# Structural organization of the bovine cathelicidin gene family and identification of a novel member

Marco Scocchi<sup>a,b</sup>, Shenglun Wang<sup>1,a</sup>, Margherita Zanetti<sup>a,b,\*</sup>

<sup>a</sup>Laboratorio Nazionale Consorzio Interuniversitario Biotecnologie, AREA Science Park, Padriciano 99, I-34012 Trieste, Italy <sup>b</sup>Dipartimento di Scienze e Tecnologie Biomediche, Università di Udine, I-33100 Udine, Italy

Received 2 October 1997

Abstract Cathelicidins are a group of myeloid antimicrobial peptide precursors found in a variety of mammalian species. Transcripts of this family show a highly conserved 5' region corresponding to the 5' untranslated region, signal peptide and propiece, and diverse 3' regions encoding structurally varied Cterminal sequences that correspond to mature antimicrobial peptides after proteolytic release from the precursors. To establish the size of the bovine gene family encoding these proteins,  $\lambda$  genomic clones were isolated by screening a bovine library with a probe based on the conserved cDNA region of bovine members. Restriction mapping, hybridization studies and sequence analysis identified 11 distinct genes that based on the intergenic distances of contiguous genes appear to be in close physical proximity. Among these, a novel gene encoding the precursor of a putative  $\alpha$ -helical antimicrobial peptide was recognized and sequenced. The novel gene appears to be expressed in bovine bone marrow myeloid cells, spleen and testis. © 1997 Federation of European Biochemical Societies.

S 1997 Federation of Baropean Bioenemean Societies

Key words. Cathelicidin; Gene; Antimicrobial peptide

#### 1. Introduction

A large number of cationic peptides are recognized to play a role in animals as a first line of defense against pathogens [1-3]. These peptides rapidly inactivate microbes by perturbing microbial membranes. In mammals, antimicrobial peptides are found in circulating phagocytes [4-6], where they contribute to the killing of engulfed microorganisms, and in epithelial surfaces, where they act as a local defense mechanism that protects anatomical compartments from microbial invasion [6-9]. Different mammalian species are equipped with different sets of peptides showing broad spectra of antimicrobial activity. The mammalian antimicrobial peptides fall within one of two major groups: the cysteine-rich  $\alpha$  and  $\beta$ defensins [10,11], and the heterogeneous group of the cathelicidin-derived peptides [12]. Peptides belonging to the latter group vary markedly in sequence and length, ranging from 12 to 79 amino acid residues, and include  $\alpha$ -helical peptides (e.g. human LL-37, rabbit CAP18), linear, proline- and arginine-rich (e.g. bovine Bac7, pig PR-39), or tryptophan-rich (bovine indolicidin) peptides, and peptides with one or two

\*Corresponding author. Fax: (39) (40) 398990. E-mail: zanetti@icgeb.trieste.it disulfide bonds (e.g. bovine dodecapeptide, pig protegrins). They are processed from precursors (prepropeptides) that are synthesized in myeloid bone marrow cells and have collectively been designated cathelicidins. Typically, precursors belonging to the cathelicidin family show an N-terminal preprosequence of approximately 130 residues which is highly conserved both intra- and inter-species (90–97% intraspecies amino acid identity in pig, 75–87% in cow), and structurally diverse C-terminal sequences corresponding to antimicrobial peptides. The propeptides are stored in the secretory granules of neutrophils, and the C-terminal antimicrobial peptides can be released from the precursors by enzymatic removal of the propiece at specific sites upon neutrophil stimulation [13–15].

Thus far, cathelicidins have only been found in mammals. More than 20 congeners with molecular masses of 16–26 kDa have been identified as deduced from myeloid cDNA in human, cow, pig, rabbit [12], sheep [16,17], mouse [18,19], and guinea pig [20], and each congener has been named after the antimicrobial domain. The number of different cathelicidins varies substantially among species. Only one is expressed in human, while several congeners have been identified in the artiodactyl species analyzed. Among these, the complete gene sequences of pig PR-39, protegrins and prophenin-2, and that of the human FALL-39/hCAP18 are known [21– 25]. All these genes show a similar structural organization with the coding sequence distributed over four exon regions.

In this study, we have defined the size of the cathelicidin gene family in cattle by screening a bovine genomic library with probes based on the conserved 5' cDNA sequence of bovine congeners. Eleven cathelicidin genes were identified. Among these is a novel gene encoding the precursor of an  $\alpha$ -helical antimicrobial peptide. This gene was found to be expressed in bone marrow myeloid cells, spleen and testis.

## 2. Materials and methods

2.1. Genomic Southern blotting

Genomic DNA was prepared from bovine blood cells as described [26]. Aliquots of 10 µg of DNA were digested with EcoRI, BamHI and HindIII restriction endonucleases (used singly or pairwise), and transferred after gel electrophoresis onto nylon filters (Genescreen plus, DuPont NEN Products, Boston, MA) that were hybridized overnight at 65°C in 0.5 M NaH<sub>2</sub>PO<sub>4</sub>, 1 mM EDTA, 5% SDS, pH 7.2, 100 µg/ml of denatured salmon sperm DNA. Filters were then washed at room temperature with 40 mM NaH<sub>2</sub>PO<sub>4</sub>, 5% SDS, 1 mM EDTA, pH 7.2 for 30 min, and at 65°C in 40 mM NaH<sub>2</sub>PO<sub>4</sub>, 0.5% SDS, 1 mM EDTA, pH 7.2 for 30-180 min (depending on the probe used). <sup>32</sup>P-labeled hybridization probes were prepared by random primer synthesis (Boehringer Mannheim, Germany). Hybridization probes included an EcoRI fragment of 396 bp from the 5' end of bovine Bac7 cDNA (Cathl probe), and PCR-generated fragments corresponding to the 3' ends of the cDNAs of bovine Bac7, Bac5, dodecapeptide, BMAP-27, BMAP-28, indolicidin, and a DNA restriction fragment of BMAP-34.

<sup>&</sup>lt;sup>1</sup>Present address: Cardiovascular Research Institute, New York Medical College, Valhalla, NY 10595, USA.

The nucleotide sequence data reported in this paper are deposited in the GenBank/EBI Data Bank under accession numbers Y12728 and Y12729.

# 2.2. Screening of genomic library and isolation of genomic clones

A bovine liver genomic library cloned into  $\lambda$  DASH II phage (Stratagene cloning system, Inc., La Jolla, CA) was screened by hybridization with the radiolabeled 5' cDNA of bovine Bac7 (Cathl probe) [26]. Phage DNA from the positive plaques was purified as described [27], digested with the restriction enzymes *NotI*, *Ban*HI, *EcoRI*, *SpeI* and *Hin*dIII, and analyzed by Southern blotting using 5'-<sup>32</sup>P-end-labeled oligonucleotide probes. The following probes were used: 5'-CAAGAATTCGGAGACTGGGGACCATG-3' and 5'-ACCGAATTCAGTGTGACTTCAAGGA-3', from the conserved 5' untranslated region and proregion of cathelicidin cDNA sequences; the oligonucleotides 5'-AGTGCTAACCTTGATGTT-3', 5'-CAGATTTCGTC-CACCAA-3', 5'-CAGATCCAGTAGCTTGAGGCC-3', 5'-TCTGAA-CAAATCAGACACTTA-3', 5'-AATTGGGCCATACTTCTCC-3', 5'-ACAGGATTCTTCCATGGGCT-3', based on the 3' cDNA ends of bovine Bac7, Bac5, dodecapeptide, indolicidin, BMAP-28 and BMAP-27, respectively.

#### 2.3. Gene cloning and sequencing, and sequence analysis

The sequence of BMAP-34 gene was obtained from three restriction fragments of clone  $\lambda$ Cl10 including a 1.8 kbp SpeI/HindIII, a 3.1 kbp HindIII/BamHI and a 3.1 kbp BamHI/SpeI fragment. Genomic fragments were subcloned into pBluescript KS+ vector (Stratagene). Sequential unidirectional deletions of the subcloned fragment were created using an exonuclease III/S1 nuclease reagent kit (Pharmacia) according to the protocol of the supplier. Sequencing was performed on both strands using universal or sequence-specific primers, with deazaguanosine and automated fluorescent DNA sequencing (EMBL fluorescent DNA sequencer, Heidelberg, Germany) as previously described [28]. Nucleotide sequence assembly and analysis, and polypeptide secondary structure prediction were conducted with the software package of the Wisconsin Genetics Computer Group (GCG). Database searches were performed using the BLAST network service of the National Center for Biotechnology Information. Consensus sites were identified by the Signal Scan Program using David Ghosh transcriptional factor Database [29].

#### 2.4. Reverse transcription-PCR and Northern analysis

Total RNA extraction from bovine tissues and reverse transcription were performed as described [30,31]. PCR was performed in a 1605 Air Thermo Cycler (Idaho Technology, Idaho Falls, ID) for 30 cycles: 15 s at 94°C, 15 s at 52–56°C, 30 s at 72°C. The reaction was carried out in a total volume of 20  $\mu$ l of PCR buffer (50 mM Tris-HCl, pH 8.3, 0.25 mg/ml BSA and 2 mM MgCl<sub>2</sub>) containing 100  $\mu$ M dNTPs, 1  $\mu$ l of a 1/5 dilution of cDNA, 0.5 units of TaqI polymerase (Perkin Elmer-Roche Mol. Syst., Branchburg, NJ) and 10 pmol each of a sense primer 5'-ACCGAATTCAGTGTGACTTCAAGGA-3', based on a sequence contained in exon 2 of BMAP-34 gene, and an antisense primer 5'-ACCTGATCCTAAGGACTTT-3' complementary to an exon 4 sequence. Amplified products were size fractionated in a 2% agarose gel, visualized by ethidium bromide staining, blotted and hybridized with a BMAP-34-specific probe. Northern analysis was performed as described [31], by using the labeled oligonucleotide probe 5'-ACCTGATCCTAAGGACTTT-3' complementary to the 3' non-coding region of BMAP-34 gene.

#### 3. Results and discussion

## 3.1. Characterization of the bovine cathelicidin gene family

On the basis of molecular cloning and sequence analysis of cDNAs, we previously reported the mRNA sequences of myeloid cathelicidins in several mammalian species. These were identified through a RT-PCR-based approach that exploits the high conservation of the 5' mRNA region including the 5' untranslated and the signal peptide and propiece coding regions [12,16].

As an initial step toward the characterization of this gene family in cattle, a Southern analysis of bovine genomic DNA digested with various restriction endonucleases was performed. The presence of genomic sequences related to cathelicidins was investigated by hybridizing blots with a 400 bp EcoRI fragment, referred to here as Cathl probe, that covers the conserved 5' region of preproBac7 cDNA [32]. The band pattern obtained indicates the presence of several cathelicidinrelated sequences (Fig. 1A). Known members of the family were then recognized through their diverse 3' regions, by hybridizing blots with the 3' cDNA sequence of each bovine



Fig. 1. A: Southern blot analysis of bovine genomic DNA digested with *Eco*RI (lane 1), *Bam*HI (lane 2), *Eco*RI/*Bam*HI (lane 3), and *Hin*dIII (lane 4). Blots were hybridized with the Cathl probe. B: Southern analysis of eight *Bam*HI/*Not*I digested genomic clones. An oligonucleotide probe from the conserved proregion of cathelicidin cDNA sequences was used. C: Schematic representation of the bovine cathelicidin genes. The restriction maps of the  $\lambda$  DNA clones and the contigs formed by partial overlapping of some of the fragments are shown. Genes (dark gray) and pseudogenes (light gray) are represented by boxes indicating the direction of transcription.

A

Þ

NcoI	SpeI	SmaI		HindIII EcoRV
	exon 1	exon 2	exon 3	exon 4

~																															
aat cct	cct tgc	.ctt aca	.gga	caa	gee ete	tgc cct	etg acc	rtga tac	icad tac	etto	ccca aqea	atgg	jaco cca	caca itga	igca lagc	iggg aac	ccc	aag	jaca	ttga:	tca ggc	ggg	gaga gtco	atg cac	ga tt	gaa cac	aco aco	cat cct	ttel	ttca agga	90 180
aqq	cac	iaas	itad	aaa	caa	atc	aac	raac	ragt	cod	Laat	taa	acc	stt	cada	itca	aac	icc.	cago	icto	rago	ata	aaa	aqq	aq	qqc	cc	ctc	agg	stga	270
aaa	αac	race	att	ta	aaa	acc	ATG	GAC	ACC	CAC	AGC	GCC	AGC	TTC	TCC	CTG	GGA	CGC	TCA	TCA	CTG	TG	<b>C</b> T2	ACT	GC	ĨĞĊ	TG	GGA	CTA	STGG	360
5-5	0				555		М	Е	т	0	R	A	S	F	s	L	G	R	S	s	L	W	L	L		L :	L	G	L	v	
										~																					
TGC	сст	CGG	GCCZ	AGC	GCC	CAG	GAC	CTC	CAGC	TAC	CAGO	GAC	GCC	GTC	<b>CT</b> T	CGT	GCT	GT	GAT	CAG	TTC	AA	rga	GAG	GT	CCT	CA	GAA	GCT	AATC	450
V	Ρ	s	А	S	А	Q	D	$\mathbf{L}$	S	Y	R	Е	А	V	L	R	А	V	D	Q	F	N	Е	R		S	S	Е	А	N	
mam	200	000	mov			000	010		iaan	000			0.20							-		~~~			~+	-+	a+.	~~~	- t- a	-	540
TCT	ACC	GCC	TCC	JUG T	JAG	UTA T	GAC	CCG	SCC1		-GAC	CAG	GA'I	gtg	jagt	tgg	gga	ggg	lâãc	tgg	gaa	ggg	<u>iga</u>	LCL	gu	cuc	сц	Jaç	auco	ceg	540
	ĩ	R	Ц	Ц	E	Ц	D	P	P	P	Е	Q	D																		
gca	cto	rttg	icco	cta	ttc	att	aag	get	ggt	tct	ctt	gto	ago	jaac	igca	ctt	ttc	CCa	icta	.cgt	ggg	itte	coce	acc	tc	ttc	tag	gga	aaco	ctcc	630
cag	acc	tgg	gto	cate	с																										645
aca	ggg	raga	icaç	jaco	cag	aga	ago	aaa	icat	aad	acco	rage	cca	igto	ccc	cca	.ctt	tga	atco	ttg	acc	ag	GTG	GAG	CA	ccc	GG	GAG	CTC	GAAA	90
						Č	55			-													V	Е	Н	Ρ	(	з.	A I	ĸк	
GCC	TGI	GAC	CTI	['CA	CAG	TGA	AGC	AGA	CCC	GIGI	rgcc	CCA	GGZ	ACAA	ICCC	CGC	AGC	CCC	CAG	AGC	AGT	GTC	GAC'	<b>FT</b> C	AA	GA	GA	ATG	GGg1	tgag	180
Р	Λ	1 5	5 E	7 S	г	V	ĸ	Е	T	V	С	Ρ	R	Т	т	Ρ	Q	Ρ	Ρ	Ε	Q	С	D	F	K	Ε	1	N (	3		
cat	aac	aac	ta	an	nan	cad	~ = =	aa+	aat	tat	-020		act		naa	aaa	act	+ _+	- 000	aaa	naa	++/	rcar	nte	cc	Faa	aa.	Fαa	aata	raaa	270
tat	aac	tac	idad	rata	att	acc	tac	aat	tto	tac	rttt	aac	tto	aac	icct	tct	ttt	CCE	act	GGT	GAA	ACA	AGT	GTG	ΤG	GGG.	AC)	AGT	CAC	CCGG	360
- 3 -	55-	- 3 3				J	-95								,				Ľ	ν	K	: ;	2 (	2	v	G	т	V	Т	R	
TAC	TGG	ATC	CAGO	GGG	TGA'	TTT	CGA	CAI	CAC	CTTC	GTAR	TAP	∖Tgt	aag	ıtgg	laca	ctt	ttç	ytac	tgt	gag	igad	ctg	cta	ga	ggg	gt	aaa	ttg	tgga	450
Y	W	Ι	R	G	D	F	Γ	) ]	1	r c	C N	1 N	I																		
	tar				t	~~~~	~ ~ ~					~~~~				+	a+ 0			+ ~~	aat	ate		-00	~~	~~~	aa		aaa	tata	540
acc	tac	tac	.ggc	suce stai	tat	yac ccc	CC8	IC LG	JCCC	:cai		uggg	icaç act	Jayo	adag	igitt iccc	euc aac	tar	1000	. ugg	aca	1000	ner	rca	cc. ra	tya ⊦+a	ta:	aac	ccc	acto	630
ana	tet	cat	raa	atc	aga	ctc ctc	cto etc	rgag	aad	raac	rcar	rect	tta	rtat	-+++	cac	aaa	ace	rtac	add	aca	tte	-+++	fca	at	act	αa.	aat	t.t.a	aaac	720
tca	ttr	act	caa	ati	cat	nao	ete	tca	act	1995	aat		ret t	CCC	aat	tta	it to	aas	atac	act	tat	aa	cct	taa	cra	aac	cc	cct	atca	acct	810
aad	ade	octi	tac	rttr	tac	g∝g tta	tet	ato	tat	aat	ato	itto	att	taa	icca	icad	act	CC5	aac	atc	ata	tad	raa	aat	aa	act	aa	aac	tcca	aaqc	900
ttc	cto	ato	rtgo	dcc.	cag	gaa	tcc	rgac	rtqt	tca	aagt	ggc	aac	iggo	atct	ggt	ttt	aat	get	gag	cca	ata	acco	cca	aa	gca	ac	tat	tte	ctcc	990
tgg	tto	aca	ıgAʻl	PTC.	AGA	GTG	CGC	GGC	TAT:	PTCC	CGCC	GGC	CTGA	AGGC	GACT	CAA	TCC	'GG/	AGAC	GTC	AGC	AG/	AAA	ATC	CT	CGA	AA	AAG	CAA	GGAG	1080
			-	E (	Q ·	S	A	G	L	F	R	R	<u>ь</u>	R	Ď	Ş	I	<u>R</u>	<u>R</u>	G	Q	0	K	Į	_L	E	-	K .	<u>A</u>	<u>R R</u>	
			~				1	-	ma		~~ * *						_						. An in i		~	o at	<b>a</b> t.	~			1170
AAT	160	-1GF	GAC	JAA.	TCA	AAG	ATF	TCI	10	1661	JGAI	.'AA∂	igeo	CCU	agg	acc	agg	L L C	auco	rgg	CĻC	aga	aut	LCL	.ya	act	οų	yaa		Laaa	11/0
Ī	C	3 F	E F	2	I	ĸ	<u>D</u>	I	F	R	G	*																			
						- <i>n</i> -															~ ~			<b>.</b>	+	t	at	~ ~ ~ ~	ant	aata	1260
tta	c CC	uaa at-	idda Hat	acai - ati	act	tda ±α⊧	50U.	laca		caaa		act		LUCI Yaar	-ucc	:LCC	.aag	+++	LLac	cag	gac	tag	yyti rati	vec var	.ud .ac	adu add	ςυ	yag	ayu	LUUU	1338
	uyi	-9.00	درياء	-y u	yuy	-yu	Jan	Jays	ryay	ومحاد	nyal	Jayo	مردران	juyo		$\sim \sim \sim$	- u u a		مىدى ل		-v y a	91	-9	ეოყ	9~	~23					

Fig. 2. A: Schematic representation of the BMAP-34 gene. Exons are symbolized by boxes: light gray, dark gray and black denote the coding regions for the signal peptide, the proregion and the mature peptide. The 5' and 3' untranslated regions are white. Unique restriction sites present in the gene are indicated. The map is drawn to scale. B: Sequence of the BMAP-34 gene. The 5' flanking, the exon 1, and the 5' regions of intron 1 are included in the nucleotide sequence 1-645, while the sequence 1-1338 includes the remaining 3' region of the gene, starting from the 3' region of intron 1. Non-coding sequences are in lowercase letters, the coding sequence is in uppercase letters. The deduced amino acid sequence of the open reading frame is indicated in single letter code. The sequence putatively corresponding to the antimicrobial peptide is underlined, the stop codon is indicated by an asterisk, the TATA-like box is in boldface, the polyadenylation signal is double underlined.

congener, i.e. Bac7 [32], Bac5 [28], indolicidin [31], dodecapeptide [33], BMAP-27 and BMAP-28 [34] (not shown). The results of this analysis are consistent with the presence in the bovine genome of single copy genes for Bac7, Bac5, BMAP-27 and BMAP-28, and of at least two copy genes for dodecapeptide and indolicidin (not shown). In an attempt to define the size of the bovine cathelicidin gene family, and reveal the presence of novel members of this family, approximately  $1 \times 10^6$  clones of a  $\lambda$  DASH II bovine genomic library were then screened with the Cathl probe. DNA from 11 hybridization-positive clones was purified. Each clone was double digested with *Not*I and *Bam*HI and

Table 1

Putative regulatory elements found in the 5' flanking region of the BMAP-34 gene

Factor or site name	Consensus sequence	Position (orientation)
NF-IL6	T G/T N N G N A A G/T	-127 (+) -196 (-)
NF-κB γIFN RE	G G G A/G A/G N C/T C C C C A/T G/T G/T A N N C/T	$ \begin{array}{c} -168(-) \\ -71(+) \\ -140(-) \\ -285(+) \\ \end{array} $

Numbering gives the position of the first nucleotide with regard to the transcription start site. + and - indicate positive and negative DNA strands, respectively.



Fig. 3. A: Southern analysis of bovine genomic DNA, digested with *Eco*RI (lane 1), *Bam*HI (lane 2), *Eco*RI/*Bam*HI (lane 3), and *Hin*dIII (lane 4) and hybridized with a DNA fragment corresponding to the exon 4 region of BMAP-34. B: RT-PCR analysis of tissue expression of BMAP-34. A product of about 250 bp was amplified from the indicated bovine tissues using BMAP-34-specific primers (upper panel) and analyzed by Southern blot analysis using a BMAP-34-specific probe (lower panel).

shown by Southern analysis to contain one or two cathelicidin-related genes (shown in Fig. 1B for eight different clones). These clones were mapped, and approximate positions of individual genes were obtained by Southern blotting using sequence-specific oligonucleotides derived from the 3' cDNAs of the bovine congeners (Fig. 1C). The gene copy number detected for each known congener was consistent with the copy number indicated by Southern analysis of genomic DNA. Three additional Cathl probe-positive sequences failed to give any detectable hybridization signal with the 3' cDNA probes used, and were cloned and sequenced. Two of these, namely Bac4 (GenBank/EBI Data Bank accession number Y12730) and CATHLP, were presumed to be pseudogenes by sequence analysis, whereas a sequence present in clone  $\lambda$ Cl10 was found to correspond to a novel functional gene and is described here in more detail. Bac7 and one copy gene of dodecapeptide, corresponding to Dod I in Fig. 1C, were also sequenced (GenBank/EBI Data Bank accession numbers Y09471 and Y09472), and shown to share the typical four exon, three intron structure of cathelicidins [21-25].

3.2. Gene structure and expression of the novel BMAP-34 gene The genomic clone  $\lambda$ Cl10 was shown to contain a cathelicidin-related sequence that did not hybridize to any of the bovine cathelicidin 3' cDNAs, and was further analyzed for the presence of a potential novel antimicrobial domain. The region of interest was sequenced as described in Section 2, revealing a novel member of this family. The gene was named BMAP-34 (bovine myeloid antimicrobial peptide of 34 residues) by analogy with two other reported members of this family [34], and its sequence is reported in Fig. 2B. BMAP-34 is located approximately 2 kbp 3' to the Dod II gene (Fig. 1C), spans about 1900 bp, and reproduces the four exon/three intron organization of cathelicidin genes (Fig. 2A). The sequence indicates that BMAP-34 is a functional gene with a TATA-like box. Several potential recognition sites for nuclear factors involved in the transcription of immune-related molecules are found in the 5' flanking region (Table 1), and a polyadenylation signal is located 52 bp from the putative stop codon. A Southern analysis of bovine genomic DNA using a PCR-generated fragment corresponding to the exon 4 region indicates that BMAP-34 is a single copy gene (Fig. 3A). Northern analysis of total RNA using a sequence-specific oligonucleotide reveals a corresponding transcript of approximately 0.7 kbp in bovine bone marrow cells (not shown). A

more widespread expression of this gene however is suggested by RT-PCR analysis of various bovine tissues, as indicated by the amplification of a BMAP-34 mRNA segment of the expected size also in spleen and testis but not in resting peripheral neutrophils, trachea, small intestine, liver, lymph nodes, Peyer's patches (Fig. 3B), kidney and heart (not shown).

The open reading frame (Fig. 2B) predicts a polypeptide of 165 amino acid residues, with a calculated mass of 18848 Da and a pI of 9.6. The coding sequence for the signal peptide and for most of the propiece is contained in exons 1-3. The remaining four residues of the propiece and a novel C-terminal sequence of 34 residues (underlined in Fig. 2B) are encoded in exon 4. The putative 34 residue antimicrobial domain is preceded by an alanyl residue. This appears to be a processing site for the maturation of several cathelicidin-derived antimicrobial peptides [12]. The novel sequence includes 12 positively and four negatively charged residues. A C-terminal glycine suggests that the corresponding mature peptide may be amidated at the C-terminus [12]. As indicated by structure prediction analysis, a peptide with this sequence may adopt an amphipathic  $\alpha$ -helical conformation (Fig. 4) which is shared by several other antimicrobial peptides and



Fig. 4. Helical wheel projection of BMAP-34. The representation shows the amphipathic structure of the 34 C-terminal residues encoded in the exon 4 region of BMAP-34 gene. Charged residues are in boldface and strongly hydrophobic residues are boxed. G1 and G34 represent the amino- and carboxy-terminal glycine residues of the sequence.

 Table 2

 Percent sequence identity in bovine cathelicidin genes

Gene pair	5' Flanking	Exon 1	Intron 1	Exon 2	Intron 2	Exon 3	Intron 3	Exon 4				
	region							coding region	non-coding region			
Bac7/Dod I	80	88	66	80	71	71	69	< 40	57			
Bac7/BMAP-34	77	90	n.d.*	89	77	72	70	$<\!40$	70			
Dod I/BMAP-34	83	86	n.d.	84	72	83	74	< 40	72			

Nucleotide homologies of pairwise comparisons within related regions of the genes are reported. For the calculation of the percent homology, gaps were not taken into consideration. A 350 bp 5' flanking region upstream of the ATG initiation codon is compared. \*Not determined, as the intron 1 region of BMAP-34 was only partially sequenced.

is thought to facilitate membrane interaction and disruption [2]. Two other  $\alpha$ -helical cathelicidin-derived peptides, named BMAP-27 and BMAP-28, have been identified in cow [34]. However, their sequences share a low level of identity (less than 30%) with that of the putative BMAP-34 peptide, and in addition, the helical segment in both peptides is followed by a highly hydrophobic C-terminal tail [34] that is absent in BMAP-34. The sequence of BMAP-34 may thus be considered another example of the structural variety exhibited by peptide members of this family within an animal species. This is further pointed out in Table 2. Comparative analysis of the genes indicates that the degree of identity of related sequences in BMAP-34 and other bovine genes drops markedly just in the segment encoding the mature peptide, which is thus the primary target for diversification.

In conclusion, the results of this analysis indicate that (i) the bovine cathelicidin gene family consists of at least 11 members. Since this number is comparable with the cathelicidinpositive signals detected in the Southern analysis of genomic DNA, the genes identified likely comprise most of the cathelicidins present in the bovine genome; (ii) all these genes may be in close physical proximity, as suggested by the identification of three contigs formed by partially overlapping clones, each carrying two or three genes within 2-10 kbp of each other. A close proximity of these genes is also supported by physical mapping, by means of fluorescence in situ hybridization, suggesting clustering of clones  $\lambda$ Cl1,  $\lambda$ Cl2,  $\lambda$ Cl8,  $\lambda$ Cl10,  $\lambda$ Cl23,  $\lambda$ Cl34, at a CATHL<sup>®</sup> locus on bovine chromosome 22q24 [35]. Finally, another important result of this analysis is the identification of a novel member of this family that appears to be expressed not only in bone marrow cells, but also in other tissues and organs, and may thus represent an important agent of the local host response to microbial infections.

Acknowledgements: We thank Dr. R. Gennaro from the Dipartimento di Scienze e Tecnologie Biomediche, University of Udine, for critically reading the manuscript. This work was supported by grants from the National Research Council, and from the Italian Ministry for University and Research (MURST 60% and 40%).

#### References

- Boman, H.G., Marsh, J. and Goode, J.A. (1994) Antimicrobial Peptides. Ciba Foundation Symposium 186. John Wiley and Sons, Chichester.
- [2] Boman, H.G. (1995) Annu. Rev. Immunol. 13, 62-92.
- [3] Goldman, M.J., Anderson, G.M., Stolzenberg, E.D., Kari, U.P., Zasloff, M. and Wilson, J.M. (1997) Cell 88, 553–560.
- [4] Levy, O. (1996) Eur. J. Haematol. 56, 263-277.
- [5] Ganz, T. and Lehrer, R. (1997) Curr. Opin. Haematol. 4, 53-58.
- [6] Selsted, M.E. and Ouellette, A.J. (1995) Trends Cell Biol. 5, 114-
- 119.

- [7] Stolzenberg, E.D., Anderson, G.M., Ackermann, M.R., Whitlock, R.H. and Zasloff, M. (1997) Proc. Natl. Acad. Sci. USA 94, 8686–8690.
- [8] Frohm, M., Agerberth, B., Ahangari, G., Stahle-Backdahl, M., Liden, S., Wigzell, H. and Gudmundsson, G.H. (1997) J. Biol. Chem. 272, 15258–15263.
- [9] Harden, J., Bartels, J., Christophers, E. and Schroder, J.-M. (1997) Nature 387, 861.
- [10] Martin, E., Ganz, T. and Lehrer, R.I. (1995) J. Leukocyte Biol. 58, 128–136.
- [11] White, S.H., Wimley, W.C. and Selsted, M.E. (1995) Curr. Opin. Struct. Biol. 5, 521–527.
- [12] Zanetti, M., Gennaro, R. and Romeo, D. (1995) FEBS Lett. 374, 1–5.
- [13] Zanetti, M., Litteri, L., Gennaro, R., Horstmann, H. and Romeo, D. (1990) J. Cell Biol. 111, 1363–1371.
- [14] Scocchi, M., Skerlavaj, B., Romeo, D. and Gennaro, R. (1992) Eur. J. Biochem. 209, 589–595.
- [15] Panyutich, A., Shi, J., Boutz, P.L., Zhao, C. and Ganz, T. (1997) Infect. Immun. 65, 978–985.
- [16] Bagella, L., Scocchi, M. and Zanetti, M. (1995) FEBS Lett. 376, 225–228.
- [17] Mahoney, M.M., Lee, A.Y., Brezinski-Caliguri, D.J. and Huttner, K.M. (1995) FEBS Lett. 377, 519–522.
- [18] Popsueva, A.E., Zinovjeva, M.V., Visser, J.W., Zijlmans, J.M., Fibbe, W.E. and Belyavsky, A.V. (1996) FEBS Lett. 391, 5–8.
- [19] Gallo, R.L., Kim, K., Bernfield, M., Kozak, C., Zanetti, M., Merluzzi, L. and Gennaro, R. (1997) J. Biol. Chem. 272, 13088–13093.
- [20] Nagaoka, I., Tsutsumi-Ishii, Y., Yomogida, S. and Yamashita, T. (1997) J. Biol. Chem. 272, 22742–22750.
- [21] Gudmundsson, G.H., Magnusson, K.P., Chowdhary, B.P., Johansson, M., Andersson, L. and Boman, H.G. (1995) Proc. Natl. Acad. Sci. USA 92, 7085–7089.
- [22] Zhao, C., Ganz, T. and Lehrer, R.I. (1995) FEBS Lett. 368, 197– 202.
- [23] Zhao, C., Ganz, T. and Lehrer, R.I. (1995) FEBS Lett. 376, 130– 134.
- [24] Gudmundsson, G.H., Agerberth, B., Odeberg, J., Bergman, T., Olsson, B. and Salcedo, R. (1996) Eur. J. Biochem. 238, 325–332.
- [25] Larrick, J.W., Lee, J., Ma, S., Li, X., Francke, U., Wright, S.C. and Balint, R.F. (1996) FEBS Lett. 398, 74–80.
- [26] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [27] Del Sal, G., Manfioletti, G. and Schneider, C. (1989) BioTechniques 7, 514–519.
- [28] Zanetti, M., Del Sal, G., Storici, P., Schneider, C. and Romeo, D. (1993) J. Biol. Chem. 268, 522–526.
- [29] Prestridge, D.S. (1991) CABIOS 7, 203-206.
- [30] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156– 159.
- [31] Del Sal, G., Storici, P., Schneider, C., Romeo, D. and Zanetti, M. (1992) Biochem. Biophys. Res. Commun. 187, 467–472.
- [32] Scocchi, M., Romeo, D. and Zanetti, M. (1994) FEBS Lett. 352, 197–200.
- [33] Storici, P., Del Sal, G., Schneider, C. and Zanetti, M. (1992) FEBS Lett. 314, 187–190.
- [34] Skerlavaj, B., Gennaro, R., Bagella, L., Merluzzi, L., Risso, A. and Zanetti, M. (1996) J. Biol. Chem. 271, 28375–28381.
- [35] Castiglioni, B., Scocchi, M., Zanetti, M. and Ferretti, L. (1996) Cytogen. Cell Genet. 75, 240–242.