This is a postprint of an article published in
Langner, K.F.A., Darpel, K.E., Denison, E., Drolet, B.S., Leibold, W.,
Mellor, P.S., Mertens, P.P.C., Nimtz, M., Greiser-Wilke, I.
Collection and analysis of salivary proteins from the biting midge
Culicoides nubeculosus (Diptera: Ceratopogonidae)
Collection and Analysis of Salivary Proteins from the Biting Midge
Culicoides nubeculosus (Diptera: Ceratopogonidae)

Kathrin F. A. Langner a,*, Karin E. Darpel b, Eric Denison b, Barbara S. Drolet c, Wolfgang
Leibold a, Philip S. Mellor b, Peter P. C. Mertens b, Manfred Nimtz d, Irene Greiser-Wilke e

a Immunology Unit, School of Veterinary Medicine, Bischofsholer Damm 15, 30173 Hannover,
Germany.

b Pirbright Laboratory, Institute for Animal Health, Ash Road, Pirbright, Surrey, GU24 0NF, UK
c United States Department of Agriculture, Agriculture Research Service, Arthropod-Borne Animal
Disease Laboratory, 1000 E. University Avenue, Laramie, Wyoming 82071, USA
d German Research Centre for Biotechnology, Mascheroder Weg 1, 38124 Braunschweig, Germany
e Institute of Virology, School of Veterinary Medicine, Buenteweg 17, 30559 Hannover, Germany

*Corresponding author. Tel.: 49-511-856-7241, fax: 49-511-856-7682
E-mail address: klangner@tiho-hannover.de (K.F.A. Langner).
Abstract

Salivary proteins of hematophagous *Culicoides* spp. are thought to play an important role in pathogen transmission and skin hypersensitivity. Analysis of these proteins, however, has been problematic due to the difficulty in obtaining adequate amounts of secreted *Culicoides* saliva. In this study, we describe the initial proteome analysis of *Culicoides nubeculosus* midge saliva obtained by a newly developed, efficient saliva collection method. Proteins of collected saliva were separated using one- and two-dimensional gel electrophoresis. Twenty predominant protein spots were selected for internal peptide sequencing by electrospray ionization time-of-flight tandem mass spectrometry (ESI-ToF MS-MS) followed by FASTA analysis and comparisons with public data bases. Peptides derived from 11 protein spots showed significant identities to deduce amino-acid sequences of salivary proteins recently obtained from a cDNA-library of *Culicoides sonorensis*. These included proteins involved in sugar meal digestion (maltase), defense, and coagulation inhibition (Kunitz protease inhibitor-like), as well as members of the D7 family and novel or unclassified salivary proteins. Peptides from four spots were related to proteins from insects other than *Culicoides* spp. including a chemical sense-related lipophilic ligand-binding homologous salivary protein from the black blowfly (*Phormia regina*) and a carboxypeptidase from *Drosophila* spp. Two proteins were assigned putative identities of non-insect peroxidases and cytochrome P450 monooxygenases. Peptides from four protein spots did not have significant identities with GenBank sequences and may indicate the presence of novel salivary components of *C. nubeculosus*. The potential role of the identified proteins for the transmission of pathogens and the induction of allergy is discussed.

Keywords: *Culicoides nubeculosus*; midge; saliva collection; proteomics; two-dimensional gel electrophoresis; mass spectrometry
1. Introduction

*Culicoides* biting midges (Diptera: Ceratopogonidae) are vectors of a wide range of economically important pathogens that affect both domestic and wild animals (Mellor et al., 2000). Midges can transmit bluetongue virus (Chandler et al., 1985), epizootic hemorrhagic disease virus (Paweska et al., 2005), vesicular stomatitis virus (Drolet et al., 2005; Perez de Leon and Tabachnick, 2006), and African horse sickness virus (Mellor and Hamblin, 2004). In addition to acting as a vector, *Culicoides* spp. are the primary cause of an extremely pruritic allergic dermatitis known colloquially as ‘sweet itch’, ‘summer eczema’, ‘insect bite hypersensitivity’, ‘Queensland itch’ and ‘Kasen’, in atopic horses worldwide (Riek, 1953; Anderson et al., 1993; Kurotaki et al., 1994; Marti et al., 1999; McKelvie et al., 1999).

It is suggested that the saliva of hematophagous arthropods plays an important role in mediating pathogen transmission (Ribeiro and Francischetti, 2003). This is illustrated by the mechanism of saliva activated transmission (SAT), first reported for tick-borne encephalitis (Jones et al., 1990). In addition, pathogens such as *Leishmania* spp., vesicular stomatitis virus, dengue fever virus, and Sindbis virus have shown enhanced or decreased infections of mammalian hosts or cell lines when co-administered with the saliva of the pathogen’s specific vector (Titus and Ribeiro, 1988; Theodos et al., 1991; Limesand et al., 2000, 2003; Ader et al., 2004; Schneider et al., 2004). Vector saliva may also be influential in arbovirus transmission from one vector to another by local co-feeding on infected non-viraemic mammalian hosts (Labuda et al., 1993, 1997; Mead et al., 2000; Higgs et al., 2005). Moreover, saliva of hematophagous insects is known to be an important source of allergens (Galindo et al., 1998; McDermott et al., 2000; Paddock et al., 2001). Salivary proteins isolated from the mosquito *Aedes aegypti* were found to be capable of inducing an allergic dermatitis in atopic humans (Peng and Simons, 2004). Bites of the cat flea *Ctenocephalides felis* cause a similar disease in cats and dogs (Lee et al., 1999; Bond et al., 2006). The saliva of assassin
bugs (Triatominae) contains a number of different allergens leading to severe allergic reactions such as urticaria, dyspnoea and life-threatening anaphylaxis in man (Moffitt et al., 2003).

Despite their relevance for pathogen transmission and hypersensitivity, individual salivary proteins of Culicoides spp. have not yet been biochemically identified. Fractionation of Culicoides sonorensis salivary glands has revealed that midge saliva possesses functional components commonly found in hematophagous insects such as preventing blood coagulation and promoting vasodilatation (Perez de Leon and Tabachnick, 1996; Perez de Leon et al., 1997, 1998). Recently a broad range of genes encoding salivary proteins were characterized from a salivary gland cDNA-library of C. sonorensis and a number of proteins involved in roles such as hemostasis, pheromone binding and digestion were identified (Campbell et al., 2005). However, since it is known that the levels of mRNA and expressed gene products often do not correspond, it is necessary to analyze the proteome itself (Anderson and Seilhamer, 1997). Thus far, the lack of sufficient quantities of purified midge saliva has restricted research on Culicoides salivary proteins. Here, we report the application of a novel, quick and efficient saliva collection method which was used in combination with one and two-dimensional electrophoresis and mass spectrometry to identify 16 salivary proteins of C. nubeculosus. Insight into the midge’s salivary proteome will help further investigations concerning the role of Culicoides spp. as an efficient disease transmission vector and as a source of allergens.

2. Materials and methods

2.1. Saliva collection and elution of proteins

A pre-existing membrane feeding system reported by Mellor (1971) was modified for saliva collection from 2-5 day old, unfed C. nubeculosus reared according to the method of Boorman (1974). The collection system consisted of an inner and an outer glass chamber (Fig. 1). The inner chamber was filled with horse blood and sealed with a double layer of Parafilm™ membrane.
(American National Can, Greenwich, CT). The outer chamber was connected to a heated water circulation system to adjust the blood temperature to 37 °C. A hydrophilic Durapore™ filter (Millipore, Eschborn, Germany), similar to that described for the collection of salivary proteins from the cat flea (Frank et al., 1996), was used in the collection system. The filter was soaked in sucrose solution (10% w/v) to encourage midge feeding and placed on top of the Parafilm™ membrane. Finally another layer of Parafilm™ was placed above the filter to prevent contamination with non-salivary products of the midges.

Over 5,000 *C. nubeculosus* (male and female) in twenty pill-boxes (Watkins and Doncaster, Kent, UK) covered with fine-mesh gauze were allowed to probe through the membrane and to deposit saliva on the filter. Each box containing 250-300 insects was placed on the saliva collection system for 20 to 30 min, two to three times a day (4-6 h intervals) over a period of three days. Filters were changed when dry and stored at 4 °C in phosphate-buffered saline (PBS) until further processing. For the elution of salivary products, 1 mM CHAPS (3-[(3-Cholamido-propyl)dimethylammonio]-1-propanesulfonate; Omnilab, Hannover, Germany) was added to the PBS and filters were shaken overnight at 4 °C. The eluate was concentrated to a volume of 1 ml using ultrafiltration tubes (Vivaspin 20, Cut-off 3 kDa; Vivascience, Hannover, Germany). The concentrate was desalted by subsequent addition of 30 ml MilliQ water and centrifugation back to a volume of 1 ml. The Bicinchoninic Acid Protein Assay (Perbio Science, Bonn, Germany) was used to determine the protein concentration. The concentrate was stored at 4 °C. Prior to gel electrophoresis the required volume was dried in a vacuum centrifuge.

2.2. One-dimensional gel electrophoresis

Vacuum dried salivary proteins (20 µg per lane) were dissolved in 1x sodium dodecyl sulfate (SDS) sample buffer (100 mM dithiothreitol (DTT), 62.5 mM Tris pH 6.8, 10% glycerine, 2% SDS, 0.001% bromophenol blue). SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli 1970) was done using 1 mm thick 15% polyacrylamide gels in a MINI PROTEAN II
Cell (Bio-Rad Laboratories, Munich, Germany). Gels were run with SDS running buffer (192 mM glycine, 25 mM Tris, 0.1% SDS) at 200 V. The Mark12 wide range ladder (Invitrogen, Karlsruhe, Germany) was used to determine molecular masses of the salivary proteins.

2.3 Two-dimensional gel electrophoresis. First dimension: isoelectric focusing (IEF)

To separate proteins by their isoelectric points (pI), vacuum dried salivary proteins (50 µg per strip) were dissolved in 125 µl rehydration buffer containing 8 M urea, 2% CHAPS, 40 mM DTT, 0.2% Bio-Lyte 3-10 (Bio-Rad) and 0.001% bromophenol blue. Samples were applied to immobilized pH gradient (IPG) strips (pH 3-10 and 3-6; 7 cm; Bio-Rad) by passive rehydration for 12 h. IEF was performed at a voltage of 4,000 V until 10,000 Vh were reached.

2.4 Two-dimensional gel electrophoresis. Second dimension: gel electrophoresis

Prior to running the SDS-PAGE for the second dimension, the IPG strips were soaked for 15 min in DTT equilibration buffer (6 M urea, 0.375 M Tris-HCl, 130 mM DTT, 20% w/v glycerol, 2% w/v SDS; pH 8.8) followed by 15 min in iodoacetamide equilibration buffer (6 M urea, 0.375M Tris-HCl, 135 mM iodoacetamide, 20% w/v glycerol, 2% w/v SDS; pH 8.8). Second dimension electrophoresis was done using 1 mm thick 15% polyacrylamide gels as described in 2.2. As above, the Mark12 wide range ladder (Invitrogen) was used to determine molecular masses of the salivary proteins.

2.5 In-gel digestion and peptide extraction

After one- and two-dimensional electrophoresis, proteins were visualized using the Gelcode Snap Silverstain kit (Perbio Science). When in-gel protein digestion was to be performed, gels were stained with Simply-Blue-Stain™ (Invitrogen). Evaluation of molecular masses and isoelectric points was done arithmetically. Protein spots stained with Simply-Blue-Stain™ were excised and transferred to 1 ml tubes that had been previously rinsed with MilliQ water and acetonitrile to remove any contaminants. Gel pieces were washed two times for 15 min with 50 mM ammonium
bicarbonate (pH 8.0) and two times for 15 min with 60% acetonitrile/0.5% formic acid until they were completely dehydrated. After drying under vacuum, protein spots were reduced with 10 mM DTT for 20 min at 56 °C and alkylated with 100 mM iodoacetamide for 20 min at room temperature in the dark. To remove residual iodoacetamide, gel spots were washed with 50 mM ammonium bicarbonate (pH 8.0) followed by 60% acetonitrile/0.5% formic acid like described above. Gel pieces were dried under vacuum. Protein digestion was performed with trypsin (Promega, Mannheim, Germany) according to manufacturer’s instructions. For peptide extraction, gel pieces were washed with 50 mM ammonium bicarbonate and 60% acetonitrile/0.5% formic acid like described above. The extracts were pooled and concentrated under vacuum to a final volume of 20 µl. ZipTip® C18 sepharose tips (Millipore) were used to desalt and concentrate the peptide samples to a volume of 10 µl as per manufacturer’s instructions.

2.6. Electrospray ionization time of flight tandem mass spectrometry (ESI-TOF-MS/MS) and protein identification

Peptides eluted from the ZipTip® (approx. 3 µl) were placed in gold-coated nanospray glass capillaries (Protana, Odense, Denmark). The tip of the capillary was placed orthogonally in front of the entrance hole of a Quadrupole time of flight (Q-TOF 2) mass spectrometer (Micromass, Manchester, UK), equipped with a nanospray ion source, and a voltage of approx. +1,000 V was applied. For collision-induced dissociation experiments, parent ions were selectively transmitted from the Q-TOF mass analyzer into the collision cell. Argon was used as the collision gas and the kinetic energy was set at -35 eV. Resulting daughter ions were then separated by an orthogonal TOF mass analyzer. The spectra of ions were computer enhanced using Maximum Entropy 3 (Micromass, Manchester, UK). Amino acid sequences were deduced from carboxyterminal fragment ion series of the y-type showing amino acid specific mass increments and were confirmed by complementary aminoterminal fragment ion series of the b-type. Amino acid sequences were used to search the protein database UNIPROT. The software Align X (Invitrogen) was used for
determining alignments, identities of peptide sequences from public databases, and calculations of molecular masses and pI data of database proteins.

3. Results and discussion

3.1. Saliva collection

Until now the lack of efficient quantities of purified midge saliva has limited investigations of the salivary proteome of Culicoides spp. In this study, we developed a modified membrane feeding apparatus to collect saliva of C. nubeculosus, the major cause for equine allergic dermatitis in Middle Europe. When midges were placed on the saliva collection system, insects immediately assembled at the gauze-Parafilm™ interface. Feeding levels were monitored by the extent to which midges crowded together at the interface. The intensity of feeding decreased significantly after 15 to 20 min and correlated with the engorgement of female midges. Less crowding was observed at the interface when the blood temperature was below 37 °C. In total, salivary proteins from over 5,000 male and female midges were collected on nine filters over a three day period for a total of 1.2 mg of eluted protein.

3.2. Gel analysis and protein identification by mass spectrometry

One-dimensional gel electrophoresis and staining of eluted proteins revealed three predominant bands at 44, 49 and 66 kDa and three weaker bands at 25, 30 and 33 kDa (Fig. 2A). A large, unresolved area was observed from 10 to 18 kDa. Two-dimensional gel electrophoresis using IPG strips with pH ranges of 3 to 10 displayed approx. 55 protein spots with pI values of 4.5 to 9.8 and molecular masses of 5 to 67 kDa (Fig. 2B). Acidic low molecular weight proteins appeared as a large unresolved area between pI 4 and 6. For better resolution of this region, samples were separated using more focused IPG strips with a pH-range from 3 to 6. Multiple resolved spots were detected between pI 4.5 and 5.7 (Fig. 2 C; pH-range 4 to 6 shown). Based on their reproducibility on multiple one- and two-dimensional gels, a total of 22 protein spots were selected for
identification by MS. The high molecular weight proteins (spot no. 1-3) were analyzed from the one-dimensional gel since their resolution on two-dimensional gels remained poor.

Reliable peptide sequences were obtained from 20 protein spots and putative identities of 16 of these proteins were determined (Table 1). Twelve protein spots showed high convergence with proteins of *C. sonorensis*, recently identified from a salivary gland cDNA-library, and were related to the D7 family of proteins (no. 16-20) as well as to proteins for sugar meal digestion (no. 1), blood coagulation inhibition (no. 8, 9), defense (no. 3) and novel or unclassified midge proteins (no. 2, 3, 6, 15). Peptides obtained from five spots (no. 4, 7, 13, 14, 20) matched with proteins from species other than *C. sonorensis*. Following peptide sequencing, four proteins remained unidentified (no. 5, 10-12). No reliable sequence data was obtained from two proteins (Fig. 2 B; x₁, x₂) that were therefore not included in the 20 spot analysis.

Comparison of observed and theoretical molecular masses and pI data were consistent for the majority of proteins with significant identity to putative *C. sonorensis* salivary proteins. This illustrates the similarity between *C. nubeculosus* and *C. sonorensis* which are both members of the subgenus *Monoculicoides*. However, the predicted mass of one *C. sonorensis* protein (Q66U72_9DIPT), similar to peptides found in two spots (no. 2, 3), was nearly two times lower than the observed *C. nubeculosus* protein mass. This finding might be due to dimerization of the analyzed protein, or significant differences in the mass of individual proteins between the two species of midges. The predicted mass of *C. nubeculosus* protein from spot no. 15, which was 95% identical to *C. sonorensis* hypothetical proteins Q66U47_9DIPT, Q66U45_9DIPT, Q66U43_9DIPT was nearly two times bigger than its observed mass and might be explained by the fragmentation of the protein during storage or species specific differences. The same protein also showed significant discrepancies for the pI data which could be attributed to post-translational modifications or multiple protein isoforms in addition to the explanations previously mentioned. The calculated and observed molecular mass and pI data of proteins related to insects other than *C.
sonorensis were not always in agreement. These differences contribute to previous findings by Campbell et al. (2005) and are likely to be related to the independent evolvement of the midge’s salivary factors.

3.3. Sugar meal digestion

One of the proteins analyzed from the one-dimensional gel (no. 1) was identified as a maltase (alpha-glucosidase). In total, ten peptides were obtained from the spot comprising 25% of the entire protein sequence of 602 amino acids (Fig. 3). Comparison of the C. nubeculosus sequence with other alpha-glucosidases (maltase-like 1) from C. sonorensis, Aedes aegypti, Aedes albopictus and Culex pipiens revealed that peptides 1 and 5 cover highly conserved regions of the enzyme. The entire C. sonorensis protein sequence (Q66UC_9DIPT) was 50% identical to both Aedes spp. (MALT_AEDAE, Q5MIZ5_AEDAL) and 40% to C. pipiens (Q95WY5_CULPI).

Maltase, amylase, and aldolase facilitate sugar meal digestion. These enzymes have commonly been found in the salivary glands of several different hematophagous insects such as mosquitoes and sand flies (Marinotti et al., 1996; Jacobson and Schlein, 2001). In our investigations we identified peptides from maltase, but none from amylase or aldolase. This might be explained by previous findings of genomic approaches in Anopheles spp. and C. sonorensis that showed maltases and/or maltase-like 1 alpha-glucosidases to be the most abundant members of the family of enzymes involved in sugar meal digestion (Valenzuela et al., 2003; Campbell et al., 2005).

Previous investigations in allergic horses revealed that skin reactivity is not only provoked by Culicoides bites, but also by other hematophagous insects like mosquitoes which may indicate the presence of species-shared allergens (Geiben, 2003). The maltase-like alpha-glucosidase from Aedes aegypti saliva (MALT_AEDAE) was identified as a major allergen in human hypersensitivity to mosquito bites (Peng and Simons, 2004). Since the maltase identified in the midge saliva showed a 50% identity to MALT_AEDAE this protein may be a candidate allergen in this insect known to be the primary cause of allergic dermatitis.
3.4. D7 family

Five protein spots (no. 16-20) were related to proteins of the D7 family. Six peptides from three spots (no. 16-18) matched with a D7 related protein (Q66UB7_9DIPT) and covered 38% of the complete sequence. Although the observed molecular masses of the spots analyzed were all 15 or 16 kDa, the isoelectric points ranged from 4.5 to 5.6 indicating the presence of multiple isoforms of the D7 related protein. Two of the peptide sequences derived from spot no. 19 and 20 were related to putative secreted D7 family salivary proteins of *C. sonorensis* (Q66U39_9DIPT, Q66U41_9DIPT), both of which are very similar in sequence and molecular mass but show slightly different pI values (Campbell et al. 2005). Since the obtained peptides were consistent with both of the *C. sonorensis* sequences, the identification of the proteins was made on the basis of the varying pI data of the spots.

Previous investigations have demonstrated proteins of the D7 family in the saliva or salivary glands of numerous female blood-sucking insects (James et al., 1991; Arcá et al., 1999; Valenzuela et al., 2002). The function of D7-related proteins currently is unknown, however, these proteins share some structural similarities to insect odorant-binding proteins (OBP) and might function as small hydrophobic ligand carriers and/or in binding of host hemostatic factors (Valenzuela et al., 2002). Proteins of the D7 family are the most abundant proteins in the saliva of female mosquitoes (James et al., 1991). The predominance of the D7 related spots of *C. nubeculosus* suggests a similar abundance. A D7 protein of *A. aegypti* is involved in mosquito allergies in man (Peng and Simons, 2004). Similarly, these proteins may function as allergens in horses.

3.5. Coagulation inhibition

Screening of a salivary gland library from *C. sonorensis* revealed three platelet aggregation inhibitors and nine other coagulation inhibitors (Campbell et al., 2005). Surprisingly, only one Kunitz protease inhibitor-like protein was found in *C. nubeculosus* (Q66U91_9DIPT; no. 8, 9). It is possible that artificial collection of saliva may result in the secretion of fewer platelet aggregation
and coagulation inhibitor proteins than would be secreted when feeding on a host. However, prolonged clotting times were seen when human platelets were incubated with collected saliva, clearly indicating the functional presence of these proteins (data not shown). Analysis of C. sonorensis salivary gland extracts revealed that antihemostatic proteins are efficient at the nanogram level (Perez de Leon and Tabachnick, 1996). Since only the most abundant spots of C. nubeculosus were sequenced, proteins involved in coagulation inhibition may be among those not investigated. In addition, the proteins that were not identified (no. 5, 10-12) or the proteins related to yet unclassified or novel midge proteins (no. 2, 3, 6, 15) may have antihemostatic function.

Salivary gland extracts of C. sonorensis were shown to inhibit blood coagulation by interference with extrinsic and intrinsic clotting pathways (Perez de Leon et al., 1998). Previously, an apyrase (platelet aggregation inhibitor) with an estimated molecular mass of 35 kDa and an inhibitor of factor Xa (inhibitor of the blood coagulation cascade) with an estimated molecular mass of 28 kDa were isolated by molecular sieving of salivary gland extracts (Perez de Leon and Tabachnick, 1996; Perez de Leon et al., 1998). The coagulation inhibitor identified in this study had an observed molecular mass of 19 kDa (no. 8, 9). Thus, it is unlikely that the identified C. nubeculosus protein represents one of those previously described even when differences in molecular masses of species-specific variations are taken into consideration. Overall, proteins involved in the inhibition of blood coagulation are considered to be important for the exacerbation of infections with arthropod-borne pathogens and in the induction of allergic reactions (Titus and Ribeiro, 1988; Perez de Leon et al., 1998; Valenzuela, 2002; Peng and Simons, 2004). Therefore, the identified coagulation inhibitor may be a good candidate for further transmission and allergen investigations.

3.6. Defense

Spot 3, taken from the one-dimensional gel, revealed one peptide that had 86% identity to two hypothetical C. sonorensis protein fragments involved in defense (Q66TX6_9DIPT, Q66TX4_9DIPT). A more defined classification was not possible since the molecular masses of the
complete proteins were not available from the database and the pI data were not obtained in this study. A second peptide obtained from spot 3 was also identified from spot 2. Sequence alignment revealed that this peptide not only displays an identity of 89% to an unclassified hypothetical protein (Q66U72_9DIPT), but also an identity of 86% to the hypothetical defense protein fragments mentioned above. Therefore, it is possible that both peptides represent the same hypothetical defense protein fragments. An additional peptide was sequenced from spot 3 with 67% identity to an unclassified C. sonorensis hypothetical protein fragment (Q66U84_9DIPT) indicating the presence of more than one protein in the spot.

Arthropods must protect themselves against a variety of ingested and environmental pathogens. Antibacterial and antifungal compounds were identified in A. aegypti, Triatoma infestans and D. melanogaster (Rossignol and Lueders, 1986; Ferrandon et al., 1998; Amino et al., 2001). Additionally, the expression of defense molecules is upregulated in different Anopheles spp. after the uptake of Plasmodium spp. (Dimopoulos et al., 1998; Shandilya et al., 1999). These molecules are assumed to induce parasite tolerance in the insects and to enhance the capability of parasite transmission. Potentially, defense molecules of Culicoides spp. could be involved in similar roles in pathogen transmission.

3.7. C. nubeculosus proteins with identities to proteins of non-C. sonorensis insect species

Peptides sequenced from three spots (no. 4, 14, 20) had identities to proteins from insects other than Culicoides spp. The peptide derived from spot 4 showed high identity with unclassified proteins from D. pseudoobscura (Q296E2_DROPS), D. melanogaster (Q9VAH5_DROME), and A. gambiæ (Q7PR47_ANOGA). Search of the databases did not reveal a reliable putative function for these proteins. The peptide sequenced from spot 14 was related to a carboxypeptidase from Drosophila spp. (Q9VL21_DROME, Q6XHU5_DROYA, Q5BI80_DROME). The potential function of a salivary carboxypeptidase is well described in ticks where it acts to destroy the pain mediator bradykinin which is released subsequent to tissue damage (Ribeiro et al., 1985; Ribeiro
and Mather, 1998). Since *Culicoides* spp. feed on their hosts for only a few minutes and cause minor tissue damage, it may be unnecessary to eliminate the host’s nociception. A more likely role is protein digestion, identified for a carboxypeptidase recently isolated from the larval salivary glands of the wheat blossom midge, *Sitodiplosis mosellana* (Mittapalli et al., 2005). A similar function was suggested for the carboxypeptidases identified in the midgut of *C. sonorensis* (Campbell et al., 2005).

A peptide sequenced from spot 20 showed 60% identity to a chemical sense-related lipophilic ligand binding protein (CRLBP) of the black blow fly, *Phormia regina* (Q5NTY8_9DIPT). Observed and predicted molecular masses were in agreement whereas pI data showed a slight difference. CRLBP proteins represent a subfamily of odorant binding proteins and have been identified in bees, beetles and non-hematophagous flies (Scafe et al., 2006). Our findings indicate that CRLBP proteins might also be present in *Culicoides* spp.

### 3.8. *C. nubeculosus* proteins with identities to proteins of non-insect species

Peptides sequenced from two proteins (no. 7, 13) had identities to enzymes from species other than insects. Although the molecular masses and pI data were not in agreement, the protein sequence from spot 7 had 58% identity to a fungal cytochrome P450 monoxygenase (P450) and 50% identity to a bacterial P450. P450’s have been shown in other insects and function in detoxification or metabolism of food components (Li et al., 2003; Tilquin et al., 2004). Moreover, they were described in human lice in relation to the development of permethrin resistance (Hemingway et al., 1999). Their role in *Culicoides* spp. remains unclear.

A peptide sequenced from spot 13 had 56% identity with a peroxidase precursor of barley (O49866_HORVU) and 50% identity with a peroxidase precursor of wheat (Q5GMP4_WHEAT). Peroxidases in the salivary glands of female mosquitoes have been reported to destroy biogenic amines in order to inhibit vasoconstriction in the host (Ribeiro and Nussenzveig, 1993). A peroxidase isolated from ticks is suggested to be involved in the detoxification of reactive oxygen...
species generated during metabolism (Das et al., 2001). In addition, this protein was shown to be an immunodominant antigen that elicits strong antibody responses. Therefore, the putative peroxidase precursor isolated from the midge saliva might be another promising candidate for both investigations in pathogen transmission and hypersensitivity.

4. Conclusions

In this study we describe the development of a novel artificial collection method that allows the production of large amounts of purified salivary proteins of Culicoides spp. Saliva obtained with this method was used for the first, partial analysis of the salivary proteome of C. nubeculosus. In total, 16 proteins were identified, five of which have potential involvement in pathogen transmission and/or allergic reactions to midge bites. These proteins might be candidates for further investigations in both the vector competence of, and the hypersensitivity to Culicoides spp. In addition, a number of as yet uncharacterized or novel midge proteins have been isolated and will require further characterization. Twelve of the proteins present in the saliva were related to salivary compounds of C. sonorensis whereas nine proteins showed identities with proteins of other species or were not identified. Considering that C. sonorensis in comparison to C. nubeculosus is capable of transmitting a large number of arboviral livestock pathogens, these differences may represent significant grounds for investigation with regard either to vector competence or in defining components that modulate host immune responses.

Acknowledgements

The authors thank Susanne Pohl and Harald S. Conradt for advice and technical assistance and Simon Carpenter and Melissa A. Stuart for critical reading of the manuscript. This work was supported by a scholarship of the Karl-Enigk-Stiftung awarded to K. F. A. Langner.
References


Table 1

Mass spectrometry identification of salivary proteins from *C. nubeculosus*

<table>
<thead>
<tr>
<th>Spot no.</th>
<th>Peptide sequences</th>
<th>% Id</th>
<th>Protein identification, GenBank accession number</th>
<th>Putative function</th>
<th>Observed&lt;sup&gt;d&lt;/sup&gt; M&lt;sub&gt;r&lt;/sub&gt;/pI</th>
<th>Predicted&lt;sup&gt;d&lt;/sup&gt; M&lt;sub&gt;r&lt;/sub&gt;/pI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SFMDSDGIDVGDGLK</td>
<td>100</td>
<td>Malate <em>Culicoides sonorensis</em> (<em>C. s.</em>), Q66UC5_9DIPT</td>
<td>Sugar meal digestion (Maltase)</td>
<td>66/n.d.&lt;sup&gt;f&lt;/sup&gt;</td>
<td>68/5.0</td>
</tr>
<tr>
<td></td>
<td>VFPQFGLSLSSDELAQFLER</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NLPTNWNVSAFR</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SSAWEWNER</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GEYLYHQFLAEQPDNLNYR</td>
<td>94</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VLMVEAYAPLTK</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TVDPQACTNPNIFHAK</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TPMWNAQK</td>
<td>78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>YGTVDYSLANDDLVVIK</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TLIAVLNLGFEQVVNLNLNER</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>FNQVYFTQEAQFSNKS</td>
<td>89</td>
<td>Hypothetical protein <em>C. s.</em>, Q66U72_9DIPT</td>
<td>Unclassified</td>
<td>49/n.d.</td>
<td>21.7/7.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>FNQVYFTQEAQFSNKS</td>
<td>89</td>
<td>Hypothetical protein <em>C. s.</em>, Q66U72_9DIPT</td>
<td>Unclassified</td>
<td>44/n.d.</td>
<td>21.7/7.8</td>
</tr>
<tr>
<td></td>
<td>QLGFTGELYIYGNK</td>
<td>86</td>
<td>Hypothetical protein [Fragment] <em>C. s.</em>, Q66TX6_9DIPT</td>
<td>Defense</td>
<td>44/n.d.</td>
<td>25.5/8.1</td>
</tr>
<tr>
<td></td>
<td>ALGVETKFDLSTGK</td>
<td>67</td>
<td>Hypothetical protein [Fragment] <em>C. s.</em>, Q66U84_9DIPT</td>
<td>Unclassified</td>
<td>44/n.d.</td>
<td>17.2/9.8</td>
</tr>
<tr>
<td>4</td>
<td>FVPTIVNEQFQELQDQLR</td>
<td>76</td>
<td>GA21779-PA [Fragment] <em>Drosophila (D.) pseudoobscura</em>, Q296E2_DROPS</td>
<td>Unclassified</td>
<td>25/5.1</td>
<td>21.9/5.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>71</td>
<td>CG9427-PA (GH19763p) <em>D. melanogaster</em>, Q9VAH5_DROME</td>
<td>Unclassified</td>
<td>25/5.1</td>
<td>23.6/6.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>67</td>
<td>ENSANGP00000018265 [Fragment] <em>Anopheles gambiae</em>, Q7PR47 ANOGA</td>
<td>Unclassified</td>
<td>25/5.1</td>
<td>17.8/4.7</td>
</tr>
<tr>
<td>5</td>
<td>LALVNLSDLFNR</td>
<td>n.i.&lt;sup&gt;s&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>21/6.3</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>GANIQYFTSLK</td>
<td>91</td>
<td>Hypothetical protein <em>C. s.</em>, Q66U35_9DIPT</td>
<td>Novel</td>
<td>19/6.9</td>
<td>17.3/6.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>91</td>
<td>Hypothetical protein <em>C. s.</em>, Q66U37_9DIPT</td>
<td>Novel</td>
<td>19/6.9</td>
<td>17.8/6.0</td>
</tr>
<tr>
<td>7</td>
<td>VFGVGLYSCSR</td>
<td>58</td>
<td>Cytochrome P450 monoxygenase <em>Penicillium pailli</em>, Q9C449_PENPX</td>
<td>Cytochrome P450 family</td>
<td>19/7.5</td>
<td>58.8/8.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>Cytochrome P450-like putative monooxygenase <em>Azorarcus sp.</em>, Q5P87_TAOSE</td>
<td>Cytochrome P450 family</td>
<td>19/7.5</td>
<td>18.5/4.7</td>
</tr>
<tr>
<td>8</td>
<td>YNVQNTNCFYSLTK</td>
<td>57</td>
<td>Putative secreted salivary protein <em>C. s.</em>, Q66U91_9DIPT</td>
<td>Coagulation inhibitor</td>
<td>19/9.0</td>
<td>22.3/9.0</td>
</tr>
<tr>
<td>9</td>
<td>YNVQNTNCFYSLTK</td>
<td>57</td>
<td>Putative secreted salivary protein <em>C. s.</em>, Q66U91_9DIPT</td>
<td>Coagulation inhibitor</td>
<td>19/9.7</td>
<td>22.3/9.0</td>
</tr>
<tr>
<td>10</td>
<td>GSYLSLDGDKETR</td>
<td>n.i.</td>
<td></td>
<td></td>
<td>16/6.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LLADSTGVVVLDDDGK</td>
<td>n.i.</td>
<td></td>
<td></td>
<td>15/6.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VAWSNVMCAGDFWLFLYDR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>QDAFVGLYSCCLR</td>
<td>n.i.</td>
<td></td>
<td></td>
<td>15/6.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LSSPATLNSR</td>
<td>n.i.</td>
<td></td>
<td></td>
<td>15/7.1</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>LLADSTGVVVLDDDGK</td>
<td>n.i.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>QDWYYTLENOQYK</td>
<td>n.i.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EANFYVQSR</td>
<td>n.i.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued on next page)
<table>
<thead>
<tr>
<th>Spot no.</th>
<th>Peptide sequences</th>
<th>% Id</th>
<th>Protein identification, GenBank accession number</th>
<th>Putative function</th>
<th>Observed&lt;sup&gt;d&lt;/sup&gt; M&lt;sub&gt;r&lt;/sub&gt;/pI</th>
<th>Predicted&lt;sup&gt;e&lt;/sup&gt; M&lt;sub&gt;r&lt;/sub&gt;/pI</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>QGDGTELSELDPQSAR</td>
<td>56</td>
<td>Peroxidase precursor <em>Hordeum vulgare</em>, O49866_HORVU</td>
<td>Peroxidase</td>
<td>12/7.4</td>
<td>36.6/5.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>Peroxidase precursor <em>Triticum aestivum</em>, Q5GMP4_WHEAT</td>
<td>Peroxidase</td>
<td>12/7.4</td>
<td>36.5/5.6</td>
</tr>
<tr>
<td>14</td>
<td>YPASSASVPWVLDEDG</td>
<td>50</td>
<td>CG7025-PA <em>D. melanogaster</em> Q9VL21_DROME</td>
<td>Carboxypeptidase</td>
<td>14/9.7</td>
<td>48.3/4.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>Similar to <em>D. melanogaster</em> CG7025 [Fragment] <em>D. yakuba</em>, Q6XHU5_DROYA</td>
<td>Carboxypeptidase</td>
<td>14/9.7</td>
<td>21.9/4.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>LP21640p [Fragment] <em>D. melanogaster</em>, Q5BI80_DROME</td>
<td>Carboxypeptidase</td>
<td>14/10</td>
<td>48.5/4.8</td>
</tr>
<tr>
<td>15</td>
<td>LYQGLLSDSYNLALEQYMK</td>
<td>95</td>
<td>Hypothetical protein <em>C. s.</em> Q66U47_9DIPT, Q66U45_9DIPT, Q66U43_9DIPT</td>
<td>Novel</td>
<td>10/9.7</td>
<td>18.3/8.5</td>
</tr>
<tr>
<td>16</td>
<td>ADELYACLLNADGQK</td>
<td>93</td>
<td>D7 related protein <em>C. s.</em> Q66UB7_9DIPT</td>
<td>D7 family</td>
<td>16/4.5</td>
<td>15.3/4.7</td>
</tr>
<tr>
<td></td>
<td>PSGDQYDTDNLLVR</td>
<td>93</td>
<td>D7 related protein <em>C. s.</em> Q66UB7_9DIPT</td>
<td>D7 family</td>
<td>16/4.9</td>
<td>15.3/4.7</td>
</tr>
<tr>
<td></td>
<td>DCEEEKDLPA</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>DEAPSGDQYDTDNLLK</td>
<td>81</td>
<td>D7 related protein <em>C. s.</em> Q66UB7_9DIPT</td>
<td>D7 family</td>
<td>15/4.7</td>
<td>15.3/4.7</td>
</tr>
<tr>
<td>18</td>
<td>PSGDQYDTDNLLK</td>
<td>100</td>
<td>D7 related protein <em>C. s.</em> Q66UB7_9DIPT</td>
<td>D7 family</td>
<td>16/4.9</td>
<td>15.3/4.7</td>
</tr>
<tr>
<td></td>
<td>VPANPTECVIDCIFQK</td>
<td>88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>YYEVENIFQAWEK</td>
<td>96</td>
<td>Putative secreted salivary protein <em>C. s.</em> Q66U41_9DIPT</td>
<td>D7 family</td>
<td>15/5.4</td>
<td>17.8/5.4</td>
</tr>
<tr>
<td>20</td>
<td>QCFICLCELTNLNILK</td>
<td>93</td>
<td>Putative secreted salivary protein <em>C. s.</em> Q66U39_9DIPT</td>
<td>D7 family</td>
<td>16/5.6</td>
<td>17.8/5.6</td>
</tr>
<tr>
<td></td>
<td>ISECNSLELYQDNCK</td>
<td>60</td>
<td>Chemical sense-related lipophilic ligand-binding homologous protein <em>Phormia regina</em> (black blowfly), Q5NTY8_9DIPT</td>
<td></td>
<td>16/5.4</td>
<td>16.4/4.9</td>
</tr>
</tbody>
</table>

<sup>a</sup> Sequence information obtained from peptide fragmentation data. The isobaric amino acids isoleucine (I) and leucine (L) cannot be distinguished using Q-TOF mass spectrometry. Hence, I and L are interchangeable in the peptide sequences reported here. When possible, translated ESTs were used to assign either I or L in the peptide sequence.

<sup>b</sup> Calculated identity of obtained peptide sequences to proteins from the database using the AlignX software (Invitrogen).

<sup>c</sup> Proteins were identified by searching the UniProt database with peptide fragmentation data using the Fasta program (European Bioinformatics Institute, Cambridge, UK).

<sup>d</sup> Observed molecular masses (M<sub>r</sub>) in kDa and isoelectric points (pI) estimated from protein spots on gel.

<sup>e</sup> Predicted molecular masses (M<sub>r</sub>) in kDa and isoelectric points (pI) from the database.

<sup>f</sup> n.d., not determined.

<sup>g</sup> n.i., not identified.
Figure Captions

Fig. 1. Saliva of *C. nubeculosus* was obtained using a modified artificial membrane feeding system. Midge deposited saliva proteins in the Durapore™ filters attracted by 37 °C horse blood in the inner glass chamber. Contamination with non-saliva products was avoided by covering the filters with Parafilm™.

Fig. 2. Representative 1D and 2D gel separation of midge saliva on SDS polyacrylamide gels. (A) One-dimensional gel electrophoresis. (B) Two-dimensional gel electrophoresis (2D) using IPG-strips in a pH range from 3 to 10. (C) 2D gel electrophoresis using IPG-strips with a pH range from 3 to 6 for a better resolution of acidic, low molecular mass proteins (pH 4 to 6 shown). Gels containing 15% polyacrylamide were used for one-dimensional gel electrophoresis and second dimension of 2D. Molecular masses in kDa are indicated on the left lane, pI data on the top lane. The assigned proteins were used for mass spectrometry. Proteins for these images were visualized by silver staining.

Fig. 3. Peptide sequence alignment of maltase, an allergen candidate identified from the saliva of *C. nubeculosus*. *C. nubeculosus* peptide sequences (bold, no. 1-10) were aligned with alpha-glucosidases from different insect species: Q66UC5_9DIPT - maltase *C. sonorensis*, MALT_AEDAE - probable maltase precursor *Aedes aegypti*, Q5MIZ5_AEDAL - probable salivary maltase *Aedes albopictus*, Q95WY5_CULPI - binary toxin-binding alpha-glucosidase *Culex pipiens*. Dark grey areas indicate sequence regions conserved in all species shown. Light grey areas indicate similarity of the obtained peptide sequences of *C. nubeculosus* to at least one of the other four alpha-glucosidases.
Box with *C. nubeculosus*

Gauze as pot lid

Inner glass chamber filled with horse blood

Outer glass chamber with warm water circulation

Outer Parafilm™

Durapore™ filter

Inner Parafilm™ (double layer)
Fig. 2