RSC Advances

PAPER

View Article Online
View Journal | View Issue

Cite this: *RSC Adv.*, 2022, **12**, 21092

Computational insights into the formation and nature of the sulfilimine bond in collagen-IV†

Anupom Roy, Taqred H. Alnakhli and James W. Gauld *

Collagen IV is essential component of basement membrane in the tissues. It provides proper cellular structure by the formation of sulfilimine bond (S=N) between methionine and lysine or hydroxylysine (cross-links) residues which can be formed with or without post-translational modification. The sulfilimine bond has critical roles in tissue development and human diseases. Peroxidasin, a basement membrane peroxidase, generates reactive halogen species including hypobromous (HOBr) acid and hypochlorous (HOCl) acid which help to form halosulfonium or haloamine. The sulfilamine bond can be formed either by the formation of halosulfonium or by the formation of halomine. The aim of the study is the investigation of the formation of sulfilimine bond and its nature in collagen IV using multi-scale approach that included MD, QM-cluster, systematic series of small models, and NBO analysis. These results suggest that sulfilimine bond can be formed either *via* brominated/chlorinated halosulfonium or haloamine pathway. The results of systematic series of small model indicate that the formation of sulfilimine complex from halosulfonium happens through the formation of positively charged halosulfonated sulfilimine complex. It also suggests that the formation of sulfilimine complex from haloamine occurs through the formation of positively charged sulfilimine complex where the S and N bond forms and halogen goes off at the same time. Furthermore, the NBO analysis suggest the S and N bond is strongly polarized toward nitrogen in both single protonated and neutral system, $N^{\delta-} \leftarrow S^{\delta+}$ and also indicate the existence of a coordinate covalent (*i.e.* dative) bond.

Received 1st April 2022
Accepted 18th July 2022

DOI: 10.1039/d2ra02105f

rsc.li/rsc-advances

1 Introduction

The basement membrane is a thin, highly specialized extracellular matrix that is essential for the proper structure and function of tissues.[1] In particular, it provides structural support for epithelial cells and plays important roles in, for example, intracellular signal transduction and cell behavior regulation.[1,2] Its importance is further underscored by the fact that its improper functioning, due to for instance genetic mutations, can result in a number of severe or even fatal diseases or conditions in organisms.[3] However, despite its physiological importance much remains unclear or even unknown including its structure.

For instance, the most abundant molecular components of basement membranes are the collagen proteins of which there are least 28 sub-types.[2,4] In particular, type IV Collagen (Collagen IV), discovered in 1966,[5] is the most abundant molecular species of the basement membrane and is a critical component of the lamina densa,[2] a sub-layer of the basement membrane.[6] In particular, the collagen IV forms an essential and complex network that provides a supportive scaffold for other components of the basement membranes.[7] Improper structures and/or defects in collagen IV are associated with several diseases including rheumatological and dermatological diseases.[8,9] Furthermore, it is believed that the potentially fatal Goodpasture syndrome, a rare autoimmune disease, is thought to result from antibodies attacking certain alpha subunits within collagen IV in the basement membranes of the renal glomeruli and pulmonary alveoli.[9] Another well-known collagen IV related pathology is the genetic disorder Alport's syndrome, caused by mutations within collagen IV, which is associated with proteinuria, progressive renal failure, ocular abnormalities, and hearing loss.[9] In addition, it has been reported that collagen IV also has critical roles in wound healing[10] and embryogenesis.[11]

While the exact structure of collagen IV remains unclear, it is known to be composed of six highly homologous polypeptide chains; α1(IV) to α6(IV), with each comprising of an (i) amino-terminal; (ii) triple-helical, and (iii) carboxy-terminal domain. Three α-chains bind together to form a triple helix, a protomer, which then aggregate end to end to form a hexamer and ultimately collagen fibrils.[9,12,13] It is known that collagen IV can undergo several post-translational modifications and that protomers covalently cross-link through a variety of enzymatic and non-enzymatic processes. While these are often to help provide

Department of Chemistry and Biochemistry, University of Windsor, Windsor, Ontario N9B 3P4, Canada. E-mail: gauld@uwindsor.ca

† Electronic supplementary information (ESI) available. See https://doi.org/10.1039/d2ra02105f

the structural integrity of the collagen IV network, some cross-links have also been associated with aging and disease.[14] However, in 2009 it was discovered that covalent cross-links could also form between an hydroxylysine (Hyl) or lysine (Lys) residue in one chain with a Methionine (Met) residue in another.[15] It is noted that this also indicates that post-translational modifications are not essential for the formation of these cross-links.[15] More specifically, this cross-linking involved the side-chain amine of Hyl/Lys and the sulfur of Met and led to the formation of a sulfilimine (S=N) crosslink. Furthermore, such inter-residue bonds are rare in biochemistry; so far the only known natural occurrence is in collagen IV.[15] Furthermore, it is held that the sulfilimine bond of collagen IV was essential for evolution of tissues and eumetazoa.[16]

Recently, experimental studies have suggested that the sulfilimine bond formation in collagen IV may be induced by the hypohalous acids, hypobromous (HOBr) and hypochlorous (HOCl) acid, though their exact roles remain unclear.[1,17] Furthermore, the efficiency of HOBr facilitated sulfilimine bond formation is 50 000-fold greater than for HOCl.[1] These acids are produced by peroxidasin, a multifunctional extracellular matrix-associated member of the heme peroxidase family.[18] Based on experimental observations it has been proposed that hypohalous acid (HOX; X = Br, Cl) mediated sulfilimine bond formation can proceed *via* either a halosulfonium or halomine intermediate as shown in Fig. 1.[15,19] More specifically, the HOX acid may react with the side chain sulfur centre of Met93 to form a halosulfonium cation. This then reacts with the side chain amine nitrogen of Hyl, and with loss of HX forms a neutral sulfilimine crosslink between Met93 and Hyl211 (or Lys211). Alternatively, if the side chain amine of Hyl211 (or Lys211) is neutral, the HOX may react with to form a haloamine. This then reacts with the sulfur of Met93 to form the neutral sulfilimine crosslink. Of these two possible reaction pathways, the halosulfonium pathway is the major pathway for the formation of sulfilimine bond.[1] While the neutral sulfilimine bond is represented as a covalent double bond, a previous computational study concluded that it is more accurately described as a polarized coordinate-covalent (dative) single bond; $N^{\delta-} \leftarrow S^{\delta+}$.[20] Meanwhile, another computational study concluded that the sulfilimine bond is protonated at pH 7.[21]

In this present study we have applied computational chemistry methods to examine the proposed mechanisms by which the sulfilimine bond formation in collagen IV is mediated by the hypohalous acids HOX (X = Cl and Br). More specifically, using molecular dynamics (MD) simulations, quantum mechanical (QM)-cluster methods to examine possible mechanisms for formation of possible sulfilimine crosslinks. In addition, we have applied Natural Bond Orbital (NBO) analyses to gain further insights into the nature of possible intermediates, as well as the nature of the product sulfilimine in the basement membrane.

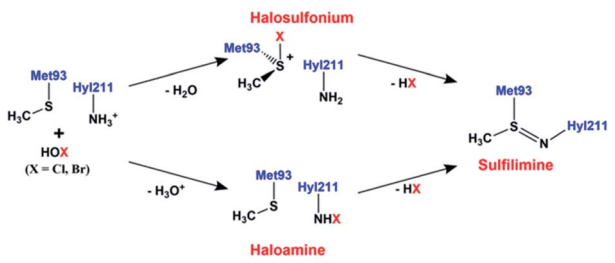

Fig. 1 The proposed[15,19] mechanism for the formation of the sulfilimine bond in collagen IV.

2 Computational methods

2.1 Molecular dynamic simulations (MD)

A suitable X-ray crystal structure of collagen IV, containing the sulfilimine Met–Lys crosslink, was obtained from the Protein Data Bank (PDB ID: 1M3D).[22] While the structure has 6 similar chains, only chains B and D were used in the subsequent MD simulations as these two are involved in the crosslink. The structure was prepared for MD simulations using the Molecular Operator Environment (MOE).[23] All parameterization, solvation, minimization, heating, equilibration and MD production runs were performed using AmberTools 16 and the AMBER 16 program.[24] First, the prepared initial structure was solvated using TIP3P water model in which the structure was surrounded by a water layer of 10 Å deep with periodic boundary conditioned enabled. Then, the solvation was followed by minimization using the molecular mechanics (MM) force field AMBER ff14SB.[25] The minimization was performed in two steps: (1) first hydrogen atoms on all heavy atoms were minimized, and then (2) the protein backbone and all heavy atoms using SHAKE algorithm. The resulting minimized system was gradually heated to 300 K using Anderson-like temperature coupling scheme (ntt = 2) at a constant pressure with periodic boundary conditions (ntp = 1) with a time step of 2 fs for 100 ps. Then, the heated system was submitted for equilibration and production run using NVT ensemble with constant temperature (300 K) and constant volume (ntb = 1) with a time step of 2 fs for 50 ns that includes 100 ps equilibration. The Particle Mesh Ewald (PME) method[26] was used for calculating coulombic interactions, with cutoffs for nonbonded long-range interactions set to 8 Å (the default). The generated conformations from the MD simulation were analyzed based on backbone atoms' root mean square deviations (RMSD) using AmberTools 16.[24] Then, the obtained RMSD values were clustered, and the structure with the shortest $_{Met93}S\cdots N_{Lys/Hyl211}$ distance was chosen for the QM-cluster calculations. All parameterization, solvation, minimization, heating, equilibration and MD production runs were performed using AmberTools 16 and the AMBER 16 program.[24] First, the prepared initial structure was solvated using TIP3P water model in which the structure was surrounded by a water layer of 10 Å deep with periodic boundary conditioned enabled. Then, the solvation was followed by minimization using the molecular mechanics (MM) force field AMBER ff14SB.[25] The minimization was performed in two steps: (1) hydrogen atoms on all heavy atoms were minimized, and then (2) the protein backbone and all heavy atoms using the SHAKE algorithm. The resulting minimized system was gradually heated to 300 K using Anderson-like temperature coupling scheme (ntt = 2) at a constant pressure with periodic boundary conditions (ntp = 1)

with a time step of 2 fs for 100 ps. The system was then submitted for equilibration and production using NVT ensemble with constant temperature (300 K) and volume (ntb = 1) and a time step of 2 fs for 50 ns that includes 100 ps equilibration. The Particle Mesh Ewald method[26] was used for calculating coulombic interactions, with cutoffs for nonbonded long-range interactions set to 8 Å (the default). The generated conformations from the MD simulation were analyzed based on backbone atoms' root mean square deviations (RMSD) using AmberTools 16.[24] The obtained RMSD values were clustered, and the structure with the shortest $_{Met93}S\cdots N_{Lys/Hyl211}$ distance was chosen for the QM-cluster calculations.

2.2 Quantum Mechanics (QM)-cluster and systematic series of small model studies

The Gaussian 09 and Gaussian 16 programs were used for all QM-cluster and small model calculations.[27,28] QM-cluster optimization and harmonic frequency calculations were performed using M06-2X/6-31G(d,p) level of theory while the surrounding protein and solvent environment was modelled using the IEF-PCM method (ε = 78.35 for water or 4.0 for protein). The M06-2X functional was chosen because of its purported greater reliability and accuracy for main group thermochemistry and description of non-covalent long-range interactions.[29,30] The results were compared where possible with those obtained at the QCISD/6-31G(d,p) or/6-311+G(2df,p) levels of theory (see text). The QM-cluster models for the formation of the haloSulfonium and haloamine intermediates consisted of 30 and 33 atoms in total, respectively. More specifically, the clusters included truncated sidechains of Met93, Lys211/Hyl211, and Glu214 (that was observed in the MD simulations to be spatially close), and 1 or more water molecules. It should be noted that Met93, Lys211/Hyl211, and Glu214 were modelled by dimethylsulfide ($S(CH_3)_2$), methylamine (CH_3NH_2), and ethanoate/ethanoic acid (CH_3COO^-/CH_3COOH), respectively. Single point energies were obtained at M06-2X/6-311+G(2df,p) level of theory, on the above geometries, with the solvent environment modelled using the IEF-PCM method. Gibb's Free Energy corrections were obtained at the M06-2X/6-31G(d,p) level of theory.

2.3 Natural Bond Orbital (NBO) analysis

The natural bond orbital (NBO) analysis was performed using the NBO program (NBO Pro 7) with the method of Weinhold and coworkers,[31] and Gaussian 09 program.[27] The NBO analyses were performed at the M06-2X/6-311G+(2df,p)//M06-2X/6-31G(d,p) level of theory.

3 Results and discussion

3.1 Molecular Dynamics (MD) results

As mentioned in the previous Computational methods section, a 50 ns MD simulation was first performed. This assisted examination of, for instance, key distances and interactions between residues, especially those involving the sidechains of Met93 and Lys211/Hyl211 in adjacent strands, as well as water availability around these side chains. It is noted that the side chain of Lys211/Hyl211 was protonated in the MD simulations.

The distances between the side chain S of Met93 (S_{Met93}) and side chain N of Lys211/Hyl211 ($N_{Lys211/Hyl211}$), in the absence of the hypohalous acids HOBr or HOCl, are shown in Fig. 2. Over the course of the MD simulation the average $_{Lys211/Hyl211}NH\cdots S_{Met93}$ distance was 6.4 Å. However, during the MD simulation the distance varied and notably, shortened to less than 4.0 Å, with the shortest distance of approximately being just 3.13 Å. In these short $_{Lys211/Hyl211}NH\cdots S_{Met93}$ distanced conformations, the protonated side-chain amine of Lys211/Hyl211 forms a weak electrostatic interaction with the side-chain S of Met93. In addition, water molecules were also observed in the immediate vicinity of the Met93 and Lys211/Hyl211 sidechains. Thus, these distances indicate the possibility for weak interactions and reasonable positioning of the two mechanistically key functional groups, as well as the availability of waters, for the required subsequent reactions.

To explore possible and proposed sulfilimine bond formation reactions involving HOBr and HOCl, as noted in the computational methods section, a suitable representative structure was then selected and hypobromous acid (HOBr) and hypochlorous acid (HOCl) manually inserted and positioned in proximity to the sidechains of Met93 and Lys211/Hyl211. It is noted that except for the addition of HOBr or HOCl, and when needed a proton in the form of H_3O^+, no additional changes were made.

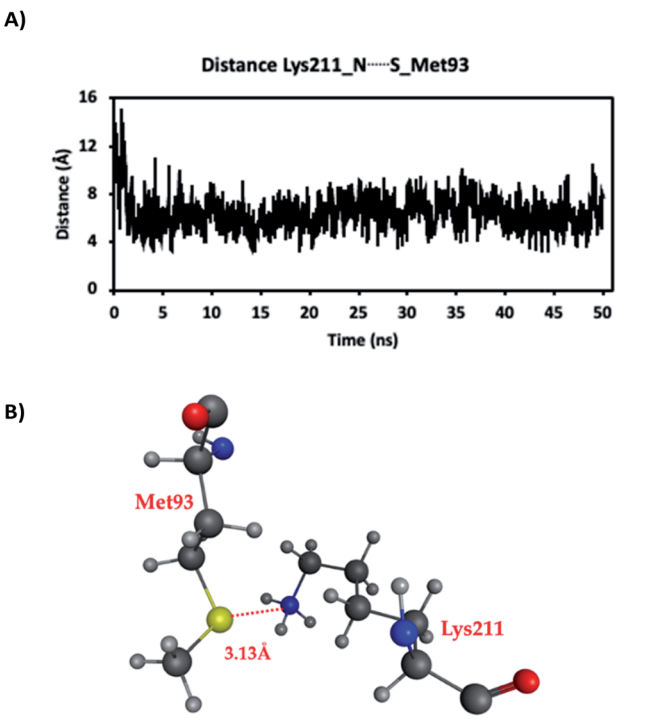

Fig. 2 (A) The distances between S_{Met93} and $N_{Lys211/Hyl211}$ during the 50 ns MD simulations, and (B) a representative structure of the conformations with the shortest $_{Met93}S\cdots N_{Lys211/Hyl211}$ distance. For clarity, all other residues and water molecules are omitted (atom colours: O = red; C = dark grey; N = blue; S = yellow; H = light grey).

3.2 QM-cluster studies on sulfilimine bond formation *via* the proposed halosulfonium or haloamine

Synthetically, the formation of sulfilimines can be achieved *via* reaction of a haloamine with a sulfide[32,33] and several possible mechanisms, based at least in part on the results of kinetic studies, have been proposed. These include nucleophilic attack of the sulfide on the haloamine nitrogen, halogenation of the sulfide by the haloamine, or oxidative addition of the haloamine to the sulfide.[32] It is noted that other possible synthetic mechanisms *via* other intermediates or related species (*e.g.*, sulfoxides) have also been suggested. However, in the mechanisms proposed[17] for collagen IV, it has been suggested that hypohalous acids produced by peroxidasin, HOCl and HOBr, could react with the sidechains of methionine (Met93) or the lysine/hydroxylysine (Lys211/Hyl211) residues to give the corresponding halosulfonium or haloamine derivatives, respectively. Subsequently, the halosulfonium could react with the sidechain amine of lysine/hydroxylysine (Lys211/Hyl211), or conversely the haloamine could react with the sidechain sulfide of Met93 to give a neutral sulfilimine crosslink. Thus, the reaction of HOCl and HOBr with the Met93 sulfur or Lys211/Hyl211 sidechain nitrogen was considered.

Gibbs free energies were obtained for the formation of the halosulfonium or haloamine, using the chemical models described in the Computational methods section, are provided in the Tables 1 and 2, and the energy diagrams are shown in the ESI (Fig. S2–S9†). Herein, only key features of the PESs obtained are described. It is noted that for formation of the halosulfonium and haloamine the sidechain of Glu214 was neutral (*i.e.*, CH_3COOH). In both cases this resulted from addition of a H_3O^+ which was needed to neutralize the hydroxide (HO^-) formed during the reactions.

3.2.1 Halosulfonium formation. For both HOBr and HOCl these reactions were found to occur in one step. More specifically, the halogen of HOX (X = Br, Cl) nucleophilically attacks the sulfur of Met93. Concomitantly, a proton is effectively transferred from the sidechain carboxylic acid of Glu214, *via* the sidechain of Lys211/Hyl211 and a bridging water molecule, onto the leaving hydroxide HO^- of HOX (Fig. S2 and S3†). A halogen (X) dependent difference in reaction barrier heights was observed. For HOBr the barrier was calculated to be 3.0 kJ mol^{-1}, while for HOCl the barrier was decidedly higher at 41.1 kJ mol^{-1}. Furthermore, the brominated product ($(CH_3)_2SBr^+$) is calculated to lie −3.6 kJ mol^{-1} lower in energy than the initial reactants (*i.e.*, the reaction is exothermic), while the chlorinated product ($(CH_3)_2SCl^+$) is calculated to lie same in energy than the initial reactants by 0.0 kJ mol^{-1} (*i.e.*, the reaction is endothermic). Importantly, however, both reactions are calculated to be thermodynamically feasible under biologically relevant aqueous conditions. Structurally, in both resulting products the halogen is positioned perpendicular to the C–S–C plane. In addition, in the bromo halosulfonium the $_{Met93}$S–Br bond length is 2.23 Å while in the chloro halosulfonium the $_{Met93}$S–Cl bond length is 2.05 Å indicating formation of strong S–X (X = Br, Cl) bonds (Fig. S2 and S3†).

3.2.2 Haloamine formation. As noted, an alternative pathway has also been proposed for the formation of the sulfilimine crosslink involving intermediate formation of a haloamine. In this pathway the Lys211/Hyl211 sidechain amine must start as neutral to provide an electron lone pair on the nitrogen for the hypohalous acid to attack. Like that observed for halosulfonium formation, haloamine formation occurs in one step and involves nucleophilic attack of the hypohalous halogen at the amine nitrogen. Like before, this attack occurs with concomitant transfer of a proton from Glu214 to the leaving hydroxide of the hypohalous acid *via* two bridging water molecules. Using the present models, the reaction barriers for formation of the bromo and chloro haloamine's were 52.0 and 29.8 kJ mol^{-1} respectively (Fig. S4 and S5†). Thus, as for halosulfonium formation, formation of haloamine is predicted to be feasible in a biologically relevant aqueous environment. Both the bromo and chloro haloamines are predicted to be lower in energy than the initial reactants by 51.2 and 95.5 kJ mol; that is, both reactions are predicted to be exothermic. Structurally, the $_{Lys211/Hyl211}$N–Br and $_{Lys211/Hyl211}$N–Cl bond lengths in the

Table 1 Gibbs free energies (kJ mol^{-1}) obtained using QM-cluster (see Methods) for the formation of the proposed halosulfonium and haloamine intermediates from hypobromous and hypochlorous acid

Reactions	Reactants	$\Delta G^{\ddagger}$ (kJ mol^{-1})	ΔG^0 (kJ mol^{-1})
Halosulfonium by HOBr	$S(CH_3)_2 + CH_3NH_2 + CH_3COOH + H_2O + HOBr$	3.0	−3.6
Halosulfonium by HOCl	$S(CH_3)_2 + CH_3NH_2 + CH_3COOH + H_2O + HOCl$	41.1	0.0
Haloamine by HOBr	$S(CH_3)_2 + CH_3NH_3 + CH_3COO^- + 2H_2O + HOBr$	52.0	−51.2
Haloamine by HOCl	$S(CH_3)_2 + CH_3NH_3 + CH_3COO^- + 2H_2O + HOCl$	29.8	−95.5

Table 2 Gibbs free energies (kJ mol^{-1}) obtained using QM-cluster (see Methods) for the formation of the proposed sulfilimine bond from a halosulfonium or haloamine intermediate

Reactions	Reactants	$\Delta G^{\ddagger}$ (kJ mol^{-1})	ΔG^0 (kJ mol^{-1})
Sulfilimine *via* bromo halosulfonium	$BrS(CH_3)_2 + CH_3NH_2 + H_2O$	118.2	−28.0
Sulfilimine *via* chloro halosulfonium	$ClS(CH_3)_2 + CH_3NH_2 + 3H_2O$	15.9	−29.9
Sulfilimine *via* bromo haloamine	$S(CH_3)_2 + CH_3N(H)Br + H_3O^+$	108.9	−96.8
Sulfilimine *via* chloro haloamine	$S(CH_3)_2 + CH_3N(H)Cl + H_3O^+$	62.4	−163.0

haloamine products are calculated to be 1.90 and 1.75 Å respectively, suggesting that they are reasonably strong bonds (Fig. S4 and S5†).

3.2.3 Sulfilimine bond formation. In the proposed[15,19] mechanisms the next step is reaction of the resultant halosulfonium with the neutral sidechain amine of Lys211/Hyl211 or, alternatively, reaction of the resulting haloamine with the sidechain sulfide of Met93. For both, multiple pathways by which the sulfilimine bond could be formed, including varying the number of water molecules required, were considered and key pathways are summarized in Fig. 3 and 4 (and in S6–S9†).

3.2.4 Further reaction of the halosulfonium to form a sulfilimine bond. Notably, it was found that formation of a protonated sulfilimine bond-containing species could occur in one step. More specifically, the Lys211/Hyl211 sidechain nitrogen can attack the halosulfonium's sulfur *trans* to the halo group which results in the alter group leaving as a halide. Concomitantly, however, a proton transfers from the attacking amine *via* at least one water molecule to the halosulfonium's leaving halide to form HX (X = Br or Cl). When the leaving halide is bromide three water molecules were required to transfer the proton from the attacking amine to the leaving bromide. In contrast, when the leaving halide was chloride, only one water was required. This reflects at least in part the larger size of the bromide compared to chloride, and differences in the length of the breaking S···X (X = Br, Cl) bond in the transition structure. The calculated reaction barriers for formation of the protonated sulfilimine species *via* the bromo and chloro halosulfonium were 118.2 and 15.9 kJ mol^{-1}, respectively. Thus, like that noted above for halosulfonium formation, a marked halogen-dependence of the calculated reaction barriers for formation of the protonated sulfilimine formation was observed. Importantly, however, both reaction barriers are thermodynamically feasible. Furthermore, the resulting protonated sulfilimine products lie 28.0 and 29.9 kJ mol^{-1} lower in energy than the initial bromo- and chloro-containing reactants, respectively. That is, formation of the protonated sulfilimine is exothermic *via* either the bromo- or chloro halosulfonium.

Building upon these results, we also examined the possible stepwise formation of the protonated sulfilimine product. It was found that it was possible to form an initial complex between the halosulfonium and amine, with the nitrogen complexed to the sulfur *trans* to the halo (Fig. 3). The resulting cationic bromo- and chloro-containing complexes were calculated to be only slightly higher in free energy than the initial reactants by 11.4 and 8.0 kJ mol^{-1}, respectively. Subsequent loss of a proton, that is formation of a neutral halo-containing species did not produce a stable species. Rather, loss of a halide to formally give the double protonated sulfilimine resulted in a species lying only 72.4 and 76.8 kJ mol^{-1} higher in free energy than the initial bromo- and chloro-containing reactants, respectively (see Fig. 3). Subsequent loss of a proton to the aqueous solution to give the single protonated sulfilimine species could then occur.

Sulfilimine bond formation *via* the haloamine intermediate was also examined. Again, it was found that the single protonated sulfilimine bond containing species described above, could be formed in one step from the initial reactants. More specifically, the sulfur of Met93 and sidechain haloamine nitrogen of Lys211/Hyl211 react directly to form an S–N interaction. Importantly, this bond formation occurs concomitantly with loss of the halide as X$^-$ from the haloamine component.

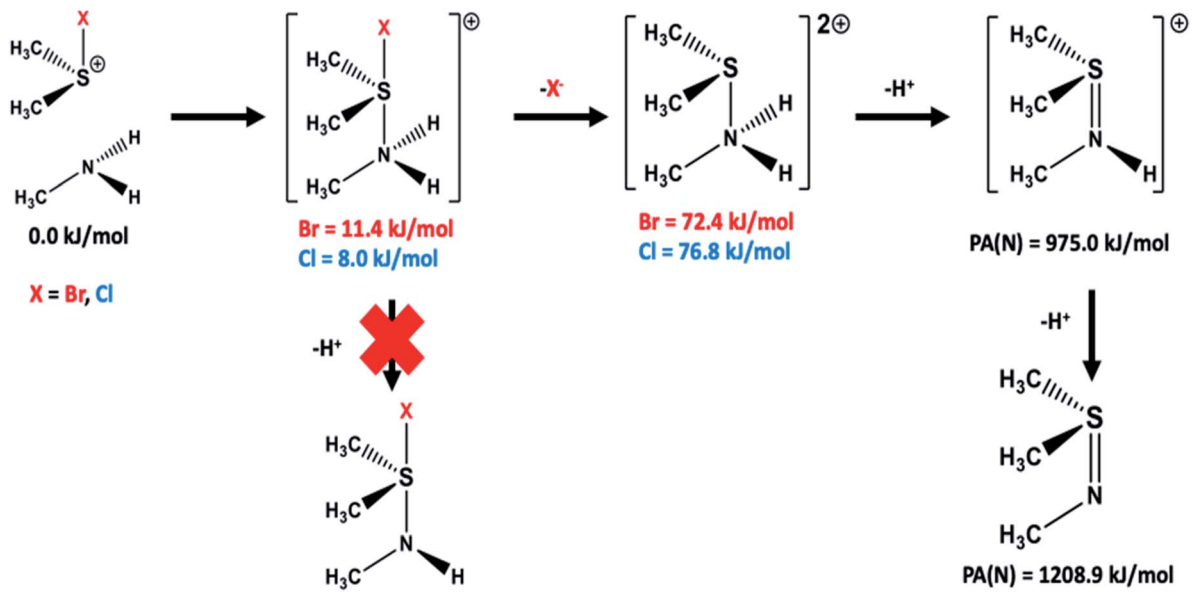

Fig. 3 Schematic illustration of the reaction pathways examined for formation of the sulfilimine crosslink by reaction of the Met93 derived halosulfonium cation ($(CH_3)_2SX^+$; X = Br, Cl) with the sidechain amine of Lys211/Hyl211 (CH_3NH_2). Gibb's free energies relative to the initial reactants are shown in red (bromo derived species) or blue (chloro derived species), while proton affinities of the sulfilimine nitrogen (PA(N)) in the neutral and single protonated species are also given. All energies are in kJ mol^{-1} and were obtained at the IEFPCM(water)-M06-2X/6-311+G(2df,p)//IEFPCM(water)-M06-2X/6-31G(d,p) with Gibb's Free Energy Corrections (see Computational methods).

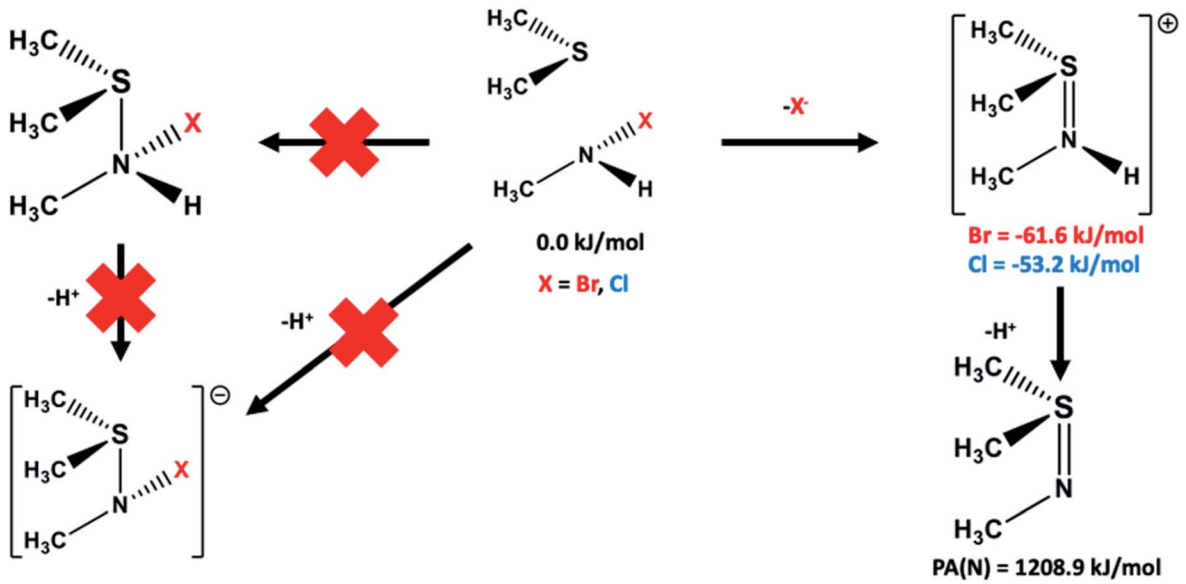

Fig. 4 Schematic illustration of the reaction pathways examined for formation of the sulfilimine crosslink by reaction of the Lys211/Hyl211 derived haloamine's (CH_3NHX; X = Br, Cl) with the sidechain sulfide of Met93 ($(CH_3)_2S$). Gibb's free energies relative to the initial reactants are shown in red (bromo derived species) or blue (chloro derived species), while proton affinities of the sulfilimine nitrogen (PA(N)) in the neutral and single protonated species are also given. All energies are in kJ mol^{-1} and were obtained at the IEFPCM(water)-M06-2X/6-311+G(2df,p)//IEFPCM(water)-M06-2X/6-31G(d,p) with Gibb's Free Energy Corrections (see Computational methods).

Within the present models used this also required the presence of a hydronium ion (H_3O^+) which gives a proton to the leaving halide group to form HX (X = Br, Cl). The calculated reaction barrier heights for formation of the single protonated sulfilimine product (see Fig. S8 and S9†) were 108.9 and 62.4 kJ mol^{-1} for the reactions involving the bromo- and chloro haloamine's, respectively. While these are both higher than those calculated using the corresponding halosulfonium cation, it is important to note that both barriers obtained using the present models are predicted to be thermodynamically feasible. Furthermore, both reactions are calculated to be exothermic with the single protonated sulfilimine species lying lower in energy than the bromo- and chloro-halomine containing reactants by 96.8 and 163.0 kJ mol^{-1}, respectively (see Fig. S8 and S9†).

Possible alternative pathways by which a bromo- or chloro-haloamine may react with a sulfide to form a sulfilimine species were also examined and are illustrated schematically in Fig. 4. In particular, the possibility of the haloamine and sulfide forming a neutral complex with an S···N interaction, or alternatively reacting with concomitant loss of a proton to form an S–N interaction containing anionic species were also considered. Neither alternate complex and pathway, at the present level of theory and with the chemical models used, was found to be feasible.

3.3 What is the likely protonated state of the sulfilimine bond-containing product?

In the feasible pathways for sulfilimine bond formation *via* either a halosulfonium reacting with a neutral alkyl amine, or a haloamine reacting with a disulfide, the resultant product was the protonated sulfilimine product (see Fig. 3 and 4). However, the sulfilimine crosslink in collagen IV is commonly represented and/or assumed to be neutral.[15,20] Thus, we further examined whether the single protonated sulfilimine may lose a proton to the aqueous environment. It is noted that other studies have suggested that at least some sulfilimines could potentially be classed as superbases with calculated proton affinities of over 1000 kJ mol^{-1}.[34] In this present study, at the IEFPCM(water)-M06-2X/6-311+G(2df,p)//IEFPCM(water)-M06-2X/6-31G(d,p) corrected for Gibb's Free Energy, the proton affinity of the nitrogen centre in the sulfilimine bond is 1208.9 kJ mol^{-1}. This is in general agreement with calculated values for other sulfilimines. Importantly, it is significantly higher than the proton affinity of water calculated at the same level of theory (989.1 kJ mol^{-1}). Meanwhile, the proton affinity of the nitrogen centre in the single protonated sulfilimine bonded species is calculated to be slightly lower than that of water at 975.0 kJ mol^{-1}. Thus, the present results indicate that under expected biological conditions the sulfilimine crosslink is most likely single protonated, in agreement with a previous computational study,[21] but has the potential to vary its protonation state; but it is unlikely neutral.

The protonation state of the sulfilimine bond's nitrogen centre significantly influences the length of the sulfilimine bond. For the current chemical models, the length of the sulfilimine bond in the optimized structures of the neutral, single- and double protonated nitrogen species are 1.606, 1.661, and 1.788 Å, respectively. That is, while the above proton affinities indicate that the single protonated species is more common *in vivo*, the neutral sulfilimine has a markedly shorter and stronger bond. We also optimized the structures of all three sulfilimine species at the considerably higher QCISD/6-311+G(2df,p) level of theory and determined the proton affinity

of the nitrogen centre in the neutral and single protonated species (Table S1†). The optimized sulfilimine bond lengths obtained for the neutral and single- and double protonated nitrogen species of 1.599, 1.650, and 1.778 Å, respectively, in agreement with the trends observed when using the above M06-2X DFT-based approaches. Similar trends in the calculated proton affinities and comparison with that of water were also observed. In particular, at the IEFPCM(water)-QCISD/6-311+G(2df,p) level corrected for Gibbs Free Energy, the proton affinities of water is calculated to be 993.8 kJ mol^{-1}, while that of the neutral sulfilimine is much higher at 1218.4 kJ mol^{-1}, and that of the single protonated sulfilimine is slightly lower than that of H_2O by 7.6 kJ mol^{-1} at 986.2 kJ mol^{-1}.

3.4 Natural Bond Orbital (NBO) analyses

To gain further insights into the nature of the sulfilimine bond in the neutral, single- and double protonated species, they were examined using NBO. Pichierri[20] has previously used NBO to characterize the nature of a neutral sulfilimine bond as proposed for collagen IV. For completeness, we considered all possible conformers and isomers of the neutral, single- and double protonated sulfilimine chemical species (Fig. 5). Specifically, there were two conformers for each species as the methyl group on the nitrogen can be on the same or opposite side as the sulfur's methyl groups. It is noted that thermodynamically the difference in conformers was not significant within the computational models used. In addition, there are also two possible isomers for each single protonated conformer Key NBO results are summarized in Table 1.

The NBO method analyzes the wavefunction in terms of localized electron pairs (*i.e.*, natural orbitals) and provides Lewis-type descriptions for molecules in terms of two-electron two-center (σ, π) bonds and lone-pairs (n). The σ_{NS} (orbital hybridization) NBO of the conformers 1 and 2 are shown in Table 3.

3.4.1 NBO analyses of conformer 1. In the double protonated system, the σ_{NS} is characterized as being formed from the interaction of a $sp^{3.90}d^{0.01}$ hybrid (79.40% p-character) on nitrogen with a $sp^{8.90}d^{0.13}$ hybrid (88.65% p-character and 0.07% of d-character) orbital on sulfur (Table 3). The two single protonated isomers show very similar, but not the same, characteristics. More specifically, for isomers 1 and 2, σ_{NS} is formed from a $sp^{4.39}d^{0.03}$ hybrid (81.07% p-character and 0.01% d-character) and $sp^{3.30}d^{0.02}$ hybrid (76.36% p-character and 0.02% d-character) orbitals on the nitrogen, respectively. These respectively interact with hybrid $sp^{6.03}d^{0.07}$ (isomer 1: 84.87% p-character and 0.04% d-character) and $sp^{4.89}d^{0.07}$ (isomer 2: 82.04% p-character and 0.04% d-character) orbitals on the sulfur. Meanwhile, in the neutral species the σ_{NS} is formed from the interaction of a hybrid $sp^{4.04}$ $d^{0.04}$ (79.50% p-character and 0.03% d-character) orbital on nitrogen with a $sp^{3.62}d^{0.04}$ hybrid (77.65% p-character and 0.04% of d-character) orbital on sulfur. As can be seen in Table 3, as the protonation state of the sulfilimine bond systematically decreases from double to neutral, the contribution from the nitrogen decreases while that of the sulfur increases by a slightly greater amount.

Table 3 Summary of key results of the NBO analyses for the orbital hybridization and σ_NS (sigma bond) characterization of the neutral, and single and double protonated derivatives of Conformer's 1 and 2 shown in Fig. 5. *P*(N) indicates the protonation state of N in sulfilimine complex while 11 and 12 denote isomer 1 and 2 of the single-protonated sulfilimine species, respectively

P(N)	σ_{NS}
Conformer 1	
2	$\sigma_{NS} = 0.8169(sp^{3.90}d^{0.01})_N + 0.5768(sp^{8.90}d^{0.13})_S$
1^1	$\sigma_{NS} = 0.7720(sp^{4.39}d^{0.03})_N + 0.6356(sp^{6.03}d^{0.07})_S$
1^2	$\sigma_{NS} = 0.7814(sp^{3.30}d^{0.02})_N + 0.6240(sp^{4.89}d^{0.07})_S$
0	$\sigma_{NS} = 0.7513(sp^{4.04}d^{0.04})_N + 0.6599(sp^{3.62}d^{0.04})_S$
Conformer 2	
2	$\sigma_{NS} = 0.8166(sp^{3.94}d^{0.01})_N + 0.5772(sp^{8.91}d^{0.13})$
1^1	$\sigma_{NS} = 0.7822(sp^{3.23}d^{0.02})_N + 0.6230(sp^{4.80}d^{0.07})_S$
1^2	$\sigma_{NS} = 0.7822(sp^{3.23}d^{0.02})_N + 0.6230(sp^{4.80}d^{0.07})_S$
0	$\sigma_{NS} = 0.7559(sp^{3.48}d^{0.04})_N + 0.6547(sp^{3.11}d^{0.04})_S$

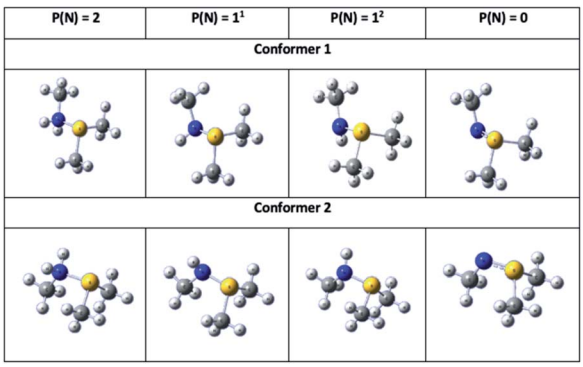

Fig. 5 Illustration of the conformers and isomers of the neutral, and single- and double-protonated sulfilimine species considered herein for NBO analysis. *P*(N) indicates the protonation state of N in sulfilimine complex while 1^1 and 1^2 denote isomer 1 and 2 of the single-protonated sulfilimine species, respectively. (Atom colours: N = blue; S = yellow; C = grey; H = white).

3.4.2 NBO analyses of conformer 2. The alternate conformer 2 shows similar trends and relative contributions to those described above for conformer 1, thus are not described in detail again here. As for conformer 1, a decreasing the protonation state of the sulfilimine bond from double to single to neutral, concomitantly decreases and increases the contribution from the nitrogen and sulfur respectively. Unlike conformer 1, however, there is negligible difference between the nitrogen and sulfur contributions in isomer 1 and 2, and no difference in the hybridization within the orbitals contributing to form the σ_{NS} bond (Table 3). Furthermore, the results for all conformer 1 and 2 species indicate that the sigma bond between N and S arises from the interaction of sp^3 hybrid orbitals on sulfur and nitrogen, with negligible sulfur d-character contributions, except in the double protonated systems. In addition, there is only a small difference in the s- and p-character observed in the NBOs of the two conformers for all protonation states considered (Fig. 6).

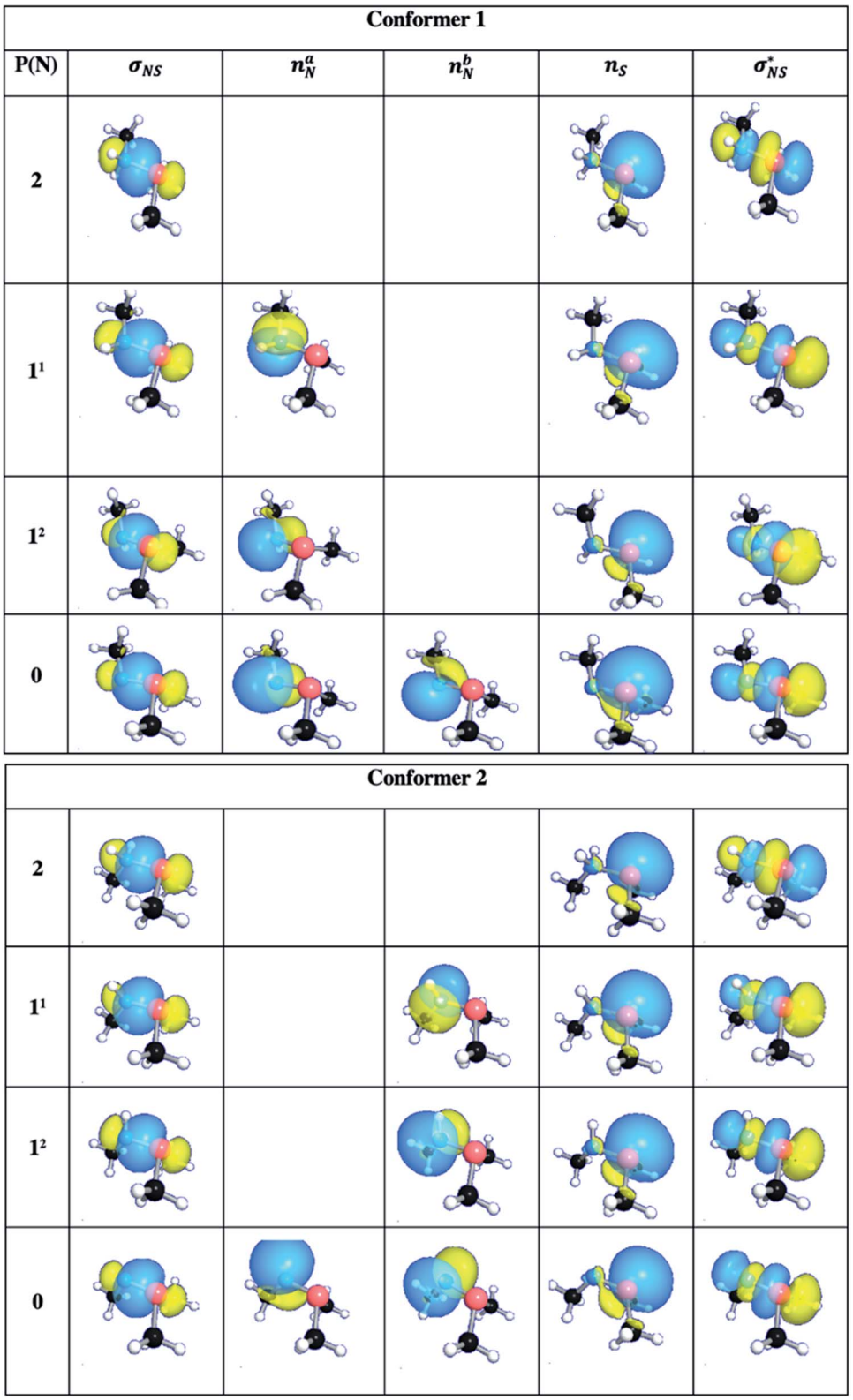

Fig. 6 Natural bond orbitals obtained for conformers 1 and 2 showing 2e–2e σ-type covalent bond between N and S as well as the lone-pairs on sulfur (n_S) and nitrogen (n_N^a and n_N^b). The σ_{NS} and σ^*_{NS} represent bonding and antibonding orbitals, respectively. P(N) indicates the protonation state of N in sulfilimine complex while 1^1 and 1^2 denote isomer 1 and 2 of the single-protonated sulfilimine species, respectively. (Atom colours: N = blue; S = rose; C = black; H = white).

The NBO analyses also provided descriptions of the electron lone pairs on nitrogen and sulfur, and additional bonding and antibonding orbitals which are illustrated in Fig. S6.† Notably for both conformers there is a lone pair on sulfur (n_S), which remained unprotonated in all possible species considered herein. Meanwhile, the nitrogen centre in both neutral conformers has two lone pairs (n_N^a and n_N^b) which are the sites of the first and second protonations (described above).

NBO analyses also provide values for the stabilization energy (E) associated with interaction(s) between filled and empty NBOs.[31] For both double protonated conformers the highest stabilization energies of 4.78 and 6.19 kcal mol^{-1} were obtained for the CH(9) $\rightarrow \sigma^*_{SN}$ of conformer 1 and CH(12) $\rightarrow \sigma^*_{SN}$ of conformer 2 interactions, respectively. For single protonated conformers and isomers, for conformer 1 stabilization energies of 5.95 and 9.48 kcal mol^{-1} were obtained for the $n_N^a \rightarrow \sigma^*_{CH(8)}$ and $n_N^a \rightarrow \sigma^*_{SC(11)}$ interactions in isomers 1 and 2, respectively. In contrast, for conformer 2 much closer stabilization energies of 9.35 and 9.48 kcal mol^{-1} were obtained for the $n_N^a \rightarrow \sigma^*_{SC(2)}$ and $n_N^a \rightarrow \sigma^*_{CH(13)}$ interactions in isomers 1 and 2, respectively. However, for the neutral (unprotonated) sulfilimine conformers 1 and 2 differences were observed between the origins and size of the largest stabilizing interaction(s). For conformer 1, it arose from the interaction of a lone pair on the nitrogen (n_N^a) and an antibonding σ S–C orbital (σ^*_{SC}), $n_N^a \rightarrow \sigma^*_{SC(11)}$, with a value of 23.25 kcal mol^{-1}. In contrast, for unprotonated conformer 2 the largest combined stabilization energy of 109.82 kcal mol^{-1} arose from interactions of both S–C bonding orbitals (SC) and the σ^*_{SN} orbital; SC(2) $\rightarrow \sigma^*_{SN}$, SC(7) $\rightarrow \sigma^*_{SN}$.

Interestingly, the NBO analyses also indicate that in both double protonated conformers 1 and 2, the S···N bond is strongly polarized towards S (*i.e.*, $N^{\delta-} \rightarrow S^{\delta+}$). However, upon loss of a proton to give the single protonated conformers the S···N bond is strongly polarized towards N (*i.e.*, $N^{\delta-} \leftarrow S^{\delta+}$). This same polarization is observed in the neutral sulfilimine conformers as previously described by Pichierri.[20] Indeed, these NBO results on all the neutral and protonated conformers support the previous conclusion of Pichierri[20] that the S···N bond in collagen IV should be described as a coordinate covalent $N^{\delta-} \leftarrow S^{\delta+}$ bond.

The charges obtained from the NBO analyses for the sulfur and nitrogen centre in each sulfilimine species (*i.e.*, conformers 1 and 2, and for all the protonation states) are given in Table 2. For both conformers, as the protonation state is systematically decreased from double to single the neutral, a corresponding decrease in the positive charge of the sulfur atom with a simultaneous increase in the negative charge on the nitrogen is observed. In the neutral system the positive and negative charge on the sulfur and nitrogen, are closest to equal in absolute value.

The dipole moments of the unprotonated (neutral), and single and double protonated sulfilimine species were also examined. The dipole moments in the double protonated species are in a distinctly different orientation to those of the single protonated and neutral sulfilimine species. Furthermore, in the neutral species the magnitude of the dipole moment is markedly higher than that of the double and single protonated species due to the changes of natural charges in S and N. This phenomenon has previously[35] been said to be a characteristic of dative bonding such as suggested for the sulfilimine bond in collagen IV (Table 4).

Table 4 Natural atomic charges were obtained from the present NBO analyses for all conformer 1 and 2 sulfilimine species considered in this present study. *P*(N) indicates the protonation state of N in sulfilimine complex while 1^1 and 1^2 denote isomer 1 and 2 of the single-protonated sulfilimine species, respectively

P(N)	Atoms & charges
Conformer 1	
2	S = 1.118, N = −0.705
1^1	S = 1.065, N = −0.831
1^2	S = 1.096, N = −0.840
0	S = 1.048, N = −1.056
Conformer 2	
2	S = 1.119, N = −0.703
1^1	S = 1.100, N = −0.840
1^2	S = 1.100, N = −0.840
0	S = 1.060, N = −1.026

4 Conclusions

We have performed a multi-scale computational investigation, using MD simulations and Quantum Mechanical (QM)-cluster methods to examine proposed mechanisms for formation of the sulfilimine bond in collagen IV between Met93 and Lys/Hyl211 residues. In addition, we have used NBO analysis to examine the nature of the possible neutral or protonated (single or double) sulfilimine bond-containing products formed.

4.1 Formation of a sulfilimine product

The present results suggest that formation of the sulfilimine crosslink *via* the proposed halosulfonium or haloamine intermediate is thermodynamically feasible. In reaction of a halosulfonium with a neutral alkyl amine, the first step involves formation of a complex in which the sulfur of the cationic halosulfonium (Met93-X^+; X = Cl, Br) interacts with the lone pair on the amine''s nitrogen (Lys/Hyl211). The halogen is then lost as a halide to give a double protonated sulfilimine species. In contrast, the first step in the formation of an S···N bond *via* reaction of a haloamine with the Met93 sulfur requires concomitant loss of the halogen as a halide. This directly gives the most likely product, the single protonated sulfilimine species.

4.2 Protonation of the sulfilimine bond

Proton affinities of the nitrogen (PA_N) in the neutral and single protonated sulfilimine species were calculated. In the neutral sulfilimine bond PA_N is >1200 kJ mol^{-1}, which is markedly higher than that of H_2O (~990 kJ mol^{-1}). In contrast, PA_N in the single protonated sulfilimine bond is slightly lower (<15 kJ mol^{-1}) than that calculated for H_2O (Table S1†). Thus, single protonated is the most likely state of the sulfilimine

crosslink *in vivo*, but conditions may be able to influence its protonation state (*e.g.*, single or double).

4.3 Nature of the sulfilimine crosslink

NBO analyses indicate that, as suggested previously for the neutral sulfilmine bond,[20] the S and N bond is strongly polarized toward nitrogen in the single protonated and neutral systems, $N^{\delta-} \leftarrow S^{\delta+}$, and that it is also well describe as a coordinate covalent (*i.e.*, dative) bond.

Author contributions

Anupom Roy: computational calculations and writing, Taqred H. Alnakhli: computational calculations of bromo halosulfonium and bromo haloamine formation, James W. Gauld: supervision, review, and editing.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

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

Notes and references

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

2 R. Kalluri, Basement membranes: structure, assembly and role in tumour angiogenesis, *Nat. Rev. Cancer*, 2003, **3**(6), 422–433.

3 M. A. Morrissey and D. R. Sherwood, An active role for basement membrane assembly and modification in tissue sculpting, *J. Cell Sci.*, 2015, **128**(9), 1661–1668.

4 J. Heino, The collagen family members as cell adhesion proteins, *Bioessays*, 2007, **29**(10), 1001–1010.

5 N. A. Kefalides, A collagen of unusual composition and a glycoprotein isolated from canine glomerular basement membrane, *Biochem. Biophys. Res. Commun.*, 1966, **22**(1), 26–32.

6 M. Paulsson, Basement-Membrane Proteins - Structure, Assembly, and Cellular Interactions, *Crit. Rev. Biochem. Mol. Biol.*, 1992, **27**(1–2), 93–127.

7 Y. Wu and G. Ge, Complexity of type IV collagens: from network assembly to function, *Biol. Chem.*, 2019, 565–574.

8 R. E. Petty, D. W. Hunt and A. M. Rosenberg, Antibodies to type IV collagen in rheumatic diseases, *J. Rheumatol.*, 1986, **13**(2), 246–253.

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

10 L. E. Tracy, R. A. Minasian and E. J. Caterson, Extracellular Matrix and Dermal Fibroblast Function in the Healing Wound, *Adv. Wound Care*, 2016, **5**(3), 119–136.

11 X. Bai, D. J. Dilworth, Y. C. Weng and D. B. Gould, Developmental distribution of collagen IV isoforms and relevance to ocular diseases, *Matrix Biol.*, 2009, **28**(4), 194–201.

12 M. D. Shoulders and R. T. Raines, Collagen structure and stability, *Annu. Rev. Biochem.*, 2009, **78**, 929–958.

13 J. Khoshnoodi, V. Pedchenko and B. G. Hudson, Mammalian collagen IV, *Microsc. Res. Tech.*, 2008, **71**(5), 357–370.

14 J. Gaar, R. Naffa and M. Brimble, Enzymatic and non-enzymatic crosslinks found in collagen and elastin and their chemical synthesis, *Org. Chem. Front.*, 2020, **7**(18), 2789–2814.

15 R. Vanacore, A. J. Ham, M. Voehler, C. R. Sanders, T. P. Conrads, T. D. Veenstra, K. B. Sharpless, P. E. Dawson and B. G. Hudson, A sulfilimine bond identified in collagen IV, *Science*, 2009, **325**(5945), 1230–1234.

16 A. L. Fidler, R. M. Vanacore, S. V. Chetyrkin, V. K. Pedchenko, G. Bhave, V. P. Yin, C. L. Stothers, K. L. Rose, W. H. McDonald and T. A. Clark, A unique covalent bond in basement membrane is a primordial innovation for tissue evolution, *Proc. Natl. Acad. Sci.*, 2014, **111**(1), 331–336.

17 G. Bhave, C. F. Cummings, R. M. Vanacore, C. Kumagai-Cresse, I. A. Ero-Tolliver, M. Rafi, J. S. Kang, V. Pedchenko, L. I. Fessler, J. H. Fessler and B. G. Hudson, Peroxidasin forms sulfilimine chemical bonds using hypohalous acids in tissue genesis, *Nat. Chem. Biol.*, 2012, **8**(9), 784–790.

18 G. Sirokmany, H. A. Kovacs, E. Lazar, K. Konya, A. Donko, B. Enyedi, H. Grasberger and M. Geiszt, Peroxidasin-mediated crosslinking of collagen IV is independent of NADPH oxidases, *Redox Biol.*, 2018, **16**, 314–321.

19 G. E. Ronsein, C. C. Winterbourn, P. Di Mascio and A. J. Kettle, Cross-linking methionine and amine residues with reactive halogen species, *Free Radical Biol. Med.*, 2014, **70**, 278–287.

20 F. Pichierri, Theoretical characterization of the sulfilimine bond: Double or single?, *Chem. Phys. Lett.*, 2010, **487**(4–6), 315–319.

21 M. Ončák, K. Berka and P. Slavíček, Novel covalent bond in proteins: calculations on model systems question the bond stability, *ChemPhysChem*, 2011, **12**(17), 3449–3457.

22 M. Sundaramoorthy, M. Meiyappan, P. Todd and B. G. Hudson, Crystal structure of NC1 domains. Structural basis for type IV collagen assembly in basement membranes, *J. Biol. Chem.*, 2002, **277**(34), 31142–31153.

23 Molecular Operating Environment (MOE), *Molecular Operating Environment*, 2020.09, Chemical Computing Group ULC, 1010 Sherbooke St, West, Suite #910, Montreal, QC, Canada, H3A 2R7, p. 2021.

24 D. Cerutti, T. Cheatham III, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, N. Homeyer, S. Izadi and P. Janowski, *Amber 2016*, University of California, San Francisco 2016.

25 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB, *J. Chem. Theory Comput.*, 2015, **11**(8), 3696–3713.

26 T. Darden, D. York and L. Pedersen, Particle mesh Ewald: An N log (N) method for Ewald sums in large systems, *J. Chem. Phys.*, 1993, **98**(12), 10089–10092.

27 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision D.01*, Gaussian, Inc., Wallingford CT, 2009.

28 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.

29 M. Walker, A. J. A. Harvey, A. Sen and C. E. H. Dessent, Performance of M06, M06-2X, and M06-HF Density Functionals for Conformationally Flexible Anionic Clusters: M06 Functionals Perform Better than B3LYP for a Model System with Dispersion and Ionic Hydrogen-Bonding Interactions, *J. Phys. Chem. A*, 2013, **117**(47), 12590–12600.

30 E. G. Hohenstein, S. T. Chill and C. D. Sherrill, Assessment of the Performance of the M05-2X and M06-2X Exchange-Correlation Functionals for Noncovalent Interactions in Biomolecules, *J. Chem. Theory Comput.*, 2008, **4**(12), 1996–2000.

31 A. E. Reed, L. A. Curtiss and F. Weinhold, Intermolecular Interactions from a Natural Bond Orbital, Donor-Acceptor Viewpoint, *Chem. Rev.*, 1988, **88**(6), 899–926.

32 T. L. Gilchrist and C. J. Moody, The chemistry of sulfilimines, *Chem. Rev.*, 1977, **77**(3), 409–435.

33 K. Tsujihara, N. Furukawa, K. Oae and S. Oae, Sulfilimine. I. synthesis and formation mechanism, *Bull. Chem. Soc. Jpn.*, 1969, **42**(9), 2631–2635.

34 E. D. Raczyńska, J.-F. Gal, P.-C. Maria and M. Szeląg, Proton Transfer Chemistry in the Gas Phase. Is a Spontaneous' Neutralization'Reaction a Myth or a Reality?, *Croat. Chem. Acta*, 2009, **82**(1), 87–103.

35 F. A. Weinhold and C. R. Landis, *Valency and Bonding: A Natural Bond Orbital Donor–Acceptor Perspective*, Cambridge University Press, New York, 2005.