**LEGEND SUPPLEMENTARY FIGURES**

**Figure S1.** PAUP\*10Altivec phylogenetic analysis of insect CSPs, DBPs, and RBPs. Mp10 and related proteins (Allergen, TIF, Mucin, WAS/WASL, and SamkC) are from Tables S2 and S3. Bommo-CSPs and Trica AAJJ0012J are from Xuan et al. [24] and Liu et al. [37-38]. Whitefly BemtaCSP1 is from Liu et al. [28, 30-31, 37, 67]. Bacterial CSPs (*E. coli* WP\_08938515) are from Liu et al. [38]. DBP amino acid sequences (transcriptional regulator, XRE, helix-turn-helix, AcrR, DGR, and Avd proteins) are from Table S3. Table S4 presents the selection of RBP proteins. **A.** Bio-NJ analysis. **B.** Bootstrap/Jackknife algorithm analysis. The red dots show orthology groupings supported by high bootstrap values (≥ 50%). The red asterisk indicates Mp10’s position (\*): grouping with WAS/WASL, Mucin, SamkC, CWP, Rho, and CSP. For comparative molecular analysis, the amino acid sequences of proteins are used. Amino acid tree (data matrix: total characters 1089, constant characters 285, variable parsimony-uninformative characters 279, parsimony-informative characters 525, all characters of type unord, all characters have equal weight): Length 5532, CI 0.653, RI 0.528, RC 0.345, HI 0.347, G-fit -361.010. Outgroup: Bommo-CS2, OIC81003, and OIC85870.

**Figure S2.** MyzpeMp10’s alignment with related proteins from the CSP, Allergen, Mucin, Rho, TIF, ASRP, and NPCP families (see Table S2). This is the “CSP” consensus sequence, below the alignment. High consensus color: red. Low consensus color: blue. Met-1 from Mp10 (Myzpe) and KAG6454812 (Manse) signal peptide (37-42 amino acids) is referred to as amino acid 1. The bacterial *A. baumannii* “CSPs” OIC81003 and OIC85870 are included for phylogenetic and comparative molecular sequence analysis of amino acid sequences (reference outgroup, see Figures 4 & S1). “CSPs” consisting of approximately 142-153 amino acids are mapped onto Mp10 to XP\_045521964 (short CSPs: Allergen group). The remaining sequences are associated with larger molecules (approximately 241 to 751 amino acids) that have a complete CSP at the protein's N-terminus (large CSPs: ASRP/Mucin group).

**Figure S3.** The modeling of the molecular structures of Mp10 and its related proteins belonging to the Allergen, Mucin, Rho, TIF, ASRP, and NPCP families: mucin-2-like, allergen Tha p 1, pherokine, acid trehalase, cell wall protein-stress response protein (CWP), PAN1-like, WAS/WASL, WASP, Rick-A, Formin-1-like, UL36, Ib alpha chain-like (PgIb), major outer envelope glycoprotein (jg5928), SamkC, LRR, and YLP motif-containing protein 1-like/NNK (see Table S2). Figure 4 illustrates how they align with “CSP” molecular sequences. In order to identify the signal peptide and cut it off based on N-terminal sequencing by Edman degradation, sequences were aligned with Bommo-CSP1 [17]. Next, using Swissmodel.expasy.org, the mature protein's amino acid sequence was subjected to molecule structure modeling as for WASP and PgIb. The molecules with the highest identity score were used as template references among 1kx9.1, 1n8u.1.A, A0A0K8TVC4.1.A, A0A182NU82.1.A, A0A1Q3FP99.1.A, A0A1V1WC08.1.A, A0A2H4ZB78.1.A, A0A345BEP2.1.A, A0A6H0D4B0.1.A, A0A6I8TVU3.1.A, A0A6I8TWU8.1.A, A0A835CMW0.1.A, D6WUE8.1.A, Q3LB93.1.A, Q6H8Y7.1.A, 9W0X2.1.A, and A7UL73.1.A. For every protein model, the Global Model Quality Estimation (GMQE) and the percentage of Sequence Identity (Seq Id) are shown. C: C-terminus, N: N-terminus. The -helices that make up the CSP prism are numbered 1 through 6. The position of the black square with dotted lines indicates that the CSP prism is always present in the N-terminus of mucin-2, CWP, TIF, Rho, NPCP, ACRP, PAN1-like, WAS/WASL, Rick-A, Formin-1-like, UL36, jg5928, SamkC, LRR, and NNK in addition to WASP and PgIb (see Figure 5). It is anticipated that in all of these intracellular protein families, the extra-long -helical C-terminal tail (up to 13 turns) represents a transmembrane domain (TMB, see Figure 5).