Adv Exp Med Biol - Protein Reviews (2021) 21: 129–141
https://doi.org/10.1007/5584_2020_582

Published online: 27 September 2020

Collagen IV Exploits a Cl^- Step Gradient for Scaffold Assembly

Sergey V. Ivanov, Ryan Bauer, Elena N. Pokidysheva, and Sergei P. Boudko

Abstract

Collagen molecules are crucial extracellular players in animal tissue development and in functions ranging from ultrafiltration to organism locomotion. Among the 28 types of collagen found in human, type IV collagen stands out as a primordial type found in all species of the animal kingdom. Collagen IV forms smart scaffolds for basement membranes, sheet-like acellular structures that isolate, coordinate, and direct cells during morphogenesis. Collagen IV is also involved in multiple functions in developed tissues. As part of the basement membrane, collagen IV scaffolds provide mechanical strength, spatially tether extracellular macromolecules and directly signal to cells via receptor binding sites. Proper assembly and structure of the scaffolds are critical for development and function of multiple types of basement membranes. Within last 5 years it was established that Cl^- concentration is a key factor for initiating collagen IV scaffold assembly. The biological role of Cl^- in multiple physiological processes and detailed mechanisms for its signaling and structural impacts are well established. Cl^- gradients are generated across the plasma and intracellular organelle membranes. As collagen IV molecules are secreted outside the cell, they experience a switch from low to high Cl^- concentration. This transition works as a trigger for collagen IV scaffold assembly. Within the scaffold, collagen IV remains to be a Cl^- sensor as its structural integrity continues to depend on Cl^- concentration. Here, we review recent findings and set future directions for studies on the role of Cl^- in type IV collagen assembly, function, and disease.

Keywords

Basement membrane · Cl^- · Collagen IV · Extracellular matrix · Kinetics · NC1 domain · Protein assembly

Abbreviations

NC1	non-collagenous domain 1 of collagen IV
7S	7 Svedberg, the dodecameric region of collagen IV

S. V. Ivanov, R. Bauer, and E. N. Pokidysheva
Department of Medicine, Division of Nephrology and Hypertension, Vanderbilt University Medical Center, Nashville, TN, USA

Vanderbilt Center for Matrix Biology, Vanderbilt University Medical Center, Nashville, TN, USA

S. P. Boudko (✉)
Department of Medicine, Division of Nephrology and Hypertension, Vanderbilt University Medical Center, Nashville, TN, USA

Vanderbilt Center for Matrix Biology, Vanderbilt University Medical Center, Nashville, TN, USA

Department of Biochemistry, Vanderbilt University, Nashville, TN, USA
e-mail: sergey.budko@vumc.org

ECM	extracellular matrix
BM	basement membrane
GBM	glomerular BM
LBM	lens capsule BM
$NC1^{sc}$	single polypeptide chain NC1 trimer
PEG	polyethylene glycol

1 Introduction

Basement membranes (BMs) are evolutionary ancient and highly conserved sheet-like structures of extracellular matrix (ECM) that coordinate cells, direct their polarization, and play a formative role in shaping tissues (Morrissey and Sherwood 2015). In adult as all animals are multicellular, BMs serve structural and protective functions by providing signaling stimuli for cell behaviors and tissue regeneration (Jayadev and Sherwood 2017; Morrissey and Sherwood 2015). Prominent cases of highly specialized BMs are glomerular basement membrane (GBM), which provides ultrafiltration in the kidney (Gunwar et al. 1998), lens capsule basement membrane (LBM), which acts as anchor points for lens cells during migration and proliferation (Lu et al. 2008), and seminiferous tubule basement membrane, which plays a role in spermatogenesis (Kahsai et al. 1997). Type IV collagen is a structural scaffold in all BMs that integrates other ECM components, such as laminins, nidogens and proteoglycans (Fidler et al. 2017).

Genetic variants in collagen IV genes, COL4A1-6, cause BM pathologies in various tissues. Central nervous system and cerebral vasculature are primary targets of pathogenic variants in COL4A1 and COL4A2 genes, while pathogenic variants in COL4A3-5 genes almost exclusively lead to kidney diseases, hearing defects and ocular abnormalities (Cosgrove and Liu 2017; Meuwissen et al. 2015; Stokman et al. 2016; Wang et al. 2018). Whereas connection between gene variant and resulting abnormalities is firmly established in most cases, the mechanism of pathogenesis remains unclear. To provide better insight into why and how these diseases develop, it is essential to understand molecular mechanisms of collagen IV scaffold assembly.

2 Structural Organization of Collagen IV

Collagen IV was discovered in kidney GBM over 50 years ago in the seminal works of Spiro and Kefalides groups (Spiro 1967; Kefalides 1968) as a novel crosslinked molecule with unusual properties. Structural organization of the collagen IV scaffold was later analyzed using acid extraction and limited proteolysis of tumor BM, which led to a discovery of hexamer and dodecamer oligomers (Fig. 1a) (Timpl et al. 1981). Subsequently, two key structural elements responsible for forming each of these complexes, the 7S dodecamer and the NC1 hexamer (Fig. 1a), were isolated using proteolytic enzymes (Weber et al. 1984; Risteli et al. 1980; Timpl et al. 1979). These and other functional domains of collagen IV can be extracted and purified to homogeneity from different tissues, tumors, and matrix deposited by cultured cells (Boudko et al. 2018). The type IV collagen family has six genetically distinct α-chains designated α1 to α6 (Fig. 1b). Each α-chain contains an N-terminal 7S region, collagenous domain, and C-terminal NC1 domain. Initially, three α chains form collagen IV protomer that further assembles into the scaffolds composed of the hexameric and dodecameric assemblies. Studies of the NC1 hexamers extracted from tissues by treatment with collagenase (Fig. 1c) led to discoveries of three compositions of NC1 hexamers: α121, α345, and α121/α556 as well as a new type of covalent cross-link, the sulfilimine bond, stabilizing these hexamers. (Fig. 1d) (Hudson et al. 2003; Hudson et al. 1994; Borza et al. 2001; Boutaud et al. 2000; Vanacore et al. 2009).

3 Approaches to Overcome Hurdles to Study NC1 Hexamer Assembly

Collagen IV is notoriously insoluble in tissue due to extensive cross-linking, though this hurdle can be partially resolved by applying β-aminoproprionitrile, a lathrytic agent that prevents the formation of cross-links, during animal development or tumor growth. Although full-

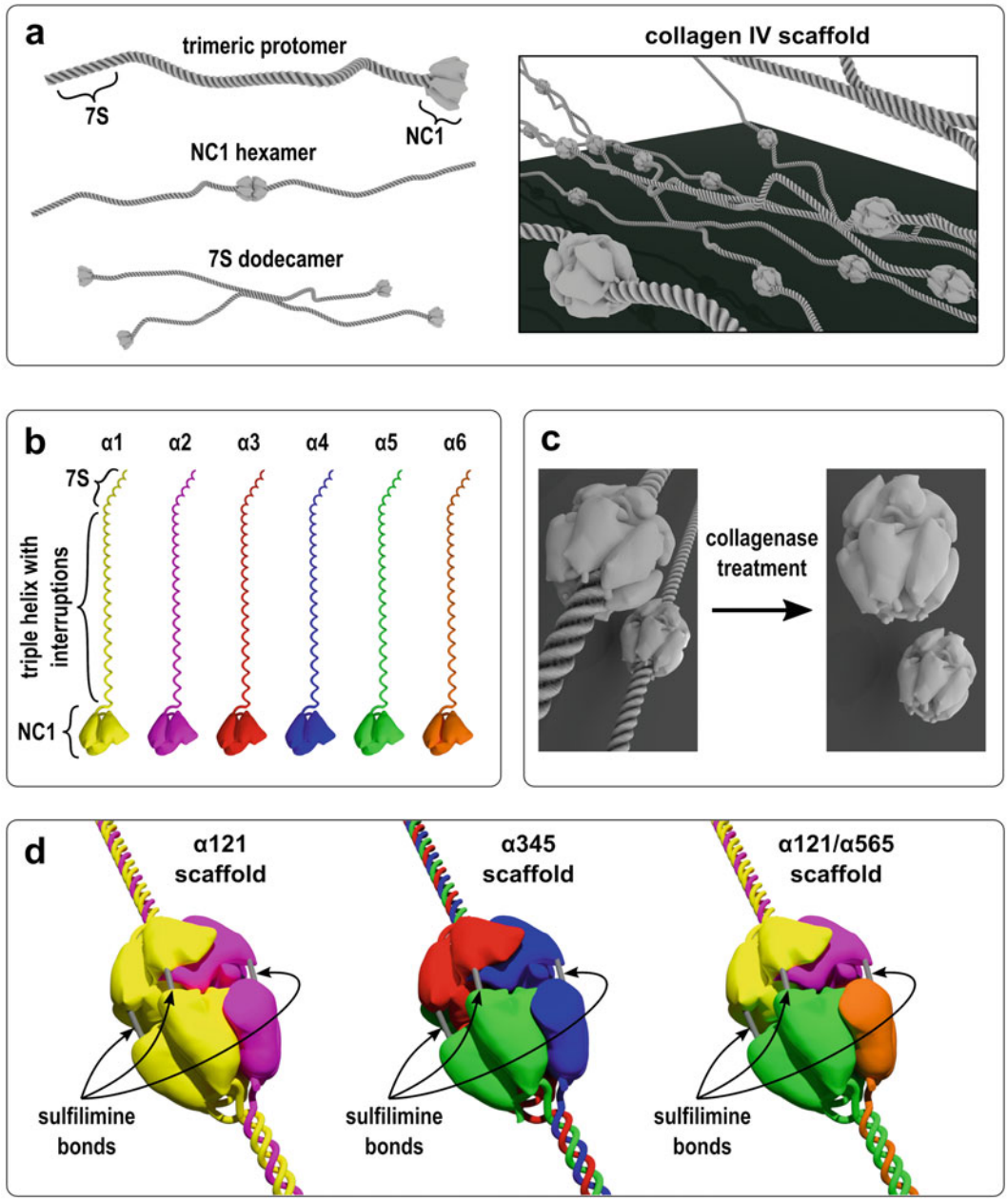

Fig. 1 **Collagen IV**. (**a**) Two key assemblies, NC1-to-NC1 hexamer formation and 7S dodecameric assembly drive the formation of collagen IV scaffolds that are further re-enforced by lateral interactions. (**b**) The six α-chains, α1-α6, of collagen IV found in humans. (**c**) The non-collagenous (NC1) domain can be solubilized from tissues and matrix deposited by cultured cells after collagenase treatment. (**d**) The three reported types of collagen IV scaffolds based on analysis of solubilized NC1 hexamers. The NC1 hexamers are often re-enforced by sulfilimine bonds between two protomers

length collagen IV protomers can be extracted from animal tissues or cell cultures, they are quite challenging to study under physiological conditions as they naturally tend to aggregate (Bachinger et al. 1982). These challenges make full-length protomers prohibitively difficult to use

to study NC1 hexamer assembly. Currently, there are three approaches available to bypass full-length protomers when studying hexamer formation (Fig. 2). The first approach (Fig. 2a) utilizes either recombinantly produced NC1 monomers (Casino et al. 2018) or collagenase liberated native NC1 from bovine LBM or PFHR9 cell culture (Bhave et al. 2012). These hexamers dissociate into monomers upon removal of Cl^- ions using dialysis or desalting columns. These monomers can then be analyzed in hexamer assembly assays by addition of Cl^-. This approach though does not reflect natural hexamer formation as it combines two steps, trimerization and hexamerization, into one process. The second approach (Fig. 2b) utilizes recombinant constructs consisting of the NC1 domain and an N-terminal stretch of 28 GXY repeats essential for forming a stable triple helix (Cummings et al. 2016). These constructs form trimeric molecules in the absence of Cl^- ions and thus, can be used for direct hexamer assembly experiments. The third approach (Fig. 2c) utilizes single-chain recombinant technology in which all three chains of the NC1 domain are tied together in a desired domain composition and registry to form a stable, single polypeptide chain NC1 trimer ($NC1^{sc}$) (Pedchenko et al. 2019). All three approaches bear unique advantages and have certain limitations, but collectively provide tools for studying assembly of the NC1 hexamer.

4 Role of Cl^- in NC1 Hexamer Assembly

The initial clue for how NC1 domains assemble into hexamers was provided by crystal structures of human and bovine $NC1^{\alpha 121}$ hexamers that presented acetate, Br^-, Cl^-, Ca^{2+}, and K^+ coordinated by residues from two opposite trimers (Sundaramoorthy et al. 2002; Than et al. 2002; Vanacore et al. 2004). These and chemically similar ions were extensively studied for their abilities to trigger or facilitate the hexamer assembly (Cummings et al. 2016). Neither Ca^{2+} or K^+

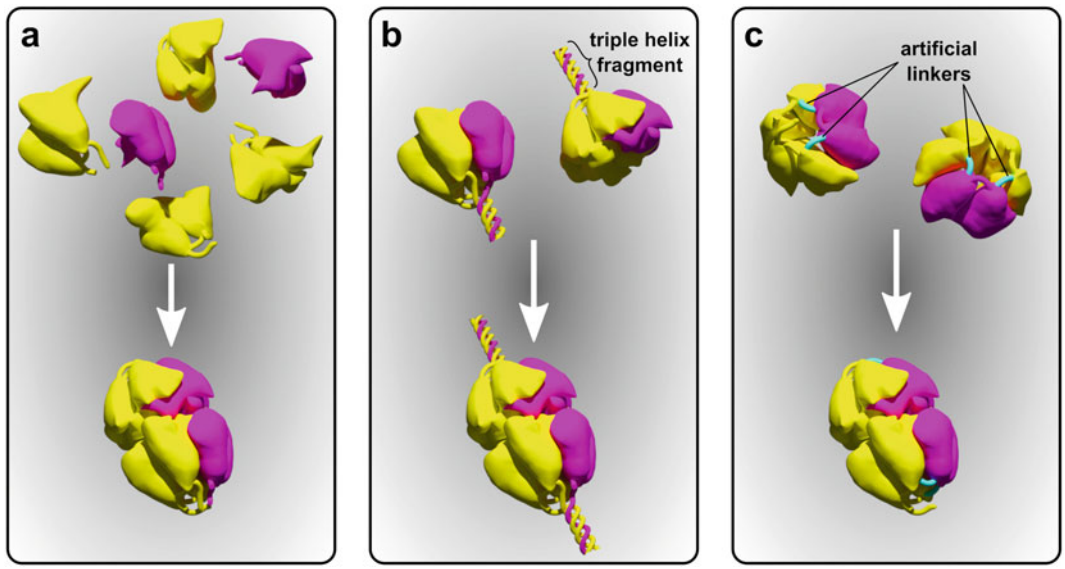

Fig. 2 Experimental models of collagen IV hexamer assembly. (**a**) NC1 monomers from either recombinant expression or from collagenase liberated NC1 extracted from bovine lens capsules or murine PFHR9 cell culture matrix. (**b**) Recombinant truncated protomers containing 28 GXY repeats necessary to form the triple helix and the whole NC1 domain. (**c**) Recombinant single-chain NC1 trimers. Three NC1 chains are linked by artificial 3-residue linkers (shown in cyan)

were able to induce the hexamer formation without presence of negatively charged ions. In contrast, halides were able to induce the hexamer assembly with the following relative efficiency: $Br^- > Cl^- >> I^- >> F^-$, where F^- had negligible effect. Whereas Br^- was the most efficient in initiating hexamerization, Cl^- was the only ion that could efficiently assemble NC1 hexamers at its physiologically relevant extracellular concentration (Cummings et al. 2016). Interestingly, addition of 1 mM Ca^{2+} in the presence of physiologically relevant (100 mM) Cl^- further increased efficiency of hexamerization suggesting complementary effect of divalent cations in this process (Cummings et al. 2016).

The role of Cl^- in assembly and stability of the hexamer can be illustrated in several ways. Size-exclusion chromatography demonstrates assembly (Fig. 3a, left) and disassembly (Fig. 3a, right) of the hexamer by the presence of hexamer or trimer peaks in elution profiles upon addition or removal of Cl^- (Pedchenko et al. 2019). Atomic force microscopy was found to be a very illustrative technique to assess the presence of hexamers in a sample (Fig. 3b) (Pedchenko et al. 2019). When a sample with NC1 trimers was imaged, individual randomly dispersed particles were seen (Fig. 3b, left). Surprisingly, when a sample with NC1 hexamers was imaged, pairs of individual particles were seen (Fig. 3b, right). Explanation of this phenomenon can be found in the method of sample preparation. After absorbing the NC1 hexamers onto the mica surface using the working solution (with

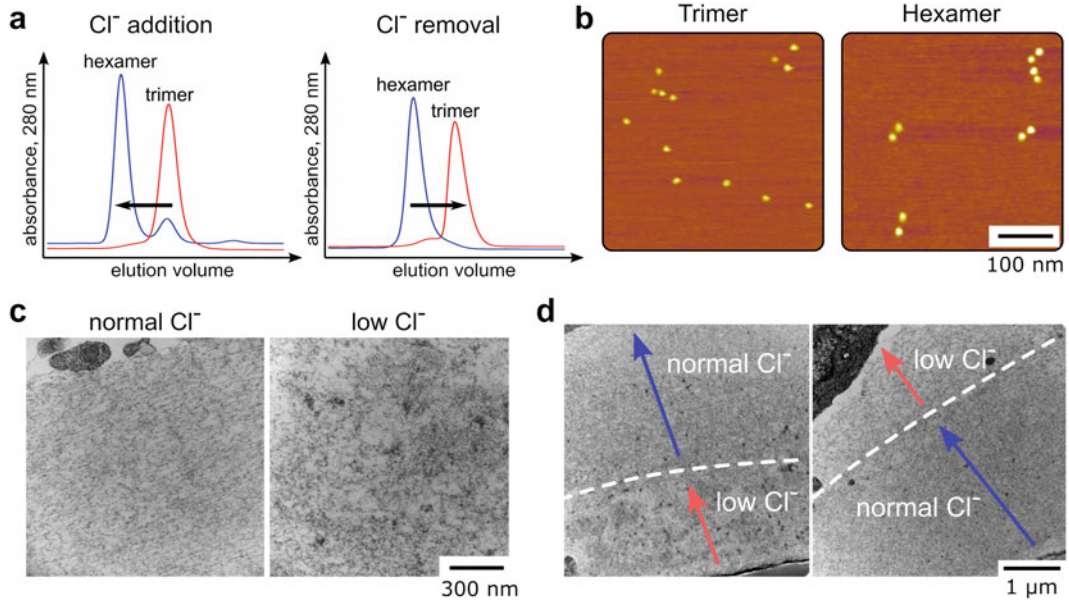

Fig. 3 Cl^- concentration affects collagen IV. (**a**) Elution profiles from size-exclusion chromatography demonstrate that single-chain NC1 trimers (red line) associate into hexamers (blue line) at high Cl^- concentration (left) but dissociate back to trimers when Cl^- are replaced with acetate (right). (**b**) Atomic force microscopy of single-chain NC1 trimers (left) or hexamers (right). Hexamers dissociate into pairs of trimers after washing the sample with water, which initiates Cl^- removal from the NC1. (**c**) Effect of Cl^- on the quality of deposited basement membrane-like matrix by cultured PFHR9 cells as visualized by transmission electron microscopy. Left, cells were grown using normal medium with physiological Cl^- concentration for 10 days. Right, cells were grown in the medium with low Cl^- concentration for 10 days. (**d**) Effects of medium changes on the matrix. Left, initially cells were grown in the medium with low Cl^- concentration for 5 days and then in the normal medium for additional 5 days. Right, initially cells were grown in the normal medium for 5 days and then switched to low Cl^- concentration for additional 5 days. Overall, the matrix deposited under low Cl^- concentration is less dense and a boarder is noticeable between two types of matrix. When cells were initially grown in the low Cl^- concentration the matrix was partially restored (patches of dense matrix) after switching to the normal medium. Arrows indicate direction of basement membrane-like matrix deposition over time

enough Cl^- concentration), the surface is briefly washed with water and immediately air dried for subsequent imaging. This short washing step removes Cl^- and thus, initiates dissociation into trimers, which remain non-covalently absorbed to the surface.

Depletion of Cl^- had drastic effect on quantity and quality of the growing BM as visualized in cell culture experiments with varying Cl^- concentration in the medium (Cummings et al. 2016). At low Cl^- concentration (~5 mM) cells deposit disorganized matrix with lesser density compared to normal conditions (Fig. 3c). Switching from low to normal Cl^- concentration partially repairs the abnormal BM (Fig. 3d, left panel). Meanwhile, switching from normal to low Cl^- concentration seems to not significantly disrupt BM deposited under normal conditions (Fig. 3d, right panel), suggesting stabilization of the NC1 hexamer through sulfilimine cross-linking (Bhave et al. 2012; Vanacore et al. 2009) and possibly complexation with other macromolecules in the BM. Nevertheless, Cl^- depletion can disrupt NC1 hexamer assembly, and may disrupt yet unknown macromolecular complexes within BM.

The development of a single-chain NC1 trimer as a tool to study the assembly of the NC1 hexamer (Fig. 2c) allowed quantitative analysis of this process and establishment of a kinetics model (Pedchenko et al. 2019). Plotting hexamer formation versus Cl^- concentration results in a sigmoidal curve (Fig. 4a). Cl^- concentrations below 10 mM (intracellularly relevant) appear to prohibit hexamer formation, whereas at 100 mM (extracellular concentration), the fraction of hexamer formed approaches a plateau. This titration curve surprisingly fits the two extremes of the Cl^- step gradient across the cell membrane (e.g. 7 mM inside the cell and 100 mM outside the cell (Andersen 2013; Armstrong 2003)) and suggests a major role for Cl^- in collagen IV scaffold assembly outside the cell. Experimentally measured kinetics of the hexamer assembly at high Cl^- concentration and protein concentration dependence revealed that this Cl^- ion driven process fits a simple bimolecular reaction model with the rate constant $k_a = 3.45 \pm 0.12\ M^{-1}\ s^{-1}$ (Pedchenko et al. 2019). Projections of the hexamer assembly demonstrate slow kinetics with half-times ranging from 20 min to 7 h at 25 to 1 mg/ml protein concentration (Fig. 4b). The correlation between this *in vitro* data and the *in vivo* processes remains unknown.

5 Cl^- Ions Are Structural Components of the NC1 Hexamer

Recently reported crystal structures of Cl^- bound hexamers $NC1^{sc\text{-}\alpha121}$ (Fig. 2c), $NC1^{\alpha111}$, $NC1^{\alpha333}$, and $NC1^{\alpha555}$ finally revealed the

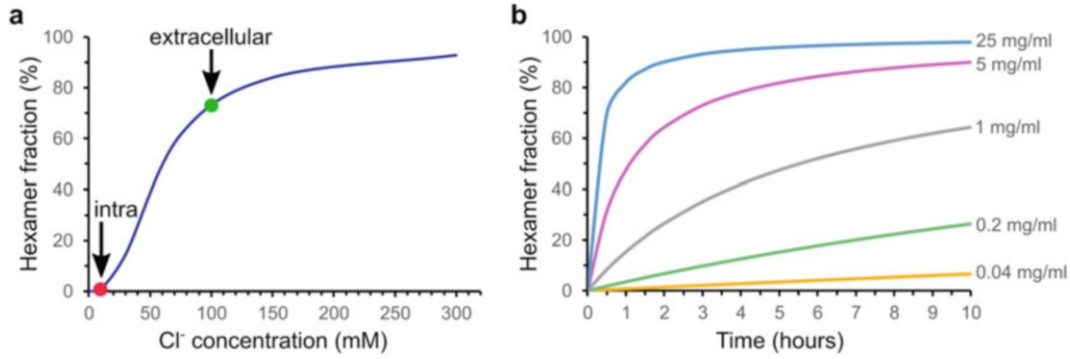

Fig. 4 Cl^- titration and kinetics of the hexamer assembly. (**a**) Cl^- titration curve, adapted from (Pedchenko et al. 2019). The amount of the hexamer formed depends on Cl^- concentration. The red dot indicates approximate intracellular Cl^- concentration, while the green dot indicates approximate extracellular Cl^- concentration under normal physiological conditions. (**b**) Calculated kinetics of hexamer formation at various protein concentrations, based on experimentally defined mechanism (Pedchenko et al. 2019)

number and positions of Cl^- within each hexamer (Casino et al. 2018; Pedchenko et al. 2019). The hexamers were specifically held at high Cl^- concentration prior and during crystallization to assure complete saturation of the structure with the Cl^- ions (Pedchenko et al. 2019). Each of these hexamer structures revealed 12 Cl^- ions at the trimer-trimer interface as shown for the α121 composition (Fig. 5). These 12 ions form a Cl^- ring, composed of two groups of 6, based on their structural environment.

In the following detailed descriptions of Cl^- environment, $NC1^{sc\text{-}\alpha121}$ and $NC1^{\alpha555}$ structures will be used as they were solved at the highest resolution. Group 1 Cl^- intercalate into the base of the NC1 trimers, where they are centered in loops formed by amino acids 74–78 and coordinated by these residues' backbone atoms (Fig. 6, top panel). In addition, an arginine residue from the opposite trimer forms a salt bridge to each of these ions. Molecular dynamics simulations also suggest that binding Group 1 ions causes the re-arrangement of the R76 side chain to break intra-trimeric salt bridges in order to form inter-trimeric salt bridges that contribute to hexamer formation (Cummings et al. 2016).

Group 2 Cl^- are located closer to the equatorial plane of the hexamer (Fig. 5). Instead of intercalating into the base of the NC1 trimers, these ions sit directly at the interface and act like a bridge that connects two protomers by accepting hydrogen bonds from each NC1 trimer (Fig. 6, bottom panel). A unique aspect of the Group 2 Cl^- that will be discussed further in the subsequent section is the fragility of this coordination. Aliphatic C-H groups from alanine and aromatic residue side chains serve as donors for the Cl^- interaction. In addition, the solution outside the hexamer can donate hydrogen bonds not only from water, but also from polyethylene glycol

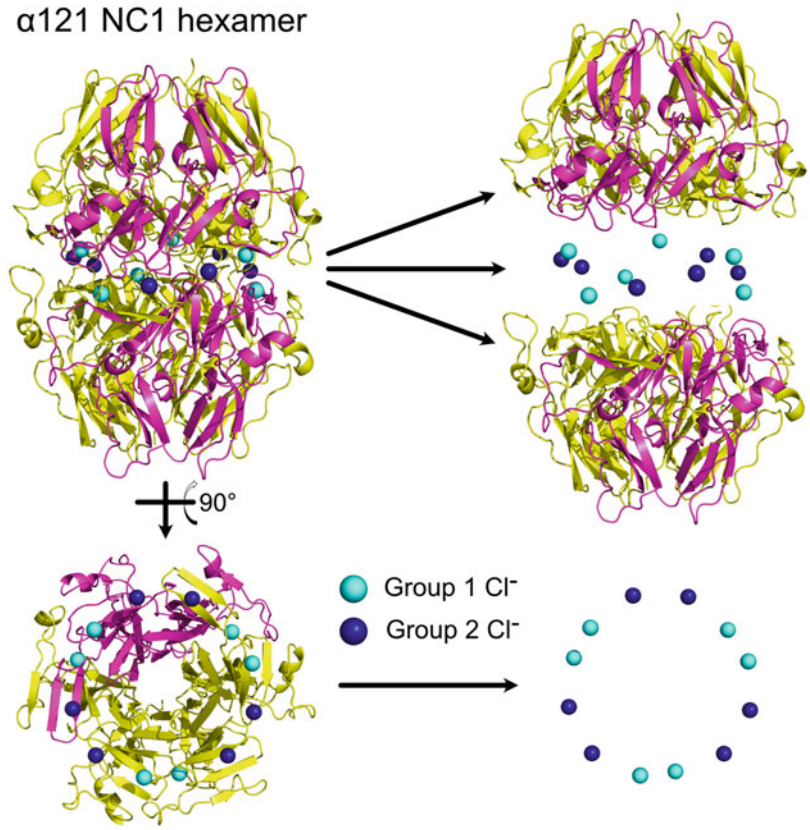

Fig. 5 Cl^- at the hexamer interface of $NC1^{\alpha121}$. Cartoon representation of the crystal structure of the $NC1^{\alpha121}$ hexamer. Coloring for α1 and α2 chains are yellow and magenta, respectively. Spaced-out image reveals 12 Cl^- (shown as spheres) sandwiched between NC1 trimers. Rotation of the hexamer presents an interfacial view that reveals the distribution of Cl^- resembles a ring around the interface. The Cl^- can be further split into Group 1 (colored cyan) and Group 2 (colored blue) depending on the nature of their coordination. The figure was generated using coordinates of the crystal structure of $NC1^{sc\text{-}\alpha121}$ (PBD code: 6MPX)

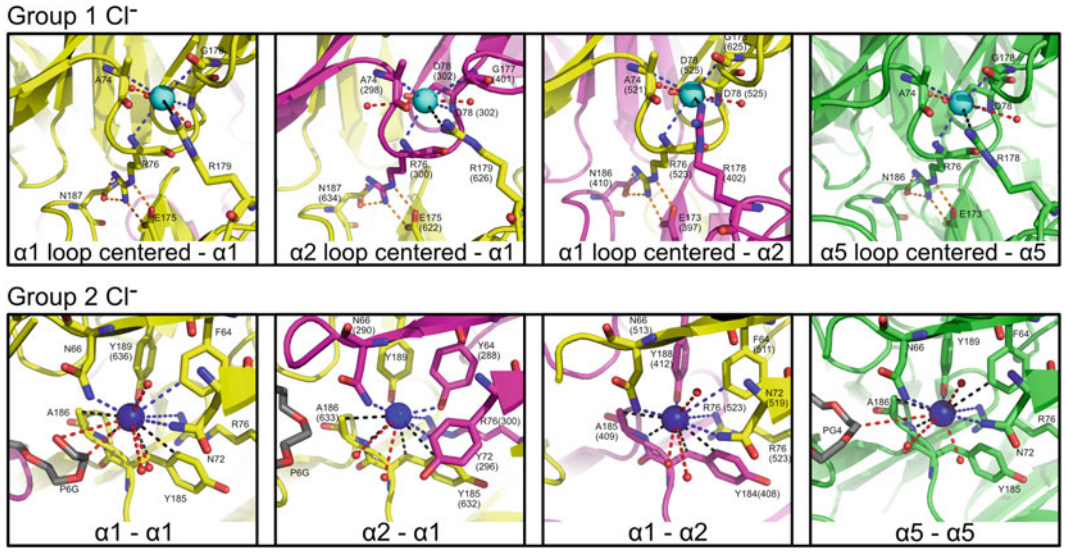

Fig. 6 Cl^- coordination at the interface between NC1 trimers. The top panel shows coordination of Group 1 Cl^- (shown as cyan spheres), while the bottom panel depicts coordination of Group 2 Cl^- (shown as blue spheres). The α1, α2, and α5 NC1 domains are colored yellow, magenta, and green, respectively. Water molecules are shown as red spheres. PEG molecules found near Group 2 ions are shown as gray carbon chains with red oxygens. Interactions between Cl^- and either water or PEG are denoted by red dashes. Blue dashes denote Cl^- interactions with one NC1 trimer, while black dashes denote Cl^- interactions with the opposite NC1 trimer. Residue numbering follows the convention used for non-single chain structures with single-chain structure numbering in parenthesis. The figure was generated using coordinates of the crystal structures of $NC1^{sc\text{-}\alpha121}$ and $NC1^{\alpha555}$ hexamers (PBD codes: 6MPX and 5NAZ)

(PEG) molecules included in the crystallization solution. Combined, these interactions represent a weak interaction network that could allow for relatively easy exchange with solvent Cl^- ions suggesting highly dynamic nature of Group 2 Cl^- coordination within the NC1 domain.

6 Surface Environment Accessibility of Cl^- Ions

Surface analysis of ions at the center of the NC1 hexamer presents a putative binding order and differentiation of roles between Group 1 and Group 2 Cl^-. In both the $NC1^{\alpha121}$ and $NC1^{\alpha555}$ hexamers, Group 1 Cl^- are visible only when one of the trimers is removed (Fig. 7). Hence, it appears the Group 1 Cl^- must bind to the NC1 before Group 2 Cl^- and thus, play a signaling role in assembly. It appears that Group 1 Cl^- ions in the $NC1^{\alpha555}$ trimer are embedded deeper into the NC1 surface compared to equivalent Cl^- in the $NC1^{\alpha121}$ trimer. Whether this indicates $NC1^{\alpha555}$ bind tighter to this Cl^- group or is also relevant to $NC1^{\alpha345}$ or $NC1^{\alpha121/\alpha565}$ hexamers is unknown currently.

Whereas the Group 1 Cl^- are sequestered inside the core of the NC1 hexamers, Group 2 Cl^- are readily accessible to solvent through portals in the hexamer (Fig. 8). Both structures contain polyethylene glycol (PEG) fragments, though their location in the structure leads to a curious observation. While both PEG and water molecules are found near the portals in odd-numbered NC1 alignments, only water was identified near the portals for $NC1^{\alpha2}$ (Fig. 8). Conclusions cannot be drawn from the $NC1^{\alpha2,\ \alpha3,\ \alpha4}$ homo-oligomer structures (PDB IDs: 5NB2, 5NB0, and 5NB1) as none were solved at sufficient resolution to observe PEG molecules included in the crystallization drop and $NC1^{\ \alpha2,\ \alpha4}$ structures form unusual non-hexameric assemblies. Meanwhile, the $NC1^{\alpha1}$ homo-hexamer

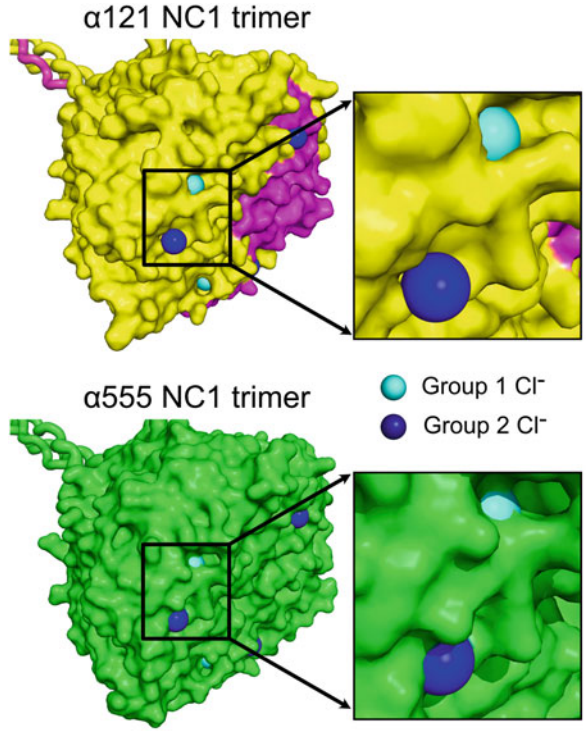

Fig. 7 Interfacial view of the surface of NC1$^{\alpha 121}$ and NC1$^{\alpha 555}$ trimers as part of a hexameric assembly. As Group 1 Cl^- (cyan) are embedded into the NC1 trimer, they are accessed only through channels leading into the interface. Group 2 Cl^-, meanwhile, lie in the interface itself, appearing as though they "sit" on the NC1 surface. The triple helical collagen segments in this figure are added to orient the reader to the interface surface and are not a part of the crystal structures. Arrows extending from each box point to an expand out of the interfacial region within the box. The figure was generated using coordinates of the crystal structures of NC1$^{sc\text{-}\alpha 121}$ and NC1$^{\alpha 555}$ hexamers (PBD codes: 6MPX and 5NAZ)

was not crystallized in the presence of PEG molecules and no structures including NC1$^{\alpha 6}$ are available. Nevertheless, observation of the position of PEG molecules raises the possibility that the surface topography of odd-numbered NC1 pairings is distinguished from the surface topography of even-numbered NC1 pairings in that it could accommodate environmental perturbations near Cl^- that are necessary for NC1 assembly.

Whereas Group 1 is solvent-accessible only in the trimer configuration (Fig. 7), Group 2 remains exposed to solvent upon hexamer assembly (Fig. 8). Thus, each Group 2 Cl^-, positioned in a pocket communicating to the outside through a portal, can therefore be in a dynamic equilibrium with free ions in the solution. Remarkably, the nature and geometry of this arrangement seems to be also conserved for all reported types of NC1 hexamers (Fig. 1d) suggesting a common sensing mechanism to Cl^- concentration.

7 Conclusions

Assembling of collagen IV structures has distinct stages both inside and outside of the cell (Fig. 9). The NC1 domain plays dual intra- and extracellular roles. Inside the cell, the domain is responsible for selecting three α chains, bringing them together and nucleating folding of triple helix in a zipper-like fashion (Soder and Poschl 2004). Outside the cell, trimeric NC1 domains from two protomers are connected by Cl^- to form hexamers, which can be then covalently

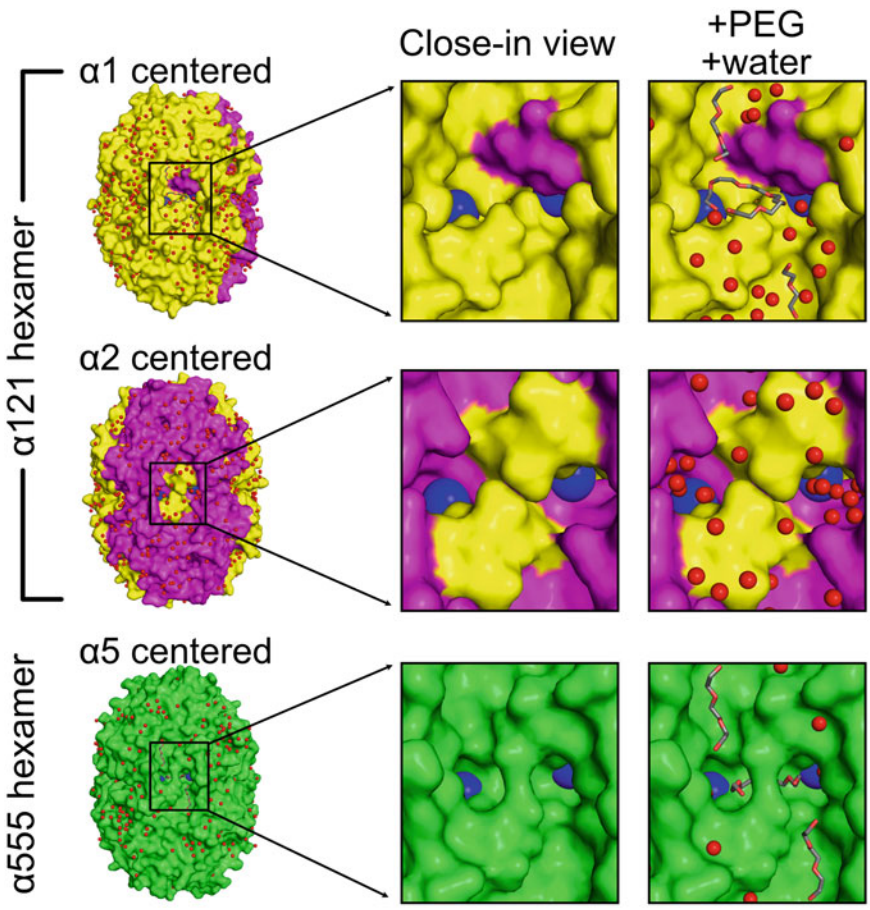

Fig. 8 Group 2 Cl^- are solvent accessible in the NC1 hexamers. Solvent accessible portals on the surface of the NC1 hexamers reveal Group 2 Cl^- (blue spheres) and unique portal environments. Group 1 Cl^- are not directly solvent accessible. Structured water molecules are shown as smaller red spheres, while structured PEG molecules are shown with gray carbon chains. Coloring for each NC1 domain is yellow for α1, magenta for α2, and green for α5. Only one $NC1^{\alpha121}$ trimer is found in the asymmetric unit of the crystal structure and hence, the two α1 centered faces of the hexamer are similar due to the twofold crystallographic symmetry rotation necessary to complete the hexamer. Meanwhile, only one $NC1^{\alpha5}$ domain is found in the asymmetric unit of the corresponding structure. Thus, each α5 centered face is identical due to the threefold and twofold crystallographic symmetry rotations necessary to complete the hexamer. The figure was generated using coordinates of the crystal structures of $NC1^{sc\text{-}\alpha121}$ and $NC1^{\alpha555}$ hexamers (PBD codes: 6MPX and 5NAZ)

connected by sulfilimine cross-links with the help of the peroxide reductase, peroxidasin (Fig. 1d). (Brown et al. 2017; Bhave et al. 2012; Vanacore et al. 2009). Until recently, the exact molecular mechanism of the NC1-driven collagen scaffold assembly remained unclear. Biochemical and structural data have now demonstrated that NC1 domains of collagen IV are equipped with a Cl^- sensing mechanism, which is capable of triggering conformational switches essential for hexamer assembly. It also appears that this mechanism has emerged early in evolution (Cummings et al. 2016; Pedchenko et al. 2019). Structural organization resulting in two groups of Cl^- ions suggests distinct steps take place during hexamer assembly. Group 1 ions bind to the trimer and re-organize the surface for trimer-trimer docking. Group 2 Cl^- bind to the formed hexamer and ultimately stabilize it (Fig. 9) (Pedchenko et al. 2019). The hexamer structure is sufficiently

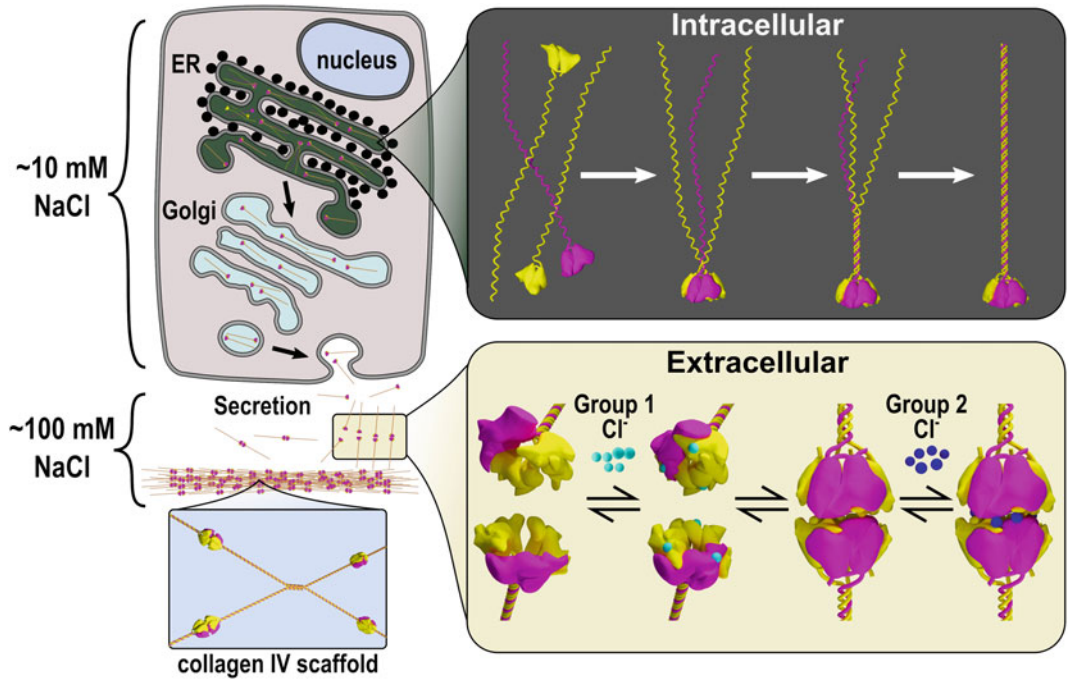

Fig. 9 Collagen IV assembly steps driven by the NC1 domain. Once collagen IV chains are translated into the ER lumen the NC1 domain drives chain selection and trimerization event. Trimerization of the NC1 domain subsequently nucleates the formation of the triple helix which propagates along the molecule from the NC1 domain to the N-termini of the chains in a zipper-like fashion. After multiple post-translational modifications and quality control events, the collagen IV protomers are secreted outside the cell into a high Cl^- concentration environment. In the current model, once in the extracellular space the NC1 trimers undergo conformational transitions upon binding Cl^- (Group 1 (shown as cyan spheres) first, followed by Group 2 (shown as blue spheres)). Subsequently, physiologically high extracellular Cl^- concentration is critical for maintaining the hexamer structure, as Cl^- depletion results in dissociation of the hexamer. This hexamer is a critical step towards collagen IV scaffold assembly

dynamic to require high Cl^- concentration in the surrounding environment as depletion of free Cl^- causes dissociation of the hexamer.

Thus far there is limited structural information on the nature of the extracellular Cl^- binding sites, and therefore, no consensus motifs or even essential residues for Cl^- binding have emerged to allow searching for similar sites in other proteins (Luscher et al. 2020). Establishing Cl^- binding sites within the NC1 hexamer of collagen IV provides a new brick for building our understanding of role of Cl^- in structural organization, functional aspects, and signaling events happening in the extracellular milieu. The step gradient of Cl^- concentration between intracellular and extracellular space can also be exploited by other ECM proteins, yet to be discovered, as a common mechanism. Involvement of Cl^- concentration in control of assembly, function and signaling can also be considered for development of new types of therapies.

There are several interesting questions to be addressed in the future studies of collagen IV scaffold assembly and function related to the NC1 hexamer and Cl^-: (1) Does the sequence similarity between α1-α6 chains lead to the same Cl^--dependent mechanism for NC1 hexamer formation for other known compositions, *i.e.* α345 and α121/α565? (2) Will the Cl^- titration curve and kinetics observed for α121 apply to other assemblies? (3) What is the mechanism of discrimination between α121 and α121/α565 hexamers? (4) What cofactors or helper proteins could accelerate the remarkably

slow rate-limiting step of the hexamer assembly (Sect. 4) and/or reduce the high protein concentration necessary for hexamer formation? (5) What mechanisms, besides sulfilimine cross-linking (Bhave et al. 2012; Vanacore et al. 2009), could protect the NC1 hexamer in the event of a Cl^- concentration drop? (6) How much variability in BM Cl^- concentration exists in health and disease? (7) Is there a cell signaling role for the NC1 domain that utilizes BM Cl^- concentration sensing?

Acknowledgements Supported by grant R01DK18381 and, in part, by start-up funding from Department of Medicine, Division of Nephrology at Vanderbilt University Medical Center to Dr. Boudko.
Figures were prepared using Blender (www.blender.org), PyMol (www.pymol.org), and Inkscape (www.inkscape.org) software.

Conflicts of Interests Authors declare no conflicts of interest.

References

Andersen O (2013) Cellular electrolyte metabolism. Encyclopedia of metalloproteins. Springer, New York

Armstrong CM (2003) The Na/K pump, Cl ion, and osmotic stabilization of cells. Proc Natl Acad Sci U S A 100(10):6257–6262. https://doi.org/10.1073/pnas.0931278100

Bachinger HP, Fessler LI, Fessler JH (1982) Mouse procollagen IV. Characterization and supramolecular association. J Biol Chem 257(16):9796–9803

Bhave G, Cummings CF, Vanacore RM, Kumagai-Cresse-C, Ero-Tolliver IA, Rafi M, Kang JS, Pedchenko V, Fessler LI, Fessler JH, Hudson BG (2012) Peroxidasin forms sulfilimine chemical bonds using hypohalous acids in tissue genesis. Nat Chem Biol 8(9):784–790. https://doi.org/10.1038/nchembio.1038

Borza DB, Bondar O, Ninomiya Y, Sado Y, Naito I, Todd P, Hudson BG (2001) The NC1 domain of collagen IV encodes a novel network composed of the alpha 1, alpha 2, alpha 5, and alpha 6 chains in smooth muscle basement membranes. J Biol Chem 276 (30):28532–28540. https://doi.org/10.1074/jbc.M103690200

Boudko SP, Danylevych N, Hudson BG, Pedchenko VK (2018) Basement membrane collagen IV: isolation of functional domains. Methods Cell Biol 143:171–185. https://doi.org/10.1016/bs.mcb.2017.08.010

Boutaud A, Borza DB, Bondar O, Gunwar S, Netzer KO, Singh N, Ninomiya Y, Sado Y, Noelken ME, Hudson BG (2000) Type IV collagen of the glomerular basement membrane. Evidence that the chain specificity of network assembly is encoded by the noncollagenous NC1 domains. J Biol Chem 275(39):30716–30724. https://doi.org/10.1074/jbc.M004569200

Brown KL, Cummings CF, Vanacore RM, Hudson BG (2017) Building collagen IV smart scaffolds on the outside of cells. Protein Sci 26(11):2151–2161. https://doi.org/10.1002/pro.3283

Casino P, Gozalbo-Rovira R, Rodriguez-Diaz J, Banerjee S, Boutaud A, Rubio V, Hudson BG, Saus J, Cervera J, Marina A (2018) Structures of collagen IV globular domains: insight into associated pathologies, folding and network assembly. IUCr J 5 (Pt 6):765–779. https://doi.org/10.1107/S2052252518012459

Cosgrove D, Liu S (2017) Collagen IV diseases: a focus on the glomerular basement membrane in Alport syndrome. Matrix Biol 57-58:45–54. https://doi.org/10.1016/j.matbio.2016.08.005

Cummings CF, Pedchenko V, Brown KL, Colon S, Rafi M, Jones-Paris C, Pokydeshava E, Liu M, Pastor-Pareja JC, Stothers C, Ero-Tolliver IA, McCall AS, Vanacore R, Bhave G, Santoro S, Blackwell TS, Zent R, Pozzi A, Hudson BG (2016) Extracellular chloride signals collagen IV network assembly during basement membrane formation. J Cell Biol 213(4):479–494. https://doi.org/10.1083/jcb.201510065

Fidler AL, Darris CE, Chetyrkin SV, Pedchenko VK, Boudko SP, Brown KL, Gray Jerome W, Hudson JK, Rokas A, Hudson BG (2017) Collagen IV and basement membrane at the evolutionary dawn of metazoan tissues. elife 6. https://doi.org/10.7554/eLife.24176

Gunwar S, Ballester F, Noelken ME, Sado Y, Ninomiya Y, Hudson BG (1998) Glomerular basement membrane. Identification of a novel disulfide-cross-linked network of alpha3, alpha4, and alpha5 chains of type IV collagen and its implications for the pathogenesis of Alport syndrome. J Biol Chem 273 (15):8767–8775. https://doi.org/10.1074/jbc.273.15.8767

Hudson BG, Kalluri R, Gunwar S, Noelken ME (1994) Structure and organization of type IV collagen of renal glomerular basement membrane. Contrib Nephrol 107:163–167. https://doi.org/10.1159/000422975

Hudson BG, Tryggvason K, Sundaramoorthy M, Neilson EG (2003) Alport's syndrome, Goodpasture's syndrome, and type IV collagen. N Engl J Med 348 (25):2543–2556. https://doi.org/10.1056/NEJMra022296

Jayadev R, Sherwood DR (2017) Basement membranes. Curr Biol 27(6):R207–R211. https://doi.org/10.1016/j.cub.2017.02.006

Kahsai TZ, Enders GC, Gunwar S, Brunmark C, Wieslander J, Kalluri R, Zhou J, Noelken ME, Hudson BG (1997) Seminiferous tubule basement membrane. Composition and organization of type IV collagen chains, and the linkage of alpha3(IV) and alpha5

(IV) chains. J Biol Chem 272(27):17023–17032. https://doi.org/10.1074/jbc.272.27.17023

Kefalides NA (1968) Isolation and characterization of the collagen from glomerular basement membrane. Biochemistry 7(9):3103–3112. https://doi.org/10.1021/bi00849a012

Lu JY, Mohammed TA, Donohue ST, Al-Ghoul KJ (2008) Distribution of basal membrane complex components in elongating lens fibers. Mol Vis 14:1187–1203

Luscher BP, Vachel L, Ohana E, Muallem S (2020) Cl(−) as a bona fide signaling ion. Am J Physiol Cell Physiol 318(1):C125–C136. https://doi.org/10.1152/ajpcell.00354.2019

Meuwissen ME, Halley DJ, Smit LS, Lequin MH, Cobben JM, de Coo R, van Harssel J, Sallevelt S, Woldringh G, van der Knaap MS, de Vries LS, Mancini GM (2015) The expanding phenotype of COL4A1 and COL4A2 mutations: clinical data on 13 newly identified families and a review of the literature. Genet Med 17 (11):843–853. https://doi.org/10.1038/gim.2014.210

Morrissey MA, Sherwood DR (2015) An active role for basement membrane assembly and modification in tissue sculpting. J Cell Sci 128(9):1661–1668. https://doi.org/10.1242/jcs.168021

Pedchenko V, Bauer R, Pokidysheva EN, Al-Shaer A, Forde NR, Fidler AL, Hudson BG, Boudko SP (2019) A chloride ring is an ancient evolutionary innovation mediating the assembly of the collagen IV scaffold of basement membranes. J Biol Chem 294 (20):7968–7981. https://doi.org/10.1074/jbc.RA119.007426

Risteli J, Bachinger HP, Engel J, Furthmayr H, Timpl R (1980) 7-S collagen: characterization of an unusual basement membrane structure. Eur J Biochem 108 (1):239–250. https://doi.org/10.1111/j.1432-1033.1980.tb04717.x

Soder S, Poschl E (2004) The NC1 domain of human collagen IV is necessary to initiate triple helix formation. Biochem Biophys Res Commun 325(1):276–280. https://doi.org/10.1016/j.bbrc.2004.10.034

Spiro RG (1967) Studies on the renal glomerular basement membrane. Preparation and chemical composition. J Biol Chem 242(8):1915–1922

Stokman MF, Renkema KY, Giles RH, Schaefer F, Knoers NV, van Eerde AM (2016) The expanding phenotypic spectra of kidney diseases: insights from genetic studies. Nat Rev Nephrol 12(8):472–483. https://doi.org/10.1038/nrneph.2016.87

Sundaramoorthy M, Meiyappan M, Todd P, Hudson BG (2002) Crystal structure of NC1 domains. Structural basis for type IV collagen assembly in basement membranes. J Biol Chem 277(34):31142–31153. https://doi.org/10.1074/jbc.M201740200

Than ME, Henrich S, Huber R, Ries A, Mann K, Kuhn K, Timpl R, Bourenkov GP, Bartunik HD, Bode W (2002) The 1.9-A crystal structure of the noncollagenous (NC1) domain of human placenta collagen IV shows stabilization via a novel type of covalent Met-Lys cross-link. Proc Natl Acad Sci U S A 99 (10):6607–6612. https://doi.org/10.1073/pnas.062183499

Timpl R, Risteli J, Bachinger HP (1979) Identification of a new basement membrane collagen by the aid of a large fragment resistant to bacterial collagenase. FEBS Lett 101(2):265–268. https://doi.org/10.1016/0014-5793(79)81022-4

Timpl R, Wiedemann H, van Delden V, Furthmayr H, Kuhn K (1981) A network model for the organization of type IV collagen molecules in basement membranes. Eur J Biochem 120(2):203–211. https://doi.org/10.1111/j.1432-1033.1981.tb05690.x

Vanacore RM, Shanmugasundararaj S, Friedman DB, Bondar O, Hudson BG, Sundaramoorthy M (2004) The alpha1.alpha2 network of collagen IV. Reinforced stabilization of the noncollagenous domain-1 by noncovalent forces and the absence of met-Lys cross-links. J Biol Chem 279 (43):44723–44730. https://doi.org/10.1074/jbc.M406344200

Vanacore R, Ham AJ, Voehler M, Sanders CR, Conrads TP, Veenstra TD, Sharpless KB, Dawson PE, Hudson BG (2009) A sulfilimine bond identified in collagen IV. Science 325(5945):1230–1234. https://doi.org/10.1126/science.1176811

Wang X, Li W, Wei K, Xiao R, Wang J, Ma H, Qin L, Shao W, Li C (2018) Missense mutations in COL4A5 or COL4A6 genes may cause cerebrovascular fibromuscular dysplasia: case report and literature review. Medicine (Baltimore) 97(30):e11538. https://doi.org/10.1097/MD.0000000000011538

Weber S, Engel J, Wiedemann H, Glanville RW, Timpl R (1984) Subunit structure and assembly of the globular domain of basement-membrane collagen type IV. Eur J Biochem 139(2):401–410. https://doi.org/10.1111/j.1432-1033.1984.tb08019.x