ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

Molecular Dynamics Investigation on the Effects of Protonation and Lysyl Hydroxylation on Sulfilimine Cross-links in Collagen IV

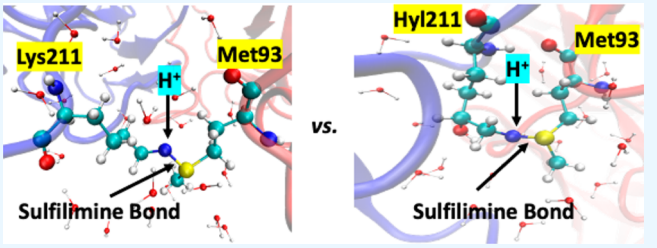
Anupom Roy and James W. Gauld*

Cite This: *ACS Omega* 2022, 7, 39680−39689

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Collagen IV networks are an essential component of basement membranes that are important for their structural integrity and thus that of an organism's tissues. Improper functioning of these networks has been associated with several diseases. Cross-links, such as sulfilimine bonds interconnecting NC1 domains, are critical for forming and mechanically stabilizing these collagen IV networks. More specifically, the sulfilimine cross-links form between methionine (Met93) and lysine/hydroxylsine (Lys211/Hyl211) residues of NC1 domains. Therefore, the dynamic nature of the sulfilimine bond in collagen IV is crucial for network formation. To understand the dynamic nature of a neutral and protonated sulfilimine bond in collagen IV, we performed molecular dynamics (MD) simulations on four sulfilimine cross-linked systems (i.e., $_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$) of collagen IV. The MD results showed that the neutral $_{Met93}S-N_{Lys211}$ system has the smallest protein backbone and showed the cross-linked residues' RMSD value. The conformational change analyses showed that the conformations of the sulfilimine cross-linked residues take on a U-shape for the $_{Met93}S-N_{Hyl211}$ and $_{Met93}S-HN_{Hyl211}^{+}$ systems, whereas the conformations of the sulfilimine cross-linked residues are more open for the $_{Met93}S-N_{Lys211}$, and $_{Met93}S-NH_{Lys211}^{+}$ systems. Protonation is a crucial biochemical process to stabilize the protein structure or the biological cross-links. Furthermore, the protonation of the sulfilimine bond could potentially influence hydrogen bond interaction with near amino acid residues, and according to water distribution analyses, the sulfilimine bond can potentially exist in one or more protonation states.



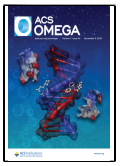
INTRODUCTION

Collagens are abundant structural proteins of the extracellular matrix found, for instance, in the connective tissues of the human body.[1] There are 28 types of collagen that have been identified and have various roles such as mechanical properties, organization, and shape of tissues in the human body.[2] Collagen has triple-helical domain structural features that contain multiple chains including isoforms of chains. As the part of the mechanical properties of collagen, during fibril formation, some collagens such as collagen III, IV, VI, and XVI also form covalent cross-links.[3]

In particular, collagen IV (also known as type IV) is a nonfibrillar collagen primarily found in the basal lamina of the basement membrane.[4] Human collagen IV has six homologous genes, COL4A1−COL4A6.[5] Each collagen IV consists of three structural domains, such as the N-terminal 7S domain, which is a short domain that contains ∼25 amino acids, a long middle triple-helical domain that contains ∼1400 amino acids, and a C-terminal (NC1) domain that contains ∼230 amino acids.[6] The NC1 domains are the noncollagenous globular structures that differ from connective tissue fibrillar collagens.[4] Self-association of any three chains of six homologous chains via their NC1 domains results in the formation of a protomer. The NC1 domain of the protomer then associates with the NC1 domain of another protomer in a head-to-head orientation to form a collagenous network.[7] Likewise, the other end of the protomer, the 7S domain, associated with the 7S domain of another protomer. As NCI and 7S domains are critical for a collagenous network, these NC1 to NC1 and 7S to 7S domain associations help to form a network that provides the structural integrity of collagen IV.

Type IV collagen is a major component of the basement membrane and is found in every tissue, especially in the skin.[8] It is also a major component of the dermo-epidermal junction, where it is thought to be important in the skin aging process.[9,10] In addition, it is crucial in a range of physiological processes[11] including cell adhesion, migration, differentiation,[12] tissue regeneration,[13] embryogenesis,[8] and would healing.[8] Furthermore, it has been shown to also regulate angiogenesis for the growth and remodeling of new vessels,[14] Due to these varied and life-crucial roles, even minor structural

Received: May 30, 2022
Accepted: October 13, 2022
Published: October 26, 2022



https://doi.org/10.1021/acsomega.2c03360
ACS Omega 2022, 7, 39680−39689

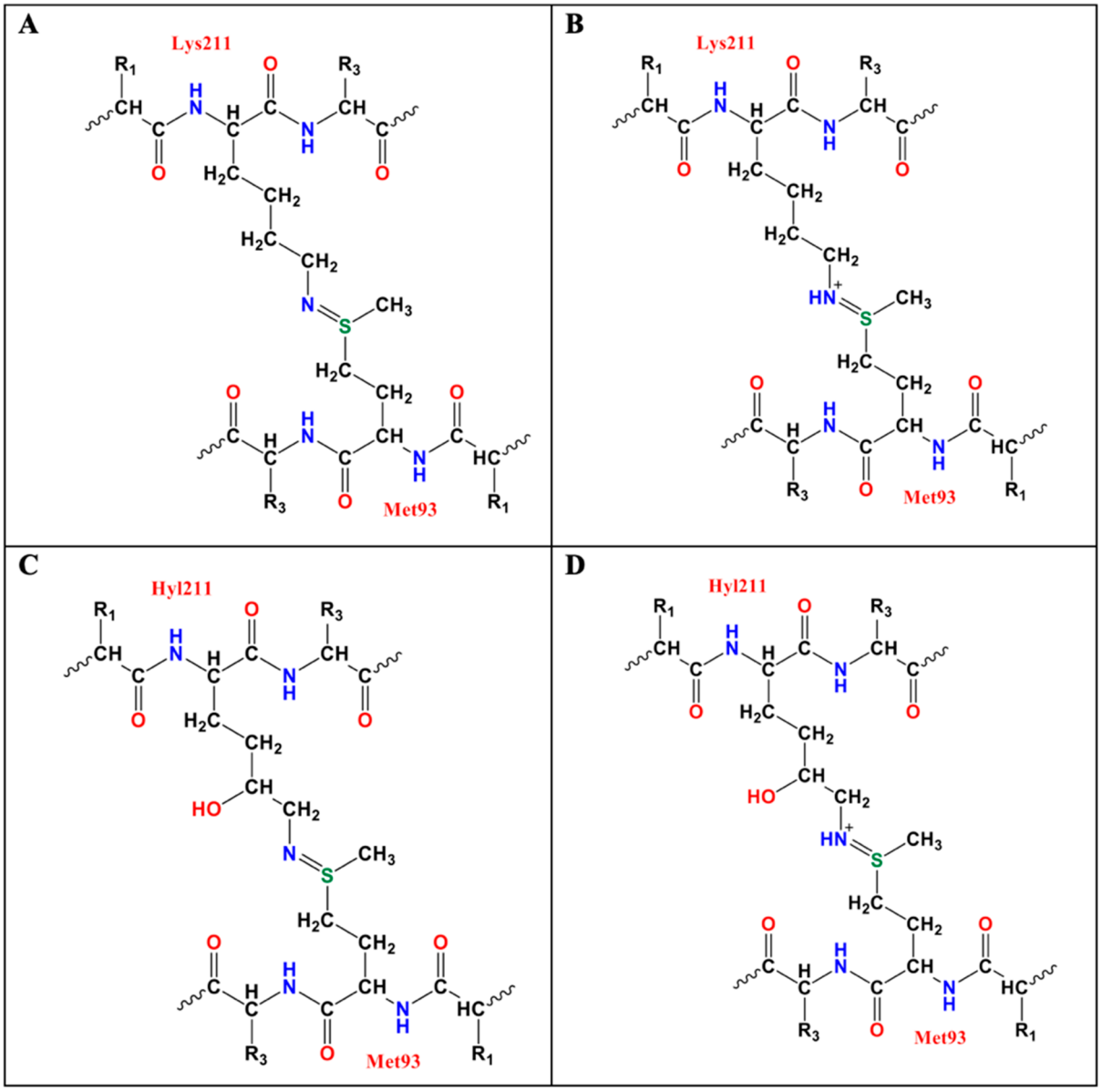


Figure 1. Schematic illustration of the sulfilimine (S═N) cross-links in collagen IV that were considered: (A) neutral link between Met93 and Lys211 residues; (B) protonated link between Met93 and Lys211 residues; (C) neutral link between Met93 and Hyl211 residues; and (D) protonated link between Met93 and Hyl211 residues.

differences or damage in collagen IV could lead to many different clinical diseases. Indeed, in Goodpasture syndrome, pathogenic autoantibodies can target and bind to its non-collagenous NC1 domains, causing conformational changes in the collagen IV network.[15] The deposition of pathogenic autoantibodies to collagen IV can also cause rheumatic diseases such as rheumatoid arthritis, connective tissue disease, systemic sclerosis, and juvenile dermatomyositis.[8] Meanwhile, Alport syndrome is a genetic disease resulting from mutations in the collagen IV COL4A3, COL4A4, or COL4A5 genes.[16] Similarly, mutations in the collagen COL4A1 gene can cause HANAC syndrome, constant nephropathy muscle cramps, and frequent intracranial aneurysms or affect the brain and/or retinal vasculature, anterior and posterior ocular structures, and the renal glomerules.[17]

As the part of mechanical properties of collagen, during fibril formation, some collagens such as collagen III, IV, VI, and XVI also form covalent cross-links.[3] In collagen IV networks, inter- or intrachain homophilic covalent cross-links can form among the four neighboring N-terminal 7S domains or two NC domains at the C-terminal.[6] In particular, as the α chains of collagen IV are cysteine and lysine rich, many of the cross-links known contain one or more of these residues. Indeed, cysteine-cysteine (disulfide), lysine-hydroxylysine, methionine-lysine, and methionine-hydroxylysine cross-links have been observed.[18] Some of these are known to be enzymatically formed, such as disulfide and lysine-hydroxylysine cross-links in the 7S domain(s), which can be formed by the action of lysyl oxidase-like 2 (LOXL2).[19]

However, these are not thought to be the major cross-links in collagen IV, especially in the C-terminal NC1 domains. Rather, the most important cross-link is held to be sulfilimine cross-links formed between methionine and lysine or hydroxylysine residues (Figure 1).[20] Importantly, the sulfilimine (S═N) bond and cross-link is unique to collagen IV, only occurring it its NC1 domains.[20−22] These cross-links are

https://doi.org/10.1021/acsomega.2c03360
ACS Omega 2022, 7, 39680−39689

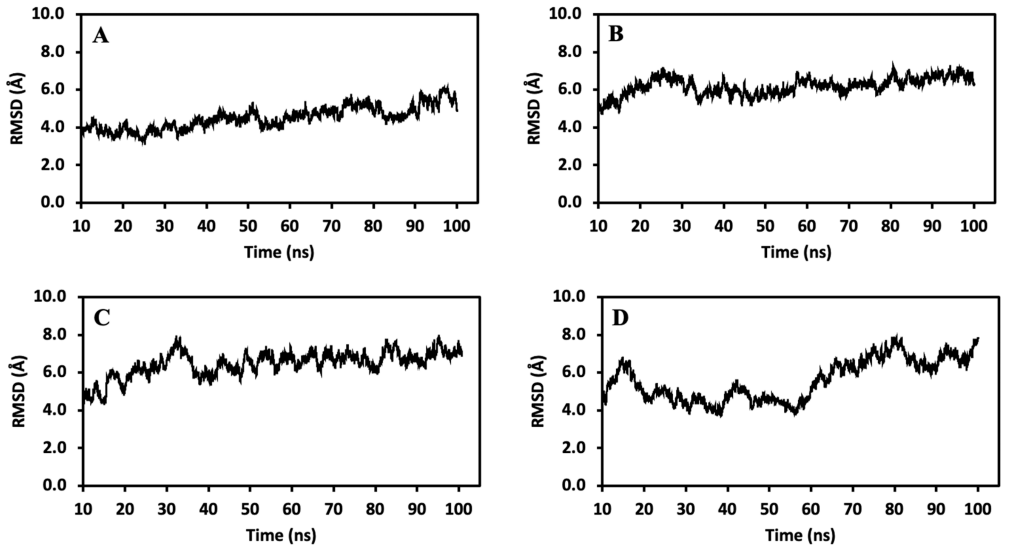


Figure 2. Plots of the RMSDs of the protein backbones of the (A) $_{Met93}S-N_{Lys211}$, (B) $_{Met93}S-NH_{Lys211}^{+}$, (C) $_{Met93}S-N_{Hyl211,}$ and (D) $_{Met93}S-NH_{Hyl211}^{+}$ systems from 10 to 100 ns production of 100 ns MD simulations.

not thought to be enzymatically formed. Instead, their formation is facilitated by hypohalous acids produced by peroxidasin, a basement membrane-bound heme peroxidase.[20,22] More specifically, peroxidasin is proposed to convert bromide to hypobromous acid.[23] It can then react with the sulfide sulfur of a methionine or side chain amine of a lysine/hydroxylysine to form a bromo halosulfonium ion or haloamine. Subsequently, these halo-intermediates react with their corresponding counterpart, the lysine/hydroxylysine amine or methionine sulfur, to form the sulfilimine bond.[23] Likewise, peroxidase can also convert chloride to hypochlorous acid, which can then also facilitate a sulfilimine bond in collagen IV but not as efficiently as hypobromous acid.[24]

As described above, covalent cross-link formation in collagen IV is immensely important to, for instance, its ability to provide mechanical and functional stability to the basement membrane.[18] Of all of its cross-links, the sulfilimine bond has been called the most important, as its biochemical appearance has been linked to the evolution of complex multicellular organisms.[25] However, despite its clear biochemical importance, the nature of the sulfilimine cross-link has remained unclear. More specifically, experimental studies have suggested or assumed that the sulfilimine bond is a neutral formal double bond between the sulfide sulfur of methionine and the side chain amine nitrogen of lysine/hydroxylysine. Previous computational studies[26] have suggested that such a bond is in fact better described as a coordinate covalent, or dative, bond. A previous computational study[27] and, more recently, our own computational studies[28] suggest that such bonds are not neutral within a biological context but instead are most likely protonated and furthermore, that this modifies the properties of the bond. It is in fact noted that protonation is one of the most important biochemical processes, especially to stabilize protein structure or biological cross-links.[29,30] However, how protonation of the sulfilimine bond may influence or effect both the behavior of the bond and collagen IV currently remains unclear or unknown.

Given the crucial role of sulfilimine cross-links on the role(s) of collagen IV, it is important to gain a greater understanding of how its occurrence and various possible states influence or effect collagen IV. Such insights are also likely to be of broader importance given that biological cross-links are essential for the proper functioning and roles of numerous biomolecules such as proteins and enzymes.[31] Molecular dynamics (MD) simulations are a powerful tool for obtaining valuable multiscale insights into, for instance, hydrogen bond interactions, residue conformational fluctuations, and protein dynamics as well as effects upon the environment (e.g., changes to solvent distribution).[32,33] Hence, in this study, we have performed a systematic series of MD simulations to investigate the effects of sulfilimine bonds in collagen IV, in several possible protonation states, on the protein fibers as well as the solvent distribution.

RESULTS AND DISCUSSION

As noted in the Computational Methods, the X-ray crystal structure of the NC1 hexamer of collagen IV (PDB ID: 1LI1), was used to construct four models of possible sulfilimine cross-linked systems. More specifically, Met93 was cross-linked with Lys211 or Hyl211, and the sulfilimine bond was either neutral or protonated. The four systems can be represented as $_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$, respectively. Molecular dynamic simulations allow conformational motion of a protein at the atomic level over a range of time. It also helps to determine the atomic fluctuations of protein and the effects of cross-links on stability.[32] Thus, the fluctuations of the protein backbones, the cross-linked residues, and the individual involved residues themselves were systematically considered.

RMSDs of the Protein Backbones. The root-mean-square deviation (RMSD) measures the difference between the structure of the protein backbone during the MD simulation from its initial structure. In addition, it can also illustrate the conformational flexibility over the duration of the 100 ns MD simulation. The initial 10 ns data from all systems were not included in the plots to allow the systems to be more fully equilibrated. Plots of the RMSD values of the protein backbones for the four systems modeled (i.e., $_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$), from 10 to 100 ns of 100 ns MD simulations, are shown in Figure 2.

https://doi.org/10.1021/acsomega.2c03360
ACS Omega 2022, 7, 39680−39689

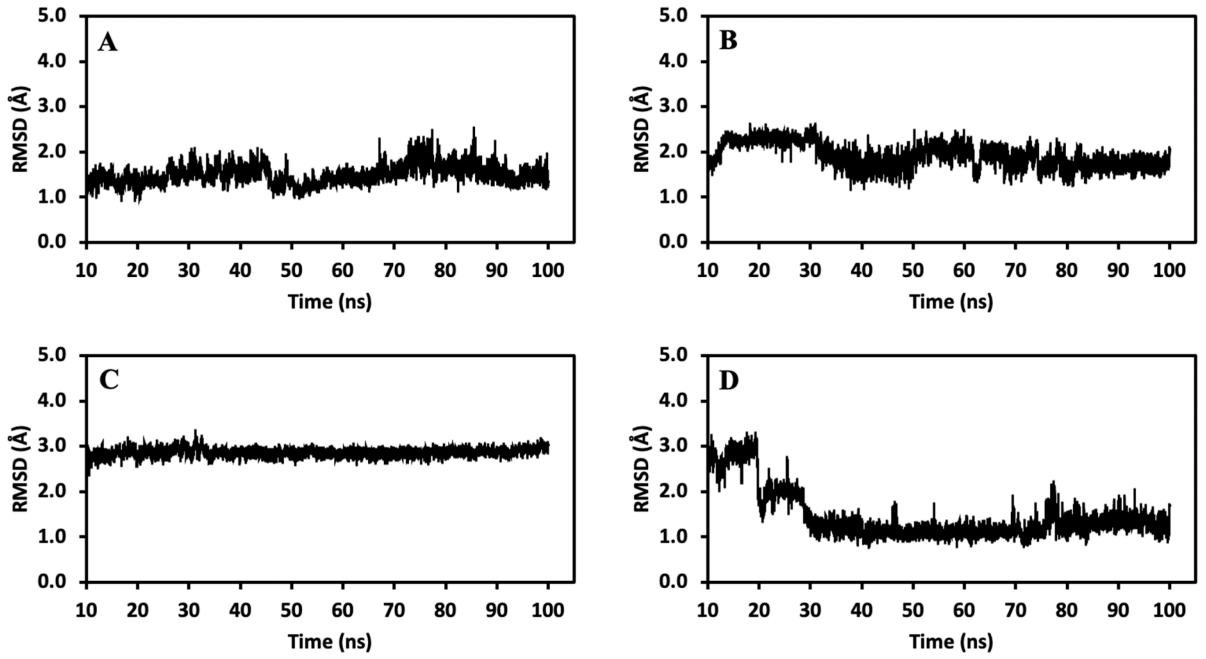


Figure 3. Plots of the RMSDs of the cross-linked residues Met93 and Lys211 or Hyl211 in the (A) $_{Met93}S-N_{Lys211}$, (B) $_{Met93}S-NH_{Lys211}^{+}$, (C) $_{Met93}S-N_{Hyl211}$, and (D) $_{Met93}S-NH_{Hyl211}^{+}$ systems from 10 to 100 ns of 100 ns MD simulations.

The average RMSD values with the standard deviation, from 10 to 100 ns, for the $_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$ systems are 4.50 ± 0.61 Å, 6.18 ± 0.46 Å, 6.47 ± 0.67 Å, and 5.59 ± 1.09 Å, respectively. These values suggest that over the course of the simulations the conformations sampled for each system did not differ too much from the corresponding initial starting structure. The smallest (4.50 ± 0.61 Å) and largest (6.47 ± 0.67 Å) deviations are observed for the unprotonated sulfilimine systems $_{Met93}S-N_{Lys211}$ and $_{Met93}S-N_{Hyl211}$, respectively. Furthermore, their RMSD values differ by 1.97 Å. In contrast, in the protonated systems $_{Met93}S-NH_{Lys211}^{+}$ and $_{Met93}S-NH_{Hyl211}^{+}$, the opposite trend is observed. More specifically, the average RMSD for the system in which hydroxylysine is involved in the sulfilimine bond, 5.59 ± 1.09 A, is notably smaller by 0.59 Å than the RMSD of the system in which lysine is involved (6.18 ± 0.46 Å). It is also noted that this difference in RMSD values between protonated systems is considerably smaller than that observed between the corresponding neutral systems. The plot of the backbone RMSD values (Figure 2D) for the $_{Met93}S-NH_{Hyl211}^{+}$ system, and corresponding standard deviation (±1.09 Å), suggest that even after 100 ns the system may still be fluctuating more than the others. Therefore, we extended the MD simulation another 90 ns (Figure S1). This showed that by around 100 ns, the RMSD values for the $_{Met93}S-NH_{Hyl211}^{+}$ system more consistently averaged around 6.05 Å with a standard deviation of ±0.95 Å (for the RMSD values obtained from 10 to 190 ns).

RMSDs of the Cross-Linked Residues. The root-mean-square deviation (RMSD) of just the cross-linked residues themselves was also analyzed from 10 to 100 ns of 100 ns MD simulations and are shown plotted in Figure 3. The resulting average RMSD values, from 10 to 100 ns simulations, for the $_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$ systems are 1.50 ± 0.21 Å, 1.91 ± 0.28 Å, 2.87 ± 0.09 Å, and 1.46 ± 0.55 Å, respectively. As observed for the protein backbone RMSD average values, the smallest (1.46 ± 0.55 Å) and largest (2.87 ± 0.09 Å) average values are observed for the hydroxyl sulfilimine systems $_{Met93}S-NH_{Hyl211}^{+}$ and $_{Met93}S-N_{Hyl211}$, respectively. Furthermore, they differ from each other by 1.41 Å with the system in which the sulfilimine bond contains a neutral hydroxylysyl, giving the markedly higher average RMSD.

The average values noted above for those systems in which a lysyl is involved in the sulfilimine bond, i.e., $_{Met93}S-N_{Lys211}$ and $_{Met93}S-NH_{Lys211}^{+}$, do generally reflect the behavior of the cross-linked residues over the course of the entire simulations. For instance, as can be seen in Figure 3, the RMSD values of the $_{Met93}S-N_{Lys211}$ system are almost consistently below 2.0 Å over the duration of the MD simulation. In contrast, in the corresponding protonated sulfilimine bond system, $_{Met93}S-NH_{Lys211}^{+}$, the RMSDs of the cross-linked residues are higher (as the above-mentioned averages suggest) and are more consistently around 2.0 Å over the duration of the MD simulation.

This noted consistency in observed conformations over the course of the MD simulations is also illustrated by, for example, comparison of the MM minimized conformation and those obtained every 10 ns (from 0 to 100 ns) over the course of the MD simulations. For the neutral and protonated systems in which the sulfilimine bond is formed between Met93 and Lys211, these structures are collated in Figures S2 and S3, respectively (see the Supporting Information). Furthermore, in both systems, the conformation of the cross-linked residues is quite open. More specifically, the positioning of Lys211 and Met93 to each other can be described as generally roughly perpendicular to more linear relative to each other.

For the systems in which it is a hydroxylysyl residue involved in the sulfilimine bond, as noted for the RMSDs of the protein backbones (Figure 2), the neutral system $_{Met93}S-N_{Hyl211}$ has a higher average RMSD value (2.87 ± 0.09 Å) than the corresponding protonated system $_{Met93}S-NH_{Hyl211}^{+}$ (1.46 ± 0.55 Å). However, these average values do not necessarily reflect the behavior of the residues over the course of the

https://doi.org/10.1021/acsomega.2c03360
ACS Omega 2022, 7, 39680−39689

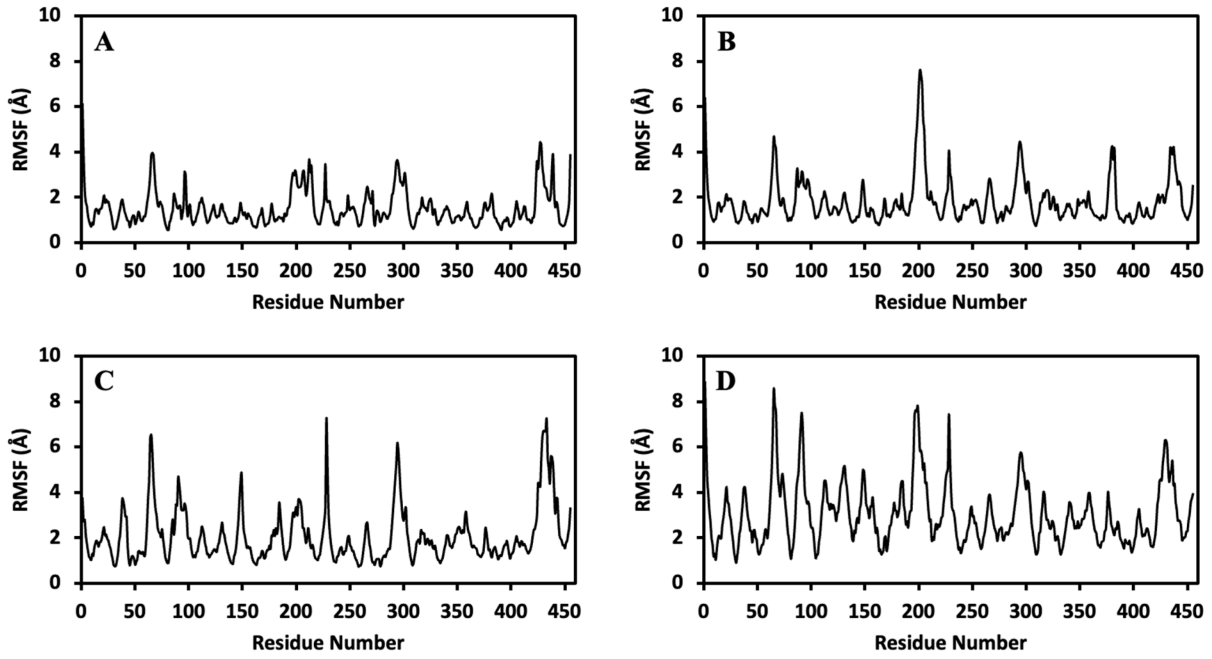


Figure 4. Plots of the RMSFs obtained for the protein backbone α-carbons of each residue in the (A) $_{Met93}S-N_{Lys211}$, (B) $_{Met93}S-NH_{Lys211}^{+}$, (C) $_{Met93}S-N_{Hyl211}$, and (D) $_{Met93}S-NH_{Hyl211}^{+}$ systems from 10 to 100 ns of 100 ns MD simulations.

simulations. Indeed, as can be seen in Figure 3, for the $_{Met93}S-NH_{Lys21}$ system, after the first ∼11 ns the Met93-Hyl211 cross-linked residues are conformationally quite stable, with RMSD values consistently near 3.0 Å. In contrast, for the protonated cross-link system $_{Met93}S-NH_{Hyl211}^{+}$, larger RMSD values are obtained during the first 30 ns, suggesting some initial shorter-lived and large RMSD conformations. Then, however, from 30 ns onward, the cross-linked residues give markedly lower RMSD values, consistently around ∼1.0 Å, suggesting they assume a longer lived and lower RMSD conformation.

As for the systems in which a lysyl is involved in the sulfilimine bond (see above), the conformational behavior over the course of the MD simulations can be illustrated by comparing the MM minimized conformation and those obtained at 10 ns intervals (from 0 to 100 ns) from the MD simulations. For the neutral and protonated systems in which the sulfilimine bond is formed between Met93 and Hyl211, these structures are collated in Figures S4 and S5, respectively (see the Supporting Information). For the neutral $_{Met93}S-N_{Hyl211}$ system, the conformation of the sulfilimine cross-linked residues quite quickly, as seen in the 10 ns snapshot (Figure S4), takes on a U-shape rather than the more open conformations seen for $_{Met93}S-N_{Lys211}$, and $_{Met93}S-NH_{Lys211}^{+}$. The "U" conformation is then consistently observed over the course of the MD simulation. In contrast, for the protonated $_{Met93}S-NH_{Hyl211}^{+}$ system. as indicated in Figure 3D, the initial snapshots (e.g., minimized, 0, 10, and 20 ns) show greater conformational variability (e.g., perpendicular to linear to eventually more U-shaped). In the snapshots taken at 30 ns and longer, the cross-linked residues in $_{Met93}S-NH_{Hyl211}^{+}$ have a more U-shaped arrangement similar to that in the corresponding neutral $_{Met93}S-N_{Hyl211}$ system, though perhaps more open (Figure S5).

Hydrogen bonds, in particular inter-residue interactions, can play a critical role in proteins' conformational structure and stability. In addition, protonation of the sulfilimine bond, whether the residue is a lysyl (Lys211) or hydroxylysyl (Hyl211), could potentially influence such interactions. Hence, we also analyzed hydrogen bond interactions between the cross-linked residues in all four systems considered ($_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$) with other protein residues, and their consistency. Interaction maps were created for the MM minimized structures, the equilibrated structures (0 ns), and then for snapshots every 10 ns over the course of the 100 ns MD simulation production runs. They are collated in the Supporting Information (Figures S2−S5) for the $_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$ systems, respectively. In all four cases, they show that the backbone amides of the cross-linked residues can form hydrogen bonds with other protein residues, but that they are inconsistent and variable. For example, in the $_{Met93}S-N_{Lys211}$ system Met317 at times hydrogen bonds to the amide −NH− of Met93. However, in the corresponding $_{Met93}S-NH_{Lys211}^{+}$ system, this is not observed. Instead, Glu213 hydrogen bonds at times to the backbone −NH− of Lys211. Different residues are then observed in the analogous hydroxylysyl systems ($_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$), see Figures S6−S9. It is also noted that in the protonated sulfilimine systems, no hydrogen bonds were observed between the cross-link and protein residues. Meanwhile, in Hyl211-containing systems, the side chain −OH group of Hyl211 did not hydrogen bond with protein residues except in MM minimized or an initial MD structure. These maps do, however, showcase the above-mentioned relative positioning of Met93 to Lys/Hyl211 in the four systems.

The RMSD results suggest that when the sulfilimine bond is unprotonated, less conformational variation is observed when a lysyl ($_{Met93}S-N_{Lys211}$) is involved rather than a hydroxylysyl ($_{Met93}S-N_{Hyl211}$). However, when the sulfilimine bond is protonated, the reverse holds. That is, greater conformational variation is observed when a lysyl ($_{Met93}S-NH_{Lys211}^{+}$) is involved rather than a hydroxylysyl ($_{Met93}S-N_{Hyl211}^{+}$). But the difference between their average RMSD values, 1.91 ± 0.28 Å

https://doi.org/10.1021/acsomega.2c03360
ACS Omega 2022, 7, 39680−39689

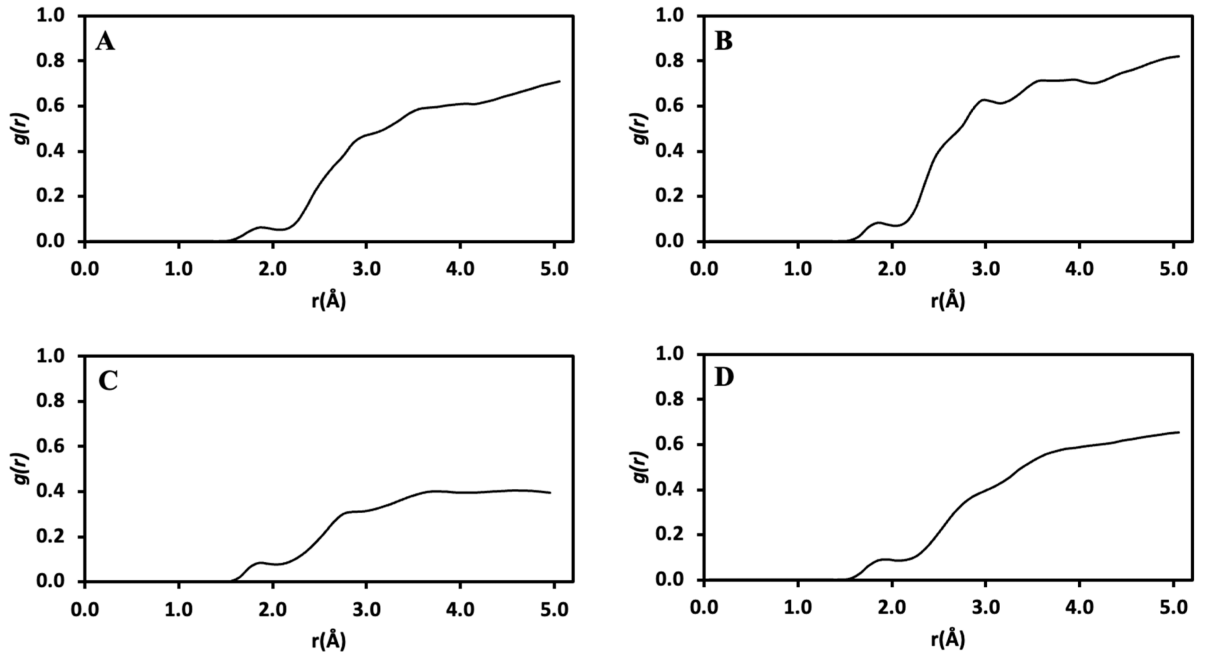


Figure 5. Plots of the radial distribution function's (RDF) for water molecules up to 5 Å from the sulfilimine cross-linked residues in the (A) $_{Met93}S-N_{Lys211}$, (B) $_{Met93}S-NH_{Lys211}^{+}$, (C) $_{Met93}S-N_{Hyl211}$, and (D) $_{Met93}S-NH_{Hyl211}^{+}$ systems as observed from 10 to 100 ns of 100 ns MD simulations.

and 1.46 ± 0.55 Å, respectively, is lower at 0.45 Å than that observed between the corresponding neutral systems $_{Met93}S-N_{Lys211}$ and $_{Met93}S-N_{Hyl211}$.

The above RMSD plots of the protein backbones, for example, illustrate the average dynamic behavior of the entire protein backbones over the course of the MD simulations. Meanwhile, those of the cross-linked residues give a more focused illustration of their conformational behavior over the course of the MD simulations. Indeed, in the above, the RMSDs are an average of, at a given point in time, the differences between the current position of all particles (e.g., atoms) and the corresponding initial starting structure. In contrast, root-mean-square fluctuations (RMSFs) are the differences between the position of a particle (e.g., atom or residue) and, in this case, its initial starting position averaged over the time duration of the MD simulations (100 ns). Like RMSD values, RMSF values can also provide insight into the dynamic behavior of a protein, residue(s) or atom(s). Thus, select RMSF values obtained from the MD simulations of each of the four systems considered were also calculated.

RMSFs of the Protein Backbone α-Carbons (C_α). As each protein system contained 455 amino acid residues, to examine the structural fluctuations of all the residues, we calculated only the RMSF values of the α-carbons of all the residues for all four systems ($_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$). The resulting calculated RMSF values for each residues C_α, for all four systems, are shown in Figure 4. For each system (i.e., $_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$), their average RMSF values for the α-carbons are 1.51 ± 0.76 Å, 1.80 ± 1.00 Å, 2.06 ± 1.19 Å, and 3.05 ± 1.38 Å, respectively. These values suggest that overall, when hydroxylysyl is involved in the sulfilimine bond larger fluctuations are observed for the α-Carbons. Furthermore, the protonation of the sulfilimine bond also corresponds to larger fluctuations in the positions of the α-carbons. As can be seen in Figure 4, for the $_{Met93}S-N_{Lys211}$ system, the largest RMSF value(s) is 6.09 Å (Asp1). Upon protonation (i.e., $_{Met93}S-NH_{Lys211}^{+}$) there does indeed appear to be an increase in α-carbon RMSF values along much of the length of the protein with the largest RMSF value now being 7.62 Å (Phe201). In contrast, in the $_{Met93}S-N_{Hyl211}$ system large spikes in α-carbon RMSF values are observed along the protein with the largest RMSF value (7.26 Å) occurring for residue Thr228. Again, protonation of the sulfilimine bond (i.e., $_{Met93}S-NH_{Hyl211}^{+}$) does again appear to cause increased α-carbon RMSF values along the peptide. But the maximum value observed is now 8.84 Å for residue Asp1 (see Figure 4D), which is markedly less than in the corresponding unprotonated system.

Distribution of Water around the Sulfilimine Cross-links. Water can play crucial roles in maintaining, for instance, the structure, dynamics, and functions of biomolecules such as proteins.[34] These roles can arise due to, for example, its electrostatic properties and/or its abilities to act as a hydrogen bond donor and acceptor. Given that the sulfilimine bond can potentially exist in one or more protonation states, and that the lysyl residue (Lys211) involved in the sulfilimine bond could instead by a side-chain hydroxylated lysyl (Hyl211), one or both of these factors could influence the water distribution around the sulfilimine bonds.

The production MD simulations for each of the four systems considered here can also potentially provide insights into the water distribution around their sulfilimine bonds, and possibly indicate the effects of protonating the sulfilimine bonds as well as the impact of the side chain hydroxyl group in Hyl. The radial distribution function (RDF) is used to define the probability of finding a particle at a distance from a fixed particle. To examine the density of water molecules around the sulfilimine cross-links in the four systems considered herein, we determined the RDF of the water molecules from the cross-linked residues (i.e., Met93 and Lys211 or Hyl211), shown in Figure 5.

https://doi.org/10.1021/acsomega.2c03360
ACS Omega 2022, 7, 39680−39689

Table 1. Calculated Solvent Accessible Surface Area (SASA) of the Cross-linked Residues Met93 and Lys211/Hyl211 in the $_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$ Systems Obtained at Every 20 ns over the 100 ns Production MD Simulations

		solvent accessible surface area (SASA)					
cross-links	residue	0 ns	20 ns	40 ns	60 ns	80 ns	100 ns
$_{Met93}S-N_{Lys211}$	Met93	95.967	111.854	146.937	136.069	157.568	141.892
	Lys211	106.270	58.326	83.860	118.059	153.182	146.592
$_{Met93}S-NH_{Lys211}^{+}$	Met93	61.139	161.228	173.543	177.457	176.950	177.899
	Lys211	68.550	117.093	65.740	108.488	126.755	122.491
$_{Met93}S-N_{Hyl211}$	Met93	22.793	67.423	87.383	65.878	85.268	52.302
	Hyl211	110.73	16.969	38.773	18.860	35.450	37.907
$_{Met93}S-NH_{Hyl211}^{+}$	Met93	79.104	115.913	181.214	66.936	166.883	152.695
	Hyl211	117.687	81.018	114.106	113.152	131.385	135.015

The RDF of water molecules indicate the zero probability of water from the sulfilimine cross-link when the distance is less than 1.45 Å in all four systems (Figure 5). The RDF's obtained for the neutral $_{Met93}S-N_{Lys211}$ and $_{Met93}S-N_{Hyl211}$ systems are shown in Figures 5A, C respectively. The only functional group difference between these two systems is the hydroxyl attached to the side chain's d-carbon. As can be seen, for the $_{Met93}S-N_{Lys211}$ system, its RDF value then increases to ∼0.05 by ∼2.0 Å and then increases even further to a value over 0.39 by 5.0 Å. In contrast, for the $_{Met93}S-N_{Hyl211}$ system, its RDF value increases to approximately 0.13 by 2.4 Å then increases to almost 0.40 by 3.65 Å, where it essentially remains as the distance from the residues increases to 5.0 Å. These plots suggest that the presence of the d-hydroxyl group in Hyl211, at least in the neutral sulfilimine cross-linked systems, helps reduce the water density around the cross-linked residues by either influencing their movement or via conformational differences between the two neutral sulfilimine cross-links.

Comparison of the RDF plots shown in Figure 5A ($_{Met93}S-N_{Lys211}$) and 5B ($_{Met93}S-NH_{Lys211}^{+}$) shows the effects on the RDF values upon protonation of the sulfilimine bond when it is formed by cross-linking Met93 and Lys211. For both systems, the RDF values again are low at around ∼0.01 until approximately 2.0 Å, although it is noted that the RDF values are slightly higher in the $_{Met93}S-NH_{Lys211}^{+}$ system. However, compared to the neutral $_{Met93}S-N_{Lys211}$ system, in the $_{Met93}S-NH_{Lys211}^{+}$ system, a much more rapid increase is observed in its RDF values, to ∼0.6, as the distance increases from 2.0 to ∼3.0 Å. The RDF values continue to rise as the distance from the cross-linked residues increases further to 5.0 Å. Notably, however, at 5.0 Å, the RDF value (∼0.8) for the $_{Met93}S-NH_{Lys211}^{+}$ system is slightly higher than the unprotonated $_{Met93}S-N_{Lys211v}$ system (see Figure 5B). Hence, protonation of the sulfilimine bond slightly increases the RDF water's values near the cross-link (i.e., out to ∼2.0 Å), but further increases the water's RDF values as the distance from the cross-linked residues increases to 5.0 Å.

The RDF plots for the corresponding neutral ($_{Met93}S-N_{Hyl211}$) and protonated ($_{Met93}S-NH_{Hyl211}^{+}$) hydroxylysyl-containing systems are shown in Figures 5C, D, respectively. As for the above lysyl211 containing systems, protonation of the sulfilimine bond increases the calculated RDF values near the cross-link out to ∼2.0 Å, this increases even further in hydroxylysyl211 (see Figure 5). Then, as the distance from the cross-linked residues goes from 2.0 to 5.0 Å, an increase in the RDF value is again observed. Similar to that described above for the above lysyl-containing systems, a much more rapid increase in the RDF values is seen in the protonated (i.e., $_{Met93}S-NH_{Hyl211}^{+}$) system than the neutral (i.e., $_{Met93}S-N_{Hyl211}$) system as we increase the distance from the cross-linked residues from 2.0 to 5.0 Å. More specifically, for the $_{Met93}S-NH_{Hyl211}^{+}$ system (Figure 5D) the RDF value rapidly rises to over 0.4 by 3.0 Å away from the residues, then continues to rise to over 0.6 as the distance further increases to 5.0 Å. This RDF value is lower than that obtained for the corresponding neutral system, $_{Met93}S-N_{Hyl211}$, at 5.0 Å. As noted above during comparison of the maximum RDF values calculated for the two neutral systems ($_{Met93}S-N_{Lys211}$ and $_{Met93}S-N_{Hyl211}$), the overall lower maximum RDF values obtained for the hydroxylysyl-containing $_{Met93}S-NH_{Hyl211}^{+}$ system compared to the $_{Met93}S-NH_{Lys211}^{+}$ system may again in part reflect conformational differences between the two sulfilimine cross-links.

Together, however, these results suggest that protonation of the sulfilimine bond, perhaps by forming a stronger interaction with the solvent waters in general or by a water molecule (e.g., by impacting their movement) reducing the RDF of the solvent waters in the immediate vicinity of the sulfilimine bond. However, as the distance from the cross-linked residues increases out to 5.0 Å, the protonated systems exhibit a more rapid and larger increase in their RDF values, i.e., an increase in the waters around the cross-linked residues, compared to that observed in the corresponding neutral systems. In addition, we further calculated the RDF values out to a distance of 10 Å so as to better include the second solvation shell, etc. It has been noticed that the RDF values remains consistent from 5 to 10 Å for all systems, which are ∼0.7, ∼0.8, ∼0.5, and ∼0.7 for the $_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$ systems, respectively (Figure S10).

Snapshots of the water distribution around the sulfilimine cross-linked residues for the $_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$ systems were also obtained at their MM minimized and equilibrated (0 ns) conformations, and then at every 10 ns over the course of the 100 ns production MD simulations and are provided in the Supporting Information (Figures S11−S15). These further highlight the distribution of waters around the sulfilimine cross-links including the greater water density around the $_{Met93}S-N_{Lys211}$ and $_{Met93}S-NH_{Lys211}^{+}$ systems compared to the $_{Met93}S-N_{Hyl211}$ and $_{Met93}S-NH_{Hyl211}^{+}$ systems.

Solvent Accessible Surface Area (SASA). SASA is a measure of the surface area of a molecule or a single residue that is accessible to a solvent such as water. In this present study, the SASA for residues involved in the sulfilimine bond (i.e., Met93 and Lys211 or Hyl211) was calculated to help

https://doi.org/10.1021/acsomega.2c03360
ACS Omega 2022, 7, 39680−39689

provide insight into the solvent exposure of the cross-link. Furthermore, in all four systems ($_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$), the residues SASA values were calculated at their initial (equilibrated: 0 ns) conformations, and then at every 20 ns of the 100 ns production MD simulations and are provided in Table 1.

The calculated SASA values suggest that in general, the Met93 residue is more solvent exposed compared to the Lys211 or Hyl211 residue in all four systems considered; that is, regardless of the protonation state of the sulfilimine bond (i.e., its nitrogen center). However, it should also be noted that in the protonated systems, both residues involved in the cross-link have higher SASA values (i.e., are more solvent exposed) than in their corresponding unprotonated (neutral) systems (see Table 1.). This again illustrates that protonation of the sulfilimine bond does influence its conformation and structure. It is noted that the SASA values calculated for the various snapshots also reflect the conformational changes described above for the Met93-Lys211/Hyl211 cross-links (see Figure 3.).

CONCLUSION

We have performed a series of MD simulations to gain insight into the dynamic behavior of sulfilimine cross-link in collagen IV. More specifically, the sulfilimine cross-link involves the sulfur of Met93 and the side chain amine nitrogen of a Lys211/Hyl211 residue. Furthermore, its nitrogen center can be protonated. Hence, we considered four possible sulfilimine cross-links: $_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$, $_{Met93}S-N_{Hyl211}$, and $_{Met93}S-NH_{Hyl211}^{+}$ (i.e., two systems containing a neutral sulfilimine cross-link, and two containing a protonated sulfilimine cross-link). The main findings of these studies can be summarized as follows:

- Protein backbone RMSD analysis showed that over the course of the 100 ns MD simulations the conformations sampled for each system did not differ too much from the corresponding initial structure. The smallest (4.50 ± 0.61 Å) and largest (6.47 ± 0.67 Å) average deviations are observed for the neutral sulfilimine systems $_{Met93}S-N_{Lys211}$ and $_{Met93}S-N_{Hyl211}$, respectively.
- Based on RMSD analysis of the cross-linked residues' only, it was observed that, like the protein backbone RMSD average values, the smallest (1.46 ± 0.55 Å) and largest (2.87 ± 0.09 Å) average values are observed for the unprotonated sulfilimine systems $_{Met93}S-NH_{Hyl211}$ and $_{Met93}S-N_{Hyl211}$, respectively.
- RMSFs of the backbone α-carbon analysis showed that, overall, when Hyl211 is involved in the sulfilimine bond, larger fluctuations are observed for α-carbons.
- Conformational change analysis showed that when the sulfilimine cross-link involves hydroxylysine (i.e., $_{Met93}S-N_{Hyl211}$, $_{Met93}S-HN_{Hyl211}^{+}$), the conformations of the cross-linked residues take on a U-shape, while those involving lysine (i.e., $_{Met93}S-N_{Lys211}$, $_{Met93}S-NH_{Lys211}^{+}$) have a more open conformation.
- Based on RDF analysis, it has been observed that the presence of the d-hydroxyl group in Hyl211, at least in the neutral sulfilimine cross-linked systems, helps reduce the water density around the cross-linked residues by either influencing their movement or, via conformational differences between the two neutral sulfilimine cross-links. However, it has also been observed that the water density increases as the RDF value increases due to the protonation of sulfilimine bond for both lysyl and hydroxylysyl systems.
- SASA analysis showed that in general, the Met93 residue is more solvent exposed compared to the Lys211 or Hyl211 residue in all four systems considered. However, it has also been observed that in the protonated systems both residues involved in the cross-link have higher SASA values (i.e., are more solvent exposed) than in their corresponding unprotonated (neutral) systems. These SASA analyses also suggest that protonation of the sulfilimine bond influences its conformation and structure.

COMPUTATIONAL METHODS

Structure Preparation. A suitable crystal structure of NC1 hexamer of collagen IV was obtained from the Protein Data Bank (PDB ID: 1LI1) that contains 6 (A, B, C, D, E, and F) chains representing NC1 hexamers.[35] Among these six chains, chains A, B, D, and E contain Met93 and Lys211 residues that may involve the sulfilimine cross-link formation; specifically, chain A cross-links with chain E and chain B cross-links with chain D. For this study, chain B and chain D were used in MD simulations. As the cross-link forms between the sulfur (S) of methionine93 of chain D and the amine nitrogen (N) of hydroxylysine/lysine211 of chain B, four cross-linked structures were selected and prepared for MD and are shown in Figure 1. All the structures were prepared using the Molecular Operator Environment (MOE) program.[36]

Molecular Dynamic Simulations (MD). An initial minimization was performed using the AMBER14SB force field[37] after preparing the structures for MD using MOE. Then, all four minimized structures were solvated such that each structure was surrounded by a water layer 10 Å deep with periodic boundary conditions enabled. All four solvated systems were again minimized using molecular mechanics (MM) AMBER14SB force field until the root-mean-square gradient fell below 0.01 kcal/(mol Å). The minimized systems were then submitted for heating in which the temperature gradually increased from 0 to 300 K at constant pressure (1 atm). The heated systems were then submitted for 500 ps equilibration with a time step of 2 fs at 300 K constant temperature. The Nosé−Poincaré thermostat[38] was coupled with the equations of motion, where a 2 fs time step was set for numerical integration. The Particle Mesh Ewald method[39] was used for calculating Coulombic interactions, with cutoffs for nonbonded long-range interactions set to 8 Å. Finally, the equilibrated systems were submitted for 100 ns production MD simulations, under the same conditions. Additionally, the production run was extended for another 90 ns (i.e., 190 ns total) for $_{Met93}S-NH_{Hyl211}^{+}$ system. All MD simulations were performed using the NAMD program.[40] This MD protocol has been successfully applied in other enzymatic studies.[41,42]

Analyses. The trajectories of the 100 ns MD simulations were saved every picosecond. However, the initial 10 ns data were not considered in the average and standard deviation calculations; allowing the molecular systems more time to fully equilibrate. Therefore, all the RMSD, RMSF, and RDF values are presented with standard deviation in the Results and Discussion section are from 10 to 100 ns of total 100 ns MD simulations. As the additional 90 ns MD simulation was done for the $_{Met93}S-NH_{Hyl211}^{+}$ system, the 10 to 190 ns of total 190

https://doi.org/10.1021/acsomega.2c03360
ACS Omega 2022, 7, 39680−39689

ns RMSD values are presented in Figure S1. The results were then analyzed and clustered according to the root-mean-square deviation (RMSD). The analyses such as RMSD, RMSF, RDF, and water density distribution were conducted using VMD.[43] The conformational changes and hydrogen bonds were analyzed using MOE. The solvent-accessible surface areas of specific residues were estimated using the PyMol program.[44]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.2c03360.

Plot of RMSD obtained for the protein backbone of $_{Met93}S-NH_{Hyl211}^{+}$; MM minimized structures of the sulfilimine cross-link and snapshots taken every 10 ns (0 to 100 ns) from the corresponding MD simulations; maps of hydrogen bonding involving the cross-linked residues; plots of the radial distribution functions (RDF) for water molecules up to 10 Å from the sulfilimine cross-link; visualization of the water distribution around sulfilimine cross-link (PDF)

Topology and coordinate files of the neutral systems $_{Met93}S-N_{Lys211}$ and $_{Met93}S-N_{Hyl211}$ (ZIP)

Topology and coordinate files of the protonated systems $_{Met93}S-NH_{Lys211}^{+}$ and $_{Met93}S-NH_{Hyl211}^{+}$ (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

James W. Gauld − *Department of Chemistry and Biochemistry, University of Windsor, Windsor, Ontario N8S 1C1, Canada;* orcid.org/0000-0002-2956-9781; Email: gauld@uwindsor.ca

Author

Anupom Roy − *Department of Chemistry and Biochemistry, University of Windsor, Windsor, Ontario N8S 1C1, Canada*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c03360

Author Contributions

A.R.: molecular dynamics simulations, analyses, and writing; J.W.G.: mentoring, reviewing, and editing.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the Natural Science and Engineering Research Council of Canada (NSERC) for financial support. We thank Compute Canada for computational resources.

■ REFERENCES

(1) Di Lullo, G. A.; Sweeney, S. M.; Korkko, J.; Ala-Kokko, L.; San Antonio, J. D. Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen. *J. Biol. Chem.* **2002**, *277* (6), 4223−31.

(2) Ricard-Blum, S. The collagen family. *Cold Spring Harb Perspect Biol.* **2011**, *3* (1), a004978.

(3) Gutsmann, T.; Fantner, G. E.; Kindt, J. H.; Venturoni, M.; Danielsen, S.; Hansma, P. K. Force spectroscopy of collagen fibers to investigate their mechanical properties and structural organization. *Biophys. J.* **2004**, *86* (5), 3186−93.

(4) Kalluri, R. Basement membranes: structure, assembly and role in tumour angiogenesis. *Nat. Rev. Cancer* **2003**, *3* (6), 422−33.

(5) Khoshnoodi, J.; Pedchenko, V.; Hudson, B. G. Mammalian collagen IV. *Microsc Res. Tech* **2008**, *71* (5), 357−70.

(6) Naylor, R. W.; Morais, M.; Lennon, R. Complexities of the glomerular basement membrane. *Nat. Rev. Nephrol* **2021**, *17* (2), 112−127.

(7) Mao, M.; Alavi, M. V.; Labelle-Dumais, C.; Gould, D. B. Type IV Collagens and Basement Membrane Diseases: Cell Biology and Pathogenic Mechanisms. *Curr. Top Membr.* **2015**, *76*, 61−116.

(8) Abreu-Velez, A. M.; Howard, M. S. Collagen IV in Normal Skin and in Pathological Processes. *N Am. J. Med. Sci.* **2012**, *4* (1), 1−8.

(9) Feru, J.; Delobbe, E.; Ramont, L.; Brassart, B.; Terryn, C.; Dupont-Deshorgue, A.; Garbar, C.; Monboisse, J. C.; Maquart, F. X.; Brassart-Pasco, S. Aging decreases collagen IV expression in vivo in the dermo-epidermal junction and in vitro in dermal fibroblasts: possible involvement of TGF-beta1. *Eur. J. Dermatol* **2016**, *26* (4), 350−60.

(10) Varani, J.; Dame, M. K.; Rittie, L.; Fligiel, S. E.; Kang, S.; Fisher, G. J.; Voorhees, J. J. Decreased collagen production in chronologically aged skin: roles of age-dependent alteration in fibroblast function and defective mechanical stimulation. *Am. J. Pathol.* **2006**, *168* (6), 1861−8.

(11) Tanjore, H.; Kalluri, R. The role of type IV collagen and basement membranes in cancer progression and metastasis. *Am. J. Pathol.* **2006**, *168* (3), 715−7.

(12) Aumailley, M.; Timpl, R. Attachment of cells to basement membrane collagen type IV. *J. Cell Biol.* **1986**, *103* (4), 1569−75.

(13) Ferreira, A. M.; Gentile, P.; Chiono, V.; Ciardelli, G. Collagen for bone tissue regeneration. *Acta Biomater* **2012**, *8* (9), 3191−200.

(14) Bahramsoltani, M.; Slosarek, I.; De Spiegelaere, W.; Plendl, J. Angiogenesis and collagen type IV expression in different endothelial cell culture systems. *Anat Histol Embryol* **2014**, *43* (2), 103−15.

(15) Pedchenko, V.; Kitching, A. R.; Hudson, B. G. Goodpasture's autoimmune disease - A collagen IV disorder. *Matrix Biol.* **2018**, *71*−*72*, 240−249.

(16) Cosgrove, D.; Liu, S. Collagen IV diseases: A focus on the glomerular basement membrane in Alport syndrome. *Matrix Biol.* **2017**, *57*−*58*, 45−54.

(17) Vahedi, K.; Alamowitch, S. Clinical spectrum of type IV collagen (COL4A1) mutations: a novel genetic multisystem disease. *Curr. Opin Neurol* **2011**, *24* (1), 63−8.

(18) Wu, Y.; Ge, G. Complexity of type IV collagens: from network assembly to function. *Biol. Chem.* **2019**, *400* (5), 565−574.

(19) Anazco, C.; Lopez-Jimenez, A. J.; Rafi, M.; Vega-Montoto, L.; Zhang, M. Z.; Hudson, B. G.; Vanacore, R. M. Lysyl Oxidase-like-2 Cross-links Collagen IV of Glomerular Basement Membrane. *J. Biol. Chem.* **2016**, *291* (50), 25999−26012.

(20) Vanacore, R.; Ham, A. J.; Voehler, M.; Sanders, C. R.; Conrads, T. P.; Veenstra, T. D.; Sharpless, K. B.; Dawson, P. E.; Hudson, B. G. A sulfilimine bond identified in collagen IV. *Science* **2009**, *325* (5945), 1230−4.

(21) Ronsein, G. E.; Winterbourn, C. C.; Di Mascio, P.; Kettle, A. J. Cross-linking methionine and amine residues with reactive halogen species. *Free Radic Biol. Med.* **2014**, *70*, 278−87.

(22) Bhave, G.; Cummings, C. F.; Vanacore, R. M.; Kumagai-Cresse, C.; Ero-Tolliver, I. A.; Rafi, M.; Kang, J. S.; Pedchenko, V.; Fessler, L. I.; Fessler, J. H.; Hudson, B. G. Peroxidasin forms sulfilimine chemical bonds using hypohalous acids in tissue genesis. *Nat. Chem. Biol.* **2012**, *8* (9), 784−90.

(23) McCall, A. S.; Cummings, C. F.; Bhave, G.; Vanacore, R.; Page-McCaw, A.; Hudson, B. G. Bromine Is an Essential Trace Element for Assembly of Collagen IV Scaffolds in Tissue Development and Architecture. *Cell* **2014**, *157* (6), 1380−1392.

(24) Cummings, C. F.; Pedchenko, V.; Brown, K. L.; Colon, S.; Rafi, M.; Jones-Paris, C.; Pokydeshava, E.; Liu, M.; Pastor-Pareja, J. C.; Stothers, C.; Ero-Tolliver, I. A.; McCall, A. S.; Vanacore, R.; Bhave, G.; Santoro, S.; Blackwell, T. S.; Zent, R.; Pozzi, A.; Hudson, B. G.

Extracellular chloride signals collagen IV network assembly during basement membrane formation. *J. Cell Biol.* **2016**, *213* (4), 479−94.
(25) Fidler, A. L.; Vanacore, R. M.; Chetyrkin, S. V.; Pedchenko, V. K.; Bhave, G.; Yin, V. P.; Stothers, C. L.; Rose, K. L.; McDonald, W. H.; Clark, T. A.; et al. A unique covalent bond in basement membrane is a primordial innovation for tissue evolution. *Proc. Natl. Acad. Sci.* **2014**, *111* (1), 331−336.
(26) Pichierri, F. Theoretical characterization of the sulfilimine bond: Double or single? *Chem. Phys. Lett.* **2010**, *487* (4−6), 315−319.
(27) Ončák, M.; Berka, K.; Slavíček, P. Novel covalent bond in proteins: calculations on model systems question the bond stability. *ChemPhysChem* **2011**, *12* (17), 3449−3457.
(28) Roy, A.; Alnakhli, T. H.; Gauld, J. W. Computational insights into the formation and nature of the sulfilimine bond in collagen-IV. *RSC Adv.* **2022**, *12*, 21092−21102.
(29) Bax, B.; Chung, C.-w.; Edge, C. Getting the chemistry right: protonation, tautomers and the importance of H atoms in biological chemistry. *Acta Crystallogr. Sec D* **2017**, *73* (2), 131−140.
(30) Silverstein, T. P. The Proton in Biochemistry: Impacts on Bioenergetics, Biophysical Chemistry, and Bioorganic Chemistry. *Front Mol. Biosci* **2021**, *8*, 764099.
(31) Wong, S. S.; Wong, L. J. Chemical crosslinking and the stabilization of proteins and enzymes. *Enzyme Microb Technol.* **1992**, *14* (11), 866−74.
(32) Pikkemaat, M. G.; Linssen, A. B.; Berendsen, H. J.; Janssen, D. B. Molecular dynamics simulations as a tool for improving protein stability. *Protein Eng.* **2002**, *15* (3), 185−92.
(33) Ali, S. A.; Hassan, M. I.; Islam, A.; Ahmad, F. A review of methods available to estimate solvent-accessible surface areas of soluble proteins in the folded and unfolded states. *Curr. Protein Pept Sci.* **2014**, *15* (5), 456−76.
(34) Levy, Y.; Onuchic, J. N. Water and proteins: a love-hate relationship. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (10), 3325−6.
(35) Vanacore, R. M.; Friedman, D. B.; Ham, A. J.; Sundaramoorthy, M.; Hudson, B. G. Identification of S-hydroxylysyl-methionine as the covalent cross-link of the noncollagenous (NC1) hexamer of the alpha1alpha1alpha2 collagen IV network: a role for the post-translational modification of lysine 211 to hydroxylysine 211 in hexamer assembly. *J. Biol. Chem.* **2005**, *280* (32), 29300−10.
(36) *Molecular Operating Environment (MOE) 2020.09*; Chemical Computing Group ULC: Montreal, Quebec, Canada, 2021.
(37) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput* **2015**, *11* (8), 3696−3713.
(38) Bond, S. D.; Leimkuhler, B. J.; Laird, B. B. The Nosé-Poincaré method for constant temperature molecular dynamics. *J. Comput. Phys.* **1999**, *151* (1), 114−134.
(39) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log (N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98* (12), 10089−10092.
(40) Phillips, J. C.; Hardy, D. J.; Maia, J. D.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **2020**, *153* (4), 044130.
(41) Ion, B. F.; Bushnell, E. A. C.; Luna, P. D.; Gauld, J. W. A molecular dynamics (MD) and quantum mechanics/molecular mechanics (QM/MM) study on ornithine cyclodeaminase (OCD): a tale of two iminiums. *Int. J. Mol. Sci.* **2012**, *13* (12), 12994−13011.
(42) Aboelnga, M. M.; Gauld, J. W. Comparative QM/MM study on the inhibition mechanism of β-Hydroxynorvaline to Threonyl-tRNA synthetase. *J. Mol. Graph Model* **2022**, *115*, 108224.
(43) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph* **1996**, *14* (1), 33−38.
(44) DeLano, W. L. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter on protein crystallography* **2002**, *40* (1), 82−92.

https://doi.org/10.1021/acsomega.2c03360
ACS Omega 2022, 7, 39680−39689