EuPA OPEN PROTEOMICS 3 (2014) 195–205

ScienceDirect

journal homepage: http://www.elsevier.com/locate/euprot

Identification of two capa cDNA transcripts and detailed peptidomic characterization of their peptide products in *Periplaneta americana*

*Susanne Neupert, Christian Derst, Sebastian Sturm, Reinhard Predel**

University of Cologne, Biocenter Cologne, Zuelpicher Strasse 47b, 50674 Cologne, Germany

ARTICLE INFO

Article history:
Received 19 October 2013
Received in revised form
20 December 2013
Accepted 14 February 2014
Available online 25 February 2014

Keywords:
Insect neuropeptide
Periviscerokinin
Pyrokinin
Single cell peptidomics
Alternative splicing
Tribolium

ABSTRACT

The first CAPA peptides of insects were identified from *Periplaneta americana* using extracts of 8000 abdominal perisympathetic organs. The corresponding gene of cockroaches and other basal insects, however, remained undiscovered. Here we identified two *capa* cDNA transcripts of *P. americana*. Single cell mass spectrometry of *capa* neurons in abdominal ganglia yielded a complete coverage of the precursor sequences of both transcripts; with intermediates of precursor processing more prominent in somata than in perisympathetic organs. The processing of the CAPA precursor in *P. americana* is compared with that in the beetle *Tribolium castaneum* which was re-evaluated in this study.

1. Introduction

CAPA peptides are typical of the neurosecretory system of the abdominal ventral nerve cord (VNC) of insects [1]. For several insects it has been shown that these peptides have strong myostimulatory or diuretic/antidiuretic effects. Osmoregulatory functions were reviewed in recent publications (e.g. [2,3]); here we present a few general findings related to CAPA peptides in a chronological order.

CAPA peptides of abdominal ganglia are hormones that are stored in abdominal perisympathetic organs (PSOs) of insects. These neurohemal organs perform the same function for the neurosecretory material of the VNC as the corpora cardiaca for neurosecretions of the brain. Since CAPA peptides are major compounds stored in abdominal PSOs [4], functions of these peptides and abdominal PSOs cannot be separated from each other. PSOs were first described from the stick insect *Carausius morosus* [5] and subsequently found in most other insect taxa [6]. Our current knowledge suggests PSOs as autapomorphy of Pterygota but similar neurohemal release sites were also found in silverfish (Zygentoma; Eckardt, Diesner, Predel, unpublished). Interspecific cardioaccelerating as well as diuretic and antidiuretic effects of extracts of abdominal PSOs have been described already in the 60 s of the last century (summarized in [7]).

Tublitz et al. focussed on the cardio-acceleratory effects of extracts from abdominal ganglia/abdominal PSOs in *Manduca sexta* and separated and characterized several factors they called cardio-acceleratory peptides (CAPs, [8]). The current designation CAPA peptides reflects this history. CAPs were not all necessarily CAPA peptides as we define them

* *Corresponding author*. Tel.: +49 221 4705817.
E-mail address: rpredel@uni-koeln.de (R. Predel).

http://dx.doi.org/10.1016/j.euprot.2014.02.005

today but a single fully identified peptide ($CAP2_b$; [9]) was later recognized as a product of the *M. sexta capa* gene [10].
The first sequenced CAPA peptide of insects was described as periviscerokinin (PVK) from the cockroach *Periplaneta americana* [11]. At that time, isolation and bioassays necessitated an extract of 8000 abdominal PSOs. The same extract was later used to identify five additional CAPA-related peptides; including the first CAPA pyrokinin (PK, [12]).
The first *capa* gene was reported from *Drosophila melanogaster* [13]; subsequent peptidomic analysis revealed a differential distribution of *capa* products in the CNS [14].
Differential sorting and packaging mechanisms of CAPA peptides were found for *capa* interneurons and neuroendocrine cells in *P. americana* [15]. In abdominal PSOs, CAPA peptides are co-localized in translucent vesicles; this was found in *P. americana* as well as in *D. melanogaster* [16,17].
Receptors for *Drosophila* CAPA peptides were described; these data confirmed that CAPA-PVKs (CAPA-1 and 2 in *D. melanogaster*) activate receptors [18,19] different from that of CAPA-PK (CAPA-3 in *D. melanogaster*; [20]). CAPA-PVK receptor expression was found to be abundant in Malpighian tubules [21].
Diuretic actions of CAPA peptides were studied in detail; including signalling pathways (reviewed in [2,3]). Two *capa* genes were reported in *Rhodnius prolixus* [22,23]; products of both genes have been identified in neurohemal tissue of the abdomen [24].
CAPA peptides were used on a large scale to resolve phylogenetic relationships within insects [25]. These data showed, for the first time, the robust phylogenetic information typical of neuropeptides co-evolving with receptors.
Suska et al. [26] reported that prospective *capa* neurons in the VNC of *D. melanogaster* (Va-neurons) are initially generated in all segments but segment-specific programmed cell death in posterior segments [27] and development of alternative identity in anterior segments reduce the expression of the *capa* gene to the abdominal segments 2–4.

In *P. americana*, the sequence elucidation of PVK-1, -2 has been accomplished by Edman degradation of HPLC-separated bioactive substances [11,28]. In these experiments, bioactivity was monitored with a highly sensitive hyperneural muscle (= ventral diaphragm of cockroaches) assay [29]. Later on, ultrastructural studies indicated that transverse nerve fibres arborizing on the hyperneural muscle rather than abdominal PSOs themselves are release sites of CAPA peptides [16]. Thus, release of PVKs would first stimulate the hyperneural muscle and this stimulation ensures a fast distribution of released hormones in the abdominal cavity. As a result of this, the peptides reach the dorsal vessel more rapidly and, as it is also known from abdominal PSO extracts of *M. sexta* [8], stimulate heart activity. In *P. americana*, cardioacceleratory effects are accompanied by an activation of closer muscles at the origin of segmental vessels [28] and both effects increase an anteriorly directed hemolymph transport.

Following the structural elucidation of four additional putative CAPA peptides (VEA-OH, YLS-NH_2, SKN-OH, and CAPA-PK, originally published as Pea-PK-5) which did not show myotropic efficacy on isolated hyperneural muscle, foregut, hindgut, dorsal vessel, and oviduct [12,30], numerous CAPA peptides of different cockroaches were identified by mass spectrometry without monitoring bioactivity [25]. Single abdominal PSOs turned out to be sufficient for sequence analysis in many cases [25,31]. The *capa* gene of cockroaches and other basal insects, however, remained undiscovered. In this study, we reduce the gap and describe two *capa* cDNA transcripts of *P. americana*. This information was used to perform a comprehensive single cell analysis of CAPA precursor products which yielded a comprehensive list of mature CAPA peptides by using MALDI-TOF mass spectrometry. To exclude the description of only species-specific findings, we re-analyzed *capa* gene products in a holometabolous insect, *Tribolium castaneum*, and compare the products of precursor processing in both species.

2. Materials and methods

2.1. Insects

Cockroaches were raised under a 12 h light, 12 h dark photoperiod at a constant temperature of 28 °C. They were fed food pellets for rats and had free access to water. Adult *P. americana* of both sexes were used for experiments. Adult specimens of *T. castaneum* were taken from a culture at the Biocenter of the University of Cologne (Siegfried Roth).

2.2. Cloning of *P. americana capa* cDNA

A BLAST search identified a partial EST clone (cloneID C1171674) of *P. americana capa* cDNA. The missing 5′ end was amplified by a nested RACE-PCR using two gene specific primers and two adapter primers (see Supplemental Material for a list of primers and primer positions). RACE-PCR was performed using AdvantageTaq 2 polymerase mixture (Clontech Laboratories, Mountain View, CA) and a *P. americana* Marathon RACE library [32] according to the manufacturers protocol with 35 cycles of 1 min at 94 °C, 1 min at 62 °C, and 3 min at 68 °C. Four different primer combinations (PCR F1-R1, F1-R2, F2-R1 and F2-R2, see Supplemental Material) were used to amplify an alternatively spliced N-terminal region. For these experiments, PCR was performed using AdvantageTaq 2 polymerase mixture with 35 cycles of 1 min at 94 °C, 1 min at 60 °C, and 1 min at 68 °C and cDNA prepared from *P. americana* abdominal ganglia using RNeasy Mini Prep Kit (Qiagen, Hilden, Germany) and Sensiscript Reverse Transcription Kit (Qiagen). RACE and PCR products were cloned into pGEM-T vector (Promega GmbH, Mannheim, Germany) for sequencing.

2.3. Retrograde filling of *capa* somata

Single abdominal ganglia from cockroaches with attached abdominal PSOs and transverse nerves were dissected and placed in a small acrylic glass container containing insect saline (pH 7.25) of the following composition: NaCl (7.50 g/l), KCl (0.20 g/l), $CaCl_2$ (0.20 g/l) and $NaHCO_3$ (0.10 g/l). Distal end of severed transverse nerves were placed in a separate vial filled with 5% dextran-tetramethylrhodamine (MW 3000, anionic, lysine-fixable, Molecular Probes, Eugene, USA). The

bridge between the two chambers was covered with vasiline. Preparations were placed in a humid chamber and kept in darkness at 4 °C. Labelling of the capa neurons via passive diffusion of the fluorescence dye was finished after 12–16 h.

2.4. Tissue preparation and cell dissection for direct MALDI-TOF mass spectrometry

Abdominal PSOs of *P. americana* were dissected as described in [33]; CAPA-peptide release sites of *T. castaneum* were dissected with the help of a glass capillary as described in [34].

For single cell preparations, isolated abdominal ganglia of *P. americana* were fixed with microneedles and rinsed with insect saline. Subsequently, the ganglionic sheath was removed using ultra-fine scissors under a stereo fluorescence microscope (SteREO Lumar.V12, Carl Zeiss AG, Goettingen, Germany) equipped with a Lumar 43 HE Cy3 filter. Without enzyme treatment, identified cells were removed step-by-step using an uncoated glass capillary (Hilgenberg GmbH, Malsfeld, Germany) and transferred to a stainless steel sample plate for MALDI-TOF mass spectrometry. The insect saline was then removed using the same glass capillary and cells were allowed to dry at room temperature. Prior to matrix application, cells were rinsed with water to remove salt contamination.

2.5. Cell preparation for peptide derivatisation

For 4-sulfophenyl isothiocyanate (SPITC, Sigma–Aldrich, Steinheim, Germany) derivatisation experiment, a single *capa* cell identified by retrograde dye filling was dissected as described above. For single cell extraction, the cell was allowed to settle and to attach on the inner side of the glass capillary. Subsequently, insect saline was replaced by approximately 150 nl purified water. After an incubation time of 30 min on ice in a humid chamber, extracted peptides were used for N-terminal derivatisation.

2.6. Peptide derivatisation using SPITC

Based on a protocol described in [35], the single cell extract was transferred in 40 μl SPITC dissolved in 20 mM $NaHCO_3$ (pH 9.0) at a concentration of 15 mg/ml. The sulfonation reaction was carried out in a 0.5 ml Eppendorf tube for 1 h at 55 °C and 300 rpm. The sample was then acidified by adding 2.5 μl 10% acetic acid followed by 50 μl 0.5% acetic acid. This sample was loaded onto an activated and equilibrated home-made StageTip Empore 3 M (IVA Analysentechnik e.K., Meerbusch, Germany) C18 column [36]. The column was rinsed with 50 μl 0.5% acetic acid and peptides were eluted from the column with solutions of 10/20/30/40/50/80% acetonitrile (0.5% acetic acid) and spotted onto a MALDI target. Before drying, 0.3 μl matrix solution was added to each sample, respectively.

2.7. Reduction and alkylation of disulfide bonds

For alkylation experiments, five abdominal PSOs were extracted in 50 μl methanol/0.5% formic acid, sonicated and centrifuged for 15 min at 13.000 rpm at 4 °C. The supernatant was concentrated in a SpeedVac to a volume of about 2 μl and refilled to 25 μl with 50 mM ammonium bicarbonate buffer (ABC, pH 8.2). Disulfide reduction was performed with 200 mM dithiotreitol in ABC in a final concentration of 10 mM at 37 °C for 1 h. Alkylation of reduced cysteines was performed for 1 h at a concentration of 40 mM iodoacetamide dissolved in ABC at room temperature. Finally, 100 μl 0.5% formic acid was added and the alkylated sample was loaded onto an activated and equilibrated home-made StageTip Empore 3 M C18 column. Samples were eluted from the column with 1.5 μl 80% methanol/0.5% formic acid and directly spotted onto a MALDI target. Samples were mixed with same quantities of matrix solution before air drying at room temperature.

2.8. MALDI matrix application

10 mg/ml 2,5-dihydroxybenzoic acid (DHB, Sigma–Aldich, Steinheim, Germany) dissolved in 20% acetonitrile/1% formic acid was loaded onto the dried samples using a glass capillary (10–20 nl for single cell preparations) or 0.1–10 μl pipettes (0.3 μl for alkylated samples and direct tissue profiling). For an even distribution of DHB matrix crystals, the samples were dried using a regular hairdryer.

2.9. MALDI-TOF mass spectrometry

Mass spectra were acquired in positive ion mode on an UltrafleXtreme TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany). All MS acquisitions were made under manual control in reflector positive mode in a detection range of 600–20,000 Da. The instrument settings were optimized for the mass ranges of m/z 600–3500, 3000–10,000, 8000–15,000, and 15,000–20,000 and calibrated using Bruker peptide and protein standard kits, respectively. Laser fluency was adjusted to provide the optimal signal-to-noise ratio. The data obtained in these experiments were processed with the FlexAnalysis 3.4 software package. MS/MS was performed with LIFT technology. LIFT acceleration was set at 1 kV. The number of laser shots used to obtain a spectrum varied from 5000 to 25,000, depending on signal quality. Peptide identities were verified using MS/MS fragmentation of the molecules, determination of the molecular mass of the fragments, and comparison of predicted (http://prospector.ucsf.edu) and experimentally obtained fragmentation patterns.

3. Results

3.1. Cloning of a P. americana capa cDNA transcript

Sequence analysis of a partial EST clone with subsequent amplification of the missing 5′-region by RACE-PCR yielded a 977 bp cDNA fragment (GenBank accession number KF709654) which encodes the entire 217 AA open reading frame of a *P. americana* CAPA precursor sequence (Fig. 1). Analysis of this sequence indicated a 19 AA signal sequence [37]. The precursor also contains, among others, sequences of five neuropeptides that were identified in earlier studies. These sequences include two paracopies of PVK-1 [11], PVK-2 [28], CAPA-PK [12], YLS-NH_2 and VEA-OH [30]. Another peptide which was identified from abdominal PSOs of *P. americana*, SKN-OH (SDLTWTYQSPGDPTNSKN-OH; [30]), is represented

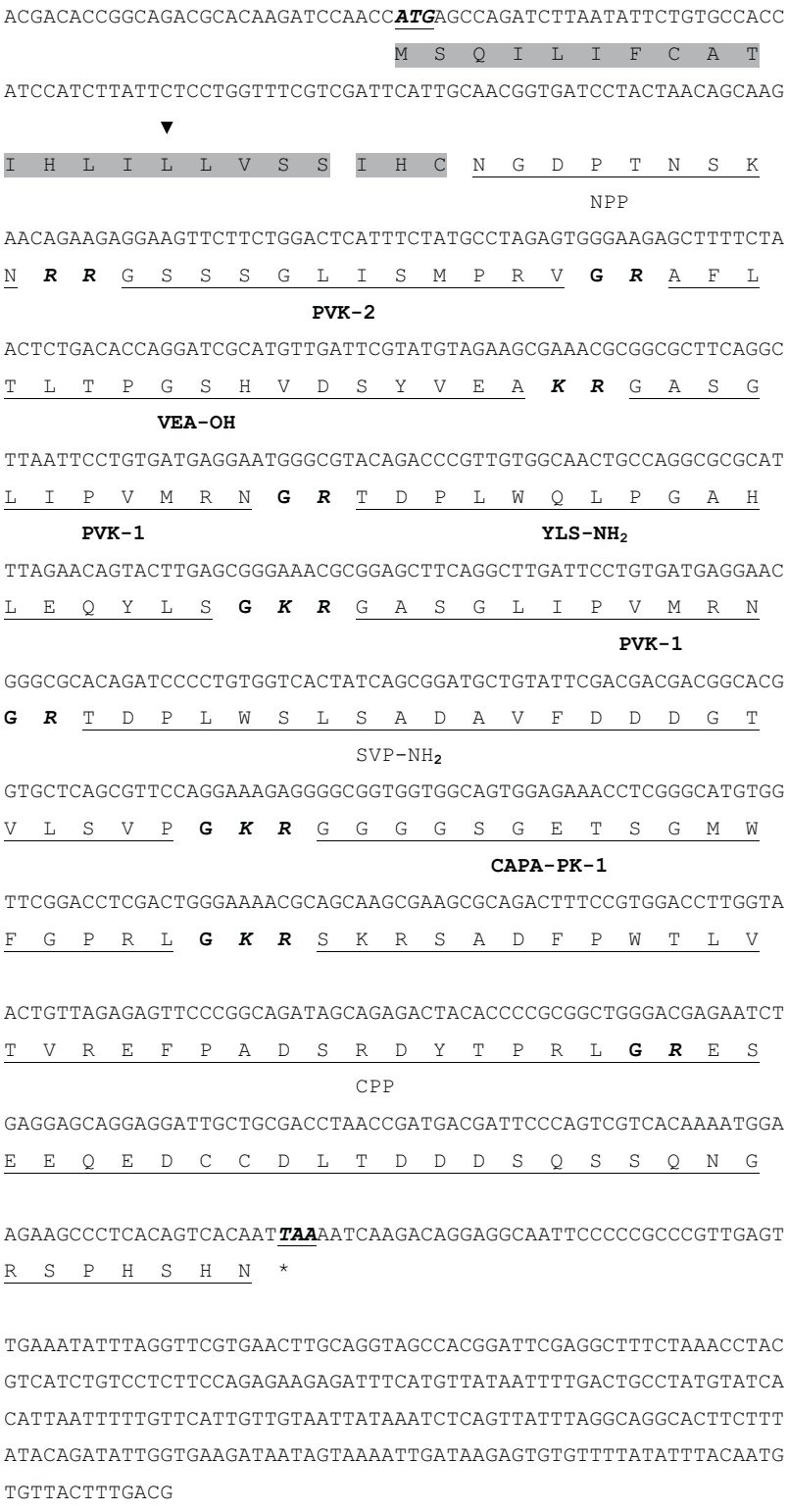

Fig. 1 – Nucleotide and AA sequences of a *P. americana capa* cDNA. Start and stop codons (italics, underscored, bold), signal sequence (highlighted in grey; 19 AAs as predicted by SignalP and additional IHC motif as suggested by peptidomics), amidation (bold)/cleavage (italics, bold) signals, and abundant mature peptides (underscored) as confirmed by mass spectrometry are indicated. Within CPP, a weaker secondary cleavage site exists, which results in low amounts of CAPA-PK-2 (CPP$^{1–28}$) and CPP$^{31–59}$. The arrowhead denotes the start of a partial EST clone; the missing 5′ end was amplified by a nested RACE-PCR. CAPA peptides identified from *P. americana* in earlier studies [12,28,30] are marked with bold letters. CPP, C-terminal precursor peptide; NPP, N-terminal precursor peptide.

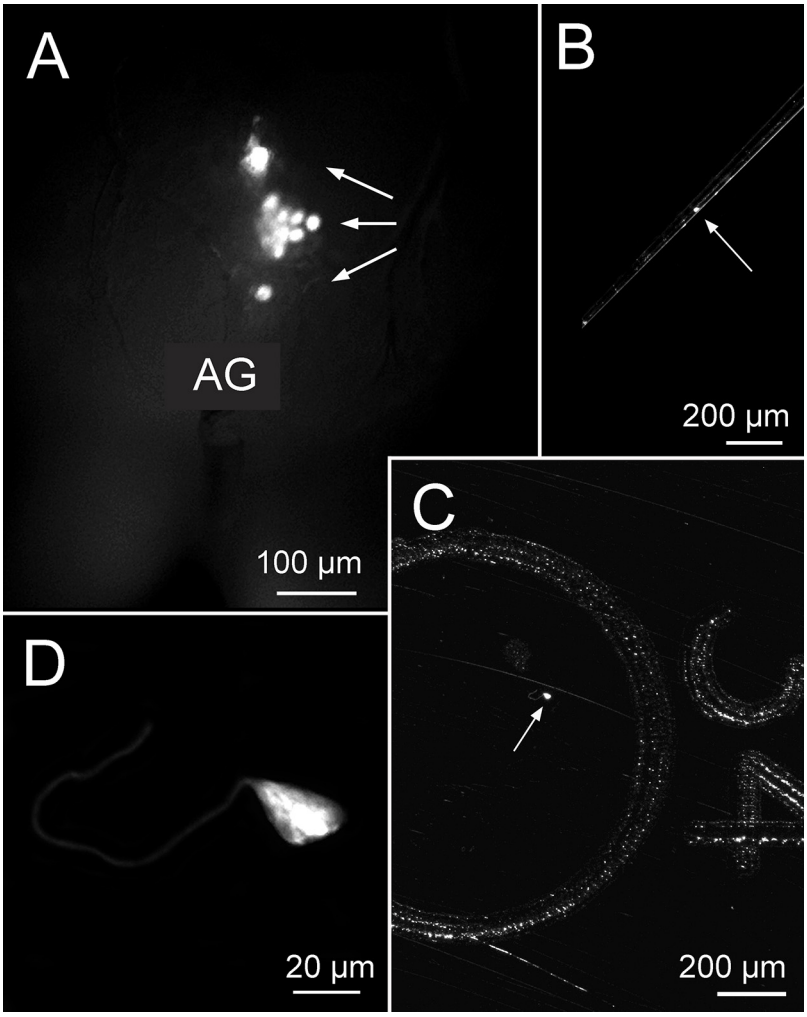

Fig. 2 – Dissection of a single *capa* neuron from abdominal ganglion 2 (Stereo-Lumar V12) following dextran-tetramethylrhodamine backfilling via transverse nerve. In abdominal ganglia posterior of transverse nerves/PSOs, only *capa* somata are stained (see [38]). (A) Overview of stained *capa* somata (arrows). (B) Transfer of an isolated *capa* neuron (arrow) by means of a glass capillary. (C) and (D) Documentation of the successful transfer of a *capa* neuron (arrow) with neurite onto the sample plate for MALDI-TOF mass spectrometry. AG, abdominal ganglion.

with only the eight C-terminal AAs in the N-terminal precursor peptide (NPP).

3.2. *Identification of capa gene products in single capa neurons of abdominal ganglia*

Capa neurons were visualized via retrograde filling of transverse nerves with fluorescence dye. Identified somata were dissected (Fig. 2) and prepared with DHB for single cell peptidomics. The detection range of *m/z* 600–20,000 was analyzed with four different settings; optimized for *m/z* 600–3500, 3000–10,000, 8000–15,000, and 15,000–20,000, respectively. Only mass spectra in the mass range *m/z* 600–8000 contained ion signals mass-identical with predictable peptides from the CAPA precursor and are presented here (Fig. 3A and B). In the mass range of *m/z* 600–2100, all peptides previously assigned as CAPA peptides were detectable by mass matches (Fig. 3A). Sequences of these peptides were confirmed by MS/MS experiments using the same samples, or using SPITC-derivatized peptides from an extract of a single *capa* cell (see Fig. 4 with MS/MS spectrum of CAPA-PK-1). Notably, the confirmed sequences included SKN-OH (*m/z* 2010.9), which sequence was only partially detectable in the identified CAPA precursor sequence and a third PVK (GSSGLISMPRV-NH_2; *m/z* 1102.7) which was first described from blaberid cockroaches as Lem-PVK-2 [31] and later reported to be present in abdominal PSOs of *P. americana* [33]. The sequence of this PVK did not exist in the identified precursor sequence. In the higher mass range, mass matches indicated the presence of further CAPA peptides which were not known so far (Fig. 3B), including SVP-NH_2 (*m/z* 2419.1) and the C-terminal precursor peptide (CPP) with an internal disulfide bond (*m/z* 6380.8; doubly charged 3190.9). Ion intensities of both peptides were sufficient for sequence confirmations (see Fig. 5 for MS/MS of CPP). Two less prominent ion signals at *m/z* 2938.5 and *m/z* 3247.2 (not labelled in Fig. 3B) indicated the

Table 1 – List of CAPA peptides identified by MALDI-TOF mass spectrometry in preparations of *capa* somata from abdominal ganglia and abdominal PSOs. Intermediates of CAPA precursor processing were generally detected in somata but not in hormone release sites (abdominal PSOs). Only the N-terminal processing intermediate of the long splice variant (NPPlong + Lem-PVK-2), which contains an endoproteolytic RR cleavage site, was regularly observed also in mass spectra of abdominal PSOs. Intermediates are given with AA position in the respective precursors. CPP, C-terminal precursor peptide; Int., intermediate of CAPA-precursor processing; long, peptide belonging to the long *capa* cDNA transcript; NPP, N-terminal precursor peptide; PK, pyrokinin; PVK, periviscerokinin; short, peptide belonging to the short *capa* cDNA transcript.

Peptide name	Sequence	*m/z*	Soma	aPSO
NPPshort	NGDPTNSKN-OH	946.42	×	x
Lem-PVK-2	GSSGLISMPRV-NH_2	1102.60	×	x
PVK-1	GASGLIPVMRN-NH_2	1113.62	×	x
PVK-2	GSSSGLISMPRV-NH_2	1189.64	×	x
NPPlong	NGEESETLSTNTNT-OH	1496.63[1]	×	x
CAPA-PK-1	GGGGSGETSGMWFGPRL-NH_2	1651.76	×	x
YLS-NH_2	TDPLWQLPGAHLEQYLS-NH_2	1967.00	×	×
VEA-OH	AFLTLTPGSHVDSYVEA-OH	1806.89	×	×
SKN-OH	SDLTWTYQSPGDPTNSKN-OH	2010.90	×	×
SVP-NH_2	TDPLWSLSADAVFDDDGTVLSVP-NH_2	2419.17	×	×
Int.short$^{1–23}$	NGDPTNSKNRRGSSSGLISMPRV-NH_2	2429.24[a,b]	×	–
Int.long$^{1–27}$	NGEESETLSTNTNTRRGSSGLISMPRV-NH_2	2892.42	×	×
CAPA-PK-2	SADFPWTLVTVREFPADSRDYTPRL-NH_2	2938.48	×	×
CPP $[M+2H]^{2+}$	SADFPWTLVTVREFPADSRDYTPRLGRESEEQEDCCDLTDDDSQSSQNGRSPHSHN-OH	3190.90	×	×
Int.short$^{12–42}$	GSSSGLISMPRVGRAFLTLTPGSHVDSYVEA-OH	3191.61[a,b]	×	–
CPP$^{31–59}$	ESEEQEDCCDLTDDDSQSSQNGRSPHSHN-OH	3247.22	×	×
Int.short$^{45–74}$; Int.long$^{83–112}$	GASGLIPVMRNGRTDPLWQLPGAHLEQYLS-NH_2	3275.71[a]	×	–
Int.long$^{30–61}$	SDLTWTYQSPGDPTNSKNRRGSSSGLISMPRV-NH_2	3493.72[a]	×	–
Int.short$^{78–113}$; Int.long$^{116–151}$	GASGLIPVMRNGRTDPLWSLSADAVFDDDGTVLSVP-NH_2	3727.87	×	–
Int.short$^{1–42}$	NGDPTNSKNRRGSSSGLISMPRVGRAFLTLTPGSHVDSYVEA-OH	4431.22[a]	×	–
Int.long$^{30–80}$	SDLTWTYQSPGDPTNSKNRRGSSSGLISMPRVGRAFLTLTPGSHVDSYVEA-OH	5495.67[a]	×	–
Int.long$^{96–151}$	TDPLWQLPGAHLEQYLSGKRGASGLIPVMRNGRTDPLWSLSADAVFDDDGTVLSVP-NH_2	6018.06[a]	×	–
CPP	SADFPWTLVTVREFPADSRDYTPRLGRESEEQEDCCDLTDDDSQSSQNGRSPHSHN-OH	6380.79	×	×

[a] Mass match.
[b] In CHCA spectra only.

presence of a secondary cleavage site within CPP, yielding CAPA-PK-2 (equivalent to amidated CPP$^{1–28}$) and CPP$^{31–59}$. The identity of both of these peptides was subsequently confirmed by using preparations of abdominal PSOs (see below). A subsequent search for possible intermediates of CAPA precursor processing revealed distinct hits at *m/z* 3275.8 (GASGLIPVMRNGRTDPLWQLPGAHLEQYLS-NH_2) and *m/z* 3727.9 (GASGLIPVMRNGRTDPLWSLSADAVFDDDGTVLSVP-NH_2) (see Fig. 3B); the latter sequence could be confirmed by fragments obtained in MS/MS experiments (not shown). These presumptive processing intermediates resulted from cleavage at the KR motifs. Regularly occurring low intensity ion signals suggested the presence of further processing intermediates at *m/z* 4431.2, 2429.2, 3191.6, and 6018.0 (for sequences see Table 1). With the exception of a substance at *m/z* 2892.4, all distinct ion signals in the mass range of *m/z* 600–8000 could be assigned to pro-CAPA-derived peptides and only the ion signal of the predicted NPP (IHCNGDPTNSKN-OH) was not found at all. Since ion signals of suggested N-terminal processing intermediates (at *m/z* 4431.2 and 2429.2) indicated a signal peptide sequence three AAs longer than predicted by SignalP [32], we searched for a potential NPP at *m/z* 946.4 (NGDPTNSKN-OH). This peptide did not ionize well and was detected with very low signal intensity (see Fig. 3A). Nevertheless, the sequence could be confirmed by MS/MS experiments (not shown).

3.3. *Identification of capa gene products in abdominal PSOs*

Abdominal PSOs of cockroaches store the products of *capa* neurons but are also crossed by other peptide-containing axons either running via the transverse nerves into the periphery [38,39] or crossing abdominal PSOs via segmental nerves/link nerves/transverse nerves (Eckert, Predel unpublished). Therefore, mass spectra obtained from direct tissue profiling of abdominal PSOs generally represent more ion signals than those obtained from preparations of *capa* cells but CAPA peptides always belong to the most abundant substances. Most CAPA peptides observed in somata of *capa* neurons were also found in preparations of abdominal PSOs (Fig. 6A, B). Ion signals which were consistently observed in *capa* somata but extremely rare in preparations of abdominal PSOs, are the suggested intermediates of CAPA precursor processing. In contrast to the pro-CAPA-derived intermediates, CPP was found to be highly abundant in preparations of abdominal PSOs as well. CAPA-PK-2 (*m/z* 2938.5) and CPP$^{31–59}$ (*m/z* 3247.2) which result from cleavage of the secondary cleavage site within CPP showed higher relative abundances in mass spectra of abdominal PSOs than in mass spectra of *capa* cells. In a series of experiments, extracts of five abdominal PSOs each were used for alkylation experiments, SPITC derivatizations, and MS/MS analyses of a number of substances

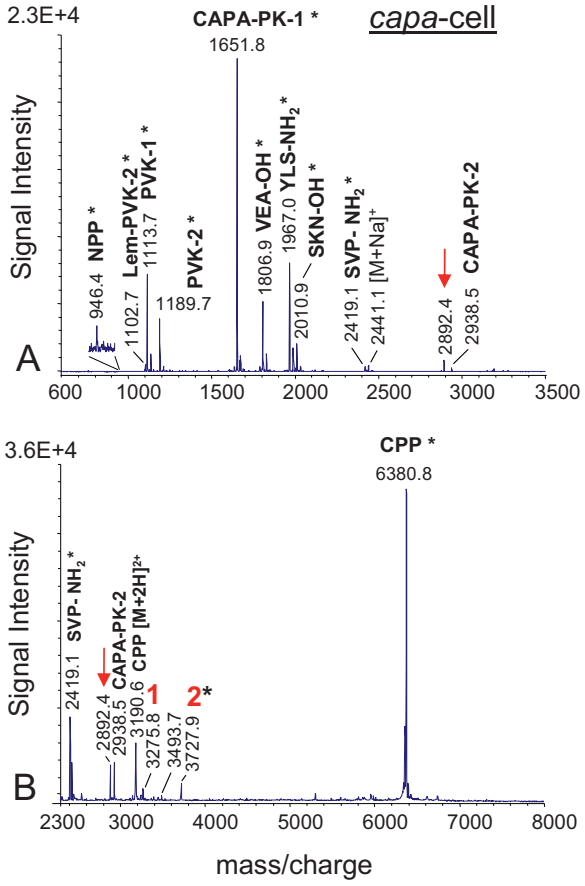

Fig. 3 – MALDI-TOF mass spectra (mass fingerprints) obtained by direct profiling of a single *capa* neuron. Ion signals mass-identical with predictable peptides from the initially identified *capa* cDNA are labelled. Two additional peptides, SKN-OH and Lem-PVK-2, are not encoded by this cDNA. Sequences of peptides marked with an asterisk could be confirmed by MS/MS experiments using a single cell extract (see Fig. 4) or directly from a sample as shown here (see Fig. 5). The ion signal at *m/z* 2892.4 (arrow) could not be assigned to a CAPA peptide. Spectra were acquired with different settings which have been optimized for the given mass ranges, respectively. (A) *m/z* 600–3500, (B) *m/z* 2300–8000; ion signals at 3275.8 (1), and 3727.9 (2), are mass identical with putative intermediates of CAPA precursor processing, resulting from cleavage at KR motifs. The identity of the sequence of the intermediate at *m/z* 3727.9 could be confirmed by MS/MS experiments (not shown). Further ion signals with low signal intensity (not marked) indicate the presence of more intermediates of CAPA precursor processing (see Table 1).

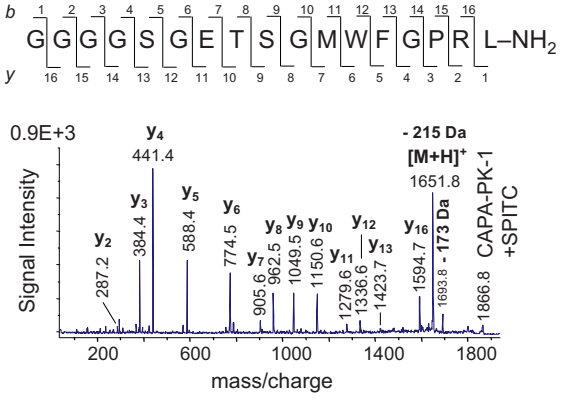

Fig. 4 – MS/MS fragment spectrum of SPITC-derivatized CAPA-PK-1 from an extract of a single *capa* cell. Derivatization induces neutral losses of the modifying group (−173 Da, −215 Da) and intensifies the generation of y-fragments. Assigned fragments are labelled and the resulting sequence is indicated.

which sequences were not successfully fragmented using single cell preparations (e.g. Fig. 7). Cysteine reduction and alkylation verified the internal disulfide bond in CPP (Fig. 8).

The unassigned peptide at *m/z* 2892.3 (see above) was found to be N-terminally extended Lem-PVK-2; linked with an unknown peptide with C-terminal Thr via a potential endoproteolytic RR cleavage site. Thus, we assumed the presence of a second and less abundant *capa* transcript in *capa* cells of abdominal ganglia. Peptidomics data suggested that this transcript includes at least SKN-OH (*m/z* 2010.9), Lem-PVK-2 (*m/z* 1102.6), and a third peptide which could represent the NPP (*m/z* 946.4) of a longer transcript. This information was used to design new PCR primers. Indeed, we succeeded in the identification of a second *capa* cDNA transcript (long transcript GenBank accession number KF836433). Alternative splicing inserts a 114 bp fragment (38 AAs) within the coding region of NPP of the short transcript (Fig. 9A; Supplemental Material). This was verified by four different primer combinations (see Supplemental Material). As expected from mass spectra, all PCR reactions verified that the longer splice variant is less abundant than the shorter transcript (Fig. 9B). The partially sequenced peptide at *m/z* 2892.3 could be assigned

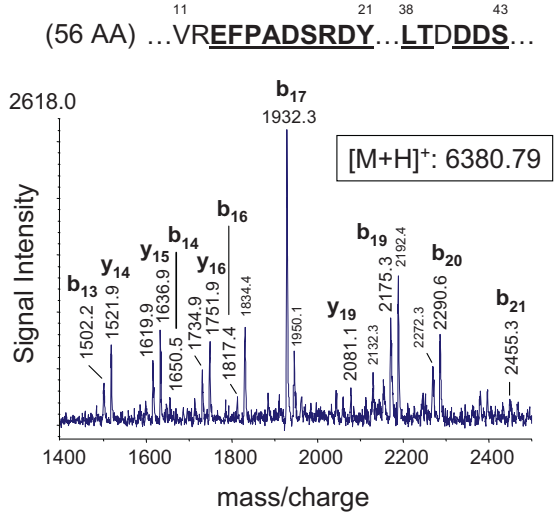

Fig. 5 – Detail of MS/MS fragment spectrum of a CPP signal at *m/z* 6380.8 from a single cell preparation (direct profiling experiment). Fragments do not cover the entire sequence but are sufficient to confirm the identity of the peptide as CPP. For confirmation of internal disulfide bond, see Figs. 7 and 8.

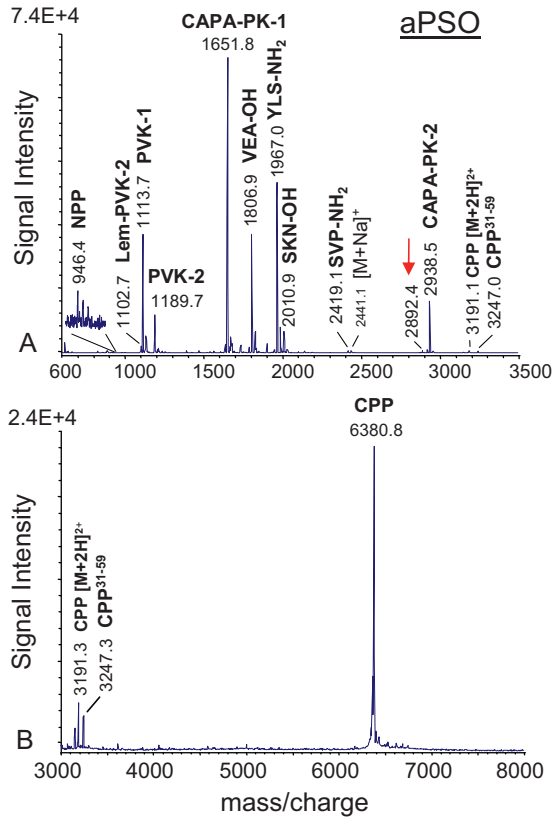

Fig. 6 – MALDI-TOF mass spectra (mass fingerprints) obtained by direct profiling of a single abdominal PSO-5 with exceptional low abundance of peptides other than CAPA peptides. CAPA peptides are labelled. Note that, in comparison with single cell spectra (see Fig. 3), ion signals of pro-CAPA-intermediates disappeared, relative signal intensity of the ion signal at *m/z* 2892.4 (arrow) decreased, and relative signal intensity of CAPA-PK-2 (equivalent to amidated CPP$^{1–28}$) and CPP$^{31–59}$ increased. Spectra were acquired with different settings which have been optimized for the given mass ranges, respectively. (A) *m/z* 600–3500, (B) *m/z* 2300–8000.

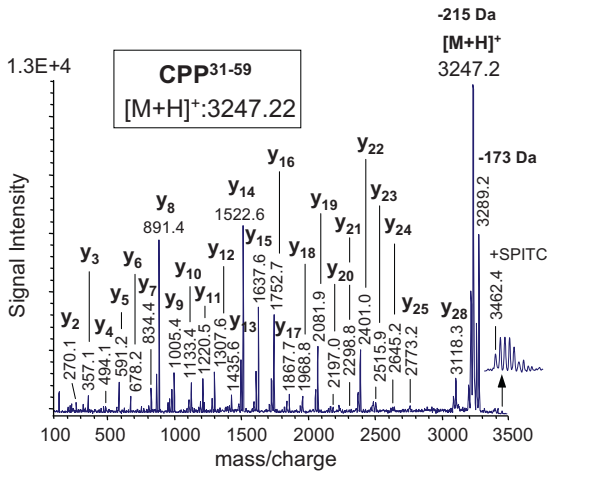

Fig. 7 – MS/MS fragment spectrum of SPITC-derivatized peptide at *m/z* 3247.2 including the neutral losses of the modifying group (−173 Da, −215 Da). Assigned fragments are labelled and reveal the sequence of CPP$^{31–59}$ (ESEEQEDCCDLTDDDSQSSQNGRSPHSHN-OH). Mass differences between y_{20}-, y_{21}-, and y_{22}-fragments indicate an internal disulfide bond (see also Fig. 8).

as N-terminal intermediate, which yields Lem-PVK-2 and the NPP of the long transcript, and also indicated that both transcripts share the same signal peptide sequence. The CAPA peptides which have been identified by MALDI-TOF mass spectrometry in *capa* somata and abdominal PSOs cover the complete prohormone sequences of both transcripts. All peptides detected, including potential processing intermediates typical of the long transcript only, are summarized in Table 1.

3.4. *Re-evaluation of capa gene products in T. castaneum*

As shown above, the CAPA-PK-2 of *P. americana*, although potentially bioactive, was rather weakly cleaved from CPP, likely due to acidic Glu residues C-terminal of the monobasic cleavage site (Arg). Homologous PKs are known from CAPA precursor sequences of a number of insects (e.g. [34,40,41] but were never confirmed to be processed. To approach this question, we analyzed *capa* gene products of the beetle *T. castaneum*; earlier peptidomic analyses failed to detect *Tribolium*-PKs at all [34]. Indeed, direct tissue profiling of abdominal neurohemal release sites revealed that the homologous CAPA-PK-2 of *T. castaneum* was not cleaved at all from the corresponding CPP (Fig. 10). In addition, CAPA-PK-1 with a C-terminal WFGPRL-NH_2 motif was not efficiently cleaved from a C-terminal VQ-OH peptide, the more frequent cleavage product was N-terminal extended CAPA-PK-1 (SESNWVPDDNSYGAQRPGANSGGMWFGPRLGRVQ-OH).

4. Discussion

The detailed evaluation of a first *capa* cDNA and mass spectra from single *capa* neurons of *P. americana* indicated that not all

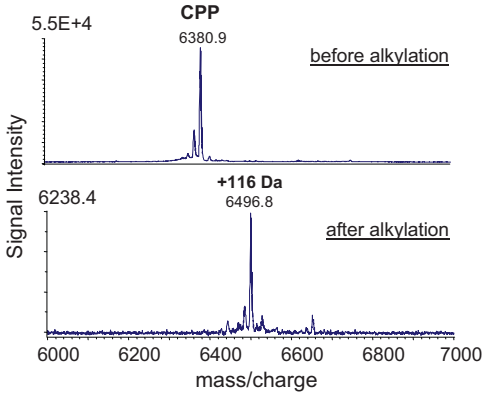

Fig. 8 – Confirmation of the internal disulfide bond in CPP (*m/z* 6380.9) of *P. americana*. The carbamidomethylation (CAM) of two cysteine residues resulted in a mass shift of 116 Da.

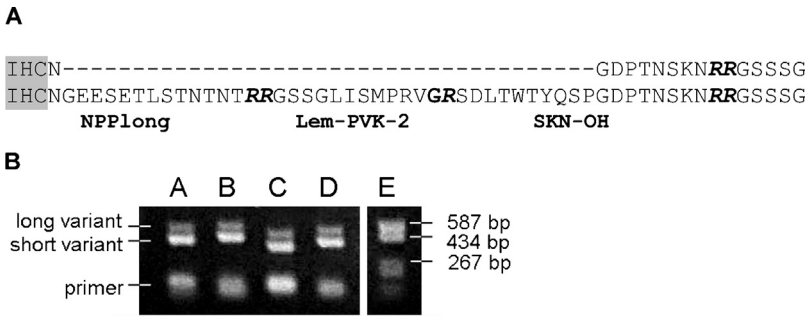

Fig. 9 – (A) Comparison of N-terminal precursor sequences (supposed C-terminal sequence of signal peptide highlighted in grey) from alternatively spliced *capa* transcripts of *P. americana*. Amino acids N-terminal and C-terminal of the depicted sequences are identical in both transcripts (see Supplemental Material). A single peptide (NPPshort) is exclusively cleaved from the short transcript; the long transcript yields three additional peptides (NPPlong, Lem-PVK-2, SKN-OH). (B) PCR analysis of alternatively spliced *capa* mRNAs. Mass spectra (see above) as well as PCR suggest that the longer splice variant is always less abundant than the shorter transcript. Four primer combinations (A, F1-R1; B, F1-R2; C, F2-R1; D, F2-R2) were used to amplify the N-terminally spliced region of *capa* cDNA (see Supplemental Material for primer positions). The position of the longer and shorter splice variants are indicated on the left. E, pBR322 DNA/BsuRI marker.

of the observed CAPA peptides are encoded by this cDNA transcript. Subsequent hypothetical assignment of the remaining CAPA peptides to the N-terminus of a longer cDNA transcript enabled the design of specific primers for PCR amplification of the intended alternatively spliced region and thereby the identification of a second splice variant. With this, all peptides previously described as potential CAPA peptides of *P. americana* could be assigned to one of the *capa* cDNA transcripts. These results also explained the different abundances of PVKs in mass spectra; a single PVK (Lem-PVK-2) occurs only in the longer and less abundant transcript.

In the course of our study, most data were obtained by performing single cell peptidomics of identified *capa* somata; this approach resulted in the most comprehensive peptidomics-based information known from any insect so far. In fact, single cell mass spectrometry yielded a complete coverage of the sequences of both transcripts (without signal peptides). Fragment analysis was supported by single cell SPITC derivatization since SPITC-derivatized peptides generally provide unmatched y-fragment series which are easy to interpret [35,42]. We found that relative ion intensities of mature CAPA peptides were very similar in mass spectra of *capa* cells and mass spectra of abdominal PSOs. This corroborates the assumption [16] that all peptides finally cleaved from the CAPA precursor are packaged into vesicles and are likely to be released into the hemolymph; including those peptides without known receptors. Evidence of intermediates of CAPA precursor processing was found in mass spectra from *capa* somata. Low signal intensity of most of these peptides prevented MS/MS experiments. However, mass matches suggested at least nine processing intermediates of both cDNA transcripts (see Table 1) that cover the complete precursor sequences. Processing intermediates resulting from cleavage at two KR cleavage sites (*m/z* 3275.8 and 3727.9) were found to be more abundant than other intermediates in *capa* somata and the sequence of one these peptides (*m/z* 3727.9) could eventually be confirmed by fragment analysis. A single intermediate of CAPA precursor processing (NPP + Lem-PVK-2; long transcript) was also detected in abdominal PSOs. This peptide contains the only RR-cleavage site of both *capa* cDNA transcripts.

In mass spectra of *capa* neurons, a co-localization of CAPA peptides with other neuropeptides was not found (*m/z* 600–8000). Truncated CAPA peptides, as occasionally detected in mass spectra of tissue extracts (not shown), were also not observed. In addition, mass spectra verified that both CAPA cDNA transcripts are present in *capa* neurons of abdominal ganglia and also confirmed that products of the short transcript were much more abundant in these cells. CAPA peptides of a number of holometabolous insects are known to be processed in a cell-specific pattern (see [1]), also in *D. melanogaster* [14]. In the fruitfly, Va-neurons in abdominal ganglia process PVKs, CAPA-PK, and the C-terminal PP (CPPB; [43]), whereas two *capa* neurons in the subesophageal ganglion express only CAPA-PK and CPPB. Such differentiation possibly evolved within Pterygota. In the basel pterygote insect *P. americana*, a cell-specific pattern of CAPA peptides was not observed so far [15,16, this study], although different transcripts are present.

Another interesting finding of this study was the presence of a second PK (CAPA-PK-2) in the CAPA precursor sequence of *P. americana*. CAPA-PK-2 with the C-terminal sequence YTPRL-NH_2 potentially activates a specific PK-receptor (PK-2 receptor in *D. melanogaster*; [44]); different from the receptor recognizing CAPA-PK-1 (PK-1 receptor in *D. melanogaster*; [20]). Bioassays have shown that myostimulatory effects of *Periplaneta* PK-4 (DHLPHDVYSPRL-NH_2), which is a product of the pyrokinin/FXPRL gene [12] and shares C-terminal sequence similarity with CAPA-PK-2, are much stronger than the effects of *Periplaneta* CAPA-PK-1 [45]. Since CAPA peptides are known to be stored in the same vesicles [16], released CAPA peptides potentially activate three different receptors. This is likely not a physiological advantage and possible receptor activations by CAPA-PK-2 have been minimized by incomplete cleavage from the precursor. The acidic Glu-residues C-terminal of the GR-motif in the precursor sequence are not well suited

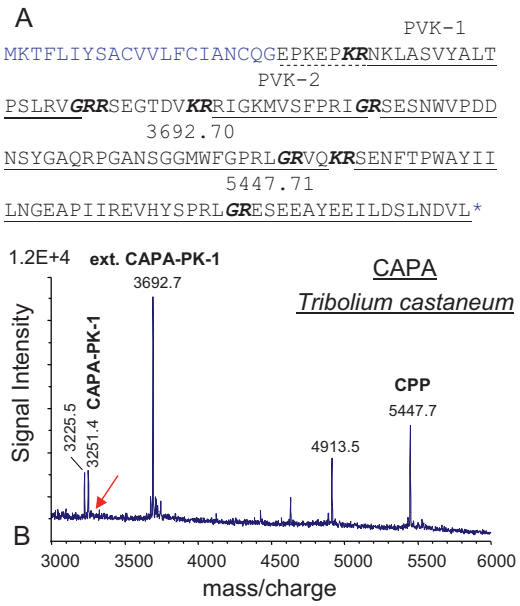

Fig. 10 – Analysis of products of the *capa* gene from the beetle *Tribolium castaneum*. (A) CAPA precursor sequence of *T. castaneum*; mature peptides as detected in [34] and this study are underlined. Earlier experiments failed to confirm the presence of PKs at all [34]. (B) MALDI-TOF mass spectrum (*m/z* 3000–6000; for lower mass range with PVKs see [34]) obtained by direct profiling of CAPA peptide release sites at the junction of abdominal segmental nerve/link nerve (evolved type of abdominal PSO; [6]). An ion signal typical of CAPA-PK-2 (arrow) is not detectable. Instead and similar to the situation we found in *P. americana*, this PK is not cleaved from CPP (*m/z* 5447.7). The spectrum also shows that CAPA-PK-1 (extended WFGPRL-NH_2) is only weakly cleaved from a C-terminal tetrapeptide (GRVQ) in *capa* cells of abdominal ganglia.

for prohormone convertases and it is remarkable that we observed exactly the same situation in the beetle *T. castaneum*, which is not closely related to the cockroach *P. americana*. Although CAPA-PK-2 of *T. castaneum* is not cleaved from the precursor, the C-terminal sequence is well conserved (see also capa gene of e.g. *Aedes aegypti*, [40]). This conservation over hundreds of millions of years is unusual for precursor sequences without functional relevance as ligands.

In summary, the peptidomics approach to study CAPA-peptides of *P. americana* revealed a number of novelties regarding different products of the *capa* gene which likely evolved within the arthropod lineage.

Acknowledgements

We acknowledge the financial assistance of the Deutsche Forschungsgemeinschaft (RP 766/9-1; 595/10-1) and also thank Michael Gonsior (Bruker Daltonik GmbH) for technical support, as well as Xin Zhou (Shenzhen, China), Shanlin Liu (Shenzhen, China) and Benjamin Wipfler (Jena, Germany) for providing EST data from the 1KITE project.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.euprot.2014.02.005.

REFERENCES

[1] Predel R, Wegener C. Biology of the CAPA peptides in insects. Cell Mol Life Sci 2006;63:2477–90.
[2] Paluzzi JP. Anti-diuretic factors in insects: the role of CAPA peptides. Gen Comp Endocrinol 2012;176:300–8.
[3] Davies SA, Cabrero P, Povsic M, Johnston NR, Terhzaz S, Dow JA. Signaling by *Drosophila* capa neuropeptides. Gen Comp Endocrinol 2013;188:60–6.
[4] Wegener C, Predel R, Eckert M. Quantification of periviscerokinin-1 in the nervous system of the American cockroach, *Periplaneta americana*. An insect neuropeptide with unusual distribution. Arch Insect Biochem Physiol 1999;40:203–11.
[5] Raabe M. Etude des phenomenes de neurosecretion au niveau de la chaine nerveuse ventrale des Phasmides. Bull Soc Zool France 1965;90:631–54.
[6] Raabe M. The neurosecretory–neurohaemal system of insects: anatomical, structural and physiological data. Adv Insect Physiol 1983;17:205–303.
[7] Provansal-Baudez A. Physiological roles of perisympathetic organs in insects. In: Gupta AP, editor. Neurohemal organs of arthropods. Springfield: C.C. Thomas; 1983. p. 581–601.
[8] Tublitz NJ, Brink D, Broadie KS, Loi PK, Sylwester AW. From behavior to molecules: an integrated approach to the study of neuropeptides. Trends Neurosci 1991;14:254–9.
[9] Huesmann G, Cheung C, Loi P, Lee T, Swiderek K, Tublitz NJ. Amino acid sequence of CAP_{2b}, an insect cardioacceleratory peptide from the tobacco hawkmoth *Manduca sexta*. FEBS Lett 1995;371:311–4.
[10] Loi PK, Tublitz NJ. Sequence and expression of the CAPA/CAP2b gene in the tobacco hawkmoth, *Manduca sexta*. J Exp Biol 2004;207:3681–91.
[11] Predel R, Linde D, Rapus J, Vettermann S, Penzlin H. Periviscerokinin (Pea-PVK): a novel myotropic neuropeptide from the perisympathetic organs of the American cockroach. Peptides 1995;16:61–6.
[12] Predel R, Kellner R, Nachman RJ, Holman GM, Rapus J, Gade G. Differential distribution of pyrokinin-isoforms in cerebral and abdominal neurohemal organs of the American cockroach. Insect Biochem Mol Biol 1999;29:139–44.
[13] Kean L, Cazenave W, Costes L, Broderick KE, Graham S, Pollock VP, et al. Two nitridergic peptides are encoded by the gene capability in *Drosophila melanogaster*. Am J Physol Reg Integr Comp Physiol 2002;282:1297–307.
[14] Predel R, Wegener C, Russell WK, Tichy SE, Russell DH. Nachman RJ Peptidomics of CNS-associated neurohemal systems of *Drosophila*: a mass spectrometric survey of peptides from individual flies. J Comp Neurol 2004;474:379–92.
[15] Pollák E, Eckert M, Molnár L, Predel R. Differential sorting and packaging of capa-gene related products in an insect. J Comp Neurol 2005;481:84–95.
[16] Eckert M, Herbert Z, Pollák E, Molnár L, Predel R. Identical cellular distribution of all abundant neuropeptides in the major abdominal neurohemal system of an insect (*Periplaneta americana*). J Comp Neurol 2002;452:264–75.
[17] Santos JG, Pollák E, Rexer KH, Molnár L, Wegener C. Morphology and metamorphosis of the peptidergic Va

neurons and the median nerve system of the fruit fly, *Drosophila melanogaster*. Cell Tissue Res 2006;326:187–99.
[18] Iversen A, Cazzamali G, Williamson M, Hauser F, Grimmelikhuijzen CJ. Molecular cloning and functional expression of a *Drosophila* receptor for the neuropeptides capa-1 and -2. BBRC 2002;299:628–33.
[19] Park Y, Kim YJ, Adams ME. Identification of G protein-coupled receptors for *Drosophila* PRXamide peptides, CCAP, corazonin, and AKH supports a theory of ligand-receptor coevolution. Proc Natl Acad Sci USA 2002;99:11423–8.
[20] Cazzamali G, Torp M, Hauser F, Williamson M, Grimmelikhuijzen CJ. The *Drosophila* gene CG9918 codes for a pyrokinin-1 receptor. BBRC 2005;335:14–9.
[21] Terhzaz S, Cabrero P, Robben JH, Radford JC, Hudson BD, Milligan G, et al. Mechanism and function of *Drosophila* capa GPCR: a desiccation stress-responsive receptor with functional homology to human neuromedin U receptor. PLoS One 2012;7:e29897.
[22] Paluzzi JP, Russell WK, Nachman RJ, Orchard I. Isolation, cloning, and expression mapping of a gene encoding an antidiuretic hormone and other CAPA-related peptides in the disease vector, *Rhodnius prolixus*. Endocrinology 2008;149:4638–46.
[23] Paluzzi JP, Orchard I. A second gene encodes the anti-diuretic hormone in the insect, *Rhodnius prolixus*. Mol Cell Endocrinol 2010;317:53–63.
[24] Neupert S, Russell WK, Russell DH, Predel R. Two capa-genes are expressed in the neuroendocrine system of *Rhodnius prolixus*. Peptides 2010;31:408–11.
[25] Roth S, Fromm B, Gäde G, Predel R. A proteomic approach for studying insect phylogeny: CAPA peptides of ancient insect taxa (Dictyoptera, Blattoptera) as a test case. BMC Evol Biol 2009;3:50.
[26] Suska A, Miguel-Aliaga I, Thor S. Segment-specific generation of *Drosophila* capability neuropeptide neurons by multi-faceted Hox cues. Dev Biol 2011;353:72–80.
[27] Gabilondo H, Losada-Pérez M, del Saz D, Molina I, León Y, Canal I, et al. A targeted genetic screen identifies crucial players in the specification of the *Drosophila* abdominal Capaergic neurons. Mech Dev 2011;128:208–21.
[28] Predel R, Rapus J, Eckert M, Holman GM, Nachman RJ, Wang Y, et al. Isolation of periviscerokinin-2 from the abdominal perisympathetic organs of the American cockroach, *Periplaneta americana*. Peptides 1998;19:801–9.
[29] Predel R, Agricola H, Linde D, Veenstra JA, Wollweber L, Penzlin H. The insect neuropeptide corazonin: immunocytochemical and physiological studies in Blattariae. Zoology (ZACS) 1994;98:35–49.
[30] Predel R, Eckert M, Holman GM. The unique neuropeptide pattern in abdominal perisympathetic organs of insects. Ann N YAcad Sci 1999;897:282–90.
[31] Predel R, Kellner R, Baggerman G, Steinmetzer T, Schoofs L. Identification of novel periviscerokinins from single neurohaemal release sites in insects. MS/MS fragmentation complemented by Edman degradation. Europ J Biochem 2000;267:3869–74.
[32] Predel R, Neupert S, Wicher D, Gundel M, Roth S, Derst C. Unique accumulation of neuropeptides in an insect: FMRFamide related peptides in the cockroach, *Periplaneta americana*. Europ J Neurosci 2004;20:1499–513.
[33] Predel R. Peptidergic neurohemal system of an insect: mass spectrometric morphology. J Comp Neurol 2001;436: 363–75.
[34] Li B, Predel R, Neupert S, Hauser F, Tanaka Y, Cazzamali G, et al. Genomics, transcriptomics, and peptidomics of neuropeptides and protein hormones in the red flour beetle *Tribolium castaneum*. Genome Res 2008;18:113–22.
[35] Wang D, Kalb SR, Cotter RJ. Improved procedures for N-terminal sulfonation of peptides for matrix-assisted laser desorption/ionization post-source decay peptide sequencing. Rapid Commun Mass Spectrom 2004;18:96–102.
[36] Rappsilber J, Mann M, Ishihama Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nature Prot 2007;2: 1896–906.
[37] Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 2011;8:785–6.
[38] Eckert M, Predel R, Gundel M. Periviscerokinin-like immunoreactivity in the nervous system of the American cockroach, *Periplaneta americana*. Cell Tissue Res 1999;295:159–70.
[39] Predel R, Eckert M, Pollák E, Molnár L, Scheibner O, Neupert S. Peptidomics of identified neurons demonstrates a highly differentiated expression pattern of FXPRLamides in the neuroendocrine system of an insect. J Comp Neurol 2007;500:498–512.
[40] Predel R, Neupert S, Garczynski SF, Crim JW, Brown MR, Russell WK, et al. Neuropeptidomics of the mosquito *Aedes aegypti*. J Prot Res 2010;9:2006–15.
[41] Badisco L, Huybrechts J, Simonet G, Verlinden H, Marchal E, Huybrechts R, et al. Transcriptome analysis of the desert locust central nervous system: production and annotation of a *Schistocerca gregaria* EST database. PLoS One 2011; 6:e17274.
[42] Zoephel J, Reiher W, Rexer KH, Kahnt J, Wegener C. Peptidomics of the agriculturally damaging larval stage of the cabbage root fly *Delia radicum* (Diptera: Anthomyiidae). PLoS One 2012;7:e41543.
[43] Wegener C, Reinl T, Jansch L, Predel R. Direct mass spectrometric peptide profiling and fragmentation of larval peptide hormone release sites in *Drosophila melanogaster* reveals tagma-specific peptide expression and differential processing. J Neurochem 2006;96:1362–74.
[44] Rosenkilde C, Cazzamali G, Williamson M, Hauser F, Sondergaard L, DeLotto R, et al. Molecular cloning, functional expression, and gene silencing of two *Drosophila* receptors for the *Drosophila* neuropeptide pyrokinin-2. BBRC 2003;309:485–94.
[45] Predel R, Nachman RJ. Efficacy of native FXPRLamides (pyrokinins) and synthetic analogs on visceral muscles of the American cockroach. J Insect Physiol 2001;47:287–93.