Chapter

The Importance of the Extracellular Matrix in HPV-Associated Diseases

Joana Sampaio, Joana Ferreira, Ana Carolina Santos, Manuel Bicho and Maria Clara Bicho

Abstract

The extracellular matrix (ECM) is the non-cellular component of the tissues of our organism. It is the dynamic element that maintains a biochemical structure capable of supporting the organization and architecture of the tissue constituents. The diversity of ECM's constituents gives it the biochemical and biophysical properties necessary to regulate its behavior and differentiation. ECM has an important role in the biology of cancer cell development and progression. Human papillomavirus infection (HPV) is the principal etiological agent of the most common sexually transmitted diseases. It is a virus that can cause lesions precursors of epithelial squamous and glandular tumors. Type 16 (HPV16) is the leading cause of pre-malignant lesions and invasive cancers in these tissues. This work will focus on HPV infection to understand the role of ECM in the invasion, spread, and pathogenesis of the lesions caused by this virus. Cancer is no longer considered a pathology explained only by uncontrolled proliferation and apoptosis but also by the deregulation of the microenvironment.

The in-depth knowledge of ECM dynamics and its complexity is central and promising, specifically in developing new targeted therapies.

Keywords: Extracellular matrix, human papillomavirus, heparan sulfate proteoglycans, metalloproteinases, heparanase

1. Introduction

The extracellular matrix (ECM) can be defined as a three-dimensional, non-cellular macromolecular network made of collagen, proteoglycans, elastin, fibronectin, laminins, and other glycoproteins structural support for the organization of cellular constituents. It is known to be a physiologically active component of living tissue. The various parts of the matrix bind to each other and cell adhesion receptors, forming a complex network on which cells rely in all tissues and organs. Through transduction of extracellular signals originating from the ECM, cell surface receptors regulate diverse cellular functions such as survival, growth, migration, and differentiation and are vital for homeostasis maintenance [1, 2].

Each organ has a unique combination of elements in its constitution so that the specific function of the tissue itself can be fulfilled. This unique composition arises

from biophysical and biochemical feedback between cellular components and the microenvironment, where they are inserted during the genesis of the tissue. The ECM continuously undergoes remodeling mediated by different decomposition enzymes, the proteinases being a highly dynamic structure. The balance between ECM degradation and secretion, orchestrated by ECM-modifying cells, is responsible for tensional homeostasis and organ-specific properties such as elasticity and compressibility [2–6].

In vitro, most animal cells have been shown to remain viable only when adherent to a substrate. Thus, the cell relies heavily on its sense of touch to survive by adhering and spatially interacting with the surrounding ECM- the concept of anoikis. The ability to attach and communicate with its environment is responsible for several growth factor receptors, and adhesion molecules arranged along the cell membrane, namely integrins. Indeed, cells have been shown to translate signals from the ECM to coordinate crucial morphological organization and signal events by regulating gene transcription. A cell converts external mechanical stimuli into a downstream intracellular chemical signal. This process is known as mechanotransduction. The sensitivity with which cells respond to biophysical and biochemical signals from the ECM demonstrates the importance of tissue homeostasis in maintaining healthy host cells.

Consequently, dysregulation of ECM remodeling has been shown to contribute significantly to cellular evolution through various fibrotic conditions, characterized by excessive ECM deposition and increased ECM stiffness. Due to increased interstitial pressure, irreversible loss of tissue homeostasis has been linked to an increased risk for various pathological conditions such as osteoarthritis, cardiovascular disease, and cancer [3, 4, 6–10].

2. ECM constitution

Various proteins that constitute the ECM result in different structures and properties. The main components of ECM include collagen, proteoglycans, laminin, and fibronectin. As each structure has a function, different subtypes and combinations of ECM molecules confer different functions essential to the correct functioning of the whole organism, **Figure 1** [11, 12].

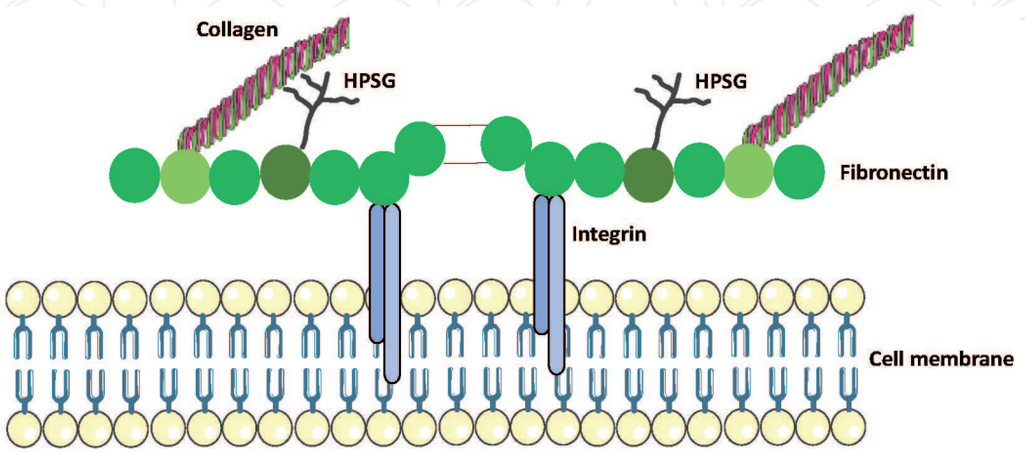

Figure 1.
ECM components and their organization. Organization of the different collagens, proteoglycans (HPSG), laminins, and fibronectin in the basement membrane and the extracellular matrix. In the basement membrane, laminin is attached to the cell, forming a fibrillar network. It is then linked to the type IV collagen network through nidogen and proteoglycans such as perlecan and agrinin. The different proteoglycans hold the fibrils together to form a collagen fiber. Fibronectin is bound to the cell by integrins and syndecans.

2.1 Collagen

Collagen is the basis of ECM architecture, is the most significant functional component, and the most abundant protein in human tissue. It can be classified into fibrillar (I-III, V, and XI) and non-fibrillar. Collagen fibers give ECM its tensile strength by limiting tissue distensibility. Collagens are trimeric molecules composed of three α-polypeptide chains that contain the sequence repeat (G-X-Y). This repeat allows the formation of a triple helix that gives the characteristic structure of this superfamily, **Figure 1**. Currently, there are 28 unique subtypes of collagen discovered. Each member of the collagen family has at least one triple helix domain. Most collagens bind to and interact with several ECM proteins forming supramolecular aggregates [3, 6, 8, 11].

Fibrillar collagens form fibrous structures often found in tendons, cartilage, skin, and cornea. Each collagen fiber is made up of several collagen subtypes in response to their tissue location.

Fundamentally we can define four classes of collagens:

1. Fibril-forming collagens (I, II, III, V, XI, XXVI, XXVII)

2. Fibril-associated collagens with interrupted helices (FACITs) (IX, XII, XIV, XVI, XIX, XX, XXI, XXII, XXIV)

3. Network-forming collagens (IV, VII, X)

4. Membrane anchored collagens (MACITs) (XIII, XVII, XXIII, XXV)

5. Short collagens (XXVI, XXVIII)

Types I, III, and V, predominantly produced by fibroblasts, are essential for the structure of the interstitial matrix. Their function as a pericellular "glue and structure" is necessary in tissue repair. Type IV is primarily located in the basement membrane and underlying epithelial or endothelial cells, ensuring their specialized polarization and function. Type IV collagen is the main constituent of basement membranes in tissues such as the lung, kidney, skin, intestine, and liver. It is mainly produced by endothelial and epithelial cells, and seen as intelligent collagen necessary for tissue repair processes that allow polarized cells (endothelial and epithelial) to survive and function, enable regular tissue function [6, 13].

2.2 Proteoglycans

Proteoglycans are the functional modifiers of ECM and provide additional properties. They are proteins characterized by being covalently linked to glycosaminoglycans (GAGs). These GAGs are long, negatively charged chains conferred by sulfate and carboxyl groups (heparan sulfate (HS), dermatan sulfate, and keratan sulfate (KS). The addition of sulfate and carboxyl groups to GAGs gives them this negative charge, enabling them to sequester water and cations (sodium and calcium) and have cellular lubrication and filling functions, **Figure 1** [6, 12, 14].

There are about three dozen extracellular matrix proteoglycans encoded in mammalian genomes, divided into several families. The two largest families include the LRR (leucine-rich repeat) proteoglycans and the hyalectans (such as versican). In addition to those mentioned, perhaps the most significant of all is perlecan (HSPG2), a multidomain protein that is part of all basement membranes. There are also two small families of transmembrane proteoglycans: glypicans and syndecans, both of which have heparan sulfate side chains, as does CD44 [12, 14].

Syndecans are encoded by four different genes and represent the most abundant transmembrane heparan sulfate proteoglycans (HSPGs). Their core protein comprises an extracellular domain, a single transmembrane domain, and a short cytoplasmic domain that interacts with the cell cytoskeleton. Glypicans are encoded by six different genes and are anchored to the cell membrane via glycosylphosphatidylinositol (GPI). The multiple organs of the human body contain different isoforms of HSPG with various polysaccharide compositions and sulfation patterns [12, 15].

The extracellular domain of syndecans is intrinsically disorganized, a feature that allows it to interact with a huge variety of molecules and perform a wide variety of biological functions. Some of these functions involve acting as co-receptors for tyrosine kinase receptors linked to growth factors. The HS chains of these proteoglycans share various ligands, such as matrix proteins, growth factors, cytokines, and chemokines, which are presented to high-affinity receptors present on the cell surface [3, 4, 6, 10].

In conclusion, proteoglycans are highly variable in their shape and structure to exert different functions in the ECM, i.e., and they are highly pleiotropic. This characteristic makes them essential in maintaining a healthy ECM, and without them, its entire structure would collapse.

2.3 Laminins

Laminins are trimeric glycoproteins formed by α, β, and γ chains often found in the basal lamina and mesenchymal compartments. The three chains form a coiled α-helical structure that builds the long arm, while the three short arms are each composed of one chain. At the end of the long arm are five laminated G-type (LG) domains of the α-chain that serve as binding points for the cell. Integrins, dystroglycan, lutheran glycoprotein, or sulfated glycolipids bind to these LG domains. At the end of each short arm are the N-terminal laminin (LN) domains that are important for laminin polymerization and the formation of the basement membrane [5, 6, 11].

Laminins have cell-specific functions such as adhesion, differentiation, migration, maintenance of phenotype, and resistance to apoptosis (anoikis). By binding to integrins, laminins can create a dynamic link between the cell and the ECM. The unique heterotrimeric laminins have the integrins as anchored partners allowing the induction of signaling pathways and the organization of the intracellular cytoskeleton. It has been observed that heparan sulfates directly mediate the interaction between laminins and collagen IV. Laminins play crucial roles in both basal membrane formation and interactions between cells and the ECM. While collagen, proteoglycans, and hyaluronic acid constitute the main structural component of the MEC, laminins are one of the molecules that bridge the cell-ECM interaction gap [6, 11, 12, 14].

2.4 Fibronectin

Fibronectin is a high molecular weight protein composed of two subunits linked together by two cysteine persulfide bonds. It is secreted in a soluble form by hepatocytes into the bloodstream or expressed in tissues by fibroblasts, forming a fibrillar network. The structure of fibronectin and its multiple post-translational modifications result in an immense variety of interactions with various ECM components (growth factors and GAGs) that mediate cell attachment and motility, ECM remodeling, host-pathogen interactions, among others. A single gene encodes this glycoprotein with 20 human isoforms resulting from alternative mRNA excisions

(primary transcript). Similar to collagen, fibronectin forms a fibrillar network in the ECM. The structure of the fibronectin matrix is mediated by selective binding to α5β1 integrins. Compact, soluble fibronectin is secreted and unfolded through these integrins, revealing specific binding sites for other fibronectin molecules to form its fibrillar network. Binding to fibronectin induces integrin aggregation, which provides high local concentrations of fibronectin on the cell surface. This phenomenon promotes fibronectin–integrin interactions through the N assembly domains of each molecule [5, 6, 16].

Once fibronectin is attached to the cell surface by integrins, the actin cytoskeleton can drag molecules to the fibronectin to change its conformation. This will affect the C-terminal regions of fibronectin, revealing binding sites for fibronectin, heparan sulfates, heparin, collagen, and other ECM molecules. Through strong non-covalent protein–protein interactions, the fibronectin network matures and becomes insoluble, although other ECM proteins can mediate mature lateral interactions between fibrils. These interactions stabilize the relatively weaker binding sites [5, 6, 16].

In conclusion, fibronectin works as a skeleton upon which the bioavailability and activity of various growth factors are orchestrated. The interaction of fibronectin with growth factors (e.g., TGF-β, PDGF, HGF, VEGF, FGF) can impact cell migration, cell proliferation, survival signals, and angiogenesis as downstream outcomes of their activation through mechanical or enzymatic activation [5, 6, 14, 16].

3. Functions of the extra cellular matrix

The countless unique molecules that are part of the constitution of ECM give it various functions that simultaneously influence biochemical and biophysical processes in the cell. Although ECM was for many years considered an inert component that only provided structure to cells in tissue formation, its role in determining cellular functions and phenotypes has been clarified in the last two decades [2, 6, 14, 17].

The many proteolytic processes that modify ECM, by the action of proteolytic enzymes, play roles in ECM remodeling and are thought to release ECM-binding growth factors and expose cryptic activities in ECM, including the release of anti-angiogenic factors. Similarly, enzymes that degrade GAGs, such as heparanases and sulfatases, can also change the properties of proteoglycans in the ECM. Remodeling of the ECM by these various processes has important effects on development and associated pathologies [2, 14, 17].

Finally, ECM is known to transmit mechanical signals to cells and activate intracellular signaling mechanisms and the cytoskeleton machinery. The importance of ECM and its varied functions in the development and maintenance of cellular balance or homeostasis is indisputable [2, 14, 17].

3.1 Migration and proliferation

Cell migration is essential for tissue development, a fact demonstrated by neural crest cells, which migrate from the periphery of the neural tube to different parts of the embryo to form the heart, nerves, skin, and skull [2, 6].

The ECM influences the path and speed of migrating cells through its topography, composition, and physical properties. Cells migrate from regions with low ECM concentration to high ECM concentrations due to an adhesion gradient called haptotaxis. Proteases that degrade ECM also play a facilitating role in cell

migration through a process involving matrix metalloproteinases (MMPs), their inhibitors, among other enzymes. It is essential to realize that constant remodeling of the ECM occurs during development [2, 4, 6].

3.2 Development

The topographic variation in the structure and elasticity of ECM provides cells with the ability to adapt and form complex morphological systems that are essential for different organs [2, 4, 6].

ECM modulates tissue growth to form complex structures that are necessary for these organs to function. In addition, it provides structural organization not only through its action as a physical barrier to cell growth and by activating intracellular signaling in a time- and context-dependent manner. ECM exerts this effect by modulating the distribution of growth factors, physical anisotropy, and anchoring [2, 4, 6, 11].

ECM also has an essential role in highlighting the influence on cell fate. This role is shown in mammary gland differentiation. Even with hormonal stimulation in vitro, mammary gland cells do not secrete milk proteins. However, after exposure of these cells to laminin-1, they begin to secrete these proteins. This phenomenon indicates to us that an appropriate ECM microenvironment is indispensable for the cell to be able to fulfill its functions [2, 4, 6].

To conclude, cells can sense the physical properties of the adjacent matrix and activate the appropriate intracellular mechanisms for their differentiation. Therefore, the physical properties of the ECM have cell differentiation capabilities.

3.3 Tissue homeostasis

ECM is a highly dynamic structure. Even after development, ECM is constantly being deposited, degraded, and modified to maintain tissue homeostasis. This is especially important for preserving cell phenotype and physiological processes such as wound healing, angiogenesis, and bone remodeling [2, 4, 6].

To maintain tissue homeostasis, cells in contact with the ECM perceive ECM properties through receptors and adhesion complexes. In turn, the cell regulates the expression of ECM components and enzymes based on the signals it receives. This leads to a feedback mechanism in which cell also influences ECM, resulting in a balance of deposition and degradation of ECM components [2, 3, 6].

The response of cells to other stimuli is ultimately influenced by the ECM components. The complexity and importance of the feedback mechanism between the ECM and the cell is essential to maintain tissue homeostasis [2, 3, 6].

The imbalance in ECM deposition and degradation leads to disease and is a hallmark of cancer and other conditions that course with fibrosis. Overall, the role of ECM in tissue homeostasis is to direct the appropriate cellular response and phenotype to maintain mechanical integrity and tissue function [6].

4. ECM disruption in cancer progression

Traditional views of cancer have been changing, and the significant role of ECM in the regulation of cell proliferation, migration, and apoptosis have been highlighted. At the microscopic level, the organization of ECM constituents forms a specific microenvironment that plays a critical role in tumor progression. ECM is constantly remodeling and actively influences cell adhesion and migration. Thus,

small changes in the homeostasis of the microenvironment can result in significant effects on cancer cell proliferation. As the main component of ECM, collagen, dictates the main properties of the matrix, changes in its deposition or degradation can lead to loss of ECM homeostasis [2–4, 6, 11].

As cancer cells proliferate, the surrounding matrix changes a dynamic interaction between cells and the microenvironment. Changes such as increased fibronectin secretion, collagen type I, II and IV, indicate that tumor progression requires continuous interaction between the ECM and tumor cells. Increased deposition of matrix proteins promotes tumor progression as it interferes with cell–cell adhesion, cell polarity, and the amplification of growth factor signaling. High collagen deposition has been shown to result in tumor progression through increased integrin signaling [6].

The increased stiffness of the matrix activates integrins as well as cytoskeletal tension, promoting cell adhesion and motility [2–4, 6, 11]. It has been observed that local cell invasion of tumors is directed along collagen fibers aligned, suggesting that this linearization of the fibers facilitates invasion. These densely aligned collagen fibers are believed to act as cues for proliferating neoplastic cells to migrate outward from the tumor [2–4, 6, 11].

Tumor tissue hydration also has some impact on ECM dysregulation. Since it is strongly influenced by specific tissue GAGs due to their anionic structure and their ability to attract water, it is known that as hydration increases, the increase in intra-tumor hydrostatic pressure also increases, altering the biomechanical properties of the tissue that are known to be crucial for invasion [2–4, 6, 11].

Elevated levels of hyaluronic acid (carboxylated and free glycosaminoglycan) in ECM correlate with an increased likelihood of malignancy and poor prognosis. As a ubiquitous linear polysaccharide, hyaluronic acid (HA) is fundamental in determining the compressive properties of most biological tissues. Combining tensile strength due to collagen conformation and compressive strength due to hyaluronic acid creates the ideal biophysical properties for tissue homeostasis. Furthermore, hyaluronic acid is an induction signal for epithelial-mesenchymal transition (EM) and a migration substrate. It is also important in regulating vascular endothelial barrier permeability by stabilizing cell–cell junctions [2–4, 6, 11, 18].

Although increased levels of collagen directly promote ECM stiffness and mechanistically drive cell motility and proliferation, the exact role of hyaluronic acid in cancer metastasis remains unclear. However, its downregulation may serve as a key biomarker for invasion and metastization [2–4, 6, 11, 18].

5. Human papillomavirus (HPV)

5.1 The virus

The human papillomavirus (Human papillomavirus-HPV) is a small (55 nm) non-enveloped virus that belongs to the Papillomaviridae family. It can be classified according to its tropism (cutaneous, mucocutaneous, and mucosal). Since the availability of cellular factors expressed in different layers of the epithelium plays a role in viral gene expression and genome amplification, the viral cycle is strictly dependent on the epithelial differentiation program HPV infections are associated with some hyperproliferative pathologies of epithelia and mucosa, and most cervical cancer and warts cases. It has been described more than 200 types of HPV. Almost 40 types exhibit a particular tropism for the anogenital region's cellular floor epithelium and mucous membranes. HPV types in this subgroup are classified as being of high or low oncogenic risk, depending on the clinical lesions they cause.

The high-risk HPV types are associated with almost all cases of cervical cancer, and the low-risk HPV types are the cause of nearly all anogenital warts and low-grade lesions with a slight tendency for malignant progression. The most prevalent high-risk HPVs are HPV16 and HPV18, while the most common low-risk types are HPV6 and HPV11. Infections with specific high-risk HPV types are etiologically related to a significant proportion of vulvar, vaginal, anal, penile, and head and neck carcinomas [19–22].

5.2 Structure of HPV

HPV is a double-stranded circular DNA virus with 6.8–8.4 kb and can be divided into three functional regions:

- Early region (E-early), consisting of the early genes (E1-E8) that encode the replication proteins and regulate the different phases of the viral cycle.

- Late region (L-late) consisting of the L1 and L2 genes that encode the capsid proteins. L1 is called the major capsid protein and is responsible not only for the specific adhesion of the virus to the cell but also for the immune response produced by the host against this virus. L2, the minor capsid protein, appears to be important in the encapsidation of viral DNA [23].

- The long control region (LCR- long control region), consisting of regulatory genes, contains the origin of replication and the E6 and E7 genes that control transcription and regulate the expression of different HPV genes [19, 24].

The two oncogenes, E6 and E7, play a major role in carcinogenesis, contribute to immortalizing normal human keratinocytes in cell culture systems, and are essential to maintain the transformed phenotype in vivo. The main role of these proteins during the HPV cycle is to generate a permissive cellular microenvironment for viral replication. This includes the induction of DNA replication machinery, immune evasion, and the downregulation of apoptosis. To achieve this, the E6 and E7 proteins give rise to critical cellular regulatory pathways, including those dominated by p53 and pRb. The E6 protein (16–19 kDa) associates with p53 (tumor suppressor protein) and promotes its degradation. E7 (10–14 kDa) inactivates the function of another tumor suppressor protein, the retinoblastoma protein (pRB). Together, these two proteins promote the mechanisms involved in the genesis of tumors caused by these viruses, favoring cell transformation and immortalization [19, 24].

5.3 Viral replication

As already described, the life cycle of this virus is synchronized with cell differentiation and division.

Whether or not the viral life cycle is complete depends on the nature of the epithelial site where infection occurs and external factors such as hormones and cytokines. It is suggested that infection requires access to the viral particles (composed of viral DNA and the capsid proteins, L1 and L2) to the basal lamina and their interaction with HSPGs laminin [23]. Structural changes in the virion capsid facilitate transfer to a secondary receptor in the basal keratinocyte, necessary for virus internalization and subsequent transfer of the viral genome to the nucleus. Once internalized, virions undergo endosomal transport and pass into the nucleus, where the capsid disassembles and DNA release occurs. The L2-DNA protein

complex ensures the correct nuclear entry of the viral genomes, while the L1 protein is retained in the endosome and ultimately subject to lysosomal degradation [20, 23, 24].

Infection is thought to require an epithelial wound or micro-wound to allow the virus access to the basal lamina. Indeed, active cell division, as occurs during wound healing, is required to enter the virus genome into the nucleus. It has been proposed that lesion formation requires the initial infection of a mitotically active cell [19, 20].

It is also known that in the basal cells of the epithelium, the expression of viral genes is repressed and expression of early (E) and late (L) genes only happens at the level of keratinocytes or the upper mucosal layers. Viral replication is associated with excessive cell proliferation of all epidermal layers except the basal layer. Since the basal cells of the stratified sidewalk epithelium are the only ones capable of dividing, they are the initial target of HPV infection [19, 20, 23].

Regardless of the nature of the infected basal cell, infection is followed by an initial phase of genome amplification and then the maintenance of the viral episode at a low copy number. The viral replication proteins E1 and E2 are considered essential for this initial amplification phase. The precise role of the HPV E6 and E7 proteins in infected basal cells is uncertain, particularly for low-risk HPV types that are not generally associated with neoplasia and whose viral DNA does not integrate into the chromosomes. They are thought to produce lesions following infection of a basal stem cell at the site of a wound. The role of the curative response in driving the initial proliferation of the infected cell may well be critical, with local microenvironment signaling influencing viral gene expression and/or protein functions. In the case of high-risk types that cause neoplasia, there is the integration of the viral genome into the chromosomal DNA, and the role of the viral E6 and E7 proteins in cell proliferation in the basal and parabasal cell layers is quite clear, especially at cervical sites where neoplasia can occur. It is clear that many functional differences exist between high- and low-risk E6 and E7 proteins and that these contribute, along with differences in promoter activity and gene expression patterns, to the different HPV-associated pathologies seen in vivo. Indeed, recent studies have suggested that a critical event in determining the neoplastic grade is the down-regulation of E6/E7 expression, even in the absence of genome integration, which is classified according to the extent to which basal-like cells extend into suprabasal epithelial layers. While such functional differences undoubtedly contribute to the respective abilities of high and low-risk HPV types to cause neoplasia and cancer, it is important to remember that a key function of the E6 and E7 proteins in most HPV types is not to promote basal cell proliferation but rather to stimulate cell cycle re-entry into the middle epithelial layers to allow genome amplification [19, 23, 24].

6. HPV-MEC interaction

As previously mentioned, the role of HSPGs in HPV infection is quite relevant [15, 25].

The HSPGs typically consist of a core protein and GAG chains. The core protein of syndecans is composed of an extracellular domain, a unique transmembrane domain, and a short cytoplasmic domain that interacts with the cytoskeleton. The glypicans are GPI-labeled HSPGs. The GAG chain comprises unbranched anionic polysaccharides composed of repeated disaccharide units formed by sulfated uronic acid and hexosamine residues [15–18].

As components of the ECM, HSPGs contribute to the organization of the basement membrane and mediate cell adhesion and motility. HSPGs bind to

cytokines, chemokines, and growth factors on the cell surface, preventing their degradation, creating temporary storage sites or morphogen gradients important in development. Still, on the cell surface, they also serve as endocytosis receptors, and regulate the lysosomal degradation of extracellular molecules and provide nutrients to the cells. In addition, they are involved in the endocytosis of cell receptors. They mediate the transcellular transport of chemokines through endothelial cells. They also serve as co-receptors for a fibroblast growth factor (FGF) and its receptor. They mediate intracellular signaling or intracellular stress through the proteolytic shedding of syndecans and play an important role in developing and maintaining stem cell niches [15, 18].

The strategic localization of HSPGs in tissues is critical to their functional role. The localization of SDCs and GPCs at the plasma membrane regulates intracellular and cell-to-ECM signaling. The localization of HSPGs in the basement membrane regulates their barrier functions and coordinates cell–cell and cell-MEC interactions [15, 18].

Degradation of heparan sulfate chains by heparanase produces heparin-like fragments that activate FGF-2 mitogenicity. Therefore, the biological role of an HSPG depends on the properties of its core protein, the number of GAG chains attached, its location in cells and tissues, and the biosynthetic modifications that its GAG chains receive in situ. A wide range of biological functions is attributed to GAGs in cancer metastasis and other biological events due to their controlled, highly heterogeneous, and complex structure that allows for the regulation of tissue-specific functions [15, 18].

6.1 HSPGs as viral receptors

HSPGs are receptors hijacked by numerous viruses to bind to host cells. This typically occurs through electrostatic interactions between the negative charges of HSPGs and the basic amino acid portions of viral surface proteins. A consistent amount of data supports the natural dependence of HSV, DENV, and HPV on HSPGs for their binding to host cells [15, 26].

HSPGs, due to highly sulfated GAG chains, exhibit an overall negative charge that can interact electrostatically with the basic residues of viral surface glycoproteins or viral capsid proteins of non-enveloped viruses. Viruses exploit these weak interactions to increase their concentration on the cell surface and increase their chances of binding to a more specific entry receptor. HSPGs directly serve as entry receptors in rare cases, as described for herpes simplex virus (HSV)-1. Another study showed that HSPGs are crucial for SARS-CoV entry. Prophylactic treatment with bacterial heparinase I or heparin showed a reduction in SARS-CoV infectivity. Thus, either loss of HS or competitive inhibition confers some protection to cells against SARS-CoV. Given the structural similarities between SARS-CoV and the novel SARS-CoV-2, it would be interesting to study the effects of removing HS on SARS-CoV-2 infections. There are multiple ways to reduce viral contact with cell surface HS (HPSE, heparinase, heparin, soluble HS, and MMPs), investigating that this binding may give insight into SARS-CoV-2 entry and possible therapeutics [15, 26, 27].

All papillomaviruses are believed to rely on HSPGs for their initial binding. However, HPV-16 is the serotype whose pathogenesis is most studied due to its oncogenic potential and prevalence [15].

As already mentioned, the HPV infection cycle starts from the basal membrane of the vaginal mucosa, exploring abrasions or lesions in the epithelium (**Figure 2A** and **B**). The entry of HPV particles into host cells is a multistep process that begins with binding to HSPGs expressed on the cell surface of basal

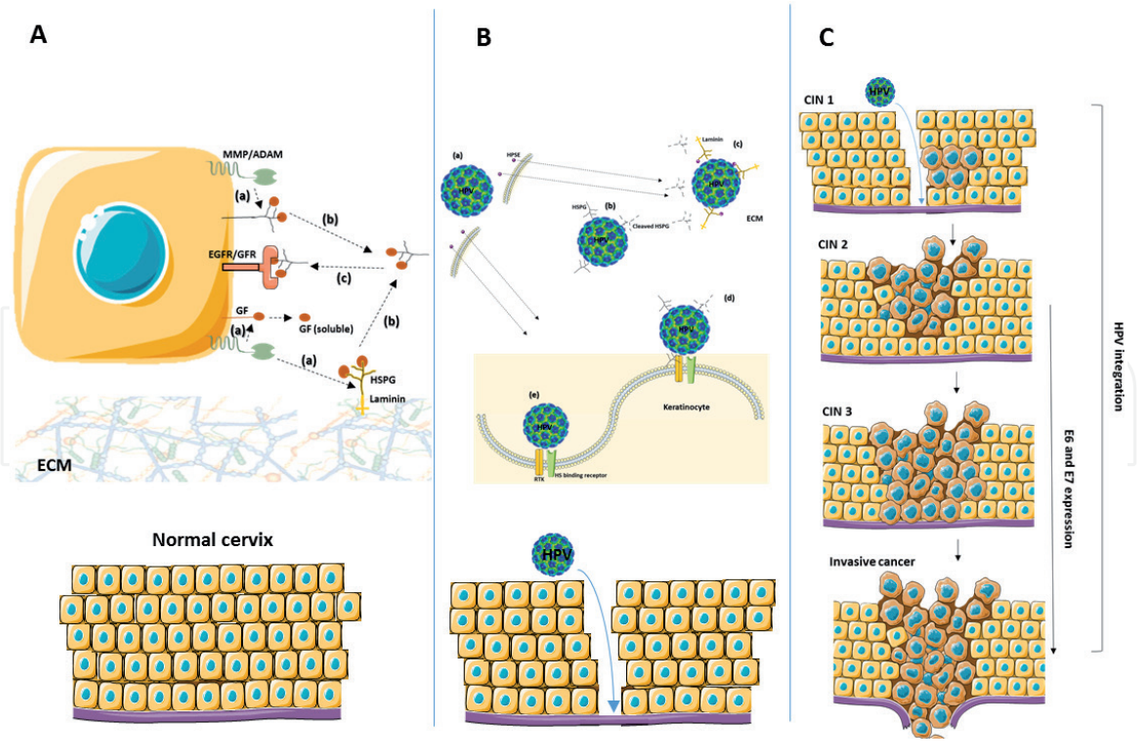

Figure 2.
Schema of the role of HSPGs in HPV16 infection. (A) Bottom is a normal cervix epithelia disrupted. Upper figure represents the details of wounded tissue healing mechanism, (a) MMP/ADAM cleaved GF (growth factor) bound to HSPG of cell membrane and also to the complex with laminin (b) shedd HSPG/GF complex, bounds to (c) adjacent cells GFR/EGFR as a co-receptor. (B) In bottom of figure HPV16 infected wound tissue release, as shown in upper figure, HPSE (heparanase) and HPV16 (a), HPV16 is coated with HSPGs and HS (heparansulfate) released respectively by MMPs/ADAM and HPSE, (b), HPV16 bounds to the complex HSPG-laminin in ECM, (c) adjacent keratinocytes through membrane EGFR-RTK and HS receptor bound as a complex (d) is endocytosed by the cell (e). (C) HPV16 integration by the infected keratinocytes during wound healing, after endocytosed in (B), progresses through CIN I, CIN II and CINIII to invasive cancer, after expression of E6 and E7 HPV oncogenic proteins.

keratinocytes or in the ECM. The interaction of the HPV16 L1 capsid protein with the HS chains of proteoglycans is well known but generally considered to have a passive role in infection. Binding to HSPGs induces conformational changes in the capsule and facilitates proteolytic cleavage of L1. This cleavage allows interactions between capsid and cyclophilin B, which results in further conformational changes that expose the N-terminus of L2. The exposed N-terminus contains a conserved consensus cleavage site for the extracellular proprotein furin convertase. This interaction is essential for successful HPV infection since cleavage of furin results in the exposure of a binding site on L1 postulated to be recognized by an unknown receptor. The described changes in virion conformation further facilitate the reduction in binding affinity to HSPGs, thus facilitating binding to the unknown receptor(s). These findings underscore the role of initial HSPG attachment to facilitate the critical step of L2 cleavage by furin and association with the putative second receptor for entry. Cleavage of furin has also been implicated in successful endosomal escape before transporting the viral L2/ADN complex into the nucleus, emphasizing the necessity of furin cleavage for successful HPV infection [15, 20].

Syndecan-1, the most abundant HSPG in keratinocytes, plays an important role in this initial binding due to its expression in epithelial cells and its overproduction during wound healing. It has been shown that syndecan-1 when released plays a major role in the infection of keratinocytes by HPV. Instead of separating it from its HS chains, HPV particles are released from the cell surface through the normal process of HSPGs, remaining bound to HS chains (heparan sulfate) and growth

factors. They can bind to the epidermal growth factor receptor (EGFR). The specificity of growth factors is the bridge for the interaction of the virus with cellular receptors, for example, tyrosine kinase, whose signaling promotes infection **Figure 2A** and **B** [15, 28].

Considering the biology of HSPGs, it stands to reason that HPV particles bound to HSPGs would associate with ECM, and indeed, free syndecan forms bind tightly to ECM via their HS chains. Many studies have shown that HPV particles accumulate in the ECM, and LN-332, a component of the ECM, is a proposed attachment factor for HPV. There is already evidence of direct protein–protein interaction between HPV16 and LN-332 **Figure 2A** and **B** [15, 25, 28].

Thus, HS chain cleavage enzymes are known to increase the release and infectivity of HPV particles bound to ECM, indicating that many virus particles bind to ECM via these HS chains. Free syndecan-1 interacts with HPV16 and LN-332, demonstrating it to be an ECM-binding factor for HPV16, in addition to its role in binding HPVs to the plasma membrane. The interaction of snd-1 with LN-332 is expected based on reports demonstrating the concentration in the EMC of free snd-1 with its native HS chains and that LN-332 binds plasma membrane resident snd-1 with high affinity and specificity through HS chains. LN-332 has been identified as a binding factor between ECM and HPV. Since LN- 332 intrinsically lacks HS chains but contains HS-binding domains, it seems more likely that HS chains bridge the gap between HPV and LN-332, and this could account for the co-localization in the ECM of HPV and LN-332 **Figure 2A** and **B** [15, 25, 28].

Models have been developed for explaining the mechanism of HPV release from HSPGs:

1. It is suggested that conformational changes in the structure of the virus capsule are caused by the binding of HSPGs, which then allows cleavage of the HPV L2 protein, and subsequently triggers HPV release from binding factors that passively accumulate virus particles on their HS chains.

2. Based on the physiological processing of HSPGs molecules, HPV particles are released from the cell surface still in complex with HS and growth factors (GFs) and signal through GFRs to promote infection [28].

None of these models addressed ECM-bound virus release and infectivity. Still, a recent study suggested another model: high-speed processing of normal HSPGs from HPVs to gain infectious entry into keratinocytes. Inhibited viral release from ECM, cellular access, and infectiousness from ECM can be easily explained by this model. Proteases and heparanase play an essential role in HPV release from primary receptors [28].

This model, in which HPV usurps the processing of HSPGs and GFR/RTK signaling to promote infection, reflects the role of epithelial injury in mediating papillomavirus infections in vivo. Consequential breaks in epithelial damage result in an influx of HRs and cytokines involved in syndecan dissemination. Snd-1 expression is enormously increased in keratinocyte migration and proliferation, and free syndecans present in wound fluids regulate the activity of GFs and MMPs. Thus, HPVs appear to have evolved to control the epithelial wound to gain access to mitotically active basal cells and take advantage of the factors and architecture that favor infection. Many intracellular pathogens of the female genital tract (HIV, herpesviruses, chlamydia, Neisseria) interact with cellular HSPGs. Thus, it is tempting to infer that these pathogens also appropriate the biology of HSPGs during infection. In summary, there is new knowledge about the transmission of oncogenic HPVs, and high-speed pathogens usually function during infection of their hosts.

These findings may point to additional targets for preventing HPV infections and potentially those of similarly acting pathogens **Figure 2A** and **B** [28].

Upon contact with HSPGs, the HPV capsid undergoes conformational changes assisted by extracellular cyclophilin B and cleavage of the capsid protein L2 by furin. This leads to a loss of affinity for HSPGs and binding to different secondary receptors. Identification of the internalization receptor is ongoing, but α6 integrins, EGFR, and the tetraspanin family may be involved. The entry kinetics of HPV appears to be asynchronous and slower than for most other viruses, but the cause is not yet fully known. Some research suggests that it may be linked to the cell cycle phase or the involvement of multiple receptors. Subsequently, the virus is internalized through endocytosis, but there are conflicting reports on different HPV types and cells **Figure 2A** and **B** [15, 20].

The main goal in developing microbicides against HPV (or any viral infection) is to block the interaction between the virion proteins and the cell surface receptors used by the virus to gain entry into cells. As discussed earlier, the initial binding of HSPG is an important step for successful HPV internalization, as its inhibition has been shown to decrease HPV infection in vitro and in vivo. Since several different viruses use HS chains as the initial receptor/corrector to bind to the cell surface, it is considered a viable drug target, particularly about producing a microbicide with broad-spectrum protection against a range of HPVs (as well as other sexually transmitted viruses). Efforts in this direction will aid the development of antiviral drugs that are effective not only on many existing viruses but also on unpredictable emerging viruses [15, 20].

6.2 Pathogenesis and immunity

Once the basal layer of keratinocytes is reached, the virus can remain latent or take advantage of cell differentiation to replicate and initiate the disease. As for the host immune response, it is known that it can eliminate the infection or silence it (latent). The virus can persist with low infectivity, survive a weak immune response (persistence) and later induce pathology [19, 22, 24].

The mark of HPV infections is the effective evasion of innate immune recognition. The viral productive life cycle is exclusively intraepithelial, there is no viremia, no virus-induced cytolysis or cell death, and viral replication and release are not associated with inflammation. HPV globally decreases innate immune signaling pathways in infected keratinocytes, and pro-inflammatory cytokines are not released, activation signals to Langerhans cells, cell migration from and recruitment of stromal dendritic cells (DCs) and macrophages are absent or inadequate [19, 24].

Despite the high impact of HPV protein expression on cellular homeostasis, these viruses are incomplete carcinogens. Therefore, further changes in the cell and its microenvironment are required for tumor establishment and progression. This process includes dysregulation of the ECM. In some cases, changes in the levels and activity of defined ECM components have been experimentally associated with the expression of HPV-specific proteins, suggesting the direct involvement of the virus in the downregulation of these factors. Other studies, mainly those performed with clinical specimens, have identified changes in the levels of ECM molecules during the progression of HPV-related diseases [21, 29].

7. ECM alterations in HPV-associated diseases

The natural history of cervical cancer development begins with a precursor lesion called cervical intraepithelial neoplasia (CIN). CIN 1 CIN 2 lesions are

classified as productive lesions, in which the viral cycle is complete. On the other hand, CIN 2 lesions and CIN 3 lesions are potential precursors of cervical cancer. The development of these lesions is mainly caused by persistent infection with oncogenic types of HPV. Intraepithelial lesions show low to moderate histological changes and may regress spontaneously within 1 to 2 years. In persistent high-risk HPV infections, high-grade precancerous lesions (CIN 2 and CIN 3) may develop within 3 to 5 years. Morphologically, CIN 3 (carcinoma in situ-CIS) represents a heterogeneous disease and can be considered a precancerous lesion of a more advanced cervical cancer **Figure 2C** [21, 30].

HPV infection has been associated with several changes in tissue organization and architecture, including downregulation of the expression and activity of MMPs. It has been shown that up-regulation of MMP-2 and MMP-9 expression and activity are associated with high-grade CIN, and their respective inhibitors have reduced expression levels in these lesions [4, 21]. On the analysis of MMP-11 and MMP-12 expression in high- and low-grade lesions, it was shown that both might be associated with the appearance of cancer precursor lesions and suggested that increased expression of these proteins may be considered an early event during the development of preneoplastic cervical lesions **Figure 2C** [21].

Alterations in other components of the ECM have also been explored in the context of cervical tissue transformation. Expression of the 67-kDa laminin receptor (LR67) was found to progressively increase with CIN grade. LR67 is associated with CIN 2 to 3 and can be considered a marker of cell proliferation in cervical tissue. These authors also demonstrated that combined analysis of LR67 and vascular endothelial growth factor-C (VEGF-C) could improve the clinical detection of high-grade CIN **Figure 2C** [21].

HPV infection has also been linked to a percentage of lesions in other epithelia of the anogenital tract, including the vulva, vagina, and anus. Vulvar, vaginal, and anal cancer precursor lesions are called vulvar intraepithelial neoplasia (VIN), vaginal intraepithelial neoplasia, and anal intraepithelial neoplasia (AIN), respectively. As with cervical precursor lesions, VIN and AIN also progress through degrees of epithelial transformation. The analysis of the expression of MMP-2, MMP-9, TIMP-1, and TIMP-2 by IHC in samples of VIN 1, 2, and 3 and invasive vulvar carcinoma suggested that overexpression of MMP-2, MMP-9, and TIMP-2 proteins may be related to the progression of VIN to invasive carcinoma **Figure 2C** [21].

8. ECM composition in HPV-associated cancers

The loss of regulation of ECM remodeling by unbalanced proteolysis plays a significant role in the loss of tissue homeostasis and pathological processes such as cancer. In cancer, this event may impact tissue tension and release chemotactic fragments of ECM components that influence the local microenvironment. It promotes cell migration and recruitment of stromal, endothelial, and immune cells to the tumor vicinity. The heterogeneous association of cancer cells and other elements observed in the tumor microenvironment, such as inflammatory infiltrate, endothelial cells, and tumor-associated fibroblasts, should also be explored to understand ECM remodeling changes fully [4, 21].

The most investigated proteases present in this process are the MMPs, as described previously. These MMPs play a central role in basement membrane breakdown and cell invasion and neoangiogenesis, and metastasis. Excess MMP activity generates topographical changes in the tumor microenvironment through

modifications associated with proteolysis of the structural skeleton of the ECM. Indeed, linearization and thickening and/or degradation of specific collagen are common events observed in areas of epithelial tissue adjacent to tumor-associated blood vessels where cancer cells invade. The activity of MMPs also regulates cell migration and release of ECM fragments with biological functions, such as growth factors. The crucial role of specific MMPs in the process of carcinogenesis has set an objective task for researchers to explore the potential of MMPs as therapeutic targets. However, the use of broad-spectrum MMPIs offered no clinical advantages due to dose-limiting side effects [4, 21, 31].

Several authors have studied ECM alterations in both structural and remodeling molecules in invasive cervical cancer. More specifically, changes in the expression/activity of galectins, collagens, proteoglycans, laminins, fibronectins, integrins, proteases, and regulators have been observed in cervical cancer samples derived cell lines [21].

The claudins (CLDNs) and occludins are families of proteins associated with tight junction establishment, epithelial cell polarity, and intercellular permeability. The expression levels of CLDN type 1, 2, 4, and 7 proteins are increased in HSIL lesions and invasive cervical tumors when compared with normal cervical tissues. On the other hand, occludin is expressed in the basal cell layer of normal cervical tissues. Its protein level is reduced in invasive cervical carcinomas compared with CIN samples. Thus, changes in cell adhesion and ECM structure are a common early feature in cervical cancer progression [21].

Expression of the high-affinity laminin-binding protein 67-kd (67LR) has also been shown to increase both CIN and cervical cancer samples compared to normal cervical tissues.

Versican, an ECM proteoglycan, was evaluated in cervical cancer samples by IHC (Immunohistochemistry) and situ hybridization (ISH). Expression of high levels of versican in tumor stromal myofibroblasts was associated with a lower frequency of CD8-positive T cells, more significant invasion and depth of tumor parametrial infiltration, and no change in cervical cancer survival. Interestingly, the beta-galactoside- galectin-1 binding protein was more expressed in stromal cells adjacent to cancer when compared to normal stroma associated with the cervical tissue. Furthermore, a higher expression level of galectin-1 was associated with increased local tumor recurrence and poor cancer-specific survival in patients with stage I-II cervical tumors after radiation treatment. However, it could not predict distant metastasis [21].

Laminin-1 and smooth muscle actin proteins (SMA) showed increased expression, mainly in the surrounding cervical tumor stroma compared to the normal cervical stroma. In addition, tumor cells especially expressed laminin integrin a6b4 receptors, and tumor-associated fibroblasts showed higher levels of laminin-a1 and laminin-b1 and lower levels of laminin-5, fibronectin, collagen III, TIMP-1, and the hyaluronic acid receptor CD44 when compared to normal fibroblasts. Finally, MMP-7 and MMP-9 expression has been shown to correlate with CD44 expression in skin cancer cells [21].

The data discussed above show that ECM composition and function alterations are common in HPV-associated lesions and cancers. Taken together, these changes highlight the complex molecular pathways that lead from initial infection to disease. For example, analysis of the impact of HPV on components of the MMP family has produced a spectrum of data that could be used for disease diagnosis and identification of targets for therapy. The data summarized here also show that, concerning the mechanisms by which HPV modulates MMP expression and activity, there is still much to learn. Finally, alterations in the ECM may impact the microenvironment of HPV-infected tissue, affecting the

recruitment of inflammatory infiltrates, altering the fate of different cell populations present in the tumor, and ultimately determining disease progression and prognosis [4, 21].

More studies are needed to understand how HPV proteins affect the dynamic balance of the ECM in associated pathologies. This will help us to understand the disease genesis and define more appropriate clinical interventions. Importantly, the vast majority of the ECM changes described have also been observed in HPV-independent tumors. Therefore, understanding the virus-mediated molecular events that lead to ECM disruption may be useful for understanding the basic mechanisms of carcinogenesis and developing more general antitumor approaches [21, 32].

8.1 Heparanase and heparan sulfate/syndecan-1 axis

As mentioned earlier, syndecans are a family of four HSPGs that can be soluble or membrane-anchored. Syndecan 1 (SDC-1) is the one that has been most studied and is found mainly on the surface of epithelial cells. Loss of syndecan-1 and E-cadherin from the cell surface is known to be a critical step in the transition to epithelial neoplasia [18, 33].

The heparanase/SDC-1 axis is a key point regulating cell signaling when tumor cells are present and in their respective microenvironment. This heparanase/SDC-1 axis modulates cell proliferation as it regulates hepatocyte growth factor (HGF) and vascular endothelial growth factor (VEGF). HGF is a cytokine that increases the growth, motility, angiogenesis of tumor cells. Free syndecan-1 binds to secreted HGF and ultimately facilitates a paracrine and autocrine signaling cascade through the cellular receptor c-Met. It binds to syndecan-1 in the ECM, stimulating angiogenesis and tumor invasion via the Erk pathway secreted VEGF. To regulate gene expression, heparanase and syndecan-1 can also be transported to the nucleus to regulate gene expression. Overall, nuclear HS chains and syndecan-1 are considered anti-proliferative and decrease gene transcription. Specifically, highly sulfated HS chains are mostly inhibitory, contrasting with free syndecan-1 that promotes angiogenesis, proliferation, and cell invasion. Conversely, heparanase present in the nucleus increases gene expression and promotes growth. Thus, syndecan-1 expression is considered a prognostic tool in solid and hematologic malignancies. A high level of stromal expression of syndecan-1 is a negative prognostic factor, and low levels of epithelial expression are indicators of advanced disease and poor prognosis. Loss of syndecan-1 is believed to represent cancer cells with high malignant and metastatic potential [4, 17, 33].

9. Heparanase

Heparanase is an endo-β-D-glucuronidase that cleaves the side chains of HS. This results in the release of bioactive HS fragments from the ECM and in structural changes. Over the past two decades, much work has been dedicated to studying the role of heparanase in cancer biology. Various analysis methods have revealed that heparanase expression is increased in numerous cancers, including hematological malignancies, carcinomas, and sarcomas. In addition, elevated heparanase levels are associated with reduced postoperative survival, increased angiogenesis, and metastasis. All of these factors have triggered the development of heparanase inhibitors as novel anti-cancer agents [4, 17, 33].

Mammalian cells express a single functional heparanase enzyme, heparanase-1. Heparanase-2, a homolog of heparanase, has been cloned but cannot perform HS

degrading activity. It can, however, regulate the activity of heparanase-1. The heparanase structure contains a TIM barrel fold, which incorporates the enzyme's active site and a distinct C-terminus domain with non-catalytic properties and is involved in the non-enzymatic signaling and secretion function of heparanase [17, 33].

The expression of heparanase is under tight regulation. In non-cancer cells, the heparanase promoter is constitutively inhibited, secondary to promoter methylation and p53 activity, which suppresses heparanase gene transcription by binding directly to its promoter. In addition, further regulation occurs during post-translational processing. Cathepsin L is required for post-translational activation of heparanase, and cathepsin L inhibitors prevent the formation of active heparanase. In non-pathological states, heparanase expression is restricted primarily to platelets, activated white blood cells, and the placenta with little or no expression in normal connective tissue or epithelium. In addition, it is most active under acidic conditions (pH 5–6), during inflammation, or within the tumor microenvironment [17, 33].

9.1 Role of heparanase in HPV viral pathogenesis

As already mentioned, increased expression of heparanase in numerous malignancies is associated with poor prognosis. And the direct role of this enzyme in neoplasms was confirmed when inhibition/silencing of heparanase in cancer cell lines resulted in a significant reduction in the invasive phenotype of the cells [4, 33].

The primary enzymatic activity of heparanase in the cleavage of the HS side chains of HSPGs and consequent release of growth factors and cytokines give rise to cell signaling pathways capable of inducing ECM remodeling. Heparanase can also release proangiogenic growth factors bound to heparan sulfates such as bFGF, HGF, PDGF, and VEGF, from the extracellular matrix to promote endothelial cell migration and proliferation indirectly. Tumors with high levels of heparanase have significantly higher microvascular density than tumors with low levels of this enzyme [4, 33].

Several heparanase capabilities have been demonstrated in the progression of cancers. These include increased cell proliferation via insulin, increased resistance to chemotherapy, expression of mesenchymal markers, increased autophagy, increased cell adhesion, and even a procoagulant function [26, 33].

Focusing on human papillomavirus infection, we remain to understand what role heparanase plays in HPV viral pathogenesis. Recent studies show that HPV16 particles bind to the ECM through HS chains. Reducing the activity of matrix metalloproteinases and heparanase drastically reduces the release of viruses from the ECM, which results in the loss of viral uptake and infection of human keratinocytes. On the other hand, exogenous heparanase promotes viral release and disease. This phenomenon may be necessary for explaining, especially at the wound site, the host healing response and RTK/GFR signaling that increases HS release, allowing an ideal environment for HPV to infect keratinocytes **Figure 2B** [3, 34].

Other research work has shown the significance of the HPV E6 gene in HPV-heparanase interaction in head and neck squamous cell carcinoma (HNSCC). The HPV E6 gene interacts with p53, decreasing its activity, leading to increased expression of heparanase since p53 is a potent inhibitor of transcription of this enzyme; expression of p21, a downstream component of the p53 pathway, correlates positively with heparanase expression on tissue section staining, confirming the heparanase-p53 signaling event. Polysaccharide segments of the HS chains serve as attachment sites for many growth factors, cytokines, chemokines, and various bioactive ligands. Cleavage of the HS chains by heparanase releases these bioactive factors increasing tumor invasion and malignancy in the case of HPV-induced HNSCC16 [3, 34].

10. The potential of future targeted therapies

Current statistics on the prevalence of HPV and cervical cancer alone underscore the need for alternative therapies to prevent and treat HPV infections. Existing HPV vaccines, while highly effective, do not offer protection against all high-risk HPV types, let alone low-risk HPV types. Furthermore, these vaccines are purely prophylactic and are not easily accessible to women in low- and middle-income countries (LMIC). For these reasons, alternative therapies that have broad-spectrum protection against HPV types, as well as other sexually transmitted infections, are worth exploring.

Since heparanase is an influential enzyme in tumor progression, it is the ideal therapeutic target. And since it is typically not expressed in healthy tissue, the side effects of its inhibition would be minimal. A series of heparanase inhibitors have been studied and produced, namely heparin, logically because it is a molecule close to HS. However, it is limited as an anti-cancer therapy because of its anticoagulant effects. Similarly, when LMWH (Low molecular weight heparins) was tried as an alternative for the same effect, but the results were controversial [4, 17, 33].

In addition to heparins, strategies have been developed to inhibit heparanase, such as HS mimetic molecules, modified heparins, etc. HS mimicking molecules have lower anticoagulant activity and greater selectivity for heparanase than heparin, allowing a wider therapeutic window. Some inhibitors already investigated are PI-88 (Mupaphosphat), PG545, SST0001 (Roneparstat), M402 (Necumparanib). In addition to HS mimetics, the non-steroidal anti-inflammatory aspirin, widely suggested to have a long-term anti-cancer effect, binds directly to the active site of HPSEs, inhibiting their enzymatic activity and preventing HPSE-dependent cancer cell migration, metastasis, and angiogenesis both in vitro and in vivo. As HPSE inhibition may seem attractive for cancer mitigation, it is important to note the critical role of the enzyme in the infiltration of activated NK cells in primary tumors and metastasis sites. Consequently, potential inhibitors must be highly selective and thoroughly investigated to limit adverse effects [4, 17, 33].

As we see, several heparanase inhibitors have entered clinical trials for various cancers, but none yet for viral diseases. Early results already suggest that using heparanase as a target may have rewarding benefits in controlling many viral infections, and thus the inhibitors listed above may have beneficial effects. Now, the promise that inhibiting this enzyme or the upstream effector, p65, could provide a novel therapeutic intervention to treat the disease. Overall, emerging knowledge about heparanase as an essential regulator of viral infections and associated morbidities could one day make a broad-spectrum antiviral drug a real possibility [34].

11. Conclusion

In this paper, we have discussed the extracellular matrix's very complex and important role in developing and progressing cancers and, more specifically, in human papillomavirus infection. Indeed, in recent years the ECM has been increasingly considered a crucial component in physiological processes such as cell proliferation, adhesion, and migration. The perspective of cancer and its progression has also changed. We no longer consider a disease caused only by dysregulated cell proliferation. Still, we give importance to the microenvironment and its changes and adaptations to cellular stress. In-depth knowledge of the ECM components, their complex interactions, and their constant and dynamic remodeling during all stages of tumor development brings some hope in developing promising targeted

therapies to combat these pathologies. Likewise, knowledge of the components of viral particles and host cell entry factors and their specific interactions may allow the design of efficient antiviral strategies [15, 22].

Indeed, the role of HSPGs in HPV interaction with the ECM is undisputed and developed throughout the work. The enzyme heparanase, which we know has a significant impact on tumor progression and metastasis. Thus, as the influence of the tumor microenvironment on cancer progression becomes more evident, the focus on inhibiting enzymes that degrade HSPGs highlights an approach to maintain normal tissue architecture, inhibit tumor progression, and block metastasis. This review addresses the role of these enzymes, namely heparanase, in the context of the tumor microenvironment and their promise as a therapeutic target for cancer treatment, particularly cervical cancer [15, 17].

Acknowledgements

The Authors would like to acknowledge the Instituto de Investigação Científica Bento da Rocha Cabral for support.

Conflict of interest

The authors declare that they have no competing interests.

Acronyms and abbreviations

CD44	Cluster of differentiation 44
CIN	cervical intraepithelial neoplasia
CLDN	claudins
ECM	extracellular matrix
EGFR/RTK	epidermal growth factor receptor/ tyrosine kinase receptor
FGF	fibroblast growth factor
GF	growth factors
GFRs	growth factor receptors
GPI	glycosylphosphatidylinositol
HGF	hepatocyte growth factor
HNSCC	head and neck squamous cell carcinoma
HPSE	heparanase
HPV	human papillomaviruses
HS	heparan sulfate
HSIL	high-grade intraepithelial lesion
HSPGs	heparan sulfate proteoglycans
LN-332	laminin 332
LRR	leucine-rich repeat
LSIL	low-grade intraepithelial lesion
MMPIs	matrix metalloproteinases
PDGF	platelet-derived growth factor
SND-1	syndecan 1
TIMPs	an inhibitor of type 2 metalloproteinases
VEGF	vascular endothelial growth factor

Author details

Joana Sampaio[1], Joana Ferreira[1,3,5], Ana Carolina Santos[1,3,5], Manuel Bicho[1,3,5] and Maria Clara Bicho[1,2,3,4]*

1 Faculty of Medicine, University of Lisbon, Lisboa, Portugal

2 Institute of Preventive Medicine and Public Health and ISAMB – Environmmental Health Institute, Faculty of Medicine, University of Lisbon, Lisboa, Portugal

3 Genetics Laboratory and ISAMB – Environmmental Health Institute – Ecogenetics and Human Health Unit, Faculty of Medicine, University of Lisbon, Lisboa, Portugal

4 Dermatology Research Unit, Instituto de Medicina Molecular, University of Lisbon, Lisboa, Portugal

5 Instituto de Investigação Científica Bento da Rocha Cabral, Lisboa, Portugal

*Address all correspondence to: mcbicho@medicina.ulisboa.pt

References

[1] Reiland J, Sanderson RD, Waguespack M, Barker SA, Long R, Carson DD, et al. Heparanase degrades syndecan-1 and perlecan heparan sulfate: Functional implications for tumor cell invasion. J Biol Chem. 2004;279(9):8047-8055.

[2] Theocharis AD, Skandalis SS, Gialeli C, Karamanos NK. Extracellular matrix structure. Vol. 97, Advanced Drug Delivery Reviews. 2016. 4-27 p.

[3] Masola V, Bellin G, Gambaro G, Onisto M. Heparanase: A Multitasking Protein Involved in Extracellular Matrix (ECM) Remodeling and Intracellular Events. Cells. 2018;7(12):236.

[4] Piperigkou Z, Kyriakopoulou K, Koutsakis C, Mastronikolis S, Karamanos NK. Key Matrix Remodeling Enzymes: Functions and Targeting in Cancer. Cancers (Basel). 2021; 13(6):1441.

[5] Streuli CH. Integrins as architects of cell behavior. Mol Biol Cell. 2016; 27(19):2885-2888.

[6] Walker C, Mojares E, Del Río Hernández A. Role of extracellular matrix in development and cancer progression. Vol. 19, International Journal of Molecular Sciences. 2018.

[7] Buchheit CL, Rayavarapu RR, Schafer ZT. The regulation of cancer cell death and metabolism by extracellular matrix attachment. ScienceDirect. 2012;23(4):402-411.

[8] Ostapoff KT, Awasthi N, Cenik BK, Hinz S, Dredge K, Schwarz RE, et al. PG545, an angiogenesis and heparanase inhibitor, reduces primary tumor growth and metastasis in experimental pancreatic cancer. Mol Cancer Ther. 2013;12(7):1190-1201.

[9] Paoli P, Giannoni E, Chiarugi P. Anoikis molecular pathways and its role in cancer progression. Biochim Biophys Acta - Mol Cell Res [Internet]. 2013;1833(12):3481-3498. Available from: http://dx.doi.org/10.1016/j.bbamcr.2013.06.026

[10] Secchi MF, Masola V, Zaza G, Lupo A, Gambaro G, Onisto M. Recent data concerning heparanase: Focus on fibrosis, inflammation and cancer. Biomol Concepts. 2015;6(5-6):415-421.

[11] Iozzo R V., Gubbiotti MA. Extracellular matrix: The driving force of mammalian diseases. Matrix Biol [Internet]. 2018;71-72:1-9. Available from: https://doi.org/10.1016/j.matbio.2018.03.023

[12] Iozzo R V., Schaefer L. Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. Matrix Biol [Internet]. 2015;42:11-55. Available from: http://dx.doi.org/10.1016/j.matbio.2015.02.003

[13] Karsdal MA, Nielsen SH, Leeming DJ, Langholm LL, Nielsen MJ, Manon-Jensen T, et al. The good and the bad collagens of fibrosis – Their role in signaling and organ function. Adv Drug Deliv Rev. 2017;121:43-56.

[14] Hynes RO, Naba A. Overview of the matrisome-An inventory of extracellular matrix constituents and functions. Cold Spring Harb Perspect Biol. 2012;4(1): 1-16.

[15] Cagno V, Tseligka ED, Jones ST, Tapparel C. Heparan sulfate proteoglycans and viral attachment: True receptors or adaptation bias? Viruses. 2019;11(7):1-24.

[16] Efthymiou G, Saint A, Ruff M, Rekad Z, Ciais D, Van Obberghen-Schilling E. Shaping Up the Tumor Microenvironment With Cellular Fibronectin. Front Oncol. 2020; 10(April):1-18.

[17] Hammond E, Khurana A, Shridhar V, Dredge K. The role of heparanase and sulfatases in the modification of heparan sulfate proteoglycans within the tumour microenvironment and opportunities for novel cancer therapeutics. Front Oncol. 2014;4 JUL(July):1-15.

[18] Elgundi Z, Papanicolaou M, Major G, Cox TR, Melrose J, Whitelock JM, et al. Cancer Metastasis: The Role of the Extracellular Matrix and the Heparan Sulfate Proteoglycan Perlecan. Front Oncol. 2020;9(January).

[19] Bicho MC. Microbiologia Médica - Cap 44. Virologia Médica (Editado por Helena Barroso, António Meliço-Silvestre, Victor Duque e Nuno Taveira). Lidel Edições Técnicas , Lisboa - 2014. ISBN:978-972-757-576-3

[20] Carse S, Bergant M, Schäfer G. Advances in targeting hpv infection as potential alternative prophylactic means. Int J Mol Sci. 2021;22(4):1-28.

[21] Herbster S, Paladino A, de Freitas S, Boccardo E. Alterations in the expression and activity of extracellular matrix components in HPV-associated infections and diseases. Clinics. 2018;73:1-7.

[22] Leto M das GP, Santos Júnior GF Dos, Porro AM, Tomimori J. Human papillomavirus infection: etiopathogenesis, molecular biology and clinical manifestations. An Bras Dermatol [Internet]. 2011;86(2):306-17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21603814

[23] R.Keiffer LGMGMB-HMSSDT. Human Papillomavirus Major Capsid Protein L1 Remains Associated with the Incoming Viral Genome throughout the Entry Process. J Virol. 2017;91(16):1-17.

[24] Doorbar J, Quint W, Banks L, Bravo IG, Stoler M, Broker TR, et al. The biology and life-cycle of human papillomaviruses. Vaccine [Internet]. 2012;30(SUPPL.5):F55–F70. Available from: http://dx.doi.org/10.1016/j.vaccine.2012.06.083

[25] Raff AB, Woodham AW, Raff LM, Skeate JG, Yan L, Da Silva DM, et al. The Evolving Field of Human Papillomavirus Receptor Research: a Review of Binding and Entry. J Virol. 2013;87(11):6062-6072.

[26] Koganti R, Suryawanshi R, Shukla D. Heparanase, cell signaling, and viral infections. Cell Mol Life Sci [Internet]. 2020;77(24):5059-77. Available from: https://doi.org/10.1007/s00018-020-03559-y

[27] Clausen TM, Sandoval DR, Spliid CB, Pihl J, Perrett HR, Painter CD, et al. SARS-CoV-2 Infection Depends on Cellular Heparan Sulfate and ACE2. Cell. 2020;183(4):1043-1057.e15.

[28] Surviladze Z, Sterkand RT, Ozbun MA. Interaction of human papillomavirus type 16 particles with heparan sulfate and syndecan-1 molecules in the keratinocyte extracellular matrix plays an active role in infection. J Gen Virol. 2015;96(8):2232-2241.

[29] Cogliano V, Baan R, Straif K, Grosse Y, Secretan B, Ghissassi F El. 4. Molecular Mechanisms of HPV-induced Carcinogenesis. IARC Monogr Vol 90. 2007;90:432-65.

[30] IARC WORKING GROUP. Human Papillomavirus (HPV) Infection [Internet]. Vol. VOLUME 90, IARC MONOGRAPHS. International Agency for Research on Cancer; 2010. 11-12 p. Available from: http://www.cdc.gov/std/treatment/2010/hpv.htm

[31] Branca M, Ciotti M, Giorgi C, Santini D, Di Bonito L, Costa S, et al. Matrix metalloproteinase-2 (MMP-2) and its tissue inhibitor (TIMP-2) are

prognostic factors in cervical cancer, related to invasive disease but not to high-risk human papillomavirus (HPV) or virus persistence after treatment of CIN. Anticancer Res. 2006;26(2 B): 1543-56.

[32] Oxford JT, Reeck JC, Hardy MJ. Extracellular matrix in development and disease. Int J Mol Sci. 2019;20(1).

[33] Heyman B, Yang Y. Mechanisms of Heparanase Inhibitors in Cancer Therapy. Physiol Behav. 2019;176(3):139-48.https:// doi. org/10.1016/j.exphem.2016.08.006

[34] Thakkar N, Yadavalli T, Jaishankar D, Shukla D. Emerging roles of heparanase in viral pathogenesis. Pathogens. 2017;6(3):1-12.