitrile and water) over 60 min. Peptide elution was monitored by absorbance at 215 nm.

Amino Acid Composition Analyses—Amino acid analysis was carried out on the T-5014 complex at the W.M. Keck Facility at Yale University (New Haven, CT) on a Beckman model 7300 ion-exchange instrument following a 16-h hydrolysis at 115 °C in 100 μl of 6 N HCl, containing 0.2% phenol. After hydrolysis, the HCl was dried in a SpeedVac, and the resulting amino acids dissolved in 100 μl of Beckman sample buffer. The instrument was calibrated with a 2-nmol mixture of amino acids and was operated via the manufacturer’s programs and with the use of internal fragmentation ions.

FIG. 2. MALDI-TOF MS analyses of T_{αβ}-complex reveals the presence of 2 component peptides (T-1414 and T-3600) and a post-translational modification of lysine to Hyl. Panel A, MALDI-TOF mass spectrum of the T_{αβ}-complex shown in Fig. 1. Arrows indicate ions at m/z 1414 and 3600 selected for MALDI-TOF/TOF tandem MS, the fragmentation spectra for which are shown in panels B and C, respectively, where b (↑) and y (↓) ions are denoted along with internal fragmentation ions.
their buffers. Data analysis was carried out on an external computer using PerkinElmer/Nelson data acquisition software.

**Edman Degradation Sequencing—NH2-terminal peptide sequencing of the T-5014 complex was carried out at the W.M. Keck Facility at Yale University on an Applied Biosystems Procise 494 cLC instrument equipped with on-line HPLC for the identification of the resulting phenylthiohydantoin-derivatives.**

**RESULTS**

**Identification of a Structural Difference between Dimers and Monomers of the NC1 Domain: Isolation of the Reinforcement Site by Trypsin Digestion—Trypsin digestion was used to identify structural differences between monomer and dimer subunits of the NC1 hexamer domain of collagen IV from bovine placenta basement membrane. To isolate monomers and dimers, the NC1 hexamers were reduced and alkylated in 4 M GdnHCl, then fractionated on a gel filtration column equilibrated with 4 M GdnHCl (Fig. 6, Ref. 9). The dimers and monomers were separately digested with trypsin under conditions for maximal cleavage. The peptide products were then fractionated over a size-exclusion chromatography column (Fig. 1). A comparison of the profiles for monomer (dotted) and dimer (solid) reveals that dimers yield an extra peak ($M_r 5000$), and therefore are designated as T5k-complex, indicative of a structural distinction between dimer and monomer subunits.

The chemical nature of the T5k-complex was characterized by mass spectrometry. Because of the large size of the T5k-complex, the MALDI-TOF MS instrument was optimized to detect ions in the $m/z$ range 500–8000. Surprisingly, the most intense ion in the spectrum was $m/z 1413.8$ and the highest mass observed was $m/z 3601.6$ (Fig. 2). The mass spectra were identical in linear or reflectron mode, which rules out a possible in-source fragmentation. Although the T5k-complex was not observed, the combined mass of T-1414 plus T-3600 peptides is ~5014, a value close to that observed by size-exclusion chromatography.

To identify the sequence of T-1414 and T-3600 peptides, they were analyzed by MALDI-TOF-TOF tandem MS. As shown in Fig. 2C, the fragmentation pattern of the T-1414 peptide is consistent with the sequence, 77NDYSYWLST-PEPMPSMAPITGENIRPFISR107, derived from the α1 NC1 domain. The mass of the T-1414 peptide was not predicted from the α1 NC1 domain sequence, therefore, it was considered to be a modified peptide. As shown in Fig. 2C, the fragmentation pattern of the T-1414 peptide is consistent with the sequence, 204KPTPSTLKAGELR216, but with 16 extra mass units attached to Lys211. It is well known that lysine residues within the collagenous domain of collagen IV are post-translationally modified by a lysyl hydroxylase catalyzed addition of a hydroxy group (12), converting lysine to 5-hydroxylysine (Hyl). Thus, the extra 16 mass units observed in the T-1414 peptide in comparison to the known primary sequence, 204KPTPSTLKAGELR216, but with 16 extra mass units attached to Lys211. It is well known that lysine residues within the collagenous domain of collagen IV are post-translationally modified by a lysyl hydroxylase catalyzed addition of a hydroxy group (12), converging lysine to 5-hydroxylysine (Hyl). Thus, the extra 16 mass units observed in the T-1414 peptide in comparison to the known primary sequence, suggests the presence of Hyl211 (confirmed below). Therefore, the results are consistent with a stable complex (T-5014), composed of T-1414 and T-3600 peptides, which dissociate before or during sample preparation for MALDI-TOF MS analyses.

To test the hypothesis that T-1414 and T-3600 peptides form a stable complex, we analyzed the T5k-complex, isolated by gel-filtration (Fig. 1), using LC-ESI/MS/MS. As shown in Fig. 3, the sample displayed two major peaks (Fig. 3A) each giving rise to a charge envelope containing $+3, +4, +5, +6$, and $+7$ ions in the mass analyzer (Fig. 3B). The ions had an experimental

![Fig. 3. T-1414 and T-3600 peptides exist as a stable binary complex (T-5014 and T-5030) as determined by LC-ESI/MS3 analysis. Panel A shows the total ion current trace for the chromatographic separation on a C18 stationary phase of the T5K-complex (Fig. 1). Panel B, full scan ESI/MS/MS spectra reveals that T5k-complex display two main components: complexes T-5014 and T-5030. Multiply charged ions of the peptide complexes are indicated above the corresponding peaks.](http://www.jbc.org/)

**FIG. 3.** T-1414 and T-3600 peptides exist as a stable binary complex (T-5014 and T-5030) as determined by LC-ESI/MS3 analysis. Panel A shows the total ion current trace for the chromatographic separation on a C18 stationary phase of the T5K-complex (Fig. 1). Panel B, full scan ESI/MS/MS spectra reveals that T5k-complex display two main components: complexes T-5014 and T-5030. Multiply charged ions of the peptide complexes are indicated above the corresponding peaks.
average peptide mass of 5014.0 and 5030.1 and were named T-5014 and T-5030 complexes, respectively. The mass of the T-5014 complex is equal to the sum of the masses of the ions detected by MALDI-TOF MS, T-1414 and T-3600 peptides (Fig. 5). Similarly, the T-5030 complex is equal to the mass of T-1414 and T-3600 peptides plus 16 mass units.

The T-5014 and T-5030 complexes were further characterized by MALDI-TOF MS analyses. The T5k-complex (Fig. 1) was fractionated by reversed-phase HPLC yielding two major components that corresponded to T-5014 and T-5030 complexes (Fig. 4A). MALDI-TOF MS analyses (Fig. 4B) showed that the T-5014 complex is composed of T-1414 and T-3600 peptides, and that the T-5030 complex is composed of T-1414 and T-3616 peptides. The only difference between T-5014 and T-5030 complexes is the oxidation of methionine in the latter, which may have occurred during chromatography. Thus, the apparent masses of the complexes equal the sum of the masses of the constituent peptides, indicating that interaction of the peptides occurs without a change in mass (non-covalent), or a change in mass (covalent linkage) that is within the experimental error of the measurement for the 5,000 mass range. Moreover, the results establish that the T-5014 and T-5030 complexes dissociate under the conditions of MALDI-TOF MS analyses, a finding that explains why they were not observed in the earlier report (9) that used this method. The presence of fragments T-1460 and T-3554, which are the result of an alternative fragmentation of the tryptic complexes, are indicated in each spectrum (Figs. 1 and 4) and will be addressed below.

The T-5014 complex, isolated by HPLC (Fig. 4A), was also characterized by conventional amino acid analysis and Edman degradation. The amino acid composition (Table I) revealed a very close correlation between the experimental and theoretical values based on the sequences of the composite T-1414 and T-3600 peptides, identified by mass spectrometry, including one residue of Hyl in the T-1414 peptide. Amino acid analyses of monomer and dimer subunits also revealed the presence of one Hyl residue in monomer and two residues in the dimer (data not shown). Twenty cycles of Edman degradation revealed the presence of two peptides with sequences of \( N\text{DYSYWLSTPEPMPMSMAPITGENIR} \) and \( 204\text{KPTPSTLK}_{197}\text{AGELR}^{216} \), which correspond to T-3600 and T-1414 peptides (Fig. 5). The latter peptide contained Hyl at the eighth position, corresponding to Hyl\(^{211} \), as suggested by the above mass spectrometry studies. Furthermore, mass spectrometry and Edman degradation analyses indicate that this Hyl is not glycosylated. Moreover, that the T-5014 complex sequenced through 20 cycles without interruption at Met\(^{93} \) or Hyl\(^{211} \) indicates that the interaction between T-1414 and T-3600 peptides is labile under the conditions of Edman degradation.
Evidence for a Covalent Cross-link in the Tryptic Complex (T-5014)—The chemical nature of the interactions between T-1414 and T-3600 peptides in the T-5014 complex was explored by LC-ESI/MS$^3$ mass spectrometry, using collision-induced dissociation for fragmentation. The fragmentation generated two modified peptides: T-1460 and T-3554 that differed in mass from the T-1414 and T-3600 peptides observed by MALDI-TOF (see above). Although the modified peptides (T-1460 and T-3554) were much less intense than T-1414 and T-3600 peptides, they were also observed in the MALDI-TOF MS spectra (Figs. 1 and 4). These results suggested that a chemical group of ~46 was transferred from the T-3600 to T-1414 peptide, possibly indicating a covalent cross-link between the two peptides. However, the location of mass changes within the two peptides could not be interpreted with confidence because of the complexity of MS$^3$ spectra because of the large masses of the peptides. To circumvent this problem, the T$_{5k}$-complex isolated in Fig. 1 was digested with a second protease to reduce its size for characterization by LC-ESI/MS$^3$ mass spectrometry.

The large number of prolyl residues in the T$_{5k}$-complex suggested the use of post-proline endopeptidase for truncation. This enzyme cleaves the peptide bond on the carboxyl side of prolyl residues (13), and it would reduce the T-5014 complex of 44 residues down to 14 residues, as depicted in Fig. 5. The truncated product would be composed of MSMAP (535.2) and STLK$_{35}$AGELR (989.6) peptides, which together exist as a complex with a theoretical mass of 1524.8 (designated as the P-1525 complex). The post-proline endopeptidase digestion product was analyzed by LC-ESI/MS$^3$. Fig. 6a (left panel) presents a full MS spectrum showing two major ions at m/z 509.1 and 762.7 corresponding to the triply and doubly charged forms of the P-1525 complex, respectively. The experimental average mass of the P-1525 complex is 1523.9, approximately 1 mass unit less than the theoretical mass of 1524.8 (Fig. 5), indicating that interaction of the two constituent peptides of the P-1525 complex results in the loss of a single hydrogen.

Further MS/MS analyses of the doubly charged ion of m/z 762.6 revealed ions of m/z 488.3 and 1036.6 (Fig. 6a, right panel). These ions differ from the theoretical ones of 536.2 (535.2 + H$^+$) and 990.6 (989.6 + H$^+$), calculated for the two constituent peptides of the P-1525 complex (Fig. 5), by values of ~48 and +46 mass units, respectively. To obtain structural information about each of these ions and the nature of the 48 and 46 masses, each ion was submitted to a third collision-induced dissociation fragmentation (MS$^3$) in the instrument. In Fig. 6b, the MS$^3$ spectrum at the m/z 488.3 fragment is shown; the fragmentation profile is consistent with the MSMAP peptide sequence, except that the b and y ion series demonstrate the loss of 48 atomic mass units of Met$^{95}$. In contrast, in Fig. 6c, the fragmentation profile of the ion at m/z 1036.6 is consistent with the SKLK$_{35}$AGELR, but the y and b ion series demonstrate that Hyl$_{211}$ gained 46 atomic mass units of fragmentation of the P-1525 complex. These fragmentation profiles indicate a loss of a CH$_3$S$^-$ group along with a proton (totaling 48 atomic mass units) from Met$^{95}$ and a gain of a CH$_3$S$^-$ group onto Hyl$_{211}$ and the loss of a proton (totaling 46 atomic mass units).

Overall, the MS results provide compelling evidence that the two peptides of the P-1525 complex are connected by a covalent bond between the side chains of Met$^{95}$ and Hyl$_{211}$, and that its formation is concomitant with the loss of 1 mass unit. A structure for the cross-link is proposed in Fig. 6d, in which the S atom of Met$^{95}$ is covalently linked to the C$^*$ atom of Hyl$_{211}$, forming a sulfonium linkage. Alternatively, the S atom could be attached directly to the N atom on C$^*$, to C$^5$, or the O atom on C$^5$, either of which are consistent with the loss of the CH$_3$S$^-$ group from Met$^{95}$ and the gain of this group onto Hyl$_{211}$ upon collision-induced dissociation fragmentation. Based on the x-ray structure of the NC1 hexamer, Than et al. (8) proposed a thioether bond that linked S$^5$ of Met$^{93}$ and C$^*$ of Lys$_{211}$, with the concomitant loss of the methyl group of Met$^{93}$ and ε-amino group of Lys$_{211}$; this structure is incompatible with all the mass spectrometry results reported herein, and the presence of Hyl$_{211}$.

**Conditions for Breakage of the Cross-link between Monomers of the NC1 Dimer**—The labile nature of the T-5014 complex during the MALDI-TOF MS analyses prompted us to explore the conditions for breaking the cross-link(s) between NC1 monomers of the PBM dimer. In our previous work (9), we

### Table I

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Theoretical</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>2</td>
<td>2.1</td>
</tr>
<tr>
<td>Arg</td>
<td>3</td>
<td>2.9</td>
</tr>
<tr>
<td>Asn + Asp</td>
<td>3</td>
<td>3.4</td>
</tr>
<tr>
<td>Asn</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Cya</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gln + Ghu</td>
<td>3</td>
<td>3.6</td>
</tr>
<tr>
<td>Gln</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>2</td>
<td>3.1</td>
</tr>
<tr>
<td>His</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ile</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Leu</td>
<td>3</td>
<td>3.2</td>
</tr>
<tr>
<td>Lys</td>
<td>2</td>
<td>1.1</td>
</tr>
<tr>
<td>Hyl</td>
<td>0</td>
<td>0.9</td>
</tr>
<tr>
<td>Met</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Phe</td>
<td>1</td>
<td>0.7</td>
</tr>
<tr>
<td>Pro</td>
<td>7</td>
<td>7.3</td>
</tr>
<tr>
<td>Ser</td>
<td>5</td>
<td>4.7</td>
</tr>
<tr>
<td>Thr</td>
<td>4</td>
<td>3.7</td>
</tr>
<tr>
<td>Trp</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Val</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

![Fig. 5. Primary structure and location of the T-5014 and P-1525 complexes in the α1 NC1 domain dimer.](http://www.jbc.org/content/29305/12/29305/full)
FIG. 6. LC-ESI/MS³ analyses of the P-1525 complex demonstrate that His⁹¹¹ and Met⁹³⁹ are covalently cross-linked. The T-5014 complex was digested with post-proline endopeptidase (Fig. 5) and analyzed by LC-ESI/MS³. A (left) shows the ESI mass spectra of the P-1525 complex constituted by MSMAP and STLKOHAGELR peptides. The observed m/z values of the doubly [M + 2H]²⁺ and triply [M + 3H]³⁺ charged ions give an experimental average peptide mass of 1523.9, which is approximately 1 atomic mass unit below the theoretical mass for the P-1525 complex (1524.8). A (right) shows the MS³ mass spectra of the doubly charged ion at m/z 762.7. The loss of 48 atomic mass units by the MSMAP peptide (m/z 536.2) and the concomitant gain of 46 atomic mass units by the STLKOHAGELR peptide (m/z 990.6) are indicated in the figure between the squared brackets. B shows that the MS³ spectrum of the ion at m/z 488.3 is consistent with the sequence MSM*AP, where M* indicates the cross-link of Type IV Collagen NC1 Domain.

**B** shows that the MS³ spectrum of the ion at m/z 488.3 is consistent with the sequence MSM*AP, where M* indicates the cross-link of Type IV Collagen NC1 Domain.

**D** shows the precursors, crosslink, and fragments. The precursors are MSXAP and STLXAGELR peptides, and the crosslink is indicated by the reaction of H⁺ with the sulffhydryl group of the MSMAP peptide. The fragments include [M-46+H⁺] and [M+46+H⁺] ions, with masses 1523.9 and 1035.6, respectively.

**Peptide mass** 1524.8  1523.9  487.3  1035.6

**Crosslink**

**Fragments**

**(- CH₃SH)**

**(+ CH₃S - H⁺)**

**MSMAP**

**MSXAP**

**STLXAGELR**

**X= -NH-CH-C-**

**1522.9**
Fig. 7. The cross-link that connects monomers forming dimer subunits is labile to DTT in the presence of protein denaturants. Panel A, the monomer and dimer ratio was determined by size-exclusion HPLC on a TSK SW3000 HPLC column, equilibrated with 4 M GdnHCl. PBM hexamers were reduced in 0.2 M Tris-HCl, pH 8.5, containing 4 M GdnHCl and 25 mM DTT by the indicated time and temperature. The samples were alkylated with 50 mM iodoacetamide and incubated in the dark for 30 min before injection. Panel B, SDS-PAGE analyses: time course experiment showing the monomer (M) to dimer (D) ratio after incubating PBM hexamers at 80 °C for 0, 15, 30, 45, 60, 90, and 120 min (lanes 2–8, respectively) in Laemmli buffer (2% SDS), which included 100 mM DTT. A control sample (lane 1) where the hexamer was heated for 2 h at 80 °C without DTT did not significantly alter the dimer:monomer ratio, showing that the conversion of dimers into monomers occurs only in the presence of the reducing agent.

established that PBM NC1 hexamers dissociate into dimer and monomer subunits (80/20 ratio) under denaturing conditions by gel-filtration chromatography (4 M GdnHCl and pH 3.0) and SDS-PAGE. In the present study, we extended those studies to investigate the effect of DTT and denaturant as a function of temperature and incubation time on hexamer dissociation. Fig. 7A shows the gel-filtration dissociation profiles of hexamer treated with DTT in GdnHCl under different incubation time and temperature conditions. The dimer:monomer ratio decreases to 50:50 after incubation of hexamers at 100 °C for 1 min in the presence of DTT and GdnHCl (panel a). The amount of dimer decreases even more, a change of ratio from 80:20 to 30:70 when the incubation at 100 °C is extended to 20 min (panel b). Similarly, incubation of hexamers for 80 min at 56 °C changes the dimer:monomer ratio to 10:90 (panel c), while extending the incubation time to 16 h, in the same conditions, achieves complete conversion of dimers into monomers (panel d). These experiments demonstrate that DTT in the presence of denaturant converts dimers into monomers, breaking the cross-link that hold dimers together.

The dissociation of PBM hexamer was also studied by SDS-PAGE with the samples treated with DTT and SDS. Fig. 7B shows a decrease in mobility of the monomer and dimer bands as expected by the reduction of intramolecular disulfide bonds (lanes 1 and 2). In addition, a gradual conversion of dimers into monomers (lanes 2–8) as a function of incubation time is clearly demonstrated. Importantly, incubation of hexamers for 2 h at 80 °C in the absence of DTT (lane 1) did not change the 80:20 dimer:monomer ratio. Thus, DTT appears to exert two effects: the reduction of disulfide bonds and the breakage of the Met-Hyl cross-link.

Mass spectrometry analysis of the breakage of dimers into monomers can also provide information about the location and nature of the cross-link. In the dimer, the Hyl(211)–Ala(212) bond is resistant to trypsin cleavage, resulting in the excision of the T-1414 peptide with an intact Hyl211–Ala212 bond (Fig. 5); such bonds are known to be susceptible to trypsin cleavage except when Hyl is modified by a carbohydrate unit (14). This poses the question of whether the Hyl211–Ala212 bond is resistant or susceptible to cleavage when present in the monomer, and the questions of whether the Hyl211–Ala212 bond is resistant or susceptible to cleavage when present in the monomer, and

loss of 48 atomic mass units from Met211. The asterisk (*) above the b(1) and y(1) ions indicates the loss of 48 atomic mass units from the expected mass value of the unmodified fragments. C shows that the MS3 spectrum of the ion at m/z 1036.6 is consistent with the sequence STLK*OHAGELR, where K*OH represents the gain of 46 atomic mass units by the Hyl211 residue. The asterisk (*) above b(1) and y(1) ions indicates the gain of 46 atomic mass units, the predicted mass value of the unmodified fragments containing the Hyl211 residue. D presents a proposed structure for the covalent linkage between the peptides containing Hyl211 and Met212 based on the mass spectrometry results. The S atom of Met212 is covalently linked to the C* atom of Hyl211 forming a sulfonium ion. Alternatively, the S atom could be attached directly to the N atom on C*, C2, or the O atom on C* of Hyl, either of which are consistent with all the LC-ESI/MS3 results. Note that the difference in mass between the precursors (1525) and the cross-linked peptides (1524) is 1 mass unit, represented by the loss of H2O in each alternative. The novel cross-link is termed S-hydroxylysyl-methionine to denote the linkage via the sulfur atom. The structures for the fragments generated by collision-induced dissociation of the P-1525 complex are shown at the right of the cross-link. A thioether bond was previously proposed by Than et al. (8) that linked S* of Met211 and C* of Lys211, but with the concomitant loss of methyl group of Met211 and ε- amino group of Lys211, a structure that is incompatible with all the mass spectrometry results reported herein.
whether it becomes susceptible after breakage of dimer into monomer by treatment with DTT. Thus, the status of the Hyl211–Ala212 bond was explored in two kinds of monomers: monomers that were derived from hexamers and fractionated by gel-filtration chromatography (Fig. 7A, panel b), and monomers that were derived from the dimers (Fig. 7A, panel b) followed by treatment with 2% SDS plus 100 mM DTT and incubated for 2 h at 80 °C.

This objective was accomplished using SDS-PAGE to separate dimers and monomers, in-gel digestion of components with trypsin to release peptides, and LC-ESI/MS/MS analyses of tryptic peptides. The data sets were interrogated for the presence of the T-1414 peptide, which contains the intact Hyl211–Ala212 bond, and T-887 peptide, which corresponds in mass to a peptide with the sequence KPTPSLKOH211 and that indicates cleavage of the Hyl211–Ala212 bond (Fig. 8). The T-1414 peptide was present in the dimer (lane 1), consistent with previous MS analyses, but not in the monomer (lane 2). Conversely, the T-887 peptide was present in the monomer but not in the dimer. The T-887 peptide was also present in monomer (lane 3) that had been treated at elevated temperatures, but with DTT. Moreover, the monomer that was derived by DTT treatment of dimer (lane 4) also revealed the presence of the T-887 peptide, but the complete absence of the T-1414 peptide. These results indicate that the Hyl211–Ala212 bond is resistant to trypsin cleavage in the dimer subunit, but susceptible to cleavage in the monomer subunit, and that the resistance is broken by treatment with DTT. The findings are consistent with a covalent cross-link involving Hyl211 rendering the Hyl211–Ala212 bond resistant to trypsin cleavage. Moreover, the presence of Hyl211 in both M- and D-hexamers suggests that the post-translational modification of Lys211 in monomers is necessary but not sufficient for cross-linking of monomers to form dimers.

**DISCUSSION**

In a recent study (9), we presented evidence for the existence of two distinct kinds of hexamers, M102-hexamers composed exclusively of monomers, and D102-hexamers composed exclusively of dimers. The two extremes reflect a process that reinforces/cross-links the interaction of monomer subunits forming D-hexamers. The proportion of M- and reinforced D-hexamers indicates that the collagen IV network of PBM is a more stable structure than that of the lens capsule basement membrane, a feature that may be an important determinant of biological function. The reinforcement is not unique to the α1α2(IV) network, but also occurs in the α3α4α5(IV) and α1α2-α5α6(IV) networks (1, 3, 4).

In the present study, the location and nature of the reinforcement/cross-link site of dimers of D102-hexamers were investigated using trypsin digestion as a strategy to excise the site and using mass spectrometry for characterization. Fortuitously, a structural feature distinguishing dimers from monomer subunits could be excised as a low molecular weight complex that was easily purified for chemical and physical characterization. The tryptic complex is composed of two short peptides, comprising residues 77–107 and 204–216 of an α1-NC1 monomer. The two peptides, respectively, contain Met93 and Lys211 post-translationally modified to Hyly211 and they correspond to one of the regions that are in close proximity at the trimer-trimer interface of the NC1 hexamer, as defined previously from the crystal structure (7–9). That the complex is derived exclusively from dimer subunits of D102-hexamers, and its primary structure corresponds to a region that connects the trimer-trimer interface, provides the first chemical evidence for the location of the reinforcement/cross-link site (Fig. 9). Its location is in agreement with that proposed by Than et al. (8) based on connectivity observed by x-ray crystallography, but differs with respect to the presence of Hyly211 instead of Lys211.

The chemical nature of the reinforcement/cross-link was elucidated by mass spectrometry of a truncated form of the tryptic complex. A second digestion with post-proline endopeptidase yielded the same T-1414 peptide, suggesting that its formation is concomitant with the loss of 1 mass unit, relative to the masses of unmodified Met93 and Hyly211, and that its formation is concomitant with the loss of 1 mass unit, relative to the masses of unmodified Met93 and Hyly211. The proposed structure, shown in Figs. 6d and 9, is a sulfonium ion in which the S atom of Met93 is covalently linked to the C atom of Hyly211. Alternatively, the S atom could be attached directly to the N atom on C211, or the O atom on C211 of Hyl, either of which are consistent with all the LC-ESI/MS3 results. The sulfonium ion linkage is consistent with the susceptibility of the cross-link to cleavage by DTT (Figs. 7–9), because methionine sulfonium ions undergo cleavage by sulphydryl agents (15). Of particular note, the linkage is also susceptible to cleavage by the conditions of Edman degradation and MALDI-TOF analysis; the latter explains why the cross-link was not observed by MALDI-TOF analysis of dimer subunits in our previous study. Several amino acid residues, including methionine, are susceptible to radiation damage by synchrotron x-rays (16). The breakage in methionine is known to occur at the CH3S– group, and its transfer to Hyl, provided compelling evidence that the two peptides of the smaller complex are connected by a covalent bond between the side chains of Met93 and Hyly211.

The novel cross-link is termed S-hydroxylysyl-methionine to denote a sulfur atom connection between Hyl and Met. At least two events occur in the formation of the cross-link: the post-translational hydroxylation of Lys211 to Hyly211 within the NC1 domain during the biosynthesis of α-chains; and the connection of Hyly211 to Met93 between the trimeric NC1 domains. The cross-link connects two adjoining triple-helical proomers, reinforcing the stability of collagen IV networks.

The presence of hydroxylysine 211 in the NC1 domain is a
novel feature. Hyl is a post-translational modification of lysine residues that typically occurs in X-Lys-Gly triplets in the collagenous domain of various types of collagens and collagen-like proteins (12). Hyl was previously noted in the preparation of NC1 domains (17), but it was thought to be a contaminant from the collagenase digestion of the collagenous domain or a residue in the two Gly-X-Y triplets at the NH2 terminus of the isolated NC1 domains (Fig. 5, Ref. 18). In certain cases, hydroxylation occurs in nonhelical peptides with sequences of X-Lys-Ala(Ser) at the end of collagen I chains that are involved in aldehyde-derived cross-links (19). This consensus sequence is identical to the X-Lys211-Ala sequence in the T-1414 peptide. Hyl residues have two important functions. They are essential for the stability of the intermolecular collagen cross-links, and their hydroxyl groups serve as attachment sites for the monosaccharide galactose or the disaccharide glucosyl-galactose (20–22). The carbohydrate units influence the lateral packing of fibril-forming collagen molecules into fibrils and may facilitate the assembly of the 7 S domain of collagen IV (23). In a recent study, lysyl hydrolysase-3 was shown to be essential for the assembly of collagen IV networks, presumably because of the absence of the hydroxylysine-linked carbohydrates (24). Conceivably, such carbohydrate units attached to Hyl211 might prevent the self-assembly of collagen IV pro- 

FIG. 9. Summary of findings regarding the location and nature of the reinforcement/cross-link site. The top panel shows the space-filling model for the reinforced/cross-linked D-hexamer (left) comprised of two trimeric caps, each composed of two α1 monomers and one α2 monomer. The juxtaposition of two opposing α1 NC1 monomers at the hexamer interface generates two sites of reinforcement: site 1 and site 2, which are defined by the tryptic 5014 complex composed of T-3600 (red) and T-1414 (green) peptides. The side chain of Met93 is colored in gold, and the side chain of Hyl211 is colored in cyan. In site 2, the peptide (not the side chains) colors are reversed. Under denaturing conditions, the D-hexamer dissociates into dimer subunits (middle). The dimer upon trypsin digestion releases T-5014 complex (right), composed of the T-3600 peptide (31 residues, red) and T-1414 peptide (13 residues, green), which constitutes the reinforcement site. The T-1414 peptide contains a post-translational modification in which Lys211 is converted to Hyl211. For mass spectrometry analyses to determine the chemical nature of the cross-link, the T-5014 complex was truncated by digestion with post-proline endopeptidase (PPE), which generated the P-1525 complex composed of the MSMAP peptide covalently cross-linked to the STLKOHAGELR peptide. The side chains of Met93 and Hyl211 are covalently connected by a sulphonium ion linkage between S of Met93 and C of Hyl211, as presented in Fig. 6d. A magnification of the proposed S-hydroxylysyl-methionine cross-link is shown. The bottom panel illustrates the non-reinforced/non-cross-linked M-hexamers (left), which dissociates into NC1 monomers (middle) upon denaturation. In this case trypsin digestion generates T-3600 and T-887 peptides that are not cross-linked in the hexamer. T-887 peptide, derived from the monomers, is a truncated version of T-1414 peptide derived from dimers. The structural relationship of these peptides indicates that in the dimer the Hyl211-Ala212 bond is blocked to trypsin cleavage, but in the monomer this bond is susceptible to cleavage. This difference in susceptibility to trypsin cleavage provides independent support for the location of a cross-link at Hyl211, and the difference is consistent with a covalent linkage between Met93 and Hyl211. The presence of Hyl211 in both M- and D-hexamers suggests that the post-translational modification of Lyl211 in monomers is necessary but not sufficient for cross-linking of monomers to form dimers.
Cross-link of Type IV Collagen NC1 Domain

mers form α1-α1, α1-α5, and α3-α5 dimers, and the α2-like monomers form α2-α2, α2-α6, and α4-α4 dimers (3, 4), indicating the presence of a cross-link. As noted previously by Than et al. (8), the sequences encompassing Met93 and Lys211 are invariant among all six human α-chains, suggesting that the same post-translational modifications could occur in each chain. Of particular importance, the reinforcement/cross-link of the α3α4α5 network was recently found to sequester B-cell epitopes within the NC1 hexamer, rendering them inaccessible to pathogenic autoantibodies in patients with Goodpasture syndrome. Thus, the cross-link represents a novel molecular mechanism for establishing immune privilege (25).

Acknowledgments—We thank Dr. Ned Porter, Dr. Terry Lybrand, Dr. Daniel Liebler, and Dr. Paul Voziyan from Vanderbilt University, and Dr. Gary Gray from the University of Minnesota for helpful discussions. We also thank Dr. Myron Crawford, director of W.M. Keck Foundation Biotechnology Resource Laboratory at Yale University, for protein sequencing and amino acid analyses. The excellent technical assistance of Parvin Todd is gratefully acknowledged.

REFERENCES