Biochimica et Biophysica Acta xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Review

Cysteine cathepsins and extracellular matrix degradation ☆

Marko Fonović [a,b,*], Boris Turk [a,b,c,*]

[a] Department of Biochemistry, Molecular and Structural Biology, Jozef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia
[b] Centre of Excellence for Integrated Approaches in Chemistry and Biology of Proteins, Jamova cesta 39, SI-1000 Ljubljana, Slovenia
[c] Faculty of Chemistry and Chemical Technology, University of Ljubljana, Slovenia

ARTICLE INFO

Article history:
Received 17 January 2014
Received in revised form 16 March 2014
Accepted 22 March 2014
Available online xxxx

Keywords:
Cathepsin
Extracellular matrix
Lysosome
Collagen
Osteoporosis
Glycosaminoglycan

ABSTRACT

Background: Cysteine cathepsins are normally found in the lysosomes where they are involved in intracellular protein turnover. Their ability to degrade the components of the extracellular matrix *in vitro* was first reported more than 25 years ago. However, cathepsins were for a long time not considered to be among the major players in ECM degradation *in vivo*. During the last decade it has, however, become evident that abundant secretion of cysteine cathepsins into extracellular milieu is accompanying numerous physiological and disease conditions, enabling the cathepsins to degrade extracellular proteins.
Scope of view: In this review we will focus on cysteine cathepsins and their extracellular functions linked with ECM degradation, including regulation of their activity, which is often enhanced by acidification of the extracellular microenvironment, such as found in the bone resorption lacunae or tumor microenvironment. We will further discuss the ECM substrates of cathepsins with a focus on collagen and elastin, including the importance of that for pathologies. Finally, we will overview the current status of cathepsin inhibitors in clinical development for treatment of ECM-linked diseases, in particular osteoporosis.
Major conclusions: Cysteine cathepsins are among the major proteases involved in ECM remodeling, and their role is not limited to degradation only. Deregulation of their activity is linked with numerous ECM-linked diseases and they are now validated targets in a number of them. Cathepsins S and K are the most attractive targets, especially cathepsin K as a major therapeutic target for osteoporosis with drugs targeting it in advanced clinical trials.
General significance: Due to their major role in ECM remodeling cysteine cathepsins have emerged as an important group of therapeutic targets for a number of ECM-related diseases, including, osteoporosis, cancer and cardiovascular diseases. This article is part of a Special Issue entitled Matrix-mediated cell behaviour and properties.

1. Introduction

Extracellular matrix (ECM) is a complex macromolecular structure present in all organ and tissue types [1]. Besides providing basic physical scaffolding it also creates biochemical and biomechanical environment, which is vital for tissue homeostasis, morphogenesis and differentiation. ECM is a dynamic structure that is constantly remodeled through a large variety of post-translational modifications, which can be enzymatic or non-enzymatic. Among the enzymatic modifications, proteolytic processing and degradation are among the most important. Through proteolysis, proteases can modify the ECM structure and also release the bioactive factors, which influence growth, morphogenesis and pathological processes, thereby restructuring extracellular matrix in a rapid and irreversible manner. Due to their extracellular localization and enzymatic stability at neutral pH, metalloproteases and serine proteinases were traditionally considered to be the main agents of ECM degradation [2,3]. In contrast to that, lysosomal cysteine proteases or cysteine cathepsins, which require acidic pH for optimum activity, were mainly regarded as intracellular proteases involved in a general protein turnover. This view was shifted by the discovery that under specific physiological and pathological conditions, the cathepsins can also be secreted into the extracellular space where they can remain proteolyticaly active and degrade various components of the extracellular matrix [4,5]. Moreover, inhibition of their extracellular activity was also shown to be highly relevant in clinical applications [6–10]. In this review, we will present an overview of cysteine cathepsins and discuss how they can remodel the extracellular matrix including their importance in disease.

2. Cysteine cathepsins

Cathepsins are a diverse group of proteases which is composed of various catalytic types. Among them are serine cathepsins (cathepsins A and G), aspartic cathepsins (cathepsins D and E) and cysteine cathepsins which are the most abundant. Cysteine cathepsins belong to the C1 family (papain family) of CA clan of cysteine proteases [11]. Human

☆ This article is part of a Special Issue entitled Matrix-mediated cell behaviour and properties.

* Corresponding authors at: Department of Biochemistry, Molecular and Structural Biology, Jozef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia. Tel.: +386 1 477 3772; fax: +386 1 477 3984.
E-mail addresses: marko.fonovic@ijs.si (M. Fonović), boris.turk@ijs.si (B. Turk).

http://dx.doi.org/10.1016/j.bbagen.2014.03.017
0304-4165/

genome sequencing data revealed the existence of 11 human cysteine cathepsins (cathepsins B, C, F, H, K, L, O, S, V, X and W) [12]. They are monomeric proteins with molecular mass in the range of 20–35 kDa, the only exception being cathepsin C, which is a tetrameric protein with molecular mass around 200 kDa. Regarding their substrate preference, cysteine cathepsins are mainly endopeptidases (cathepsins F, K, L, O, S, V and W), whereas the aminopeptidase cathepsin C and the carboxypeptidase cathepsin X are strict exopeptidases. Although primarily exopeptidases, cathepsins B and H can also act as endopeptidases (Table 1). The majority of cathepsins are ubiquitously expressed in human tissues, while cathepsins K, S, V and W show more tissue-specific distribution [4,9,13–19]. Such diverse tissue-specific expression of some cysteine cathepsins suggests that their action is not limited solely to the bulk protein degradation but is also involved in the regulation of specific cellular functions. This was confirmed by the finding that active cysteine cathepsins are not localized strictly in the lysosomes but can also be found in other cellular compartments, including the nucleus, where they perform very specific functions [20–22]. All these help the cathepsins to regulate numerous important physiological processes, among others intracellular protein turnover, immune response, bone remodeling and prohormone processing [5–7,14,23].

2.1. Structure and specificity

Cysteine cathepsins have a characteristic papain-like fold, which is composed of two domains: left (L-), which is formed by three α-helices, and right (R-), which has a β-barrel structure. The two domains form an active site cleft, where the two active-site residues (Cys25 and His159, papain numbering) are located (Fig. 1a). The two residues form a thiolate-imidazolium ion pair, which is critical for the proteolytic activity. Binding surfaces for the substrate residues are provided by both domains and substrate binds along the active-site cleft with its side chains alternately oriented towards the L- and R- domains [4,5,14]. According to the Schechter and Berger nomenclature, substrate amino acid residues (P) and substrate binding pockets (S) are numbered according to their position relative to the scissile peptide bond (Fig. 1b). The positions towards the N-terminus are denoted as the "non-prime" residues and sites (P1–Pn; S1–Sn), whereas those towards the C-terminus are denoted as the "prime" residues and sites (P1′–Pn′ and S1′–Sn′) [24]. Structural studies have shown that in cysteine cathepsins only the S2, S1 and S1′ binding sites are well defined and that the S2 and S1′ sites are the main determinants of the substrate specificity [25]. Contrary to some other cysteine proteases such as caspases or legumain, cysteine cathepsins exhibit much broader specificity [25]. This suggests that their specificity is determined by a cumulative effect of interaction with multiple binding sites and is not governed by a single site. Even the S2 subsite, which is structurally the most stringently defined (it is the only one which forms an actual pocket"), can be occupied by various hydrophobic or aromatic amino acid residues, such as leucine, isoleucine, valine, tyrosine and phenylalanine [25–28]. In addition, profiling substrate specificity of cysteine cathepsins using a combinatorial peptide library revealed a unique specificity for cathepsin K for Pro and Gly in the P2 and P3 positions, respectively, which matches well the cathepsin K specificity on ECM proteins such as collagen [27]. In addition, cysteine cathepsins that exhibit exopeptidase activity (cathepsins B, C, H and X), have additional structural features, which limit the substrate access to the active site and regulate exo- and endopeptidase activities of cathepsins B and H [5,29–35].

2.2. Localization and regulation of cathepsin activity

Cysteine cathepsins are expressed as inactive preproenzymes in the endoplasmic reticulum. The signal peptide is proteolyticaly removed during the transport through the lumen of the rough endoplasmic reticulum, followed by procathepsin glycosylation in the Golgi. In the trans-Golgi network, glycosylated procathepsins are tagged with the mannose-6-phosphate marker, which enables their sorting to the endolysosomal compartments [36]. However, lysosomal sorting of cathepsins can also be independent of the mannose-6-phosphate pathway [37]. Inside the lysosome, procathepsins are then activated by proteolytic removal of their prodomains. Their activation can be autocatalytic or catalyzed by other proteases [14,38]. However, even during autocatalytic activation, activation is always in trans, *i.e.* one cathepsin or procathepsin molecule activating another procathepsin molecule [39]. In contrast to endopeptidases, exopeptidases cathepsins C and X cannot be autocatalytically activated but require an endopeptidase such as cathepsin L or S for their activation [40,41]. A major role in facilitating autocatalytic activation of cysteine cathepsins also has various glycosaminoglycans (GAGs) and other negatively charged surfaces [42–47]. It seems that the molecular mechanism involves a conformational change in the cathepsin zymogen, induced by GAG binding, which loosens the interaction between the propeptide and the mature part of the enzyme, thereby enabling easier processing of another procathepsin molecule [44]. Moreover, GAG and other negatively charged polymers such as dextran sulfate also facilitated autocatalytic activation of the cathepsins at pH values close to neutral, which may have an important function in disease states [43,48]. In addition, GAG binding was also found to have a stabilizing role thereby protecting the cathepsins against pH-induced inactivation [49]. However, there is no conserved GAG-binding surface that would be common to all cathepsins.

For their optimal activity, cathepsins require reducing and slightly acidic environment such as found in the lysosome. This linked with their expression in the form of inactive zymogens and their endo/lysosomal compartmentalization led to the general idea that cathepsins are primarily endolysosomal enzymes involved in general protein turnover. Moreover, their localization inside the lysosomes, a general instability at neutral pH and a requirement for reducing conditions were long believed to protect the cell and limit unwanted proteolysis. If any of these would fail, activity of cysteine cathepsins that are released from the lysosomes or secreted is additionally regulated by the endogenous

Table 1
Localization and specificity of cysteine cathepsins.

Name	Gene symbol	Substrate cleavage mode	Extracellular localization	Tissue and cell specificity
Cathepsin B	CTSB	Carboxydipeptidase endopeptidase	Secreted membrane bound	Ubiquitous
Cathepsin C	CTSC	Aminodipeptidase	n.d.	Ubiquitous
Cathepsin F	CTSF	Endopeptidase	Secreted	Ubiquitous
Cathepsin H	CTSH	Aminomonopeptidase endopeptidase	Secreted	Ubiquitous
Cathepsin L	CTSL	Endopeptidase	Secreted	Ubiquitous
Cathepsin K	CTSK	Endopeptidase	Secreted	Predominant in osteoclasts and synovial fibroblasts
Cathepsin O	CTSO	n.d.	n.d.	ubiquitous
Cathepsin S	CTSS	Endopeptidase	Secreted	Predominant in antigen presenting cells
Cathepsin V	CTSL2	Endopeptidase	n.d.	Thymus, testis
Cathepsin W	CTSW	n.d.	n.d.	Cytotoxic lymphocytes
Cathepsin X	CTSZ	Carboxymonopeptidase	Secreted membrane bound	Ubiquitous

Several properties of cysteine cathepsins are shown, including their preference to cleave substrates as endo- or exopeptidases, their tissue and/or cell specific expression preference, as well as information about their potential extracellular localization (n.d., not determined). Original references can be found within [4,6,9,14,23,52,83].

Please cite this article as: M. Fonović, B. Turk, Cysteine cathepsins and extracellular matrix degradation, Biochim. Biophys. Acta (2014), http://dx.doi.org/10.1016/j.bbagen.2014.03.017