Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic Chemistry 31 (2003) 425–436
BIOORGANIC CHEMISTRY
www.elsevier.com/locate/bioorg

Minireview

Cathelicidin family of antimicrobial peptides: proteolytic processing and protease resistance

Ann Eisenberg Shinnar,* Kathryn L. Butler, and Hyon Ju Park

Chemistry Department, Barnard College, Columbia University, New York, NY 10027, USA

Received 10 October 2002

Abstract

Cathelicidins are a gene family of antimicrobial peptides produced as inactive precursors. Signal peptidase removes the N-terminal signal sequence, while peptidylglycine α-amidating monooxygenase often amidates and cleaves the C-terminal region. Removal of the cathelin domain liberates the active antimicrobial peptide. For mammalian sequences, this cleavage usually occurs through the action of elastase, but other tissue-specific processing enzymes may also operate. Once released, these bioactive peptides are susceptible to proteolytic degradation. We propose that some mature cathelicidins are naturally resistant to proteases due to their unusual primary structures. Among mammalian cathelicidins, proline-rich sequences should resist attack by serine proteases because proline prevents cleavage of the scissile bond. In hagfish cathelicidins, the unusual amino acid bromotryptophan may make the active peptides less susceptible to proteolysis for steric reasons. Such protease resistance could extend the pharmacokinetic lifetimes of cathelicidins in vivo, sustaining antimicrobial activity.

Keywords: Antimicrobial peptides; Cathelicidins; Cathelin; Proteolytic processing; Proline-rich sequences; Bromotryptophan; Carboxyamidation

1. Introduction

Over the last 25 years, a wide array of peptides exhibiting broad-spectrum antimicrobial activity has been discovered in plants, insects, and higher animals [1].

* Corresponding author. Fax: 1-212-854-2310.
E-mail address: ashinnar@barnard.edu (A.E. Shinnar).

0045-2068/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0045-2068(03)00080-4

These polycationic peptides inactivate microbes through a general mechanism, by binding preferentially to their membranes and causing membrane disruption. Antimicrobial peptides play a fundamental role in providing innate immunity, prior to stimulating the adaptive immune system of higher animals, which involves production of antibodies and specialized cells [2].

These gene-encoded peptides are first produced as inactive precursors, mostly by epithelial cells and populations of phagocytic cells, and are released via proteolytic processing with specific enzymes [3]. Although varying in primary structure and length, the mature peptides readily adopt amphipathic α-helical or β-sheet secondary structures whereby segregation of cationic and hydrophobic amino acids facilitates membrane binding and penetration [1,4,5]. These bioactive peptides could be inactivated, however, if they encounter a variety of proteases. Hence, the lifetime and efficacy of antimicrobial peptides are dependent upon two opposing processes, i.e., specific protease activation and indiscriminate degradation. A review of the cathelicidin antimicrobial peptides reveals that primary structure features of this family may counterbalance these opposing proteolytic processes.

2. Proteolytic activation of cathelicidins

Cathelicidins are a family of antimicrobial peptides with a distinctive gene structure. The precursor region is highly conserved, consisting of a signal peptide followed by a domain with homology to the cathelin protein. The cationic antimicrobial peptides released from the C-terminal region, in contrast, are highly variable in sequence. The cathelicidin family name is based on its limited homology to cathelin, a protein acronym for cathepsin-L-inhibitor, and a member of the cystatin family of cysteine protease inhibitors. Cathelicidins have been detected in a variety of mammals, where they are expressed in neutrophils, hematopoietic cells, and even some epithelial tissues [6–10].

Proteolytic processing of the prepropeptide must occur in order to release the antimicrobial peptide in its active form (Fig. 1). Signal peptidase first removes the signal sequence, which directs the translocation and secretion of the polypeptide during biosynthesis. This universal processing enzyme recognizes the distinctive primary and secondary structure motifs characteristic of the signal peptide domain [11,12]. The action of peptidylglycine α-amidating monooxygenase (PAM) can account for the processing in the C-terminal domain for many cathelicidins [13–15]. Through hydroxylation and oxidative cleavage, this bifunctional enzyme transforms glycine and converts the antimicrobial peptide into a carboxyamide.

Release of the active antimicrobial peptides depends upon a processing site between the cathelin and antimicrobial peptide domains. For most of the bovine and porcine cathelicidins, this is accomplished by elastase, present in neutrophils. The specificity pocket of this serine protease is known to accommodate small, neutral sidechains such as glycine, alanine, serine, and valine. In mammalian cathelicidins, valine appears most frequently at this cleavage site (Table 1); alanine and isoleucine are also found [6]. Recent X-ray crystallographic studies of protegrin-3, a porcine

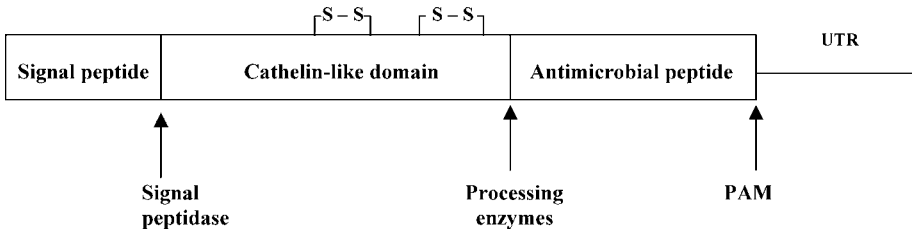

Fig. 1. Precursor structure of cathelicidin antimicrobial peptides. Cathelicidins exhibit a highly conserved signal peptide and cathelin-like domain, followed by a highly variable antimicrobial peptide domain. The 3D motif for the cathelin-like domain of protegrin, recently revealed by X-ray crystallography, shows the characteristic cystatin family fold, consisting of an α-helix packed against four strands of antiparallel β-sheet, cross-linked by two disulfide bridges [16]. Some antimicrobial peptides in this gene family are carboxyamidated by PAM and have a 3′-untranslated region (UTR), such as observed for the hagfish cathelicidins [19].

cathelicidin, have revealed that its elastase-sensitive sequence is located in a β-hairpin turn, within a solvent-accessible loop [16,17]. This region may serve as a hinge that exposes the cleavage site to elastase upon conformational change [16].

Elastase is not effective in cleaving human CAP-18, where threonine is situated between domains (Table 1). Instead, proteinase 3 liberates LL-37 as the active peptide, after exocytosis [18]. Hagfish cathelicidins, the most ancient member of this gene family, display four arginines positioned between the cathelin domain and the antimicrobial sequences [19]. Some prohormone convertases (PC; also known as Kex2 or furin proteases) are specific for arginine tetrads, and process latent zymogens [13,14]. Presumably such a putative enzyme operates in the specialized cells of the hagfish intestinal submucosa. Thus, this critical step in processing cathelicidins is governed by different enzymes in different species.

The biological function of the cathelin-like domain, released from this precursor structure by processing enzymes, is still unknown. One early suggestion that the cathelin domain plays a role as an inhibitor of cysteine proteases has not been thoroughly explored by experimental studies. Since this cathelin-like sequence is rich in acidic residues, it has also been proposed that this anionic domain binds electrostatically to the polycationic peptide, preventing its antimicrobial function until time of release. The availability of the cathelin domain from overexpression of porcine protegrin [17] should prompt research to elucidate the biological role of this enigmatic domain [10].

Once liberated from their precursor structures, the cationic antimicrobial peptides are attracted through electrostatic interactions to the cell surface of many microorganisms [4,5]. In vitro studies show that cathelicidins exhibit broad-spectrum antimicrobial activity against bacteria and fungi, including multidrug resistant (MDR) pathogens [20]. In addition, a variety of recent studies suggest that some cathelicidins are involved in other biologically important functions, including chemoattraction of phagocytic neutrophils, antisepsis through binding of bacterial endotoxin, and stimulation of components of the adaptive immune system [2]. These biological activities, in addition to the prominent anti-infection function, can be diminished as the active

Table 1
Mammalian cathelicidins: cleavage sites and proline content

Name	Genus	Mature antimicrobial peptide sequence (↓ cleavage site)	Proline content
Bactenecins			
Bac4	*Bos*	ELQSV RRLH**P**QHQRF**P**RER**P**W**P**K**P**LSL**P**L**P**R**P**G**P**R**P**W**P**K**P**L	**31**% (12/36)
Bac5	*Bos*	ELQSV RFR**PP**IRR**PP**IR**PP**FY**PP**FR**PP**IR**PP**IF**PP**IR**PP**FR**PP**LG**P**F**P**G	**47**% (20/43)
Bac7	*Bos*	ELQSV RRIR**P**R**PP**RL**P**R**P**R**P**RL**P**FF**P**R**P**G**P**R**P**IP**R**PL**P**FF**P**R**P**G**P**R**P**IP**R**PL	**47**% (28/60)
Bac5	*Capra*	ELQSV RFR**PP**IRR**PP**IR**PP**FN**PP**FR**PP**VR**PP**FR**PP**FR**PP**FR**PP**IG**P**F**P***G*	**47**% (20/43)
Bac5	*Ovis*	ELQSV RFR**PP**IRR**PP**IR**PP**FR**PP**FR**PP**VR**PP**IR**PP**FR**PP**FR**PP**IG**P**F**P***G*	**47**% (20/43)
Bac6	*Ovis*	EFQSV RRLR**P**RHQHF**P**SER**P**W**P**K**P**L**P**L**P**LR**P**G**P**R**P**W**P**K**P**L**P**L**P**L**P**R**P**GLR**P**W**P**K**P**L	**38**% (20/52)
Bac7.5	*Ovis*	ELQSV RRLR**P**RR**P**RL**P**R**P**R**P**R**P**R**P**RSL**P**LR**P**Q**P**RRIR**P**RILL**P**WR**PP**R**P**IR**P**Q**P**Q**P**IR**P**WL	**38**% (23/60)
Bac11	*Ovis*	ELQSV RRLR**P**RR**P**RL**P**R**P**R**P**R**P**R**P**RSL**P**LR**P**K**P**R**P**IR**P**L**P**L**P**R**P**R**P**K**P**IR**P**L**P**L**P**R**P**R **P**RRIR**P**L**P**L**P**R**P**R**P**IR**P**L**P**L**P**Q**P**Q**P**S**P**IR**P**L	**45**% (42/94)
Prophenins and proline-rich porcine peptides			
Prophenin-1	*Sus*	CNEGV RRF**P**WWW**P**FLRR**P**RLRRQ‡AF**PPP**NF**P**G**P**RF**PPP**NV**P**G**P**RF**PPP**NF**P**G**P**RF**PPP**NF**P**G**P** RF**PPP**NF**P**G**PP**F**PPP**IF**P**G**P**WF**PPPPP**FR**PPP**FG**PP**RF**P***G*	**53**% (42/79)
Prophenin-2	*Sus*	CNEGV RRF**P**WWW**P**FLRR**P**RLRRQAF**PPP**NV**P**G**P**RF**PPP**NV**P**G**P**RF**PPP**NF**P**G**P**RF**PPP**NF**P**G**P** RF**PPP**NF**P**G**PP**F**PPP**IF**P**G**P**WF**PPPPP**FR**PPP**FG**PP**RF**P***G*	**46**% (45/97)
PR-39	*Sus*	EIQSV RRR**P**R**PP**YL**P**R**P**R**PPP**FF**PP**RL**PP**RI**PP**GF**PP**RF**PP**RF**P**G	**49**% (19/39)
Tryptophan-rich peptide			
Indolicidin	*Bos*	ELQSV IL**P**WKW**P**WW**P**WRRG	**23**% (3/13)
Myeloid antimicrobial peptides (MAPs)			
BMAP-27	*Bos*	ELQSV GRFKRFRKKFKKLFKKLS**P**VI**P**LLHLG	**8**% (2/26)
BMAP-28	*Bos*	V**P**QSV GGLRSLGRKILRAWKKYG**P**IIV**P**IIRI*G*	**7**% (2/27)
BMAP-34	*Bos*	NIQSA GLFRRLRDSIRRGQQKILEKARRIGERIKDIFR*G*	**0**% (0/33)
eCath-1	*Equus*	E**P**QRV KRFGRLAKSFLRMRILL**P**RRKILLAS	**4**% (1/26)
eCath-2	*Equus*	KPQ**P**I KRRHWF**P**LSFQEFLEQLRRFRDQL**P**F**P**	**11**% (3/27)
eCath-3	*Equus*	E**P**QRV KRFHSVGSLIQRHQQMIRDKSEATRHGIRIITR**P**KLLLAS	**3**% (1/40)
OMAP-29	*Ovis*	E**P**QSV RGLRRLGRKIAHGVKKYG**P**TVLRIIRIA*G*	**4**% (1/28)
OMAP-34	*Ovis*	DIQNV GLFGRLRDSLQRGGQKILEKAERIGDRIKDIFR*G*	**0**% (0/33)

PMAP-23	*Sus*	QLQSV RIIDLLWRVRR**P**QK**P**KFVTVWVR	**9%** (2/23)
PMAP-36	*Sus*	EIQSV GRFRRLRKKTRKRLKKIGKVLKWI**PP**IVGSI**P**LGCG	**9%** (3/35)
PMAP-37	*Sus*	EIQSV GLLSRLRDFLSDRGRRLGEKIERIGQKIKDLSEFFQS	**0%** (0/37)
Cationic antibacterial peptides (CAPs) and primate peptides			
CAP11	*Cavia*	FRRMV GLRKKFRKTRKRIQKLGRKIGKTGRKVWKAWREYGQI**P**Y**P**CRI	**5%** (2/43)
CAP18	*Oryct.*	S**PE**P**T** GLRKRLRKFRNKIKEKLKKIGQKIQGLL**P**KLA**P**RTDY	**5%** (2/37)
CRAMP	*Mus*	ISRLA GLLRKGGEKIGEKLKKIGQKIKNFFQKLV**P**Q**P**EQ	**6%** (2/34)
LL-37	*Homo*	NKRFA LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLV**P**RTES	**3%** (1/37)
RhLL-37	*Macaca*	NRRSA RLGNFFRKVKEKIGGGLKKVGQKIKDFLGNLV**P**RTAS	**3%** (1/37)
Disulfide-rich peptides			
Cyclic-12	*Bos*	**PP**QAA RLCRIVVIRVCR	**0%** (0/12)
Cyclic-12	*Ovis*	**PP**QAA RICRIIFLRVCR	**0%** (0/12)
Protegrin-1	*Sus*	EVQGV RGGRLCYCRRRFCVCVGR*G*	**0%** (0/18)
Protegrin-2	*Sus*	EVQGV RGGRLCYCRRRFCICV*G*	**0%** (0/16)
Protegrin-3	*Sus*	EVQGV RGGGLCYCRRRFCVCVGR*G*	**0%** (0/18)
Protegrin-4	*Sus*	EVQGV RGGRLCYCRGWICFCFGR*G*	**0%** (0/18)
Protegrin-5	*Sus*	EVQGV RGGRLCYCR**P**RFCVCVGR*G*	**6%** (1/18)

Primary structures for 35 mammalian cathelicidins are grouped by their common names within a genus and ranked by proline content. Sequences were retrieved from the Antimicrobial Sequence Data Base (AMSDb; www.bbcm.univ.trieste.it/~tossi/antimic.html) or GenBank, and then aligned beginning with five residues before the amino terminus of the mature antimicrobial peptide ([19]). In 27 of these 35 examples, valine is the residue just prior to the cleavage site, indicated by an arrow (↓). For prophenin-1, a second processing site (‡) has been proposed to release the mature peptide [48]. The mature peptides range in length from 12 to 94 amino acids (average length = 37 residues). All prolines are shown in bold, but the proline content was calculated only for the antimicrobial peptide (#proline/#AA in mature peptide). Carboxyamidation is implicated by the presence of glycine at the C-terminus for 18 peptides, based on cDNA sequencing; these extra glycines, shown in italics, are not included in the count of AA in the mature peptide. Of these antimicrobial peptides, 26 contain proline while 9 do not (74 and 26% of 35 sequences, respectively). Modest proline content of 3–11% is observed in 14 sequences (40% of 35). One-third of the sequences (12 of 35) have remarkably high proline content of 23–53%. Many antimicrobial peptides isolated from insects are also rich in proline, with an average of 24% proline (based on 17 peptides surveyed from AMSDb, [21]). Peptide sequences come from the following mammalian species: *Bos taurus* (cattle), *Capra hirca* (goat), *Cavia porcellus* (guinea pig), *Equus caballus* (horse), *Homo sapiens* (human), *Macaca mulatta* (rhesus monkey), *Mus musculus* (mouse), *Oryctolagus cuniculus* (rabbit), *Ovis aries* (sheep), and *Sus scrofa* (pig). Some of the peptide nomenclature are acryonyms (based on original reports cited in [9]): bactenecin (bacterium necare, from Latin, to kill); prophenin (proline phenylalanine rich); indolicidin (indole-rich cidal peptide); protegrin from protegre (Latin, to protect); cationic antimicrobial peptides (CAPs); cathelin-related antimicrobial peptide (CRAMP); myeloid antimicrobial peptides (MAPs).

peptides encounter various endogenous proteases, from extracellular and intracellular sources, which can ultimately hydrolyze and thereby inactivate these biologically important peptides.

Some members of the cathelicidin family of antimicrobial peptides show an abundance of particular amino acids and post-translational modifications [5,6]. Their primary structure may afford special stability and resistance to proteolysis, as proposed below.

3. Protease resistance features of cathelicidins

3.1. Arginine/proline-rich mammalian cathelicidins

Cationic sidechains are a universal feature of the antimicrobial peptides and are essential for electrostatic interactions with the anionic lipids of bacterial membranes [1]. All of the mammalian cathelicidins are rich in arginine [21]. Although the majority of mature cathelicidin peptides (26/35 or 74%) contain proline, varying in content from 3 to 53% (Table 1), the proline content is modest (3–11%) in 40% of these antimicrobial peptides, and is thus similar to the average value (5%) typical of proteins [22]. The bactenecins, prophenins, and PR-39, however, have remarkably high proline content of 31–53%; these 11 sequences account for about one-third of all known mammalian cathelicidins. Because proline thwarts proteolysis of peptidic bonds, these exceptionally high proline contents may help prevent degradation of these peptides.

Protease specificity depends both on the nature of the peptide substrate (P sites) and the subsites (S sites) within the active site of the enzyme (see Fig. 2 for scheme based on [23,24]). Cleavage of the scissile bond between P_1 and P_1' is dictated by the identity of P_1, and is also affected by neighboring sidechains. It is widely known that

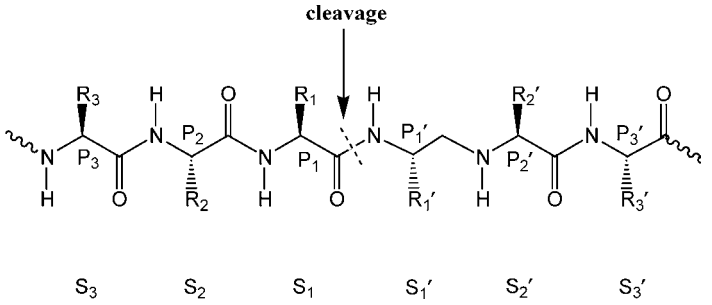

Fig. 2. Schematic of protease specificity. Protease specificity is dependent both upon the amino acid sidechains of the peptidic substrate (P_3 through P_3') and the surface contacts with the enzyme's subsites (S_3 through S_3'), according to the scheme of Schecter and Berger [23]. In serine proteases, the presence of proline in the P_1' and P_2' positions prevents the acyl transfer step from occurring [26]. The activity of elastase is further restricted when proline occupies the P_3 position [27]. Proline also restricts the activity of some aminopeptidases [30].

proline in the P_1' position prevents serine proteases from cleaving the peptide bond [25]. Studies using model peptides show that the presence of proline in both the P_1' and P_2' positions actually prevents the acyl transfer step from occurring with chymotrypsin, trypsin, α-lytic protease, and cercarial protease [26]. Furthermore, studies on elastase show that proline in the P_3 position also restricts this enzyme from effecting productive cleavages, due to poor contacts with the S_3 subsite [27].

For cathelicidins rich in proline, the mature antimicrobial peptides would thus become very poor substrates for proteolysis, particularly via members of the serine protease family. Prolines should dramatically prevent the action of elastase, an endogenous enzyme to which many cathelicidins would be susceptible after their initial release. Inspection of the primary structure for the bactenecins and prophenins reveals glycine, isoleucine and valine residues with proline in the adjacent P_1' position (Table 1). In porcine prophenin (so named because the peptide is rich in both proline and phenylalanine), numerous proline residues are found in the P_1' register with respect to phenylalanines (and even tryptophan). This pattern strongly suggests that serine proteases with specificity pockets large enough to accommodate aromatic residues, such as cathepsin G and chymase, would not degrade prophenin significantly [2,28]. In humans, cathepsin G is abundant in neutrophils, where it is thought to play a role in triggering signals involved in both innate and adaptive immune responses [2]. Prophenin isolated from porcine leukocytes should be naturally resistant to proteolysis by similar proteases.

Similarly for PR-39, where only a single isoleucine is present, many of the prolines are positioned next to or near bulkier hydrophobic residues, such as leucine and phenylalanine. For PR-39, resistance to intracellular proteases is particularly important. In some strains of *Escherichia coli*, PR-39 interferes with synthesis of DNA and proteins [29], implying that these peptides enter the bacterial cytoplasm after penetrating the membrane. Other members of the mammalian cathelicidins have only 2 or 3 proline residues, but their position in the sequence might be significant in restricting proteolytic degradation. For example, in PMAP-36 and BMAP-27, proline follows serine and isoleucine residues, which normally would be susceptible to cleavage by elastase. Many of the antimicrobial peptides isolated from insects are also proline-rich (21) and perhaps serve a similar protease-resistance function.

The role of proline in preventing indiscriminate proteolytic degradation has also been proposed in other contexts. Proline in the penultimate N-terminal position of proteins could restrict those aminopeptidases with broad specificity from achieving rapid proteolysis [30]. Proline is also overrepresented at positions immediately following dibasic sequences that often are recognized by processing enzymes as cleavage sites for prohormones and proproteins, thus limiting undesirable proteolysis at the newly exposed N-terminus [31].

Proline has a pronounced influence on the conformation of the peptide backbone as well. For the cathelicidins with only 2 or 3 proline residues, this could cause a kink in the secondary structure. For cathelicidins with higher proline content, secondary structure studies show that they adopt a backbone conformation like polyproline type II [32,33]. Such a rigid conformation might not bind well within the active site of proteases, also contributing to the stability of these proline/arginine-rich peptides.

3.2. Bromotryptophan in hagfish cathelicidins

Recently, cathelicidins have been discovered in the Atlantic hagfish (*Myxine glutinosa*), which expresses potent, broad-spectrum hagfish intestinal antimicrobial peptides (HFIAPs) [19,34]. HFIAP-1, -2, and -3 are carboxyamidated, but have no proline (Fig. 3). All three of these cationic peptides contain bromotryptophan residues, apparently as a result of post-translational modification [35]. The biological role of bromotryptophan in peptides is generally unknown. Bromination does not alter the antimicrobial activity of HFIAPs. Minimum inhibitory concentration (MIC) values for synthetic peptides prepared with both native tryptophan and 5-bromotryptophan are quite comparable ([34]; Shinnar et al., manuscript in preparation).

Bromination of tryptophan might also make HFIAPs less susceptible to proteolytic degradation. The presence of bromine could transform the tryptophanyl sidechain into a poorer substrate, with sub-optimal binding for the specificity pocket. In a preliminary experiment in vitro, partially purified HFIAP-3 (containing two residues of monobrominated tryptophan) retained partial antimicrobial activity even after treatment with proteinase K, a nonspecific protease; in contrast, a synthetic control peptide (PGLa from frogs) rapidly and completely became inactive in this qualitative assay (Shinnar, unpublished observations). Protection against proteases present in the intestinal submucosa of hagfish might extend effectively the pharmacokinetic lifetime of these peptides and increase their efficacy in vivo. Since biosynthesis of HFIAPs occurs in blood-forming cells [19], it is also possible that these peptides are playing a role not just in local defense, but in systemic circulation as

HFIAP		# of AA	Mass (amu)
1	GFFKKA(Br)WRKVKHAGRRVLDTAKGVGRHYVNN(Br)WLNRYR	37	4643.3
2	GFFKKA(Br)WRKVKHAGRRVLDTAKGVGRHYVNNWLNRYR	37	4564.0
3	(Br)G(Br)WFKKAWRKVKNAGRRVL...KGVGIHYGVGLI	30	3551.9

Fig. 3. Chemical sequences of hagfish intestinal antimicrobial peptides (HFIAPs). Primary structures of HFIAPs were determined using automated Edman sequencing and electrospray mass spectrometry. Additional mass spectral studies of both native and synthetic peptides confirmed identity of bromotryptophan (Br-W). HFIAP-1 is the major species isolated, while HFIAP-2 and HFIAP-3 are minor. Alignment using cDNA sequences of cloned genes shows that HFIAP-3 is a deletion mutant (indicated by the dots). In the isolated peptides, the C-termini are amidated (based on mass analysis) and the cDNA sequence shows codons for glycine prior to the stop codons [19,34]. Although antimicrobial peptides show great sequence diversity, primary structure trends of alternating hydrophobic and polar residues have been observed statistically, particularly in peptides that adopt α-helical secondary structures [4,49]. All three HFIAPs show these general trends in the N_t domain, including aromatic residues at positions 2 and 7 in the sequence (Shinnar et al., manuscript in preparation); Br-W in these positions might be poor substrates for serine proteases with specificity pockets suitable for aromatic sidechains. The exact consensus sequence recognized by the enzyme involved in brominating tryptophan in HFIAPs has not yet been established.

well. Cathepsin L, a member of the ubiquitous cysteine protease family [36], has been identified in the intestinal tissues of Atlantic hagfish (GenBank Accession No. AAF19630). In mammals, cathepsin L serves as a housekeeping enzyme and shows a preference for aromatic residues in the P_2 position (Fig. 2). It seems likely that this endoprotease in hagfish could degrade the N_t domain of HFIAPs, where aromatic residues are clustered.

To test our working hypothesis that bromotryptophan confers protease resistance to the HFIAP family, our laboratory has first undertaken a molecular modeling study [37]. The sidechains of native tryptophan and regioisomers of bromotryptophan have been docked into the specificity pocket of chymotrypsin, using the crystallographic coordinates for this well-studied serine protease. Steric clashes, measured as overlap of van der Waals radii between the bromine and seven amino acids lining this pocket, are observed for all possible bromotryptophan regioisomers (2-, 4-, 5-, 6-, and 7-Br-trp), whereas native tryptophan produced few unfavorable interactions. The increase in molecular volume due to bromination of the indole ring is significant, making all isomers of bromotryptophan a poor fit for the active site of chymotrypsin. As a complementary experimental approach, our laboratory is undertaking the synthesis of *N*-α-protected *p*-nitroanilides of tryptophan and bromotryptophan. The ability of chymotrypsin to hydrolyze these amides can be studied spectrophotometrically via the release of *p*-nitroaniline. Enzyme kinetic studies under way in our laboratory should reveal how tightly (K_M) and efficiently (k_{cat}) the brominated tryptophanyl amides perform as a substrate for chymotrypsin in comparison with natural tryptophan [38].

Although many hundreds of marine natural products have halogenated indole rings [39], relatively few are peptidic and, of these, only several contain bromotryptophan in a linear peptide [40–43]. Styelin D is another gene-encoded antimicrobial peptide, recently isolated from tunicates. In addition to 6-bromotryptophan, styelin's primary structure includes dihydroxyarginine and dihydroxylysine. Synthetic styelin, prepared with unmodified amino acid residues, is less potent than the isolated peptide. The native modified residues may enhance biological activity by making the peptides less susceptible to common proteases. Carnivorous marine cone snails (*Conus* spp.) produce potent venom with neurotoxic peptides that are punctuated by unusual amino acids, including 6-bromotryptophan, D-tryptophan, pyroglutamic acid, γ-carboxyglutamic acid, and hydroxyproline. These various post-translational modifications could make these *Conus* peptides less susceptible to the proteases found in their prey because of their poor fit into specificity pockets of common proteases. Here, too, we propose that bromination of tryptophan thereby extends the pharmacokinetic lifetimes of these bioactive peptides by making them poorer substrates for endogenous proteolytic enzymes.

3.3. Carboxyamidation

In antimicrobial peptides, the presence of a C_t amide instead of a free acid increases the net cationic charge, which plays a fundamental role in the electrostatic

attraction to target membranes [4,5,44]. C-terminal amidation, achieved through the post-translational action of PAM, is a general feature that might also provide resistance to C-terminal exopeptidases [4,30]. Carboxypeptidases, as exemplified by carboxypeptidase A (CPA), recognize peptide substrates with free C_t-carboxylates, mooring them via a guanidinium salt link and hydrogen bonds (Arg 145, Asn 144, and Tyr 248 in CPA) [45]. Carboxyamidation of antimicrobial peptides would remove some of these specific anchors.

3.4. Pharmaceutical applications

There is a general need for new classes of anti-infective agents, especially with the serious increase in pathogens resistant to conventional antibiotic treatments [20]. Several major drawbacks of oligopeptides as possible drugs are their proteolytic susceptibility, limited bioavailability, and possible immunigenicity, in addition to cost. Despite these challenges, some pharmaceutical companies have pioneered the development of antimicrobial peptides as drugs, using the structure activity relationship (SAR) approach for optimization [44,46]. IntraBiotics Pharmaceuticals has progressed to Phase II/III clinical trials with a synthetic peptide analog of protegrin (Protegrin IB-367 or iseganan hydrochloride; www.intrabiotics.com) for the treatment of oral mucositis [46]. Iseganan administered within the oral cavity avoids the usual problems of digestive proteases from stomach and intestine as well as intestinal absorption. For the future, design of peptidomimetic structures may offer promising alternatives to overcome the inherent limitations of peptide drugs [47].

As reviewed here, the primary structure of many cathelicidins suggests that this family of antimicrobial peptides has evolved features that confer some natural protection against certain proteases. The presence of proline along with some post-translational modifications (halogenation and carboxyamidation) may thwart rampant proteolytic degradation just enough to allow these bioactive peptides to attack microbes successfully.

4. Summary

Antimicrobial peptides from the cathelicidin gene family must be released from their inactive precursor form through several proteolytic steps. Once liberated, these peptides can serve their antimicrobial function as long as they are not degraded by ubiquitous proteases. Among the mammalian cathelicidins, natural resistance to proteolytic degradation could be afforded by sequences rich in proline, a cyclic residue known to block the mechanism of serine proteases. In hagfish cathelicidins, bromination of tryptophan may serve a similar biological role, by making these modified residues poorer substrates for common proteases. Although the mature cathelicidins exhibit very diverse primary structures, this family of antimicrobial peptides may have evolved structural features conferring resistance to proteolytic degradation in order to sustain their antimicrobial activity.

Acknowledgments

We thank Professor Ann McDermott of the Chemistry Department at Columbia University for guidance on the molecular modeling studies of bromotryptophan and critical reading of the manuscript. Dr. Thomas Uzzell of the Philadelphia Academy of Natural Sciences provided insight for aligning sequences of mammalian cathelicidins. For technical assistance with the manuscript, we thank Dr. Toby Berger Holtz, SuQing Liu, and Laura Shifley.

References

[1] M. Zasloff, Nature 415 (2002) 389–395.
[2] D. Yang, O. Chertov, J.J. Oppenheim, Cell. Mol. Life Sci. 58 (2001) 978–989.
[3] T. Ganz, R.I. Lehrer, in: Development of Novel Antimicrobial Agents: Emerging Strategies, Horizon Scientifc Press, Wymondham, UK, 2001, pp. 139–147.
[4] A. Tossi, L. Sandri, A. Giangaspero, Biopolymers (Peptide Science) 55 (2000) 4–30.
[5] Y. Shai, Biopolymers (Peptide Science) 66 (2002) 236–248.
[6] R. Gennaro, M. Zanetti, Biopolymers (Peptide Science) 55 (2000) 31–49.
[7] M. Zanetti, R. Gennaro, D. Romeo, FEBS Lett. 374 (1995) 1–5.
[8] R.I. Lehrer, T. Ganz, Curr. Opin. Hematology 9 (2002) 18–22.
[9] B. Ramanathan, E.G. Davis, C.R. Ross, F. Blecha, Microb. Infect. (2002) 361–372.
[10] M. Zaiou, R.L. Gallo, J. Mol. Med. 80 (2002) 549–561.
[11] H. Nielsen, J. Engelbrecht, S. Brunak, G. von Heijne, Prot. Eng. 10 (1997) 1–6.
[12] M. Paetzel, A. Karla, N.C.J. Strynadka, R.E. Dalbey, Chem. Rev. 102 (2002) 4549–4579.
[13] D.F. Steiner, Curr. Opin. Chem. Biol. 2 (1998) 31–39.
[14] N.C. Rockwell, D.J. Krysan, T. Komiyama, R.S. Fuller, Chem. Rev. 102 (2002) 4525–4548.
[15] S.T. Prigge, A.S. Kolhekar, B.A. Eipper, R.E. Mains, L.M. Amzel, Science 278 (1997) 1300–1305.
[16] J.F. Sanchez, F. Hoh, M.P. Strub, A. Aumelas, C. Dumas, Structure 10 (2002) 1363–1370.
[17] J.F. Sanchez, F. Wojcik, Y.S. Yang, M.P. Strub, J.M. Strub, A. van Dorsselaier, M. Martin, R. Lehrer, T. Ganz, A. Chavanieu, B. Calas, A. Aumelas, Biochemistry 41 (2002) 21–30.
[18] O.E. Sørensen, P. Follin, A.H. Johnsen, J. Calafat, G.S. Tjabringa, P.S. Hiemstra, N. Borregaard, Blood 97 (2001) 3951–3959.
[19] T. Uzzell, E.D. Stolzenberg, A.E. Shinnar, M. Zasloff. Hagfish cathelin-associated antimicrobial peptides and genes, Patent Application, PCT # WO 03012044, pub. February 13, 2003, manuscript submitted for publication.
[20] M. Zanetti, R. Gennaro, Skerlavaj, L. Tomasinsig, R. Circo, Curr. Pharm. Design 8 (2002) 779–793.
[21] R. Gennaro, M. Zanetti, M. Benincasa, E. Podda, M. Miani, Curr. Pharm. Design 8 (2002) 763–778.
[22] R.F. Doolittle, in: Prediction of Protein Structure and the Principles of Protein Conformation, Plenum Press, New York, 1989, pp. 599–623.
[23] I. Schechter, A. Berger, Biochem. Biophys. Res. Commun. 27 (1967) 157–162.
[24] L. Hedstrom, Chem. Rev. 102 (2002) 4501–4523.
[25] H.U. Jörnvall, Pettersson, L. Philipson, Eur. J. Biochem. 48 (1974) 179–192.
[26] V. Schellenberger, C.W. Turck, W.J. Rutter, Biochemistry 33 (1994) 4251–4257.
[27] R.C. Thompson, E.R. Blout, Biochemistry 12 (1973) 51–57.
[28] J. Polanowska, I. Krokoszynska, H. Czapinska, W. Watorek, M. Dadlex, J. Otlewski, Biochim. Biophys. Acta 1386 (1998) 189–198.
[29] H.G. Boman, B. Agerberth, A. Boman, Infect. Immun. 61 (1993) 2978–2984.
[30] A. Yaron, Biopolymers 26 (1987) S215–222.
[31] H. Jörnvall, B. Persson, Biosci. Rep. 3 (1983) 225–232.
[32] V. Cabiaux, B. Agerberth, J. Johansson, F. Homble, E. Goormaghtigh, J. Ruysschert, Eur. Biochem. 224 (1994) 1019–1027.

[33] T. Niidome, H. Mihara, M. Oka, T. Hayshi, T. Saiki, K. Yoshida, H. Aoyagi, J. Peptide Res. 51 (1998) 337–345.
[34] A.E. Shinnar, T. Uzzell, M.N. Rao, E. Spooner, W.S. Lane, M.A. Zasloff, in: P. Kaumaya, R. Hodges (Eds.), Peptides: Chemistry and Biology, Mayflower Scientific, Leiden, 1996, pp. 189–191.
[35] A.E. Shinnar, S. Fu, J. Bautista, FASEB 14 (2000) A1488.
[36] M.E. McGrath, Annu. Rev. Biophys. Biomol. Struct. 28 (1999) 181–204.
[37] K. Curran, Bromotryptophan: biochemical and biophysical properties of an unusual amino acid in antimicrobial peptides isolated from hagfish, Senior honors thesis, Department of Chemistry, Barnard College, New York, NY, 2002.
[38] H.J. Park, Biological role and physical properties of bromotryptophan in hagfish intestinal antimicrobial peptides, Senior honors thesis, Department of Chemistry, Barnard College, New York, NY, 2003.
[39] G.W. Gribble, Accounts Chem. Res. 31 (1998) 141–152.
[40] A.G. Craig, E.C. Jimenez, J. Dykert, D.B. Nielson, J. Gulyas, F.C. Abogadie, J. Porter, J.E. Ribier, L.J. Cruz, B.M. Olivera, J.M. McIntosh, J. Biol. Chem. 272 (1997) 4689–4698.
[41] E.C. Jimenez, A.G. Craig, M. Watkins, D.R. Hillyard, W.R. Gray, J. Gulyas, J.E. Rivier, L.J. Cruz, B.M. Olivera, Biochemistry 36 (1997) 989–994.
[42] S.W. Taylor, B. Kammerer, G.J. Nicholson, K. Pusecker, T. Walker, E. Bayer, S. Scippa, M. Vincentis, Arch. Biochem. Biophys. 348 (1997) 278–288.
[43] S.W. Taylor, A.G. Craig, W.H. Fischer, M. Park, R.I. Lehrer, J. Biol. Chem. 275 (2000) 38417–38426.
[44] W.L. Maloy, U.P. Kari, Biopolymers (Peptide Science) 37 (1995) 105–122.
[45] D.W. Christianson, W.N. Lipscomb, Accounts Chem. Res. 22 (1989) 62–69.
[46] J. Chen, T.J. Falla, M.A. Hurst, C.A. Fujii, D.A. Mosca, J.R. Embree, D.J. Loury, P.A. Radel, C.C. Chang, L. Gu, J.C. Fiddes, Biopolymers (Peptide Science) 55 (2000) 88–98.
[47] J.A. Patch, A.E. Barron, Curr. Opin. Chem. Biol. 6 (2002) 872–877.
[48] S.S.L. Harwig, V.N. Kokryakov, K.M. Swiderek, G.M. Aleshina, C. Zhao, R.I. Lehrer, FEBS Lett. 362 (1995) 65–69.
[49] G. Basañez, A.E. Shinnar, J. Zimmerberg, FEBS Lett. 532 (2002) 115–120.