



OPEN

## An odorant-binding protein in the elephant's trunk is finely tuned to sex pheromone (Z)-7-dodecenyl acetate

Valeriia Zaremska<sup>1</sup>, Giovanni Renzone<sup>2</sup>, Simona Arena<sup>2</sup>, Valentina Ciaravolo<sup>2</sup>, Andreas Buberl<sup>3</sup>, Folko Balfanz<sup>3</sup>, Andrea Scaloni<sup>2</sup>, Wolfgang Knoll<sup>1,4</sup> & Paolo Pelosi<sup>1</sup>✉

Chemical communication in elephants has been well studied at the chemical and behavioural levels. Pheromones have been identified in the Asian elephant (*Elephas maximus*), including (Z)-7-dodecenyl acetate and frontalinalin, and their specific effects on the sexual behaviour of elephants have been accurately documented. In contrast, our knowledge on the proteins mediating detection of pheromones in elephants remains poor and superficial, with only three annotated and reliable entries in sequence databases, two of them being odorant-binding proteins (OBPs), and the third a member of von Ebner's gland (VEG) proteins. Proteomic analysis of trunk wash extract from African elephant (*Loxodonta africana*) identified one of the OBPs (LafrOBP1) as the main component. We therefore expressed LafrOBP1 and its Asian elephant orthologue in yeast *Pichia pastoris* and found that both recombinant proteins, as well as the natural LafrOBP1 are tuned to (Z)-7-dodecenyl acetate, but have no affinity for frontalinalin. Both the natural and recombinant LafrOBP1 carry post-translational modifications such as O-glycosylation, phosphorylation and acetylation, but as these modifications affect only a very small amount of the protein, we cannot establish their potential effects on the ligand-binding properties of OBP1.

The order Proboscidea is phylogenetically separated from all other mammals and includes three living species, the Asian elephant (*Elephas maximus*) and the two African species, the savanna and the forest elephant (*Loxodonta africana* and *L. cyclotis*, respectively)<sup>1</sup>, besides the extinct mammoths. In this work, we focus on the savanna African elephant *L. africana* as no chemistry on chemical signalling by the forest African elephant *L. cyclotis* has been published to our knowledge.

Communication between sexes in elephants has been mainly studied at the chemical and behavioural level, thanks to the long and accurate work of the late Elizabeth Rasmussen and her collaborators, who continued her research producing several interesting and unique results<sup>2</sup>. Two sex pheromones were first identified in the Asian elephant: (Z)-7-dodecenyl acetate, which is released in the urine of pre-ovulatory females, and frontalinalin {1,5-dimethyl-6,8-dioxabicyclo[3.2.1]octane}, the main component of the male temporal gland secretion (TGS) during musth<sup>3-8</sup>. Intriguingly, both compounds are also pheromones for a large number of Lepidoptera and Coleoptera, respectively<sup>9</sup>. Later work reported the presence of the same compounds in the urine and in the TGS of female African elephants, together with a number of other volatiles, such as *exo*- and *endo*-brevicomin {(1R,5S,7R)-7-ethyl-5-methyl-6,8-dioxabicyclo[3.2.1]octane and (1R,5S,7S)-7-ethyl-5-methyl-6,8-dioxabicyclo[3.2.1]octane, respectively}, (*E*)- $\beta$ -farnesene [(6E)-7,11-dimethyl-3-methylene-1,6,10-dodecatriene] and (*E,E*)- $\alpha$ -farnesene [(3E,6E)-3,7,11-trimethyldodeca-1,3,6,10-tetraene], all known also as insect pheromones<sup>9</sup>. Moreover, frontalinalin was also detected in the blood of musth males<sup>10</sup>. An interesting aspect is the role of chirality played in the action of frontalinalin. In young males, the (+)-enantiomer is predominant in TGS, but the presence of its optical antipode increases with age reaching the composition of a racemate at full maturity<sup>11</sup>. On the contrary, females produce almost pure (–)-frontalinalin<sup>10</sup>. Also both enantiomeric forms of *endo*- and *exo*-brevicomin occur in blends of the four stereoisomers in variable ratios<sup>9,10</sup>.

<sup>1</sup>Austrian Institute of Technology GmbH, Biosensor Technologies, Konrad-Lorenz Straße, 24, 3430 Tulln, Austria. <sup>2</sup>Proteomics, Metabolomics and Mass Spectrometry Laboratory, ISPAAM, National Research Council, 80055 Portici, Italy. <sup>3</sup>Vienna Zoo, Maxingstraße 13B, Vienna, Austria. <sup>4</sup>Department of Physics and Chemistry of Materials, Faculty of Medicine/Dental Medicine, Danube Private University, Krems, Austria. ✉email: ppelosi.obp@gmail.com; paolo.pelosi@ait.ac.at

EmaxOBP1		MLEEPLLDDEYCS	EISGTWYTIYEASANIEVL
LafrOBP1	<u>MKVLVLTLLGLVYA</u>	DPLEEPLLDDEYCLE	EISGTWYTIYEASANIEVI
LafrOBP2	<u>MOTLFLTIMFSLITLQA</u>	QDPLSFAQEEQNI	TGTWYVKAMVASKNLPL
LafrVEG	<u>MOTLFLTLVLGLIAILQAQA</u>	NPDLYKNVSADGREGGWGFP	SPDREGLLQVLGRWYLKAEIADKDVS-
EmaxOBP1	SENSPLRGYFRLIK	CHPDGETLLVIFYTKENGT	CQLYNKQGQRIDENGYTTNYEGKVDFSF
LafrOBP1	SENSPLSGYFRHIE	CHPDGETLLVIFYTKENGI	CQLYNKQGQRIDENWYTTNYEGKVDFSF
LafrOBP2	EKRPKVSPLTFTALK	--GGSLEISFTFLKKDQC	HENRIVMESVGEPEGKYS
LafrVEG	EKNVEVVGPMITALE	--GGNLEAKNTILMNGQC	QEVKLVLEKTHEPRKFTAYGGERHVFIERTHVKD
EmaxOBP1	LLIHAINKNEEGNVFEVVGALAREKDISEENYQAFLEFAVENGIPKENIVK	---	VIDTDTCPETLT
LafrOBP1	LLIYVINKNEEGNVFEVVGALAREKDISEENYQAFLEFAVENGIPKENIVK	---	VIDTDTCPETPT
LafrOBP2	HYIILGECQHDGRTLVRGKLLGRSLDVNAEAL	EEFRKFTQOKGFQOENIT	---
LafrVEG	HYILYCEGQLKGQOVRMAKLVGEWTS	GALQPLPPP-PLSSTPSFS	SHNFLRIGFCSSSSQCPGQV

**Figure 1.** Two orthologous sequences in the Asian elephant *Elephas maximus* (EmaxOBP1: AAN17333) and in the African elephant *Loxodonta africana* (LafrOBP1: XP\_023395442) share 93% of their residues and are annotated as OBP1. Another sequence, found only in the African elephant (LafrOBP2: SP\_tr[G3U771]) is only 15% identical with LafrOBP1. In addition, a member of the VEG proteins, found only in the databases of the African elephant (LafrVEG: XP\_023398512.1), is 13% and 31% identical with LafrOBP1 and LafrOBP2, respectively. Predicted signal peptides are indicated in italics and underlined. The sequence of EmaxOBP1 is found in the databases as reported, without a signal peptide. Cysteine residues are highlighted.

Behaviour has been extensively investigated, also using the identified pheromones, and generated a wealth of interesting and novel information<sup>6,10,12</sup>. This can be briefly summarised as follows:

1. Males are attracted to the urine of females in estrus and communicate their readiness with frontalin secreted by TGS<sup>13</sup>. While musth is observed regularly in males, estrous in females occurs only every 4–5 years, after a two-year gestation and a long weaning period.
2. Frontalin produces different responses in males and females depending on age and physiological condition. At the beginning, male elephants produce a TGS containing floral smelling compounds, but no frontalin, which is only synthesised with the onset of maturity and increases its levels many folds as the animal gets older<sup>11,14</sup>.
3. Frontalin does not produce significant behaviour in old males nor in females in the luteal phase. Instead, young males and pregnant females get apprehensive and try to avoid this pheromone. On the contrary, females in the follicular phase show attraction to this chemical<sup>8</sup>.
4. Elephants can discriminate between the two enantiomers of frontalin. It has also been observed that females respond much stronger to racemic frontalin than to different mixtures of the two enantiomers. In fact, the racemate is a sign of maturity, whereas the (+)-enantiomer predominates in young males, and the (–)-enantiomer is the only one produced by females<sup>10,12</sup>.

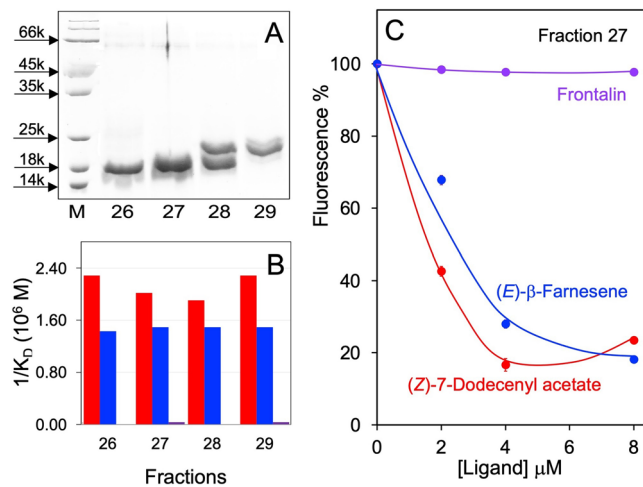
In this work, we have investigated the role of soluble odorant-binding proteins (OBPs) in sexual chemical communication of two elephant species, the Asian and the African elephants, and we also have measured their affinities to the pheromone components. In mammals, such proteins mediate chemical communication between sexes by a dual mechanism; in the nose, they carry pheromones to membrane-bound receptors, while in specialised glands and secretions, they solubilise the same pheromones for their delivery in the environment<sup>15</sup>. OBPs are small soluble polypeptides present in large amounts in the nasal mucus<sup>16–18</sup>, as well as in body fluids such as urine, saliva, seminal fluid and other external secretions<sup>19,20</sup>. They bind species-specific pheromones with dissociation constants in the upper nanomolar/lower micromolar range. Although tuned to their target pheromones, their specificity is rather broad, exhibiting weaker binding activity to some structurally related compounds. Recent work has shown that post-translational modifications (PTMs), such as glycosylation and phosphorylation play important roles in tuning the specificity of OBPs to their target ligands by increasing affinity, narrowing specificity and modulating the binding spectrum<sup>21</sup>.

An OBP was detected in the trunk wash of the Asian elephant and shown to bind (*E*)-7-dodecenyl acetate<sup>22</sup>. The adopted assay involved separation of a crude protein extract of the trunk wash on an electrophoretic gel after incubation with tritium-labelled (*E*)-7-dodecenyl acetate. Radioactivity associated with the OBP band indicated specific binding, but did not allow evaluation of the dissociation constant.

Here, we report on the proteomic analysis of the trunk wash from the African elephant, the expression of two orthologue OBPs from the Asian and the African elephants in the yeast *Pichia pastoris*, and the characterization of their binding activity.

## Results and discussion

**Sequence analysis.** The OBPs of elephants are not completely annotated. A search in different sequence databases (NCBI and UniProtKB) returned two orthologous sequences for the African and the Asian elephant, sharing 93% of their residues, which we here indicate as LafrOBP1 and EmaxOBP1 (Fig. 1). In the African



**Figure 2.** (A) Fractionation of a crude extract of trunk wash from the African elephant by anion-exchange chromatography. SDS-PAGE analysis of the fractions showing electrophoretic bands in the low molecular weight range (15–25 kDa). (B) Summary of the affinities to (Z)-7-dodecenyl acetate, (E)-β-farnesene and frontaline measured with crude samples of the same fractions. All of them bound with good affinity both (Z)-7-dodecenyl acetate and (E)-β-farnesene. Also farnesol, proved to be as good a ligand as (E)-β-farnesene, whereas *exo*-brevicomin did not show any affinity. Dissociation constants were evaluated using the procedure adopted when dealing with solutions of a pure protein; as this is not the case, the values have to be taken as a crude indication of the affinity of the ligands for the protein. (C) Examples of the competition curves measured with the three ligands on fraction 27.

elephant, two additional sequences are annotated. The first, which here we have named LafrOBP2, presents only 15% identity with LafrOBP1. The second (LafrVEG), 13% identical to LafrOBP1, bears a significant similarity (30–40% identity) to a group of lipocalins known as VEG proteins, tear prealbumins or lipocalin-1 (LCN1). A function of the latter polypeptides as carriers for taste compounds had been suggested at the beginning<sup>23</sup> but this idea was not strongly supported by experimental evidence and alternative roles were proposed, such as scavenger for toxic chemicals and transporter for vitamins and nutrients<sup>24</sup>, without however disproving the original hypothesis. The sequences of these four lipocalins are aligned in Fig. 1. In addition, some isoforms of both LafrOBP1 and LafrOBP2 were found in the databases, but their sequences were not fully reliable, being incomplete or too long at their C-terminus.

**Proteomic analysis of whole trunk wash.** We first decided to analyse a sample of the trunk wash from the African elephant to identify the proteins secreted, with particular attention to OBPs. Therefore, the sample was filtered, concentrated and separated on an electrophoretic gel in denaturing conditions (SDS-PAGE) (Figure S1). Abundant bands were observed in the low molecular weight region (20–25 kDa), which represented most of the proteins in the trunk wash. Slices were cut from the gel, reduced, alkylated with iodoacetamide, digested with trypsin and subjected to nanoLC-ESI-Q-Orbitrap-MS/MS analysis. We identified three abundant lipocalins corresponding to the sequences XP\_023395442.1, XP\_023395433.1 and XP\_023398512.1. The first one coincides with LafrOBP1 annotated in the database, the second is an isoform of LafrOBP1 bearing the amino acid substitutions Leu13Ser, Ile32Leu, Glu78Lys, Trp80Gly, Glu116Lys, and additional sequence portions at both N- and C-terminal region while the third one corresponds to LafrVEG mentioned above. We could not detect the presence of LafrOBP2. A summary of the results of this proteomic analysis is reported in supplementary Table S1.

**Fractionation of soluble proteins from trunk wash and analysis of PTMs in LafrOBP1.** Given the predominant presence of LafrOBP1 in the trunk wash of the African elephant, we attempted purification of this protein from a crude sample of this secretion by anion-exchange chromatography. Relevant fractions (#26–29) were analysed by SDS-PAGE, (Fig. 2A) and nanoLC-ESI-Q-Orbitrap-MS/MS to identify the proteins present and detect possible PTMs. The same fractions were also tested in ligand-binding assays with some of the two elephants' pheromones and other potential semiochemicals to detect any binding activity (Fig. 2B–C). All fractions bound (Z)-7-dodecenyl acetate with low dissociation constants (0.44–0.52 μM) and showed a good affinity to (E)-β-farnesene (K<sub>D</sub>: 0.67–0.70 μM). On the contrary, none of them appreciably bound frontaline. We also tested farnesol, which proved to be as good a ligand as (E)-β-farnesene, and *exo*-brevicomin, which did not exhibit any affinity (data not shown).

Proteomic analysis of fractions 26–29 demonstrated the concomitant occurrence in each fraction of LafrOBP1, its isoform and LafrVEG, highlighting the difficulty to resolve these proteins based on anion-exchange chromatography. An independent bioinformatic search of mass spectrometric data including various putative PTMs allowed identifying modified peptides from native LafrOBP1, including acetylated, phosphorylated and O-glycosylated ones (Table S2). Abundant non-modified corresponding peptides were also identified in each

fraction. Exemplificative tandem mass spectra of modified peptides are reported in Fig. 3. As expected, slight differences were observed in the representation of modified peptides in the different chromatographic fractions (Table S2). In the whole, acetylated peptides (60–71), (62–75), (76–98), (88–98), (124–146), (124–151) and (147–163), and phosphorylated components (44–61) and (152–163) were observed in the digests of native LafrOBP1. Fragmentation spectra of above-mentioned species definitively assigned acetylation at Lys61, Lys71, Lys88, Lys125, Lys146 and Lys151, as well as phosphorylation at Tyr59 and Thr157 (Table S2 and Figure S2). Fractions 27–29 also showed the occurrence of the O-glycosylated (124–146) peptide bearing a tetra-*N*-acetyl-lactosamine-containing core-2 O-glycan<sup>25</sup> (Table S2), which was linked to Ser128, according to its tandem mass spectrum (Fig. 3).

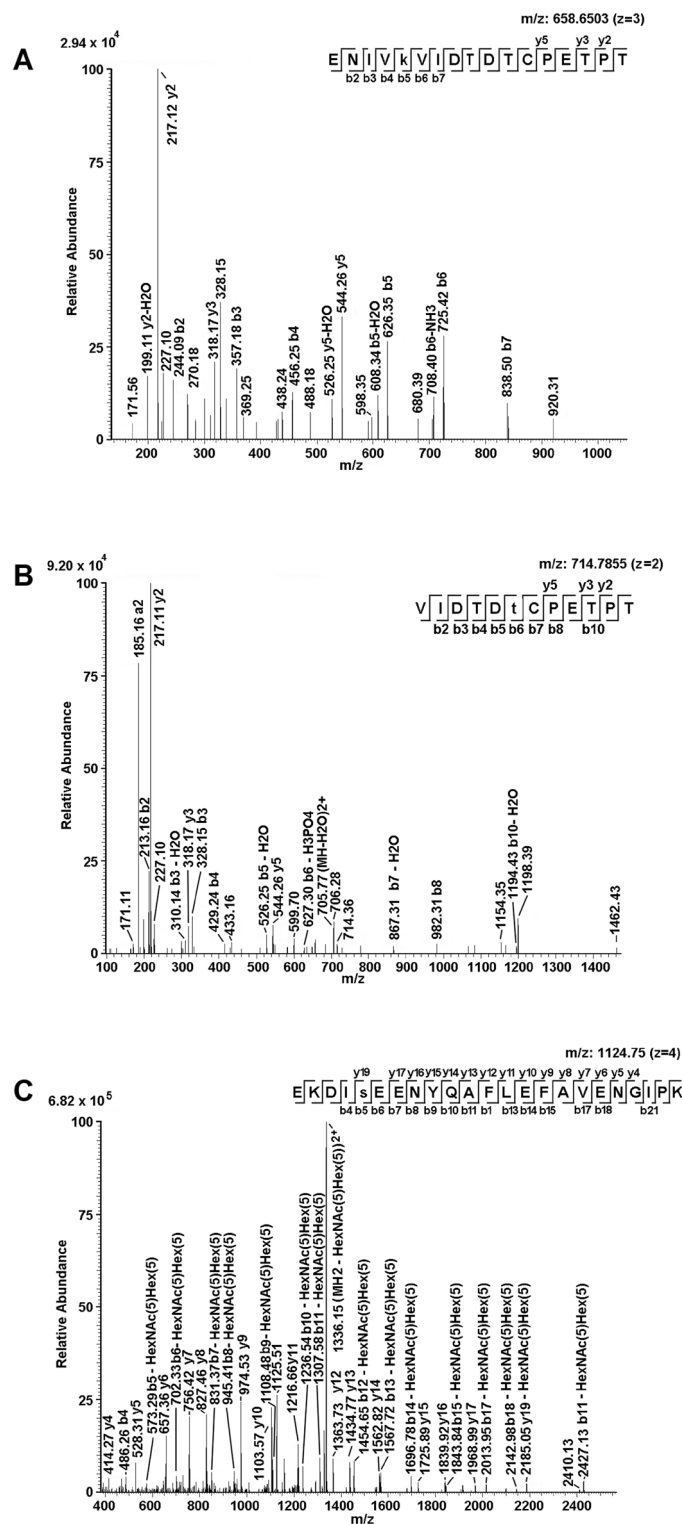
**Expression and purification of LafrOBP1 and EmaxOBP1.** Based on above-reported proteomic analysis and on the available sequences in the protein databases, we decided to focus our functional study on the OBP1 of both species. Thus, LafrOBP1 and EmaxOBP1 were expressed in the yeast *Pichia pastoris* using the protocol reported in the Materials and Methods section. We chose to perform functional studies on recombinant proteins because attempts to purify natural LafrOBP1 from its contaminants were unsuccessful (data not shown). The recombinant proteins were secreted in the culture medium, concentrated by ultrafiltration and purified by anion-exchange chromatography, followed by a step of hydrophobic interaction chromatography on Phenyl Sepharose. Figure S3 reports the SDS-PAGE analysis relative to the expression and purification of the two proteins. The expression of EmaxOBP1 produced a single major band migrating with an apparent molecular weight (around 20 kDa), compatible with the calculated mass for this protein (18,506.7 Da). Instead, the expression of the African elephant orthologue afforded two major bands, the first migrating at the same level as EmaxOBP1, the second with an apparent mass around 25 kDa. This second band was also purified, but, unlike the lower band, did not show any affinity to the fluorescent probe *N*-phenyl-1-naphthylamine (1-NPN), therefore, it was not considered further in our study.

**Structure analysis and PTMs of recombinant LafrOBP1.** To identify PTMs in the recombinant LafrOBP1 and compare them with those detected in the natural protein, the first was also subjected to proteomic analysis. We found that most of the modifications identified in the natural LafrOBP1 were conserved in the recombinant product. In fact, acetylated species (60–71), (62–75), (124–146) and (124–151), phosphorylated peptides (44–61), (62–75) and (152–163), as well as the O-glycosylated (124–146) component bearing the tetra-*N*-acetyl-lactosamine-containing core-2 were also detected in the digest of recombinant LafrOBP1 (Table S2 and Figure S2). Abundant non-modified corresponding peptides were also found in recombinant LafrOBP1. Fragmentation spectra of above-mentioned modified peptides further confirmed acetylation at Lys61, Lys71, Lys125 and Lys146, phosphorylation at Tyr59, Tyr69 and Thr157, and O-glycosylation at Ser128 (Table S2 and Figure S2). The slight modification differences observed with respect to the natural protein (Figure S2) were rationalized based on the distinct modification machineries present in the African elephant and *P. pastoris*.

LafrOBP1 and EmaxOBP1 contain four cysteines occurring at conserved positions in two elephant species (Fig. 1), but differing in number and distribution from what is observed in other mammalian OBPs<sup>26</sup>. To ascertain their oxidation state and their potential pairing in disulfide bonds, recombinant LafrOBP1 was subjected to extensive alkylation with iodoacetamide under non-reducing denaturing conditions; then, it was digested first with trypsin and then by chymotrypsin. Resulting peptides were analyzed by nanoLC-ESI-Q-Orbitrap-MS/MS, and peptides linked by disulfide bridges were identified by dedicated bioinformatic procedures<sup>27–29</sup>. Detection of peptides (12–19)S–S(43–55) and (60–69)S–S(152–163) (Fig. 4A) demonstrated that all cysteines present in LafrOBP1 are involved in disulfide bridges in the pattern C1–C2 and C3–C4. This disulfide motif involves three conserved cysteines present in some homologous proteins from other mammals; the fourth is peculiar to elephant OBP1.

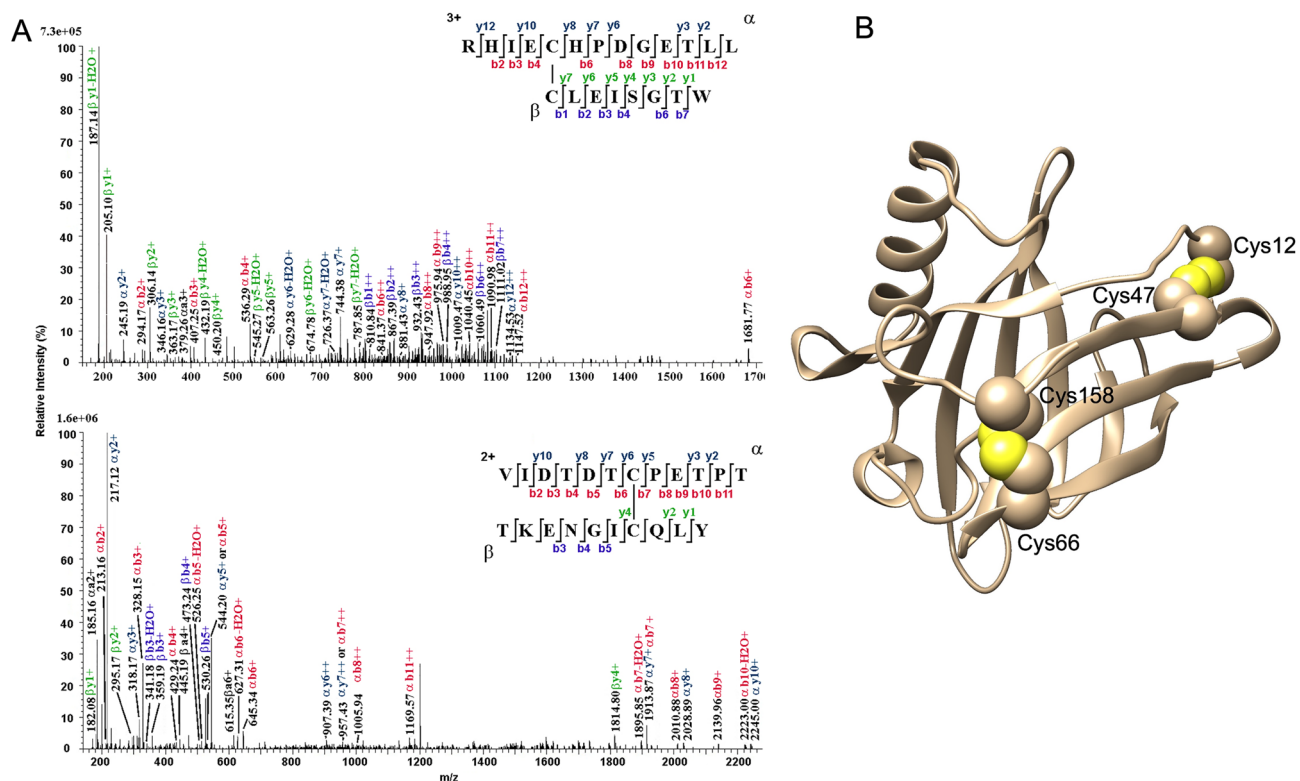
A structural model of LafrOBP1 was generated using the on-line programme Swiss Model<sup>30–32</sup> and the allergen Bosd2 (PDB: 4WFU.1.A) as a template, whose amino acid sequence is 31% identical to LafrOBP1. Apart from the good sequence similarity between the two proteins, the reliability of this model was also supported by the experimental evidence for the presence of the two disulfide bridges in LafrOBP1. Figure 4B reports the model of LafrOBP1 showing the four cysteines involved in disulfide bridges. As expected, various hydrophobic amino acids occurred in the protein binding pocket (data not shown), in agreement with what was observed in structurally related OBP 1F from the brown rat *Rattus norvegicus* (PDB: 3FIQ), aphrodisin from the golden hamster *Mesocricetus auratus* (PDB: 1E5P), allergen lipocalin Cav p1 from the guinea pig *Cavia porcellus* (PDB: 8A0D) and lipocalin allergen Bos d2 from the cattle *Bos taurus* (PDB: 1BJ7).

**Ligand-binding assays.** Recombinant LafrOBP1 and EmaxOBP1 were then tested in ligand-binding assays using the Asian elephant pheromones (*Z*)-7-dodecenyl acetate and frontalinal, and the insect pheromone components found in elephants, *exo*-brevicomine and (*E*)- $\beta$ -farnesene, as well as structurally related compounds, and chemicals of different size and structure, which could contribute to estimate the specificity of these proteins. The results are reported in Fig. 5. Both LafrOBP1 and EmaxOBP1 bound the fluorescent probe 1-NPN with good affinity (Fig. 5A), thus enabling the use of the fluorescent displacement assay to evaluate the affinities of other ligands. Some representative curves of competitive binding experiments for LafrOBP1 are shown in Fig. 5B–C, while in the histogram of Fig. 5D the binding spectra of LafrOBP1 and EmaxOBP1 are compared. The specificity of LafrOBP1 can be better appreciated from the graph of Fig. 6, where the affinities of esters to the protein are plotted as a function of the ligand's size. In addition to the ligands reported in Fig. 5, the following chemicals did not show significant binding to either of the two proteins: (*Z*)-3-hexenyl acetate, (*Z*)-11-hexadecenyl



**Figure 3.** Examples of post-translationally modified peptides assigned to LafrOBP1 from trunk wash of the African elephant. An acetylated peptide, a phosphorylated peptide and the O-glycosylated peptide are shown. Identical peptides and corresponding fragmentation spectra were observed also in the recombinant protein expressed in the yeast *Pichia pastoris* (see Table S2). Panel A shows the fragmentation spectrum of the acetylated peptide (147–163) bearing modification at Lys151. Panel B shows the fragmentation spectrum of the phosphorylated peptide (152–163) bearing modification at Thr157. Panel C shows the fragmentation spectrum of the O-glycosylated peptide (124–146) bearing a tetra-*N*-acetyl-lactosamine-containing core-2 O-glycan linked to Ser128.





**Figure 4.** (A) Fragmentation spectra of disulfide-bridged peptides identified in the tryptic-chymotryptic digest of recombinant LafrOBP1. The upper and the lower panels show independent data relative to Cys12-Cys47 and Cys66-Cys158 disulfide bridges. The fragment ions are reported in different color depending on the peptide present in S-S-linked species, and corresponding b and y ion series. (B) Three-dimensional model of LafrOBP1 showing the pairing of the four cysteines. The model was generated with the on-line software Swiss Model using the structure of the allergen Bosd2 (PDB: 4WFU.1.A) as a template.

acetate, ethyl myristate,  $\gamma$ -tridecalactone, octanal, decanal, dodecanal, (Z)-9-hexadecenal, (Z)-11-hexadecenal, (Z)-13-octadecenal, eucalyptol, borneol and eugenol.

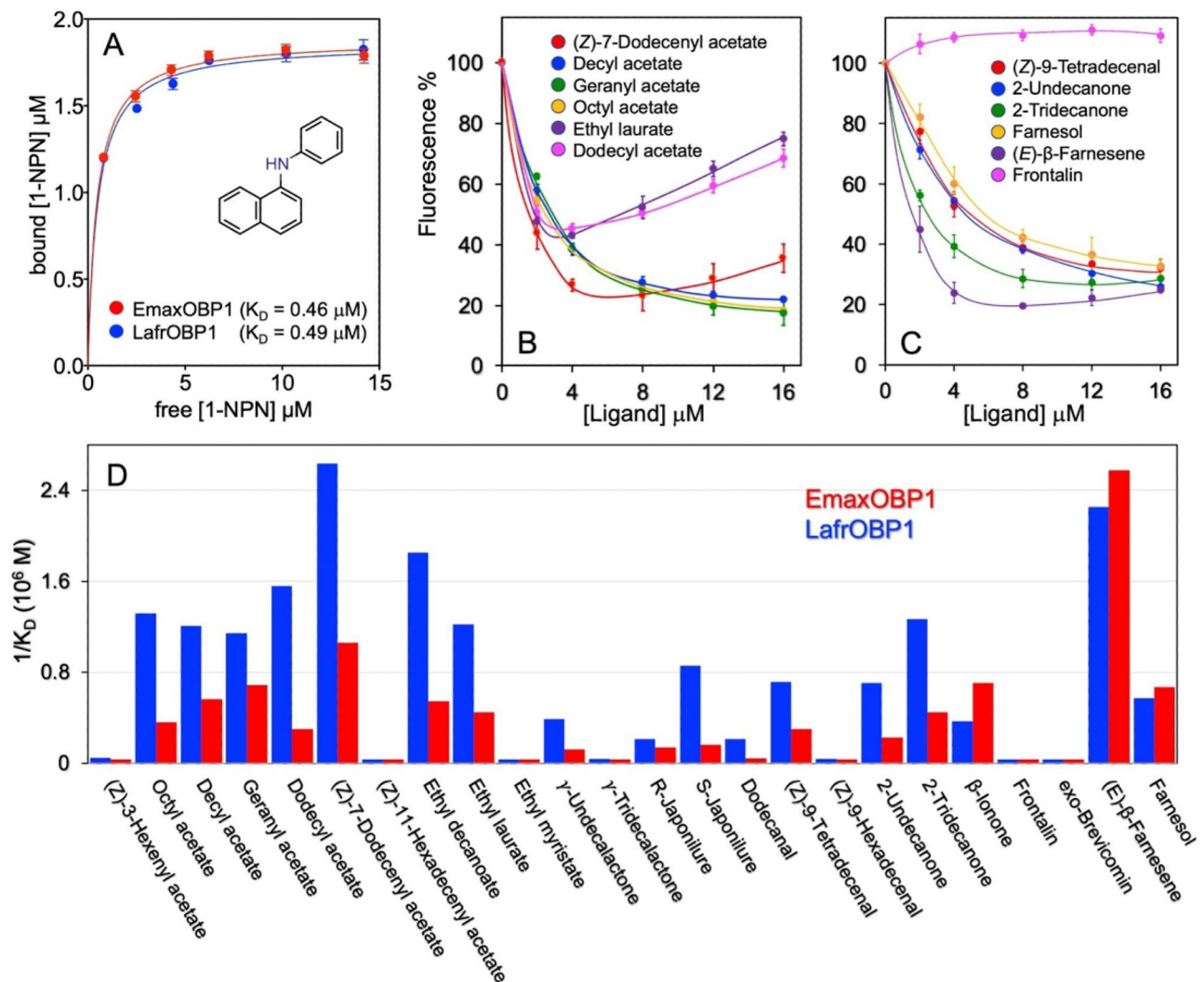
To evaluate to what extent phosphorylation and glycosylation could account for the binding activity of LafrOBP1, fractions 27 and 29 from trunk wash and a sample of the protein expressed in *P. pastoris* were treated with Lambda phosphatase and O-glycosidase. In no case did we observe any shift in the apparent molecular weight of the protein. Figure S4A shows the results relative to treatment of the recombinant protein. We also found that after digestion with phosphatase the binding of (Z)-7-dodecanyl acetate to recombinant LafrOBP1 was not affected (Figure S4B).

In this work we have shown that:

1. Two OBPs, OBP1 and OBP2, are encoded in the genome of the elephants apart from VEG protein, whose functional similarity with OBPs has been suggested, but is not supported by strong evidence. In addition, some isoforms of both OBPs can be identified, but their sequences are not fully reliable.
2. The trunk wash of the African elephant contains OBP1, one of its isoforms and the VEG protein in significant concentrations.
3. The OBP1 isolated from the trunk wash of the African elephant occurs in several fractions separated by anion-exchange chromatography, all equally binding the Asian elephant sex pheromone (Z)-7-dodecanyl acetate, as well as  $\beta$ -farnesene and farnesol, but do not show any affinity to the musth pheromone frontalinal or to *exo*-brevicommin.
4. Post-translational modifications, such as O-glycosylation, acetylation and phosphorylation, have been detected in both the protein extracted from the trunk wash and in the yeast-expressed LafrOBP1. However, as such modifications affect only a small fraction of the protein population, their effects on ligand-binding properties cannot be established.

## Materials and methods

**Materials.** Trunk wash was collected from one male (Tembo, born 1985) and five female (Tonga, 1984; Numbi, 1992; Mongu, 2003; Iqhw, 2013; Kibali, 2019) African elephants at the Vienna Zoo during routine procedures. Briefly, 100 mL of a sterile 0.9% saline solution is injected in each nostril of the trunk, which is kept in a lifted position, so that the solution is running up to the base of the trunk. The mixture of the solution and

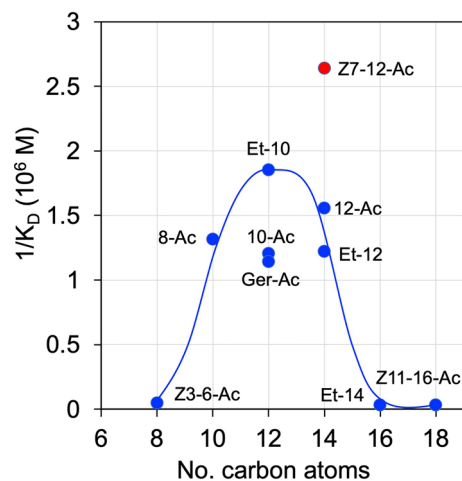


**Figure 5.** Binding of recombinant LafrOBP1 and EmaxOBP1 to the elephants' pheromones and structurally related compounds. **(A)** Both proteins bound the fluorescent probe N-phenyl-1-naphthylamine (1-NPN), whose structure is reported as an inset, with strong and very similar affinities. **(B)** Binding of the pheromone (Z)-7-dodecyl acetate and few structurally related esters to recombinant LafrOBP1. The increase of the fluorescence at higher concentrations of ligand is due to formation of micelles by some amphiphilic ligands, which could entrap molecules of 1-NPN<sup>35–35</sup>. **(C)** Binding of representative compounds among the other tested chemicals. Interestingly, (E)-β-farnesene was found to be a good ligand, despite its structural difference with the Asian elephant pheromone (Z)-7-dodecyl acetate. Instead, frontalin, a second Asian elephant pheromone, did not exhibit any affinity for the protein. **(D)** Graphical comparison of the affinities of selected ligands to the OBPs of both elephant species. No major differences were observed in the binding spectra between LafrOBP1 and EmaxOBP1. In addition to the reported ligands, the following compounds did not exhibit significant affinity for neither of the two proteins: hexanal, octanal, decanal, (Z)-11-hexadecenal, (Z)-13-octadecenal, eucaliptol, *p*-cresol, eugenol.

trunk mucus is collected in sterile plastic bags by active blowing of the elephant. Chemicals were all from Merck, Austria, unless otherwise stated. Restriction enzymes and kits for DNA extraction and purification were from New England Biolabs, USA. Oligonucleotides and synthetic genes were custom synthesised at Eurofins Genomics, Germany.

**Ethics declaration.** We confirm that the trunk wash performed to provide a sample of the mucus was carried out as a routine procedure to monitor the health of elephants at the Vienna Zoo and in accordance with relevant guidelines and regulations.

**Trunk wash fractionation.** Trunk wash was centrifuged for 1 h at 10,000 g, the supernatant was dialyzed against 50 mM Tris-HCl buffer, pH 7.4 and concentrated by ultrafiltration in the Amicon stirred cell, then frac-



**Figure 6.** Affinities of some esters to LafrOBP1 plotted as a function of the number of carbon atoms. The pheromone (*Z*)-7-dodecenyl acetate was the strongest ligand, while some other esters of 10–14 carbon atoms also bound with lower affinities. Homologous esters of less than 10 or more than 14 carbon atoms did not show significant binding for LafrOBP1. Z3-6-Ac: (*Z*)-3-hexenyl acetate; 8-Ac: 1-octyl acetate; 10-Ac: 1-decyl acetate; Ger-Ac: geranyl acetate; Et-10: ethyl decanoate; Z7-12-Ac: (*Z*)-7-dodecenyl acetate; 12-Ac: 1-dodecyl acetate; Et-12: ethyl dodecanoate; Et-14: ethyl tetradecanoate; Z11-16-Ac: (*Z*)-11-hexadecenyl acetate.

tionated by anion-exchange chromatography on HiPrep-Q 16/10 column, 20 mL (Bio-Rad), along with standard protocols.

**Protein alkylation and digestion, and mass spectrometry analysis.** SDS-PAGE gel portions of proteins from whole elephant trunk wash (for component identification), chromatographic fractions of the elephant trunk wash (for PTMs analysis) or SDS-PAGE gel bands of LafrOBP1 expressed in *P. pastoris* were in parallel tritigated, washed with water, *in gel*-reduced, S-alkylated, and digested with trypsin (Sigma, sequencing grade). Resulting peptide mixtures were desalted by  $\mu$ Zip-TipC18 (Millipore) using 50% (v/v) acetonitrile, 5% (v/v) formic acid as eluent, vacuum-dried by SpeedVac (Thermo Fisher Scientific, USA), and then dissolved in 20  $\mu$ L of aqueous 0.1% (v/v) formic acid for subsequent MS analyses by means of a nanoLC-ESI-Q-Orbitrap-MS/MS system, comprising an UltiMate 3000 HPLC RSLC nano-chromatographer (Thermo Fisher Scientific) interfaced with a Q-ExactivePlus mass spectrometer (Thermo Fisher Scientific) mounting a nano-Spray ion source (Thermo Fisher Scientific). Chromatographic separations were obtained on an Acclaim PepMap RSLC C18 column (150 mm  $\times$  75  $\mu$ m ID; 2  $\mu$ m particle size; 100  $\text{\AA}$  pore size, Thermo Fisher Scientific), eluting the peptide mixtures with a gradient of solvent B (19.92/80/0.08 v/v/v water/acetonitrile/formic acid) in solvent A (99.9/0.1 v/v water/formic acid), at a flow rate of 300 nL/min. In particular, solvent B started at 3%, increased linearly to 40% in 45 min, then achieved 80% in 5 min, remaining at this percentage for 4 min, and finally returned to 3% in 1 min. The mass spectrometer operated in data-dependent mode in positive polarity, carrying out a full MS1 scan in the range  $m/z$  345–1350, at a nominal resolution of 70,000, followed by MS/MS scans of the 10 most abundant ions in high energy collisional dissociation (HCD) mode. Tandem mass spectra were acquired in a dynamic  $m/z$  range, with a nominal resolution of 17,500, a normalized collision energy of 28%, an automatic gain control target of 50,000, a maximum ion injection time of 110 ms, and an isolation window of 1.2  $m/z$ . Dynamic exclusion was set to 20 s<sup>36</sup>.

**Bioinformatics for peptide identification and post-translational modification assignment.** Raw mass data files were searched by Proteome Discoverer v. 2.4 package (Thermo Fisher Scientific), running the search engine Mascot v. 2.6.1 (Matrix Science, UK), Byonic™ v. 2.6.46 (Protein Metrics, USA) and Peaks Studio 8.0 (BSI, Waterloo, Ontario, Canada) software, both for peptide assignment/protein identification and for post-translational modification analysis.

In the first case, analyses were carried out against a customized database containing protein sequences downloaded from NCBI (<https://www.ncbi.nlm.nih.gov/>) for superorder Afrotheria (consisting of 192,838 protein sequences, December 2021) plus the most common protein contaminants and trypsin. Parameters for database searching were fixed carbamidomethylation at Cys, and variable oxidation at Met, deamidation at Asn/Gln, and pyroglutamate formation at Gln. Mass tolerance was set to  $\pm 10$  ppm for precursors and to  $\pm 0.05$  Da for MS/MS fragments. Proteolytic enzyme and maximum number of missed cleavages were set to trypsin and 3, respectively. All other parameters were kept at default values. In the latter case, raw mass data were analyzed against a customized database containing LafrOBP1 (XP\_023395442.1) protein sequence plus the most common protein contaminants and trypsin, allowing to search Lys-acetylation ( $\Delta m = +42.01$ ), Ser/Thr/Tyr-phosphorylation ( $\Delta m = +79.97$ ), and the most common mammals N-linked glycans at Asn and O-linked glycans at Ser/Thr/Tyr, using the same parameters previously set. The max PTM sites per peptide was set to 2.



Proteome Discoverer peptide candidates were considered confidently identified only when the following criteria were satisfied: (i) protein and peptide false discovery rate (FDR) confidence: high; (ii) peptide Mascot score: > 30; (iii) peptide spectrum matches (PSMs): unambiguous; (iv) peptide rank (rank of the peptide match): 1; (v) Delta CN (normalized score difference between the selected PSM and the highest-scoring PSM for that spectrum): 0. Byonic peptide candidates were considered confidently identified only when the following criteria were satisfied: (i) PEP 2D and PEP 1D: <  $1.0 \times 10^{-3}$ ; (ii) FDR: 0; (iii) q-value 2D and q-value 1D: <  $1.0 \times 10^{-5}$ . An FDR value of 1% was specified as the cut-off of false discovery rate for peptides identification by Peaks Studio. Manual interpretation and verification of the candidate spectra were always performed.

**Yeast expression and purification of proteins.** The gene encoding OBP1 was cloned between XhoI and NotI restriction sites into the pPIC9 vector downstream of the  $\alpha$ -factor secretion signal peptide. The following set of primers was used: 5'-AACTCGAG AAA AGA ATG CTG GAA GAA CC-3'; 5'-AAGCGGCCGC TTA GGT CAG TGT TTC CGG-3' both for EmaxOBP1 and LafrOBP1. The construct was linearized with BglII and used to transform electrocompetent yeast cells (strain GS115) by electroporation (1.5 kV, 25  $\mu$ F). Transformed cells were cultured on minimal dextrose (MD) agar plates (1.5% agar, 1.34% w/v yeast nitrogen base (YNB), 2% glucose,  $4 \times 10^{-5}$ % biotin) for 48 h at 30 °C. Then, colonies were transferred onto minimal methanol (MM) agar plates (1.5% agar, 1.34% w/v yeast nitrogen base (YNB), 0.5% methanol,  $4 \times 10^{-5}$ % biotin) and incubated at 30 °C for further 48 h. Several colonies were selected from MM plates and screened for protein expression. First, colonies were grown in 10 mL of buffered minimal glycerol (BMGY) medium (1% w/v yeast extract, 2% w/v peptone, 1.34% w/v yeast nitrogen base with ammonium sulfate without amino acids (YNB),  $4 \times 10^{-5}$ % biotin, 100 mM potassium phosphate, pH 6.0, 1% v/v glycerol) at 29 °C, shaking at 240 rpm for 24 h. To induce protein expression, cells were harvested by centrifugation (15 min, 3,000 g, room temperature), re-suspended in 50 mL of buffered minimal methanol (BMM) medium (1.34% w/v YNB,  $4 \times 10^{-5}$ % biotin, 100 mM potassium phosphate, pH 6.0, 1% v/v methanol) and incubated at 30 °C for 72 h. To maintain protein expression, every 24 h methanol was added in amounts of 1.5% v/v. At the same intervals, samples of the supernatants were analyzed on SDS-PAGE for protein expression. On the basis of the results obtained, large-scale expression was performed under the above conditions for 72 h. The supernatant containing the recombinant OBP was centrifuged for 1 h at 10,000 g, dialyzed against 50 mM Tris-HCl buffer, pH 7.4, and concentrated by ultrafiltration in the Amicon stirred cell. Purification was performed by anion-exchange chromatography on HiPrep-Q 16/10, 20 mL (Bio-Rad) column, followed by elution from a Phenyl Sepharose column with the linear gradient of 0.6–0 M ammonium sulfate in 50 mM Tris-HCl, pH 7.4.

**Disulfide assignment.** A sample of recombinant LafrOBP1 (20  $\mu$ g) was dissolved in 0.1 M tetraethylammonium bicarbonate (TEAB), pH 6.5, containing 4 M guanidinium chloride, and then treated with iodoacetamide (0.5 M final concentration) for 30 min, in the dark. The protein was precipitated by the addition of 6 vol of cold acetone at -20 °C, overnight. After centrifugation at 12,000 rpm at 4 °C, for 20 min, supernatant was removed, and the recovered pellet was dried. Recovered protein was dissolved in 0.05 M TEAB, pH 6.5 (2  $\mu$ g/ $\mu$ L final concentration), treated with trypsin (1:30 w/w enzyme/substrate) for 16 h, at 37 °C, and then with chymotrypsin (1:30 w/w enzyme/substrate) for 16 h, at 37 °C. Protein digest was desalted with ZipTip C18 (Millipore, USA) and directly analyzed with a UltiMate 3000 HPLC RSLC nano-chromatographer (Thermo Fisher Scientific) linked to a Q-ExactivePlus mass spectrometer (Thermo Fisher Scientific), as reported above.

Dedicated BioPharma Finder v. 4.0 (Thermo Fisher Scientific) and pLink v. 2.3.9 software were used for database searching of mass spectrometric data, enabling the specific function of disulfide-linked peptides attribution, and applying the additional settings reported above for Proteome Discoverer, Byonic and Peaks Studio examinations. A confidence score > 95 for BioPharma Finder and/or an E-value <  $1.0^{-10}$  for pLink results were considered for a reliable identification of disulfide-bridged peptides. Candidate spectra were always verified by manual interpretation.

**Protein Model.** A three-dimensional model of LafrOBP1 was obtained with the Swiss Model software<sup>30–32</sup>, using the allergen Bosd2 structure (PDB: 4WFU.1.A) as a template. The figure was generated with Chimera software<sup>37</sup>.

**Ligand-binding assays.** Affinities of the fluorescent probe N-phenyl-1-naphthylamine (1-NPN) to EmaxOBP1 and LafrOBP1 (both expressed in yeast) were evaluated by titrating a 2  $\mu$ M solution of each protein in 50 mM Tris-HCl, pH 7.4 with 1 mM 1-NPN in methanol to final concentrations of 2–16  $\mu$ M. Dissociation constants of other ligands were measured by displacement of the fluorescent reporter 1-NPN from the complex. Accordingly, solutions of protein and 1-NPN, both at the concentration of 2  $\mu$ M in 50 mM Tris-HCl, pH 7.4, were titrated with 1 mM methanol solutions of each ligand to final concentrations of 2–16  $\mu$ M. Emission spectra were recorded on a PerkinElmer FL 6500 spectrofluorimeter in a right-angle configuration, at room temperature, with slits of 5 nm for both excitation and emission, using 1 cm path quartz cuvettes. The excitation wavelength was set at 337 nm and intensities were recorded in correspondence with the peak maximum, around 390 nm. Data for 1-NPN were processed with Prism software.

Dissociation constants of other ligands were calculated from the corresponding  $[IC]_{50}$  values (the concentration of each ligand halving the initial value of fluorescence), using the equation:  $K_d = [IC]_{50}/1 + [1-NPN]/K_{NPN}$ , where  $[1-NPN]$  is the concentration of free 1-NPN and  $K_{NPN}$  the dissociation constant of the complex OBP/1-NPN.

**Digestion with phosphatase and glycosidase.** Fractions 27 and 29 from trunk wash and a sample of LafrOBP1 expressed in *P. pastoris* were treated in parallel with Lambda phosphatase, O-glycosidase and PNGase along with the protocols provided by the manufacturers. The digestion products were analysed on SDS-PAGE. The product after digestion of recombinant LafrOBP1 with phosphatase was assayed for binding activity to (Z)-7-dodecyl acetate in competitive fluorescent binding experiments, as reported above.

### Data availability

The datasets generated and/or analysed during the current study are available as follows: the mass spectrometry-based proteomic data have been deposited to the ProteomeXchange Consortium via the PRIDE<sup>38</sup> partner repository with the dataset identifier PXD036751. All other data are reported in the manuscript files and in Supplementary information.

Received: 24 September 2022; Accepted: 11 November 2022

Published online: 21 November 2022

### References

- Roca, A. L., Georgiadis, N., Pecon-Slattery, J. & O'Brien, S. J. Genetic evidence for two species of elephant in Africa. *Science* <https://doi.org/10.1126/science.1059936> (2001).
- Schulte, B. A. & LaDue, C. A. The chemical ecology of elephants: 21st century additions to our understanding and future outlooks. *Animals* **11**, 2860 (2021).
- Rasmussen, L. E. L. Chemosensory responses in two species of elephants to constituents of temporal gland secretion and musth urine. *J. Chem. Ecol.* **14**, 1687–1711 (1988).
- Rasmussen, L. E. L., Lee, T. D., Daves, G. D. & Schmidt, M. J. Female-to-male sex pheromones of low volatility in the Asian elephant, *Elephas maximus*. *J. Chem. Ecol.* **19**, 2115–2128 (1993).
- Rasmussen, L. E. L., Lee, T. D., Zhang, A., Roelofs, W. L. & Daves, G. D. Jr. Purification, identification, concentration and bioactivity of (Z)-7-dodecyl-1-yl acetate: Sex pheromone of the female Asian elephant *elephas maximus*. *Chem. Senses* **22**, 417–437 (1997).
- Rasmussen, L. E. L. & Schulte, B. A. Chemical signals in the reproduction of Asian (*Elephas maximus*) and African (*Loxodonta africana*) elephants. *Anim. Reprod. Sci.* **53**, 19–34 (1998).
- Goodwin, T. E. *et al.* African elephant sesquiterpenes. *J. Nat. Prod.* **62**, 1570–1572 (1999).
- Rasmussen, L. E. L. & Greenwood, D. R. Frontalin: A chemical message of musth in Asian elephants (*Elephas maximus*). *Chem. Senses* **28**, 433–446 (2003).
- Goodwin, T. E. *et al.* Insect pheromones and precursors in female African elephant urine. *J. Chem. Ecol.* **32**, 1849–1853 (2006).
- BetsRasmussen, L. E. L., Greenwood, D. R., Goodwin, T. E. & Schulte, B. A. Asian Elephant Reflections: Chirality Counts. In *Chemical Signals in Vertebrates 13* (eds Schulte, B. A. *et al.*) 229–244 (Springer International Publishing, 2016).
- Greenwood, D. R., Comeskey, D., Hunt, M. B. & Rasmussen, L. E. L. Chirality in elephant pheromones. *Nature* **438**, 1097–1098 (2005).
- LaDue, C. A., Goodwin, T. E. & Schulte, B. A. Concentration-dependent chemosensory responses towards pheromones are influenced by receiver attributes in Asian elephants. *Ethology* **124**, 387–399 (2018).
- Bagley, K. R., Goodwin, T. E., Rasmussen, L. E. L. & Schulte, B. A. Male African elephants, *Loxodonta africana*, can distinguish oestrous status via urinary signals. *Anim. Behav.* **71**, 1439–1445 (2006).
- Rasmussen, L. E. L., Riddle, H. S. & Krishnamurthy, V. Mellifluous matures to malodorous in musth. *Nature* **415**, 975–976 (2002).
- Pelosi, P. & Knoll, W. Odorant-binding proteins of mammals. *Biol. Rev.* **97**, 20–44 (2022).
- Pelosi, P., Baldaccini, N. E. & Pisanelli, A. M. Identification of a specific olfactory receptor for 2-isobutyl-3-methoxypyrazine. *Biochem. J.* **201**, 245–248 (1982).
- Bignetti, E. *et al.* Purification and characterisation of an odorant-binding protein from cow nasal tissue. *Eur. J. Biochem.* **149**, 227–231 (1985).
- Pevsner, J., Trifiletti, R. R., Strittmatter, S. M. & Snyder, S. H. Isolation and characterization of an olfactory receptor protein for odorant pyrazines. *Proc. Natl. Acad. Sci. USA* **82**, 3050–3054 (1985).
- Cavaggioni, A. & Mucignat-Caretta, C. Major urinary proteins,  $\alpha$ 2U-globulins and aphrodisin. *Biochim. et Biophys. Acta (BBA) Protein Struct. Molecul. Enzymol.* **1482**, 218–228 (2000).
- Tegoni, M. *et al.* Mammalian odorant binding proteins. *Biochim. et Biophys. Acta (BBA) Protein Struct. Molecul. Enzymol.* **1482**, 229–240 (2000).
- Nagnan-Le Meillour, P. *et al.* Binding specificity of native odorant-binding protein isoforms is driven by phosphorylation and O-N-Acetylglucosaminylation in the Pig *Sus scrofa*. *Front. Endocrinol.* **9**, 816 (2019).
- Lazar, J., Greenwood, D. R., Rasmussen, L. E. L. & Prestwich, G. D. Molecular and functional characterization of an odorant binding protein of the Asian elephant, *Elephas maximus*: Implications for the role of lipocalins in mammalian olfaction. *Biochemistry* **41**, 11786–11794 (2002).
- Schmale, H., Holtgreve-Grez, H. & Christiansen, H. Possible role for salivary gland protein in taste reception indicated by homology to lipophilic-ligand carrier proteins. *Nature* **343**, 366–369 (1990).
- Glasgow, B. J. Tear lipocalin and lipocalin-interacting membrane receptor. *Front. Physiol.* **12**, 1231 (2021).
- Freire-de-Lima, L. Sweet and sour: The impact of differential glycosylation in cancer cells undergoing epithelial-mesenchymal transition. *Front. Oncol.* **4**, 59 (2014).
- Scaloni, A. Chapter Five—Analysis of post-translational modifications in soluble proteins involved in chemical communication from mammals and insects. In *Methods in Enzymology* (eds Pelosi, P. & Knoll, W.) 103–124 (Academic Press, 2020).
- Amigues, B. *et al.* A new non-classical fold of varroa odorant-binding proteins reveals a wide open internal cavity. *Sci. Rep.* **11**, 13172 (2021).
- Zhu, J. *et al.* Reverse chemical ecology: Olfactory proteins from the giant panda and their interactions with putative pheromones and bamboo volatiles. *Proc. Natl. Acad. Sci. USA* **114**, E9802–E9810 (2017).
- Zaremska, V. *et al.* Reverse chemical ecology suggests putative primate pheromones. *Molecul. Biol. Evolut.* msab338 (2021).
- Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L. & Schwede, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci Rep* **7**, 10480 (2017).
- Studer, G. *et al.* ProMod3—A versatile homology modelling toolbox. *PLoS Comput. Biol.* **17**, e1008667 (2021).
- Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
- Sun, Y. F. *et al.* Two odorant-binding proteins mediate the behavioural response of aphids to the alarm pheromone (E)- $\beta$ -farnesene and structural analogues. *PLoS ONE* **7**, e32759 (2012).

34. Leal, G. M. & Leal, W. S. Binding of a fluorescence reporter and a ligand to an odorant-binding protein of the yellow fever mosquito *Aedes aegypti*. *F1000Res* **3**, 305 (2015).
35. Pelosi, P., Zhu, J. & Knoll, W. From radioactive ligands to biosensors: Binding methods with olfactory proteins. *Appl. Microbiol. Biotechnol.* **102**, 8213–8227 (2018).
36. Arena, S., Renzone, G. & Scaloni, A. A multi-approach peptidomic analysis of hen egg white reveals novel putative bioactive molecules. *J. Proteomics* **215**, 103646 (2020).
37. Pettersen, E. F. *et al.* UCSF chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
38. Perez-Riverol, Y. *et al.* The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).

## Acknowledgements

We thank Bruce Schulte of Western Kentucky University, USA, for helpful suggestions and discussion. We are also grateful to the whole elephant keeper team at Tiergarten Schönbrunn for their flexibility and for making this work possible.

## Author contributions

P.P. conceived and designed the study. V.Z. performed all laboratory experiments regarding protein expression and purification, ligand binding and characterization. V.Z. and P.P. analysed the biochemical data and prepared the relative figures. G.R., S.A., V.C. and A.S. performed the mass spectrometry experiments, analysed the data and prepared the presentation of the relative results. F.B. provided the samples of trunk wash, A.B. coordinated the keeper team, trained the elephants to willingly cooperate in trunk wash procedures and took the samples. W.K. contributed with ideas and discussion during the development of the work. P.P. wrote the first draft of the manuscript, which was implemented with specific sections by V.Z. and A.S. All authors contributed to the final version, critically read and approved the manuscript in its final form.

## Funding

This work was supported by FWF Austria Project P32472-BBL and Lower Austrian Provincial Government – Department K3 Science and Research, grant no. K3-F-915/001-2020.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-24214-5>.

**Correspondence** and requests for materials should be addressed to P.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022