The EMBO Journal vol.12 no.4 pp.1569–1578, 1993

Apidaecin multipeptide precursor structure: a putative mechanism for amplification of the insect antibacterial response

Katty Casteels-Josson[1], Theresa Capaci, Peter Casteels and Paul Tempst

Molecular Biology Program, Memorial Sloan-Kettering Cancer Center and Cornell University Graduate School of Medical Sciences, 1275 York Avenue, New York, NY 10021, USA

[1]Corresponding author

Communicated by M.van Montagu

Apidaecins are the most prominent components of the honeybee humoral defense against microbial invasion. Our analysis of cDNA clones indicated that up to 12 of these short peptides (2 kDa) can be generated by processing of single precursor proteins; different isoforms are hereby linked in one promolecule. Assembly of the multipeptide precursors and the putative three-step maturation are strongly reminiscent of yeast alpha-mating factor. Bioactive apidaecins are flanked by the two 'processing' sequences, EAEPEAEP (or variants) and RR; joined together, they form a single unit that is repeated numerous times. The number of such repeats is variable and was reflected in the observed diversity of transcript lengths. Each such transcript is likely to be encoded by a different gene, forming a tight gene cluster. While transcriptional activation upon bacterial challenge is not exceptionally fast, the multigene and multipeptide precursor nature of the apidaecin genetic information allows for amplification of the response, resulting in a real overproduction of peptide antibiotic. Enhanced efficiency of the 'immune' response to bacterial infection through such a mechanism is, to our knowledge, unique among insects.

Key words: amplification/antibiotic/apidaecin/insect immunity/multipeptide

Introduction

Infection of insects, or introduction of foreign material in their hemolymph (blood), triggers a cellular and humoral 'immune' response that functions to eliminate the invading bacteria. As part of the humoral defense, a series of different peptides are usually produced; in combination they are responsible for killing a wide spectrum of bacteria. Most of them are strongly induced or upregulated upon infection; regulation seems to be located at the transcriptional level. Different sets of antibacterial peptides have been isolated from immunologically challenged ('immune') insects, representing such major orders as the moths/butterflies (Lepidoptera), flies (Diptera), bees/wasps (Hymenoptera) and beetles (Coleoptera). For an overview of this area of research, we refer to reviews by Hoffmann and Hoffmann (1990) and Boman *et al.* (1991), and a report by Bulet *et al.* (1991) and references therein.

While some of these antibacterial peptides, such as lysozyme, defensins and cecropins, are distributed among several insect orders and have even been found in mammalian tissues (Lee *et al.*, 1989; Lehrer *et al.*, 1991; Sun *et al.*, 1991a), the presence of others seems to be restricted to just one group of insects. A case in point are the honeybee peptides apidaecin, abaecin and hymenoptaecin which, at least until now, have never been isolated from insects other than Hymenoptera (Casteels *et al.*, 1989, 1990, 1993). In combination with defensins (royalisin) (Fujiwara *et al.*, 1990), these peptides provide wide-spectrum protection to the honeybee by virtue of complementarity.

Apidaecins are small (18 amino acids), proline-rich antibacterial peptides produced in 'immune' bees. Activities of all three known isoforms seem to be targeted against the same, predominantly Gram-negative bacteria, with minimal inhibitory concentrations in the 10^{-8} to 10^{-6} M range (Casteels *et al.*, 1989). As found for several other insect-derived antibiotics, the existence of multiple isoforms usually arises from multigene families (Ando and Natori, 1988; Reichhart *et al.*, 1989; Boman *et al.*, 1991) or different alleles (Reichhart *et al.*, 1989; Kylsten *et al.*, 1990). Inactive apidaecin precursors, containing additional amino acids at the N-terminus, have also been isolated from bee lymph (Casteels *et al.*, 1989; P.Casteels, unpublished), suggesting that, like cecropins (Dickinson *et al.*, 1988; Boman *et al.*, 1989; Kylsten *et al.*, 1990), sarcotoxin II (Ando and Natori, 1988), diptericins (Wicker *et al.*, 1990) and melittin precursors (Kreil *et al.*, 1980), apidaecin precursors need an aminopeptidase for maturation. Incidentally, a dipeptidyl aminopeptidase, necessary for melittin maturation, has already been purified from honeybees (Kreil *et al.*, 1980).

In this report, we describe the isolation and the characterization of several apidaecin cDNA clones and show that multiple mature peptides are generated from single precursor molecules. The number and nature (isoform) of the peptides derived from single precursors are variable; this is the result, most likely, of the existence of an apidaecin gene cluster. The contribution of the unique precursor structures and kinetics of induction to the 'immune' response of the honeybee are discussed.

Results

Cloning and sequence analysis of apidaecin cDNA

Based on the amino acid sequence of apidaecin Ib (Casteels *et al.*, 1989), a 29 nucleotide (nt) degenerate probe (B6-1) was constructed for screening a cDNA library derived from immunologically induced bees. Ten positive clones (out of 74) were arbitrarily selected and analyzed; insert lengths varied between 600 nt and 1100 nt: pApid22 and pApid64 (600 nt), pApid40 and pApid31 (660 nt), pApid14 (700 nt), pApid18 and pApid26 (860 nt), pApid43 (880 nt), pApid73 (1000 nt) and pApid44 (1100 nt). Initial sequence analysis revealed that all clones contained nearly identical sequences at the 5′ and at the 3′ ends, indicating that the observed

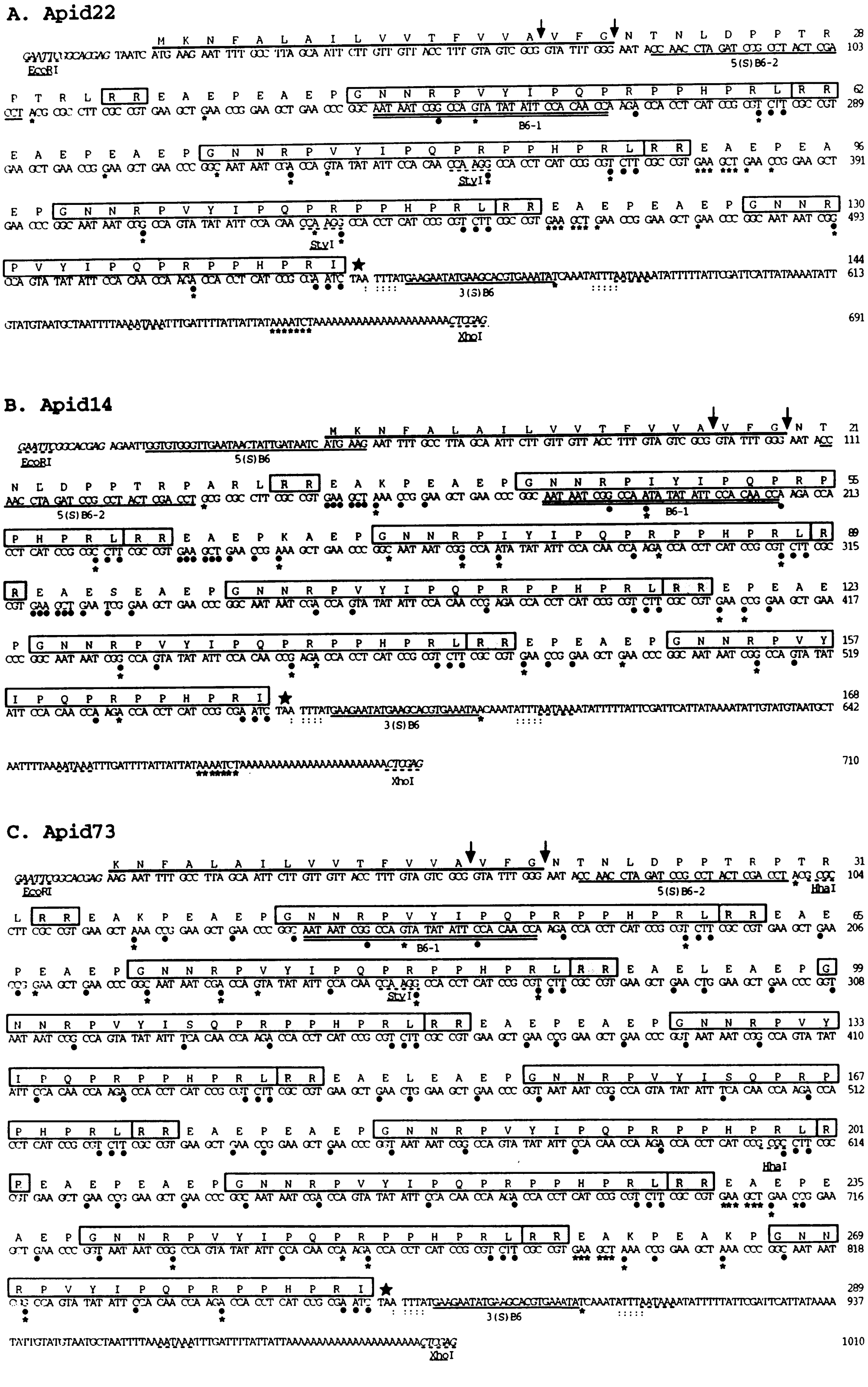


Fig. 1. Nucleotide sequences and predicted amino acid sequences of prepro-apidaecin in pApid22 (A), pApid14 (B) and pApid73 (C) cDNA clones. Numbers at the right indicate positions of the nucleotides and amino acids. The underlined amino acids locate the two possible, predicted signal sequences; arrows indicate their putative processing sites. The mature peptides are indicated by an open box, the basic (RR) dipeptides (see text) by a dotted box. Amino acids which vary between units within the same clone, as described in detail in Figure 2, are indicated with a dot; variable sequences in units between different clones (for details see Figure 3) are indicated with a star; both symbols are located under the nucleotide sequence. The large star symbol above the nucleotide sequences indicates the translational stop signal (TAA). The sequence matching the degenerate oligonucleotide (B6-1), used to screen the cDNA library, is doubly underlined; it is shown only in the first unit of each clone. Primers used for sequencing and primer extension [5′(S)B6, 5′(S)B6-2 and 3′(S)B] are indicated under the cDNA sequence by a single line, their orientations are described in the respective experiments and shown in detail in the Materials and methods section. Restriction sites and poly(A) adenylation signals (AATAAA) are indicated by a dashed line. ATTTA sequences (see text for functional importance) are underlined by a doubly dotted line.

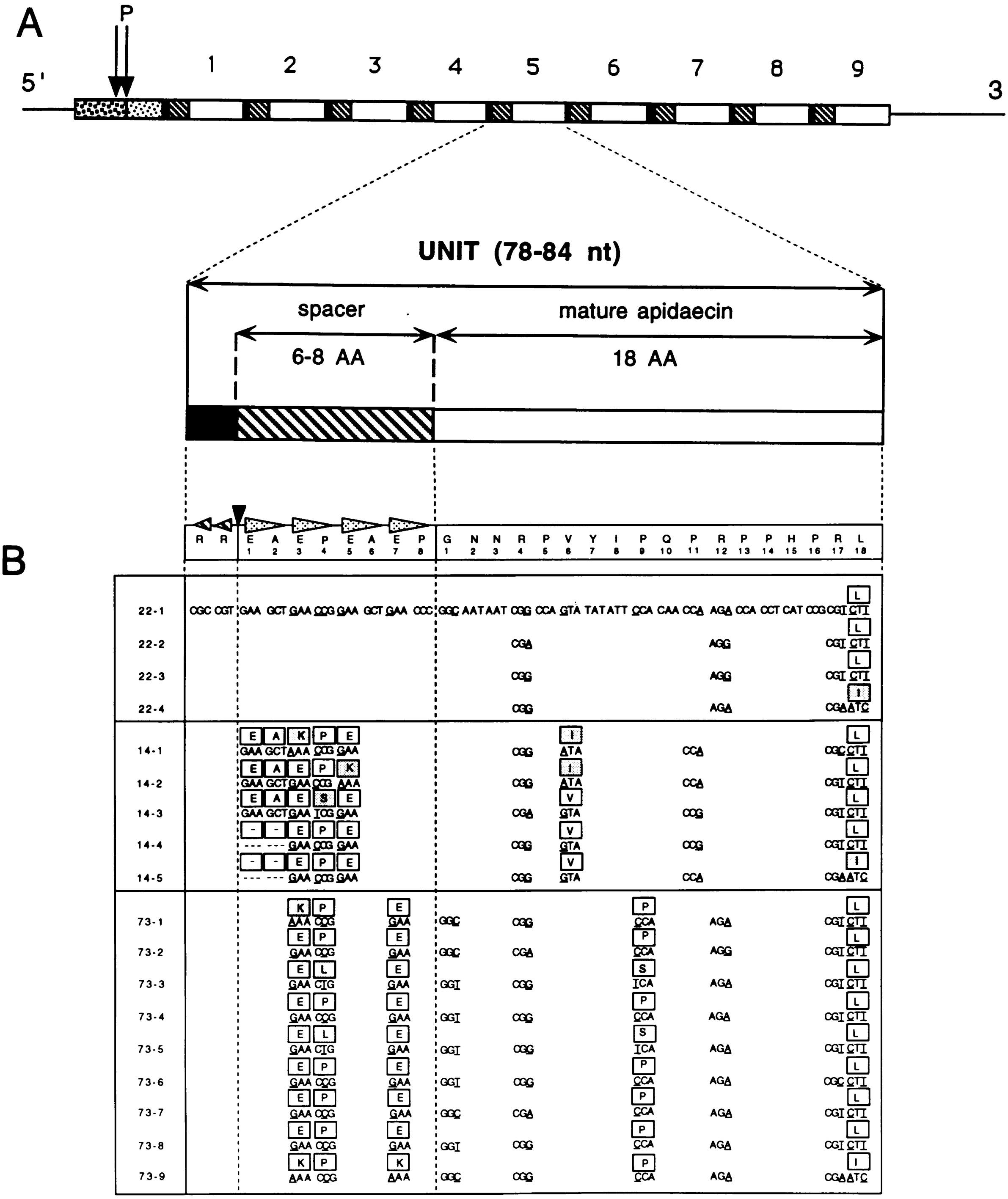


Fig. 2. Schematic structure of apidaecin precursors and sequence variability of all units determined in cDNA clones pApid22, 14 and 73. (**A**) General and schematic representation of prepro-apidaecin. The prepro-signal consists of a signal sequence (dark dotted box) and a pro-sequence (light stippled box). P indicates the putative cleavage sites for secretion. A single unit consists of a mature apidaecin (open box), preceded by a spacer region (hatched box) and an RR dipeptide (black box) and is repeated several times. The number of repetitions varies and is clone dependent. For instance, the one shown at the top of the figure contains 9 units (as in clone 73). (**B**) Detailed analysis of sequence variation in all units from clones 22, 14 and 73 of prepro-apidaecin polypeptides and their corresponding genes. The coding region and corresponding aa sequence of the first unit of clone 22 is shown in full. For all other units only those codons are shown that differ from unit 22-1. The nucleotide substitution responsible for the variation is underlined and the corresponding amino acids are boxed. Amino acids occuring least frequently in a particular position are shown in a lightly stippled box. Small dashes indicate the absence of an amino acid or nucleotide. Putative processing sites for prepro-apidaecin by an endopeptidase (▼) and a carboxypeptidase (◀), and processing sites for dipeptidyl aminopeptidase (▷) are indicated above the upper amino acid sequence. Triangles indicate the direction of processing and their length indicate the number of amino acids processed in a single step.

differences in length were located in the middle of the fragments. Therefore, a small (pApid22), a mid-sized (pApid14) and a large clone (pApid73) were selected for complete sequence analysis (see Figures 1, 2 and 3). From the other clones only the 5′- and 3′-regions were analyzed. Clone pApid14 contained the most extensive 5′ non-coding

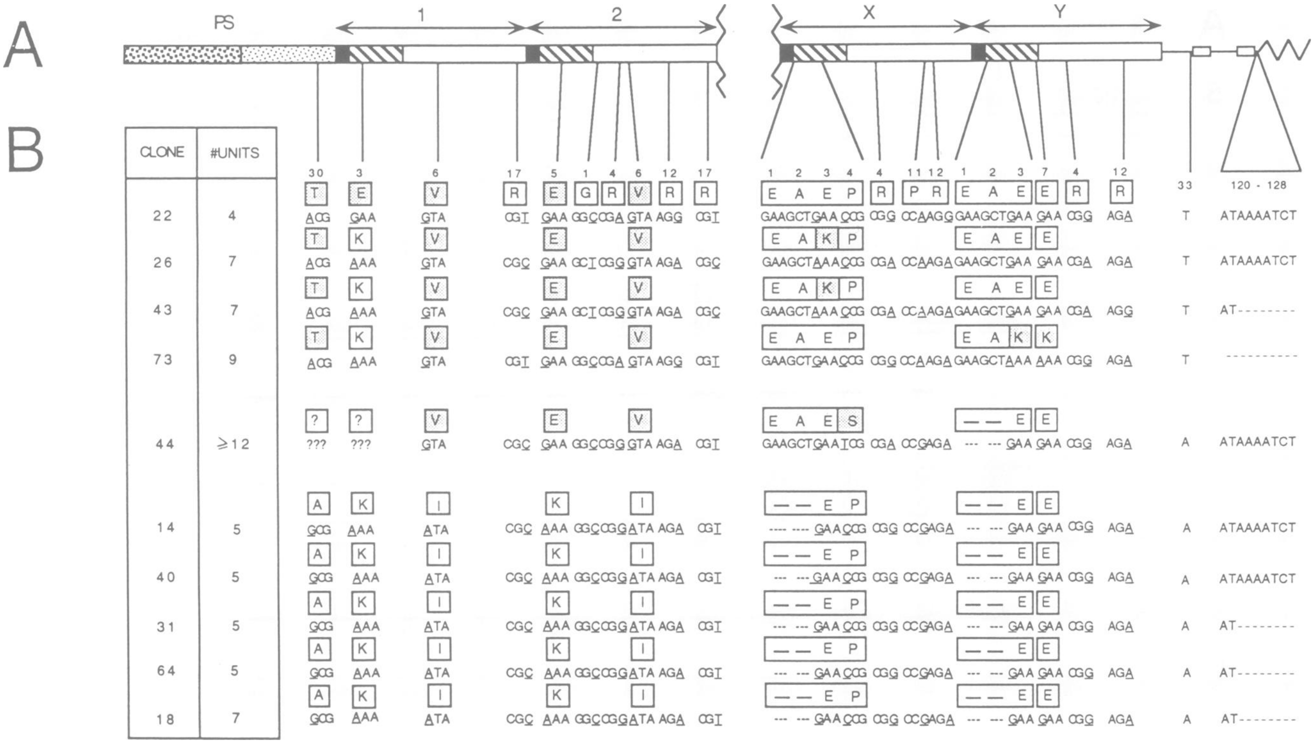


Fig. 3. Partial analysis of the nucleotide and corresponding amino acid variability of 10 different apidaecin cDNA clones. (**A**) Schematic representation of analyzed apidaecin sequences. Large boxes indicate the coding region; each unit contains the different features as described in Figure 2A. The 3′ non-coding region is shown by a line; small open boxes indicate the poly(A) adenylation signals and the horizontal sawtooth line indicates the poly(A) tail. Different units are indicated by arrows and numbered 1, 2 for the first two and X, Y for the last two units. (**B**) Detailed analysis of the variable sequences. Clones are grouped based on mutual sequence patterns. The figure is organized as Figure 2B. Vertical lines descending from the schematic figure indicate where variations have occurred. For each clone, the number of units is shown.

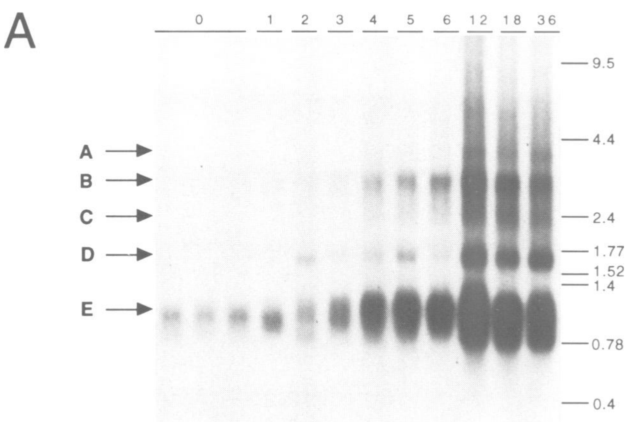
sequence. Only one ATG initiation codon could be found and was located 34 nt from the 5′ end; the presence of a stop codon (TAA) in the same reading frame and upstream of this initiation codon (−15 nt) argues that apidaecin precursor molecules indeed start at the proposed point. This initiation codon is enclosed within a sequence similar to the translational initiation region of *Drosophila* nuclear proteins ($^{C}/_{A}AA^{C}/_{A}\underline{ATG}$) (Cavener, 1987), except for the third nucleotide which is a T instead of an A in the apidaecin sequence. Inserts of pApid64, 40, 26, 43, 73 and 44 clones started immediately within the preproapidaecin coding region. The last amino acid, isoleucine, is conserved for all precursor molecules. Two poly(A) adenylation signals, separated by 50 nt, were found at the 3′ end of the transcript. Variability was detected in the number of nucleotides located between the second polyadenylation signal and the poly(A)-tail: 15 nt (pApid73), 17 nt (pApid43, 31, 64, 18) and 24 nt (pApid22, 26, 44, 14, 40) (Figure 3). Additionally, two consensus sequences (ATTTA), known to affect the stability of corresponding messenger RNA (Shaw and Kamen, 1986), were observed (Figure 1).

Further analysis revealed that each cDNA clone encodes a precursor sequence containing a prepro-sequence and a variable number of repeated, 84 nt long, almost identical units (Figure 2A). Each unit consists of a coding sequence for the mature apidaecin peptide, preceded by coding regions for a spacer sequence and a basic (RR) dipeptide. Some clones seemed to contain smaller units of only 78 nt; this was accounted for by the presence of a smaller spacer region (Figures 2B and 3B). Variation in the number of repeating units resulted in the observed differences in insert length for the different cDNA clones. Based on the insert length and the known length of 5′ and 3′ untranslated regions, of the presignal coding region length and of the unit length (generally 84 nt), the number of units could be calculated for each clone (except for pApid44). This number varied between 4 (pApid22), 5 (pApid14, 31, 40, 64), 7 (pApid18, 26, 43) and 9 (pApid73). Because the 5′ sequence of pApid44 started within a unit, the absolute number of units for this clone could not be determined, but was estimated to be 12.

Hydrophobicity prediction analysis of pApid22 apidaecin precursor suggested a small hydrophobic region (defining the signal sequence) and a hydrophilic part (containing a prosequence and the repeated apidaecin units). Using the method of Von Heijne (1986), two possible processing sites were found, resulting in putative signal peptides of 13 or 16 amino acids in length (see Figure 1). The normalized probability of these sites is 1.0 and 0.74 for the first and the second site, respectively [Miscellaneous Protein Analysis (sigseq), Rockefeller University Computing Services, New York, NY].

Sequence variation between apidaecin units

The nucleotide sequences of all characterized units were found to display several variations; some of them resulted in amino acid sequence changes of the corresponding polypeptide precursor molecules. Differences were observed, both between corresponding units of different clones (e.g. unit 1 of clones 22, 14 and 73) and between different units within single clones, as is illustrated in Figure 2. In this figure, fully sequenced clones (pApid22, pApid14 and



B

induction factor RNA A	1*	1*	1*	1	4	4	9	8	8	44	35	62
B	1	1	1	1	2	3	13	15	25	52	52	44
C	1*	1*	1*	2	3	4	7	23	44	99	82	73
D	1*	1*	1*	3	18	5	16	33	11	150	81	246
E	1	1	1	1	1	2	4	6	7	13	12	17
pep	1	1	1	1	1	1	1	1	1	4	5	10

Fig. 4. Kinetics of apidaecin mRNA and (mature) peptide synthesis following bacterial infection of bees with *E.coli* cells. (**A**) Northern blot analysis using an Apid73 cDNA-derived RNA probe. Time points (in hours between the time of infection and RNA sampling) are indicated above the corresponding lanes. RNA samples (2.5 μg each) were taken from a pool of 20 bees at different time points. Positions of RNA molecular weight standards (in kilobases) are shown on the right. (**B**) Induction factors of different 'apidaecin-like' mRNA species and of mature peptide. Relative induction factors (see Materials and methods for calculations) for all bands (A−E) are shown for each time point under corresponding lanes of the blot. Small star symbols indicate that for some bands no blank values were found; instead the grey background was arbitrarily assigned the value of 1. Production of mature apidaecin in bee lymph was monitored by reversed phase HPLC as described in Materials and methods. The relative induction factor of the peptide in function of time is given in the bottom row (pep).

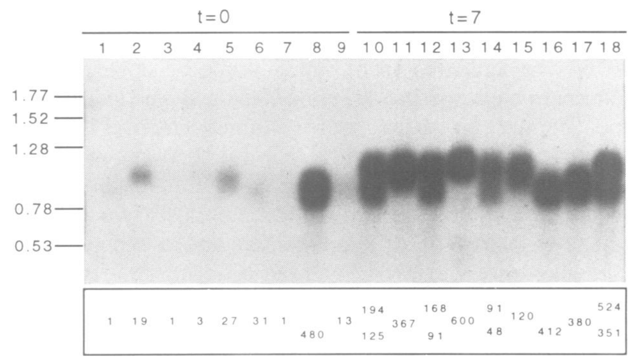


Fig. 5. Transcription of apidaecin mRNA in single bees. Thorax tissue from immune bees was sampled 7 h post-induction (t = 7), RNA (10 μg) was isolated and compared with that of several non-induced bees (t = 0). Positions of RNA molecular weight standards (in kilobases) are shown on the left. Resulting relative induction factors (see Materials and methods) are shown under their respective lanes; the top and bottom line correspond to the top and bottom band on the blot, respectively.

pApid73) are shown; the first unit of clone pApid22 was used as the reference sequence. Variations are not random in nature; the most consistent one was the substitution of the conserved C-terminal Leu residue by an Ile in the last unit, and the last unit only, of every clone.

Further detailed sequence analysis of similarly positioned units within 10 different clones indicated additional trends in amino acid preference for some positions (Figure 3). Within mature apidaecin, Ile (position 6) replaced Val in the first two units of some, but not all, clones. If present, Ile 6 was always found in both the first and second unit, but never in the last two; Ile6 in the second unit is uniquely correlated with a Glu to Lys substitution at position 5 in the corresponding spacer sequence. A variety of additional differences in the spacers, including dipeptide deletions, were also observed. A single change in the pro-sequences (at position 30, Ala/Thr) was detected.

Apidaecin mRNA population

Following the observation of high variability in apidaecin cDNA clone lengths, we were interested to determine if and how this was reflected in the mRNA population. It is a well-documented fact that insect-derived immune peptides are regulated at the transcriptional level following bacterial infection (reviewed by Faye, 1990). A series of experiments, aimed at characterizing apidaecin mRNA, was therefore carried out with RNA isolated from bees at timepoints of one or more hours post-infection (PI).

Northern blot analysis. Most of the apidaecin transcripts ranged in size from 680 to 1400 nt (band E, Figure 4). Because no distinct bands were detected, the upper and lower edge of the polydisperse hybridization band were used to calculate the lengths of corresponding RNA molecules. This might have resulted in an over- and underestimation in length for largest and smallest transcripts, respectively. Longer exposure of the autoradiogram revealed apidaecin mRNA of as little as 420 nt in length (data not shown). Thus, apidaecin transcripts encode polypeptide precursor molecules containing a number of units varying from 1 (420 nt mRNA) to 12 units (1400 nt mRNA). Some high molecular weight RNA species of 1750 nt (D), 2400 nt (C), 3300 nt (B) and 4000 nt (A) were also found to hybridize with the apidaecin probe (Figure 4). Since these bands did not comigrate with the bee ribosomal RNA molecules (running as a single band of 2000 nt, data not shown) we assumed that they contained apidaecin(-like) sequences.

PCR. Although it became clear that apidaecin mRNA was a mixed population, we were unable to analyze the real variation of unit repetition in greater detail because of the absence of discrete bands on Northern blots. Subtle variations in poly(A)-tail length and/or the sizes of 3′ and 5′ untranslated regions might account for this observation by causing indistinct boundaries between otherwise separated bands. To eliminate these possible shortcomings, a PCR reaction was started using the 5′SB6, 3′B6 oligonucleotides (Figure 1) as primers and mRNA from a pool of bees at 15 h PI as described by Mahbubani *et al.* (1991). The resulting amplified fragments ranged from 230 to 738 nt in size (data not shown). These fragments encoded one to at least seven units and differed from each other by at least a single unit (84 nt). No larger discrete fragments could be detected. Hybridization with an inner *Hha*I fragment, which does not contain any 5′ or 3′ flanking sequence (see Figure 1C), confirmed that all PCR fragments contain genuine apidaecin sequences.

mRNA from single insects. To determine the possible contribution of allelic differences to the mRNA diversity, RNA was prepared from thorax tissue of single bees for Northern blot experiments. Nine blank and nine induced (7 h post-induction) bees were studied (see Figure 5). Comparative analysis of mRNA profiles between individual, 'immune' bees clearly indicated five different patterns which could be classified into two groups: those with two apidaecin mRNA species of different length and a second group showing only one RNA band but differing in length. The lengths of all these transcripts varied between 880 and 1150 nt (corresponding to ~6−9 units, respectively). In addition, extended exposure of the film revealed small transcripts of as little as 410 nt and an additional band of 2400 nt hybridizing with the apidaecin probe (data not shown). Comparison of apidaecin expression (band E) in thorax tissue (sample 13) and abdomen tissue at the same timepoint (7 h PI) indicated a 9 times greater apidaecin expression for the abdomen than for the thorax (data not shown).

Primer extensions. With variability of unit numbers in the apidaecin mRNA population established, possible additional differences at the 5′ end were analyzed. Primer extension experiments using unfractionated mRNA, pooled from several insects and with the 5′B6 oligonucleotide as a primer, were carried out to this end. A single band of ~62 nt was observed after gel electrophoresis (data not shown). Because this primer overlaps with the coding sequence for only 6 nt (Figure 1), the 5′ untranslated region of apidaecin transcripts is likely to be 56 nt long.

Kinetics of apidaecin mRNA and peptide induction

A group of 180 bees was induced by injection of *Escherichia coli* cells in the hemocoel. Groups of ~20 bees were taken at different time points (from 1 to 36 h PI) (Figure 4). Both apidaecin mRNA in abdominal cells and mature peptide level in lymph were determined, in parallel, by Northern blot analysis and by reversed-phase HPLC, respectively. Three groups of uninduced bees were used as controls. Northern hybridization revealed the presence of low amounts of apidaecin mRNA in control bees (Figure 4A). Indeed, analysis at the level of single insects and by using thorax mRNA from control bees (see Figure 5), showed low (lanes 1, 3, 4, 7; only detected at long exposure of the autoradiogram), medium (lanes 2, 5, 6, 9) or even high (lane 8) synthesis of apidaecin transcripts. Bees with low apidaecin expression were considered as true blank; bees with higher expression had probably been induced by previous accidental wounding and/or infection in the wild. This also explains the presence of rather high levels of apidaecin mRNA in blank bees when a pool of bees was used for study.

Following induction, the levels of apidaecin transcripts increased. All major bands (from A to E) were scanned by computer and IOD values were determined (see Materials and methods). These values were normalized based on the IOD values of the rRNA bands detected on the corresponding agarose gel. The induction factor for each band as a function of time is given in Figure 4B. While only marginal levels of one apidaecin mRNA species (labeled E in Figure 4A) were observed in 'blank' insects, this transcript increased steadily and significantly after the 3 h timepoint. In addition, this particular band became very broad on the autoradiogram

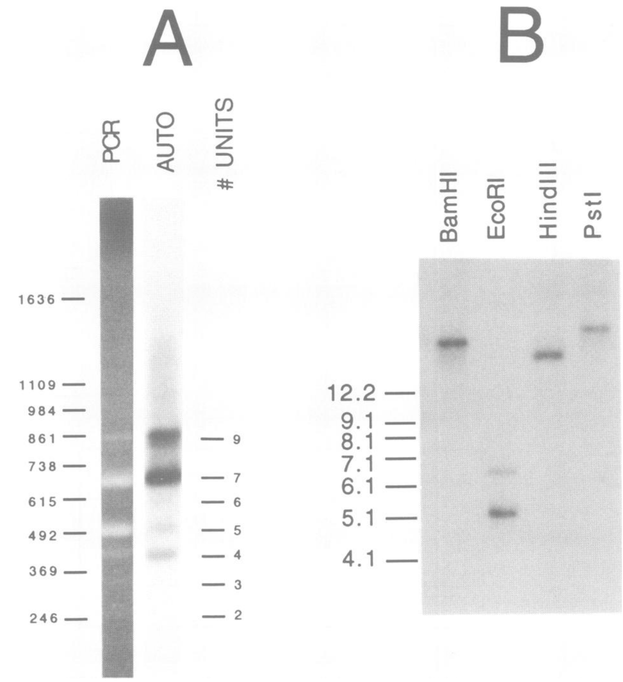


Fig. 6. Genomic information of apidaecin. (**A**) PCR amplification of genomic sequences. 0.05 μg of genomic DNA was amplified in 50 μl using 5′SB6-2 and 3′B6 as described by Kainz *et al.* (1992) using the Hot Tub polymerase (Amersham). The DNA was first denatured for 3 min at 93°C followed by 30 two-step amplification cycles (2 min at 63°C followed by 4 s at 93°C). The procedure was terminated by a long polymerization reaction of 15 min at 63°C; 20 μl was subsequently analyzed on a 1.5% agarose gel (labeled PCR), transferred to a Hybond-N+ membrane and hybridized using the inner *Hha*I fragment as a probe (labeled AUTO). (**B**) DNA blot hybridization of *Apis mellifera* (bee) genomic DNA. *Apis* DNA (5 μg) was digested with different restriction enzymes as indicated above each lane. Nick-translated pApid73 was used as probe. Numbers indicate mobilities of molecular size markers in kilobases.

with increasing timepoints, ending in a smear 12 h PI and beyond; this result indicated increased induction of smaller apidaecin transcripts over time. Moreover, at least four additional hybridizing bands also became visible after 4−12 h PI and most of them continued to gain intensities until 36 h PI. The predominant apidaecin mRNA species (E) increased by a factor of 17 after infection. However, induction factors in thoraxes of single bees could reach values of up to 600 times the lowest detectable levels of apidaecin transcription before infection.

Induction at the level of mRNA was nicely paralleled by the appearance of increased amounts of mature apidaecin peptide in the hemolymph of infected bees (Figure 4B). As was the case with at least two mRNA species, significant amounts of apidaecin peptide could also be found in control bees (30 μg/ml). Apidaecin concentrations were determined based on the peak heights during HPLC analysis (see Materials and methods). An apparent induction of the apidaecin peptide was detected 12 h PI (4 times) and increased >10 times at 36 h PI; peak concentrations in bee lymph reached up to 360 μg/ml.

Genomic organization of apidaecin

The marked variability in length (number of units) notwithstanding, apidaecin transcripts showed a very high degree of sequence identity. Since allelic differences could not fully account for the presence of the observed heterogeneity, other likely explanations could be: (i) the existence of a multigene family; (ii) differential splicing of large transcripts from a single gene; or (iii) a combination of the two. Therefore, a preliminary study was initiated to define the chromosomal organization of the apidaecin gene(s). Genomic DNA was prepared from 10 bees. A PCR experiment was performed with optimized conditions for the amplification of large fragments (Kainz *et al.*, 1992). Amplification reactions were started using oligonucleotides located in unique 5′ and 3′ flanking sequences: 5′SB6/3′B6 and 5′SB6-2/3′B6, respectively (see Figure 1). One experiment yielded a ladder of amplified fragments ranging from 250 to 861 nt in length (Figure 6A). These fragments could encode apidaecin precursors containing 2–9 units. The corresponding blot was then probed with a labeled inner fragment (repetitive unit sequence only). All bands were revealed on the resulting autoradiogram; those representing genes with 7 and 9 units gave the strongest signal. Longer exposures also indicated the presence of a smaller band representing a gene with a single unit (data not shown). A near identical banding pattern was found when similar PCR conditions were used with template DNA prepared from single bees (data not shown). These results suggest that different apidaecin genes were present. To determine whether these genes are dispersed over the chromosome or whether they are clustered, Southern blots were performed using pApid73 as a probe (Figure 6B). This hybridization revealed that apidaecin genes are located within a 15 kb *Hin*dIII fragment. *Eco*RI cuts this apidaecin gene cluster into two fragments of 5.1 and 6.4 kb. Hybridization using probes specific to the unique 3′ and 5′ flanking sequences gave similar results as with the whole apidaecin probe (data not shown). This indicated that the both *Eco*RI fragments contain entire apidaecin genes and confirmed the presence of several clustered genes.

Discussion

Multipeptide apidaecin precursors

Apidaecins are the major immune components induced in honeybees upon bacterial infection. Characterization of 10 different cDNA clones revealed dramatic differences in the length of the open reading frames and the corresponding apidaecin precursor proteins. More specifically, the three clones that were sequenced in their entirety encoded precursor molecules of 144, 168 and 289 amino acids. With the N-terminal (first 60 amino acids) and C-terminal (last 24 amino acids) ends of the open reading frame being strictly conserved among all 10 clones, differences in length were correlated to a variable number of internal repeats. Each such repeated unit consists of a mature apidaecin peptide (18 aa) preceded by an acidic spacer region (6–8 aa) and an Arg-Arg dipeptide, which amounts to a total unit length of 26–28 amino acids.

Analysis of the mRNA population of 'immune' bees by Northern blotting and PCR revealed that the majority of apidaecin transcripts were in the 420–1400 nt size range, thereby differing from one another by about 80 nt, the size of a single unit. While at the transcript level the number of units varied from 1 to 12, in single unit intervals, only 4, 5, 7, 9 and 12 units were observed in the cDNA clones. This is likely due to the small sample size (10 clones) and may reflect a prevalence of certain repeat numbers (5 and 7) among transcripts at 6 h PI.

In any event, as judged from the cDNA structures, apidaecin precursors contain multiple copies of the mature, biologically active peptides. The multi-unit part of the apidaecin precursor is preceded by a more common prepro structure at the N-terminus. Two possible signal sequences were found, consisting of 16 or 19 amino acids. The pro-fragment (13–16 aa) of the apidaecin precursor is hydrophilic and is followed by an Arg-Arg dipeptide sequence (part of unit 1).

This particular multipeptide precursor structure is reminiscent of the one described for yeast alpha-mating factor (Singh *et al.*, 1983). Detailed studies of this precursor molecule have revealed that the mature component is released by a three-step mechanism: first, endoprotease 'Kex2' cleaves at the carboxy end of the basic dipeptide(s) (Fuller *et al.*, 1989); second, carboxypeptidase 'Kex1' successively removes both basic amino acids, resulting in spacer-mature peptides (Dmochowska *et al.*, 1987). The latter are then, in turn, processed by a dipeptidyl aminopeptidase that specifically removes X-A dipeptides, also in a stepwise manner (Julius *et al.*, 1983).

By virtue of their near identical precursor assemblies, we speculate that mature apidaecins are produced by a processing system comparable to the one established for the yeast pheromone. This hypothesis is substantiated by some experimental evidence. Indeed, apidaecin precursors, isolated from 'immunized' bee larvae, consisted of the mature peptide preceded by the spacer sequence EAKPEAKP (Casteels *et al.*, 1989) or EAEPEAEP (P.Casteels, unpublished). Intermediates containing six (EPEAEP), four (EAEP) or two (KP) extra amino acids at the N-terminus have also been isolated since then (P.Casteels, unpublished). The presence of these additional amino acids all but abolished antibacterial activity (P.Casteels and P.Tempst, unpublished). Taken together, these findings confirm that dipeptidyl amino-peptidase processing is a final and essential step during apidaecin maturation. It is noteworthy that the occurrence of this enzyme in honeybees had already been documented in the context of melittin processing (Kreil *et al.*, 1980).

Most insect antibacterial peptides are synthesized as long precursor molecules. Kex2-like proteases are known to function in the maturation process of attacin and defensin (Kockum *et al.*, 1984; Matsuyama and Natori, 1988; Dimarcq *et al.*, 1990) and dipeptidyl aminopeptidases in the processing of cecropin, diptericin, sarcotoxin IIA and melittin precursors (Kreil *et al.*, 1980; Ando and Natori, 1988; Dickinson *et al.*, 1988; Kylsten *et al.*, 1990; Wicker *et al.*, 1990; Gudmundsson *et al.*, 1991). However, to our knowledge, the existence of multipeptide precursors has neither been reported nor postulated. Thus, apidaecin precursor structure and maturation are unique among insect antibiotics.

However, the multipeptide aspect of precursor proteins has also been observed for antibacterial peptides secreted by frog skin. Magainin (Zasloff, 1987; Terry *et al.*, 1988), xenopsin (Sures and Crippa, 1984), caerulein (Richter *et al.*, 1986) and levitide (Poulter *et al.*, 1988) also require endo- and exoproteases for maturation, but unlike apidaecins, processing results in the formation of additional short, acidic peptides (Terry *et al.*, 1988) or an amphipathic peptide (Poulter *et al.*, 1988) of an as yet unknown function.

Apidaecin transcript and peptide diversity

Four amino acid substitutions were detected in mature apidaecins, effectively yielding four different isoforms. Three of these peptides have been previously isolated and were designated Ia (V6, I18), Ib (V6, L18) and II (I6, L18) with relative distribution in 'immune' bee lymph of 5% (Ia), 80–90% (Ib) and 5–15% (II) (Casteels *et al.*, 1989; P.Casteels, unpublished). In the current study, we find that all three isoforms are represented in single precursor molecules (e.g. clone 14). However, some precursors contain only two isoforms. In general, isoform ratios predicted from cDNA clone analyses correspond well to the peptide distribution in the lymph.

Although cDNA sequence data suggest the existence of a fourth apidaecin isoform, with Ser at position 9, no such peptides have ever been detected in bee lymph, let alone shown to be active. It may not be coincidental that the precursor structures of exactly these peptides (clone 73, units 3 and 5) contain Leu in their spacers, at position 4, which may preclude maturation into bioactive molecules.

Apidaecin induction

A sharp increase in apidaecin transcript levels occurred 4–6 h after experimental infection, followed by a steady rise for several more hours. An induction factor of 13 (for the major band) was detected after 12 h PI, high concentrations were then sustained throughout the entire 36 h monitoring period. Similar induction kinetics were observed at the peptide level, albeit with a delayed (8 h) onset; peak concentrations in bee lymph reached 360 μg/ml at 36 h PI. It should be mentioned that peptide levels at a certain time PI depend on physical conditions of the bees, such as age and stress (F.Jacobs, personal communication). Since minimal apidaecin concentrations for growth inhibition of susceptible bacteria in rich medium are < 1 μg/ml (Casteels *et al.*, 1989), 360 μg/ml could be considered a seriously excessive response. Nevertheless, low numbers of viable bacteria (100 cells/μl) were persistently present in 'immune' honeybee lymph, even several days PI (Casteels *et al.*, 1988). While this is a rather puzzling observation, it could help explain the continued presence of apidaecin transcripts over a fairly long period of time. Indeed, there is mounting evidence that transcriptional activation of insect peptide antibiotic genes is the end-point of a signaling pathway that has bacteria, or more specifically lipopolysaccharides (LPS), as initiating agents (Reichhart *et al.*, 1992; Sun and Faye, 1992). Thus, lingering bacteria in the honeybee hemocoel and/or tissue may keep apidaecin transcription in the 'on' state for days.

Similar long-term gene activations have been observed for giant moth (*Hyalophora, Manduca*) antibacterial peptides such as cecropins A, B and D (Dickinson *et al.*, 1988; Gudmundsson *et al.*, 1991) and attacins (Sun *et al.*, 1991b). In contrast, transcription of *Drosophila* cecropin (Kylsten *et al.*, 1990) and diptericin (Wicker *et al.*, 1990), flesh fly sarcotoxin Ia/cecropin (Matsumoto *et al.*, 1986) and black blowfly diptericin and defensin (Dimarcq *et al.*, 1990) was completely turned off well before 24 h PI. This suggests a different regulatory mechanism and/or a stronger bactericidal capacity for these peptides. Alternatively, swift disappearance of transcripts may be the result of compromised mRNA stability. It has been speculated that an [ATTTA] motif in the 3'-untranslated region could be indicative for rapid mRNA turnover (Shaw and Kamen, 1986). It is somewhat unexpected then, that we also find this motif in apidaecin transcripts. The longevity of the apidaecin response would then be solely explained by assuming continued gene activation.

Larger size mRNA species of 1.7, 2.4, 3.3 and 4 kb, hybridizing to the apidaecin probe, were coordinately induced after infection. We have no structural information on these larger RNA molecules as yet, but they may represent incompletely spliced intermediates. Similarly large transcripts have also been reported for the insect immune peptides diptericin, defensin and hemolin (Dimarcq *et al.*, 1990; Wicker *et al.*, 1990; Ladendorff and Kanost, 1991).

Also surprising were the unmistakable differences in apidaecin transcript profiles of single bees; at least five distinct patterns were observed in thoraxes of individual insects. To our knowledge, no such individually specified transcription has ever been observed for other antibacterial peptides.

Origins and implications of apidaecin multiplicity

Our results indicate that multiple apidaecin genes are clustered within a 15 kb region of the honeybee genome. Following experimental infection, there may be differential activation of these genes, both in terms of timing and quantitative aspects of transcription. A precedent for this kind of genomic variation has been described for the yeast alpha-mating factor where two genes, coding for precursor proteins with two or four units, were isolated (Singh *et al.*, 1983). Added diversity of apidaecin transcripts could possibly result from allelic differences and/or differential splicing.

It is a legitimate assumption that the apidaecin multigene family and the variable multipeptide character of the precursors are the result of gene amplifications and rearrangements. The potential evolutionary benefits from such events could be (i) generation of novel polypeptides with properties advantageous to the long-term survival and prosperity of the species and (ii) a quantitative activity effect through increased and/or more rapid production of the gene products. As regards the first possibility, no functional differences between the three characterized apidaecin isoforms have been observed so far (Casteels *et al.*, 1989). There is a remote possibility that the predicted, fourth isoform would possess either improved or other, useful activities. Probing the function awaits chemical synthesis of this hypothetical peptide.

Finally, the multigene and multipeptide precursor aspects of the apidaecin genetic information seem to result in an enormous production of mature peptide some 24–48 h after infection, rather than ensuring a very swift response. In contrast to the 3–4 h it takes for apidaecin transcripts to appear following infection, far more rapid inductions (1 hour) have been reported for *Drosophila* cecropins (Kylsten *et al.*, 1990). However, for those antibacterials the classical rule of 'one transcript—one peptide' applies, whereas apidaecin production is amplified. Thus, enhanced efficiency of the honeybee immune response to infection, through a mechanism of multipeptide antibiotic precursors, seems to be unique among insects.

Materials and methods

Immunization and large-scale RNA isolation of honeybees

Humoral immunity of 440 honeybees was induced, by inoculation with live *E.coli* cells, as described by Casteels *et al.* (1989). Total RNA from bees

(6 h PI) was extracted using guanidinium thiocyanate as described by Sambrook *et al.* (1989). After crushing in a mortar under N_2, the cells were lysed in 16 ml homogenization buffer and mixed using an Ultra-Turrax T25 (Ika-Works Inc., Cincinnati, OH) with a G18 probe for 2 min at 13 500 r.p.m. Large contaminating structures were removed by low-speed centrifugation (3000 r.p.m. for 10 min) and small chitin parts were eliminated by fiberglass filtration. Centrifugation at 16 000 r.p.m. during 20 min cleared the solution completely. This suspension was then loaded onto a CsCl cushion; further cleanup of the RNA pellet was as previously described (Sambrook *et al.*, 1989). Half of the obtained RNA sample (total 2.2 mg) was taken for mRNA preparation, using the PolyATtract™ mRNA isolation system (Promega, Madison, WI). From this step 16 μg of mRNA were recovered.

Isolation and sequencing of apidaecin cDNA clones

Five μg of mRNA were used for construction of the immune bee cDNA library with the ZAP-cDNA™ synthesis kit (Stratagene, San Diego, CA). DNA was packaged into phage using the GigapackII Gold packaging extract (Stratagene). The titer, after amplification, was 2.9×10^9 p.f.u./ml of recombinant phages and 4.1% of background phages; 3×10^5 p.f.u. were plated and screened for apidaecin cDNA clones with a degenerate oligonucleotide (B6-1) of 29 bases in length [5′-GG$^C/_T$TGIGG$^G/_A$AT$^G/_A$ TAIACIGGICG$^G/_A$TT$^G/_A$TT-3′ (see Figure 1)] that was labeled by kinasing with [^{32}P]ATP (ICN Biomedicals, Inc., Costa Mesa, CA) (Sambrook *et al.*, 1989). This oligonucleotide was designed, based on the known peptide sequence (Casteels *et al.*, 1989) and using the codon usage frequencies of *Drosophila*. Those frequencies were obtained by applying the algorithm from the GCG software package (University of Wisconsin, Madison, WI) to a combined number of 40 kb of *Drosophila* cDNA sequences. The hybridization was carried out in sodium chloride/sodium citrate as described by Duby *et al.* (1988) at 40°C with 10^6 c.p.m./ml of kinased probe. At least 74 positive clones were found in the first screen. After three rounds of plaque purification, 10 cDNA clones were excised and used for further analysis. Clones were sequenced by the dideoxy method (Sambrook *et al.*, 1989) using the T7Sequencing™ kit (Pharmacia, Piscataway, NJ). Exonuclease deletions of pApid73 were made using the pBluescript ExoII/mung bean nuclease DNA sequencing kit (Stratagene). Subclones pApid73.13 and pApid73.1 were constructed by deleting the *Sty*I−*Kpn*I and *Bam*HI−*Sty*I fragment, respectively, from the pApid73 clone.

Synthetic oligonucleotides

Oligonucleotides were made at the MSKCC Microchemistry core facility using a model 394 ABI (Foster City, CA) automated instrument. 3′SB6, 5′SB6 and 5′SB6-2 are sense oriented, 3′B6, 5′B6 and 5′B6-2 anti-sense (Figure 1). Sequences were as follows: 5′SB6, 5′-GGTGTGGGTTGAA-TAACTATTGATAATCATGAAG-3′ (34 nt); 5′B6, 5′-CTTCATGAT-TATCAATAGTTATTCAACCCACACC-3′ (34 nt); 5′SB6-2: 5′-CCAA-CCTAGATCCGCCTACTCGACCT-3′ (26 nt); 5′B6-2, 5′-AGGTCGA-GTAGGCGGATCTAGGTTGG-3′ (26 nt); 3′SB6, 5′-GAAGAATATG-AAGCACGTGAAATA-3′ (24 nt); 3′B6, 5′-TATTTCACGTGCTTCA-TATTCTTC-3′ (24 nt).

Small-scale RNA isolation and Northern blotting

RNA was purified as described by Dorsett *et al.* (1989) with the following additional changes. Urea buffer was added to the frozen abdomens and mixed with the Ultra Turrax for 2 min at room temperature, sodium dodecyl sulphate was immediately added to a final concentration of 2%. For much smaller RNA preparations, such as from individual thoraxes, tissues were crushed with a small pestle in an Eppendorf tube under dry ice conditions and urea buffer was subsequently added; further processing was as above. RNA gels, blotting and hybridization techniques were as described by Dorsett *et al.* (1989) except that Hybond-N+ membranes (Amersham Corp., Arlington Heights, Il) were used and blotting was done for 2 h with 0.05 M NaOH. RNA probes were made from the *Pst*I-linearized pApid73 clone using T7 RNA polymerase (New England Biolabs, Beverly, MA) incorporating [^{32}P]UTP (NEN Research products, Boston, MA; NEG007X). The (pre)hybridizations and washes of blots were performed at 65°C.

Quantitation of RNA from gels and autoradiographs

Quantitation and normalization of RNA levels in Northern blots were performed as previously described with modifications (Masters *et al.*, 1992). Hybridized messenger RNA levels were determined from autoradiographs and normalized to ribosomal RNA determined from ethidium bromide-stained gels. Briefly, images of electrophoresis gels were obtained by digitizing autoradiographs and photographs of ethidium bromide-stained gels using the charged coupled device (CCD) camera of the BioImage Visage 110 computerized imaging system from Millipore BioImage Products (Ann Arbor, MI).

Gels were then analyzed using the BioImage Whole Band Analysis subroutines. Lanes were defined, bands quantified and integrated optical density (IOD) values determined. Several exposures of each gel were evaluated and band quantitation was performed on lanes that did not exhibit film saturation. When it was necessary to compare lanes from films exposed for different lengths of time, normalization was performed using control lanes that were not saturated on either film. Ribosomal RNA bands were analyzed in a similar manner, except that the gray scale values of the image were inverted using the Image Processor subroutine prior to quantitation.

DNA isolation and Southern blot analysis

Nuclei were isolated from ground thorax tissue as described by Greenberg and Bender (1990) and were subsequently lysed and cleaned according to the procedure of Walldorf *et al.* (1984). The suspension was clarified by a short centrifugation step at low speed and DNA was precipitated using 10 vols of 70% ethanol. Southern hybridizations were performed at high stringency conditions (hybridization at 42°C and wash at 65°C) using 50% formamide as described by Sambrook *et al.* (1989). The Apid73 probe was labeled via oligolabeling; the *Hha*I probe was made from the pApid73 clone (see Figure 1). 5′ and 3′ probes, used for oligo-hybridizations at 30°C as described by Duby *et al.* (1988), were made using PCR as described in the next section.

PCR procedures and analysis

PCR reactions were carried out using the GeneAmp PCR reagent kit as described by the supplier (Perkin Elmer Cetus, Norwalk, CT) using the GeneAmp PCR system 9600 (Perkin Elmer Cetus). $MgCl_2$ concentration was always kept at a final concentration of 1.5 mM. The 5′-probe for Southern blot analysis was made via PCR using M13 reverse primer and 5′B6-2 primer with 50 ng of linearized pApid14 as template; the 3′ probe was made similarly using 3S′B6 and T7-primers. Both M13 and T7 oligos were provided via the Pharmacia sequenase kit. The PCR reaction was initiated with 4 μl of [^{32}P]dCTP (NEN, NEG013H) and 0.5 mM cold dCTP. PCR reactions were started with one cycle at 95°C for 5 min, 37°C for 1 min and 72°C for 2 min, followed by 30 cycles at 95°C for 2 min, 37°C for 1 min, at 72°C for 2 min and a final cycle at 95°C for 2 min, 37°C for 1 min and 72°C for 10 min. A 1 μl aliquot was analyzed on an 8% polyacrylamide gel (Sambrook *et al.*, 1989) which confirmed the length of the expected probe.

Primer extensions

Primer extensions on induced bee RNA were done as described by Sambrook *et al.* (1989) using the 5′B6 primer. The resulting mixture was analyzed on a 8% urea−polyacrylamide gel together with a kinased 123 bp and a 1 kb DNA ladder (Gibco BRL, Burlington, Ontario).

Quantitative peptide analysis

Lymph from control and *E.coli*-injected insects was harvested at several timepoints by puncturing the abdomen with a glass capillary, aspirating the hemolymph and clearing, heat-treating and acidifying as described previously (Casteels *et al.*, 1989). Diluted lymph samples were then taken for RP-HPLC analysis using a model 150A ABI (Foster City, CA) LC system with a Vydac C4 (214 TP54) 4.6 mm column (The Separations Group, Hesperia, CA). Solvents and gradient elution conditions were also as in Casteels *et al.* (1989). Quantitation of apidaecin levels in honeybee lymph, as a measure of peak height, was accomplished using a calibration curve obtained from HPLC analysis of a dilution series of synthetic apidaecin. The calibration was linear in the 0.05−50 μg/ml range, with a maximal error margin of 15%. The error margin was determined from calculation of peak height (mAU@214 nm) per μg over a 10 injection experiment.

Acknowledgements

The authors are grateful to Bob Cornetto for the generous gift of honeybees, to Natacha Muckova and Mary Lui for oligonucleotide synthesis, and to Dr Wayne Patton for help with scanning of autoradiograms and useful discussions. We thank Dr Martin Wiedman for critical reading of the manuscript. We acknowledge Mark Naughton and Yuri Hrynyszyn for the final prints of the figures. P.C. was supported, in part, by a NATO (Belgium) postdoctoral stipend; P.T. is indebted to the Society of MSKCC for financial support.

References

Ando,K. and Natori,S. (1988) *Biochemistry*, **27**, 1715–1721.

Boman,H.G., Boman,I.A., Andreu,D., Li,Z. Merrifield,R.B., Schlenstedt,G. and Zimmermann,R. (1989) *J. Biol. Chem.*, **264**, 5852–5860.

Boman,H.G., Faye,I., Gudmundsson,G.H., Lee,J.-Y. and Lidholm,D.-A. (1991) *Eur. J. Biochem.*, **201**, 23–31.

Bulet,P., Cociancich,S., Dimarcq,J.-L., Lambert,J., Reichhart,J.-M., Hoffmann,D., Hetru,C. and Hoffmann,J.A. (1991) *J. Biol. Chem.*, **266**, 24520–24525.

Casteels,P., Van Steenkiste,D. and Jacobs,F. (1988) In Cavalloro,R. (ed.), *European Research on Varroatosis Control.* A.A. Balkema, Rotterdam, pp. 105–111.

Casteels,P., Ampe,C., Jacobs,F., Vaeck,M. and Tempst,P. (1989) *EMBO J.*, **8**, 2387–2391.

Casteels,P., Ampe,C., Riviere,L., Van Damme,J., Elicone,C., Flemming,M., Jacobs,F. and Tempst,P. (1990) *Eur. J. Biochem.*, **187**, 381–386.

Casteels,P., Ampe,C., Jacobs,F. and Tempst,P. (1993) *J. Biol. Chem.*, in press.

Cavener,D.R. (1987) *Nucleic Acids Res.*, **15**, 1353–1361.

Dickinson,L., Russell,V. and Dunn,P.E. (1988) *J. Biol. Chem.*, **263**, 19424–19429.

Dimarcq,J.-L., Zachary,D., Hoffmann,J.A., Hoffmann,D. and Reichhart,J.-M. (1990) *EMBO J.*, **9**, 2507–2515.

Dmochowska,A., Dignard,D., Henning,D., Thomas,D.Y. and Bussey,H. (1987) *Cell*, **50**, 573–584.

Dorsett,D., Viglianti,G.A.,Rutledge,B.J. and Meselson,M. (1989) *Genes Dev.*, **3**, 454–468.

Duby,A., Jacobs,K.A. and Celeste,A. (1988) In Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (eds), *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley-Interscience, USA, Unit 6.4.

Faye,I. (1990) In Hoffmann,D. and Hoffmann,J.A. (eds), *34th Forum in Immunology—Cellular and Molecular Aspects of Insect Immunity. Res. Immunol.*, **141**, 927–932.

Fujiwara,S., Imai,J., Fujiwara,M., Yaeshima,T., Kawashima,T. and Kobayashi,K. (1990) *J. Biol. Chem.*, **265**, 11333–11337.

Fuller,R.S., Brake,A. and Thorner,J. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 1434–1438.

Greenberg,M.E. and Bender,T.P. (1990) In Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (eds), *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley-Interscience, USA, Unit 4.10.

Gudmundsson,G., Lidholm,D.-A., Asling,B., Gan,B. and Boman,H.G. (1991) *J. Biol. Chem.*, **266**, 11510–11517.

Hoffmann,D. and Hoffman,J.A. (1990) *34th Forum in Immunology—Cellular and Molecular Aspects of Insect Immunity. Res. Immunol.*, **141**, 895–960.

Julius,D., Blair,L., Brake,A., Sprague,G. and Thorner,J. (1983) *Cell*, **32**, 839–852.

Kainz,P., Schmiedlechner,A. and Strack,H. B. (1992) *Anal. Biochem.*, **202**, 46–49.

Kockum,K., Faye,I., Hofsten,P.V., Lee,J.-Y. Xanthopoulos,K.G. and Boman,H.G. (1984) *EMBO J.*, **3**, 2071–2075.

Kreil,G., Haiml,L. and Suchanek,G. (1980) *Eur. J. Biochem.*, **111**, 49–58.

Kylsten,P., Samakovlis,C. and Hultmark,D. (1990) *EMBO J.*, **9**, 217–224.

Ladendorff,N.E. and Kanost,M.R. (1991) *Arch. Insect Biochem. Physiol.*, **18**, 285–300.

Lee,J.-Y., Boman,A., Chuanxin,S., Andersson,M., Jornvall,H., Mutt,V. and Boman,H.G. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 9159–9162.

Lehrer,R.I., Ganz,T. and Selsted,M.E. (1991) *Cell*, **64**, 229–230.

Mahbubani,M.H., Bej,A.K., Miller,R.D., Atlas,R. M., DiCesare,J.L. and Haff,L.A. (1991) *BioTechniques*, **10**, 48–49.

Masters,D.B., Griggs,C.T. and Berde,C.B. (1992) *BioTechniques*, **12**, 902–911.

Matsumoto,N., Okada,M., Takahashi,H., Ming,Q.X., Nakajima,Y., Nakanishi,Y., Komano,H. and Natori,S. (1986) *Biochem. J.*, **239**, 717–722.

Matsuyama,K. and Natori,S. (1988) *J. Biol. Chem.*, **263**, 17117–17121.

Poulter,L., Terry,A.S., Williams,D.H., Giovannini,M.G., Moore,C.H. and Gibson,B.W. (1988) *J. Biol. Chem.*, **263**, 3279–3283.

Reichhart,J.-M., Essrich,M., Dimarcq,J.-L., Hoffmann,D., Hoffmann,J.A. and Lagueux,M. (1989) *Eur. J. Biochem.*, **182**, 423–427.

Reichhart,J.-M., Meister,M., Dimarcq,J.-L., Zachary,D., Hoffmann,D., Ruiz,C., Richards,G. and Hoffmann,J.A. (1992) *EMBO J.*, **11**, 1469–1477.

Richter,K., Egger,R. and Kreil,G. (1986) *J. Biol. Chem.*, **261**, 3676–3680.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Shaw,G. and Kamen,R. (1986) *Cell*, **46**, 659–667.

Singh,A., Chen,E.Y., Lugovoy,J.M., Chang,C.N., Hitzeman,R.A. and Seeburg,P.H. (1983) *Nucleic Acids Res.*, **11**, 4049–4063.

Sun,S.-C. and Faye,I. (1992) *Eur. J. Biochem.*, **204**, 885–892.

Sun,S.-C., Asling,B. and Faye,I. (1991a) *J. Biol. Chem.*, **266**, 6644–6649.

Sun,S.-C., Lindstrom,I., Lee,J.-Y. and Faye,I. (1991b) *Eur. J. Biochem.*, **196**, 247–254.

Sures,I. and Crippa,M. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 380–384.

Terry,A.S., Poulter,L., Williams,D.H., Nutkins,J.C., Giovannini,M.G., Moore,C.H. and Gibson,B.W. (1988) *J. Biol. Chem.*, **263**, 5745–5751.

Von Heijne,G. (1986) *Nucleic Acids Res.*, **14**, 4683–4690.

Walldorf,U., Richter,S., Ryseck,R-P., Steller,H., Edstrom,J.E., Bautz,E.K.F. and Hovemann,B. (1984) *EMBO J.*, **3**, 2499–2504.

Wicker,C., Reichhart,J.-M., Hoffman,D., Hultmark,D., Samakovlis,C. and Hoffmann,J.A. (1990) *J. Biol. Chem.*, **265**, 22493–22498.

Zasloff,M. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 5449–5453.

Received on October 30, 1992; revised on December 16, 1992