Multiscale modelling of the extracellular matrix

Hua Wong, Jean-Marc Crowet, Manuel Dauchez, Sylvie Ricard-Blum, Stéphanie Baud, Nicolas Belloy

MULTISCALE MODELLING OF THE EXTRACELLULAR MATRIX

Hua Wong[1], Jean-Marc Crowet[1], Manuel Dauchez[1], Sylvie Ricard-Blum[3], Stéphanie Baud[1,2], Nicolas Belloy[1,2].

1. Université de Reims Champagne Ardenne, CNRS, MEDyC UMR 7369, 51097 Reims, France

2. Université de Reims Champagne Ardenne, Plateau de Modélisation Moléculaire Multi-Echelle (P3M), Maison de la Simulation de Champagne-Ardenne (MaSCA), 51097 Reims, France

3. Univ. Lyon, University Claude Bernard Lyon 1, ICBMS, UMR 5246 CNRS, 69622 Villeurbanne Cedex, France

Corresponding authors:

Hua Wong and Nicolas Belloy

Université de Reims Champagne Ardenne, CNRS, MEDyC UMR 7369, 51097 Reims, France

hua.wong@univ-reims.fr, nicolas.belloy@univ-reims.fr

HIGHLIGHTS

- We introduce a set of theoretical modelling tools based on rigid body dynamics.
- Basement membrane components are modelled as articulated chains of rigid bodies.
- User-defined constraints can be used to investigate and modulate self-assembly.
- Sampled conformations can be exported to all-atom representation.

KEYWORDS

Extracellular matrix, modelling, simulation, rigid bodies, mesoscopic scale, basement membrane

ABBREVIATIONS

CG, coarse-grained; Cryo-EM, cryogenic electron microscopy; DOF, degrees of freedom; ECM, extracellular matrix; EGF, epidermal growth factor; FEM, finite element method; MD, molecular dynamics; NC, non-collagenous; NMR, nuclear magnetic resonance; SAXS, small-angle X-ray scattering.

ABSTRACT

The extracellular matrix is a complex three-dimensional network of molecules that provides cells with a complex microenvironment. The major constituents of the extracellular matrix such as collagen, elastin and associated proteins form supramolecular assemblies contributing to its physicochemical properties and organization. The structure of proteins and their supramolecular assemblies such as fibrils have been studied at the atomic level (e.g., by X-ray crystallography, Nuclear Magnetic Resonance and cryo-Electron Microscopy) or at the microscopic scale. However, many protein complexes are too large to be studied at the atomic level and too small to be studied by microscopy. Most extracellular matrix components fall into this intermediate scale, so-called the mesoscopic scale, preventing their detailed characterization. Simulation and modelling are some of the few powerful and promising approaches that can deepen our understanding of mesoscale systems. We have developed a set of modelling tools to study the self-organization of the extracellular matrix and large motion of macromolecules at the mesoscale level by taking advantage of the dynamics of articulated rigid bodies as a mean to study a larger range of motions at the cost of atomic resolution.

INTRODUCTION

The extracellular matrix (ECM) is a three-dimensional network of proteins, proteoglycans and glycosaminoglycans, which are found in different isoforms in tissues. Collagens, laminins, fibronectin, thrombospondins, and proteoglycans belong to the core matrisome, whereas ECM regulators, ECM-affiliated proteins, and secreted factors are referred to as matrisome-associated proteins [1]. ECM components form insoluble supramolecular assemblies, which provide tissues with mechanical properties, namely tensile strength for collagen fibrils, and resistance to compression for proteoglycans. The organization of the ECM is tissue-specific, and varies depending on the developmental stage, and the physio-pathological context. Basement membranes are thin sheets of ECM, which underlie epithelial and endothelial cells, compartmentalize tissues, and surrounds several cell types [2]. ECM molecular entities are of various sizes, ranging from soluble, globular proteins of 10 nm in diameter or less to macromolecules such as collagens which are several hundred nanometers in length and to fibrils and fibers up to several µm-long [3,4]. Collagen IV molecules form tetramers and dimers via their N- and C-termini respectively, which then self-assemble into a network interacting with the other major components of basement membranes, namely perlecan, laminins and nidogens.

It remains difficult to analyze full-length, multi-domain, ECM macromolecules by X-ray crystallography or NMR and most structural data available at the atomic resolution are those of individual domains which do not reflect the size and shape of the full-length proteins deposited in the ECM. On the other hand, supramolecular assemblies such as fibrils can be studied at the microscopic scale. However, several ECM large proteins (e.g., collagens, laminins, and fibronectin) and protein complexes fall into an intermediate scale, the so-called biological mesoscopic scale as defined in [5] that makes them difficult to observe directly at atomic resolution. Indeed, experimental direct observations of ECM components in the range of 10^{-7}m to 10^{-8}m are often at the lower limit of microscopy and upper limit of X-Ray crystallography. While cryo-EM continues to increase in accuracy and resolution [6], proteins with molecular weight lower than 50 kDa and biomolecules or complexes longer than 500 nm (*i.e.*, the electron beam penetration limit) are still difficult to image [7].

All-atom molecular dynamics (MD) is an established method for studying the dynamics and modifications of ECM-derived peptides. It does so by modelling each atom of a molecule as a discrete point in space called particle, with a given set of parameters that are used to compute the next state of the system at each timestep. It is a proven method to study the ectodomains of membrane proteins [8] and/or protein domains [9–13]. It performs calculations on all the atoms of the model and is especially useful for molecules up to 10-nm in length. In classical MD, as the size of the system increases, so does the number of particles whose features (coordinates, velocity, forces, energies) must be updated at each step of computation, typically every 1-2 fs [14]. For example, the C-terminal non-collagenous (NC) domain 1 of native trimeric collagen IV, a major component of basement membranes, is made up of ~5000 atoms. An all-atom model of collagen IV includes roughly 28 000 atoms without considering any solvent or ions molecules. In addition to collagen IV, a model of a basement membrane should comprise several collagen types (e.g., collagens IV and XVIII) as well as laminins and the proteoglycan perlecan [15]. It should also include links to ECM cell surface receptors such as integrins. The simulation of one atomistic model of a large biomolecule such as collagen IV and a basement membrane made of a combination of the above macromolecules would necessitate huge computational resources to perform calculations and store simulation data even in the exascale computing (10^{18} calculations per second) and big data era. The current record for an atomistic model (*i.e.*, the model of HIV-1 capsid) is the simulation of over 64 million atoms for over 1 µs, which required some of the largest existing supercomputers [16].

For these systems and large ECM supramolecular assemblies (e.g., in the µm-range and beyond such as collagen fibers), coarse-grained (CG) models, requiring fewer computational resources, and discrete or finite elements methods (FEM) are used [17,18]. In contrast to fine-grained models providing atomistic resolution), coarse-grained MD (CGMD) aims at representing complex systems by a reduced number of subcomponents. CGMD clusters atoms in larger beads which simulate the properties of the individual atoms they are made of. Using the CGMD approach, the 5,000 atoms of the NC1 domain of collagen IV could be represented by ten times fewer atoms, thus reducing the degrees of freedom (DOF) of the system and allowing longer integration steps around 10-20 fs. This approach has been used successfully to study collagen molecules [19] [20] and tropoelastin self-assembly into nascent fibrils [21]. At the nanoscale ($\leq 10^{-9}$m), water can be discretized as individual molecules. As we get closer to the microscale, there are so many water molecules that they can be approximated into a continuous medium (or *continuum*) without discrete separation. This concept of *continuum* is referred to as implicit solvation and is also called *continuum* solvation in MD. While in molecular dynamics, atoms position and velocity are updated at each time point according to Newton's equation of motion, in rigid body dynamics, protein domains are considered as rigid domains whose volumes never change or deform during the simulation [22].

In contrast with all-atom MD or CG simulations, FEM tools, such as OpenFOAM [23][24] or ANSYS require the computing power of a desktop workstation but will only work at a scale where matter can be considered a *continuum* and cannot predict the behavior of discrete atoms [25] or sub-domains. Furthermore, FEM is well suited for large systems ranging from the macroscale to the microscale from 1 m to 10^{-6} m.

We have developed a set of tools within the DURABIN project (Developing Utilities for Nanometric Interactions in Biochemistry with Augmented Reality) [26] to help users without expertise in modeling to build models of large macromolecules or multimolecular complexes at the mesoscopic scale in order to bridge the gap between the microscopic and the macroscopic scales, and to study their "collective" behavior in different biological conditions. Here we report the use of these recently developed tools with the ability to export atomistic models from rigid body ones, hinting at true multiscale capability, to build models of individual full-length ECM proteins (*i.e.*, collagen IV, laminin-111, and nidogen-1) and proteoglycan (perlecan) found in basement membranes, and to generate a molecular network to build a three-dimensional model of a basement membrane-like ECM. Basement membranes play a crucial role in delimiting tissue boundaries [27], as a filter in kidneys and in tumor metastasis [28]. Our model will thus be of major interest to investigate the effects on

the organization of basement membrane of various biological contexts and diseases, which could be mimicked by adding constraints to the model (e.g., changes in the number of a particular molecule in a model to reflect up- or down-regulation of this molecule in diseases).

MATERIAL AND METHODS

Unity 3D, a multi-platform game engine

Game engines are suites of software and tools that let users import assets (3D objects, 2D textures, game controller hardware interfaces...) in a unified environment to build simulations for entertainment purposes (video games) or more serious applications (industrial training, scientific visualization...) [29]. Unity3D is a multi-platform game engine, which could be used by operating system agnostics and has been already used for scientific visualization [30]. It provides scripting tools that will be used both for simulation and programming interactions between biomolecules in the same way video game characters interact with their environment or with other characters. These scripts are embedded as components of each object in Unity and program how these objects move or react in specific situations.

Rigid body dynamics

While in molecular dynamics, atoms position and velocity are updated at each time point according to Newton's equation of motion, in rigid body dynamics, protein domains are considered as rigid domains whose volumes never change or deform [22]. While particles have three translational degrees of freedom, meaning that their motion can be described in terms of x, y, z translation, rigid bodies add three rotational degrees of freedom around the x, y and z-axis. To transform our all-atom structure gathered from the Protein Databank [31], we simply apply the x,y,z, translation and rotation of the rigid body to the whole PDB file, paving the way for getting from a rigid body representation back to an atomistic model as described in the "Results" section. Simulations are run using physics engines which, like other simulation tools (GROMACS, NAMD), are software packages that provide means to simulate the behavior of physical systems within a given set of parameters/rules (force field). In most physics engines rigid body dynamics is implemented using two components. One holds the physics information used by the physics engine to update the simulation (mass, position, orientation, velocities, forces), whereas the second, called a collider, holds the actual spatial volume. Colliders are used to define the space occupied by a biomolecule as well as to operate collisions between objects or the environment. Colliders are simple geometric shapes (spheres, cylinders, cubes, planes) also called primitives, as this speeds up collision physics. Complex mesh colliders exist but have significant computing costs. While rod-like helical or

fibrous molecules can be modelled as capsules or cylinders, more complex forms will require a compound collider made of a mix of primitive colliders. It should be noted that several figures of this article show detailed surface representations of the biomolecules, but this is, for all intent and purpose, only as a visualization proxy for the physics entities (primitive colliders) used in the simulation.

Chains of rigid bodies are created by articulating each rigid body with positional constraints, preventing them to drift from one another. At each step of the simulation, rigid body positions in the simulation are taken into account and constraints are applied to keep objects linked together [32]. Positional constraints are also used to simulate interactions between molecules. When two rigid body collide, the whole process is managed by a transient positional constraint, which can be modified to be permanent or last a given amount of time that can be modulated to mimick e.g., transient interactions or domain affinity. Random forces are applied to the rigid bodies during the simulation. For this, a Langevin equation with a viscosity term [33] is used. This mimics a fluid at thermal equilibrium and the viscosity term allows for the solvent to be simulated without explicitly modelling the solvent molecules individually (implicit solvation or continuum solvation), thus reducing the number of objects/particles for which interactions must be calculated. Other programs such as CellPACK [5] have made use of simplified representation like rigid bodies to create macromolecular ensembles with a high packing ratio but the current version of CellPACK does not seem to produce large-scale motion simulations. The data presented in this study aims to demonstrate the potential of the DURABIN toolkit to help decipher ECM assembly and dynamics.

Membrane simulation

A restraint field based on IMPALA (Integral Membrane Proteins and Lipids Association) [34] was implemented in DURABIN to simulate how transmembrane helices anchor in the cell membrane. This potential was used to simulate more realistic interactions between integrins and the cell membrane in our model. This approach can be used for any type of transmembrane component interacting with the ECM. The cell lipid membrane was described as a continuum in the following function:

Equation 1:

$$C_z = 0.5 - \frac{1}{1 + e^{\alpha(|z| - z_0)}}$$

where z is the depth at which the helix collider is relative to the center of the membrane, z_0 the depth at which $C_z = 0$. With α set to 1.99 and z_0 set to 15.75 Å, $C_z = 0.49$ at +/- 18 Å and $C_z = -0.49$ at +/-13.5 Å. These values were chosen to represent the thickness of a 1,2-dipalmitoylphosphatidylcholine (DPPC) bilayer in the fluid phase with a smooth transition to the hydrophobic core of the membrane [34,35]. This thickness can be adapted to other lipids or membrane compositions. A collider of 36 Å in height, spanning the whole "membrane", detects if a helix collider enters the "membrane" collider which applies a force proportional to C_z pulling the transmembrane domain in the membrane. The result is a vector pointing toward the inside of the membrane when the helix is in contact with the membrane but still in the water phase and approaches 0 as the helix gets closer to the hydrophobic region of the bilayer. Thus, any rigid body set in the simulation as being transmembrane, regardless of shape, will be attracted inside the membrane *continuum* field, but will still be free to move around once inside the membrane.

Collecting or modeling 3D structures of interest

Building an accurate model of multi-domains protein complexes requires the collection of the highest number of structural data available. As mentioned above, many components of the extracellular matrix are too large to have their 3D structure determined by X-ray or NMR as full-length proteins. High-molecular weight proteins are indeed difficult to crystallize [36] or to analyze by NMR [37]. Most 3D atomic structures available for ECM proteins or proteoglycans are those of individual domains or pairs of domains and not those of full-length molecules. Low-resolution techniques such as AFM, SAXS and rotary-shadowing electron microscopy give insights into the domain organization of ECM proteins and their global size, which are useful to build models. Although recent developments in the field of cryo-EM increased the resolution, it is still costly [38] and subject to limitations (size of the complex, among other things) [39]. Also of note, while recent, AI based techniques to solve the 3D structure of proteins have made large contributions toward providing plausible structures of previously unresolved or unresolveable structures [40,41], although one has to be aware of the limitations of such technique especially regarding disordered regions [42]. Ideally, data collected by both high- and low-resolution techniques should be combined to build our models.

The main source of experimentally solved protein structures is the Protein Data Bank [43] but data collected by cryo-electron microscopy are also available via the public repository Electron Microscopy Data Bank [44][45]. When no 3D structure experimentally determined is

available, modelling approaches (e.g., threading, *ab initio* or homology modelling [46–48]) are used to predict the 3D structure of a protein or a domain from its primary sequence.

3D structure importation and model building in Unity 3D

The coordinates from the PDB being not centered, it is necessary to treat the PDB data so that the center of mass corresponds to the origin of the coordinates using any tool allowing the manipulation of 3D coordinates. Our approach uses Python scripts running in Blender (https://blender.org/), a free and open-source 3D software that supports the modelling and rendering of 3D objects, but any tool allowing the transformation of the coordinates of a PDB file could be used. When an axis is obvious in the molecule (e.g., a rod-like triple-helical domain), the main axis is aligned along the Y-axis which is by default the up direction in Unity3D. The modified PDB file is then opened in VMD, where the molecular surface representation is exported in Wavefront .obj file format. This file format is human readable and can be used both by many applications including both Blender and Unity3D. The molecular surface is used as a mesh representation for the underlying rigid bodies. VMD produces very dense mesh, which can affect interactive rendering speed. Blender is used to simplify the object, lowering the polygon count, while keeping the overall shape intact. Once imported into Unity3D, the required components for the physics engine such as the rigid body components themselves and the associated constraints are added.

Simulating in Unity 3D

We define a simulation box with a base area of 400 nm x 400 nm to fit the longest model of the selected proteins. The thickness of the box varies with the nature of the studied system. For example, the basement membrane thickness varies with age, disease and/or the methods used to measure it. The thickness of basement membrane varies depending on tissues and on the physiopathological context, but most basement membranes are 50–100 nm thick [49]. The thickness thus was fixed at 100 nm, which in addition accomodates the length of laminin molecules. The total size of the simulation box was 0.016 μm^3 including less than 600 rigid bodies, all of them representing domains of larger macromolecules. In contrast, an atomistic model of basement membrane made solely of integrin (one isoform), collagen IV (one isoform), laminin and nidogen, would account for approximately 865 000 atoms.

While CGMD reduces the DOF of residues, the rigid body approach further reduces the DOF of protein domains and represents large macromolecular complexes as articulated chains of

protein domains where individual domains are the beads of a very coarse model. This approach has been used to study larger mesoscopic systems such as the yeast interphase nucleus [33]. Using this approach, collagen IV was modelled as a chain of 44 rigid bodies (one NC1 C-terminal domain and 43 triple-helical domains of 9-nm length and a total length of ~390 nm [50]. Details on the constraints used to articulate the chain are given further in the paper. The crystal structure of the collagen triple helix model $[(\text{Pro-Pro-Gly})_{10}]_3$ (PDB 1K6F, ***Table 1***) was used as an individual triple-helical domain. This all-atom representation would correspond to ~ 28 000 atoms (44*1K6F (528 atoms) + 1LI1(5253 atoms)). The number of domains being far lower than the number of residues or atoms, the amount of computing resources required is that of a desktop PC. It can even be run interactively, which allows the addition of functionalities such as user feedback and interactions in real-time as the simulation runs. By importing the model in the Unity3D game engine, it is possible to make a representation of the collagen network running in real-time with few computing resources [26].

All-atom reconstruction: back-mapping

In Unity3D, a C# script in the project's assets (SaveSnapshot.cs) exports the transformation matrix of the rigid bodies from the molecule models in the ongoing simulation. The original atomic coordinates are transformed using the matrix and the coordinates are saved as separate files formatted as GROMACS .gro files. Because these are separate files, further work is necessary to create one single molecule that can be used with a MD package like GROMACS. But by loading all the separate .gro files in the visualization software VMD it is possible to visualize the all-atom model, nonetheless.

RESULTS

Building of rigid body models of individual proteins

The 3D structures used to build the proteins of basement membrane models are listed in Table 1. There are for examples several structures of the NC1 domain of collagen IV issued from different collagen IV species and isoforms (e.g., homo-oligomers or hetero-oligomers), cross-linked or not, from different tissues, at various resolution. However, from a "rigid body" point of view and at the mesoscopic scale, the NC1 domains are similar in term of general shape (half-oval shaped geometry, same diameter, height). The 1LI1 structure has been

selected because it comprises two alpha1 chains and one alpha2 chains of collagen IV, which corresponds to the major isoform of collagen IV in most tissues, and is cross-linked by the Met-Lys cross-links. The crystal structure of the collagen triple helix model 1K6F, although not specific to collagen IV, was used because the triple helix is a common feature of all collagens [51]. A bacterial coiled-coil structure (2WPQ) was used to represent the coiled-coil region of laminin-111 [52,53]. No 3D structure of the G1 domain of nidogen-1 was available, and a threading approach was used to build a model of this domain [54]. Basement membranes surround several cell types and several basement membrane constituents (e.g., laminin-111 and perlecan) interact with integrins. Integrins were thus included in the simulation. 3FCS was first chosen because of its ability to interact with molecules such as thrombospondins [55] in globular basement membrane, but as the focus of the project shifted toward more general purpose simulation of the basement membrane, this integrin variant was kept as the default integrin model because of the mechanical similarities of the ectodomain between integrins. The extracellular part of the integrin, compared to other structures in ***Table 1***, has all the domains solved in the same PDB file instead of being spread throughout the PDB. To be exploitable in our simulation it was necessary to break down the extracellular part of the integrin into domains to generate separate PDB files. The integrin α subunit for example, is made of residues 1 to 452, the thigh (α subunit) goes from residue 453 to residue 604. The integrin β subunit was similarly sliced into domains [56].

Proteins	**Domains**	**PDB entries**
Collagen triple helix	Crystal Structure of the Collagen Triple Helix Model [(Pro-Pro-Gly)$_{10}$]$_3$	1K6F
Collagen IV	The hexameric noncollagenous domain 1 of human placenta collagen IV (α1α1α2 ισοφορλ)	1LI1
Laminin EGF-like modules	3 consecutive laminin-type EGF-like (LE) modules of laminin gamma1 chain harboring the nidogen binding site	1KLO
Laminin coiled-coil domain	A coiled-coil motif (Salmonella enterica SadA 479-519 fused to GCN4 adaptors)	2WPQ
Laminin (α5 chain)	Laminin α5 chain N-terminal fragment	2Y38
Laminin β1 (short arm)	Laminin β1 LN-LE1-4 structure	4AQS
Laminin γ1 (short arm)	Laminin γ1 LN-LE1-2 structure	4AQT
Laminin α2	laminin α2 subunit L4b Domain	4YEQ
Laminin α1 chain	Mouse laminin alpha1 chain, domains LG4-5	2JD4
Nidogen-1 (EGF domain)	EGF	1JL9
Nidogen	G1 threading model [54] (no available crystal structure)	
Nidogen-1 domain G2	Domain G2 of mouse nidogen-1	1H4U
Perlecan (LDL receptor	2nd repeat of the LDL receptor ligand-binding domain	1LDR

domain)	(domain mediates interactions of the receptor with two lipoprotein apoproteins, apo E and apo B-100)	
Perlecan (sea-urchin sperm protein, enterokinase and agrin domain)	SEA domain of human mucin 1	2ACM
Perlecan (LG like domain 3)	Laminin G like domain 3 from human perlecan	3SH4
Integrin ectodomain	Structure of complete ectodomain of integrin αIIBβ3	3FCS
Integrin transmembrane domain	Integrin αIIBβ3 transmembrane complex	2K9J
Complexes		**PDB entries**
Laminin-111 (integrin-binding domain)	The heterotrimeric integrin-binding region of laminin-111 (50 residues of α1β1γ1 coiled coil and the first 3 laminin G-like (LG) domains of the α1 chain)	5MC9
Laminin-Nidogen complex	Nidogen/Laminin Complex (a 6-bladed β-propeller domain in nidogen laminin epidermal-growth-factor-like (LE) modules III3-5)	1NPE
Perlecan/nidogen complex	Nidogen-1 G2/Perlecan IG3 Complex	1GL4

Table 1: List of known PDB entries of proteins, glycosaminoglycans and protein complexes of the basement membrane and integrins

Known biomolecular interactions were integrated and parameterized for nidogen-1, where two out of the three main domains, G2 and G3, interact with collagen IV and laminin-111 respectively [57]. Integrins interact with laminin-111 globular domain [58]. Collagen IV dimerizes through its non-collagenous (NC1) C-terminal domain and the 7S N-terminal domain to form a network [59]. Laminin-111 interacts with itself through the N-terminus of α, β and γ chains [60]. Perlecan interacts with a wide range of ECM molecules and growth factors. In basement membranes, perlecan connects laminin and collagen IV, while interacting with integrin α2β1 [61].

The process to go from existing data reported in the literature to a fully rigged rigid body model that can be used in a rigid body simulation is summarized in Figure 1 using nidogen as an example. The process is mainly manual, with helper scripts in Python for Blender and the Editor functions for Unity3D. Although we reckon the task could be automated, this was not the focus of this work. Nidogen-1 is a linear molecule made of three domains, G1, G2, G3 connected by EGF-like domains as represented in **Figure 1**A, based on biochemical [62] or biophysical [63] data. If we assemble the structural bits according to the experimentally determined domain organization the resulting s maximum size of nidogen is ~25 nm, roughly

the maximum particle size determined by SAXS experiments [63]. The curation of existing data allowed us to determine the available structures and the part of the protein to be modelled (*Table 1*). It should be noted that low-resolution data such as cryo-EM and SAXS envelopes can be used for simulations, although *in fine* a true atomistic structure will be needed. The domains were then converted to rigid body models with matching colliders as shown for G1 and G3 domains (**Figure 1**D). The slightly oblong EGF and G2 domains were matched with capsule colliders (cylinder capped with spheres). The rigid bodies were assembled according to the organization of domains by joining the C-terminus of a domain with the N-terminus of the next domain. As explained in the passage on rigid body, joints are positional constraints updated at each timesteps with customizable rotational degrees of freedom. Unity3D allows for different presets of these constraints [64]. We used the configurable joints as it was the most polyvalent and is highly configurable. The rotational and positional constraints between rigid bodies were defined at this stepto fine-tune flexibility by modulating rotational constraints. In most cases, in absence of information on the rotational freedom between domains, we use the simplest joint, positional constraint without rotational constraints. For laminin and collagen domains, we add twist constraint, reflecting how super-coiled structures can be more constrained in their helical axis, the reality being more complex [65].

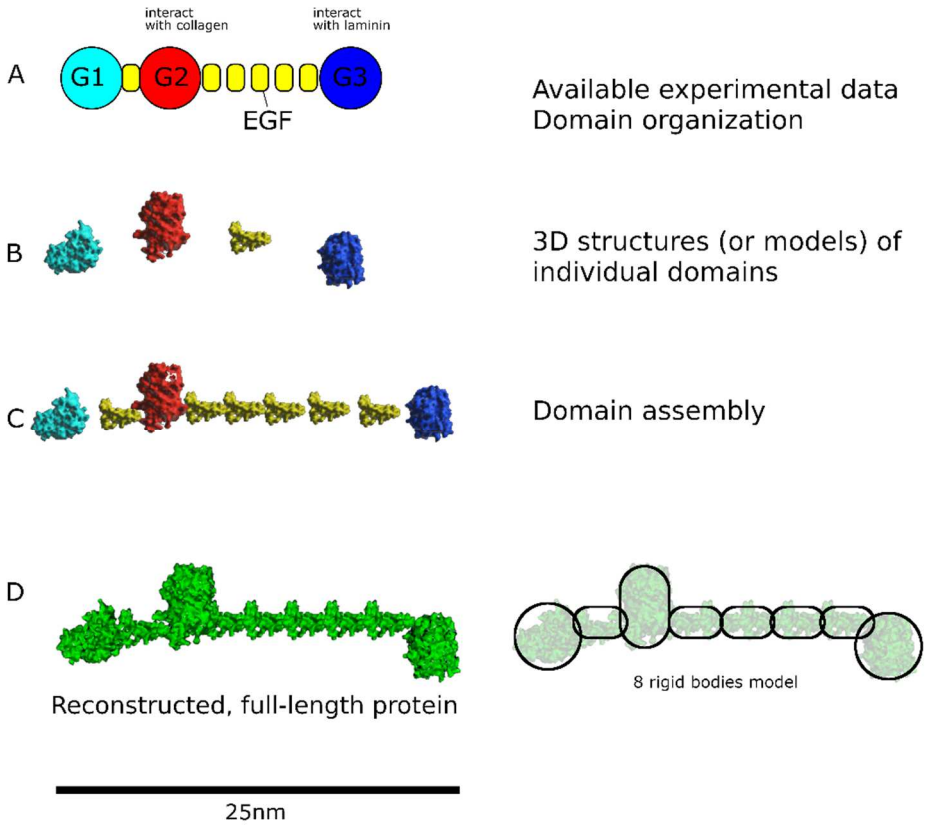

Figure 1: **A rigid body model based on nidogen-1 for rigid body simulation**. A Domain structure of nidogen-1 including the G1, G2, G3 and EGF-like domains. B: Solvent Accessible Surface representation of the individual domains. C: Construction of the rigid body model based on literature data. D: final model used in the simulation, (right) schema of the rigid body colliders used in the simulation.

All-atom reconstruction

One of the goals of the DURABIN as a project is to develop convenient tools to do simulation from all atoms to rigid body and back to all atoms (called back-mapping in CGMD) in a true multiscale way. While rigid body dynamics can provide a convenient way to sample and study large-scale motions of molecular complexes, the details of biomolecular interactions lie in the all-atoms realms of small range motions and residue reorientation.

In **Figure 2**, we use custom scripts and user interaction to generate a couple of collagen IV molecules at 1/10th their final size and we then scale them back to their final real size. This allows us to fit long polymers in a very constrained spaces, like sphere, which does not have biological relevance for collagen IV but shows the ease of building all-atom models of a large protein in a very tightly confined space, which could be of practical use when modelling a crowded ECM or integrin trafficking [66]. In CGMD, the first step of back-mapping is to get

back the coordinates of the original atoms used to build the CG model, followed by an optimization step where the model is relaxed. In the case of rigid-body models, atomic coordinates are back-mapped to fit the newly oriented and positioned rigid body (**Figure 2**B). The minimization/relaxation step is far less trivial to achieve. It is a crucial phase for using the model in MD simulations, especially to correct problems such as atom superimposition and unwanted kinks.

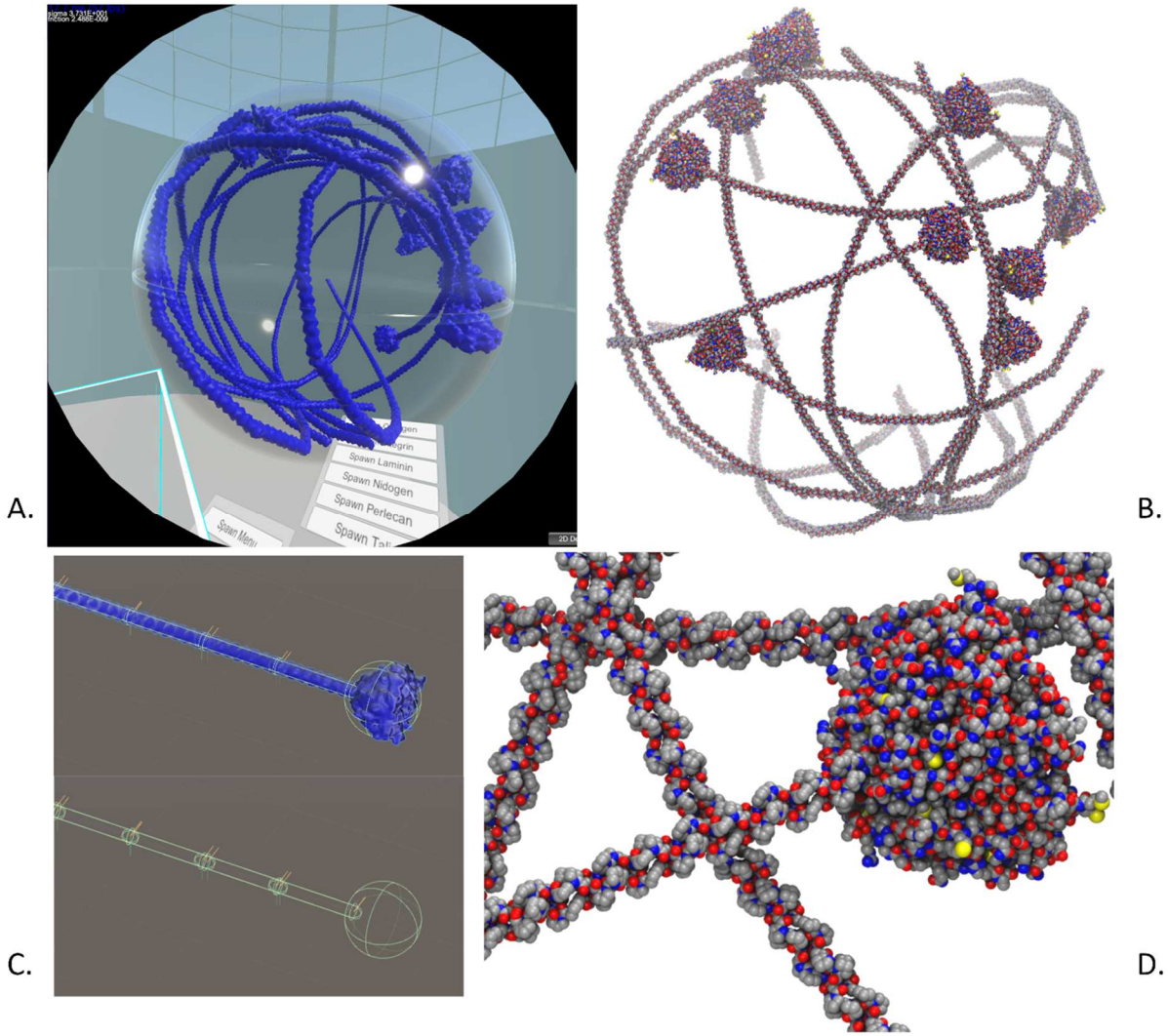

Figure 2: Example of atomistic model construction based on rigid body simulation. Example of a complex model that can be made with DURABIN. A. Real-time rigid body models of the collagen-IV heterotrimer α1α1α2 constrained to a sphere. B. Atomistic model built using PDB data (Table 1, 1K6F, 1LI1) and rigid bodies spatial orientation data extracted from A. C. Top: 3D mesh representation of the molecular surface (blue) lacking atomic details, overlaid or superimposed on the rigid bodies (green,). Bottom: rigid bodies alone showing how a slight overlap leads a gapless chain. D. Close up view of the all-atoms globular NC1 domain of collagen IV. Carbon: gray, oxygen: red, nitrogen: blue, Sulphur: yellow. Hydrogen atoms are not represented for the sake of clarity.

Modelling a basement membrane

Setting up the simulation.

Models of basement membrane components, namely collagen IV, laminin-111, nidogen-1 and perlecan, were generated in a simulation box described in the Material and Methods section (400 nm x 400 nm x 100 nm). While the components of the basement membrane are known, the proportion of each component is still subject to debate [67]. In our model, we have chosen a stoichiometry of 1:1 so that in theory, no molecule in our simulation is left orphaned during self-assembly (one collagen interacting with one laminin and so on). The simulation was then run, and the rigid body model allowed to diffuse in an implicit medium which viscosity is modulated by the use of Langevin dynamics, under the influence of random thermal forces linked to a thermostat (higher temperatures mean larger forces) in order to self-assemble into a basement membrane-like ECM (Video Supplementary material). Any interactions at this stage resulted from random/chance encounter during the simulation (Figure 5).

Influence of collagen IV flexibility.

In our previous study [26], we investigated the behaviour of laminins and nidogen. The collagen IV chains contain 21–26 interruptions of various lengths in their triple helix [68], which provide collagen IV molecules with flexibility. By tweaking the joint rotational constraints (**Figure 3**A) between the 43 individual triple helical domains used to build the model of collagen IV, it was possible to stiffen or loosen the collagen molecule (see **Figure 3**B and **Figure 3**C).

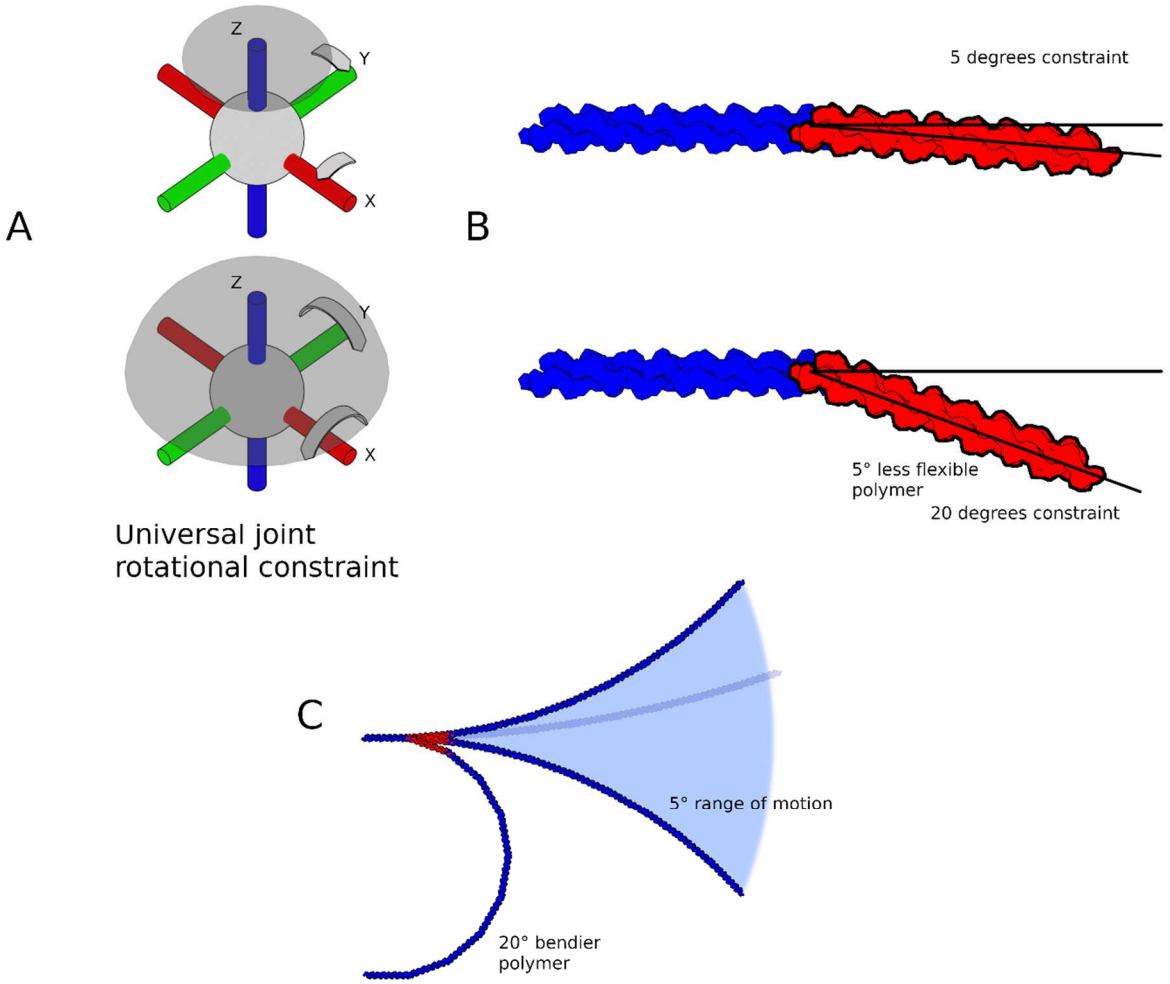

Figure 3: Tweaking joint rotational constraint changes the flexibility of a polymer model. A. How the rotation constraint affects the joint. X and Y axis is more limited in Top A than in Bottom A, the grey cap helps visualize the resulting Z axis exploration limits. B. Close up view isolating the rotation constraint. C. Difference in turn radius between a stiff (5 degrees rotation constraint) and a bendy (20 degrees rotation constraint) polymer.

A simple simulation of the basement membrane showed the influence of collagen IV flexibility on the final structures formed (***Figure 5***). When collagen IV was rigid or moderately flexible, laminin and nidogen molecules tended to spread because the collagen rigidity helped to maintain a minimal distance (***Figure 5***A and B) and to form an irregular polygonal network during the simulation as previously observed experimentally [69].

In contrast, when collagen IV was flexible, distances decreased, and the overall structure looked like an aggregate. If the flexibility was set to the maximum value, allowing the formation of kinks and U-turns, clumps formed, and the self-organization appeared to be lost. Increasing the flexibility of the collagen IV model leads to the formation of a loose network

[69]. The ability to modulate the self-assembly of basement membrane from organized to disorganized networks could be useful to mimic what happens in some diseases and to investigate the underlying molecular events associated with and/or triggering the disease [70]. The introduction of cross-links mimicking physiological cross-linked ECM would be useful to build an ECM model exhibiting various mechanical properties modulated by the extent of crosslinking [71].

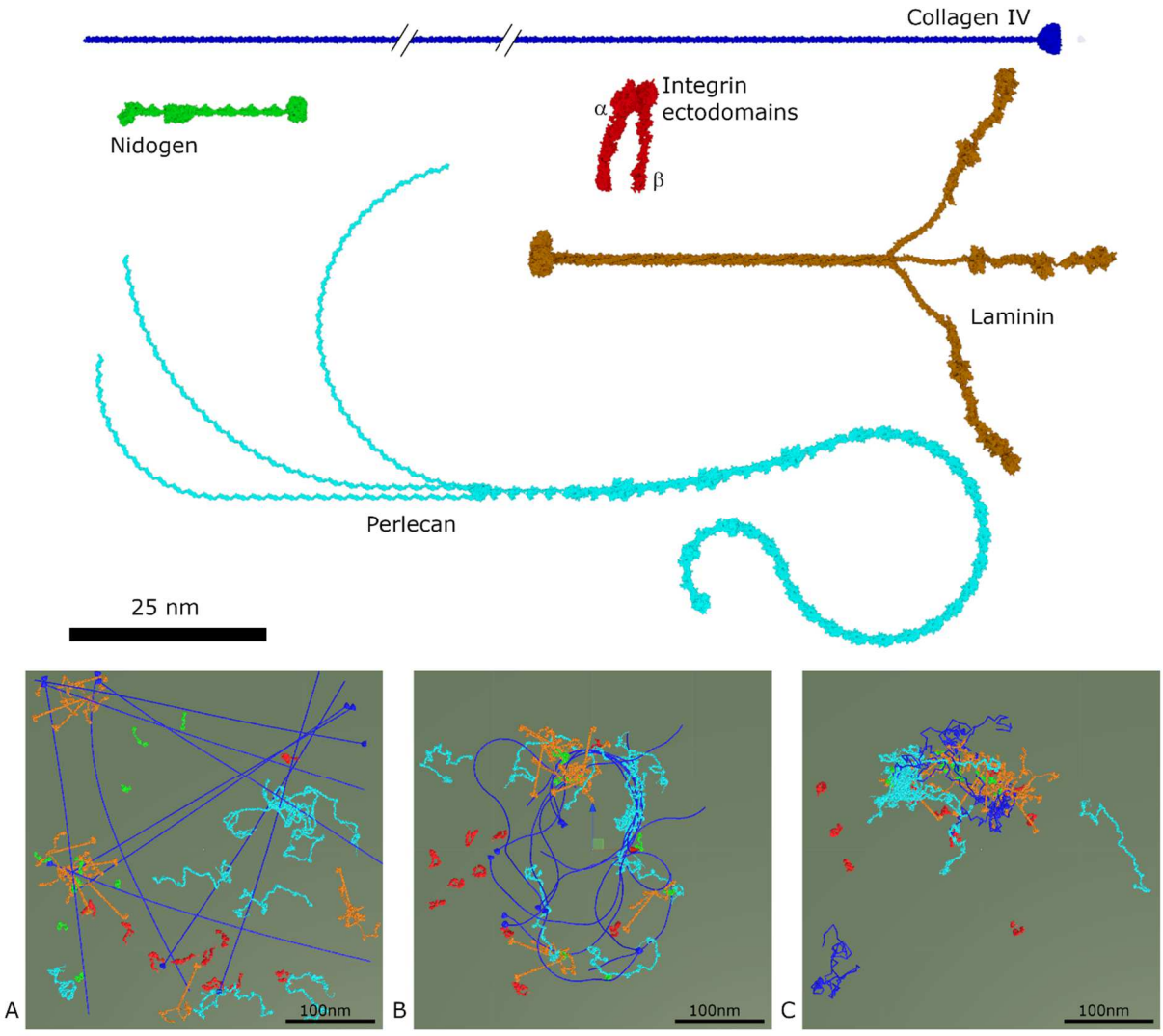

Figure 5: Top, the main components of the basement membrane and integrins. Bottom A to C: Effect of increasing flexibility of the collagen IV model from the stiffest to most flexible molecule on the organization of the basement membrane during simulation

The proportions of each molecule in the basement membrane model are listed in Table 2 and compared to the Matrigel basement membrane composition [72].

	Matrigel [72]	Basement membrane model	Molecular weight of the basement membrane

			components
Laminin-111	60%	62.33%	900 kDa [73]
Collagen IV	30%	28.04%	405 kDa [74]
Nidogen	8%	9.63%	139 kDa [63]
			1444 kDa

Table 2**:** Relative proportions calculated according to the molecular weight of the components of the Matrigel and the basement membrane model.

The numbers seem to support equal proportions between laminins, collagen and nidogen. The model is based on a 1:1 ratio stoichiometry. This ratio was chosen in our model because it maximizes networking and interactions while minimizing the number of molecules who cannot find a partner to interact with. These proportions were similar to those of the basement membrane extract Matrigel

Membrane proteins: integrins

Lipid bilayers and integrins are in close contact with basement membranes and define the boundaries of our modelled and simulated system. Until now integrins were locked in a separate 2D space with a rigid plane mimicking a membrane [26]. The constraints used did not have a physical justification. We used here a restraint field, as described in the Material and Methods section (membrane simulation), which can be used to simulate any transmembrane molecules. Forces calculated from Equation 1, for instance, maintain the transmembrane helix of the integrin models inside the simulated lipid bilayer (Figure 6). By using this newly implemented restraint field, the integrins behave as we anticipated. The restraint field is two sided, and while we only model the extracellular part, it will allows for intracellular content to be simulated as well in the future.

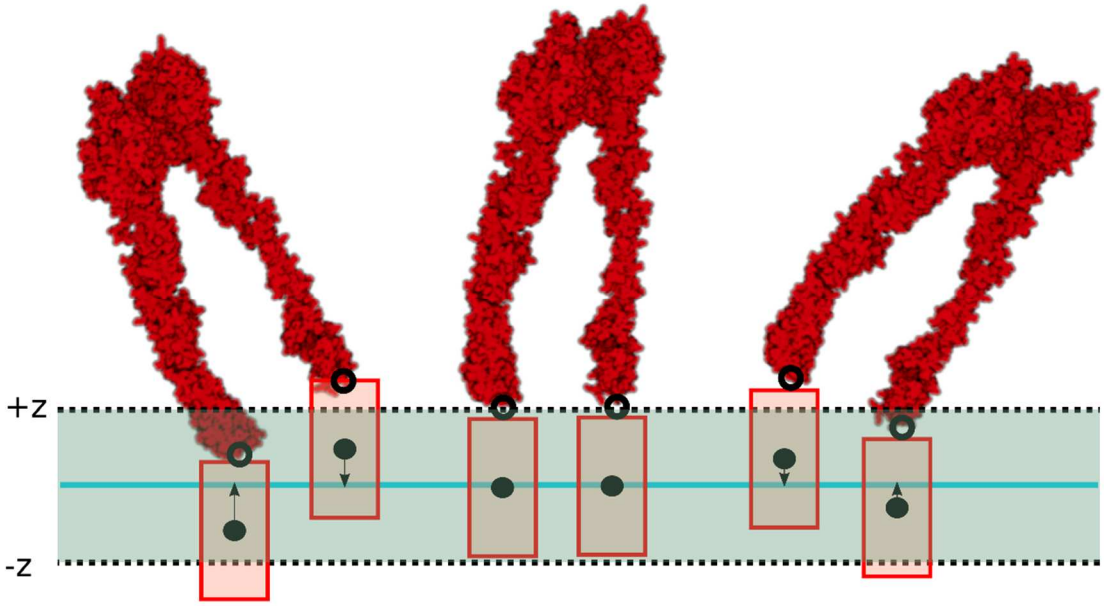

Figure 6: Integrin models embedded in a model of lipid bilayer. The black dotted lines represent the boundaries of the lipid bilayer (-z, +z), and the cyan line corresponds to the center of the membrane bilayer. The rigid bodies representing the hydrophobic transmembrane helices are represented as red rectangles. The arrows represent the direction of C_z as a function of the position of the center of mass of the rigid bodies in the IMPALA field (a force-field specific to lipid membranes simplified as a *continuum)*.

DISCUSSION

DURABIN was developed as a first contribution to fill the gap between microscopic and macroscopic observations. Experimental methods allowing the direct observation of biomolecular objects at the mesoscale are still lacking and simulation offers a tentative glimpse at systems that would otherwise elude us. We demonstrated that the combination of two incomplete datasets, namely the atomic details of microscopic observations of protein domains, with the global observation at the macroscopic scale is perfectly feasible for large molecular-weight protein. We have demonstrated that the tools we developed for DURABIN can simulate large motions in mesoscale systems like the basement membrane and hope it will help matrix biology researchers gain insight into the molecular structure of the basement membrane components at the atom level in a true multiscale fashion, all tools being simple to use and interactive thanks to virtual/augmented reality technologies.

Our rigid body to all-atom approach is promising but could certainly be improved. While transforming atomic coordinates using rigid body positions and rotations works, joining the heterogeneous bits before doing some minimization/optimisation on the resulting atomic model remains challenging. This two-stage approach is already commonly used in CGMD and the "Backward" back-mapping tool [75]. It uses a geometrical reconstruction approach starting from the backbone and library of geometrical rules. Our rigid body approach allows us to populate simulation with multiple copies of the rigid body model of the components of the basement membrane, but we plan to improve this approach. One of these improvements could come from position-based physics or particle-based rigid bodies. It is a recent development where rigid bodies are defined not by collider primitives but by a set of strongly constrained particles allowing for arbitrary shapes and unlocking computing parallelism [76]. This opens the possibility to perform simulations with a higher population of models in the simulation box. The N- terminal arms of the α, β and γ chains of different laminin molecules interact to form a laminin network [77,78]. This was indeed the case during the simulation of our model but the short arms of the β and γ chains belonging to the same were also able to interact. This could be prevented in the future by making the short arms of the β and γ chains rigid enough so that the β and γ arms of the same laminin molecule never come in contact.

This is another strength of the DURABIN approach. The rigid body models are easy to modify and adapt to the ever-changing large amount of research information available. The more information we get, the more accurate the simulation will be. It would be interesting to use DURABIN to test if and how the modifications of the basement membrane composition affect molecule diffusion, and basement membrane stability and physical properties, mimicking physio-pathological changes occurring *in vivo*. The model reported here is the first step towards the building of an extracellular matrix mesoscope, which will provide the opportunity to investigate molecular functions of the ECM and the biological processes it is involved in.

Basement membranes are usually associated with cells through interactions with adhesion receptors, sulfated glycolipids and others. We have not yet included in our model the interactions between basement membrane components and glycolipids, or dystroglycan which should provide new insights on the role of cell surface receptors in basement membrane self-assembly. Furthermore, the polarization of proteins in the basement membrane should also be considered. For example, the C-terminal globular domains of laminins interact with the cell surface, whereas the N-terminus of their three chains form the laminin network in the

basement membrane [79]. Collagen XVIII also has a polarized orientation in basement membranes with the N-terminus facing the fibrillar matrix and the C-terminus orientated towards the plasma membrane [80]. Last, collagen IV C-terminus dimerization was implemented, but we still have to fully model the tetramerization of the 7S domain, which is much more complex since it involves 12 chains, is cross-linked by lysyl oxidase like-2 but is currently modelled as a simple constraint.

Biomolecular non-covalent intermolecular bonds are not programmed to dissociate during the simulation, meaning that they are stable once created. The challenge is to find a way to translate values of the equilibrium dissociation constants into a programmable parameter in the simulation. This could be expressed as a contact frequency defining the probability a non-covalent bond/interaction could form or dissociate at each timestep.

Our approach developed with the Unity platform benefits from the native support of haptic and augmented/virtual reality (AR/VR) devices. The advantage of VR is that it makes easier for non-expert users to navigate the computer environment. From simply walking around the visualization to quite literally grabbing the model and turning it around, it offers very intuitive and new means of looking at molecular models [81,82]. Unity also provides real-time visual feedback to the user. DURABIN provides an easy way to apprehend visual medium for the users to observe and interact with complex mesoscale systems, but DURABIN lacks advanced molecular representation features like ribbons or cartoon representation, unlike Unitymol [30] or other visualization software such as VMD or PyMOL. DURABIN provides tools to import, simulate, manipulate, and export molecular models, but it does not aim at performing all-atom MD and at replacing well-known MD packages such as Gromacs, NAMD, Amber or LAMMPS. DURABIN speeds up the study of large molecular systems at the mesoscale level by facilitating the building and sampling of complex molecular systems such as ECM large multi-domains molecules.

ACKNOWLEDGEMENTS

This work was supported by Grand Reims (France) via the Chaire MAgICS (Modélisation moléculaire et Agroressources: Ingrédients, Cosmétique et Santé) to MD.

AUTHORS CONTRIBUTION

HW: methodology, software, formal analysis, investigation, data curation, visualization, writing – original draft. **JM**: writing - review & editing. **MD**: conceptualization, writing - review & editing, supervision, funding acquisition. **SRB:** writing - review & editing. **SB**:

writing - review & editing, supervision. **NB**: conceptualization, writing - review & editing, resources, project administration, funding acquisition.

REFERENCES

[1] A. Naba, K.R. Clauser, H. Ding, C.A. Whittaker, S.A. Carr, R.O. Hynes, The extracellular matrix: Tools and insights for the "omics" era, Matrix Biol. J. Int. Soc. Matrix Biol. 49 (2016) 10–24. https://doi.org/10.1016/j.matbio.2015.06.003.

[2] J.H. Miner, N.M. Nguyen, Extracellular Matrix: Basement Membranes, in: S.M. Janes (Ed.), Encycl. Respir. Med. Second Ed., Academic Press, Oxford, 2022: pp. 130–136. https://doi.org/10.1016/B978-0-08-102723-3.00046-9.

[3] D.F. Holmes, Y. Lu, T. Starborg, K.E. Kadler, Chapter Three - Collagen Fibril Assembly and Function, in: E.S. Litscher, P.M. Wassarman (Eds.), Curr. Top. Dev. Biol., Academic Press, 2018: pp. 107–142. https://doi.org/10.1016/bs.ctdb.2018.02.004.

[4] D. Rouède, E. Schaub, J.-J. Bellanger, F. Ezan, J.-C. Scimeca, G. Baffet, F. Tiaho, Determination of extracellular matrix collagen fibril architectures and pathological remodeling by polarization dependent second harmonic microscopy, Sci. Rep. 7 (2017) 12197. https://doi.org/10.1038/s41598-017-12398-0.

[5] G.T. Johnson, L. Autin, M. Al-Alusi, D.S. Goodsell, M.F. Sanner, A.J. Olson, cellPACK: a virtual mesoscope to model and visualize structural systems biology, Nat. Methods. 12 (2015) 85–91. https://doi.org/10.1038/nmeth.3204.

[6] T. Nakane, A. Kotecha, A. Sente, G. McMullan, S. Masiulis, P.M.G.E. Brown, I.T. Grigoras, L. Malinauskaite, T. Malinauskas, J. Miehling, T. Uchański, L. Yu, D. Karia, E.V. Pechnikova, E. de Jong, J. Keizer, M. Bischoff, J. McCormack, P. Tiemeijer, S.W. Hardwick, D.Y. Chirgadze, G. Murshudov, A.R. Aricescu, S.H.W. Scheres, Single-particle cryo-EM at atomic resolution, Nature. 587 (2020) 152–156. https://doi.org/10.1038/s41586-020-2829-0.

[7] K. Murata, M. Wolf, Cryo-electron microscopy for structural analysis of dynamic biological macromolecules, Biochim. Biophys. Acta BBA - Gen. Subj. 1862 (2018) 324–334. https://doi.org/10.1016/j.bbagen.2017.07.020.

[8] W. Chen, J. Lou, J. Hsin, K. Schulten, S.C. Harvey, C. Zhu, Molecular Dynamics Simulations of Forced Unbending of Integrin αVβ3, PLOS Comput. Biol. 7 (2011) e1001086. https://doi.org/10.1371/journal.pcbi.1001086.

[9] J. Djajamuliadi, T.F. Kagawa, K. Ohgo, K.K. Kumashiro, Insights into a putative hinge region in elastin using molecular dynamics simulations, Matrix Biol. 28 (2009) 92–100. https://doi.org/10.1016/j.matbio.2008.12.001.

[10] J.W. Bourne, P.A. Torzilli, Molecular simulations predict novel collagen conformations during cross-link loading, Matrix Biol. 30 (2011) 356–360. https://doi.org/10.1016/j.matbio.2011.03.010.

[11] A. Tarakanova, G.C. Yeo, C. Baldock, A.S. Weiss, M.J. Buehler, Tropoelastin is a Flexible Molecule that Retains its Canonical Shape, Macromol. Biosci. 19 (2019) 1800250. https://doi.org/10.1002/mabi.201800250.

[12] A. Nash, M. Notou, A.F. Lopez-Clavijo, L. Bozec, N.H. de Leeuw, H.L. Birch, Glucosepane is associated with changes to structural and physical properties of collagen fibrils, Matrix Biol. Plus. 4 (2019) 100013. https://doi.org/10.1016/j.mbplus.2019.100013.

[13] J. Ozsvar, A. Tarakanova, R. Wang, M.J. Buehler, A.S. Weiss, Allysine modifications perturb tropoelastin structure and mobility on a local and global scale, Matrix Biol. Plus. 2 (2019) 100002. https://doi.org/10.1016/j.mbplus.2019.03.001.

[14] Molecular dynamics simulations in biology | Nature, (n.d.). https://www.nature.com/articles/347631a0 (accessed June 1, 2021).

[15] A. Pozzi, P.D. Yurchenco, R.V. Iozzo, The nature and biology of basement membranes, Matrix Biol. 57–58 (2017) 1–11. https://doi.org/10.1016/j.matbio.2016.12.009.
[16] J.R. Perilla, K. Schulten, Physical properties of the HIV-1 capsid from all-atom molecular dynamics simulations, Nat. Commun. 8 (2017) 15959. https://doi.org/10.1038/ncomms15959.
[17] F. Feng, W.S. Klug, Finite element modeling of lipid bilayer membranes, J. Comput. Phys. 220 (2006) 394–408. https://doi.org/10.1016/j.jcp.2006.05.023.
[18] C. Bae, P.J. Butler, Finite element analysis of microelectrotension of cell membranes, Biomech. Model. Mechanobiol. 7 (2008) 379–386. https://doi.org/10.1007/s10237-007-0093-y.
[19] A. Gautieri, A. Russo, S. Vesentini, A. Redaelli, M.J. Buehler, Coarse-Grained Model of Collagen Molecules Using an Extended MARTINI Force Field, J. Chem. Theory Comput. 6 (2010) 1210–1218. https://doi.org/10.1021/ct100015v.
[20] A. Gautieri, S. Vesentini, A. Redaelli, R. Ballarini, Modeling and measuring visco-elastic properties: From collagen molecules to collagen fibrils, Int. J. Non-Linear Mech. 56 (2013) 25–33. https://doi.org/10.1016/j.ijnonlinmec.2013.03.012.
[21] A. Tarakanova, J. Ozsvar, A.S. Weiss, M.J. Buehler, Coarse-grained model of tropoelastin self-assembly into nascent fibrils, Mater. Today Bio. 3 (2019). https://doi.org/10.1016/j.mtbio.2019.100016.
[22] An Introduction to Physically Based Modeling, (n.d.). https://www.cs.cmu.edu/~baraff/pbm/pbm.html (accessed May 24, 2021).
[23] OpenFOAM | Free CFD Software | The OpenFOAM Foundation, (n.d.). https://openfoam.org/ (accessed November 19, 2021).
[24] H.G. Weller, G. Tabor, H. Jasak, C. Fureby, A tensorial approach to computational continuum mechanics using object-oriented techniques, Comput. Phys. 12 (1998) 620–631. https://doi.org/10.1063/1.168744.
[25] ANSYS System Hardware Requirements, Ozen Eng. ANSYS. (n.d.). https://www.ozeninc.com/ansys-system-hardware-requirements/ (accessed May 18, 2021).
[26] H. Wong, J. Prévoteau-Jonquet, S. Baud, M. Dauchez, N. Belloy, Mesoscopic Rigid Body Modelling of the Extracellular Matrix Self-Assembly, J. Integr. Bioinforma. 15 (2018). https://doi.org/10.1515/jib-2018-0009.
[27] K. Böse, R. Nischt, A. Page, B.L. Bader, M. Paulsson, N. Smyth, Loss of Nidogen-1 and -2 Results in Syndactyly and Changes in Limb Development *, J. Biol. Chem. 281 (2006) 39620–39629. https://doi.org/10.1074/jbc.M607886200.
[28] J.A. Engbring, H.K. Kleinman, The basement membrane matrix in malignancy, J. Pathol. 200 (2003) 465–470. https://doi.org/10.1002/path.1396.
[29] G. Wheeler, S. Deng, N. Toussaint, K. Pushparajah, J.A. Schnabel, J.M. Simpson, A. Gomez, Virtual interaction and visualisation of 3D medical imaging data with VTK and Unity, Healthc. Technol. Lett. 5 (2018) 148–153. https://doi.org/10.1049/htl.2018.5064.
[30] S. Doutreligne, C. Gageat, T. Cragnolini, A. Taly, S. Pasquali, P. Derreumaux, M. Baaden, UnityMol: interactive and ludic visual manipulation of coarse-grained RNA and other biomolecules, in: 2015 IEEE 1st Int. Workshop Virtual Augment. Real. Mol. Sci. VARMSIEEEVR, 2015: pp. 1–6. https://doi.org/10.1109/VARMS.2015.7151718.
[31] S.K. Burley, C. Bhikadiya, C. Bi, S. Bittrich, L. Chen, G.V. Crichlow, C.H. Christie, K. Dalenberg, L. Di Costanzo, J.M. Duarte, S. Dutta, Z. Feng, S. Ganesan, D.S. Goodsell, S. Ghosh, R.K. Green, V. Guranović, D. Guzenko, B.P. Hudson, C.L. Lawson, Y. Liang, R. Lowe, H. Namkoong, E. Peisach, I. Persikova, C. Randle, A. Rose, Y. Rose, A. Sali, J. Segura, M. Sekharan, C. Shao, Y.-P. Tao, M. Voigt, J.D. Westbrook, J.Y. Young, C. Zardecki, M. Zhuravleva, RCSB Protein Data Bank: powerful new tools for exploring

3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences, Nucleic Acids Res. 49 (2021) D437–D451. https://doi.org/10.1093/nar/gkaa1038.
[32] M. Wojtyra, Joint reactions in rigid body mechanisms with dependent constraints, Mech. Mach. Theory. 44 (2009) 2265–2278. https://doi.org/10.1016/j.mechmachtheory.2009.07.008.
[33] H. Wong, H. Marie-Nelly, S. Herbert, P. Carrivain, H. Blanc, R. Koszul, E. Fabre, C. Zimmer, A Predictive Computational Model of the Dynamic 3D Interphase Yeast Nucleus, Curr. Biol. 22 (2012) 1881–1890. https://doi.org/10.1016/j.cub.2012.07.069.
[34] P. Ducarme, A. Thomas, R. Brasseur, The optimisation of the helix/helix interaction of a transmembrane dimer is improved by the IMPALA restraint field, Biochim. Biophys. Acta BBA - Biomembr. 1509 (2000) 148–154. https://doi.org/10.1016/S0005-2736(00)00290-X.
[35] Z.V. Leonenko, E. Finot, H. Ma, T.E.S. Dahms, D.T. Cramb, Investigation of Temperature-Induced Phase Transitions in DOPC and DPPC Phospholipid Bilayers Using Temperature-Controlled Scanning Force Microscopy, Biophys. J. 86 (2004) 3783–3793. https://doi.org/10.1529/biophysj.103.036681.
[36] P. Parois, R.I. Cooper, A.L. Thompson, Crystal structures of increasingly large molecules: meeting the challenges with CRYSTALS software, Chem. Cent. J. 9 (2015). https://doi.org/10.1186/s13065-015-0105-4.
[37] M.P. Foster, C.A. McElroy, C.D. Amero, Solution NMR of large molecules and assemblies, Biochemistry. 46 (2007) 331–340. https://doi.org/10.1021/bi0621314.
[38] M.A. Herzik Jr, Cryo-electron microscopy reaches atomic resolution, Nature. 587 (2020) 39–40. https://doi.org/10.1038/d41586-020-02924-y.
[39] M. Wu, G.C. Lander, How low can we go? Structure determination of small biological complexes using single-particle cryo-EM, Curr. Opin. Struct. Biol. 64 (2020) 9–16. https://doi.org/10.1016/j.sbi.2020.05.007.
[40] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature. 596 (2021) 583–589. https://doi.org/10.1038/s41586-021-03819-2.
[41] AlphaFold Protein Structure Database, (n.d.). https://alphafold.ebi.ac.uk/ (accessed November 19, 2021).
[42] K.M. Ruff, R.V. Pappu, AlphaFold and Implications for Intrinsically Disordered Proteins, J. Mol. Biol. 433 (2021) 167208. https://doi.org/10.1016/j.jmb.2021.167208.
[43] RCSB PDB: Homepage, (n.d.). https://www.rcsb.org/ (accessed May 23, 2021).
[44] C.L. Lawson, A. Patwardhan, M.L. Baker, C. Hryc, E.S. Garcia, B.P. Hudson, I. Lagerstedt, S.J. Ludtke, G. Pintilie, R. Sala, J.D. Westbrook, H.M. Berman, G.J. Kleywegt, W. Chiu, EMDataBank unified data resource for 3DEM, Nucleic Acids Res. 44 (2016) D396–D403. https://doi.org/10.1093/nar/gkv1126.
[45] The Electron Microscopy Data Bank, (n.d.). https://www.ebi.ac.uk/pdbe/emdb/ (accessed May 23, 2021).
[46] J. Söding, A. Biegert, A.N. Lupas, The HHpred interactive server for protein homology detection and structure prediction, Nucleic Acids Res. 33 (2005) W244–W248. https://doi.org/10.1093/nar/gki408.

[47] S. Wu, J. Skolnick, Y. Zhang, Ab initio modeling of small proteins by iterative TASSER simulations, BMC Biol. 5 (2007) 17. https://doi.org/10.1186/1741-7007-5-17.
[48] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815. https://doi.org/10.1006/jmbi.1993.1626.
[49] R.W. Naylor, M.R.P.T. Morais, R. Lennon, Complexities of the glomerular basement membrane, Nat. Rev. Nephrol. 17 (2021) 112–127. https://doi.org/10.1038/s41581-020-0329-y.
[50] J. Khoshnoodi, V. Pedchenko, B.G. Hudson, Mammalian collagen IV, Microsc. Res. Tech. 71 (2008) 357–370. https://doi.org/10.1002/jemt.20564.
[51] S. Ricard-Blum, The Collagen Family, Cold Spring Harb. Perspect. Biol. 3 (2011). https://doi.org/10.1101/cshperspect.a004978.
[52] A. Takaya, T. Yamamoto, K. Tokoyoda, Humoral Immunity vs. Salmonella, Front. Immunol. 10 (2020). https://doi.org/10.3389/fimmu.2019.03155.
[53] C. Wagner, M. Polke, R.G. Gerlach, D. Linke, Y.-D. Stierhof, H. Schwarz, M. Hensel, Functional dissection of SiiE, a giant non-fimbrial adhesin of Salmonella enterica, Cell. Microbiol. 13 (2011) 1286–1301. https://doi.org/10.1111/j.1462-5822.2011.01621.x.
[54] P. Lössl, K. Kölbel, D. Tänzler, D. Nannemann, C.H. Ihling, M.V. Keller, M. Schneider, F. Zaucke, J. Meiler, A. Sinz, Analysis of Nidogen-1/Laminin γ1 Interaction by Cross-Linking, Mass Spectrometry, and Computational Modeling Reveals Multiple Binding Modes, PLoS ONE. 9 (2014) e112886. https://doi.org/10.1371/journal.pone.0112886.
[55] G.P. Tuszynski, M.A. Kowalska, Thrombospondin-induced adhesion of human platelets, J. Clin. Invest. 87 (1991) 1387–1394. https://doi.org/10.1172/JCI115144.
[56] J. Zhu, B.-H. Luo, T. Xiao, C. Zhang, N. Nishida, T.A. Springer, Structure of a complete integrin ectodomain in a physiologic resting state and activation and deactivation by applied forces, Mol. Cell. 32 (2008) 849–861. https://doi.org/10.1016/j.molcel.2008.11.018.
[57] R. Colin Hughes, Chapter 14 - Adhesive glycoproteins and receptors, in: J. Montreuil, J.F.G. Vliegenthart, H. Schachter (Eds.), New Compr. Biochem., Elsevier, 1997: pp. 507–570. https://doi.org/10.1016/S0167-7306(08)60627-4.
[58] A.M. Belkin, M.A. Stepp, Integrins as receptors for laminins, Microsc. Res. Tech. 51 (2000) 280–301. https://doi.org/10.1002/1097-0029(20001101)51:3<280::AID-JEMT7>3.0.CO;2-O.
[59] E. Petitclerc, A. Boutaud, A. Prestayko, J. Xu, Y. Sado, Y. Ninomiya, M.P. Sarras, B.G. Hudson, P.C. Brooks, New Functions for Non-collagenous Domains of Human Collagen Type IV: NOVEL INTEGRIN LIGANDS INHIBITING ANGIOGENESIS AND TUMOR GROWTHIN VIVO *, J. Biol. Chem. 275 (2000) 8051–8061. https://doi.org/10.1074/jbc.275.11.8051.
[60] P.D. Yurchenco, Y.S. Cheng, Self-assembly and calcium-binding sites in laminin. A three-arm interaction model, J. Biol. Chem. 268 (1993) 17286–17299.
[61] R. Sekiguchi, K.M. Yamada, Chapter Four - Basement Membranes in Development and Disease, in: E.S. Litscher, P.M. Wassarman (Eds.), Curr. Top. Dev. Biol., Academic Press, 2018: pp. 143–191. https://doi.org/10.1016/bs.ctdb.2018.02.005.
[62] M. Paulsson, R. Deutzmann, M. Dziadek, H. Nowack, R. Timpl, S. Weber, J. Engel, Purification and structural characterization of intact and fragmented nidogen obtained from a tumor basement membrane, Eur. J. Biochem. 156 (1986) 467–478. https://doi.org/10.1111/j.1432-1033.1986.tb09605.x.
[63] T.R. Patel, C. Bernards, M. Meier, K. McEleney, D.J. Winzor, M. Koch, J. Stetefeld, Structural elucidation of full-length nidogen and the laminin–nidogen complex in solution, Matrix Biol. 33 (2014) 60–67. https://doi.org/10.1016/j.matbio.2013.07.009.

[64] U. Technologies, Unity - Manual: Joints, (n.d.). https://docs.unity3d.com/Manual/Joints.html (accessed November 17, 2021).
[65] S. Neukirch, A. Goriely, A.C. Hausrath, Chirality of Coiled Coils: Elasticity Matters, Phys. Rev. Lett. 100 (2008) 038105. https://doi.org/10.1103/PhysRevLett.100.038105.
[66] P. Moreno-Layseca, J. Icha, H. Hamidi, J. Ivaska, Integrin trafficking in cells and tissues, Nat. Cell Biol. 21 (2019) 122–132. https://doi.org/10.1038/s41556-018-0223-z.
[67] W. Halfter, J. Candiello, H. Hu, P. Zhang, E. Schreiber, M. Balasubramani, Protein composition and biomechanical properties of in vivo-derived basement membranes, Cell Adhes. Migr. 7 (2013) 64–71. https://doi.org/10.4161/cam.22479.
[68] J. Bella, A first census of collagen interruptions: Collagen's own stutters and stammers, J. Struct. Biol. 186 (2014) 438–450. https://doi.org/10.1016/j.jsb.2014.03.022.
[69] P.D. Yurchenco, Basement membrane structure in situ: evidence for lateral associations in the type IV collagen network, J. Cell Biol. 105 (1987) 2559–2568. https://doi.org/10.1083/jcb.105.6.2559.
[70] A.M. Abreu-Velez, M.S. Howard, Collagen IV in Normal Skin and in Pathological Processes, North Am. J. Med. Sci. 4 (2012) 1–8. https://doi.org/10.4103/1947-2714.92892.
[71] C. Añazco, A.J. López-Jiménez, M. Rafi, L. Vega-Montoto, M.-Z. Zhang, B.G. Hudson, R.M. Vanacore, Lysyl Oxidase-like-2 Cross-links Collagen IV of Glomerular Basement Membrane, J. Biol. Chem. 291 (2016) 25999–26012. https://doi.org/10.1074/jbc.M116.738856.
[72] Corning application note, (n.d.). https://www.corning.com/catalog/cls/documents/application-notes/Application_Note_CLS-DL-AN-414_Matrigel_Matrix_3D_In_Vitro_Protocol.pdf.
[73] E. Hohenester, Structural biology of laminins, Essays Biochem. 63 (2019) 285–295. https://doi.org/10.1042/EBC20180075.
[74] R. Timpl, H. Wiedemann, V. van Delden, H. Furthmayr, K. Kühn, A network model for the organization of type IV collagen molecules in basement membranes, Eur. J. Biochem. 120 (1981) 203–211. https://doi.org/10.1111/j.1432-1033.1981.tb05690.x.
[75] T.A. Wassenaar, K. Pluhackova, R.A. Böckmann, S.J. Marrink, D.P. Tieleman, Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models, J. Chem. Theory Comput. 10 (2014) 676–690. https://doi.org/10.1021/ct400617g.
[76] Chapter 29. Real-Time Rigid Body Simulation on GPUs, NVIDIA Dev. (n.d.). https://developer.nvidia.com/gpugems/gpugems3/part-v-physics-simulation/chapter-29-real-time-rigid-body-simulation-gpus (accessed June 3, 2021).
[77] R.W. Naylor, M.R.P.T. Morais, R. Lennon, Complexities of the glomerular basement membrane, Nat. Rev. Nephrol. 17 (2021) 112–127. https://doi.org/10.1038/s41581-020-0329-y.
[78] P.D. Yurchenco, Basement Membranes: Cell Scaffoldings and Signaling Platforms, Cold Spring Harb. Perspect. Biol. 3 (2011) a004911–a004911. https://doi.org/10.1101/cshperspect.a004911.
[79] R. Jayadev, D.R. Sherwood, Basement membranes, Curr. Biol. 27 (2017) R207–R211. https://doi.org/10.1016/j.cub.2017.02.006.
[80] R. Heljasvaara, M. Aikio, H. Ruotsalainen, T. Pihlajaniemi, Collagen XVIII in tissue homeostasis and dysregulation — Lessons learned from model organisms and human patients, Matrix Biol. 57–58 (2017) 55–75. https://doi.org/10.1016/j.matbio.2016.10.002.

[81] T. Davison, F. Samavati, C. Jacob, LifeBrush: Painting, simulating, and visualizing dense biomolecular environments, Comput. Graph. 82 (2019) 232–242. https://doi.org/10.1016/j.cag.2019.05.006.
[82] X. Martinez, M. Krone, N. Alharbi, A.S. Rose, R.S. Laramee, S. O'Donoghue, M. Baaden, M. Chavent, Molecular Graphics: Bridging Structural Biologists and Computer Scientists, Structure. 27 (2019) 1617–1623. https://doi.org/10.1016/j.str.2019.09.001.