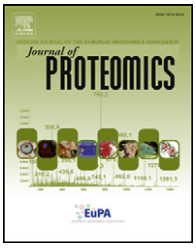
Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jprot



Proteomic protease specificity profiling of clostridial collagenases reveals their intrinsic nature as dedicated degraders of collagen☆,☆☆

Ulrich Eckhard[a,b], *Pitter F. Huesgen*[a], *Hans Brandstetter*[b], *Christopher M. Overall*[a,*]

[a]*Centre for Blood Research, Department of Oral Biological and Medical Sciences, University of British Columbia, 2350 Health Sciences Mall, Vancouver, British Columbia V6T 1Z3, Canada*

[b]*Division of Structural Biology, Department of Molecular Biology, University of Salzburg, Billrothstr, 11, 5020 Salzburg, Austria*

ARTICLE INFO

Available online 11 October 2013

Keywords:
Clostridia
Collagenase
MMPs
PICS
Mass spectrometry

ABSTRACT

Clostridial collagenases are among the most efficient degraders of collagen. Most clostridia are saprophytes and secrete proteases to utilize proteins in their environment as carbon sources; during anaerobic infections, collagenases play a crucial role in host colonization. Several medical and biotechnological applications have emerged utilizing their high collagenolytic efficiency. However, the contribution of the functionally most important peptidase domain to substrate specificity remains unresolved. We investigated the active site sequence specificity of the peptidase domains of collagenase G and H from *Clostridium histolyticum* and collagenase T from *Clostridium tetani*. Both prime and non-prime cleavage site specificity were simultaneously profiled using Proteomic Identification of protease Cleavage Sites (PICS), a mass spectrometry-based method utilizing database searchable proteome-derived peptide libraries. For each enzyme we identified >100 unique-cleaved peptides, resulting in robust cleavage logos revealing collagen-like specificity patterns: a strong preference for glycine in P3 and P1′, proline at P2 and P2′, and a slightly looser specificity at P1, which in collagen is typically occupied by hydroxyproline. This specificity for the classic collagen motifs Gly-Pro-X and Gly-X-Hyp represents a remarkable adaptation considering the complex requirements for substrate unfolding and presentation that need to be fulfilled before a single collagen strand becomes accessible for cleavage.

Biological significance

We demonstrate the striking sequence specificity of a family of clostridial collagenases using proteome derived peptide libraries and PICS, Proteomic Identification of protease Cleavage Sites. In combination with the previously published crystal structures of these proteases, our results represent an important piece of the puzzle in understanding the complex mechanism underlying collagen hydrolysis, and pave the way for the rational design of specific test substrates and selective inhibitors.

☆☆ This article is part of a Special Issue entitled: Can Proteomics Fill the Gap Between Genomics and Phenotypes?

* *Corresponding author at:* Centre for Blood Research, Department of Oral Biological and Medical Sciences, University of British Columbia, 4.401 Life Sciences Centre, 2350 Health Sciences Mall, Vancouver, BC V6T 1Z3, Canada. Tel.: +1 604 822 2958; fax: +1 604 822 7742.
E-mail address: chris.overall@ubc.ca (C.M. Overall).

1874-3919/$ – see front matter
http://dx.doi.org/10.1016/j.jprot.2013.10.004

This article is part of a Special Issue entitled: Can Proteomics Fill the Gap Between Genomics and Phenotypes?

1. Introduction

Collagens are by far the most abundant proteins in mammals that constitute up to 90% of the extracellular matrix of a tissue, thus imparting its structural integrity. Collagen is an astonishing molecule of remarkable primary, secondary, tertiary and quaternary structure, fascinating scientists for decades [1–4]. Interstitial collagen is built up by three polypeptide strands with high proline and hydroxyproline (Hyp) content (combined occurrence >20%) and a glycine at every third position. Pro-Hyp-Gly thus forms the most common triplet (10.5%). Each strand possesses the conformation of a left-handed polyproline II helix, twisted together into a rigid, rod shaped, right-handed super helix of approximately 300 nm in length and 1.5 nm in diameter, burying the peptide bonds within the interior of the helix. The three alpha-chains are held together by numerous inter-chain hydrogen bonds, and highly ordered hydration networks surround the triple helices. Collagen molecules spontaneously self-assemble into larger quaternary assemblies of great tensile strength, first as fibrils and then as bundles of fibrils to fibers up to 2 μm in length and 300 nm in diameter. Additional intermolecular interactions increase collagen thermal stability and afford its great tensile strength and long half-life of 15 years for skin collagen and of over 100 years for cartilage collagen [5–9].

Equally fascinating as collagen is the repertoire of proteases, often termed collagenases, capable of degrading native triple helical collagen. Human proteases such as the zinc-dependent matrix metalloproteinases (MMP) 1, 8, 13 and 14, and the cysteine protease cathepsin K, play crucial roles in the dynamic remodeling of connective tissue under physiological and/or pathological conditions [10,11]. Notably, all enzymes show distinct preferences towards the different fibrillar collagen substrates and cleave only at well-defined sites; *e.g.* MMP1, MMP8, MMP13 and MMP14 all share the unique ability to cut the native super-helix into 3/4 and 1/4 length fragments at a single peptide bond (775–776) between a glycine residue at P1 and a leucine (on type I collagen α2) or isoleucine (on type I collagen α1) at P1′. However, their distinct preferences for different types of interstitial collagens indicate different functional roles in the human body [12]. Notably, the hemopexin carboxyl-domain exosites bind collagen and are essential for cleavage [13–16]. Exosites represent specialized substrate binding sites located on domains outside the active site cleft, providing substrate interactions not influenced by the primary specificity pockets, thus refining substrate affinity and proteinase specificity [11]. For cathepsin K, five distinct cleavage sites have been identified within type I collagen, and one in type II collagen. In all cases, subsequent collagenolysis is performed by gelatinases (MMP2, MMP9) and other synergistic proteinases only after the initial dismantling of the collagen triple helix that exposes more generally susceptible proteolytic sites upon denaturation [10,17,18]. However, MMP2 also has a weak native collagenase activity imparted by its three fibronectin type II repeats [15] which function as exosites to localize the enzyme on the substrate.

Bacterial collagenases are secreted by saprophytic clostridia to utilize collagen as a carbon source. Pathogenic strains such as *Clostridium histolyticum* and *Clostridium tetani* use these enzymes to facilitate host invasion, colonization and toxin diffusion during anaerobic infections [19,20]. In contrast to the tightly regulated natural turnover of collagen in the host organism, clostridial collagenases circumvent all constraints for collagen degradation and are capable of efficiently digesting most, if not all, types of collagen [21–23]. Their enzymatic activity is independent of the quaternary collagen assembly, which is especially remarkable given collagen's triple helical structure that renders it resistant to most other proteases [24–26]. Clostridial collagenases not only initiate hydrolysis by multiple cleavages within the triple helical region, but also completely degrade the obtained fragments into a mixture of small oligopeptides [22,24,27]. This ability is often referred to as broad substrate specificity of clostridial collagenases [21–23], although no natural substrate other than collagen has been reported to our knowledge. These properties place clostridial collagenases among the most efficient enzymes degrading all types of collagen, which is also exploited by a diverse spectrum of medical and biotechnological applications. For instance, collagenases from *C. histolyticum* are approved therapeutic agents for breaking down the tough collagen cords in Morbus Dupuytren [28,29] and are widely used in enzymatic wound debridement [30,31] and tissue dissociation experiments [32,33].

Clostridial collagenases are large multi-modular zinc-metalloproteinase of approximately 115 kDa, consisting of four to six domains [19,26,34–36] (Fig. 1). As members of the gluzincin superfamily of metalloproteinases they share the common HExxH zinc binding motif within the peptidase domain of the N-terminal collagenase unit, complemented by an additional zinc-coordinating glutamate 28 to 30 amino acid downstream. Two to four accessory domains of approximately 10 kDa each form a C-terminal collagen recruitment unit of variable composition (Fig. 1), providing important exosites for native collagenolysis, responsible for collagen binding [37] and swelling [38] (Fig. 1). Although most clostridial strains possess only one collagenase, *C. histolyticum* encodes for two with complementary characteristics: collagenase G (ColG) exhibits high collagenolytic and low peptidolytic activity, whereas its homologue collagenase H (ColH) shows low collagenolytic and high peptidolytic activity [39].

Recently, we established a flexible expression and purification platform for clostridial collagenases [35], undertook a 'one substrate, one inhibitor'-based biochemical characterization of the collagenase units of ColG, ColH and the *C. tetani* collagenase (ColT) [40], and reported the crystal structures of the ColG collagenase unit [36], the ColG Polycystic Kidney Disease (PKD)-like domains [37], and of the ColH and ColT peptidase domains [41]. The ColG crystal structure showed a saddle-shaped architecture of the N-terminal collagenase unit (Tyr119-Gly790) capable of degrading native triple-helical collagen even in the absence of its accessory collagen recruitment domain [36]. The N-terminal activator domain (Tyr119-Asp388) was required for