Cuticular Structure Proteomics in the Pea Aphid Acyrthosiphon pisum Reveals New Plant Virus Receptor Candidates at the Tip of Maxillary Stylets

Maelle Deshoux, Victor Masson, Karim Arafah, Sebastien Voisin, Natalia Guschinskaya, Manuella van Munster, Bastien Cayrol, Craig Webster, Yvan Rahbé, Stéphane Blanc, et al.

Journal of proteome research

pubs.acs.org/jpr

Article

Cuticular Structure Proteomics in the Pea Aphid *Acyrthosiphon pisum* Reveals New Plant Virus Receptor Candidates at the Tip of Maxillary Stylets

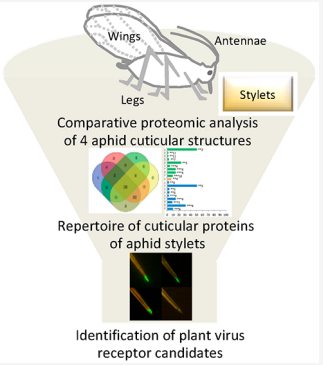
Maëlle Deshoux,[‡] Victor Masson,[‡] Karim Arafah, Sébastien Voisin, Natalia Guschinskaya, Manuella van Munster, Bastien Cayrol, Craig G. Webster, Yvan Rahbé, Stéphane Blanc, Philippe Bulet,*[,‡] and Marilyne Uzest*[,‡]

Cite This: https://dx.doi.org/10.1021/acs.jproteome.9b00851

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Aphids are phloem-feeding insects known as major pests in agriculture that are able to transmit hundreds of plant viruses. The majority of these viruses, classified as noncirculative, are retained and transported on the inner surface of the cuticle of the needle-like mouthparts while the aphids move from plant to plant. Identification of receptors of viruses within insect vectors is a key challenge because they are promising targets for alternative control strategies. The acrostyle, an organ discovered earlier within the common food/salivary canal at the tip of aphid maxillary stylets, displays proteins at the cuticle−fluid interface, some of which are receptors of noncirculative viruses. To assess the presence of stylet- and acrostyle-specific proteins and identify putative receptors, we have developed a comprehensive comparative analysis of the proteomes of four cuticular anatomical structures of the pea aphid, stylets, antennae, legs, and wings. In addition, we performed systematic immunolabeling detection of the cuticular proteins identified by mass spectrometry in dissected stylets. We thereby establish the first proteome of stylets of an insect and determine the minimal repertoire of the cuticular proteins composing the acrostyle. Most importantly, we propose a short list of plant virus receptor candidates, among which RR-1 proteins are remarkably predominant. The data are available via ProteomeXchange (PXD016517).



KEYWORDS: *cuticle, cuticular proteins, stylets, acrostyle, virus receptor, proteomics, aphid*

INTRODUCTION

Aphids are phloem-feeding insects, well-known as major pests in agriculture. More than 5,000 aphid species have been described. They colonize countless plant species and have been reported on 300 plant families ranging from gymnosperms to angiosperms.[1,2] Aphids have a complex life cycle alternating between sexual and asexual reproduction and seasonal host changes. They are and have long been extensively studied, not only because of interesting life traits such as reproductive and wing polyphenisms but also because they transmit numerous plant diseases.[3] With more than 300 species transmitted, aphids are one of the most efficient and important vectors of plant viruses and are the best studied-model to characterize the mechanisms of vector-mediated virus transmission (recently reviewed by Whitfield and colleagues).[4−7] Most aphid-transmitted viruses bind reversibly to retention sites on the inner cuticle of insect mouthparts to which they remain attached during transport to a new host plant.[4] This so-called noncirculative virus transmission is particularly difficult to limit and control in field conditions, as the viruses can be acquired and inoculated by nonresident aphids mostly within a single puncture lasting only a few seconds.[8] During this process, an intimate association occurs between viruses and their vectors. This interaction is highly specific and involves the capsid proteins or viral-encoded proteins called helper components, together with poorly characterized molecular compounds in the insect cuticle acting as receptors.[9−12] These vector molecules are promising targets for alternative viral control strategies, and their identification would help characterizing further the molecular mechanisms of virus−vector interaction.

Aphids have piercing-sucking mouthparts, composed of the short triangular labrum covering the base of the stylet bundle, and the labium, a segmented organ which contracts to facilitate stylets penetration into plant tissues. The stylet bundle arises in the head from its secreting glands and extends outside the head in a dorsal groove of the labium.[13,14] A pair of external mandibular stylets innervated by two dendrites, surround two (noninnervated) inner maxillary stylets and together form the stylet bundle. Specific anatomical features are visible on both types of stylets. Barb-like ridges are present at the tip of

Received: December 20, 2019
Published: January 28, 2020

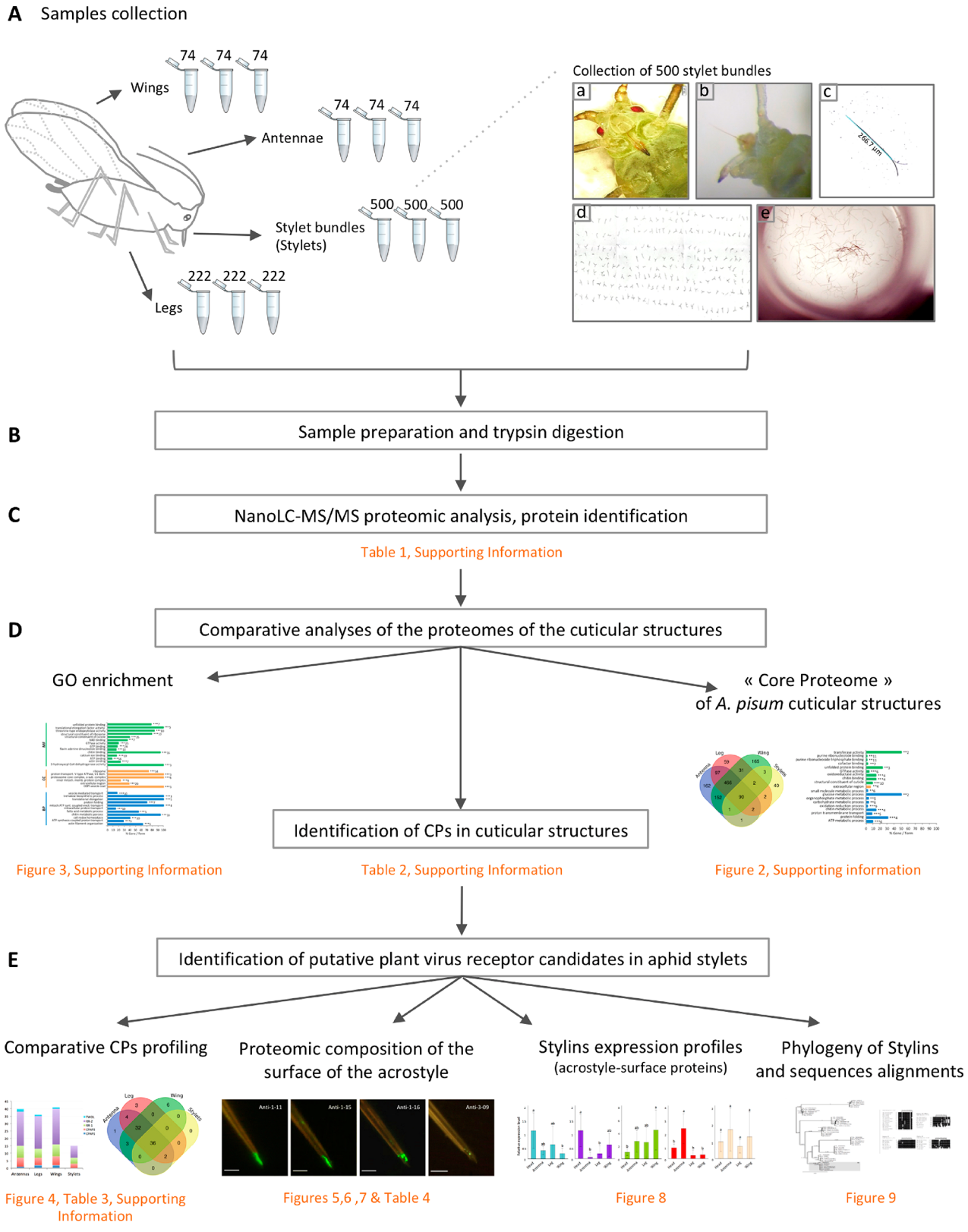


Figure 1. Experimental workflow to characterize the proteomics composition of the acrostyle and identify plant virus receptor candidates within aphid stylets.

mandibular stylets to facilitate penetration into plant tissues.[15,16] Maxillary stylets are tightly interlocked by a series of longitudinal ridges and grooves, which enclose the food and salivary canals formed by opposing grooves. Food and salivary canals fuse a few micrometers from the tip, leading to the confluent common food/salivary canal, that has long been believed to harbor receptors of noncirculative viruses.[8,17−20] The existence of such receptors has definitely been evidenced with *Cauliflower mosaic*

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

virus (CaMV), which was shown to bind to cuticular proteins (CPs) accessible at the tip of maxillary stylets, at the surface of a specific organ designated the acrostyle.[21] This organ is a distinct anatomical structure restricted to a region of the common canal. Its surface properties are different from the rest of the stylet cuticle.[22] Antibodies targeting CPs from the CPR family—the largest CP family widespread among arthropods and containing the Rebers and Riddiford (RR) consensus sequence—allowed the detection of several peptides from RR-1 and RR-2 subgroups in the acrostyle.[22−26] More specifically, two highly homologous RR-1 proteins, Stylin-01 and Stylin-02, have been localized in the organ with a peptide corresponding to their common C-terminus sequence directly accessible at the surface. *In vitro* competition assays and *in vivo* silencing approaches indicated that Stylin-01 was involved in CaMV transmission.[26] However, its role in the transmission of other noncirculative viruses could not be established, and there is currently no evidence for a common receptor of all noncirculative viruses. The acrostyle has a complex proteomic composition and contains at least two RR-1 and several RR-2 proteins which could not be specifically identified.[25] Many RR-2 have nearly identical sequences that are not distinguishable by specific antibodies, highlighting the limit of an immunolabeling approach to comprehensively determine the proteome of aphid stylets, and the need for larger scale proteomic characterization.

Annotated CPs in the genome of the pea aphid *Acyrthosiphon pisum* account for 150 proteins. They have been assigned to six distinct CP families according to their conserved consensus motifs: 125 CPR (15 RR-1 and 110 RR-2), 11 CPAP1, 8 CPAP3, 3 TWDL, 2 CPF, and 1 CPCFC proteins.[27] Up to now, mass-spectrometry (MS)-based proteomics on stylet bundles was inconceivable considering the size of these appendages. Moreover, studies of the proteomic composition of larger cuticular structures have not been reported for aphids. However, the recent advances in MS approaches allowing the extraction of CPs from small quantities of insect tissues now make stylet proteomics a realistic objective.[27]

Our aim in this proteomics study was to identify through a bottom-up approach proteins and related peptides accessible at the surface of aphid stylets that might play a key role in plant virus binding. To draw-up a list of stylet-specific CPs, we first developed a comparative analysis of the proteomes of four distinct pea aphid tissues partially composed of cuticle: antennae, legs, wings, and stylets. In a second phase, we produced antibodies targeting peptides from the stylet CPs identified by proteomics to refine their localization within stylets through *in situ* immunolabeling. As a result, we here provide a short list of plant virus receptor candidates highly conserved among aphid species and characterize their expression pattern in various body parts and throughout aphid developmental stages. These candidate receptors, named Stylins, have a patchy distribution within and along aphid stylets but are emerging at the surface of the acrostyle. Most have a RR-1 type chitin-binding domain, highlighting the remarkable accessibility of CPs of this subgroup at the surface of the cuticle.

EXPERIMENTAL PROCEDURES

The experimental design and the organization of the presentation of the results are summarized in the workflow shown in Figure 1.

Chemical Reagents

For sample preparation, all reagents were purchased from Sigma-Aldrich (St. Louis, MO) except *Rapi*Gest SF surfactant purchased from Waters (Milford, MA). For LC-MS/MS analysis, formic acid was obtained from Fluka (Sigma-Aldrich, St. Louis, MO), Milli-Q water from Merck (Merck Millipore, Billerica, MA), and all other chemicals were purchased from Carlo-Erba Reagents (Val de Reuil, France). Modified sequencing-grade trypsin (Promega Corporation, Madison, WI) was used for protein digestion.

Aphid Clones

A colony of *Acyrthosiphon pisum* (LL01) was maintained on *Vicia faba* cv Robin Hood in an environmental growth chamber at 23/18 °C (day/night) with a photoperiod of 16/8 h (day/night). *Myzus persicae* Sulzer was maintained on *Solanum melongena* cv Barbentane in the same temperature/photoperiod conditions.

Tissue Collection for Liquid Chromatography Coupled to Tandem Mass Spectrometry (LC-MS/MS) Analyses

Alate *A. pisum* adults collected from *Vicia faba* plants were starved for 1 h and stored at −20 °C for several hours before dissection. Collection of antennae, legs, and wings did not present any technical difficulties. These appendages were collected in triplicate for each structure, each replicate comprising antennae, wings, or legs from 37 insects. Stylet bundles, hereafter named stylets, are approximately 700 μm in length and 3 μm in diameter for *A. pisum* adults. They are anchored in aphid heads in the glands from which they are secreted, the retort organs (described by Ponsen and Guschinskaya and colleagues).[14,28] To extract exploitable information from MS analysis, hundreds of stylets had to be pooled for each trial as free as possible of other contaminating aphid tissues. In addition, we had to reduce electrostatic charges that too often induced stylets fly off and loss. For these reasons, a "clean-stylet" dissection protocol compatible with MS standards had to be specifically designed (Figure 1A). Frozen aphids were first glued on their back onto a microscope slide using double-sided bonding tape (Figure 1Aa). Stylets were then pulled out from the labium under a stereomicroscope using tungsten insect pins (Figure 1Ab), separated from the head with microscissors (Figure 1Ac), and then transferred onto a conductive glass slide (ITO, Bruker Daltonik, Germany), allowing secured accumulation of stylets (Figure 1Ad). Dissecting tools were carefully washed with water, followed by 70% ethanol, and finally dried on lint-free paper after each stylet collection. Once 500 stylet bundles had been accumulated onto the ITO-slide, they were carefully transferred with a fine needle into a single glass microtube, avoiding static charge effects, allowing control of the number of appendages deposited per tube (Figure 1Ae). This protocol was systematically used to collect three independent batches of 500 stylets each.

Sample Preparation

All cuticular structures were prepared according to the same protocol. Samples were first washed for 15 min with Milli-Q water and 50% acetonitrile (ACN), both acidified with 0.1% trifluoroacetic acid (TFA) (v/v, final concentration). Supernatants were then removed after centrifugation, and the cuticular structures dried out by centrifugation under vacuum (Labconco, Kansas City, USA). All anatomical structures were then treated as described previously.[27] Briefly, proteins were extracted by successive incubations in pure hexafluoroisopopa-

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

nol (HFIP) and in 50 mM ammonium bicarbonate supplemented with 0.1% *Rapi*Gest. After reduction and alkylation, proteins were submitted to trypsin digestion. Finally, digested samples were dried out and resuspended in 2% ACN/0.1% TFA (v/v) prior to mass spectrometry (MS) analysis.

Nano Liquid Chromatography Coupled to Tandem Mass Spectrometry (nanoLC-MS/MS) Analysis

NanoLC-MS/MS was carried out using an Ultimate 3000 nanoHPLC (Thermo Scientific, Germany) for the separation, hyphenated to a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific). For chromatography, the digested samples were loaded, concentrated, and washed at 10 μL/min for 6 min with 2% ACN and 0.05% TFA on a microconcentrating column (300 μm × 5 mm PepMap 100, C_{18}, 5 μm, Thermo Scientific). The separation was performed on a reversed-phase column (75 μm × 250 mm Acclaim PepMap 100, C_{18}, particle size 3 μm nanoviper column from Thermo Scientific). The LC mobile phases for the separation were water (A) and ACN (B), each supplemented with 0.1% formic acid (v/v). Separation was achieved at a flow rate of 300 nL/min using a biphasic linear gradient from 2% to 32% B in 100 min and from 32% to 65% B in 5 min. MS analysis was carried out in positive ion and data-dependent modes. The voltage applied to the nanotips (Nano Objective, USA) was approximately 2.0 kV, and the header was at 300 °C. Full scan (MS) spectra were obtained from 380 to 2,000 m/z (70,000 resolution, AGC target 3×10^6, maximum IT 200 ms), and for each full-scan the most intense ions (Top 10) were fragmented in MS^2 (17,500 resolution, AGC target 2×10^5, maximum IT 100 ms, intensity threshold 4×10^4, excluding charge-unassigned ions, Normalized Collision Energy selected at 27). Parent ions were then excluded from MS/MS for the next 15 s. The softwares Chromeleon Xpress (Thermo Fisher Scientific) was used to control the HPLC, and Xcalibur 2.2 (Thermo Fisher Scientific) to control the mass spectrometer.

Database Searching and Protein Identification

The Sequest HT searching algorithm was run by Proteome Discoverer 1.4 (Thermo Fisher Scientific) to match the acquired MS/MS spectra to a protein database, with the following settings: trypsin digest with two maximum missed cleavages, 6 and 144 amino acids as minimum and maximum peptide length, respectively, a tolerance of 10 ppm/0.02 Da for precursors and fragment ions, respectively. Cysteine carbamidomethylation was set as a fixed modification; C-terminal protein amidation and methionine and tryptophan oxidation were set as variable modifications.

Searches were performed against combined proteome databases (Table S1) including: (i) the pea aphid proteome (http://aphidbase.com, version 2.0 containing 33,291 protein-coding sequences); (ii) an eubacterial reference proteome setup for putative symbionts and aphid microbiome-related sequences (116,983 sequences); (iii) a set of sequences of entomopathogenic fungi (28,431 sequences); (iv) a set of viral sequences including aphid viruses and aphid-vectored broad-bean viruses were also selected (6,614 sequences); (v) a mixed plant proteome assembled from available *Fabacae* proteomes (66,777 sequences); and (vi) a set of common contaminants (116 sequences).

A homemade *A. pisum* CP database comprising 150 annotated CPs was constructed using the Cuticular Protein Family Prediction Tool CutProtFam-Pred (http://aias.biol.uoa.gr/CutProtFam-Pred/home.php) to accurately detect and classify putative CPs present in the pea aphid proteome v2.0, thanks to known consensus and semiconsensus sequences (Table S2).[29] Identification of CPs in cuticular structures was carried out with this homemade database.

Data Presentation

To identify the proteins that are present in pea aphid cuticular anatomical structures, pools of antennae, wings, legs, and stylet bundles were collected in triplicate on *A. pisum* adults. Proteins consistently identified in all three biological replicates with at least a same peptide in each replicate, regardless of being "Unique" (peptide present in only one protein) or "Shared" (peptide present in multiple different proteins), were considered as confidently identified and constituted whole proteomes of the cuticular structures. In these first data sets, proteins sharing a same identified peptide were grouped under a single protein identifier (accession number corresponding to the protein with the highest score and highest percentage of coverage of the protein. These parameters can differ from one sample to another, and a same "shared" peptide may be assigned to different protein identifiers by the algorithm). Comparative analyses of whole proteomes were carried out using these data sets.

Minimal lists of CPs were then retrieved from whole proteomes using homemade CP database (Table S2). *Minimal lists* represent the smallest number of CPs that can explain all identified MS peptides and will be used to design peptides for antibodies production and immunolabeling of dissected stylets. However, CPs can share high sequence similarity, particularly proteins of the CPR family.[26,30−33] Several proteins can be identified with a same tryptic fragment ("shared" peptide). In this case, we cannot tell how many of these proteins are actually incorporated into the cuticle of the anatomical structures. To be fully exhaustive in our quest for virus receptor candidates, lists of CPs *potentially present* in cuticular structures were extended to all proteins containing these "shared" peptides.

Functional Classification

Gene Ontology (GO) terms were assigned to AphidBase 2.0 (http://bipaa.genouest.org/is/aphidbase/acyrthosiphon_pisum/downloads/) via Blast2GO software version 5.1.12 using the InterProScan with the Hidden Markov Models (HMMs) present in the PFAM protein family database.[34] The number of proteins identified in each cuticular structure was compared to all corresponding functionally GO annotated terms in the entire pea aphid (*A. pisum*). GO term enrichment was determined via Fisher's Exact test at an FDR, p-value $\leq$0.05; the reduced enrichment was determined at an FDR, p-value $\leq$0.01.

Antibodies

A total of 26 antibodies were used to immunolabel aphid stylets. Eleven antibodies were already available from previous studies.[22,25,26] In addition, 15 peptides of 13−16 amino acids length with sequences originating from peptides retrieved from MS analyses (either "Unique" or "Shared" peptides), or originating from the full-length sequences of the proteins identified in this study were synthesized by Eurogentec (http://www.eurogentec.com). Eurogentec also carried out the production of antibodies in rabbit and the affinity purification of the antisera. Alexa Fluor 488-conjugated antirabbit IgG (A11070, Thermo Fisher Scientific,Waltham, MA) were used as secondary antibodies.

Immunolabeling of Dissected Stylets

A. pisum and *M. persicae* stylets were dissected and immunolabeled according to Webster and colleagues, using

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

primary antibodies at dilutions of 1:200 and secondary Alexa fluor-conjugated antibody at a dilution of 1:400.[25] Stylets were either untreated or treated with 2 U/mL of Chitinase from *Streptococcus griseus* (Sigma-Aldrich, St. Louis, MO) before their incubation with primary antibodies, as described by Uzest and colleagues, to eventually reveal epitopes that could be localized under the surface layers of the stylet cuticle.[21] When no labeling was observed after 15 min of digestion, the Chitinase treatment was extended to 30 and 60 min. Two independent repeats were performed per condition and per antibody.

Quantification of Stylin Transcripts in Various Tissues and at Different Developmental Stages

Transcripts of stylins were quantified in the heads of all *A. pisum* nymphal instars (N1 to N4) and adult stages and in different tissues. Insect tissues (antennae, legs, wings, and antenna-free heads as a proxy for stylet synthesizing glands) were collected from adults under a stereomicroscope. Six pools of 10 aphids/tissues were tested per condition. Total RNA was extracted from whole aphids or dissected tissues using an RNeasy minikit (Qiagen, Hilden, Germany). A total of 80−100 ng of RNA was treated with RQ1 RNase-free DNase I (Promega Corporation, Madison, WI). First-strand cDNA was synthesized using Moloney murine leukemia virus (MMLV) reverse transcriptase (Promega Corporation) according to the manufacturer's instructions, with oligo(dT) as primers. All RT-qPCRs were performed in duplicates on a LightCycler 480 instrument (Roche, Basel, Switzerland) using 1:4 dituted cDNAs and a LightCycler 480 SYBR green I master mix (Roche, Penzberg, Germany) according to the manufacturer's recommendations with gene-specific primers (Table S3). Two internal reference genes encoding actin and elongation factor 1α (EF1α) from *A. pisum* were used for normalization. Amplification efficiencies were analyzed with LinRegPCR free software (v. 2014.5). Relative expressions were calculated using the threshold cycle ($2^{-\Delta\Delta CT}$) method.[35]

Alignments and Phylogenetic Analyses

Sequences encoding stylin proteins were retrieved from the seven genomes available on AphidBase (http://bipaa.genouest.org/is/aphidbase/)—*Myzus persicae* (Mp; clone G006), *Myzus cerasi* (Mce), *Rhopalosiphum padi* (Rp), *Diuraphis noxia* (Dn), *A. pisum* (Ap), *Daktulosphaira vitifoliae* (Dv), *Aphis glycines* (Ag; Biotype 4)—using CutprotFam-Pred with standard settings.[29,36−39] Among these sequences, 17 were manually curated for the present study.

After removal of the predicted signal peptides using the SignalP-5.0 server, CPR_RR-1 mature protein sequences were used in phylogenetic analyses.[40] For CPAP3 proteins, full sequences were used in phylogenetic analyses. Sequences were aligned using MUSCLE (v3.8.31) configured for highest accuracy (MUSCLE with default settings).[41] Ambiguous regions (i.e., containing gaps and/or poorly aligned) were removed with Gblocks (v0.91b) using the following parameters: minimum length of a block after gap cleaning: 10, no gap positions were allowed in the final alignment and all segments with contiguous nonconserved positions bigger than eight were rejected, minimum number of sequences for a flank position: 85%. Phylogenetic trees were reconstructed using the maximum likelihood method implemented in the PhyML program (v3.1/3.0 aLRT). The WAG matrix, which works as empirical amino-acid substitution model to simulate the biological sequence evolution with flexibility, was selected assuming an estimated proportion of invariant sites (of 0.009) and four gamma-distributed rate categories to account for rate heterogeneity across sites. The gamma shape parameter was estimated directly from the data (gamma = 3.517). Reliability for internal branch was assessed using the aLRT test (SH-Like). Graphical representation and edition of the phylogenetic trees and alignments were performed with TREEDYN (v. 198.3) and the T-Coffee software, respectively.[42−44]

RESULTS

Comparative Analyses of the Proteome of Four Cuticular Structures of the Pea Aphid A. Pisum

To identify stylet-specific CPs, we first characterized and compared the proteomes of four cuticular structures: two muscular articulated segmented tissues, antennae and legs; wings which are outgrowths of the exoskeleton; and stylet bundles comprising two mandibular stylets harboring dendrites and two maxillary stylets that are cell-free cuticular structures.

6,574 peptides were identified in antennae, 6,025 in legs, and 7,242 in wings, which were specifically assigned to 971, 749, and 910 proteins. In stylets, a smaller number of 1,118 peptides were identified and assigned to 141 proteins (Table 1, Tables S4−S7). Proteins identified in legs, antennae, and wings represent 2.2−2.9% of the whole *A. pisum* proteome, while proteins identified in stylets represent only 0.4%. Of the 1,273 proteins identified in this study, 90 were found in all four cuticular structures (7%) and constitute what we hereafter refer to as the "core proteome" of cuticular structures. 426 proteins were found to be structure-specific (33%), with 162, 59, 165, and 40 proteins solely identified in antennae, legs, wings, and stylets, respectively (Figure 2A, Table S8).

Table 1. Summary of Proteins Identified in *Acyrthosiphon pisum* Cuticular Structures: (A) All Peptides and Proteins Recovered from LC-MS/MS Analyses; (B) Peptides and Proteins Assigned to *A. pisum* Proteome (*A. pisum* ID) Recovered from LC-MS/MS Analysis; (C) Cuticular Proteins (CPs) Assigned to *A. pisum* Recovered from LC-MS/MS Analysis[a]

		Antennae	Legs	Wings	Stylets
A	Total # of peptides identified	6590	6040	7260	1119
	Total # of unique peptides identified	6304	5786	6942	1056
	Total # of proteins identified	987	762	927	142
B	Total # of peptides identified (*A. pisum* ID)	6574	6025	7242	1118
	Total # of unique peptides identified (*A. pisum* ID)	6288	5771	6924	1055
	Total # of proteins (*A. pisum* ID)	971	749	910	141
C	Total # of CPs identified−minimal list[b]	39	36	41	15
	Total # of CPs identified−ungrouped[c]	78	77	77	38
	Ratio CPs[b]/Total # of proteins (*A. pisum* ID) %	4.0	4.8	4.5	10.6

[a]#, number; CPs, Cuticular Proteins; ID, identifier. [b]Smallest number of cuticular proteins that can explain all observed peptides recovered from LC-MS/MS analysis. [c]All cuticular proteins recovered from LC-MS/MS analysis identified with "Unique" peptides and "Shared" peptides.

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

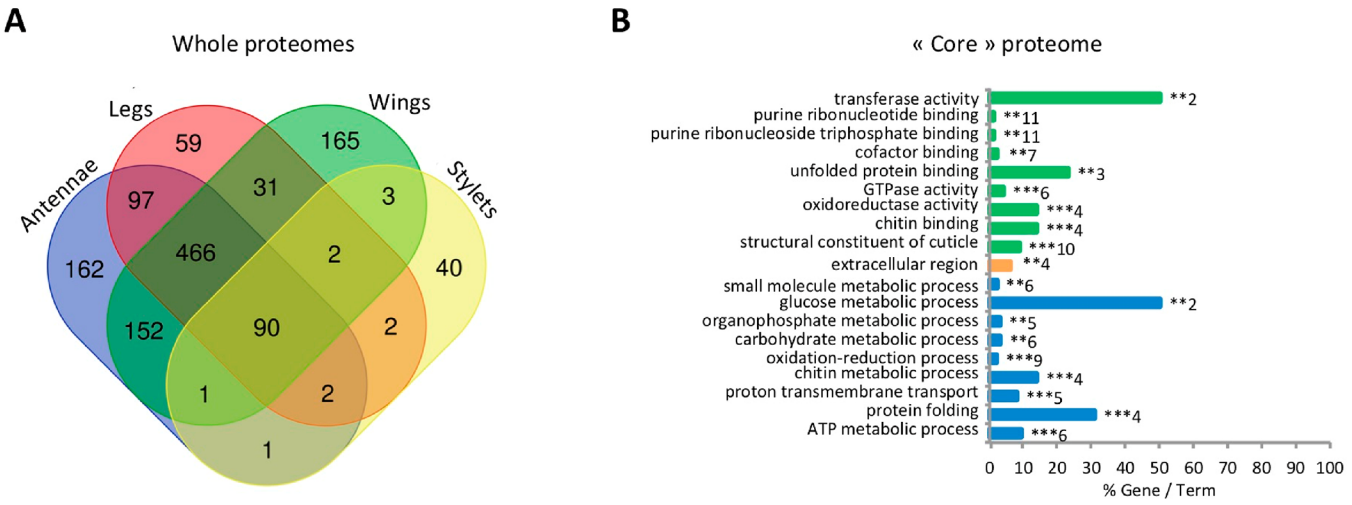


Figure 2. Core proteome of *A. pisum* cuticular structures. (A) Venn diagram showing the proteins identified in the four cuticular structures. 90 proteins found in all structures represent the "Core Proteome". (B) GO terms enriched in the Core Proteome set. Histograms represent significantly enriched functional terms in Molecular Function (green), Cellular Component (orange), and Biological Process (blue). Bars represent the proportion of genes (%) enriched in the corresponding functional groups. Numbers of genes enriched for each functional group are displayed on the bars. *, $p < 0{,}05$; **, $p < 0{,}01$; ***, $p < 0{,}001$. See also Table S9.

As expected for cuticle-containing insect body parts, the GO terms overrepresented in the core proteome were related to the extracellular region and were significantly enriched in cuticle-related functions and the sclerotization process. Hence, categories such as structural constituent of cuticle, chitin binding, carbohydrate metabolic process, chitin metabolic process, oxidoreductase activity, and oxidation−reduction process were overrepresented (Figure 2B, Table S9). Other categories related to unfolded protein binding or small molecules binding, metabolic processes, proton transmembrane transport, transferase or GTPase activities were also identified.

Proteomes of antennae, legs, and wings were enriched in functions related not only to the extracellular region but also to categories associated with intracellular compartments such as ribosome, proteasome, mitochondria, and vesicles (Figure 3A−C, Tables S10−S12). More categories were overrepresented in these three structures compared to aphid stylets (31−41 categories vs 16 categories, respectively; see Figure 3A−D). Notably, the chitin metabolic process category is not specially enriched in stylets (Figure 3A−D), emphasizing the strictly extracellular origin of stylet cuticle, while the other appendages contain the chitin-synthesizing epithelia and its intracellular functions. Apart from cuticle-related functions, antennae, legs, and wings contained proteins with commonly enriched functions associated with protein synthesis, protein folding, intracellular trafficking, and small molecules binding. In comparison, categories enriched in stylets were mostly related to cuticle synthesis or sclerotization pathways (Figure 3D, Table S13). Not surprisingly, our data sets highlighted differences in proteomic composition between nearly exclusively cuticle-based structures (i.e., stylets) and only partly cuticle-based tissues also including cellular machineries (i.e., antennae, legs, and wings) (Figure 3E).

Stylet-Specific Proteins

Forty stylet-specific proteins have been identified in our MS data sets (Figure 2A, Table S8), most of which were not previously annotated. Although not possible to ascertain, at least 13 of these stylet-specific proteins (32.5%) likely came from the saliva remaining in the salivary canal after dissection of stylet bundles. Consistent with this hypothesis, 12 out of these 13 proteins have already been reported in the pea aphid sialome (Table S8), and one is a predicted secreted peroxidase, an enzyme found in aphid saliva.[45−49] Two cuticle-related proteins were only present in the stylets data sets, one trehalase (ACYPI007462) reported to play a role in several processes, including chitin biosynthesis in other insects, and one RR-2 protein (ACYPI006670).[50] Whereas the saliva-related proteins and trehalase were identified with unique peptides not detected in the three other cuticular structures, the RR-2 protein was solely identified from two shared peptides (GSYSLLEADGSTR and TVEYTADDYNGFNAVVK), also identified in antennae, legs, and wings. Because these peptides have been assigned to ACYPI006670 in the stylets, but to other RR-2 proteins in antennae, legs, and wings by the data search program (Tables S14−S17), we cannot strictly conclude that this protein (ACYPI006670) is stylet-specific.

Distribution of the Identified CPs in the Different Investigated Cuticular Anatomical Structures

Minimal lists of CPs recovered by MS analyses comprise 39 CPs in antennae, 38 CPs in legs, 41 CPs in wings, and 15 CPs in stylets (Tables 1 and 2, Tables S14−S17). No peptides from the CPF and CPCFC families were identified in any cuticular structure. The representation of distinct CPs families was similar in antennae, legs, and wings (Figure 4A, Table 3), with proteins of the CPR (RR-1 and RR-2), CPAP1, CPAP3, and TWDL families. In stylets, only two protein families were represented, CPR (RR-1 and RR-2) and CPAP3.

RR-2 proteins were the most abundant in all structures and were often identified with shared peptides (Table 2). RR-1 and TWDL proteins were identified either with shared or with unique peptides. Finally, CPAP1 and CPAP3 were only identified with unique peptides (Table 3). The number of CPs unequivocally identified with unique peptides was 26, 26, 30, and 9 in antennae, legs, wings, stylets, respectively. Comparative analysis of CPs *potentially present* in cuticular structures indicated that only a few proteins were restricted to a single structure: one CPAP3 in antennae (ACYPI001579), one CPAP1 (ACYPI004632), and two RR-2 in legs (ACYPI006712 and ACYPI003698), one CPAP1

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

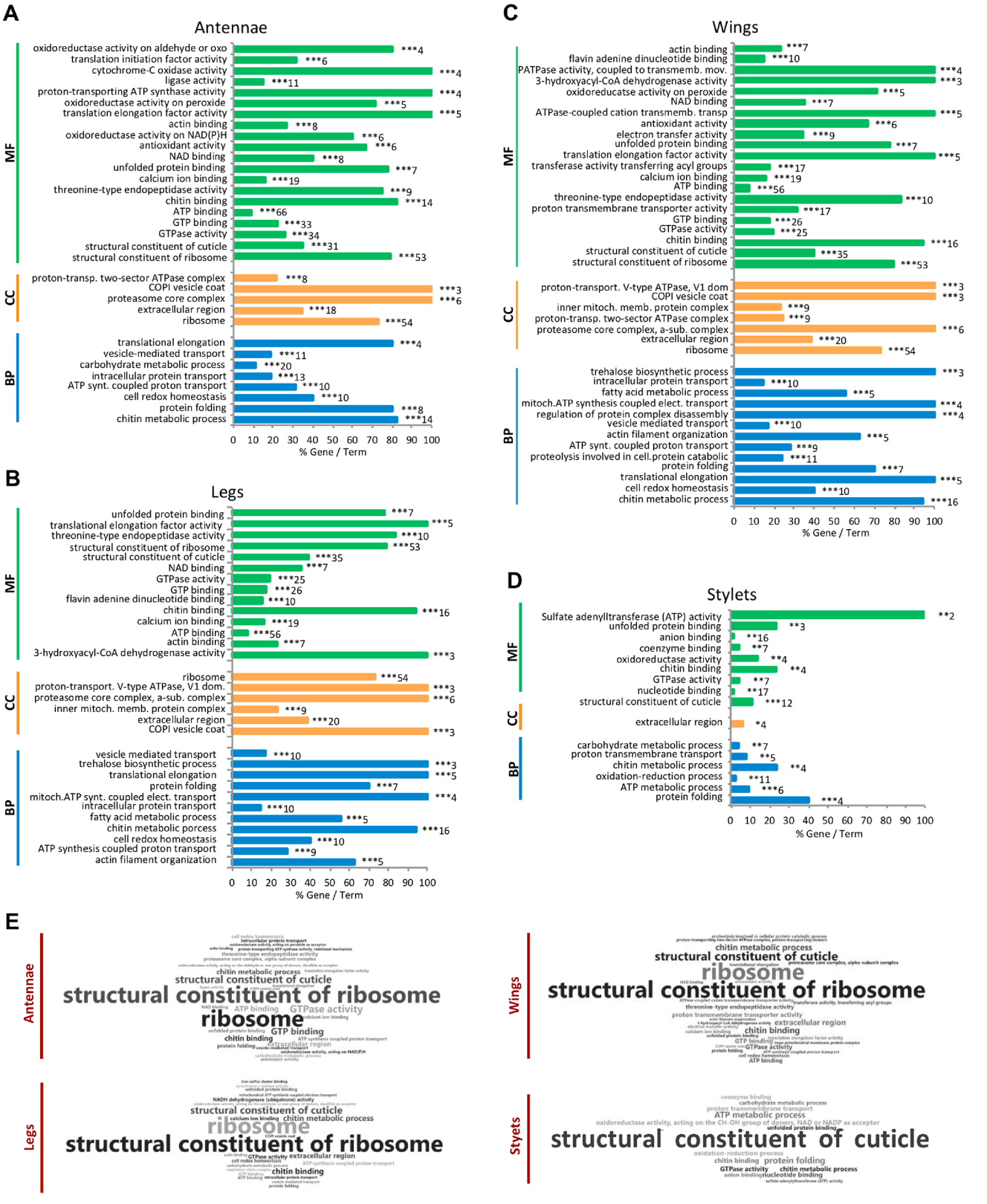


Figure 3. Comparative analysis of GO terms enriched by identified proteins in antennae, legs, wings, and stylets of *A. pisum*. (A–D) Comparison of identified proteins in the four cuticular structures for Molecular Function (MF, green), Cellular Component (CC, orange), and Biological Process (BP, blue). Histograms represent significantly enriched functional terms in antennae (A), legs (B), wings (C), and stylets (D). Bars represent the proportion of enriched genes encoding the identified proteins compared to all genes involved in the specified GO terms (%Gene/Term). Numbers of genes enriched for each functional group are displayed on the bars. *, $p < 0{,}05$; **, $p < 0{,}01$; ***, $p < 0{,}001$. (E) WordClouds summarize the functional profiles of each cuticular anatomical structure. See also Tables S10–S13.

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX–XXX

Table 2. Exhaustive List of Cuticular Proteins Identified in the Four Cuticular Structures[a]

Protein ID	Antennae	Legs	Wings	Stylets	Stylin ID
ACYPI000308-PA	S	S		S	Stylin-04bis
ACYPI000461-PA	S	U	S		
ACYPI000583-PA			U		
ACYPI000670-PA			U		
ACYPI000849-PA	U	U	U		
ACYPI000889-PA	U	U	U		
ACYPI000890-PA	U	U	U		
ACYPI000961-PA	S	S			
ACYPI000983-PA	S	S	S	S	
ACYPI001278-PA	U	U	U		
ACYPI001451-PA	S	S	S		
ACYPI001579-PA	U				
ACYPI001599-PA	U	U	U	U	
ACYPI001610-PA	U	U	U	U	Stylin-03
ACYPI001644-PA	S	S	S	S	
ACYPI001681-PA			U		
ACYPI001775-PA			U		
ACYPI002106-PA	S	S	S		
ACYPI002243-PA	S	S	S	S	
ACYPI002781-PA	U	U	U		
ACYPI002877-PA	S	S		S	Stylin-04
ACYPI002889-PA	S	S	S	S	
ACYPI003649-PA	U	U	U	U	Stylin-02
ACYPI003698-PA		S			
ACYPI004093-PA	U	U	U		
ACYPI004114-PA	S	S	U	S	
ACYPI004632-PA		U			
ACYPI004810-PA	S	S	S	S	
ACYPI004893-PA	S	S	S	S	
ACYPI004943-PA	S	S	S	S	
ACYPI005387-PA	S	S	S	S	
ACYPI006015-PA			U		
ACYPI006031-PA	U	U	U	U	
ACYPI006045-PA	S	S	S	S	
ACYPI006175-PA	S	S	S		
ACYPI006519-PA	U	S	U		
ACYPI006670-PA	S	S	S	S	
ACYPI006712-PA		U			
ACYPI006791-PA	S	S	S		
ACYPI007858-PA	U	U	U	U	
ACYPI007860-PA	U	U	U	U	
ACYPI007911-PA	U	U	U	U	Stylin-05
ACYPI007927-PA	S	S	S	S	
ACYPI008113-PA	U		U		
ACYPI008534-PA	S	S	S	S	
ACYPI008570-PA	U	U	S	S	
ACYPI008661-PA	S	S	S		
ACYPI009006-PA	U	U	U	U	Stylin-01
ACYPI009152-PA	U	U	U		
ACYPI009260-PA	S	S	S		
ACYPI009356-PA	U	U	U		
ACYPI009491-PA	U	U	U		
ACYPI009786-PA	U	U	U		
ACYPI009803-PA	S	S	S	S	
ACYPI009804-PA	S	S	S	S	
ACYPI009812-PA	S		S		
ACYPI060832-PA	S	S	S	S	
ACYPI060836-PA	S	S	S		
ACYPI060840-PA	S	S	S	S	
ACYPI062590-PA	S	S			
ACYPI064288-PA	U	U	U	S	
ACYPI064297-PA	S	S	S	S	
ACYPI064298-PA	S	S	S		
ACYPI065206-PA	S	S	S	S	
ACYPI065224-PA	S	S	S		
ACYPI066081-PA	S	S			
ACYPI066082-PA	S		S		
ACYPI066095-PA	U	U	U	U	
ACYPI070504-PA	S	S	S		
ACYPI071344-PA	S	S	S	S	
ACYPI071387-PA	S	S	S	S	
ACYPI072004-PA	S	S	S		
ACYPI072111-PA	S	S	S		
ACYPI073031-PA	S	S	S	S	
ACYPI073034-PA	S	S	S		
ACYPI073043-PA	S	S	S		
ACYPI073896-PA	S	S	S		
ACYPI26371-PA	S	S	S	S	
ACYPI35881-PA	S	S	S		
ACYPI45536-PA			U		
ACYPI48369-PA	U	U			
ACYPI56614-PA	S	S	S		
ACYPI56619-PA	S	S	S		
ACYPI56617-PA	S	S	S		
ACYPI56620-PA	S	S	S	S	
ACYPI56622-PA	S	S	S	S	
ACYPI56625-PA	U	U	U		

CPR_RR-1
CPR_RR-2
CPAP1
CPAP3
TWDL

[a]U = unique peptide found in protein. S = peptide shared among different proteins. The 15 CPs identified by mass spectrometry in aphid stylets are indicated in red (see also Table S17). RR-1 proteins are highlighted in green, RR-2 in grey, CPAP1 in yellow, CPAP3 in orange, and Tweedle in blue. ID, identifier.

(ACYPI45536), one CPAP3 (ACYPI000583), one RR-1 (ACYPI001775), three RR-2 in wings (ACYPI000670, ACYPI001681, ACYPI006015), and none in the stylets (Figure 4B, Table 2, Table S18).

Localization of RR-1 Proteins in the Stylets

To confirm the presence of the 38 CPs identified by MS in aphid stylets and to give information on their localization and accessibility, *A. pisum* stylets were dissected and immunolabeled with antibodies targeting peptides from these proteins (Table 4). All CPs located in the acrostyle with peptides directly accessible at its surface will be renamed stylins.

Four to five RR-1 proteins were identified in our MS analyses (Table S17) including Stylin-01 (ACYPI009006) and Stylin-02 (ACYPI003649). These two stylins have previously been detected in *A. pisum* stylets using specific antibodies.[26]

A third RR-1 protein, ACYPI001610, was identified with a unique peptide of 13 AA covering 10.8% of the mature protein (Table 4, Table S17). This peptide could not be detected with an anti-1-03 specific antibody in our experimental conditions despite extensive Chitinase digestion treatment (Figure 5A). Three additional antibodies targeting distinct peptides of ACYPI001610, anti-1-04, anti-1-15, and anti-1-17 were then used to detect this protein in aphid stylets (Table 4). Anti-1-04 and anti-1-17 antibodies revealed their corresponding peptides at the tip of aphid maxillary stylets in the common canal and on the lateral edge of the stylet. Labeling was only observed after preincubation of stylets with Chitinase, indicating that these two peptides were embedded in the chitin matrix (Figure 5A). A stronger labeling was observed when using anti-1-15 antibody, which labeled evenly the acrostyle without the need for Chitinase treatment. Noticeably, the targeted peptide VEGGY-SYTAPDGTPI is part of the RR-1 chitin-binding domain, and its labeling decreased upon Chitinase digestion (Figure 5A). The protein ACYPI001610 is thus accessible at the surface of the acrostyle and was renamed Stylin-03.

The fourth RR-1 protein, ACYPI002877, was identified by MS analyses with two shared peptides also found in ACYPI000308 (Table 4, Table S17). These two proteins are 96.5% identical and cannot be distinguished by immunolabeling. Anti-1-14 antibody, which targets the peptide GSYTFGYQ-SADGTQR, labeled the acrostyle and the lateral edge tip of maxillary stylets of the stylet tip after a Chitinase treatment. The other identified peptide detected in our MS data sets (Table 4, Table S17) and present within the N-terminus of the two proteins was directly accessible at the surface of the upper part of the acrostyle in close vicinity of the food canal and on a lateral edge by anti-1-16 antibody (Figure 5A). At least one of these two RR-1 proteins is thus present in the acrostyle, and they have consequently been renamed Stylin-04 and Stylin-04bis.

Localization of RR-2 Proteins in the Stylets

Eight to 30 RR-2 proteins have been identified in stylets, among which only three were identified by unique peptides. Due to high sequence similarities, even unique peptides were sometimes nearly identical to others present in distinct RR-2 proteins and could not be distinguished by immunological approaches.

Solely two RR-2 proteins were unambiguously detected in the acrostyle. ACYPI066095, identified with four peptides, was detected with anti-2-16 specific antibody. The unique peptide targeted by this antibody was found embedded in the acrostyle, only revealed after Chitinase treatment (Figure 5B, Table 4). The second RR-2 protein, ACYPI007858, was identified by MS with a single peptide of 11 AA covering 2% of the mature protein

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

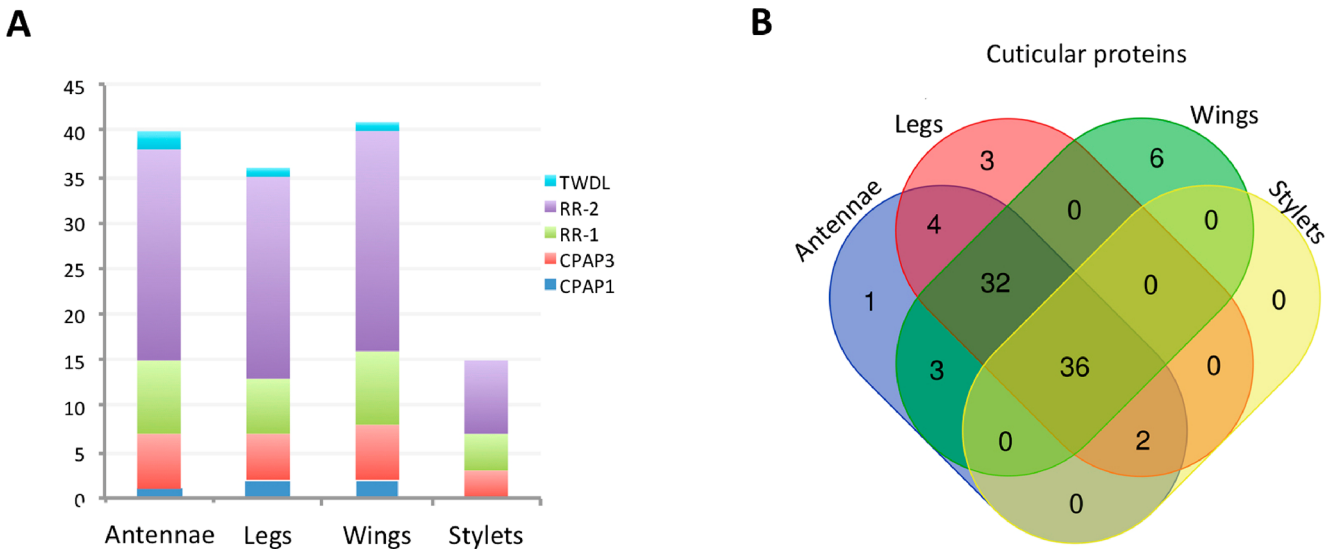


Figure 4. Comparative CP profiling in *A. pisum* cuticular structures. (A) Distribution of the different CPs from minimal lists in antennae, legs, wings, and stylets. (B) Venn diagram indicates the number of all cuticular proteins potentially present in antennae, legs, wings, and stylets.

Table 3. Classification of Cuticular Proteins of the Different Structures[a]

	CPAP1	CPAP3	CPR_RR1[b]	CPR_RR-2[b]	TWDL[b]	Total[b]
Antennae	1	6	(8)-10	(22)-59	2	(39)-78
Legs	2	5	(6)-7	(22)-61	(1)-2	(36)-77
Wings	2	6	(8)-9	(24)-58	(1)-2	(41)-77
Stylets	0	3	(4)-5	(8)-30	0	(15)-38
Whole genome	11	8	15	110	3	147

[a]Number of proteins with peptides identified by MS analyses assigned in the different CP families for each cuticular structure. Total number of CPs identified in the structures is indicated. [b]The smallest numbers of CPs that can explain all observed peptides recovered from MS analysis (minimal list) are indicated in parentheses.

and targeted by anti-2-13 antibody. This peptide was also embedded in the acrostyle, as well as in the lateral edge of the stylets (Figure 5B, Table 4). Attempts to detect another unique peptide of this CP with anti-2-14 antibody failed whatever the conditions used. A third RR-2 protein, ACYPI001599 was detected by MS with a unique peptide, but no specific antibody could be produced.

Four peptides included in the chitin-binding domain and common to several identified RR-2 proteins have earlier been reported as embedded within the acrostyle—pepL, pepS, Ap2-05, and Ap2-08.[22,25] We here extend the listing of RR-2 protein peptides detectable in stylets with four additional antibodies. Anti-Ap2-04, anti-Ap2-06, and anti-Ap2-07 are directed against peptides of the conserved RR-2 chitin-binding domain, whereas anti-2-10 targets one peptide present in the C-terminus of a few proteins. All four antibodies labeled maxillary stylets after extended Chitinase treatments. Strong fluorescent labeling appeared as dots along maxillary stylets when using anti-2-06 and anti-2-10 antibodies, indicating that RR-2 proteins are widely present under the surface (Figure 5B, Table 4).

ACYPI066095 and ACYI007858 formally detected within the acrostyle with antibodies targeting unique peptides lack several shared peptides detected in the organ, among which pepL, pepS, and Ap2-08. Therefore, another RR-2 protein containing these shared peptides may be present in the organ. We can conclude that at least three RR-2 proteins are detected in the acrostyle (Figure 5B).

Localization of CPAP3 Proteins in the Stylets

Three CPAP3 have been identified in the stylet data sets. With the set of antibodies produced and used against these proteins, the labeling was generally weak and visible as dots at the tip of maxillary stylets (Figure 5C, Table 4). ACYPI007860 was only barely revealed in the acrostyle after Chitinase treatment with a single antibody, anti-3-02, targeting a peptide that was not detected by MS. No labeling could be visualized when using anti-3−01 antibody targeting one unique MS-identified peptide of this protein. Its location within aphid stylets could therefore not be definitely stated. ACYPI006031 and ACYPI007911 were detected in the acrostyle, each with two distinct antibodies. Two peptides of ACYPI006031, respectively targeted by anti-3-06 and anti-3-07 were found embedded within the cuticle. For ACYPI007911, the peptide targeted by anti-3-03 was detected only after Chitinase treatment, whereas that targeted by anti-3-09 antibody was detected directly at the surface of the acrostyle (Figure 5C). This protein was renamed Stylin-05.

Repertoire of Cuticular Proteins Present in the Acrostyle

Of the 15 CP groups identified in stylet bundles, nine proteins were unambiguously detected in the acrostyle. These proteins belong to two CP families, including seven CPR proteins (4 RR-1 and 3 RR-2) and two CPAP3 (Figure 6). Four peptides have been shown to be directly accessible at the surface: three of them are found in RR-1 proteins and are either part of the chitin-binding domain (Stylin-03) or present at the N-terminus and C-terminus of the proteins (Stylin-04/-04bis and Stylin-01/-02, respectively), and one belongs to a CPAP3 protein and is part of

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

Table 4. Correspondence between Peptides/Proteins Identified in *A. pisum* Stylets by MS Analysis and Peptides Used for Antibody Production[a]

CP ID	Peptides identified by MS or present in the CP	Peptides used for immunization	Antibody ID	Labeling / Accessibility
ACYPI001599	- VVEYTADNYGFNAEVK	n.d.	N/A	N/A
	- APYSAPAPAYK	n.d.	N/A	N/A
	- GSYSLLEADGSTR	GSYSLLEADGSTRTVE	Anti-pepL[22]	+ / Em
	- *FEYSVNDPSTYDVKS*	FEYSVNDPHTYDVKS	Anti-pepS[22]	+ / PAc
ACYPI001610 (Stylin-03)	- YENDGVNFDGSYK	YENDGVNFDGSYK	Anti-1-03	-
	- *QAQESGSVQPAQNP*	QAQESGSVQPAQNP	Anti-1-04	+ / Em
	- *VEGGYSYTAPDGTPI*	VEGGYSYTAPDGTPI	Anti-1-15	+ / Ac
	- *QNPNESVLNVEG*	QNPNESVLNVEG	Anti-1-17	+ / PAc
ACYPI001644	- EGTPSYSSAPAYKPAYK	KEGTPSYSSAP	Anti-Ap2-10[25]	+ / Em
	- IVEYTADDYNGFVAEVK	TRTVEYTADDYNG	Anti-Ap2-06[25]	+ / Em
		EYTADDYNGFNAV	Anti-Ap2-07[25]	+ / Em
	- GSYSLVEPDGTKR	ND	N/A	N/A
	- *KSQSEYSDGNGYVKG*	KSQSEYSDGNGYVKG	Anti-Ap2-04[25]	+ / Em
ACYPI002877 (Stylin-04)	- FYNNGQNPAYNGQFVPILQQSF DLSPEGSYTFGYQSADGTQR	GSYTFGYQSADGTQR	Anti-1-14	+ / PAc
	- YPAVQGYPPVMAVQGSYSAITPEGIP IEVSYVADENGYQPTGPGV HPAIQR	n.d.	N/A	N/A
	- *FNGFRPNGAYPQQYI*	FNGFRPNGAYPQQYI	Anti-1-16	+ / Ac
ACYPI003649 (Stylin-02)	- AVILSQEQEVNFDGNFK	SQEQEVNFDGNFKNK	Anti-1-07[26]	+ / PAc
	- *KLLASLPSTPEPQYQ*	RYLASLPSTPEPKYQ	Anti-1-11[26]	+ / Ac
ACYPI004943	- APYSAPAPAYK	ND	N/A	N/A
	- *KSQSEYSDGNGNVKG*	KSQSEYSDGNGYVKG	Anti-Ap2-04[26]	+ / Em
	- *EYTADDYNGFNAE*	EYTADDYNGFNAV	Anti-Ap2-07[26]	+ / Em
	- *FEYSVNDPSTYDVKS*	FEYSVNDPHTYDVKS	Anti-pepS[22]	+ / PAc
ACYPI006031	- DALTDGFTCPDGDVVGPNGR	DGDVVGPNGRILPHP	Anti-3-06	+ / Em
	- *GQRTVQQEPKPSKG*	GQRTVQQEPKPSKG	Anti-3-07	+ / Em
ACYPI006670	- TVEYTADDYNGFNAVVK	TRTVEYTADDYNG	Anti-Ap2-06[25]	+ / Em
		EYTADDYNGFNAV	Anti-Ap2-07[25]	+ / Em
	- GSYSLLEADGSTR	GSYSLLEADGSTRTVE	Anti-pepL[22]	+ / Em
	- *FEYSVNDPHTYDVKS*	FEYSVNDPHTYDVKS	Anti-pepS[22]	+ / PAc
ACYPI007858	- LPLQGATAPAR	LPLQGATAPARQQPL	Anti-2-13	+ / Em
	- *NAGRPKPKKNRPTTT*	NAGRPKPKKNRPTTT	Anti-2-14	-
ACYPI007860	- ENCDYLHNVDCGAR	ENCDYLHNVDCGAR	Anti-3-01	-
	- SQLEPAIGGPHCPR	n.d.	N/A	N/A
	- YYICMEGTAR	n.d.	N/A	N/A
	- *RYQCSPGLAYDRE*	RYQCSPGLAYDRE	Anti-3-02	+ / Em
ACYPI007911 (Stylin-05)	- SELQNPQPSYLCPR	GERSELQNPQPSYL	Anti-3-03	+ / Em
	- EQGCSTGEVFNEESQK	n.d.	N/A	N/A
	- GQNVAHPVFAHPDDCQK	CPKDKAFNSRGQNVAH	Anti-3-09	+ / Ac (dots)
	- *CPKDKAFNSRGQNVAH*	CPKDKAFNSRGQNVAH	Anti-3-09	+ / Ac (dots)
ACYPI009006 (Stylin-01)	- AAILVQDSAPNADGSFK	ILVQDSAPSADGSLK	Anti-1-01[26]	+ / PAc
	- *YIADENGYQPYGAHL*	YVADENGYQPYGAHL	Anti-1-10[26]	+ / Em
	- *RYLASLPSTPEPKYQ*	RYLASLPSTPEPKYQ	Anti-1-11[26]	+ / Ac
	- *LPTPPPIPAEIQESLRY*	TPPPIPAEIQESLRY & LPTPPPIPAEIQESL	Anti-1-13[26]	+ / Em
ACYPI066095	- APAYAAPAYSAPAYK	n.d.	N/A	N/A
	- EGTPSYSSAPAYKPAYK	KEGTPSYSSAP	Anti-Ap2-10[25]	+ / Em
	- APYSAPSYSAPAYK	n.d.	N/A	N/A
	- GSYSLVEADGTKR	n.d.	N/A	N/A
	- *NFDYSVHDDSTYD*	NFDYSVHDDSTYD	Anti-2-16	+ / Em
	- *KSQSEYSDGNGYVKG*	KSQSEYSDGNGYVKG	Anti-Ap2-04	+ / Em
ACYPI071387	- SQNEYADANGYVK	n.d.	N/A	N/A
	- *KRTVEYTADDYNG*	TRTVEYTADDYNG	Anti-Ap2-06[25]	+ / Em
	- *EYTADDYNGFNAE*	EYTADDYNGFNAV	Anti-Ap2-07[25]	+ / Em
ACYPI26371	- GSYSLLEADGSTR	GSYSLLEADGSTRTVE	Anti-pepL[22]	+ / Em
	- *EYTADDYNGFNAE*	EYTADDYNGFNAV	Anti-Ap2-07[25]	+ / Em
	- *FEYSVNDPHTYDVKS*	FEYSVNDPHTYDVKS	Anti-pepS[22]	+ / PAc

[a]AA, amino acid; CP, Cuticular protein; ID, identifier; n.d., not done; N/A, not applicable; (+): labeling detected; (−): no labeling observed under our experimental conditions; Ac, epitope exposed at the surface and directly accessible; PAc, epitope poorly accessible (labeling visible as dots without Chitinase treatment); Em, epitope embedded, not accessible at the surface of the stylets. CP identified with unique peptides are in red; peptides not identified by LC-MS/MS analysis for which antibodies were already available in our laboratory are in italics. Differences in the AA sequence are indicated in blue. Antibody ID is followed by a reference number when the corresponding antibody was described in a previous study.

one of its chitin-binding domain (Figure 7). Noticeably, no RR-2 protein (the most numerous CP class) was detected at the untreated surface of the acrostyle.

Spatial and Temporal Expression of Stylins in *A. pisum*

None of the five proteins detected at the surface of the acrostyle was found to be stylet-specific. They have been identified in all four cuticular structures characterized in this study, except for Stylin-04/-04bis which was/were absent from the wing's proteome (Table 4). To characterize possible difference in *stylin* genes expression in different tissues, we compared their transcript levels in antennaless-heads containing the stylet-synthesizing glands, and in antennae, legs, and wings by real-time RT-qPCR analyses using stylin-specific primers for Stylin-01, -02, -03, and -05 (Table S3). For Stylin-04/-04bis, we could

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

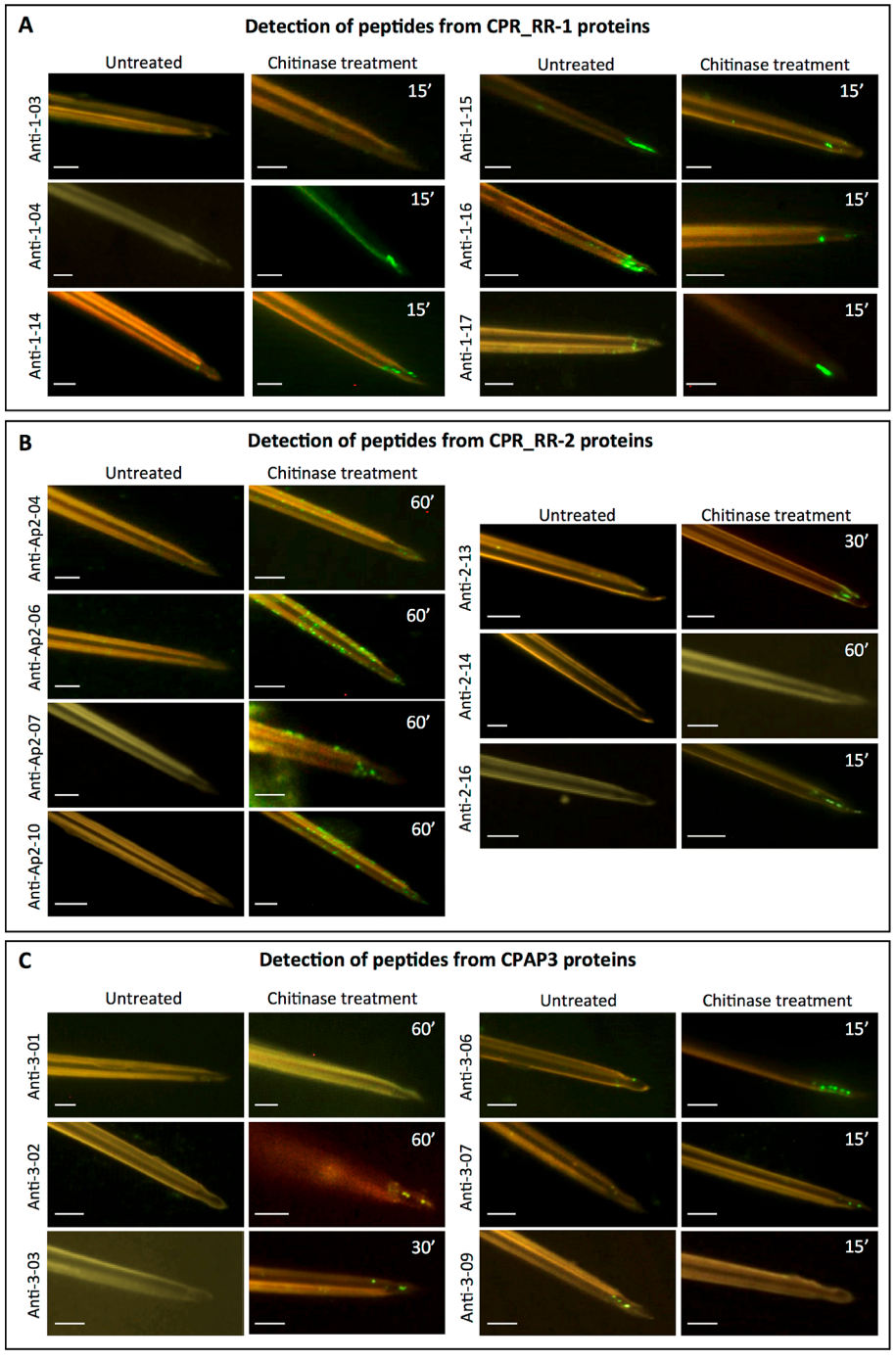


Figure 5. Detection of cuticular proteins in *A. pisum* stylets by immunolabeling. Immunolabeling of *A. pisum* maxillary stylets with antibodies targeting peptides of cuticular proteins identified by the LC-MS/MS analyses. Representative images of labeling observed for antibodies targeting peptides from CPR-RR-1 proteins (A), CPR-RR2 proteins (B), and CPAP3 proteins (C) are shown for untreated stylets and stylets treated with Chitinase prior to immunolabeling. Incubation times with Chitinase ranging from 15 to 60 min are indicated on the top right of images. Scale bars of 5 μm are included.

not design primers able to specifically amplify a single transcript. The observed patterns are thus difficult to interpret and provided for information purposes only.

Consistent with the proteomic data, *stylin-01*, *-02*, *-03*, and *-05* were expressed not only in the head but also in other body parts. They exhibited different expression patterns, with *stylin-01* and *stylin-02* transcripts being more expressed in the head, whereas *stylin-03* was less expressed and *stylin-05* highly variably but equally expressed in all four tissues (Figure 8A).

Stylin expression profiles were also analyzed during *A. pisum* development at five different time-points: during the four larval stages and in adults. *rr-1 stylin* genes showed similar expression patterns with an increase in successive larval stages peaking at the fourth instar (8 to 20 times higher than the expression level in the first instar) and then decreasing in adults. Noticeably, expression levels were highly variable in the fourth instar, the developmental stage of longest duration in *A. pisum*.[51] Expression levels of *stylin-05* gene were comparable in the four larval stages and were significantly lower in adults (Figure 8B).

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

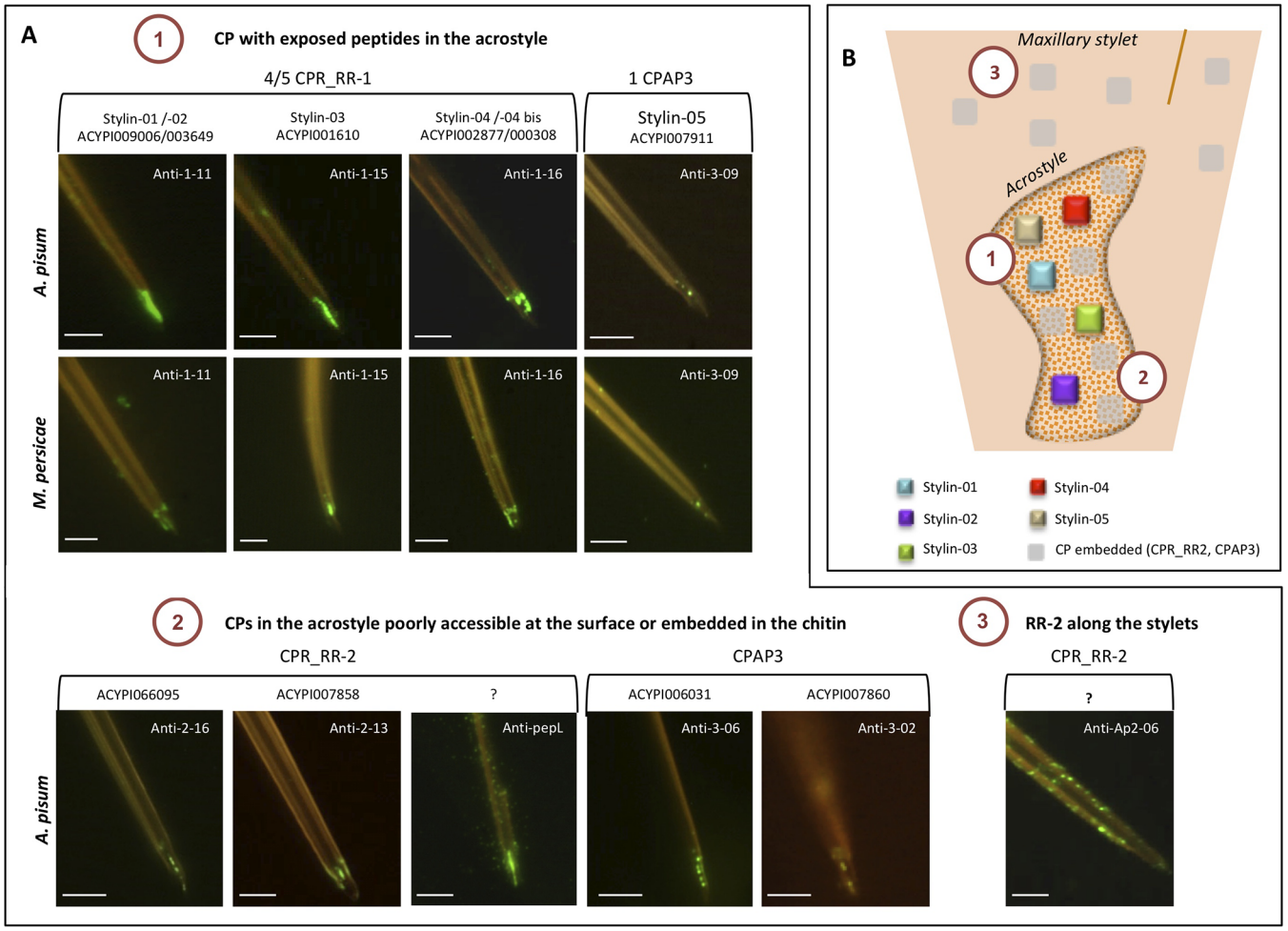


Figure 6. Repertoire of cuticular proteins in the acrostyle. (A) Summary of peptides and proteins detected in *A. pisum* maxillary stylets at the surface of the acrostyle (1), embedded in the organ (2), or distributed all along the stylets (3). (B) Schematic representation of the distribution of the CPs identified in maxillary stylets. Scale bars of 5 μm are included.

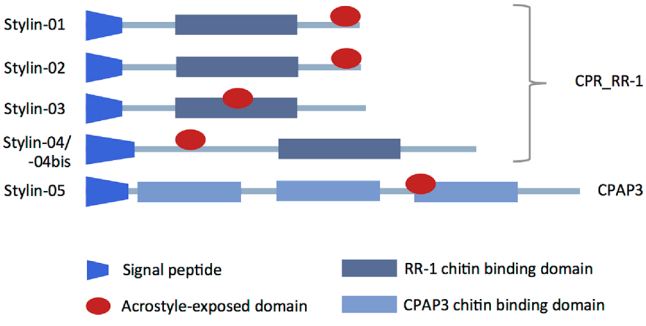


Figure 7. Domain organization of stylins. Schematic representation of stylin domains with predicted signal peptide, and RR-1 or type 2 chitin-binding domains. Acrostyle-surface exposed domains are indicated in red.

Stylins are Conserved among Aphid Species

We constructed two phylogenetic trees for RR-1 and CPAP3 protein sequences identified in the six genomes of aphids and in the phylloxera *Daktulosphaira vitifoliae*, all available in AphidBase (Figure 9A,B). The phylloxera belongs to Phylloxeroidea, a superfamily considered to be the nearest sister taxon of the Aphidoidea, which probably diverged 250 My ago.[52] The homologues of each Stylin grouped in a separated clade, except for homologues of Stylin-01 and Stylin-02 which were already shown to group in a single clade.[26] Noticeably, only one protein of *D. vitifoliae* grouped with Stylin-01 and Stylin-02 homologues, probably reflecting gene duplication in aphids after their divergence with phylloxerids. Regarding Stylin-04 and Stylin-04bis found in *A. pisum* genome, phylogenetic analyses showed that only one of these two proteins was present in other aphid species and in *D. vitifoliae*. A gene duplication of *stylin-04* gene likely occurred only in *A. pisum*, as previously described in other conserved gene families in this species.[36]

Stylins are highly conserved among aphid species. Remarkably, the domains exposed at the surface of the acrostyle are nearly identical in the six aphid species (Figure 9C). The most divergent sequences were *D. vitifoliae* homologues of Stylin-04/-04 bis and Stylin-05. In addition, as previously shown in *A. pisum*, and so suggesting a high conservation of their function, all five stylins were detected at the tip of the maxillary stylets of *M. persicae*, the most important vector of plant viruses (Figure 6A).

DISCUSSION

The aphid stylets are composed of a biomaterial with unique surface properties ensuring binding, retention, and release of plant viruses during their journey from one host to another. This transport of viruses is driven at least by cuticular proteins emerging at the surface of the cuticle, in direct contact with endogenous and exogenous compounds flowing in and out

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

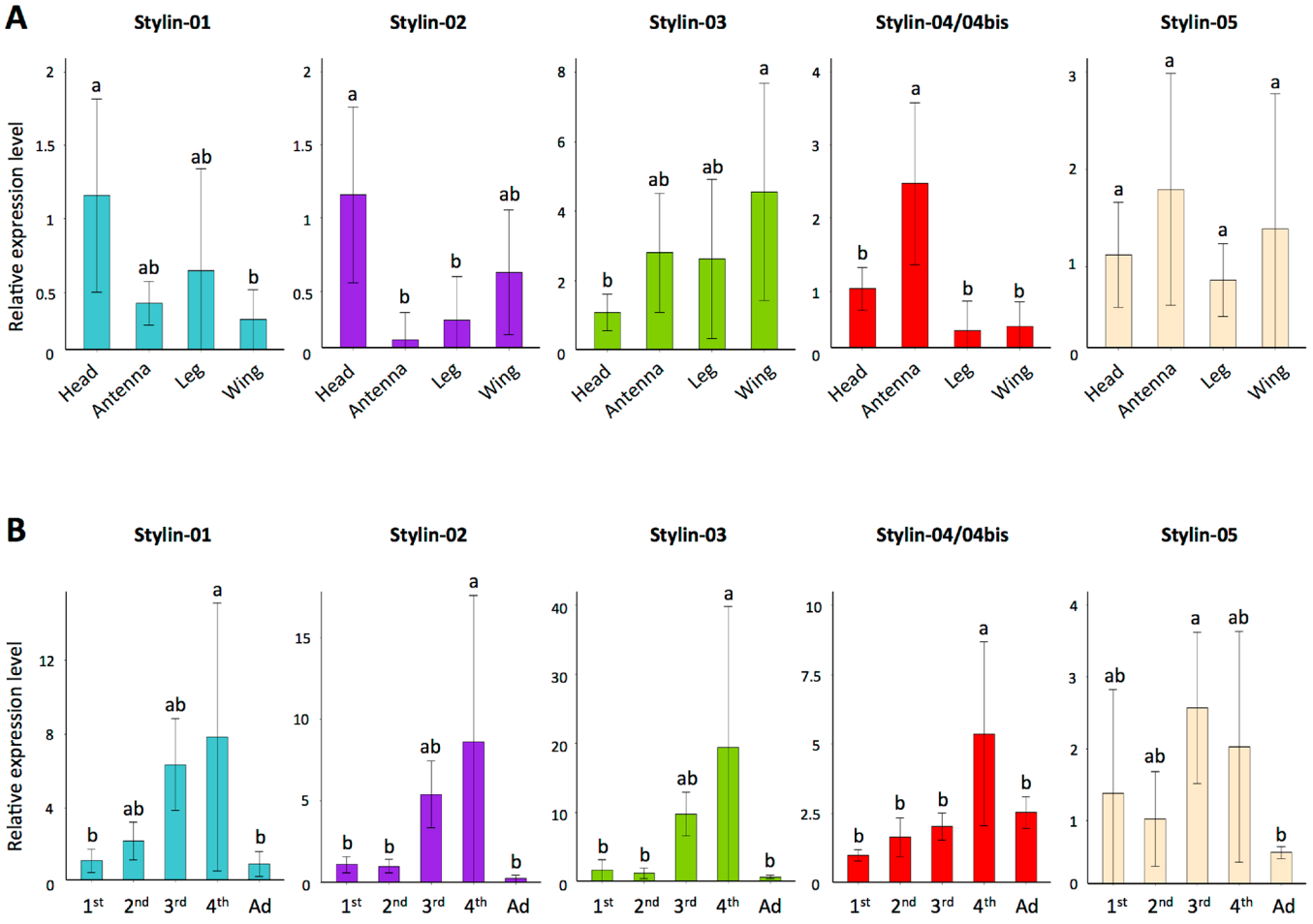


Figure 8. Stylin relative expression patterns in different *A. pisum* body parts and in different nymphal instars. (A) Stylin genes expression quantified by qRT-PCR in head, antenna, leg, and wing relative to stylins expression in Heads. (B) Stylin genes expression quantified by qRT-PCR in first (1st), second (2nd), third (3rd), and fourth (4th) nymphal stages and in adults (Ad) relative to stylins expression in first instars.*actin* and *EF1α* genes were used for data normalization. Results are reported as means ± SD for 3 independent biological replicates. Lowercase letters indicate significant differences between samples (TukeyHSD, $p < 0.005$).

maxillary stylets. Although not characterized when we started our study, we could speculate that various processes might specifically functionalize the cuticular surface at the tip of maxillary stylets in the region described as the acrostyle. These speculated processes could be, for example, the local protein composition of the cuticle or the degree of protein sclerotization. We thus initially presumed that some cuticular proteins would be specific to the stylets, even to the tip of the maxillary stylets, and absent from other anatomical structures. Our comparative proteomics analysis was designed to identify these stylet-specific CPs if there were any, because they represent prime candidate receptors of plant viruses.

Common Set of CPs to Form Aphid Cuticles

We have here experimentally generated the first proteome of four cuticular structures (antennae, legs, wings, and stylets) of the pea aphid. These four proteomes contain multiple CPs of which a shared subset likely represents the building blocks of all aphid cuticles, as suggested for the common subset of CP genes identified in a comparative analysis of seven anatomical structures of the mosquito *Anopheles gambiae*.[33,53] The pea aphid proteome profiles were globally similar in antennae, legs, and wings, with higher complexity/depth than the stylets proteome. However, a few CPs were specifically found in antennae, legs, or wings, and whether or how they contribute to specific cuticular properties/functions in these respective anatomic structures has not been investigated.

Fewer CPs Identified in Stylets

The repertoire of CPs identified in aphid cuticles was half in stylets compared to in other structures, reflecting either the absence of some CPs, their significantly reduced accumulation, or extractability. It is notable that the cuticle of antennae, legs, and wings originates from epithelial cells, while that of the stylets is synthesized by the highly specialized retort glands entirely dedicated to cuticle production.[13,14,54] This sole ontology difference may lead to cuticles with different degrees of protein complexity. However, a recent characterization of transcripts in the retort glands revealed a more complex set of expressed CP genes than in other previously characterized cuticular transcriptomes.[28] Our MS data show that over 50% of the stylet CPs were identified by a single peptide, sometimes covering less than 2% of the mature protein, suggesting that CPs are hardly extractable from this organ and perhaps explaining the low complexity of the corresponding proteome.

CPs Are Not Evenly Distributed in Aphid Stylets

Our labeling experiments showed a variable distribution of CPs within maxillary stylets. Several RR-2 peptides were displayed under the superficial layers of the cuticle all along maxillary

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

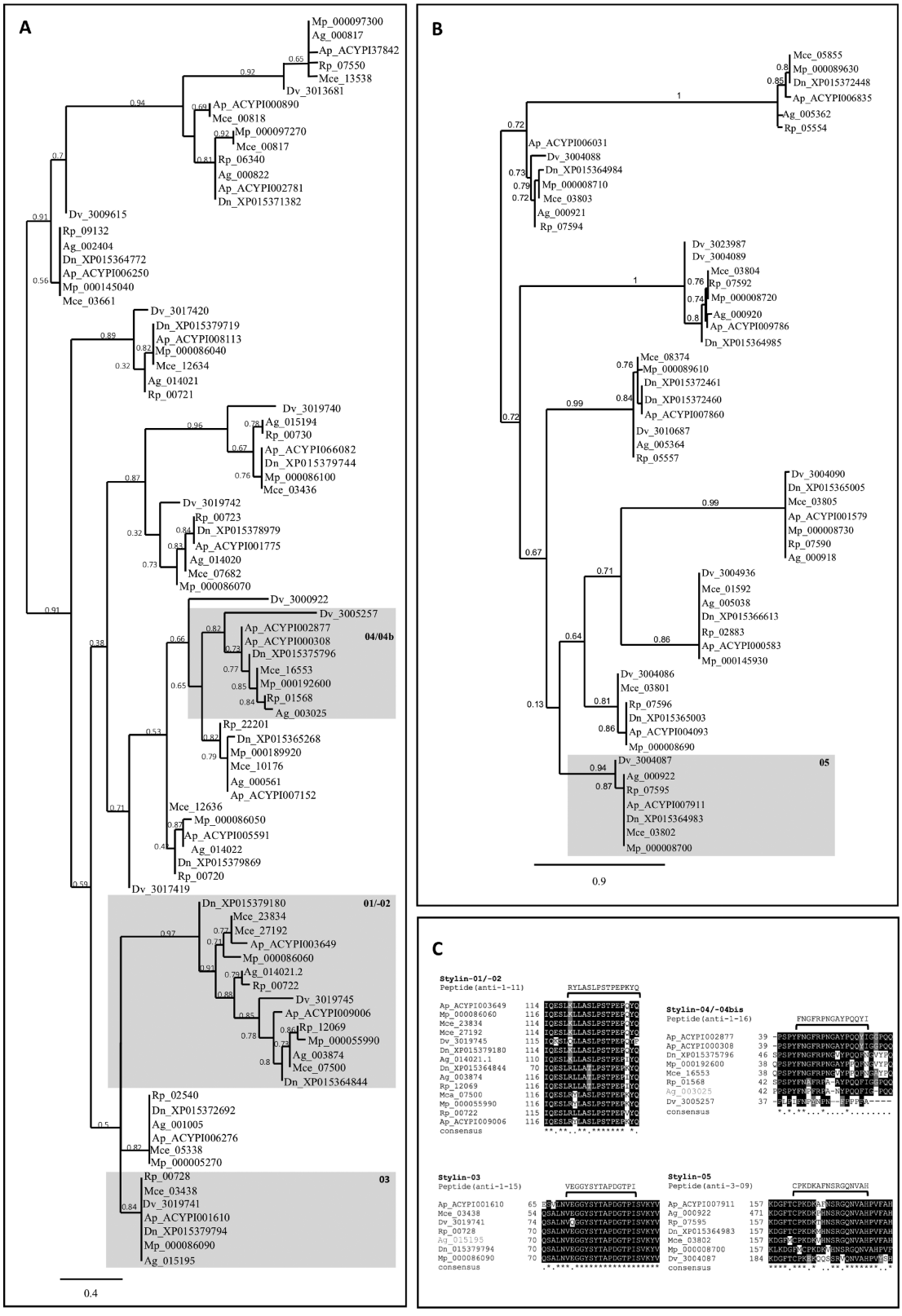


Figure 9. Stylins and acrostyle surface peptides are conserved in aphid vectors. Phylogenetic relationships of (A) CPR_RR-1 proteins with trimmed signal peptides and (B) full CPAP3 proteins of *Acyrthosiphon pisum* (Ap), *Aphis glycines* (Ag), *Daktulosphaira vitifoliae* (Dv), *Diuraphis noxia* (Dn), *Myzus cerasi* (Mce), *Myzus persicae* (Mp), *Rhopalosiphum padi* (Rp). AphidBase accession numbers for each gene are indicated on the right of the species abbreviation. It is noteworthy that to retrieve *M. persicae* sequences from AphidBase, numbers should be preceded by MYZPE13164_G006_v1.0_. Branch support values are indicated at the node, and the scale bar represents probabilities of change from one amino acid to another in terms of a unit, which is an expected 1% change between two amino acid sequences. Shaded portions represent stylin clusters and are numbered from 01 to 05 according to corresponding stylin names. (C) Conservation of exposed domains of the five stylins identified in *M. persicae* (Mp), *A. pisum* (Ap), *M. cerasi* (Mce), *R. padi* (Rp), *A. glycines* (Ag), *D. noxia* (Dn), and *Daktulosphaira vitifolia* (Dv). Alignments were performed

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

Figure 9. continued

using T-Coffee software. Shading was done with BOXSHADE 3.21 software. Identical residues are shaded in black; similar residues are shaded in gray. The consensus sequence is shown at the bottom, with periods indicating conserved substitutions and asterisks indicating identities.

stylets. Some of them were also detected under the surface of the acrostyle. With the antibodies used here and in previous studies, we could not evidence their direct accessibility at the surface of the cuticle.[22,25] However, we cannot strictly rule out such a possibility. Actually, repeat-rich sequences containing alanine, proline, tyrosine, and serine residues commonly found in aphid RR-2 proteins and present in several MS-detected peptides are poorly immunogenic and could not be considered for peptide synthesis and antibody production.[25] Therefore, their accessibility in stylet cuticle could not be assessed. By contrast, all RR-1 and CPAP3 have been exclusively detected at the distal extremity of maxillary stylets (Figure 6), with the strongest labeling being observed for RR-1, that likely decorates the entire surface of the acrostyle. Despite their apparent specific localization, we must consider that RR-1 and CPAP3 peptides may also be evenly distributed all over the stylets because these protein families do not seem to be overrepresented in the core stylet material when compared to other aphid cuticular anatomical structures (Figure 4, Table 3), and so are likely generally present in any cuticle. In this hypothesis, what would appear specific would be their accessibility at the surface of the acrostyle, and this could reveal a general property of the cuticle that would be locally functionalized by surface modifications allowing access to normally embedded CPs.

Intriguingly, no labeling was ever observed on mandibular stylets in our experimental conditions. This striking observation suggests that proteins in this structure may be tightly cross-linked and no longer detectable with our antibodies or, as hypothesized above, that these stylets have surface properties that do not give access to the embedded cuticular proteins. Both hypotheses assume a comparable protein constitution for maxillary and mandibular stylets, consistent with their undistinguishable transcript profiles reported earlier.[28]

Receptor Candidates

As expected for structural components of arthropod cuticle, stylins are highly conserved in all aphid species where sequence data are available. Receptors of plant viruses may also be conserved in vectors because the same viral species can sometimes be transmitted by dozens of different aphid species.[55,56] Our results indicate that the protein domains accessible at the surface of the acrostyle share high degrees of identity with their homologues in other aphid species (Figure 9C). Interestingly, the most divergent protein sequences were those of *D. vitifoliae*, the grapevine pest species, for which transmission of plant viruses or of other pathogens has not been reported so far.[57,58]

RR-1 proteins are largely represented at the surface of the acrostyle. Proteins from RR-1 subgroup are generally found in soft cuticle, but a few of them have also been described in hard cuticle.[24,59−61] Their role in virus transmission has recently emerged in the literature, and they are associated with both circulative and noncirculative virus transmission.[26,62,63] CPR1, a RR-1 protein of the small brown planthopper *Laodelphax striatellus*, was first shown to interact with pc3, a viral nucleocapsid protein of rice stripe virus (RSV) (62). RSV is a tenuivirus transmitted in a circulative propagative manner.[64] By contrast with noncirculative viruses, these viruses are not retained on their vector mouthparts. They are internalized in insect body, transit through the hemolymph to different internal tissues, and reach the salivary glands from which they can be inoculated together with egestion of saliva into a new host plant.[4] Knockdown of CPR1 transcripts resulted in a decrease in RSV transmission. The authors proposed that CPR1 could assist viral movement within the insect body, a totally unexpected function for cuticular protein.[62] For noncirculative viruses, *in vitro* competition experiments and transmission phenotypes associated with *stylin* genes knockdown demonstrated the role of Stylin-01 in CaMV transmission by its aphid vector. Decreasing Stylin-01 transcripts in *M. persicae* resulted in a reduced CaMV transmission capacity.[26] In an independent study, Stylin-01, named previously Mpcp4 in *M. persicae*, was shown to interact in yeast with the coat protein of *Cucumber mosaic virus* (CMV), another aphid-transmitted noncirculative virus.[65] However, its role in CMV transmission still lacks direct evidence.[66] In addition to Stylin-01, Stylin-02, -03, and -04/-04bis, distributed over the surface of the acrostyle, now stand as prime candidate receptors of plant viruses.

CPAP3 are Cuticular Proteins Analogous to Peritrophins, previously known as *"gasp"* or *"obstructor"* family.[67,68] They have been described in all insect, and possess three type 2 chitin-binding domains (ChtBD2).[69] They have been detected in different tissues. However, they were missing from the list of CPs identified in the proboscis of *An. gambiae*.[53] They have never been reported associated with virus transmission so far and are described here for the first time in aphid stylets. They are involved in cuticle formation and in structural integrity of cuticles.[69−71] Antibodies produced in our study target less than 7% of their amino acid sequences and further effort will be required to determine if some CPAP3 domains could be better exposed at the surface of aphid stylets. Determining if they play a role in virus transmission would also warrant further investigation. As they belong to a distinct structural class of chitin-binding proteins relative to CPR (harboring cysteine-bridged Chitin-Binding Domain vs cysteine-free Chitin-Binding Domains for CPR), their functional properties might be distinct from the canonical CPR proteins identified so far as *stylins* and as active players in plant virus transmission.

The RR-2 proteins represent the great majority of CPs identified in *A. pisum* stylets. These proteins are generally found in hard and rigid cuticles, and some of them have been shown to be essential for cuticle integrity, integument structure, insect development and could be involved in cuticle formation.[24,72−77] RR-2 distribution within aphid stylets suggests that they likely constitute the main building blocks of this type of cuticle. Evidence of their accessibility at the surface of aphid stylets is still lacking.[22,25] So, although not prime candidates, the role of RR-2 proteins in virus transmission cannot be discarded. They have been cited in several studies as interacting molecules of both noncirculative and circulative viruses.[78] For noncirculative viruses, three RR-2 of *M. persicae* were shown to interact *in vitro* with HC-Pro, the ligand protein of a potyvirus *Zucchini yellow mosaic virus*.[65] One of these proteins is homologous to ACYPI006670 that contains four peptides detected in maxillary stylets (Table 4). Another one has been characterized recently which reduction of transcripts correlated with a decrease in

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

Potyvirus Y transmission.[79] However, its putative ortholog in *A. pisum* was not in our MS data set, and its presence within *M. persicae* stylets, embedded in the chitin or at the surface of the cuticle, remains to be confirmed.

CONCLUSIONS

We provide through this study the first comparative proteomics analysis of four aphid cuticular anatomical structures, namely antennae, legs, wings, and stylets. Our data gives preliminary evidence that a great number of CP proteins are common to antennae, legs, and wings, while a few CP proteins seem specific to each appendage. The stylet, which was of main interest to better understand the vector/virus interaction was found to be distinct in composition compared to the other three studied appendages. We determined the repertoire of CPs of aphid stylets and precisely mapped their accessibility in maxillae at the surface of the acrostyle. Further characterization for this short list of proteins showed that they are highly conserved in aphid species and thus all represent good candidate receptors of plant viruses. These data contribute to a better characterization of aphid mouthparts, a crucial insect feeding appendage, but also point out at the surface specificities of the cuticle and at the distribution of cuticular protein accessibility, which may be relevant for local functionalization of this tissue. Beyond feeding appendage and virus receptor candidates, our proteomic data sets may contribute to future investigations of other important physiological functions in aphids such as chemoreception and sensory system.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.9b00851.

Table S1. Database generated for protein identification including 252,212 entries. Table S2. Home-made *A. pisum* CP database identified using CutProtFamPred. Table S3. List of oligonucleotides used in this study. Table S4. *A. pisum* proteins identified in antennae. Table S5. *A. pisum* proteins identified in legs. Table S6. *A. pisum* proteins identified in wings. Table S7. *A. pisum* proteins identified in stylets. Table S8. Proteins distribution in *A. pisum* cuticular structures. Table S9. Analysis of GO terms enrichment for proteins identified in the Core Proteome of antennae, legs, wings, and stylets. Table S10. GO terms enrichment in antennae. Table S11. GO terms enrichment in legs. Table S12. GO terms enrichment in wings. Table S13. GO terms enrichment in stylets. Table S14. Cuticular proteins identified in antennae. Table S15. Cuticular proteins identified in legs. Table S16. Cuticular proteins identified in wings. Table S17. Cuticular proteins identified in stylets. Table S18. Cuticular protein distribution in *A. pisum* cuticular structures. (ZIP)

Accession Codes

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier PXD016517.[80]

AUTHOR INFORMATION

Corresponding Authors

Philippe Bulet − *Plateforme BioPark d'Archamps, 74160 Archamps, France; CR University of Grenoble-Alpes, Institute for Advances Biosciences, Inserm U1209, CNRS UMR 5309, 38058 Grenoble, France*; Phone: +(33) 450 432 521; Email: philippe.bulet@univ-grenoble-alpes.fr

Marilyne Uzest − *BGPI, University of Montpellier, INRAE, 34000 Montpellier, France*; orcid.org/0000-0002-7459-8434; Phone: +(33) 499 624 857; Email: marilyne.uzest@inra.fr

Authors

Maëlle Deshoux − *BGPI, University of Montpellier, INRAE, 34000 Montpellier, France*

Victor Masson − *Plateforme BioPark d'Archamps, 74160 Archamps, France; CR University of Grenoble-Alpes, Institute for Advances Biosciences, Inserm U1209, CNRS UMR 5309, 38058 Grenoble, France*

Karim Arafah − *Plateforme BioPark d'Archamps, 74160 Archamps, France*

Sébastien Voisin − *Plateforme BioPark d'Archamps, 74160 Archamps, France*

Natalia Guschinskaya − *INRAE, INSA Lyon, UMR5240 MAP CNRS-UCBL, 69622 Villeurbanne, France*

Manuella van Munster − *BGPI, University of Montpellier, INRAE, 34000 Montpellier, France*

Bastien Cayrol − *BGPI, University of Montpellier, INRAE, 34000 Montpellier, France*

Craig G. Webster − *BGPI, University of Montpellier, INRAE, 34000 Montpellier, France*

Yvan Rahbé − *BGPI, University of Montpellier, INRAE, 34000 Montpellier, France; INRAE, INSA Lyon, UMR5240 MAP CNRS-UCBL, 69622 Villeurbanne, France; University of Lyon, 69007 Lyon, France*

Stéphane Blanc − *BGPI, University of Montpellier, INRAE, 34000 Montpellier, France*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.9b00851

Author Contributions

‡M.D. and V.M. contributed equally to this study. P.B. and M.U. jointly supervised this work. All authors contributed to the work and the manuscript presented here. All authors have approved the final version of the manuscript.

Funding

This work was supported by grants from the Bill and Melinda Gates Foundation (GCEag, OPP1130147) and the French National Research Agency (ANR15-CE20-0011). Funding for the sequencing of *Myzus persicae* and *Rhopalosiphum padi* was provided by ERC Starting Grant APHIDHOST-310190 awarded to Jorunn Bos at the James Hutton Institute, United Kingdom. Funding for *Daktulosphaira vitifoliae* clone Pcf genomic sequencing was provided by Inra (AIP Bioressources) and BGI Biotech in the frame of i5k initiative. Parts of the transcriptomic resources were obtained within the 1KITE projects (Bernhard Misof, Bonn, Germany).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the Association Plateforme BioPark d'Archamps (France) for its technical facilities through its Research & Development program and Mr. Hervé Breton from the society Chromoptic for providing us with the glass microtubes.

ABBREVIATIONS

AA, amino acid(s); ACN, acetonitrile; CaMV, cauliflower mosaic virus; ChtBD2, type 2 chitin-binding domains; CMV, cucumber mosaic virus; CP, cuticular proteins; CPR, cuticular proteins with the Rebers and Riddiford consensus sequence; CPAP3, cuticular protein analogous to peritrophins 3; FDR, false discovery rate; GO, gene ontology; h, hour; HFIP, hexafluoroisopropanol; HMMs, Hidden Markov Models; HPLC, high-performance liquid chromatography; ID, identifier; ITO, Indium−Tin Oxide; LC-MS/MS, Liquid chromatography coupled to tandem mass spectrometry; min, minutes; MS, mass spectrometry; RR, Rebers and Riddiford consensus sequence; RSV, rice stripe virus; TFA, trifluoroacetic acid; v/v, volume/volume; U, Unit

REFERENCES

(1) Dixon, A. F. G. K.; Leps, P. J.; Holman, J. Why there are so few species of aphids, especially in the tropics. *Am. Nat.* **1987**, *129* (4), 580−592.

(2) Favret, C. 2017, Aphid species file. http://aphid.speciesfile.org.

(3) Ogawa, K.; Miura, T. Aphid polyphenisms: trans-generational developmental regulation through viviparity. *Front. Physiol.* **2014**, *5*, 1.

(4) Hogenhout, S. A.; Ammar; el, D.; Whitfield, A. E.; Redinbaugh, M. G. Insect vector interactions with persistently transmitted viruses. *Annu. Rev. Phytopathol.* **2008**, *46*, 327−359.

(5) Katis, N. I.; Tsitsipis, J. A.; Stevens, M.; Powell, G. Transmission of Plant Viruses. In *Aphids as Crops pests*; van Emden, H. F., Harrington, R., Eds.; CABI, Wallingford, United Kingdom, 2007; pp 353−390.

(6) Nault, L. R. Arthropod transmission of plant viruses: a new synthesis. *Ann. Entomol. Soc. Am.* **1997**, *90*, 521−541.

(7) Whitfield, A. E.; Falk, B. W.; Rotenberg, D. Insect vector-mediated transmission of plant viruses. *Virology* **2015**, *479−480*, 278−289.

(8) Martin, B.; Collar, J. L.; Tjallingii, W. F.; Fereres, A. Intracellular ingestion and salivation by aphids may cause the acquisition and inoculation of non-persistently transmitted plant viruses. *J. Gen. Virol.* **1997**, *78* (10), 2701−2705.

(9) Perry, K. L.; Zhang, L.; Palukaitis, P. Amino acid changes in the coat protein of cucumber mosaic virus differentially affect transmission by the aphids Myzus persicae and Aphis gossypii. *Virology* **1998**, *242* (1), 204−210.

(10) Moreno, A.; Hebrard, E.; Uzest, M.; Blanc, S.; Fereres, A. A single amino acid position in the helper component of cauliflower mosaic virus can change the spectrum of transmitting vector species. *J. Virol.* **2005**, *79* (21), 13587−13593.

(11) Chen, B.; Francki, R. I. B. Cucumovirus transmission by the aphid Myzus persicae is determined solely by the viral coat protein. *J. Gen. Virol.* **1990**, *71*, 939−944.

(12) Pirone, T. P.; Blanc, S. Helper-dependent vector transmission of plant viruses. *Annu. Rev. Phytopathol.* **1996**, *34*, 227−247.

(13) Davidson, J. The structure and biology of Schizoneura lanigera Hausmann or Woolly Aphis of the Apple tree. I. The Apterous viviparous female. *Q Jl. microsc. Sci.* **1913**, *58*, 653−702.

(14) Ponsen, M. B. The site of potato leafroll virus multiplication in its vector, Myzus persicae: an anatomical study. *Meded. Landbouwhogesch. Wageningen.* **1972**, *72* (16), 1−147.

(15) van Hoof, H. A. An investigation of the biological transmission of a non-persistent virus. Doctoral thesis. Wageningen Agr. Univ. Meded. Inst. plziektenk, Onderz 161, van Putten, Ortmeijer, Alkmaar, the Netherlands, 1958, 96 pp.

(16) Forbes, A. R. Electron microscope evidence for nerves in the mandibular stylets of the green peach aphid. *Nature* **1966**, *212*, 726.

(17) Forbes, A. R. The stylets of the green peach aphid, Myzus persicae (Homoptera: Aphididae). *Can. Entomol.* **1969**, *101*, 31−41.

(18) Bradley, R. H.; Ganong, R. Y. Evidence that potato virus Y is carried near the tip of the stylets of the aphid vector Myzus persicae (sulz.). *Can. J. Microbiol.* **1955**, *1* (9), 775−782.

(19) Berger, P. H.; Pirone, T. P. The effect of helper-component on the uptake and localization of Potyviruses in *Myzus persicae*. *Virology* **1986**, *153*, 256−261.

(20) Powell, G. Intracellular salivation is the aphid activity associated with inoculation of non-persistently transmitted viruses. *J. Gen. Virol.* **2005**, *86* (Pt 2), 469−472.

(21) Uzest, M.; Gargani, D.; Drucker, M.; Hébrard, E.; Garzo, E.; Candresse, T.; Fereres, A.; Blanc, S. A protein key to plant virus transmission at the tip of the insect vector stylet. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (46), 17959−17964.

(22) Uzest, M.; Gargani, D.; Dombrovsky, A.; Cazevieille, C.; Cot, D.; Blanc, S. The ″acrostyle″: a newly described anatomical structure in aphid stylets. *Arthropod Struct. Dev.* **2010**, *39* (4), 221−229.

(23) Rebers, J. E.; Riddiford, L. M. Structure and expression of a Manduca sexta larval cuticle gene homologous to Drosophila cuticle genes. *J. Mol. Biol.* **1988**, *203* (2), 411−423.

(24) Willis, J. H. Structural cuticular proteins from arthropods: annotation, nomenclature, and sequence characteristics in the genomics era. *Insect Biochem. Mol. Biol.* **2010**, *40* (3), 189−204.

(25) Webster, C. G.; Thillier, M.; Pirolles, E.; Cayrol, B.; Blanc, S.; Uzest, M. Proteomic composition of the acrostyle: Novel approaches to identify cuticular proteins involved in virus-insect interactions. *Insect Sci.* **2017**, *24* (6), 990−1002.

(26) Webster, C. G.; Pichon, E.; van Munster, M.; Monsion, B.; Deshoux, M.; Gargani, D.; Calevro, F.; Jimenez, J.; Moreno, A.; Krenz, B.; Thompson, J. R.; Perry, K. L.; Fereres, A.; Blanc, S.; Uzest, M. Identification of Plant Virus Receptor Candidates in the Stylets of Their Aphid Vectors. *J. Virol.* **2018**, *92* (14), DOI: 10.1128/JVI.00432-18

(27) Masson, V.; Arafah, K.; Voisin, S.; Bulet, P. Comparative Proteomics Studies of Insect Cuticle by Tandem Mass Spectrometry: Application of a Novel Proteomics Approach to the Pea Aphid Cuticular Proteins. *Proteomics* **2018**, *18* (3−4), 1700368.

(28) Guschinskaya, N.; Ressnikoff, D.; Arafah, K.; Voisin, S.; Bulet, P.; Uzest, M.; Rahbé, Y. Insect mouthpart transcriptome unveils extension of cuticular protein repertoire and complex organization. *iScience.* **2020**, *23*, 100828.

(29) Ioannidou, Z. S.; Theodoropoulou, M. C.; Papandreou, N. C.; Willis, J. H.; Hamodrakas, S. J. CutProtFam-Pred: detection and classification of putative structural cuticular proteins from sequence alone, based on profile hidden Markov models. *Insect Biochem. Mol. Biol.* **2014**, *52*, 51−59.

(30) He, N.; Botelho, J. M. C.; McNall, R. J.; Belozerov, V.; Dunn, W. A.; Mize, T.; Orlando, R.; Willis, J. H. Proteomic analysis of cast cuticles from Anopheles gambiae by tandem mass spectrometry. *Insect Biochem. Mol. Biol.* **2007**, *37* (2), 135−146.

(31) Cornman, R. S.; Togawa, T.; Dunn, W. A.; He, N.; Emmons, A. C.; Willis, J. H. Annotation and analysis of a large cuticular protein family with the R&R Consensus in Anopheles gambiae. *BMC Genomics* **2008**, *9*, 22.

(32) Cornman, R. S.; Willis, J. H. Extensive gene amplification and concerted evolution within the CPR family of cuticular proteins in mosquitoes. *Insect Biochem. Mol. Biol.* **2008**, *38* (6), 661−676.

(33) Zhou, Y.; Badgett, M. J.; Bowen, J. H.; Vannini, L.; Orlando, R.; Willis, J. H. Distribution of cuticular proteins in different structures of adult Anopheles gambiae. *Insect Biochem. Mol. Biol.* **2016**, *75*, 45−57.

(34) Conesa, A.; Götz, S. Blast2GO: A Comprehensive Suite for Functional Analysis in Plant Genomics. *Int. J. Plant Genomics* **2008**, *2008*, 619832.

(35) Livak, K. J.; Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* **2001**, *25* (4), 402−408.

(36) Mathers, T. C.; Chen, Y.; Kaithakottil, G.; Legeai, F.; Mugford, S. T.; Baa-Puyoulet, P.; Bretaudeau, A.; Clavijo, B.; Colella, S.; Collin, O.; et al. Rapid transcriptional plasticity of duplicated gene clusters enables a clonally reproducing aphid to colonise diverse plant species. *Genome Biol.* **2017**, *18* (1), 27.

(37) Nicholson, S. J.; Nickerson, M. L.; Dean, M.; Song, Y.; Hoyt, P. R.; Rhee, H.; Kim, C.; Puterka, G. J. The genome of Diuraphis noxia, a global aphid pest of small grains. *BMC Genomics* **2015**, *16* (1), 429.

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

(38) Gallot, A.; Rispe, C.; Leterme, M.; Gauthier, J. P.; Jaubert-Paussamai, S.; Tagu, D. Cuticular proteins and seasonal photoperiodism in aphids. *Insect Biochem. Mol. Biol.* **2010**, *40*, 235−240.
(39) Wenger, J. A.; Cassone, B. J.; Legeai, F.; Johnston, J. S.; Bansal, R.; Yates, A. D.; Coates, B. S.; Pavinato, V. A. C.; Michel, A. Whole genome sequence of the soybean aphid, Aphis glycines. *Insect Biochem. Mol. Biol.* **2017**, *18* (1), 27.
(40) Almagro Armenteros, J. J.; Tsirigos, K. D.; Sønderby, C. K.; Petersen, T. N.; Winther, O.; Brunak, S.; von Hejine, G.; Nielsen, H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **2019**, *37* (4), 420−423.
(41) Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32* (5), 1792−1797.
(42) Chevenet, F.; Brun, C.; Banuls, A. L.; Jacq, B.; Christen, R. TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinf.* **2006**, *7*, 439.
(43) Notredame, C.; Higgins, D. G.; Heringa, J. T-coffee: a novel method for fast and accurate multiple sequence alignment11Edited by J. Thornton. *J. Mol. Biol.* **2000**, *302* (1), 205−217.
(44) Di Tommaso, P.; Moretti, S.; Xenarios, I.; Orobitg, M.; Montanyola, A.; Chang, J. M.; Taly, J. F.; Notredame, C. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res.* **2011**, *39* (suppl_2), W13−W17.
(45) Carolan, J. C.; Fitzroy, C. I. J.; Ashton, P. D.; Douglas, A. E.; Wilkinson, T. L. The secreted salivary proteome of the pea aphid Acyrthosiphon pisum characterised by mass spectrometry. *Proteomics* **2009**, *9* (9), 2457−2467.
(46) Carolan, J. C.; Caragea, D.; Reardon, K. T.; Mutti, N. S.; Dittmer, N.; Pappan, K.; Cui, F.; Castaneto, M.; Poulain, J.; Dossat, C.; Tagu, D.; Reese, J. C.; Reeck, G. R.; Wilkinson, T. L.; Edwards, O. R. Predicted Effector Molecules in the Salivary Secretome of the Pea Aphid (Acyrthosiphon pisum): A Dual Transcriptomic/Proteomic Approach. *J. Proteome Res.* **2011**, *10* (4), 1505−1518.
(47) Vandermoten, S.; Harmel, N.; Mazzucchelli, G.; De Pauw, E.; Haubruge, E.; Francis, F. Comparative analyses of salivary proteins from three aphid species. *Insect Mol. Biol.* **2014**, *23* (1), 67−77.
(48) Boulain, H.; Legeai, F.; Guy, E.; Morlière, S.; Douglas, N. E.; Oh, J.; Murugan, M.; Smith, M.; Jaquiéry, J.; Peccoud, J.; White, F. F.; Carolan, J. C.; Simon, J. C.; Sugio, A. Fast Evolution and Lineage-Specific Gene Family Expansions of Aphid Salivary Effectors Driven by Interactions with Host-Plants. *Genome Biol. Evol.* **2018**, *10* (6), 1554−1572.
(49) Madhusudhan, V. V.; Miles, P. W. Mobility of salivary components as a possible reason for differences in response of alfalfa to the spotted alfalfa aphid and pea aphid. *Entomol. Exp. Appl.* **1998**, *86*, 25−39.
(50) Nardelli, A.; Vecchi, M.; Mandrioli, M.; Manicardi, G. C. The Evolutionary History and Functional Divergence of Trehalase (treh) Genes in Insects. *Front. Physiol.* **2019**, *10*, 62.
(51) Febvay, G.; Delobel, B.; Rahbé, Y. Influence of the amino acid balance on the improvement of an artificial diet for a biotype of *Acyrthosiphon pisum* (Homoptera: Aphididae). *Can. J. Zool.* **1988**, *66*, 2449−2453.
(52) Heie, O. E. Paleontology and phylogeny. In *Aphids: their biology, natural enemies, and control*; Minks, A. A. K., Harrewijn, P., Eds.; Elsevier: Amsterdam, The Netherlands, 1987; Vol. *2a*, pp 367−391.
(53) Zhou, Y.; Badgett, M. J.; Orlando, R.; Willis, J. H. Proteomics reveals localization of cuticular proteins in Anopheles gambiae. *Insect Biochem. Mol. Biol.* **2019**, *104*, 91−105.
(54) Moussian, B. Molecular model of skeletal organisation and differentiation. In *Extracellular composite matrices in arthropods*; Cohen, E., Moussian, B., Eds.; Springer: Switzerland, 2016; pp 67−87.
(55) Blackman, L. M.; Eastop, V. F. Taxonomic Issues. In *Aphids as Crop Pests*; van Emden, H. F., Harrington, R., Eds.; CABI: Wallingford, United Kingdom, 2007; pp 1−29.
(56) Uzest, M.; Blanc, S. Non circulative virus-vector interactions. In *Microbe-Arthropode vector interactions*; Brown, J., Ed.; USA: Academic Press: San Diego, 2016; pp 59−72.
(57) Rispe, C.; Legeai, F.; Papura, D.; Bretaudeau, A.; Hudaverdian, S.; Le Trionnaire, G.; Tagu, D.; Jaquiéry, J.; Delmotte, F. De novo transcriptome assembly of the grapevine phylloxera allows identification of genes differentially expressed between leaf- and root-feeding forms. *BMC Genomics* **2016**, *17* (1), 219.
(58) Wistrom, C. M.; Blaisdell, G. K.; Wunderlich, L. R.; Botton, M.; Almeida, R. P. P.; Daane, K. M. No evidence of transmission of grapevine leafroll-associated viruses by phylloxera (Daktulosphaira vitifoliae). *Eur. J. Plant Pathol.* **2017**, *147*, 147.
(59) Andersen, S. O. Amino acid sequence studies on endocuticular proteins from the desert locust, Schistocerca gregaria. *Insect Biochem. Mol. Biol.* **1998**, *28* (5−6), 421−434.
(60) Vannini, L.; Willis, J. H. Localization of RR-1 and RR-2 cuticular proteins within the cuticle of Anopheles gambiae. *Arthropod Struct. Dev.* **2017**, *46* (1), 13−29.
(61) Noh, M. Y.; Muthukrishnan, S.; Kramer, K. J.; Arakane, Y. Tribolium castaneum RR-1 cuticular protein TcCPR4 is required for formation of pore canals in rigid cuticle. *PLoS Genet.* **2015**, *11* (2), e1004963.
(62) Liu, W.; Gray, S.; Huo, Y.; Li, L.; Wei, T.; Wang, X. Proteomic Analysis of Interaction between a Plant Virus and Its Vector Insect Reveals New Functions of Hemipteran Cuticular Protein. *Mol. Cell. Proteomics* **2015**, *14* (8), 2229−2242.
(63) Badillo-Vargas, I. E.; Chen, Y.; Martin, K. M.; Rotenberg, D.; Whitfield, A. E. Discovery of Novel Thrips Vector Proteins That Bind to the Viral Attachment Protein of the Plant Bunyavirus Tomato Spotted Wilt Virus. *J. Virol.* **2019**, *93* (21), e00699−00619.
(64) Falk, B. W.; Tsai, J. H. Biology and molecular biology of viruses in the genus Tenuivirus. *Annu. Rev. Phytopathol.* **1998**, *36*, 139−163.
(65) Dombrovsky, A.; Gollop, N.; Chen, S.; Chejanovsky, N.; Raccah, B. In vitro association between the helper component-proteinase of zucchini yellow mosaic virus and cuticle proteins of Myzus persicae. *J. Gen. Virol.* **2007**, *88* (Pt 5), 1602−1610.
(66) Liang, Y.; Gao, X. W. The Cuticle Protein Gene MPCP4 of Myzus persicae (Homoptera: Aphididae) Plays a Critical Role in Cucumber Mosaic Virus Acquisition. *J. Econ. Entomol.* **2017**, *110* (3), 848−853.
(67) Barry, M. K.; Triplett, A. A.; Christensen, A. C. A peritrophin-like protein expressed in the embryonic tracheae of Drosophila melanogaster. *Insect Biochem. Mol. Biol.* **1999**, *29* (4), 319−327.
(68) Behr, M.; Hoch, M. Identification of the novel evolutionary conserved obstructor multigene family in invertebrates. *FEBS Lett.* **2005**, *579* (30), 6827−6833.
(69) Jasrapuria, S.; Specht, C. A.; Kramer, K. J.; Beeman, R. W.; Muthukrishnan, S. Gene families of cuticular proteins analogous to peritrophins (CPAPs) in Tribolium castaneum have diverse functions. *PLoS One* **2012**, *7* (11), e49844.
(70) Petkau, G.; Wingen, C.; Jussen, L. C. A.; Radtke, T.; Behr, M. Obstructor-A Is Required for Epithelial Extracellular Matrix Dynamics, Exoskeleton Function, and Tubulogenesis. *J. Biol. Chem.* **2012**, *287* (25), 21396−21405.
(71) Pesch, Y. Y.; Riedel, D.; Behr, M. Obstructor A Organizes Matrix Assembly at the Apical Cell Surface to Promote Enzymatic Cuticle Maturation in Drosophila. *J. Biol. Chem.* **2015**, *290* (16), 10071−10082.
(72) Andersen, S. O. Studies on proteins in post-ecdysial nymphal cuticle of locust, Locusta migratoria, and cockroach, Blaberus craniifer. *Insect Biochem. Mol. Biol.* **2000**, *30* (7), 569−577.
(73) Arakane, Y.; Lomakin, J.; Gehrke, S. H.; Hiromasa, Y.; Tomich, J. M.; Muthukrishnan, S.; Beeman, R. W.; Kramer, K. J.; Kanost, M. R. Formation of Rigid, Non-Flight Forewings (Elytra) of a Beetle Requires Two Major Cuticular Proteins. *PLoS Genet.* **2012**, *8* (4), e1002682.
(74) Dittmer, N. T.; Hiromasa, Y.; Tomich, J. M.; Lu, N.; Beeman, R. W.; Kramer, K. J.; Kanost, M. R. Proteomic and Transcriptomic Analyses of Rigid and Membranous Cuticles and Epidermis from the Elytra and Hindwings of the Red Flour Beetle, Tribolium castaneum. *J. Proteome Res.* **2012**, *11* (1), 269−278.

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX

(75) Noh, M. Y.; Kramer, K. J.; Muthukrishnan, S.; Kanost, M. R.; Beeman, R. W.; Arakane, Y. Two major cuticular proteins are required for assembly of horizontal laminae and vertical pore canals in rigid cuticle of Tribolium castaneum. *Insect Biochem. Mol. Biol.* **2014**, *53*, 22−29.
(76) Qiao, L.; Xiong, G.; Wang, R. X.; He, S. Z.; Chen, J.; Tong, X. L.; Hu, H.; Li, C. L.; Gai, T. T.; Xin, Y. Q.; Liu, X. F.; Chen, B.; Xiang, Z. H.; Lu, C.; Dai, F. Y. Mutation of a cuticular protein, BmorCPR2, alters larval body shape and adaptability in silkworm, Bombyx mori. *Genetics* **2014**, *196* (4), 1103−1115.
(77) Pan, P. L.; Ye, Y. X.; Lou, Y. H.; Lu, J. B.; Cheng, C.; Shen, Y.; Moussian, B.; Zhang, C. X. A comprehensive omics analysis and functional survey of cuticular proteins in the brown planthopper. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (20), 5175−5180.
(78) Deshoux, M.; Monsion, B.; Uzest, M. Insect cuticular proteins and their role in transmission of phytoviruses. *Curr. Opin. Virol.* **2018**, *33*, 137−143.
(79) Bahrami-Kamangar, S.; Christiaens, O.; Taning, C.; De Jonghe, K.; Smagghe, G. The cuticle protein MPCP2 is involved in Potato virus Y transmission in the green peach aphid Myzus persicae. *J. Plant Dis. Prot.* **2019**, *126*, 1−7.
(80) Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **2019**, *47* (D1), D442−D450.

https://dx.doi.org/10.1021/acs.jproteome.9b00851
J. Proteome Res. XXXX, XXX, XXX−XXX