



**HAL**  
open science

## **A new non-classical fold of varroa odorant-binding proteins reveals a wide open internal cavity**

Beatrice Amigues, Jiao Zhu, Anais Gaubert, Giovanni Renzone, Philippe Leone, Simona Arena, Isabella Maria Fischer, Harald Paulsen, Wolfgang Knoll, Andrea Scaloni, et al.

### ► To cite this version:

Beatrice Amigues, Jiao Zhu, Anais Gaubert, Giovanni Renzone, Philippe Leone, et al.. A new non-classical fold of varroa odorant-binding proteins reveals a wide open internal cavity. Scientific Reports, 2021. hal-03358153

**HAL Id: hal-03358153**

**<https://hal.science/hal-03358153v1>**

Submitted on 29 Sep 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

**A new non-classical fold of varroa odorant-binding proteins reveals a wide open internal cavity**

Beatrice Amigues<sup>1#</sup>, Jiao Zhu<sup>2,4#</sup>, Anais Gaubert<sup>1</sup>, Giovanni Renzone<sup>3</sup>, Philippe Leone<sup>1</sup>,  
Simona Arena<sup>3</sup>, Isabella Maria Fischer<sup>2</sup>, Harald Paulsen<sup>4</sup>, Wolfgang Knoll<sup>2,5</sup>, Andrea  
Scaloni<sup>3</sup>, Alain Roussel<sup>1</sup>, Christian Cambillau<sup>1\*</sup>, Paolo Pelosi<sup>2\*</sup>

<sup>1</sup>*Architecture et Fonction des Macromolécules Biologiques (AFMB, UMR 6098), Centre National de la Recherche Scientifique (CNRS) & Aix-Marseille Université (AMU), Campus de Luminy, Case 932, 13288 Marseille Cedex 09, France*

<sup>2</sup>*Austrian Institute of Technology GmbH, Biosensor Technologies, Konrad-Lorenz Straße, 24, 3430 Tulln, Austria*

<sup>3</sup>*Proteomics & Mass Spectrometry Laboratory, ISPAAM, National Research Council, 80147 Napoli, Italy.*

<sup>4</sup>*Faculty of Biology, Institute of Molecular Physiology, Johannes Gutenberg-Universität, 55099 Mainz, Germany*

<sup>5</sup>*Department of Physics and Chemistry of Materials, Faculty of Medicine/Dental Medicine, Danube Private University, Krems, Austria*

# Contributed equally to the work

\*Correspondence to CC ([christian.cambillau@univ-amu.fr](mailto:christian.cambillau@univ-amu.fr)) or PP ([paolo.pelosi@ait.ac.at](mailto:paolo.pelosi@ait.ac.at); [ppelosi.obp@gmail.com](mailto:ppelosi.obp@gmail.com))

26 **Abstract**

27

28 Insect odorant-binding proteins (OBPs) are not present in other arthropods nor in other living  
29 organism. However, ticks, mites, spiders and millipedes contain genes encoding protein with  
30 sequence similarity to insect OBPs. In this work we have explored the structure and function  
31 of such non-insect OBPs in the mite *Varroa destructor*, a major pest for honey bees. *Varroa*  
32 OBPs present six cysteines, paired into three disulphide bridges, but their positions in the  
33 sequence and their connections are different from those of their insect counterparts.

34 VdesOBP1 structure was determined in two crystal forms closely related and is likely in a  
35 monomeric state. Its structure assembles 5  $\alpha$ -helices linked by three disulphide bridges, one  
36 of them exhibiting a different connection as compared to their insect counterparts.

37 Comparison with classical OBPs reveals that the second of the six  $\alpha$ -helices is lacking in  
38 VdesOBP. Ligand-binding experiments revealed molecules able binding only to specific  
39 OBPs with a moderate affinity, suggesting that either unknown compounds have to be  
40 identified as optimal ligands, or post-translational modifications present in the native proteins  
41 may be essential for modulating protein binding activity, or else these OBPs represent a  
42 failed attempt in evolution and are not used by the mites.

43

44

45

46

47

48

49

50

51

52 **Key words:** *Varroa destructor*, Odorant-binding protein, Three-dimensional structure;  
53 Disulphide bridges; Ligand-binding assays; Evolution.

54

55

56

57

58

## 59 Introduction

60

61 Odorant-binding proteins (OBPs) are carriers of pheromones and odours in vertebrates and  
62 insects<sup>1-4</sup>. Insect OBPs, which represent a class of proteins structurally different<sup>5,6</sup> from  
63 those of vertebrates<sup>7,8</sup>, have been detected so far only in Hexapoda<sup>4,9</sup>. They are  
64 polypeptides of 120-140 amino acids, folded in a compact structure generally made of six  $\alpha$ -  
65 helical domains. A pattern of six conserved cysteines paired in three interlocked disulphide  
66 bridges is a key feature of these proteins<sup>10-12</sup>. A single hydrophobic pocket binds small  
67 organic volatile compounds, such as pheromones and general odorants, with dissociation  
68 constants of the micromolar order<sup>13</sup>. Besides OBPs, other soluble carriers such as  
69 chemosensory proteins (CSPs), Nieman-Pick C2 (NPC2) proteins and lipocalins, might be  
70 likely involved in chemical communication of arthropods<sup>4,9,14-16</sup>. In particular, NPC2 proteins,  
71 are expressed with several members in all arthropods and represented until a couple of  
72 years ago the only known class of putative semiochemical carriers in Chelicerata and  
73 Crustacea.

74

75 However, recent work unveiled the presence of proteins with some sequence similarity to  
76 insect OBPs in spiders, ticks and mites, as well as in Myriapoda, and have been tentatively  
77 named "OBP-like"<sup>17-21</sup>. Such polypeptides exhibit values of amino acid identity with insect  
78 OBPs around 10-15%. However, such low similarity can be significant when compared with  
79 comparably low values of identities between the OBPs of different insect Orders. Given their  
80 phylogenetic relationship with their insect counterparts, we shall simply refer to these  
81 proteins as OBPs. Members of this family have not been found so far in Crustacea nor in  
82 other species of living organisms.

83

84 In order to understand the role of OBPs in Chelicerata, we have focused our attention on the  
85 mite *Varroa destructor*, the well-known honey bee parasite. The genome of this species  
86 contains 5 genes encoding members of the OBP family<sup>18,19,22</sup>. The sequences of the five  
87 encoded proteins are aligned in **Figure 1a**. At the mature protein level, VdesOBP1 and  
88 VdesOBP2 are very similar to each other with 72% identity. Also VdesOBP3 and VdesOBP4  
89 are closely related sharing 60% of their residues, but the identity levels drop to 15% or less  
90 between two groups and with VdesOBP5. Orthologues of these sequences are found also in  
91 other mites. Some representative examples are reported in the phylogenetic tree of **Figure**  
92 **1b**, where the three clusters (OBP1-2, OBP3-4 and OBP5) are clearly separated. A number  
93 of sequences from other ticks and mites fall into different clusters. In particular, all four OBPs  
94 of the spider mite *Tetranychus urticae* belong to a group well distinct from the three  
95 highlighted for varroa. Besides, the four OBPs of *T. urticae* are much more similar to each

96 other with identities around 40%. This suggests that different mites may have adopted  
97 different sub-classes of OBPs for their chemoreception systems, although we cannot  
98 exclude that such differences might be partly due to incomplete annotation. In any case,  
99 OBPs are present in each species of Chelicerata with a small number of genes, poorly  
100 conserved between species and highly divergent within the same species. Unfortunately,  
101 due to only few genomes available and their poor annotation, the picture provided by this  
102 tree is to be considered as partial and preliminary. The amino acid sequences used to build  
103 the phylogenetic tree are reported in [Supplementary Table S1](#) online.

104

105 Transcriptomic studies reported VdesOBP2 (XP\_022645714) as the member with the  
106 highest expression in the forelegs<sup>18</sup>, while the gene encoding VdesOBP3 (XP\_022653293)  
107 was found to be upregulated in the forelegs, where chemosensory appendages are located,  
108 with respect to the second pair of legs<sup>19</sup>. At the proteomic level, four OBPs were detected:  
109 VdesOBP1 (XP\_022672530), VdesOBP2, VdesOBP3 and VdesOBP4 (XP\_022653281), the  
110 last two being more highly expressed in the forelegs than in hindlegs<sup>22</sup>.

111

112 In this work, we have expressed and characterised the OBPs of this mite. We have also  
113 solved the crystal structure of VdesOBP1, which exhibits a folding markedly different from  
114 those of insect OBPs.

115

116

## 117 **Results**

118

### 119 *Protein expression and purification*

120 First we confirmed the five OBP sequences reported in the NCBI database by cloning the  
121 five genes from RNA extracted from whole individuals. Sequences were identical with the  
122 published ones at the amino acid level, while few single base substitutions were likely due to  
123 differences between the population of mites used in our study with respect to those used for  
124 the sequences stored in the NCBI database.

125

126 The proteins were expressed in bacteria, following standard protocols, as reported in the  
127 Methods section. We have successfully obtained VdesOBP1, VdesOBP2, VdesOBP4 and  
128 VdesOBP5 in medium to good yields. Instead, we were unable to express VdesOBP3 in  
129 amounts adequate for purification and consequently for its characterization. The proteins  
130 contained a starting methionine as the only addition to the natural sequence, and were  
131 purified by two or three chromatographic steps on anion-exchange columns, such as DE-52  
132 and HiPrep-Q. [Figure 2](#) reports a summary of expression and purification for the four OBPs.

133

#### 134 *Cysteine pairing assessment*

135 VdesOBP1-5 contain six cysteines occurring at conserved positions (Figure 1), although in a  
136 pattern different from the one typical of insect OBPs. This fact suggested possible  
137 differences in the cysteine pairing with respect to insects<sup>10-12</sup>. To assess this issue,  
138 VdesOBP1, VdesOBP2 and VdesOBP5 were subjected to extensive alkylation with  
139 iodoacetamide under non-reducing denaturing conditions; after purification, they were then  
140 digested with trypsin followed by a second treatment with chymotrypsin. Resulting digests  
141 were analyzed by nanoLC-ESI-Q-Orbitrap-MS/MS, and S-S-crosslinked peptides were  
142 assigned using dedicated bioinformatic procedures (Supplementary Table S2 online).  
143 Examples of mass spectra of disulphide-bridged peptides from VdesOBP1 and VdesOBP5  
144 digests are reported in Figure 3. These results and the absence of carboxyamidomethylated  
145 peptides in the protein digests demonstrated that all cysteines present in VdesOBPs are  
146 involved in disulphide bridges. Surprisingly, the ascertained, conserved disulphide pattern of  
147 VdesOBPs was unlike that of insect OBPs. In fact, a C1-C6, C2-C3 and C4-C5 arrangement  
148 was identified in VdesOBPs, which was completely different from the C1-C3, C2-C5, C4-C6  
149 typical of insect OBPs<sup>10-12</sup>. This phenomenon brings two major consequences in the  
150 structure of VdesOBPs: i) the pairing of adjacent cysteines should make *V. destructor*  
151 proteins more flexible and less compact than insect OBPs; ii) linking C1-C6 brings together  
152 the protein N- and C-termini, while these two ends point generally in opposite directions in  
153 insect OBPs.

154

#### 155 *Protein characterization in solution*

156 The protein oligomeric state of VdesOBP1 was characterized in solution at different pH  
157 values, namely 4.5, 7.0 and 9.7 using HPLC with multi-angle laser light scattering /UV/  
158 refractometry (SEC-MALLS-RI) coupled detectors (Wyatt, Santa-Barbara, USA). The results  
159 point to a monomeric state, with measured masses of 15,708, 16,143 and 15,306 Da  
160 (16.496 kDa theoretical mass), respectively (Supplementary Fig. S1 online)

161

#### 162 *Overall structure of VdesOBP1*

163 The three-dimensional structure of VdesOBP1 was solved by sulphur single-wavelength  
164 anomalous dispersion (SSAD) with a wavelength  $\lambda=1.512 \text{ \AA}$  on beamline proxima 2A at  
165 SOLEIL synchrotron (Saint-Aubin, France). Data were collected at  $1.59 \text{ \AA}$  resolution on a  
166 P3<sub>2</sub>21 crystal form (Table 1). Another crystal form was collected, with space group P2<sub>1</sub> at  
167 1.2 Å resolution, and its structure was determined by molecular replacement using the P3<sub>2</sub>21  
168 crystal form. VdesOBP1 crystal forms P3<sub>2</sub>21 and P2<sub>1</sub> 2 were crystallized at pH values of 7.5  
169 and 9.5, respectively. They contain one and two molecules in the asymmetric unit,

170 respectively. After refinement, the final R/Rfree values were 18.9/23.0 % and 18.3 / 20.5 %  
171 for the two crystal forms, respectively (Table 1).

172

173 The cloned protein is 147 residues long and possesses 6 cysteine residues (Fig. 4). The  
174 chain was built in the electron density map of the P2<sub>1</sub> crystal form with the exception of  
175 residues 1–6 at the N-terminus while the P3<sub>2</sub>21 crystal form starts at residue 11 and lacks  
176 loop 50-53. We checked the integrity of the protein in the crystal and our analysis showed  
177 that the protein is intact, confirming that the N-terminal segment is present but not ordered in  
178 the crystal. The rmsd value between the structures of the two VdesOBP1 crystal forms is  
179 1.55 Å.

180

181 Being the structure best defined, our description will apply to the P2<sub>1</sub> form. VdesOBP1 is  
182 mostly helical (Fig. 4). In VdesOBP1 five  $\alpha$ -helices are observed in contrast with all classical  
183 insect OBP/PBPs such as BmorPBP<sup>5</sup>, Lush<sup>23</sup> or LmadPBP<sup>24</sup> that assemble six  $\alpha$ -helices.  
184 The N- and C-terminal stretches are elongated forming a pseudo  $\beta$ -sheet of two strands  
185 (Fig. 4). Three disulphide bridges (21/141;SS-N-5, 41/68;SS1-2 and 111/132;SS4-5, Fig. 4a)  
186 strengthen the 3D fold of VdesOBP1.

187

188 It has been reported that OBPs may form functional dimers<sup>25,26</sup>. However, it cannot be  
189 concluded firmly for or against dimerization of OBPs in their biological context. In the crystal,  
190 dimers have been observed in the case of AgamOBP1, but this observation may not apply to  
191 solutions, due to an insufficient contact area<sup>27</sup>. In the case of VdesOBP1, we observe a  
192 monomer using multi-angle static light scattering (MALS)/dynamic light scattering and RI  
193 coupled on-line with a size-exclusion chromatography (SEC) column (MALS/UV/RI/SEC) at  
194 three pH values (Supplementary Fig. S1 online). In contrast, the P2<sub>1</sub> form contains two  
195 molecules in the asymmetric unit (Fig. 4b). A dimer can be reconstituted in the P3<sub>2</sub>21 form,  
196 using a symmetrical molecule in the crystal packing (Fig. 5b). However, the second  
197 monomer of one form does not display the same orientation relative to that of the second  
198 form when the first monomers are superimposed (Fig. 5b,c). Finally, the PISA server  
199 reported an ambiguous response for the P2<sub>1</sub> form (score=0.49) but discarded the possibility  
200 of a dimer in form P3<sub>2</sub>21 (score=0). All in all, it seems highly probable that the biologically  
201 relevant form is a monomer.

202

### 203 *The binding crevice*

204 The binding crevice, as detected by CavityPlus<sup>28</sup>, is large and wide open in the monomer,  
205 with dimensions of the floor of ~18x18 Å (Fig. 6a). This large surface implies that numerous

206 side-chains form the walls of the binding site (Fig. 6b). Although most side chains are from  
207 hydrophobic/aromatic residues (16/24) two basic and four acidic residues complete the wall.  
208 The active site is empty, except in form P2<sub>1</sub> where a buffer molecule, 2-(N-cyclohexylamino)-  
209 ethane sulfonic acid (NCES), is docked against a wall of the binding site (Fig. 6a).  
210 Examination of the binding site in the dimer reveals that the N-terminus of each monomer  
211 protrudes deeply in the binding site of the other monomer, drastically reducing its size. This  
212 effect explains the serendipitous binding of NCES as this ligand interacts with the cell wall of  
213 one monomer and the C-terminal extension of the other one (Fig. 6c). The sulfonic acid  
214 moiety of NCES forms a ionic bond with Arg31A side chain (atom NH<sub>2</sub>) and with the main  
215 chain N-H of Lys12B.

216

#### 217 *Comparison of VdesOBP1 with other OBPs/PBPs structures*

218 A search on the PDB with the DALI server reported hits with good statistics, taking into  
219 account the Z-score and the r/m.s.d.<sup>29</sup>. The best hit is that of entry 5dic (Z=7; r.m.s.d. 2.9 Å  
220 on 96 / 115 residues), the fatty acid binding protein OBP56a from the oral disk of the blowfly  
221 *Phormia regina* (PDB ID: 5dic, by Ishida, Y., Leal, W.S., Wilson, D.K., unpublished) with a  
222 non-classical OBP fold. This FBP/OBP has been shown to bind the C16 palmitic acid (but  
223 not the acetate), as well as the C18 stearic, oleic and linoleic acids, but not decanoic acid<sup>30</sup>.  
224 The structure is reported in complex with two molecules of palmitic acid in its wide cavity.  
225 Superposition of VdesOBP1 and PregOBP56a shows that VdesOBP1 helices 1, 2, 3, 4 and  
226 5 superimpose well with PregOBP56a helices 1, 3, 4, 5 and 6. Helix 2 of PregOBP56a  
227 closes its binding cavity, while a similar effect is not observed in VdesOBP1. This  
228 PregOBP56a helix 2 is replaced in VdesOBP1 by a loop between helices 1 and 2, projected  
229 out of the protein core (Fig. 7a). Another good fit (1ow4; Z=5.4; r.m.s.d 3.3 Å on 8-90/120  
230 residues) is observed with the PBP of the cockroach *Leucophaea maderae* (LmadPBP)  
231 (3X). In this case again, LmadPBP exhibits an extra helix instead of a loop between helices  
232 1 and 2 in VdedOBP1 (Fig. 7b). Two disulphide bonds are conserved between the two  
233 proteins, the first bridging helices 1 and 2, the second bridging helices 4 and 5 of VdesOBP1  
234 (Fig. 4a and Fig. 7b), although as result of different (C1-C6, C2-C3, C4-C5 in VdesOBP1  
235 and C1-C3, C2-C5, C4-C6 in LmadPBP) cysteine connectivities. The third disulphide bond  
236 that bridges helices 3 and 6 in LmadPBP is not observed in VdesOBP1. Instead, the third  
237 disulphide bridge is observed between helix 5 (corresponding to helix 6 in LmadPBP) and  
238 the N-terminus extended stretch in VdeOBP1 (Fig. 7b).

239

#### 240 *Ligand-binding assays*



241 The purified VdesOBP1 did not show any affinity to neither of the most common probes  
242 used in ligand-binding assays with OBPs, N- phenyl-1-naphthylamine (1-NPN) and 1-  
243 aminoanthracene (1-AMA). Extraction with dichloromethane to remove possible endogenous  
244 ligands did not prove effective. We also subjected the protein to a denaturing treatment with  
245 8 M urea and 1 mM DTT, followed by slow renaturation by extensive dialysis, but we still  
246 could not observe any binding with neither of the two probes. We also tried 8-phenylamino-  
247 naphthalenesulfonic acid (ANS) as well as larger fluorescent reporters<sup>31</sup>, such as N-  
248 naphthyl-1-naphthylamine (1-NNN) and N-phenyl-1-aminoanthracene (PAA) with no effect.  
249

250 It is still possible that the fluorescent probe binds the protein, but the event is not  
251 accompanied by a change in the emission spectrum. In fact, when adding 1-NPN to the  
252 protein, we observed a dose-dependent quenching of the intrinsic tryptophan fluorescence,  
253 suggestive of some affinity of this probe for VdesOBP1 (data not shown). In any case,  
254 competitive binding using a fluorescent reporter could not be used with VdesOBP1 nor with  
255 VdesOBP2. Therefore, we attempted to find ligands for these two proteins monitoring the  
256 quenching of intrinsic tryptophan fluorescence. This method, on the other hand, is limited to  
257 chemicals able to accept the energy transferred by the tryptophan (therefore structures with  
258 a region of conjugated  $\pi$ -electrons), and is not suitable to evaluate binding affinities. Instead,  
259 with VdesOBP4 and VdesOBP5 we obtained clear binding curves using 1-NPN as the  
260 fluorescent reporter (Figure 8a). Therefore, we performed competitive binding experiments  
261 only with these two OBPs. Figures 8c and 8d report some of the displacement curves  
262 obtained with VdesOBP4 and VdesOBP5, respectively. The structures of the ligands used in  
263 these experiments are shown in Figure 8b. We can notice that only coniferyl aldehyde is  
264 able to decrease the 1-NPN fluorescence to 50% at the maximum concentration used,  
265 qualifying for a modest-poor ligand. None of the other chemicals used in the experiments  
266 (listed in Supplementary Table S3 online) with the four proteins proved to be a better ligand.  
267 Representative curves relative to quenching of the intrinsic fluorescence of tryptophan in  
268 VdesOBP1 and VdesOBP2 are reported in Supplementary Fig. S2 online.

269

270 Summarising these data, we can first observe that some of the tested chemicals proved to  
271 be ligands for VdesOBP4 and VdesOBP5 in 1-NPN displacement assays, while they  
272 determined significant quenching of intrinsic tryptophan fluorescence in VdesOBP1 and  
273 VdesOBP2. The same few compounds (coniferyl aldehyde, safranal and vanillin) were the  
274 best ligands with all four proteins, although with poor dissociation constants between 10 and  
275 20  $\mu$ M. These ligands are associated with plants, in particular with leaf damage, and do not  
276 bear any relationship to the varroa habits. In particular none of the honey bee pheromones,

277 such as 9-ketodecanoic acid, homovanillyl alcohol or fatty acid esters were able to bind any  
278 of the tested proteins.

279

280

## 281 **Discussion**

282

283 This work reports the original structural and functional characterization of four out of the five  
284 OBPs encoded in the genome of the mite *Varroa destructor*, a major pest for honey-bee.

285 Mass spectrometric analysis demonstrated that the six cysteines VdesOBPs, whose  
286 sequence position do not match the pattern in insect OBPs, are paired into three disulphide  
287 bridges, (C1-C6, C2-C3 and C4-C5) different from that (C1-C3, C2-C5, C4-C6) of their  
288 insect counterparts. The overall folding, determined for VdesOBP1, is also novel and  
289 markedly different from the conserved structure of insect OBPs, although some similar  
290 aspects can be recognised.

291

292 VdesOBP1 structure was determined in two crystal forms closely related. Its structure  
293 assembles 5  $\alpha$ -helices linked by three disulphide bridges. Comparison with classical OBPs  
294 reveals that the second of the six  $\alpha$ -helices found in them is lacking in VdesOBP. Although  
295 dimer formation of VdesOBP1 is observed in the crystal, it most probably does not represent  
296 the functional form *in vivo*. This dimerization is probably driven by the presence of a wide  
297 and fully open cavity, able to accommodate the extended N- and C-terminal stretches of the  
298 opposite monomer. The large cavity opening can be ascribed to the absence of helix 2 in  
299 OBPs classical fold, that restrains and closes the internal binding site. In VdesOBP1, this  
300 helix is absent and replaced by a loop projected outside the protein core. Two disulphide  
301 bridges are conserved as compared to classical OBPs, although associated with different  
302 sequential cysteine connectivities. These observations suggest that VdesOBP1 possess an  
303 original fold remotely related to classical OBPs. Comparison with PregOBP56a suggests  
304 that two to three fatty acids could bind inside the VdesOBP1 cavity. Such an arrangement of  
305 three ligands was observed in *Mamestra brassicae* CSP (CSPMbraA6)<sup>30</sup>.

306

307 The unique structure of VdesOBP1 and the lack of strong ligands pose questions that, at this  
308 stage, can only be addressed by formulating hypotheses and models to be verified by further  
309 experimental work. A first idea, based on the limited binding activity of these proteins and on  
310 the small number of genes (4-5) expressed in *V. destructor*, as well as in other mites and  
311 ticks, might suggest that these genes represent by-products of evolution, which are not used  
312 any more. In fact, at least two more families of carrier proteins could be active in Chelicerata:

313 NPC2 proteins and lipocalins, apart from a third family (CCPs) that so far seems to be  
314 specific of spiders<sup>21</sup>. NPCs and lipocalins are represented by with 6 and 9 members,  
315 respectively in varroa, while in other Chelicerata often 10-20 members of NPC2s have been  
316 identified, with exceptions of nearly 50 genes, as in the spider mite *Tetranychus urticae*.  
317 Although making sense for the reason reported above, the idea that OBPs in Chelicerata  
318 might represent a dead end in evolution and a failed attempt to make efficient carrier  
319 proteins contrasts with proteomic findings showing significant expression of such proteins<sup>22</sup>.  
320 In fact, four (OBP1-OBP4) of the five VdesOBPs were identified in different parts of *V.*  
321 *destructor*, with one of them (OBP4) particularly abundant in sensory organs, namely the  
322 first pair of legs and the mouth parts.

323

324 If then the expression of OBPs at the protein level indicate that they are functional, then,  
325 along with a second hypothesis, their poor affinity might be related to the absence of post-  
326 translational modifications. It has been recently shown that the pig OBP1 can bind fatty acids  
327 or steroids depending on the phosphorylation and glycosylation status of the protein<sup>32</sup>. Our  
328 varroa OBPs, being expressed in bacteria, are devoid of any post-translational modification.  
329 On the other hand, we do not have any information on possible phosphorylation and/or  
330 glycosylation of these proteins in their native forms. This is certainly an interesting aspect,  
331 although not easy to investigate, not only for the tiny size of sensory organs, but also  
332 because phosphorylation in particular may be present only in certain time windows related to  
333 the physiology and the behaviour of the mite. Indeed, protein phosphorylation is highly  
334 related to signal transduction events regulating essential physiological process. Thus, we  
335 searched for putative phosphorylation sites in the five OBP sequences, and found a very  
336 large number of serine and threonine residues that could be potentially phosphorylated  
337 ([Supplementary Fig. S3](#) online). In addition we found few tyrosine residues as potential  
338 phosphorylation sites (2 in OBP1, 1 in OBP2 and OBP3, 4 in OBP5).

339

340 In the three-dimensional structure of VdesOBP1 no serine or threonine are found inside the  
341 binding pocket. However, phosphorylation of serine and threonine residues outside the  
342 binding pocket (as well as glycosylation of the protein) can affect the folding of the protein  
343 and influence its binding affinities. Further speculation would be inappropriate based on the  
344 available experimental data. Certainly, a structural study of the native proteins, regarding  
345 their post-translational modifications would be essential before proposing alternative modes  
346 of action.

347

348 **Materials and Methods**

349

350 *Biological material*

351 *V. destructor* mites were collected from the cells of an infected honeybee hive and used  
352 immediately for RNA extraction.

353

354 *Reagents*

355 Chemical compounds for buffers and solutions, as well as ligands for binding assays were of  
356 analytical grade from Merck, Austria, and. Ligands were dissolved in spectroscopic grade  
357 (Uvasol) methanol. Oligonucleotides and synthetic genes were custom synthesised at  
358 Eurofins Genomics, Germany. Enzymes and kits for DNA extraction and purification were  
359 from New England Biolabs, USA.

360

361 *RNA extraction and gene cloning*

362 Total RNA was extracted using Tri-reagent (Sigma-Aldrich) along with the protocol provided.  
363 cDNA was prepared from 2  $\mu$ L of total RNA using the kit qScript cDNA SuperMix (Quanta  
364 Bio) and the manufacturer's procedure. For sequence verification, PCR was performed on  
365 cDNA using specific primers at the 5' end and at the 3' end. The primers, which were also  
366 used for cloning into the expression vector, also contained the restriction sites NdeI and  
367 EcoRI at the 5' and the 3' ends, respectively. After digestion, the PCR products were  
368 inserted into pET30 vector previously linearized with the same enzymes. This procedure  
369 was adopted for all the genes of *V. destructor* except for *V. destructor* OBP1 (used for  
370 structural studies) that was produced from a synthetic gene (Eurofins Genomics), whose  
371 nucleotide sequence was optimized for *E. coli* expression.

372

373 *Gene cloning*

374 Synthetic genes encoding VdesOBP1 and VdesOBP2 were custom synthesised at Eurofins  
375 Genomics, Germany, and optimised for *E. coli* expression. The sequences encoding  
376 VdesOBP4 and VdesOBP5, instead, were amplified from cDNA, digested and cloned into  
377 the pET-30 expression vector (Novagen, Darmstadt, Germany). The constructs encoded  
378 only the mature protein sequences preceded by an additional methionine, which is not  
379 considered here for amino acid numbering. Expression was performed in *E. coli* BL-21  
380 (DE3) competent cells. The sequences of the primers used are reported in [Supplementary](#)  
381 [Table S4](#) online.

382

383 *Protein expression and purification*

384 Proteins synthesis was induced by 0.4 mM IPTG of the *E. coli* culture, grown for two  
385 additional hours at 37 °C, then harvested by centrifugation, sonicated and centrifuged at  
386 high speed. The Vdes OBPs were obtained in good yield in the supernatant, except for

387 VdesOBP4, which was produced in very poor yield and was present in the pellet, from which  
388 it was solubilized with 8 M urea and 1 mM DTT. The protein was refolded by dialysis for  
389 three against 50 mM Tris-HCl, pH 7.4. All OBPs were purified 2-3 chromatographic steps on  
390 anion-exchange columns DE-52 and HiPrep-Q 16/10, 20 mL (Bio-Rad).

391

#### 392 *Protein alkylation and digestion, and mass spectrometry analysis*

393 Samples of *V. destructor* OBP1, OBP2 and OBP5 (20-50 µg) dissolved in 0.1 M  
394 tetraethylammonium bicarbonate (TEAB), pH 6.5, containing 4 M guanidinium chloride, were  
395 treated with iodoacetamide (0.5 M final concentration) for 30 min, in the dark; then, proteins  
396 were insolubilized by addition of 6 vol of cold acetone, leaving the samples at -20 °C,  
397 overnight. After centrifugation at 12,000 rpm at 4° C, for 20 min, supernatants were  
398 removed, and recovered pellets were dried with a SpeedVac device (ThermoFisher  
399 Scientific, USA). Recovered proteins were dissolved in 0.05 M TEAB, pH 6.5 (2 µg/µL final  
400 concentration), treated with trypsin (1:10 w/w enzyme/substrate) for 16 h, at 37 °C, then with  
401 chymotrypsin (1:8 w/w enzyme/substrate) for 16 h, at 37 °C. Protein digests were desalted  
402 with ZipTip® C18 (Millipore, USA) and directly analyzed with a UltiMate 3000 HPLC RSLC  
403 nano-chromatographer (ThermoFisher Scientific) linked on-line through a nano-Spray ion  
404 source (Thermo Fisher Scientific) to a Q-ExactivePlus mass spectrometer (Thermo Fisher  
405 Scientific)<sup>33</sup>. Peptides were separated on an Acclaim PepMap RSLC C18 column (150 mm  
406 × 75 µm ID; 2 µm particle size; 100 Å pore size) (Thermo Fisher Scientific), at a flow rate of  
407 300 nL/min, using a gradient of solvent B (19.92/80/0.08 v/v/v water/acetonitrile/formic acid)  
408 in solvent A (99.9/0.1 v/v water/formic acid). Solvent B started at 3%, increased linearly to  
409 40% in 45 min, then raised to 80% in 5 min, where it remained for additional 4 min, before  
410 rapidly returning to 3%. Mass spectrometer worked in positive polarity using a data-  
411 dependent mode, performing a full MS1 scan in the range  $m/z$  345-1350, at a nominal  
412 resolution of 70,000, followed by MS/MS scans of the 10 most abundant ions in high energy  
413 collisional dissociation mode<sup>34</sup>. MS/MS spectra were acquired in a dynamic  $m/z$  range, with  
414 a nominal resolution of 17,500, a normalized collision energy of 28%, an automatic gain  
415 control target of 50,000, a maximum ion injection time of 110 ms, an isolation window of 1.2  
416  $m/z$  and a dynamic exclusion setting of 20 s.

417

#### 418 *Bioinformatics for peptide identification*

419 Firstly, raw mass data files were analyzed with Proteome Discoverer v. 2.4 package  
420 (Thermo Fisher Scientific), running by Mascot v. 2.6.1 (Matrix Science, UK) and Byonic™ v.  
421 2.6 (Protein Metrics, USA) software. Database searching was performed against a database  
422 containing the sequences of *V. destructor* OBP1, OBP2 and OBP5, trypsin, chymotrypsin  
423 and common protein contaminants. Parameters for database searching were variable

424 oxidation at Met, deamidation at Asn/Gln, pyroglutamate formation at Gln and  
425 carbamidomethylation at Cys. Mass tolerance setting included  $\pm 10$  ppm for precursor ions  
426 and  $\pm 0.05$  Da for fragmentation ions<sup>35</sup>. Proteolytic enzyme and maximum number of missed  
427 cleavages were set to trypsin, chymotrypsin and 5, respectively. Proteome Discoverer  
428 peptide candidates were considered confidently identified only when the following criteria  
429 were satisfied: i) protein and peptide false discovery rate (FDR) confidence: high; ii) peptide  
430 Mascot score:  $> 30$ ; iii) peptide spectrum matches (PSMs): unambiguous; iv) peptide rank  
431 (rank of the peptide match): 1; v) Delta CN (normalized score difference between the  
432 selected PSM and the highest-scoring PSM for that spectrum): 0. Byonic peptide candidates  
433 were considered confidently identified only when the following criteria were satisfied: i) PEP  
434 2D and PEP 1D:  $< 10 \times 10^{-5}$ ; ii) FDR: 0; iii) q-value 2D and q-value 1D:  $< 10 \times 10^{-5}$ . Finally,  
435 disulphide bridge assignment was obtained using dedicated BioPharma Finder v. 4.0  
436 (Thermo Fisher Scientific) and pLink v. 2.3.9 software<sup>36</sup>. Both programs were used for  
437 database searching, enabling the specific function of disulphide-linked peptides attribution,  
438 and applying the additional settings reported above for Proteome Discoverer and Byonic  
439 examinations. Disulphide-bridged peptide identification was considered reliable when  
440 BioPharma Finder and pLink results showed a “confidence score”  $> 95$  and/or an “E-value”  $<$   
441  $1.0^{-10}$ , respectively. Manual interpretation and verification of the candidate spectra were  
442 always performed.

443

#### 444 *Crystallization, data collection and structure determination of the P3<sub>21</sub> form*

445 The protein was concentrated at 20 mg/ml. Screening experiments were performed using  
446 several commercial kits. The sitting-drop method in Swisssi plates was used. The reservoirs  
447 of the Swisssi Plates were filled using a Genesis II (TECAN) pipetting robot. The nanodrops  
448 were dispensed by a MOSQUITO Crystal (TTP Labtech). Respectively 300nl, 200 nl and  
449 100nl of protein were dispensed onto the central of three-sitting-drop shelves. 100nl of the  
450 reservoir solution were added to the protein drops. The plates were sealed with a  
451 transparent film and stored in a R 1000 crystallization hotel (Formulatrix). Small crystals  
452 were obtained in 0.2M Cadmium Chloride, 2.2 ammonium sulfate. Those crystals were used  
453 for micro-seeding. The sitting-drop method in Greiner plates was used and new crystals  
454 were obtained by doing a re-screen in the commercial screen PEGII with the initial condition  
455 in which the first crystals were obtained. The droplet with the thicker crystals contained 2/3  
456 of the initial condition (0.2M Cadmium Chloride, 2.2 ammonium sulfate) + 1/3 of PEG II  
457 screen (0.2M calcium chloride, HEPES 0.1M pH 7,5; PEG 4000 30% w/v).  
458 The Crystal was cryo-cooled at 100K within this crystallization liquor with the addition of 10%  
459 glycerol as cryoprotectant. Structure determination used SAD-phasing method at remote  
460 wavelengths with sulphur atom anomalous dispersion on native crystals. A strong theoretical

461 Bijvoet ratio of 1.12% was expected from the 10 sulphur atoms of the protein that contribute  
462 to the anomalous signal at wavelength of 1.77 Å. The sulphur-SAD data set was collected  
463 on the PROXIMA-2 beamline (SOLEIL) equipped with a high performance micro-  
464 diffractometer with a mini-kappa (MD2), an X-ray fluorescence detector (KETEK), and a fast,  
465 low-noise, photon-counting, area detector EIGER X 9M (238 fps in 9M mode, 750 fps in 4M  
466 mode). The wavelength was optimized to 1.77 Å and a redundant data set was collected at  
467 1.81 Å resolution. including 3600 images of 0.1° rotation. The data set was processed with  
468 XDS<sup>37</sup>. The data sets were indexed and integrated with Pointless and scaled with CCP4  
469 SCALA<sup>38</sup>. VdesOBP1 crystallized in space group P3<sub>2</sub>21 with cell dimensions a= 53.84 Å, b =  
470 53.84 Å, and c = 91.74 Å and and γ= 120 with one molecule in the asymmetric unit,  
471 corresponding to a solvent content of 50.87%. Six sulfur atom sites were found by using  
472 PHENIX Autosol<sup>39</sup>. The results from PHENIX include a PDB containing the anomalous  
473 scatterers, an MTZ file containing the experimentally determined phases, and the optimized  
474 phases from density modification. Structure refinement was performed with REFMAC5<sup>40</sup> and  
475 AutoBUSTER<sup>41</sup> alternated with rebuilding using COOT<sup>42</sup>.

476  
477 *Crystallization data collection and structure determination of the P2<sub>1</sub> form*

478 The protein at 20 mg/ml was subjected to screening experiments were performed using  
479 several commercial kits. VdesOBP1 was set in 2/3 of the initial crystallization conditions  
480 (CHES pH9.5, 0.1M, 30% PEG3000 w/v ) and 1/3 of ammonium sulfate screen (0.2 M  
481 sodium nitrate, 2.2M ammonium sulfate). Crystals of VdesOBP1 were briefly soaked in  
482 crystallization solution supplemented with 10 (v/v) polyethylenglycol. Diffraction data were  
483 collected at 1.2 Å resolution on beamline Proxima-1 at SOLEIL, Paris, France. The data sets  
484 were integrated with XDS<sup>37</sup> and were scaled with CCP4 SCALA<sup>38</sup>. VdesOBP1 crystallized  
485 in space group P2<sub>1</sub> with cell dimensions a= 31.6Å, b = 112.96 Å, and c = 41.82 Å and  
486 β= 98.95 with two molecules in the asymmetric unit, corresponding to a solvent content of  
487 44.48%. The structure was solved by molecular replacement with MOLREP<sup>43</sup> using the  
488 structure of VdesOBP1 form P3<sub>2</sub>21 as starting model. Refinement was performed with  
489 REFMAC5<sup>40</sup> and AutoBUSTER<sup>41</sup> and the structures were corrected with COOT<sup>42</sup>.

490  
491 *Multi-Angle Light Scattering*

492 100 μL of purified VdesOBP1 were analyzed by SEC-MALS using a HPLC system (Thermo  
493 Scientific) equipped with a Superdex 75 increase coupled with a multi-angle light scattering  
494 device (Dawn 8, Wyatt Technology) and an Optilab TrEX differential refractive index detector  
495 (Optilab TrEX, Wyatt Technology). At first, the column was first equilibrated with a buffer  
496 containing 0.02 % w/v NaN<sub>3</sub> for several hours, until the base line of the Optilab TrEX signal  
497 was stable. Then, light scattering and concentration data were analysed in ASTRA to

498 determine molecular weight and size distribution. Inter-detector band broadening and  
499 alignment were applied using the algorithms in ASTRA 7 (Wyatt Technology, USA) using  
500 BSA protein standard according to manufacturer's recommended methods. Three different  
501 buffers were used to understand the behaviour of the protein at different pH values, namely  
502 50 mM CHES pH 10, 50 mM Hepes pH 7 and 50 mM sodium acetate pH.5). The  
503 temperature controlled-autosampler was kept at 10 °C.

504

#### 505 *Ligand-binding assays*

506 Direct binding of fluorescent probes was monitored by adding aliquots of 1 mM methanol  
507 solution of the probe to a 2 µM solution of the protein in 50 mM Tris-HCl, pH 7.4 to final  
508 concentrations of 2 to 16 µM. The excitation wavelength was 337 nm and intensities were  
509 measured on the maximum of the peak, usually between around 410 and 420 nm. For  
510 VdesOBP1 and VdesOBP2 binding of ligands was evaluated by monitoring the intrinsic  
511 tryptophan quenching of the protein. Excitation wavelength was 295 nm and intensities were  
512 recorded around 335-340 nm. For VdesOBP4 and VdesOBP5 competitive binding  
513 experiments were titrating a solution of protein and 1-NPN both at 2 µM in Tris-HCl buffer,  
514 pH 7.4 with 1 mM methanolic solutions of ligands to final concentrations of 2-16 µM. All  
515 measurements were performed with a FL 6500 spectrofluorometer (PerkinElmer); slits were  
516 set at 5 nm for both excitation and emission, and 1 cm path quartz cuvettes were used.  
517 Dissociation constants for 1-NPN were evaluated using Prism software. Affinities of other  
518 ligands were calculated from the corresponding  $[IC]_{50}$  values (the concentration of each  
519 ligand halving the initial value of fluorescence), from the equation:

$$520 \quad K_d = [IC]_{50}/1 + [1-NPN]/K_{NPN}$$

521 where  $[1-NPN]$  is the concentration of free 1-NPN and  $K_{NPN}$  is the dissociation constant of  
522 the complex OBP/1-NPN.

523

524 **Author contributions:** J.Z., P.P. and C.C. designed research; J.Z., B.A., S.A, P.L., G.R. and  
525 A.G. performed research; P.P., C.C., A.R., J.Z., B.A., S.A, G.R., A.S. H.P. and W.K. analysed  
526 data and discussed results; P.P., A.S., A.R. and W.K. contributed all reagents, analytical tools  
527 and laboratory equipment; C.C., P.P. and A.S. wrote the paper. All authors read, revised and  
528 approved the final manuscript.

529 **The authors declare no conflict of interest.**

530 **Data deposition:** The three-dimensional structure of VdesOBP1 reported in this paper has  
531 been deposited in the database with PDB ID 7NYJ and 7NZA

532



533

534

## 535 **References**

536

- 537 1. Pelosi, P., Baldaccini, N. E. & Pisanelli, A. M. Identification of a specific olfactory  
538 receptor for 2-isobutyl-3-methoxypyrazine. *Biochemical Journal* **201**, 245–248 (1982).
- 539 2. Vogt, R. G. & Riddiford, L. M. Pheromone binding and inactivation by moth antennae.  
540 *Nature* **293**, 161–163 (1981).
- 541 3. Leal, W. S. Odorant Reception in Insects: Roles of Receptors, Binding Proteins, and  
542 Degrading Enzymes. *Annual Review of Entomology* **58**, 373–391 (2013).
- 543 4. Pelosi, P., Iovinella, I., Zhu, J., Wang, G. & Dani, F. R. Beyond chemoreception: diverse  
544 tasks of soluble olfactory proteins in insects: Soluble olfactory proteins in insects. *Biol*  
545 *Rev* **93**, 184–200 (2018).
- 546 5. Sandler, B. H., Nikonova, L., Leal, W. S. & Clardy, J. Sexual attraction in the silkworm  
547 moth: structure of the pheromone-binding-protein–bombykol complex. *Chemistry &*  
548 *Biology* **7**, 143–151 (2000).
- 549 6. Tegoni, M., Campanacci, V. & Cambillau, C. Structural aspects of sexual attraction and  
550 chemical communication in insects. *Trends in Biochemical Sciences* **29**, 257–264  
551 (2004).
- 552 7. Bianchet, M. A. *et al.* The three-dimensional structure of bovine odorant binding protein  
553 and its mechanism of odor recognition. *Nature structural biology* **3**, 934 (1996).
- 554 8. Tegoni, M., Ramoni, R., Bignetti, E., Spinelli, S. & Cambillau, C. Domain swapping  
555 creates a third putative combining site in bovine odorant binding protein dimer. *Nature*  
556 *Structural Biology* **3**, 863–867 (1996).
- 557 9. Pelosi, P., Iovinella, I., Felicioli, A. & Dani, F. R. Soluble proteins of chemical  
558 communication: an overview across arthropods. *Front. Physiol.* **5**, (2014).
- 559 10. Scaloni, A. Analysis of post-translational modifications in soluble proteins involved in  
560 chemical communication from mammals and insects. in *Methods in Enzymology* vol.  
561 642 103–124 (2020).
- 562 11. Scaloni, A., Monti, M., Angeli, S. & Pelosi, P. Structural Analysis and Disulfide-Bridge  
563 Pairing of Two Odorant-Binding Proteins from *Bombyx mori*. *Biochemical and*  
564 *Biophysical Research Communications* **266**, 386–391 (1999).
- 565 12. Leal, W. S., Nikonova, L. & Peng, G. Disulfide structure of the pheromone binding  
566 protein from the silkworm moth, *Bombyx mori*. *FEBS Letters* **464**, 85–90 (1999).
- 567 13. Pelosi, P., Zhu, J. & Knoll, W. From Gas Sensors to Biomimetic Artificial Noses.  
568 *Chemosensors* **6**, 32 (2018).

- 569 14. Angeli, S. *et al.* Purification, structural characterization, cloning and immunocytochemical  
570 localization of chemoreception proteins from *Schistocerca gregaria*. *Eur J Biochem* **262**,  
571 745–754 (1999).
- 572 15. Zhu, J. *et al.* Niemann-Pick C2 Proteins: A New Function for an Old Family. *Front.*  
573 *Physiol.* **9**, 52 (2018).
- 574 16. Ishida, Y. *et al.* Niemann–Pick type C2 protein mediating chemical communication in the  
575 worker ant. *PNAS* **111**, 3847–3852 (2014).
- 576 17. Renthal, R. *et al.* The chemosensory appendage proteome of *Amblyomma americanum*  
577 (Acari: Ixodidae) reveals putative odorant-binding and other chemoreception-related  
578 proteins. *Insect Science* **24**, 730–742 (2017).
- 579 18. Eliash, N. *et al.* Chemosensing of honeybee parasite, *Varroa destructor* : Transcriptomic  
580 analysis. *Scientific Reports* **7**, 1–11 (2017).
- 581 19. Eliash, N. *et al.* *Varroa* chemosensory proteins: some are conserved across Arthropoda  
582 but others are arachnid specific. *Insect Molecular Biology* **28**, 321–341 (2019).
- 583 20. Vizueta, J. *et al.* Evolution of Chemosensory Gene Families in Arthropods: Insight from  
584 the First Inclusive Comparative Transcriptome Analysis across Spider Appendages.  
585 *Genome Biol Evol* **9**, 178–196 (2017).
- 586 21. Vizueta, J., Rozas, J. & Sánchez-Gracia, A. Comparative Genomics Reveals Thousands  
587 of Novel Chemosensory Genes and Massive Changes in Chemoreceptor Repertoires  
588 across Chelicerates. *Genome Biol Evol* **10**, 1221–1236 (2018).
- 589 22. Iovinella, I. *et al.* Proteomic analysis of chemosensory organs in the honey bee parasite  
590 *Varroa destructor*: A comprehensive examination of the potential carriers for  
591 semiochemicals. *Journal of Proteomics* **181**, 131–141 (2018).
- 592 23. Kruse, S. W., Zhao, R., Smith, D. P. & Jones, D. N. M. Structure of a specific alcohol-  
593 binding site defined by the odorant binding protein LUSH from *Drosophila melanogaster*.  
594 *Nature Structural & Molecular Biology* **10**, 694–700 (2003).
- 595 24. Lartigue, A. *et al.* The Crystal Structure of a Cockroach Pheromone-binding Protein  
596 Suggests a New Ligand Binding and Release Mechanism. *J. Biol. Chem.* **278**, 30213–  
597 30218 (2003).
- 598 25. Qiao, H. *et al.* Cooperative interactions between odorant-binding proteins of *Anopheles*  
599 *gambiae*. *Cell. Mol. Life Sci.* **68**, 1799–1813 (2011).
- 600 26. Biessmann, H. *et al.* The *Anopheles gambiae* Odorant Binding Protein 1 (AgamOBP1)  
601 Mediates Indole Recognition in the Antennae of Female Mosquitoes. *PLOS ONE* **5**,  
602 e9471 (2010).
- 603 27. Lartigue, A. *et al.* Sulfur Single-wavelength Anomalous Diffraction Crystal Structure of a  
604 Pheromone-Binding Protein from the Honeybee *Apis mellifera* L\*. *Journal of Biological*  
605 *Chemistry* **279**, 4459–4464 (2004).

- 606 28. Xu, Y. *et al.* CavityPlus: a web server for protein cavity detection with pharmacophore  
607 modelling, allosteric site identification and covalent ligand binding ability prediction.  
608 *Nucleic Acids Research* **46**, W374–W379 (2018).
- 609 29. DALI and the persistence of protein shape - Holm - 2020 - Protein Science - Wiley  
610 Online Library. <https://onlinelibrary.wiley.com/doi/full/10.1002/pro.3749>.
- 611 30. Campanacci, V. *et al.* Moth chemosensory protein exhibits drastic conformational  
612 changes and cooperativity on ligand binding. *PNAS* **100**, 5069–5074 (2003).
- 613 31. Mastrogiacomo, R., Iovinella, I. & Napolitano, E. New fluorescent probes for ligand-  
614 binding assays of odorant-binding proteins. *Biochemical and Biophysical Research  
615 Communications* **446**, 137–142 (2014).
- 616 32. Nagnan-Le Meillour, P. *et al.* Binding Specificity of Native Odorant-Binding Protein  
617 Isoforms Is Driven by Phosphorylation and O-N-Acetylglucosamylation in the Pig *Sus  
618 scrofa*. *Front. Endocrinol.* **9**, (2019).
- 619 33. Zhu, J. *et al.* Reverse chemical ecology: Olfactory proteins from the giant panda and  
620 their interactions with putative pheromones and bamboo volatiles. *Proc Natl Acad Sci  
621 USA* **114**, E9802–E9810 (2017).
- 622 34. Arena, S., Renzone, G. & Scaloni, A. A multi-approach peptidomic analysis of hen egg  
623 white reveals novel putative bioactive molecules. *Journal of Proteomics* **215**, 103646  
624 (2020).
- 625 35. Arena, S. & Scaloni, A. An Extensive Description of the Peptidomic Repertoire of the  
626 Hen Egg Yolk Plasma. *J. Agric. Food Chem.* **66**, 3239–3255 (2018).
- 627 36. Chen, Z.-L. *et al.* A high-speed search engine pLink 2 with systematic evaluation for  
628 proteome-scale identification of cross-linked peptides. *Nature Communications* **10**, 3404  
629 (2019).
- 630 37. Kabsch, W. XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132, (2010).
- 631 38. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta  
632 Crystallogr D Biol Crystallogr* **67**, 235-242 (2011).
- 633 39. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for  
634 macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221  
635 (2010).
- 636 40. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal  
637 structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-367 (2011).
- 638 41. Blanc, E. *et al.* Refinement of severely incomplete structures with maximum  
639 likelihood in BUSTER-TNT. *Acta Crystallogr D Biol Crystallogr* **60**, 2210-2221 (2004).
- 640 42. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of  
641 Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).

642 43. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr D*  
643 *Biol Crystallogr* **66**, 22-25 (2010).  
644  
6453  
646  
647  
648

**Table 1: Data collection and refinement statistics.**

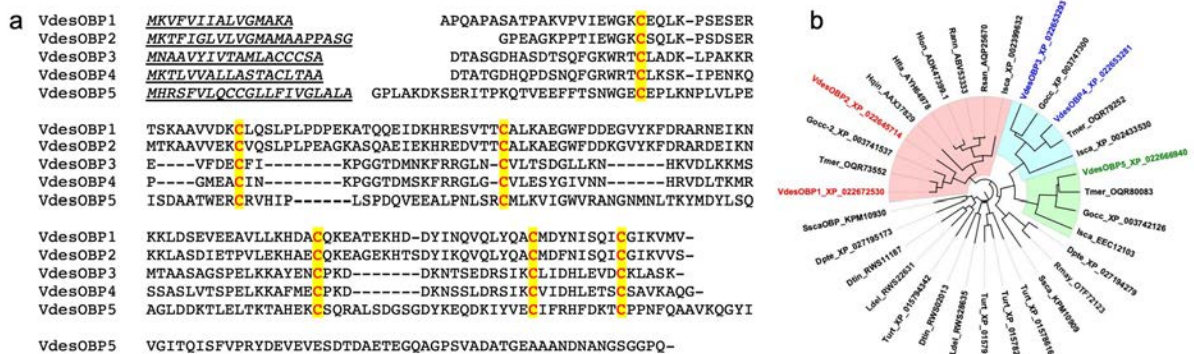
Data collection		
Beamline	SOLEIL Proxima2	SOLEIL Proxima1
PDB ID	7NYJ	7NZA
Space group	P3 <sub>2</sub> 21	P2 <sub>1</sub>
Cell parameters, a, b, c (Å), $\gamma = 120^\circ$	53.8, 53.8, 91.7	31.6, 113.0, 41.8 $\beta = 98.95^\circ$
Wavelength (Å)	1.770	0.9786
Resolution (Å)	46.6-1.81 (1.95-1.81)	40.09-1.20 (1.27-1.20)
Rmerge <sup>a</sup> (%)	7.9 (107)	6.5 (88.5)
Mean( $ I /sd(I)$ ) <sup>a</sup>	13.26 (1.87)	13.15 (1.19)
cc(1/2)	99.9 (74.2)	99.8 (78.8)
No. of observations <sup>a</sup>	284352 (57758)	634776 (86086)
No. of unique reflections <sup>a</sup>	27234 (5620)	87891 (13358)
Completeness <sup>a</sup> (%)	100.0 (100.0)	97.2 (91.4)
Multiplicity <sup>a</sup>	10.4 (10.3)	7.0 (6.4)
Refinement		
Resolution (Å)	46.59-1.81 (1.857-1.810)	38.8-1.20 (1.23-1.20)
No. of reflections <sup>a</sup>	13827 (1009)	83470 (5555)
atoms:protein/ligands/solvent	2289/36/393	2325 / 36 / 340
No. of test set reflections	728 (53)	4402 (308)
R <sub>work</sub> / R <sub>free</sub> <sup>a</sup> (%)	18.6 (28.6) / 22.6 (31.0)	18.0 (33.0) / 20.1(34.1)
rmsd bonds (Å) / angles (°)	0.015 / 1.83	0.018/ 2.20
B-factors (Å <sup>2</sup> )	40.5	19.0
Ramachandran: favored / allowed (%)	98.5 / 1.5	96.0/ 4.0

<sup>a</sup> Refers to the highest-resolution bin

652 **Figures legends**

653

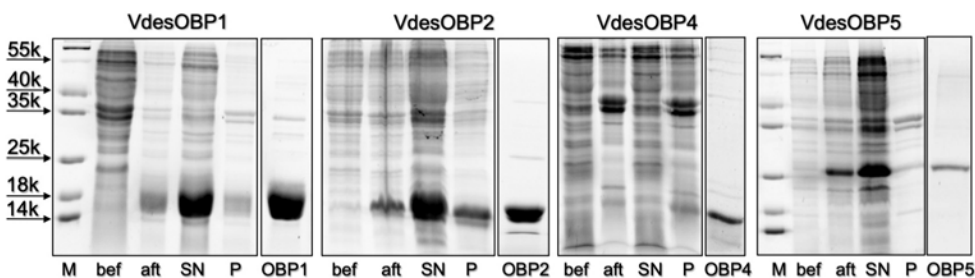
654 **Figure 1. (a) Alignment of the five OBPs of *V. destructor*.** The six conserved cysteines  
 655 are highlighted. OBP1 and OBP2 are very similar (72% identity), as OBP3 and OBP4 (60%),  
 656 but proteins of the two groups share only about 15% of their residues between them as well  
 657 as with OBP5. **(b) Phylogenetic tree of OBPs on ticks and mites.** Vdes: *Varroa*  
 658 *destructor*; Turt: *Tetranychus urticae*; Hqin: *Haemaphysalis qinghaiensis*; Hfla:  
 659 *Haemaphysalis flava*; Hlon: *Haemaphysalis longicornis*; Isca: *Ixodes scapularis*; Gocc:  
 660 *Galendromus occidentalis*; Tmer: *Tropilaelaps mercedesae*; Rann: *Rhipicephalus annulatus*;  
 661 *Rsan: Rhipicephalus sanguineus*; Ldel: *Leptotrombidium delicense*; Dtin: *Dinotrombidium*  
 662 *tinctorium*; Dpte: *Dermatophagoides pteronyssinus*; Ssca: *Sarcoptes scabiei*; Rmay:  
 663 *Euroglyphus maynei*. All sequences used to build the tree are reported in **Supplementary**  
 664 **Table S1** online.



665

666

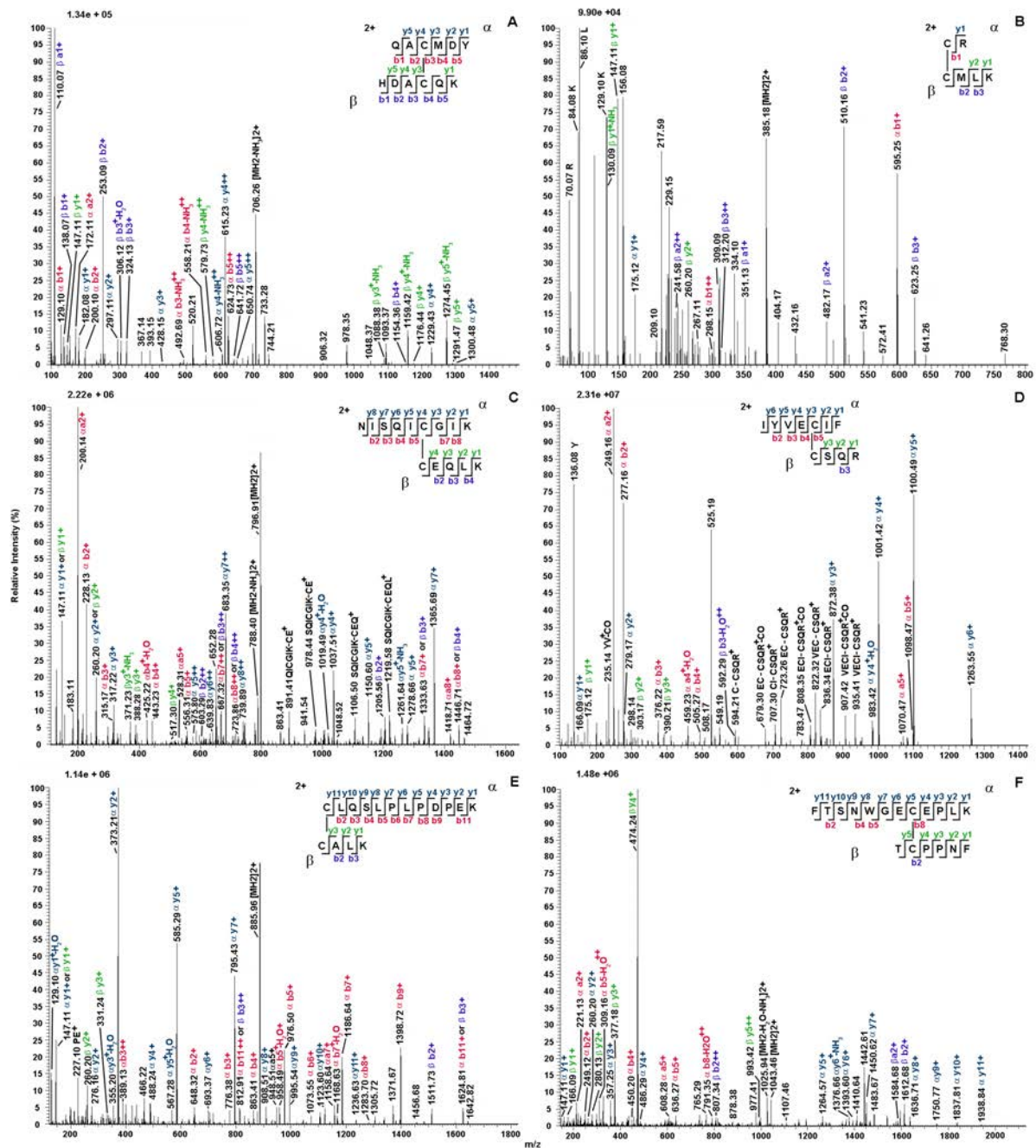
667 **Figure 2. Expression and purification of four OBPs of *V. destructor*.** While OBP1, OBP2  
 668 and OBP5 were easily produced in bacteria, OBP3 and OBP4, which are structurally related,  
 669 proved difficult to express. In fact, OBP4 was obtained only in very low yield, while we could  
 670 not observe expression of OBP3 in any of the different constructs adopted. M: Molecular  
 671 weight markers; bef: crude sample before culture induction; aft: crude sample after induction  
 672 with IPTG; SN: supernatant; P: pellet; OBP: sample of purified protein.



673

674

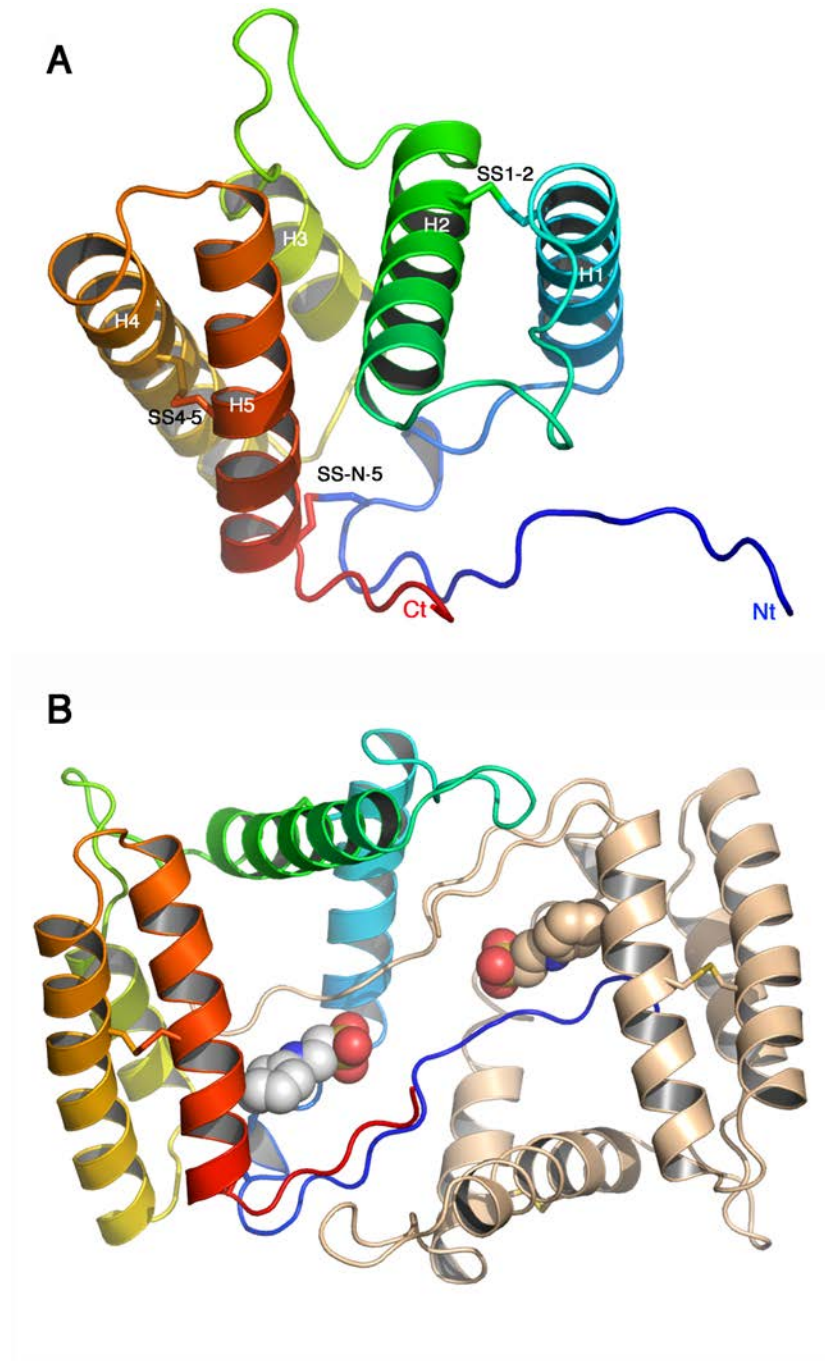
675 **Figure 3. Exemplificative fragmentation spectra of disulfide-bridged peptides**  
 676 **identified in the tryptic-chymotryptic digest of *V. destructor* OBP1 and OBP5 as**  
 677 **revealed by nanoLC-ESI-Q-Orbitrap-MS/MS analysis.** Panels A, C and E report  
 678 peptides from OBP1 digest, while panels B, D and F report peptides  
 679 from OBP5 digest. The fragments are highlighted in different colour depending on peptide  
 680 present in S-S-linked species, and corresponding b and y ion series. Complete data on  
 681 disulfide-bridged peptides are reported in Table S1.



682  
 683  
 684 **Figure 4. Crystal structure of *Vdes*OBP1 form P2<sub>1</sub>.** (a) Ribbon view of a monomer  
 685 displaying helices 1-5 and disulphide bridges. The amino-acid chain is rainbow coloured,



686 from blue (N-terminus) to red (C-terminus). The disulphide are identified by the secondary  
687 structures they belong to. **(b)** Ribbon view of the dimer. One monomer is rainbow coloured,  
688 the other one is coloured brown. The serendipitously bound buffer molecule 2-(N-  
689 cyclohexylamino)-ethane sulfonic acid (NCES) is displayed as atomic spheres (C: white, N:  
690 blue and O: red).

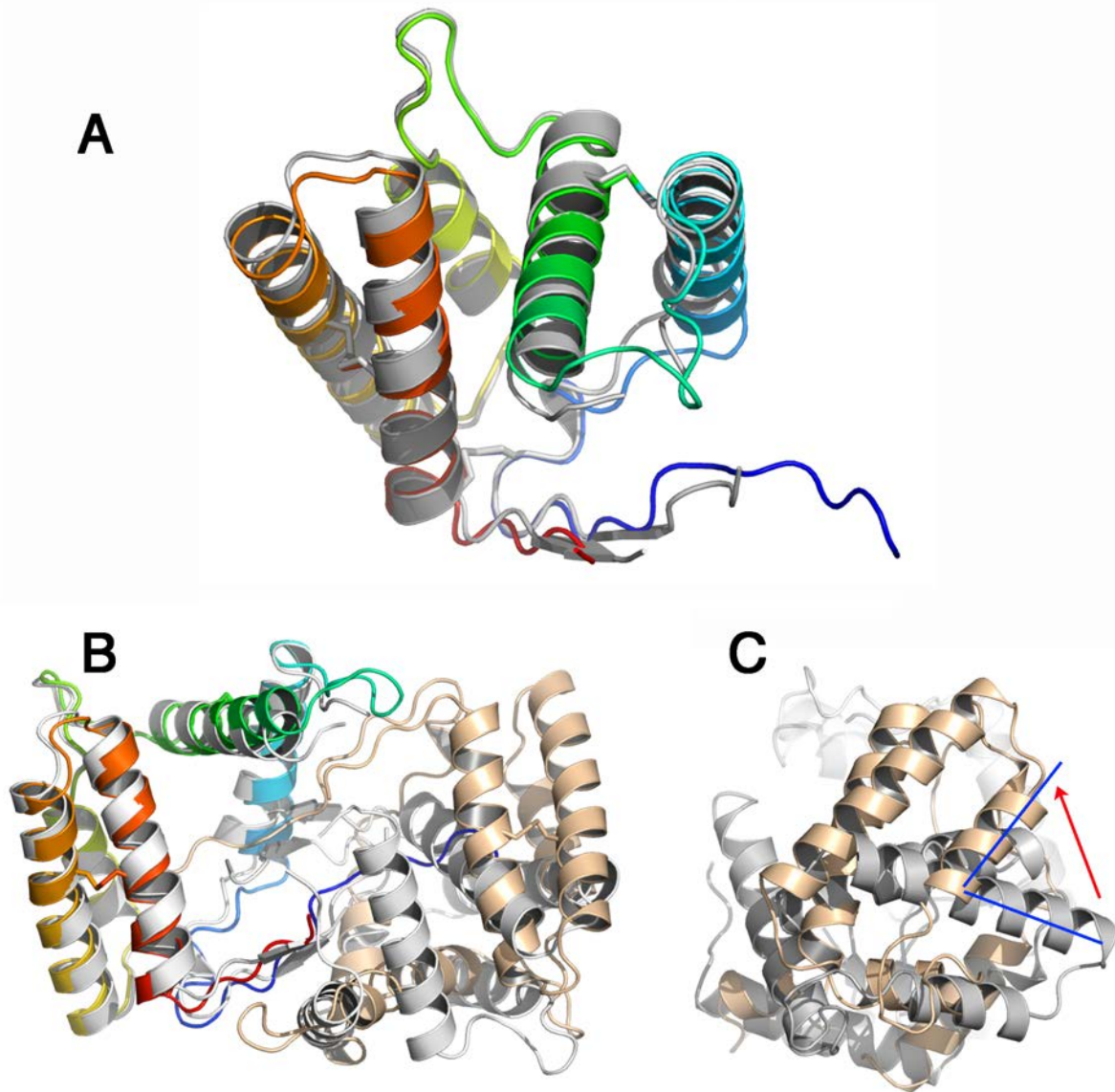


691  
692

693 **Figure 5. Comparison of crystal structures of VdesOBP1 forms P2<sub>1</sub> and P3<sub>221</sub>.** (a)  
694 Ribbon view of the superposition of monomers belonging to forms P2<sub>1</sub> (rainbow coloured)  
695 and P3<sub>221</sub> (grey). (b) Ribbon view of the dimers of forms P2<sub>1</sub> and P3<sub>221</sub>. In form P2<sub>1</sub> one



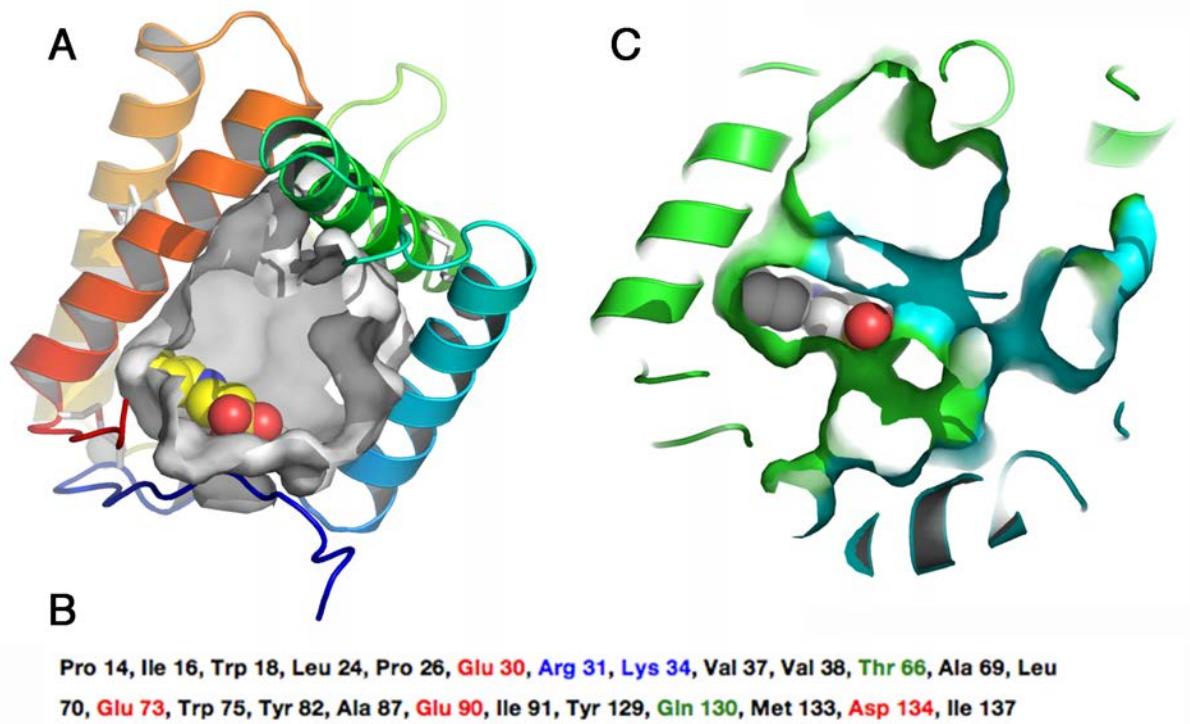
696 monomer is rainbow coloured, the other one is coloured brown. The dimer of P<sub>3</sub><sub>2</sub>21 form is  
 697 coloured light grey. One monomer of each form has been superimposed (left). The second  
 698 monomers do not superimpose. B/same view rotated 90° around a vertical axis. Note the  
 699 rotation of ~80° of the second monomer of form P<sub>3</sub><sub>2</sub>21 relative to the second monomer of  
 700 form P<sub>2</sub><sub>1</sub>.



701  
 702

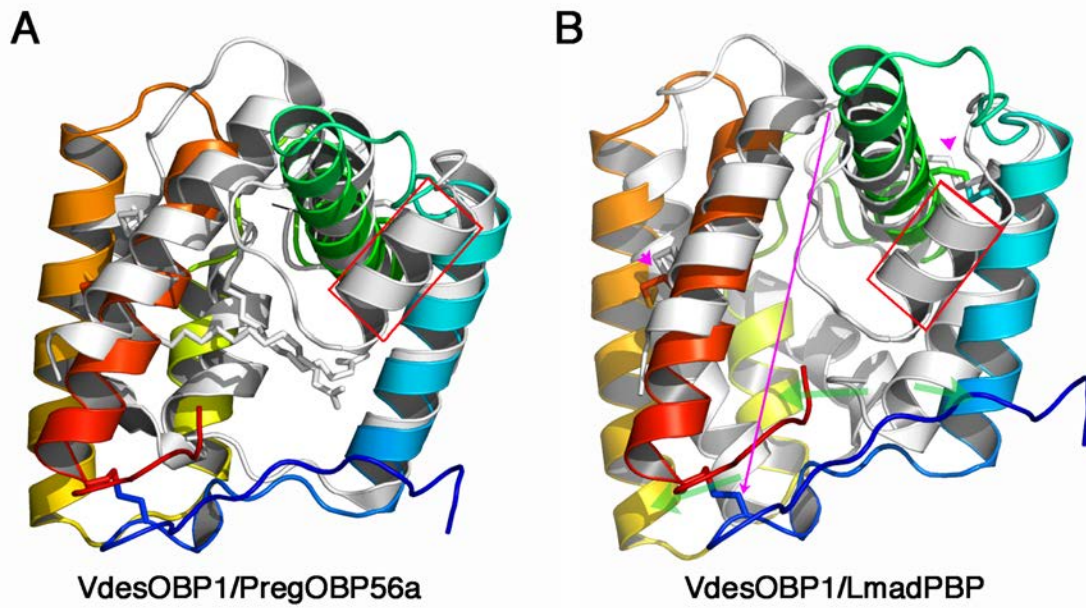
703 **Figure 6. The internal cavity of VdesOBP1 form P<sub>2</sub><sub>1</sub>.** (a) Ribbon view of a monomer  
 704 rainbow coloured. The cavity is represented by a grey surface and is wide open. The  
 705 serendipitously bound buffer molecule 2-(N-cyclohexylamino)-ethane sulfonic acid (NCES) is  
 706 displayed as atomic spheres (C: white, N: blue and O: red), and is stacked against the cavity  
 707 wall. (b) List of residues forming the cavity wall. Hydrophobic residues are coloured black,  
 708 acidic residues are in red and basic are blue. hydrophilic non-charged residues are coloured  
 709 green. (c) Ribbon view of a dimer rainbow coloured. The cavity is represented by a green

710 surface originating from monomer 1 and a blue surface from monomer 2. The  
 711 serendipitously bound buffer molecule NCES is displayed as atomic spheres (C: white, N:  
 712 blue and O: red), and is stacked against the cavity walls from both monomers.



713  
 714

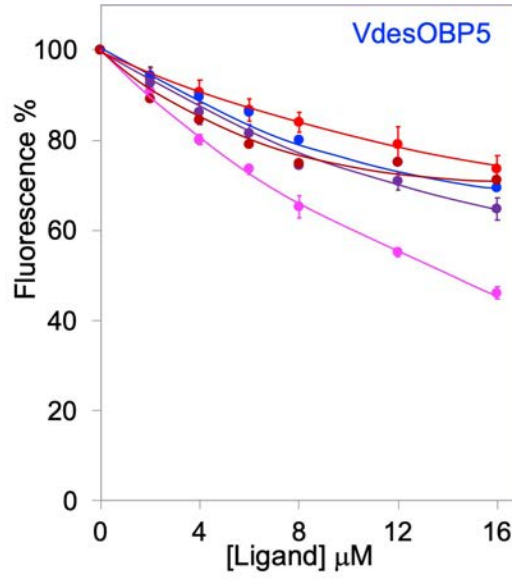
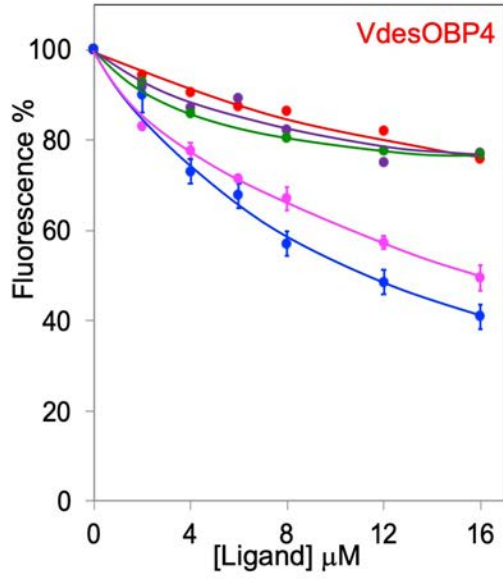
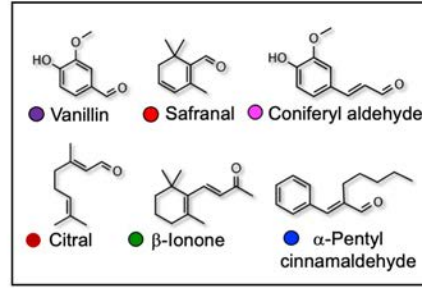
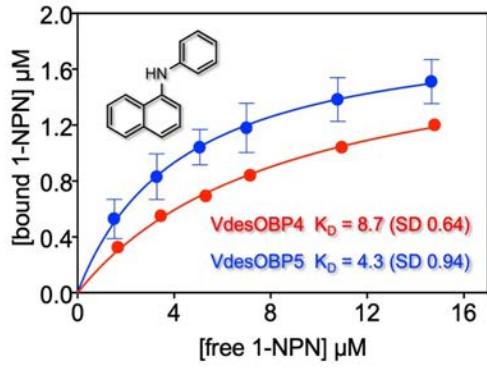
715 **Figure 7. Comparison of crystal structures of VdesOBP1 with the closest related**  
 716 **OBPs from *Phormia regina* (PregOBP56) and *Leucophera madereae* (LmadPBP).** (a)  
 717 Ribbon view of the superposition of monomers from VdesOBP1 (rainbow coloured) and  
 718 PregOBP56 (light grey) with its two fatty acid ligands (grey sticks). Note the absence of the  
 719 classical OBP fold helix 2 (red squared) in VdesOBP1. (b) Ribbon view of the superposition  
 720 of monomers from VdesOBP1 (rainbow coloured) and Lmad (light grey) with its two fatty  
 721 acid ligands (grey sticks). Note the absence of the classical OBP fold helix 2 (red squared) in  
 722 VdesOBP1 as compared to LmadPBP (red squared) and the displacement of one disulphide  
 723 bridge (red arrow). The 3D positions of the two other disulphide bridges are conserved  
 724 (purple arrow), in the final fold, although as result of different C1-C6, C2-C3, C4-C5  
 725 (VdesOBP1) and C1-C3, C2-C5, C4-C6 (LmadPBP) connectivities.



726

727

728 **Figure 8. Ligand-binding assays. (a) Binding of the fluorescent probe N-**  
 729 **phenyl-naphthylamine (1-NPN) to VdesOBP4 and VdesOBP5.** With the other two proteins  
 730 expressed (VdesOBP1 and VdesOBP2) we could not observe the blue shift usually  
 731 associated with binding of the fluorescent probe. **(c, d) Competitive binding assays.** Only  
 732 few chemicals were able to displace 1-NPN from the complex with VdesOBP4 and  
 733 VdesOBP5. The same chemicals were also the best ligands for the other VdesOBPs  
 734 ([Supplementary Fig. S2](#) online). **(b) Structures of the ligand used in the competitive binding**  
 735 **experiments.**



736