Paneth Cells of the Human Small Intestine Express an Antimicrobial Peptide Gene*

Douglas E. Jones§ and Charles L. Bevins††

From the ‡Division of Human Genetics and Molecular Biology, Children's Hospital of Philadelphia and the †Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104

(Received for publication, March 5, 1992)

Mucosal surfaces of several organ systems are important interfaces for host defense against microbes. Recent evidence suggests that antimicrobial peptides contribute to the defense of these surfaces. Defensins are one family of antimicrobial peptide, but their known distribution in humans has been limited to four members found in cells of myeloid origin. We sought to determine if the human defensin family was more complex. We found that the family of human defensins is diverse and is not restricted to expression in leukocytes. Southern blot and genomic clone analyses reveal that numerous defensin-related sequences are present in the human genome. A gene for a new human defensin family member was characterized. This gene, designated human defensin-5, is highly expressed in Paneth cells of the small intestine. This is the first example of an antimicrobial peptide gene expressed in an epithelial cell in humans. The data support the hypotheses that epithelial defensins equip the human small bowel with a previously unrecognized defensive capability which would augment other antimicrobial defenses.

Multicellular organisms utilize a variety of mechanisms to defend against microbial infiltration. These include anatomical and chemical barriers, as well as numerous cell-mediated and humoral responses. Peptide-based antimicrobial defenses are conserved components of host defense, and are found in both the animal and plant kingdoms (for reviews see Boman and Hultmark (1987), Bevins and Zasloff (1990), Spitznagel (1990), Boman (1991), Lehrer et al. (1991), and Zasloff (1992)). The sequence and structure of these antimicrobial peptides show significant diversity, but, in general, they are membrane-active amphipathic molecules with a net positive charge at neutral pH. Antimicrobial peptides may be grouped into several broad families based on certain characteristic structural features. Limited comparative studies suggest there is substantial variation in the distribution and expression of these peptide families between species (Boman, 1991; Zasloff, 1992).

Defensins are one family of well-characterized antimicrobial peptides which are 30–34 amino acids in length and have been isolated from myeloid-derived cells of several mammalian species (for recent reviews see (Ganz et al., 1990; Lehrer et al., 1991)). Defensins have in vitro antimicrobial activity against bacteria, fungi, and enveloped viruses (reviewed by Ganz et al. (1990)). Their activity is related to their ability to selectively disrupt membranes (Lehrer et al., 1989; Kagan et al., 1990; Lichtenstein, 1991). Four myeloid-derived human defensins have been isolated and characterized (Ganz et al., 1985; Selsted et al., 1985; Singh et al., 1988; Gabay et al., 1989; Wilde et al., 1989). These peptides are major constituents of neutrophil granules (Ganz et al., 1985; Selsted et al., 1985; Rice et al., 1987). Although the 16 defensins identified from several species are characterized by 11 invariant residues within the sequence, other residues in the peptides diverge. Thus, there is very limited similarity at the nucleotide level in the regions encoding these peptides, a feature that has hindered comparative studies by ordinary molecular biological approaches.

The expression of antimicrobial peptides by epithelial cells has recently attracted much interest (Zasloff, 1992). The synthesis of defense peptides by epithelial cells indicates that in addition to many other biological functions, these cells may actively participate in host defense, not simply constitute a barrier for this purpose. There have been several reports of the identification of antimicrobial peptides in mammalian epithelial tissues (Lee et al., 1989; Agerberth et al., 1991; Bateman et al., 1991; Diamond et al., 1991), but only two have clearly showed that epithelial cells per se, as opposed to leukocytes within the tissue, are the source of peptide synthesis (Ouellette et al., 1989; Diamond et al., 1991). No clear pattern has emerged yet to indicate which mammalian mucosal surfaces are protected by epithelial defense peptides, and there is no data that such peptides are synthesized by any epithelial cells in primates.

Paneth cells were first described over 100 years ago (Schwalbe, 1872; Paneth, 1888), yet their precise physiological role(s) has not been established. Many, but not all, mammalian species have Paneth cells, pyramidal shaped cells which reside at the base of the crypts of Lieberkühn in the small intestine. Paneth cells from various species share the properties of a distinctively staining cytoplasm and ultrastructural features consistent with a secretory cell (Trier, 1963; Benke and Moe, 1964). One product of the Paneth cells from numerous species is lysozyme (Paterson and Watson, 1961; Erlandsen et al., 1974; Klockars and Reitamo, 1975; Peeters and Vantrappen, 1975), an enzyme that would suggest these cells participate in host defense. However, there is notable variation in subtle morphology, ontogeny, and distribution between species (Satoh et al., 1990). These distinctions suggest there may be differences in function(s) of Paneth cells between species. Pioneering studies by Ouellette and colleagues...
leagues have beautifully demonstrated that mouse Paneth cells express the defensin-related mRNA encoding cryptdin, including inflammatory bowel disease (Paterson and Watson, 1961). An analysis of specific gene products of these secretory cells in humans will aid defining their physiological role.

The current study was undertaken to ask two fundamental questions about antimicrobial defenses in humans. First, is the human defensin family of antimicrobial peptides large, or is it limited to the four myeloid derived defensins currently recognized? Second, if the defensin family is large, are any of the newly discovered family members expressed in epithelial cells? In order to effectively approach these questions we needed to develop a molecular biological approach that would allow us efficiently identify new family members, despite very modest nucleotide conservation that the peptide encoding region. Using a new approach, we find data to indicate that the family of human defensins is much more diverse than currently appreciated and to demonstrate that their expression is not restricted to granular leukocytes. We report the first cloning of a human defensin gene and show it is expressed in Paneth cells, epithelial cells of small intestine. These observations support the hypotheses that antimicrobial defense peptides contribute to defense of mucosal surfaces in humans and that Paneth cells are involved in antimicrobial defense of the human small intestine.

**EXPERIMENTAL PROCEDURES**

*General Methodology—* All reagents were reagent grade from Sigma unless otherwise noted. Restriction enzymes were purchased from Bethesda Research Laboratories (Gaithersburg, MD) or Boehringer Mannheim (Indianapolis, IN). Oligonucleotides were made by the Nucleic Acid and Protein Core Facility, Department of Pediatrics, University of Pennsylvania School of Medicine. Oligonucleotide probes were end-labeled to a specific activity of approximately 10⁷ cpm/µg using [γ-³²P]ATP (800 Ci/µmol; Du Pont) and T4 polynucleotide kinase (Stratagene, La Jolla, CA). Double-stranded DNA probes were labeled to a specific activity of approximately 10⁶ cpm/µg using [α-³²P]dCTP (800 Ci/µmol; Du Pont) and T7 DNA polymerase with random oligonucleotide primers (Stratagene). Purified plasmid DNA was sequenced using the dideoxy-termination method (Sanger et al., 1977) with T7 DNA polymerase (United States Biochemicals) or using a thermal cycling method with fluorescently labeled primer oligonucleotides and Taq DNA polymerase (Applied Biochemical Systems, Foster City, CA). PCR products, purified by gel extraction (Bio101, La Jolla, CA) were incubated in a standard fill-in reaction with T4 DNA polymerase (Pharmacia LKB Biotechnology Inc.) and then subcloned by ligation to a blunt-end plasmid DNA (Bluescript, Stratagene). Sequence data were analyzed using DNA and protein analysis software MacVector (IBI, New Haven, CT).

*Northern Blot Analysis—* Genomic DNA was digested to completion with restriction enzymes according to the recommendations of the supplier. DNA samples were size-fractionated by agarose membranes using standard techniques (Reed and Mann, 1985; Sambrook et al., 1989). Hybridization with the S'/³²P-labeled end-labeled oligonucleotide probes D5 oligo (CTTGGCCATGTCGTTCATGTG) and HSIA262s (CTCTACAGACTCTGCTGTCGCTGAGCTTCCTAGAAG), HSIA262s (CCCTGCTTCTCTCAAGGAAAT), and HSIA309s (TTGCTGGGTCTGGGCTGGACCA) was in 20% formamide, 5 × SSC for hybridization and 50°C, 2 × SSC for high stringency wash. Plaques were taken through three or four rounds of purification at progressively lower densities. Phage DNA was isolated using Lambda-sorb (Promega). Phage insert DNA was subcloned by ligation into the multiple cloning site of Bluescript II (Stratagene) or using a thermal cycling method with fluorescently labeled primer oligonucleotides and Taq DNA polymerase (Applied Biochemical Systems, Foster City, CA). PCR products, purified by gel extraction (Bio101, La Jolla, CA) were incubated in a standard fill-in reaction with T4 DNA polymerase (Pharmacia LKB Biotechnology Inc.) and then subcloned by ligation to a blunt-end plasmid DNA (Bluescript, Stratagene). Sequence data were analyzed using DNA and protein analysis software MacVector (IBI, New Haven, CT).

*RACE-PCR protocol was modified from Frohman (Frohman et al., 1988). Total RNA (10 µg) and a poly(dt) primer (10 pmol) were used in the reverse transcription step. The DNA product was then labeled with dATP and terminal transferase. This DNA product was then used as a template in a PCR using T7RACE (TACGACTCATAGTATTATTTTTTTTTTTTTT), commercially available T7 oligonucleotide (AATACGACTCACTATAG), and HSIA262s primers according to the method described (Frohman et al., 1988).

*cdDNA and Genomic Cloning—* Lifting was made using Colony/Plaque Screen filters (Du Pont), and the filters were screened using standard techniques (Sambrook et al., 1989). The standard conditions for annealing and washing (Sambrook et al., 1989) were modified: 42°C, 20% formamide, 5 × SSC for hybridization and 50°C, 2 × SSC for high stringency wash. Plaques were taken through three or four rounds of purification at progressively lower densities. Phage DNA was isolated using Lambda-sorb (Promega). Phage insert DNA was subcloned by ligation into the multiple cloning site of Bluescript II (Stratagene) or using a thermal cycling method with fluorescently labeled primer oligonucleotides and Taq DNA polymerase (Applied Biochemical Systems, Foster City, CA). PCR products, purified by gel extraction (Bio101, La Jolla, CA) were incubated in a standard fill-in reaction with T4 DNA polymerase (Pharmacia LKB Biotechnology Inc.) and then subcloned by ligation to a blunt-end plasmid DNA (Bluescript, Stratagene). Sequence data were analyzed using DNA and protein analysis software MacVector (IBI, New Haven, CT).

*Northern Blot Analysis—* Genomic DNA was digested to completion with restriction enzymes according to the recommendations of the supplier. DNA samples were size-fractionated by agarose gel electrophoresis and blotted to nylon membranes using standard techniques (Reed and Mann, 1985; Sambrook et al., 1989). Hybridization with the S'/³²P-labeled end-labeled oligonucleotide probes D5 oligo (CTTGGCCATGTCGTTCATGTG) and HSIA262s (CCCTGCTTCTCTCAAGGAAAT), and HSIA309s (TTGCTGGGTCTGGGCTGGACCA) was in 20% formamide, 5 × SSC for hybridization and 50°C, 2 × SSC for high stringency wash. Plaques were taken through three or four rounds of purification at progressively lower densities. Phage DNA was isolated using Lambda-sorb (Promega). Phage insert DNA was subcloned by ligation into the multiple cloning site of Bluescript II (Stratagene) or using a thermal cycling method with fluorescently labeled primer oligonucleotides and Taq DNA polymerase (Applied Biochemical Systems, Foster City, CA). PCR products, purified by gel extraction (Bio101, La Jolla, CA) were incubated in a standard fill-in reaction with T4 DNA polymerase (Pharmacia LKB Biotechnology Inc.) and then subcloned by ligation to a blunt-end plasmid DNA (Bluescript, Stratagene). Sequence data were analyzed using DNA and protein analysis software MacVector (IBI, New Haven, CT).

*Blot Analysis—* Total RNA (obtained from Clontech) was fragmented by agarose gel electrophoresis in the presence of the presence of formaldehyde and blotted to nylon membranes (Zetabind, Cuno, Inc., Meriden, CT) by the capillary technique (Sambrook et al., 1989). RNA size standards (Bethesda Research Laboratories) were run in parallel lanes. The filter containing polyA-enriched RNA was similarly prepared (Clontech). Radioactively labeled DNA probes were hybridized to the immobilized RNA in 50% (v/v) formamide, 5 × SSC, 5 × Denhardt’s, 1% (w/v) SDS at 37°C, and washed in 1 × SSC, 0.1% SDS at 55°C, the same conditions of stringency as used in the in situ hybridization protocol (see below). The oligonucleotide probes used in these experiments were D5 (GAGCTGCTGCGCTGAGGCAAG), HSIA262s (CTTGGCCATGTCGTTCATGTG), HSIA309s (TGCTTTGGTTTCTATCTATA), and HSIA262s (GAGCTGCTGCGCTGAGGCAAG). Autoradiography exposure time at 4°C with intensifying screen was 3–14 days. Blots were stripped of oligonucleotide probes by treatments with 0.1 × SSC, 0.1% SDS at 65°C and then re-exposed to film to document removal of prior signal.

*In Situ Hybridization—* The method described by Young et al., 1986 was used for in situ hybridization histochemistry. Preparation of paraaffin-embedded tissue sections for use in this analysis was performed as described by Gilman (1987). The oligonucleotides HSIA262s and HSIA309s were 3′-end labeled to a specific activity of 10⁶ cpm/µmol using [α-³²P]dATP (100 Ci/mmol, Du Pont) and terminal deoxynucleotidyltrans-
RESULTS

Southern Blot Analysis—Our strategy for identifying new defensin family members was to utilize a nucleotide sequence that was common to several defensins based on limited sequence information in the literature as a probe for Southern blot analysis and for screening a human genomic library. Cloning at the genomic level offered potential advantages for defining diversity in this gene family, since expression of a novel defensin might be inducible, present in an unexpected tissue or developmentally regulated. The significant amino acid differences in nonconsensus residues of the mature peptide precluded the potential use of probes from this region to identify novel defensins.

We (Bevins and Zasloff, 1990) and others (Ganz et al., 1989; Kyllsten et al., 1990; Ouellette and Lualdi, 1990) have observed that within various families of antimicrobial peptides there is remarkably high conservation of mRNA sequence in the 5' region, and this similarity is particularly striking within the defensin antimicrobial peptide family. The nucleotides encoding the signal sequence of rabbit defensin-1 and -2 (Ganz et al., 1989) are 95% identical to that of human defensin-1 and -3 (54/57 identical nucleotides, Fig. 1) (Daher et al., 1988; Mars et al., 1988; Wiedemann et al., 1989). The nucleotide identity drops in other regions, remaining low over the segment encoding the mature peptides (53% nucleotide identity). We designed a 43-base oligonucleotide (D5'oligo) based on the sequence within the region of identity (Fig. 1). We also constructed a double stranded probe (CB587) which spans residues −29 to +184 (numbering relative to the first nucleotide of the putative initiating codon) of the human defensin-1 cDNA (Daher et al., 1988). About half of the DNA sequence of CB587 consists of the region of high conservation, the other half is more specific to human defensin-1 and -3. A search of the GenBank data base (release 60.0) using the University of Wisconsin Genetics analysis software (Devereux et al., 1984) found no sequences with significant similarity to these probes other than the known defensins.

We used these two probes in a series of experiments to test the hypothesis that the family of defensins is more diverse than currently recognized. Fig. 2A shows a Southern blot of human DNA digested to completion with each of six restriction enzymes and probed with D5'oligo. Multiple hybridization bands of similar intensity are seen in each of the first four lanes, and hybridization to high molecular weight DNA was seen for samples digested with SalI and XhoI. The activity of these two enzymes is sensitive to the methylation status of DNA, suggesting that the probe is hybridizing to a DNA region that is highly methylated in this sample. The hybridization was in 20% formamide, 5 × SSC at 42 °C, and the final wash was in 2 × SSC at 60 °C. Comparable results are seen when the final wash conditions were 2 × SSC at 63 °C (data not shown). Qualitatively similar results were also observed when the blot was sequentially stripped of probe and rehybridized with CB587 or with another oligonucleotide (HNP367a) from the 5'-untranslated region (data not shown).

In contrast to these observations, hybridization of the same blot to an oligonucleotide (HNP367a) corresponding to sequences in the 3'-untranslated region of human defensin-1 and -3 cDNAs yielded single bands in these restriction digests under similar conditions of stringency (Fig. 2B). Single bands of hybridization were seen with another oligonucleotide (HNP317s) from a portion of the coding region of the human defensin-1 and -3 cDNA (data not shown). These latter control experiments indicated that the conditions of stringency used in this series of experiments were adequate to identify highly complementary sequences in genomic DNA (as further addressed below). Fig. 2C shows a Southern blot of human DNA and that from each of six other species of animals digested to completion with HindIII and probed with CB587. The conditions used were 25% formamide, 5 × SSC at 42 °C for the hybridization and 58 °C in 2 × SSC for the final wash. Several strong bands are seen in both human and monkey samples, and weaker bands are seen in all of the other species, except perhaps mouse where the signals are very faint.
...digested to completion with each of six restriction enzymes, size-digested to completion with HindIII, and subjected to Southern blot probed with defensin related sequences.

portion of defensin mRNA, and the conservation of this exposure time was fractionated by agarose gel electrophoresis, blotted to a nylon filter, and the final wash was in 2 X SSC at 60 °C. Autoradiographic exposure was 14 days. The blot was the same as shown in Fig. 2A following two intervening hybridization experiments and probe stripplings (see text). Evidence for complete digestion in the Sau and Xhol samples was obtained from hybridization to the probe HNP19s (data not shown). C, genomic DNA from each of seven species was digested to completion with HindIII, and subjected to Southern blot analysis using the probe CB587 (see text for description of this probe). The conditions used were 25% formamide, 5 X SSC at 42 °C for the hybridization and 68 °C in 2 X SSC for the final wash. The autoradiographic exposure was 14 days.

gather these results indicated that human DNA contains numerous sequences with significant similarity to a conserved portion of defensin mRNA, and the conservation of this sequence extends between species. This suggested there is a larger family of defensin peptides than currently recognized, a notion that is further supported by our subsequent experiments.

Genomic Cloning—We screened an unamplified human genomic library (kindly provided by Drs. M. Budarf and B. Emanuel (The Children’s Hospital of Philadelphia) (McDermid et al., 1989)) with D5' oligo under the same conditions of stringency as for the Southern blot in Fig. 2. From approximately four genome equivalents of individual clones, we obtained 35 relatively strong signals on single filters. Twelve clones were purified and then categorized by a combination of restriction enzyme, hybridization, and partial sequence analysis. All 12 genomic clones had inserts in the range of 12-15 kb, and many contained more than one restriction fragment which contained a defensin-related sequence (data not shown). Hybridization properties to a panel of oligonucleotides and partial sequence analysis indicated that five of these clones contained sequences consistent with genes corresponding to the bone marrow derived defensins-1 and -3. These clones were temporarily set aside. Partial characterization of several other clones revealed that they also contained defensin sequences (data not shown). One clone, HG-2, was selected for in-depth characterization. An EcoRI restriction fragment that contained the defensin-like sequence within this clone was isolated, and the nucleotide sequence was determined (Fig. 3). Sequence analysis revealed two open reading frames that appear to encode portions of a putative preprodefensin molecule. We refer to this putative gene as defensin-5.

cDNA Cloning—In order to rapidly survey the possible tissue expression of this putative gene, oligonucleotides were selected for use as primers in a PCR (Saiki et al., 1988). The upstream sense primer was chosen from one defensin-related open reading frame and the downstream antisense primer was from the other. These primers were chosen so the amplification product would include an intron when the template was genomic DNA, a possible contaminant in a pool of cDNA. Fig. 4A shows the results of a PCR reaction using cDNA from eight tissues and genomic DNA as templates. A band of approximately 1.2 kb was seen in the genomic DNA sample, consistent with amplification across the intron. A band of 233 nucleotides was seen in the small intestine sample. No bands were seen in the other cDNA samples. A control amplification using a-tubulin primers shows that all cDNA samples contained amplifiable template (Fig. 4B). Subsequent hybridization of the amplification products to a probe spanning most of the putative exon 2 under high stringency confirmed the authenticity of the amplified DNA, and suggested high relative abundance of this sequence in the small intestine cDNA pool (Fig. 4C). On longer exposure there is some signal in fallopian tube and placenta (data not shown). A second control experiment using similarly designed primer pairs from the sequence of the human defensin-1 and -3 cDNA yielded strong etidium bromide staining band from the bone marrow cDNA and genomic DNA templates, but not from the small intestine or other cDNAs.2

We screened 2.5 x 10⁶ λ cDNA clones of a human small intestine library with D5' oligo (Chen et al., 1987) (generous gift of Dr. Lawrence Chan, Baylor). We observed 40 duplicate signals on primary screening. Twelve clones were taken through three rounds of plaque purification and isolation of phage DNA. Ten of these clones hybridized strongly with probes D5' oligo and CB587, and within these ten, two classes of clones were evident from hybridization patterns. Fig. 5A shows the hybridization pattern seen for two clones representative of each class when probed with D5' oligo and then washed at progressively higher temperature. The inserts from these four clones were subcloned and sequenced completely in both directions. Sequence analysis indicated clone 34 was identical to clone 25, except that it extended nine bases further at the 5' terminus (to nucleotide -19), and clone 14 overlapped clone 30 in a similar fashion. Fig. 5B shows the sequence alignment of each class of small intestine clones with the probe, D5' oligo. Clones S.I.25 and S.I.34 had sequence which corresponded exactly to the oligonucleotide probe; clones S.I.14 and S.I.30 were 85% identical. The sequence of clones S.I.25 and S.I.34 indicate that a messenger RNA expressed from the defensin-5 gene (Fig. 3) is expressed in the small intestine (Fig. 5C). The 3'-unique portion of this cDNA detects single bands of hybridization in Southern blot analysis (Fig. 5D), consistent with a single copy of this gene in the human genome. Clones S.I.14 and S.I.30 also have an open reading frame.3

To obtain additional nucleotide sequence of the 5' segment of the defensin-5 mRNA, a pool of small intestine cDNA was amplified using the rapid amplification of cDNA ends (RACE) technique (Frohman et al., 1988). The downstream primer (HSIA220a) was an antisense oligonucleotide from the puta-
Human Paneth Cells Express a Defensin Gene

Fig. 3. Nucleotide sequence of HG2-3e, a genomic subclone containing the human defensin-5 gene and flanking sequences. 

A, a partial restriction enzyme map of the 2.9-kb EcoRI fragment, HG2-3e, that encodes human defensin-5 (Eco, EcoRI; Xho, XhoI; Xba, XbaI; Hin, HindIII). Hash marks equal 100 base pairs. Thickened line shows position of two exons. The arrows indicate the sequencing strategy used to analyze this clone. 

B, the nucleotide sequence of HG2-3e with numbering in reference to the first nucleotide adjacent to the EcoRI site. Exon sequences are shown in upper-case lettering, and deduced amino acid sequence of the coding region is shown in three-letter code. The TATA box is underlined, and the CAAT box is double underlined. The consensus splice junction residues are shown in bold. The polyadenylation signal is boxed.
upstream open reading frames, and HSIA26la was an antisense oligonucleotide from the downstream open reading frames. Pools of genomic sequence. HNP63s was a sense oligonucleotide from the defensin-5 gene in eight human tissues.

A standard protocol for amplification was used as described for defensin-5 cDNA. The sequence represents a composite of human defensin-5 cDNA. The sequence upstream from the 5' terminus of the two most nucleotides identical to the probe. C, nucleotide sequence of human defensin-5 cDNA. The sequence represents a composite from the sequence of two λ phage clones (S1.25, nucleotides -10 to +415; S1.34, -19 to +413) and two clones from primer extension/RACE PCR (pDJ117-4 and pDJ117-5, -19 to -40), and the putative initiating methionine codon is assigned +1 to +3. The deduced amino acid sequence of the open reading frame is indicated in single-letter code. The polyadenylation addition signal is boxed. D, Southern blot hybridization of human genomic DNA probe with pS125-3’Mbo2, a segment of human defensin-5 cDNA (nucleotides +161 to +450). The filter was same as that used in Fig. 2B following stripping of probe. Hybridization was in 50% formamide, 5× SSC at 42 °C, and the high stringency wash was in 0.1× SSC at 65 °C for 30 min. The autoradiographic exposure was 14 days.

polyadenylation (Proudfoot and Brownlee, 1976) (Fig. 3, boxed). There is a TATA box at nucleotides 1328–1334, 24 nucleotides upstream from the 5’ terminus of the two most extended cDNAs identified by RACE-PCR (Fig. 3, underlined). A CAAT box is seen 87 nucleotides upstream from the termini of the extended cDNAs, at position 1267–1271 (Fig. 4).

FIG. 4. PCR analysis of possible expression of the human defensin-5 gene in eight human tissues. A, two oligonucleotide primers were chosen from potential open reading frames of the genomic sequence. HNP63s was a sense oligonucleotide from the upstream open reading frames, and HSIA26la was an antisense oligonucleotide from the downstream open reading frames. Pools of genomic sequence. HNP63s was a sense oligonucleotide from the defensin-5 gene in eight human tissues.

A standard protocol for amplification was used as described for defensin-5 cDNA. The sequence represents a composite of human defensin-5 cDNA. The sequence upstream from the 5' terminus of the two most nucleotides identical to the probe. C, nucleotide sequence of human defensin-5 cDNA. The sequence represents a composite from the sequence of two λ phage clones (S1.25, nucleotides -10 to +415; S1.34, -19 to +413) and two clones from primer extension/RACE PCR (pDJ117-4 and pDJ117-5, -19 to -40), and the putative initiating methionine codon is assigned +1 to +3. The deduced amino acid sequence of the open reading frame is indicated in single-letter code. The polyadenylation addition signal is boxed. D, Southern blot hybridization of human genomic DNA probe with pS125-3’Mbo2, a segment of human defensin-5 cDNA (nucleotides +161 to +450). The filter was same as that used in Fig. 2B following stripping of probe. Hybridization was in 50% formamide, 5× SSC at 42 °C, and the high stringency wash was in 0.1× SSC at 65 °C for 30 min. The autoradiographic exposure was 14 days.

polyadenylation (Proudfoot and Brownlee, 1976) (Fig. 3, boxed). There is a TATA box at nucleotides 1328–1334, 24 nucleotides upstream from the 5’ terminus of the two most extended cDNAs identified by RACE-PCR (Fig. 3, underlined). A CAAT box is seen 87 nucleotides upstream from the termini of the extended cDNAs, at position 1267–1271 (Fig. 4).

FIG. 5. Characterization of the human defensin-5 cDNA. A, Southern blot hybridization of four λ gt11 phage inserts isolated from a human small intestine cDNA library probed with D5’::oligo. Final washes in 2× SSC were at the indicated temperatures. B, partial sequence of four clones shown in A in the region corresponding to the probe, which is nucleotides identical to the probe. C, nucleotide sequence of human defensin-5 cDNA. The sequence represents a composite from the sequence of two λ cDNA clones (S1.25, nucleotides -10 to +415; S1.34, -19 to +413) and two clones from primer extension/RACE PCR (pDJ117-4 and pDJ117-5, -19 to -40), and the putative initiating methionine codon is assigned +1 to +3. The deduced amino acid sequence of the open reading frame is indicated in single-letter code. The polyadenylation addition signal is boxed. D, Southern blot hybridization of human genomic DNA probe with pS125-3’Mbo2, a segment of human defensin-5 cDNA (nucleotides +161 to +450). The filter was same as that used in Fig. 2B following stripping of probe. Hybridization was in 50% formamide, 5× SSC at 42 °C, and the high stringency wash was in 0.1× SSC at 65 °C for 30 min. The autoradiographic exposure was 14 days.

polyadenylation (Proudfoot and Brownlee, 1976) (Fig. 3, boxed). There is a TATA box at nucleotides 1328–1334, 24 nucleotides upstream from the 5’ terminus of the two most extended cDNAs identified by RACE-PCR (Fig. 3, underlined). A CAAT box is seen 87 nucleotides upstream from the termini of the extended cDNAs, at position 1267–1271 (Fig. 4).

FIG. 6. Northern blot hybridization of defensin expression in human tissues. A, total RNA (10 μg) from adult human small intestine was size-fractionated in a standard formaldehyde/agarose gel, capillary-blotted to a nylon membrane, and probed with HSIA309a, a defensin-5 probe. The conditions of stringency for hybridization and final wash were identical to those used in the in situ hybridization experiments (Fig. 7) (see "Experimental Procedures"). The size markers correspond to RNA standards from a parallel lane. The autoradiographic exposure was 2 days. B, a Northern blot of enriched polyadenylated RNA (2 μg) from eight adult human tissues was hybridized and washed in the same solutions as in A. The exposure was 10 days. This blot was subsequently used in the positive control experiments shown in C and D. C, control hybridization of the Northern blot shown in B to an actin probe (Gunning et al., 1983), following the experiments described in both B and D. The exposure was for 3 days. D, hybridization of the same Northern blot as in A with the antisense signal sequence oligonucleotide SIG68a. A band in the lung lane is readily apparent, but a faint band of the same size in the placenta lane did not reproduce well. The exposure was for 10 days.

3, double underlined). There is little nucleotide similarity between the 5’-flanking region of this gene and the previously reported rabbit defensin genes (data not shown). The cDNA sequence has an open reading frame of 282 nucleotides which potentially encodes a defensin-like prepropeptide. The context of the putative initiating methionine codon (CAGCCATGA) is identical to that found in the other two human defensin cDNAs and is consistent with a favorable translation start sequence (Kozak, 1991).

Northern Blot Analysis—RNA samples from small intestine and several other human tissues were subjected to Northern blot analysis using an oligonucleotide probe from this cDNA sequence (HSIA309a) (Fig. 6). The probe recognized an abundant message of approximately 600 nucleotides in the small intestinal RNA (Fig. 6A; 10 pg of total RNA, 3-day exposure). Under the same experimental conditions, no message was detected in pancreas, kidney, skeletal muscle, liver, lung, placenta, brain, or heart samples (Fig. 6B; 2 μg of poly(A) RNA, 10-day exposure), despite the presence of intact RNA in these lanes as demonstrated by hybridization to a β-actin probe (Fig. 6C). A control experiment using a signal sequence oligonucleotide showed a defensin-related mRNA in the lung sample under very similar conditions of stringency (Fig. 6D). A much fainter signal at this position (approximately 550 nucleotides) of migration was also detected in the placenta sample, but this did not reproduce well in this figure. The presence of a defensin-related sequence in human lung tissue is consistent with prior Northern blot analysis using a probe for defensin-1 and -3 (Daher et al., 1988), recent protein data from fetal lung tissue (Bateman et al., 1991) and investigations in our laboratory which led to the cloning of human defensin-1 cDNA from a human lung library.4

In Situ Hybridization—The cellular localization of the de-
defensin message was determined by in situ hybridization. Tissue sections of adult human intestinal mucosa were probed with sense and antisense 35S-labeled oligonucleotides. Strong signal was observed with the antisense oligonucleotide probe, HSIA309a in epithelial cells at the base of small intestinal crypts in sections from the adult ileum (Fig. 7, A and D). No signal was observed in parallel sections if the sense oligonucleotide probe, HSIA2628, was used (Fig. 7B) or if the sections were first treated with ribonuclease prior to hybridization with HSIA309a (Fig. 7C). Control experiments demonstrated that both the sense and antisense oligonucleotides were equally effective at hybridizing to 5.1.25 double stranded plasmid DNA under these experimental conditions (data not shown). Crypt cells with similar if not identical morphological characteristics stained strongly with phloxine-tartrazine, a histochemical stain commonly used to detect Paneth cells (Lendrum, 1947) (Fig. 7E). Furthermore, we find similar cells in parallel sections also express lysozyme (data not shown), a protein previously localized to the Paneth cell (Erlandsen et al., 1974; Peeters and Vantrappen, 1975). Eosinophils present in the lamina propria of numerous small intestinal sections also appeared to be weakly positive with the antisense oligonucleotide probe used in these experiments (for example, Fig. 7D), however the signal was not attenuated with pretreatment with RNase, and the sense oligonucleotide appeared equally positive (data not shown). The simplest explanation from these control experiments is that the signal in these white cells is probably artifactual as previously described (Patterson et al., 1989) and not from hybridization to specific RNA. No other cells in these sections hybridized to any of the probes.

DISCUSSION

The Human Defensins Are a Multiple Gene Family—We have developed a strategy for identifying new members of the defensin antimicrobial peptide gene family. Our strategy exploited the unusually high conservation of nucleotide se-

Fig. 7. Detection of defensin-5 mRNA in the Paneth cells of the human small intestine by in situ hybridization. A–C, low-power view of paraffin-embedded section of adult ileum hybridized with [35S]dATP-labeled defensin-5 oligonucleotide probes (bar equals 100 μm) and stained with hematoxylin and eosin. A, hybridization with the antisense oligonucleotide HSIA309a. B, hybridization of a serial section to that shown in A hybridized to the complementary sense probe HSIA261s. C, hybridization of a serial section as in A except that the tissue was pretreated with RNase A (10 μg/ml) for 15 min at room temperature. D, high-power view of selected region of slide shown in A (bar equals 40 μm). Note that weak signal in leukocytes in lamina propria is probably artifactual (see text). E, serial section at high magnification stained with phloxine-tartrazine, a histochemical stain which brightly stains Paneth cell cytoplasmic granules (Lendrum, 1947). See “Experimental Procedures” for experimental details. Exposure to photographic emulsion was for 14 days.
quence across the signal peptide coding region and extending into the 5'-untranslated region which is common to previously characterized defensin cDNAs. Using probes from this conserved region in Southern blot analyses (Fig. 2) and genomic cloning experiments, we find numerous defensin-related sequences present in the genomes of humans. This assertion rests, in part, on the specificity of the hybridization experiments, and several points should be made. 1) The final wash conditions for the oligonucleotide hybridizations would allow for relatively little mismatch. Under the stringency conditions used in the Southern blot analysis, the hybridization signal becomes significantly attenuated when there is a 15% mismatch with the 5's oligo (Fig. 5A). 2) With oligonucleotide probes to unique sequences, this level of stringency yielded discrete hybridization signals (Figs. 2B, 6, and 7A). 3) From analysis of five distinct defensin-related sequences that we have cloned under these conditions of stringency using this 5's oligo probe, four of five of these sequences are different from the previously cloned human defensins, and all five have >85% nucleotide identity to the probe.

Human Defensin-5 Is Expressed in Small Intestinal Paneth Cells—The defensin-5 gene is specific in its tissue expression (Figs. 4 and 6) and is expressed in the Paneth cell of the small intestine (Fig. 7). The physiological role of Paneth cells is not well defined (Sandow and Whitehead, 1979). Paneth cells are found throughout the bowel, but they are especially abundant in the small intestinal ileum (Hertzog, 1937). Paneth cells have abundant rough endoplasmic reticulum, an elaborate Golgi apparatus, and large secretory vesicles typical of a secretory cell (Trier, 1963; Benke and Moe, 1964). Several lines of evidence suggest these cells have a role in antimicrobial defense. Paneth cells have been shown to express tumor necrosis factor mRNA (Keshav et al., 1990) and to contain lysozyme (Paterson and Watson, 1961; Erlandsen et al., 1974; Klockars and Reitamo, 1975; Peeters and Vantrappen, 1975). The lysozyme is localized to secretory granules in these cells (Deckx et al., 1967; Peeters and Vantrappen, 1975), and various stimuli have been shown to effect degranulation of these cells (Klockars and Reitamo, 1975; Peeters and Vantrappen, 1975; Satoh et al., 1986; Satoh and Vollrath, 1986; Satoh et al., 1989). Pioneering studies by Ouellette and colleagues identified the expression of defensin-related mRNA in the Paneth cells of the mouse small intestine. Using differential hybridization technique, these investigators had identified this defensin-related clone as one of several messages that are differentially expressed in the small intestine soon after birth (Ouellette and Cordell, 1988). The localization of mouse cryptdin (Ouellette et al., 1989) and now human defensin-5 to the Paneth cell is consistent with an antimicrobial function of these cells.

Possible Physiological Roles of Small Intestinal Defensins—Two possible physiological roles of small intestinal defensins are suggested. First, the defensins might regulate the level of luminal microbiological flora. The high density of Paneth cells near the distal ileum might contribute to a barrier restricting the abundance of intestinal microbiological flora to the colon. Second, the defensins could be important in mucosal defense from microbial invasion. An effective host defense system in the small bowel that did not require significant inflammation could preserve the integrity of the villus epithelium, and thereby maintain the critical function of nutrient absorption. Small intestinal defensins might contribute to such a defense. These two possibilities are not mutually exclusive. In support of these speculations, several studies have identified virulence factors of enteric pathogens that are possibly targeted at membrane-active host defenses, such as defensins (Fields et al., 1988; Groisman et al., 1989; Parsot et al., 1991).

Antimicrobial Defense at Numerous Wet Mucosal Surfaces May Utilize Defensin-like Peptides—The studies presented here extend to three the number of well characterized examples of defensin-related expression by epithelial cells of mammals. Previous studies by Ouellette and colleagues beautifully demonstrated that defensin-encoding mRNA is present in the mouse small intestinal crypt cells (Ouellette et al., 1989). Recent studies in our laboratory have found that the epithelial cells at a different anatomic location, the tracheal mucosa of the cow, also express a defensin-like peptide named "tracheal antimicrobial peptide" which has broad-spectrum antimicrobial activity. The findings of the present study, together with previous observations in the mouse and cow, allow us to generalize that vertebrate gastrointestinal and respiratory epithelium express defensin-related sequences, and similar expression might also be found in other mucosal epithelia of mammals.

Structural Features of the Predicted Defensin-5 Gene Product—Comparison of the deduced amino acid sequences of defensin-5 cDNA with the previously reported preprodefensins shows significant similarity with respect to size and charge distribution (Fig. 8). The carboxyl and amino termini of the putative mature defensin-5 peptide can be tentatively inferred from patterns emerging from analysis of the two rabbit and two human defensins where peptide and cDNA data are available. An in-frame stop codon follows the last residue of all of these peptides. The amino-terminal amino acid of the mature peptide is 2 residues from a conserved cationic amino acid in the putative prepropeptide (Fig. 8).

The high resolution crystal structure of human defensin-3 has been recently determined (Hill et al., 1991). Several residues were argued to be of key importance for defining the tertiary structure of this defensin. The deduced amino acid sequences of human defensin-5 and mouse cryptdin have conservation of all of these critical residues. The small intestinal defensins also have similar composition of charge and hydrophobic amino acid residues in the other variant positions. Based on these sequence similarities, it is reasonable to predict that the small intestinal defensins will have tertiary and quaternary structure comparable to human defensin-3, and will have similar biological activities to the other defensins. However, these assertions will need to be tested when small intestinal defensin peptides become available. Because of the complex disulfide array of these peptides, chemical synthesis of this peptide is well beyond the scope of this study.

The Significance of the High Nucleotide Conservation in the 5' Segment of the Defensin Message—Sequence conservation is observed in the signal peptide region of the putative preprodefensins (Fig. 1), suggesting an evolutionary pressure to maintain sequence. The striking sequence conservation extends to the 5'-untranslated mRNA region, and relatively little silent codon substitution is observed as one might expect if evolutionary conservation was maintained primarily at the level of the protein sequence. An attractive possibility is that the high sequence conservation in this region results from

1 G. Diamond and C. L. Bevins, manuscript in preparation.
2 These residues include the 6 cysteines which participate in disulfide bonds, arginine 6, glutamic acid 14, and glycine 24. Certain residues were thought to be key in the dimeric quaternary structure: cysteine 5, glycine 18, cysteine 20, and hydrophobic residues at positions 22 and 28 (159; Hill (1991)). All of the aforementioned residues are conserved in 14 of the 15 defensin peptides isolated to date, all of which have broad spectrum antimicrobial activity. The guinea pig defensin was the exception, having substitutions in 2 of these residues.
**Fig. 8.** Comparison of deduced amino acid sequence of human preprodefensin-5 and the putative mature defensin-5 with known defensin family members. The deduced amino acid sequence of six preprodefensins are from published cDNAs, +, cationic residues in human preprodefensin-5; --, anionic residues. Residues identical to those of preprodefensin-5 are indicated by "=" in the human (Daher et al., 1988; Mars et al., 1988; Wiedemann et al., 1989), rabbit (Ganz et al., 1989), and mouse (Ouellette and Lualdi, 1990) sequences. Human preprodefensins-1 and -3 are identical except for the single amino acid indicated, and the two rabbit sequences are identical except at indicated residues. The amino-terminal residues of the human defensins-1 and -3 and rabbit defensins-1 and -2 are indicated with Δ based on available peptide data (Selsted et al., 1985, 1985a, 1985b). No peptide data is yet reported for defensin-5 or mouse cryptdin, and the amino terminus indicated is speculation based on cleavage of 2 residues from the conserved cationic amino acid in the propeptide as in the other defensins. Short gaps (indicated by "-" ) were included in the sequences to aid in alignment. A consensus sequence for reported preprodefensins is presented, where an upper-case letter denotes identity in all six sequences and lower-case letters are conserved in four (or five) of the six.

**Fig. 9.** The primary sequence of mature defensin peptides from five mammalian species. The format is similar to that presented previously (Hill et al., 1991; Lehrer et al., 1991). Solid boxes indicate sequence identity from published sequence data (Selsted et al., 1985, 1985a, 1985b; Selsted and Harwig, 1987; Singh et al., 1988; Eisenhauer et al., 1989; Ouellette et al., 1989; Wilde et al., 1989), and hashed boxes indicate near-perfect consensus. The amino-terminal residues of human defensin-5 and mouse cryptdin are double underlined to indicated that they are based on cDNA analysis, not peptide data.

Selective pressure on the nucleotide sequence, possibly to affect mRNA stability or maintain mRNA structure. Secondary structure is known to affect translatability of mRNA and several eucaryotic genes are regulated at this level (for reviews see Jagus et al., 1981, Klauser and Harford (1989), Kozak (1991)). Studies on defensin gene expression, translation, and intracellular trafficking will help discriminate between these interesting possibilities.

Acknowledgments—First and foremost, the generous support and encouragement of Dr. Michael Zasloff is gratefully acknowledged. We also thank several members of the Division of Human Genetics for helpful discussions and advice, especially Drs. Gill Diamond, Dorothy Reilly, Patrick O'Connor, and Michael Zasloff. We value the informative discussions concerning this work with Drs. Michael Selsted, Andy Ouellette, and Susan Weil. We thank Dr. Lawrence Chan for an aliquot of human small intestine cDNA and Drs. Marcia Budarf and Beverly Emanuel for use of a human genomic library. We thank Dr. Douglas Clark for access to blocks of human small intestine for in situ hybridization experiments. We thank Rita Cliff and Joshua Schaffzin for technical help. The efforts of the Nucleic Acid and Protein Core Facility at Children's Hospital, especially those of Dr. Eric Rappaport, Stephanie Friscia, and Steve Mahoney were invaluable to these studies.

Note Added in Proof—After submission of this paper, Ouellette et al. (Ouellette, A. J., Miller, S. L., Henschen, A. H., and Selsted, M. E. (1991) FEMS Lett. 304, 146–149) published the sequence of the mouse cryptdin peptide isolated from small intestinal extracts. The sequence of the isolated peptide is LRDLVYCRRGKCRERRMNGTCRKGHLTYTLCR. Thus, the predicted sequence shown in Fig. 8, which was based on our analysis of published cDNAs and a comparison to myeloid-derived defensin peptides, is incomplete at the amino terminus. It should be extended by 4 residues. This finding suggests that the amino terminus of human defensin-5 may also have an extended amino terminus from what is shown in Fig. 9.

REFERENCES


Note Added in Proof—After submission of this paper, Ouellette et al. (Ouellette, A. J., Miller, S. L., Henschen, A. H., and Selsted, M. E. (1991) FEMS Lett. 304, 146–149) published the sequence of the mouse cryptdin peptide isolated from small intestinal extracts. The sequence of the isolated peptide is LRDLVYCRRGKCRERRMNGTCRKGHLTYTLCR. Thus, the predicted sequence shown in Fig. 8, which was based on our analysis of published cDNAs and a comparison to myeloid-derived defensin peptides, is incomplete at the amino terminus. It should be extended by 4 residues. This finding suggests that the amino terminus of human defensin-5 may also have an extended amino terminus from what is shown in Fig. 9.