Conserved chemosensory proteins in the proboscis and eyes of Lepidoptera

Jiao Zhu[1#], Immacolata Iovinella[2,3#], Francesca Romana Dani[3], Yu-Ling Liu[2], Ling-Qiao Huang[2], Yang Liu[1], Chen-Zhu Wang[2], Paolo Pelosi[1], Guirong Wang[1]

SUPPLEMENTARY FIGURES

```
HarmCSP4     RPDGATYTDKYDNVDLDEILGNRRLMVPYIKCMLDQGKCAPDAKELKEHIKEALE
HarmCSP11    RPDGGTYTTKYDNVDLDEILANDRLLIPYIKCLLDEGKCAPDAKELKEHIREALE
PxylCSP729   RPNDSHYTDRYDNVNLDELISNRRLLVPYVKCVLDQGKCSPDGKELKEHIQEALE
PmacCSP648   ----EQYTDRYDNVDLDEILSNRRLLVPYINCILEEGKCSPDGKELKSHIKEALE
CsupCSP555   --LPQQYTDKYDSVDLTEILSNRRLLVPYLNCILTTGKCSPDGKELRSHIKEALE
                   ** :**.*:* *::.* **::**::*:*  ***:**.***:.**:****

HarmCSP4     NECGKCTEAQKKGTRRVIGHLINHEADFWNELTAKYDPERKYTTKYEKELKEVKA
HarmCSP11    NGCAKCTDKQKEGTRRVIAHLIKHKNADWQKLKAKYDPEGKYTHKYEKELEEVQH
PxylCSP729   NNCGKCTDKQREGTRKMIGHLINHEQEFWDQLIAKYDPERKYVSKYEKELKEVKA
PmacCSP648   NNCEKCTDTQKSGTRKVIGHLINNEKEYWGKLTAKYDPERKYVTKYETDLRKIAA
CsupCSP555   NYCAKCTETQKNGTRRVIGHLINNEADYWKKLVDKYDPERKYVVKYEEELKTVAA
             * * ***: *:.***::*.***:::   * :*  ***** **. *** :*. :


PpolCSP971   KPASTYTDKWDYINIDEILESQRLLKGYVDCLMDKGRCTADGKTLKETMPDALENE
CsupCSP460   ---NTYTDKWDHINVDEILESQRLQKGYVDCLLERGRCTPDGKALKETLPDALEHE
HarmCSP2     --QDQYESANDNFDISEVIGNDRLLHAYANCLLNKGPCTPEVKQVKEKLPEALETR
                . * .  * :::.*:: .:** :.*.:**:::* **.: * :**.:*:*** .

PpolCSP971   CSKCTQKQKEGSDKVIRFLINKRPELWKELATKYDPDNVYQQRYKDKIEAVKEH-
CsupCSP460   CSKCTEKQKTASDKVIRHLVNKRPDLWTELAAKYGPDNMYQQRYKDKIEEVKGKQ
HarmCSP2     CAKCTDKQKQMGKALAQEVKKNHPDIWKQLVAMYDPQGKYQQAWKDFLQE-----
             *:***:***  .. : : : :::*::*.:*.: *.*:. *** :** ::
```

Figure S1. Alignment of the eigth CSPs identified in the proboscis of the four species by proteomic analysis. Five sequences share about 70% of identical amino acids and can be considered as orthologues. The others are around 50% identical between each other and with the members of the first group. Harm: *Helicoverpa armigera*; Pxyl: *Plutella xylostella*; Pmac: *Papilio machaon*; Csup: *Chilo suppressalis*.

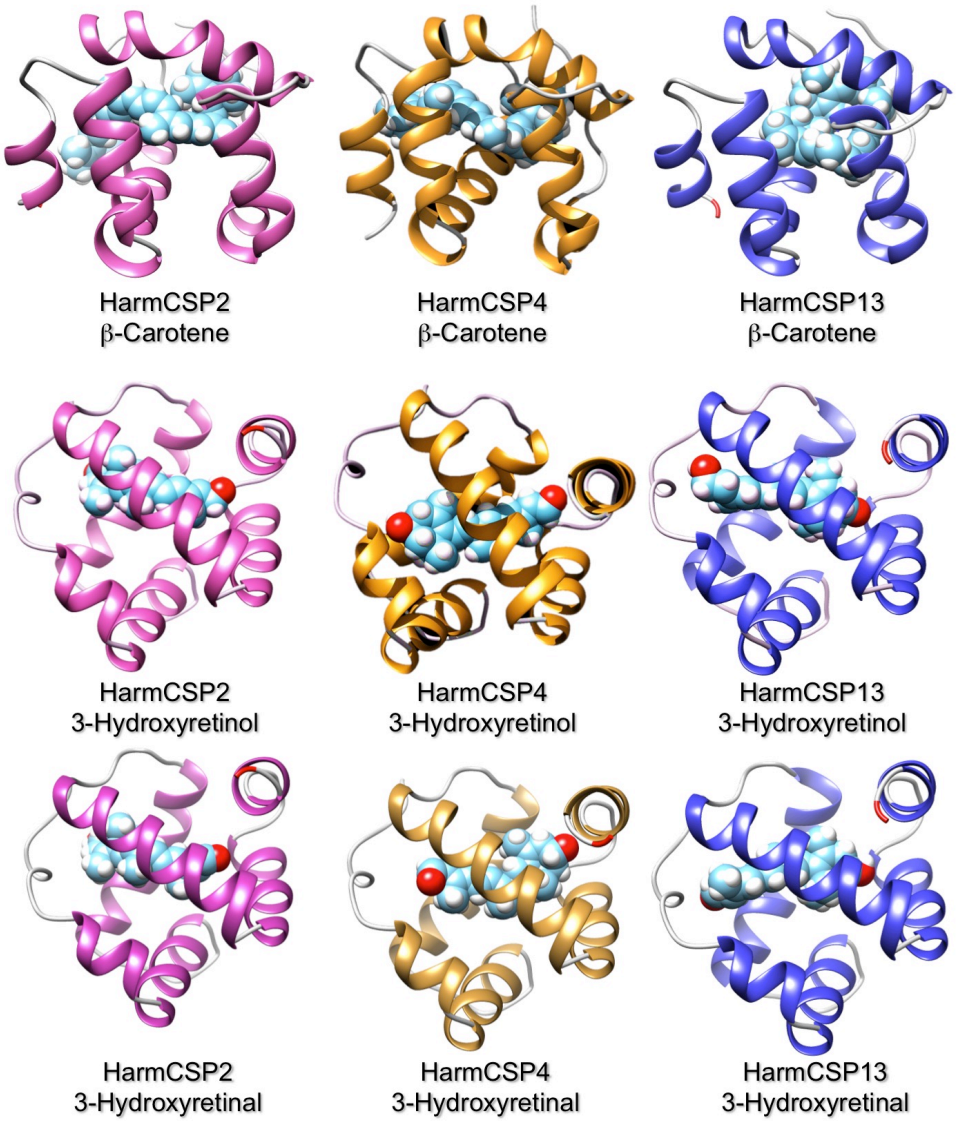

Figure S2. Docking simulations performed with models of three CSPs and three selected ligands, β-carotene, 3-hydroxyretinol and 3-hydroxyretinal. Results relative to HarmCSP11 are almost superimposable to those obtained with HarmCSP4 and are not shown. All four proteins were modeled on the CSP of *M. brassicae*, as a template, in its complexed form with three molecules of 12-Br-dodecanol [28] (PDB id: 1N8V) using the online programme SWISS MODEL [25-27]. Docking was performed by the on-line programme SWISS DOCK using default parameters [29]. Models were visualised with the UCSF Chimera package [3].

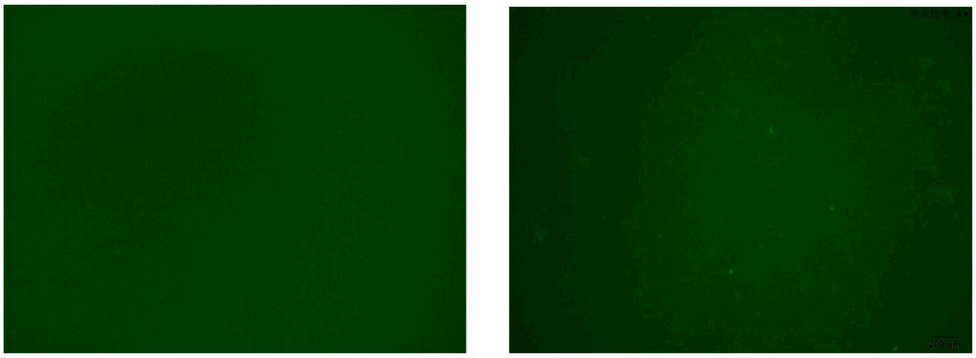

Figure S3. Control for the "drink blot" experiments. After allowing adults of *P. xylostella* drink from a nitrocellulose membrane soaked in a sugar solution, the membrane was treated with pre-immune serum in the place of primary antiserum and then with FITC-linked secondary antiserum. The Figure shows two examples of such experiments.

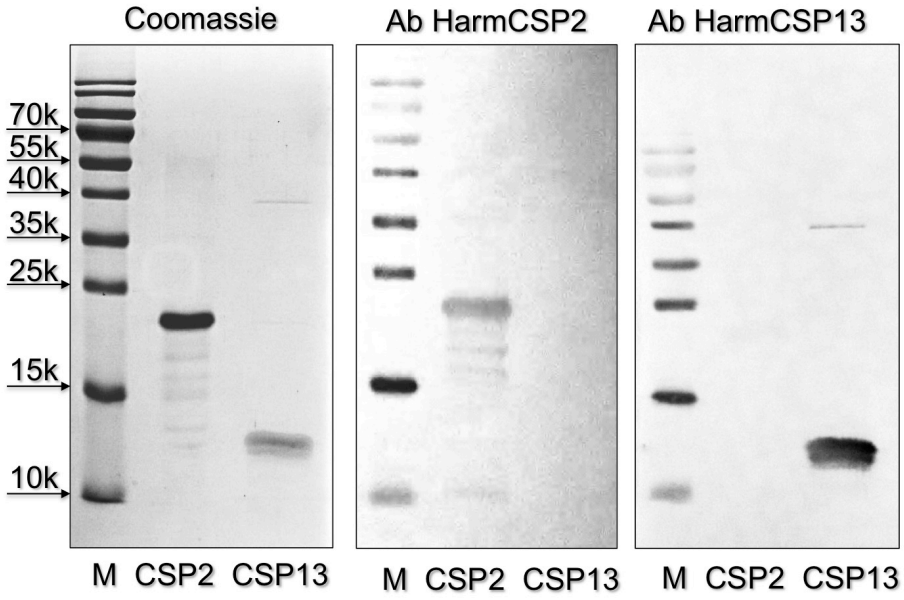

Figure S4. Western blot experiments to verify the quality and specificity of the two antisera prepared in this work. Each antiserum recognizes only the relative protein expressed in bacteria and purified. The stained bands correspond to the predicted molecular masses of the proteins. HarmCSP2 was expressed with a His-tag segment at its N-terminus, while CSP13 contained an additional methionine residue at the N-terminus as the only modification with respect to the mature sequence.

Table S1. Primers used for the expression of the four CSPs of *Helicoverpa armigera*.

Protein	PRIMERS (fw/rv)	Restriction sites	Added residues
HarmCSP2	AAAGAATTCCAAGACCAGTACGAGTCTGC AAACTCGAGTTATTCTTGAAGGAAGTC	EcoRI/XhoI	His-tag at N-terminus
HarmCSP4	CGCGTCCTGACGGCGCCACATA AAGAATTCTTAAGCCTTGACTTCTTT	MscI/EcoRI	Ala at N-terminus
HarmCSP11	CGCGCCCTGATGGTGG AACTCGAGTTAATGTTGAACTTCTTC	MscI/XhoI	Ala at N-terminus
HarmCSP13	CATATGGATGACAAGTACACGGA AAGAATTCTTATTCGGGGATCTGGATGCC	NdeI/EcoRI	Met at N-terminus