

OPEN

Aedes aegypti Odorant Binding Protein 22 selectively binds fatty acids through a conformational change in its C-terminal tail

Jing Wang¹, Emma J. Murphy^{1,4}, Jay C. Nix³ & David N. M. Jones^{1,2*}

Aedes aegypti is the primary vector for transmission of Dengue, Zika and chikungunya viruses. Previously it was shown that Dengue virus infection of the mosquito led to an increased expression of the odorant binding protein 22 (AeOBP22) within the mosquito salivary gland and that siRNA mediated knockdown of AeOBP22 led to reduced mosquito feeding behaviors. Insect OBPs are implicated in the perception, storage and transport of chemosensory signaling molecules including air-borne odorants and pheromones. AeOBP22 is unusual as it is additionally expressed in multiple tissues, including the antenna, the male reproductive glands and is transferred to females during reproduction, indicating multiple roles in the mosquito life cycle. However, it is unclear what role it plays in these tissues and what ligands it interacts with. Here we present solution and X-ray crystallographic studies that indicate a potential role of AeOBP22 binding to fatty acids, and that the specificity for longer chain fatty acids is regulated by a conformational change in the C-terminal tail that leads to creation of an enlarged binding cavity that enhances binding affinity. This study sheds light onto the native ligands for AeOBP22 and provides insight into its potential functions in different tissues.

A critical step in disease transmission by hematophagous mosquitoes is the location of a human host for a blood meal by the female mosquito. Host location and selection of biting sites is driven by the perception of chemosensory stimuli that requires the interplay of a number of factors including chemosensory receptors and odorant binding proteins (OBPs)^{1,2}. *Ae. aegypti* OBP22 (AeOBP22) is a member of the OBP family of proteins that has been directly implicated in regulating these feeding behaviors³. AeOBP22 is unusual in that it is expressed in multiple chemosensory tissues, including in the antenna, the proboscis of the females, the thoracic spiracles⁴, in the male reproductive glands where it is transferred to the females during mating^{4,5}, and in the salivary glands^{3,6,7}. Surprisingly, it was discovered that, in combination with other chemosensory genes, its expression in the salivary glands is up regulated in response to Dengue virus (DENV) infection and that knockdown of AeOBP22 using dsRNA approaches led to reduced blood feeding behaviors³.

OBPs were first identified through their role as pheromone binding proteins (PBPs), which are components of the chemosensory apparatus that are secreted into the lymph fluid surrounding the neuronal dendrites of the chemosensory sensilla⁸. PBPs are a subgroup of OBPs that bind preferentially to pheromones. OBPs are essential for many aspects of chemosensory signal transduction^{9,10}. In the lymph it is proposed that OBPs function to transport hydrophobic ligands across the aqueous lymph fluid and deliver them to chemosensory receptors^{1,2,11,12}. Many studies have provided support for this hypothesis, particularly for perception of pheromonal compounds^{10,13–17}. However, other studies have suggested that lymph fluid is in fact an emulsion of fatty acids and that these can enhance dispersion of pheromones into the lymph and promote interactions with the OBP¹⁸. Additional evidence suggests that OBPs may rather function to sequester ligands to regulate the gain/sensitivity of the chemosensory response^{19,20}.

¹Dept. of Pharmacology, University of Colorado School of Medicine, Anschutz Medical Campus, 12801 East 17th Ave, Aurora, CO, 80045, USA. ²Program in Structural Biology and Biochemistry, University of Colorado School of Medicine, Anschutz Medical Campus, 12801 East 17th Ave, Aurora, CO, 80045, USA. ³Molecular Biology Consortium, Beamline 4.2.2, Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, California, USA. ⁴Present address: Alzheimer's Research UK Oxford Drug Discovery Institute, NDM Research Building, University of Oxford Old Road Campus, Roosevelt Drive, Oxford, OX3 7FZ, USA. *email: david.jones@cuanschutz.edu

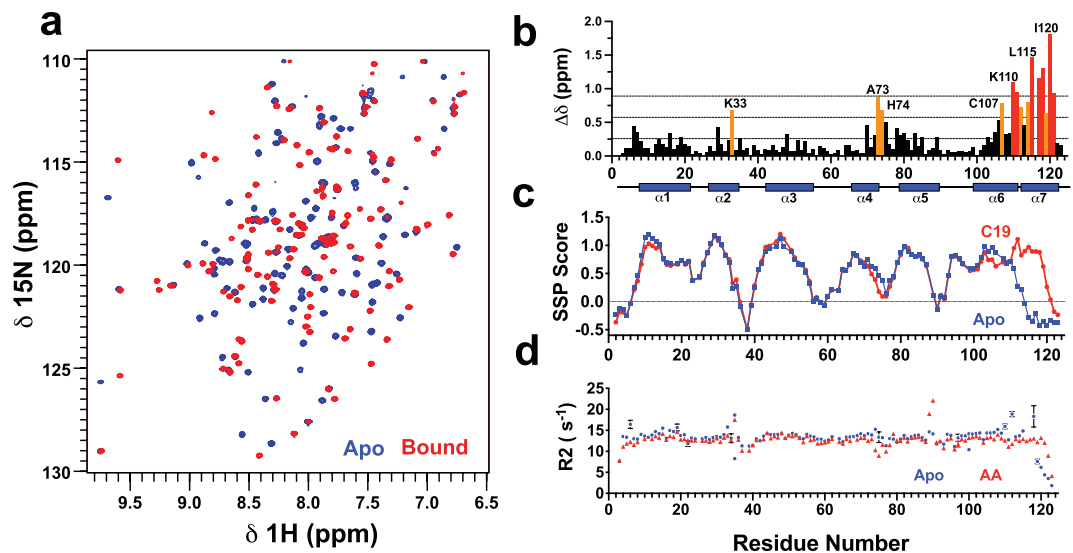


Figure 1. Conformational changes on binding of fatty acids to AeOBP22 **(a)** Region of the ^1H - ^{15}N HSQC of apo-AeOBP22 (blue) and bound to nonadecanoic acid (red). **(b)** Plot of normalized chemical shift changes between apo-AeOBP22 and bound to C19 fatty acid. Significant chemical shift changes are color coded as greater than 1 s.d. above the mean (orange) and 2 s.d. above the mean (red). Horizontal dashed lines indicate the position of the mean, +1 s.d. and +2 s.d. from bottom to top. The location of the α -helical regions is shown as blue cylinders below. **(c)** Plot of secondary structure propensity scores (SSP)⁴¹ for apo AeOBP22 (blue) and the complex with C19 (red). **(d)** Plot of NMR ^{15}N R_2 relaxation rates for apo (blue) and bound to arachidonic acid (red) recorded at a ^1H frequency of 900 MHz. Error bars are shown in black.

Increasingly it is apparent that OBPs are also expressed in outside of the primary chemosensory tissues including in hemolymph²¹, reproductive tissue^{4,5,22–24}, as components of mosquito eggshells^{23,25–27}, and as secreted components of the salivary glands of multiple insects^{28,29}. In particular it has been proposed that the D7 family of OBP related proteins in mosquito saliva function to limit inflammation^{30–32}, and blood clotting through their ability to sequester pro-inflammatory signals including biogenic amines and cysteinyl-leukotrienes³³. The expression of AeOBP22 in the antenna and proboscis clearly implicates it in regulating responses to host-derived odors that emanate from skin and/or sweat that drive blood feeding behaviors, while its expression in the salivary gland suggests the potential for it to be transferred to the human host during a blood meal. It is well established that components of salivary gland extracts (SGEs) dramatically impact blood feeding behaviors and viral infectivity^{33–36}. In order to better understand the potential roles of AeOBP22, we have undertaken a structural and biophysical characterization of the protein with the aim of understanding its ligand binding properties. These studies suggest that AeOBP22 may have evolved to bind to a range of fatty acids and that binding selectivity for longer chain fatty acids (>12 carbon atoms) is achieved through a conformational change in the C-terminal tail that leads to the formation of an expanded ligand binding pocket.

Results

NMR spectroscopy identifies long chain fatty acids as ligands for AeOBP22. We used NMR spectroscopy to screen compounds from human sweat and skin, repellents, fatty acids and bioactive lipids for binding to AeOBP22. We prepared purified, delipidated protein^{37–39} by extensively washing isolated inclusion bodies using buffer containing 1 M urea followed by refolding⁴⁰. Retrospectively, we determined that “non-delipidated” samples contain a mixture of the apo-state of the protein and the complex formed with palmitic acid (16 carbons) (Supplementary Fig. 1). The ^1H - ^{15}N HSQC spectrum of the fully delipidated sample (Fig. 1a) shows 119 of the expected 120 peaks for the backbone amides. In NMR screening, we consistently found that longer chain fatty acids (C16–C20 carbons) produce large changes in the appearance of the NMR spectrum at low concentrations (Fig. 1a). Other compounds can produce the same magnitude of NMR chemical shift changes, notably geraniol, citronellol and benzaldehyde (Supplementary Fig. 2). However, this only occurred at high ligand concentrations (typically > 1 mM) and binding is weak as exemplified by the concentration dependence of the chemical shift changes. In contrast, fatty acids bind with high affinity as evidenced by the presence of peaks from the free and the bound states present in slow exchange on the NMR timescale when the ligand is present in sub-stoichiometric quantities (Supplementary Fig. 3). From the NMR chemical shift assignments of the apo-protein and the complexes with nonadecanoic (C19) acid and arachidonic acid (AA)⁴⁰ we determined that fatty acid binding has the largest impact on residues 106–121 in the C-terminus (Fig. 1b). Further calculations of the secondary structural propensities (SSPs) from the chemical shift data⁴¹ predict that the C-terminal tail adopts an extended conformation in the apo-state and adopts an α -helical conformation when bound to longer chain fatty acids (Fig. 1c)⁴⁰.

Crystal structures of AeOBP22 complexes. Next we used X-ray crystallography to better define the conformational changes that occur on binding fatty acids and we determined structures in three different

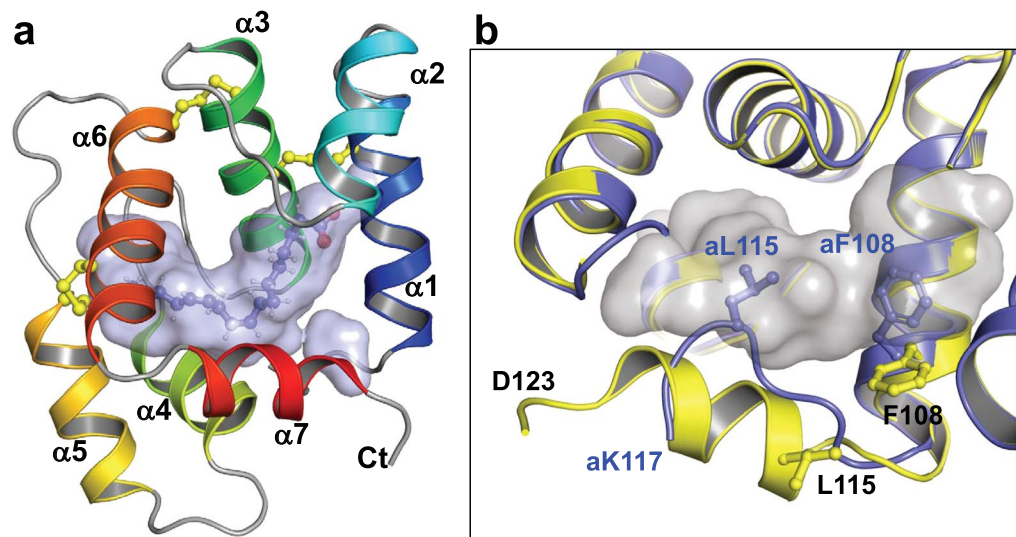


Figure 2. Structure of AeOBP22. **(a)** Ribbon diagram of the AeOBP22-linoleic acid complex solved by X-ray crystallography at 1.85 Å. Helices are color coded from blue to red, N to C terminus. The binding pocket is shown as a surface representation and linoleic acid is shown as sticks. **(b)** Comparison of the C-terminal tail in the linoleic acid complex (yellow) and in the apo-state (blue). In apo AeOBP22, L115 and F108 insert into the pocket (shown in grey) to occlude binding of larger ligands. Residue numbers for the apo-state are preceded with an “a”. Residues 118–123 of the apo-state are not observable in the crystal.

crystal forms. Initial crystals formed in the $P3_121$ space group, and the structure was solved using single wavelength anomalous dispersion (SAD) of a tantalum bromide (Ta_6Br_{12}) soaked crystal refined at 2.6 Å resolution. A native data set collected on the same crystal was refined at a resolution of 1.9 Å to an R/R_{Free} of 18.3/20.1% (Supplementary Fig. 4a). In this crystal form the protein forms a domain swapped dimer, with residues 116–122 of one molecule forming an anti-parallel β -sheet with residues 37–41 in a symmetry related molecule (Supplementary Fig. 4b).

The second crystal form was solved in the $P3_1$ space group and contains nine molecules in the asymmetric unit arranged in a pseudo three-fold arrangement of “trimers” consisting of domain swapped dimer and a separately packed monomer (Supplementary Fig. 5). The dimers are identical to those observed in crystal form 1 and superimpose with an average pairwise RMSD of 0.47 Å. The individual monomers superimpose to each other with an average pairwise RMSD of ~0.2 Å. Ligands are observed bound at the center of each monomer (Fig. 2a and 3a) with residues 112–121 forming an α -helix that forms one edge of the ligand-binding pocket (Fig. 2b). No ligand is observed in the dimeric components. In this form we solved the structures with palmitic acid (C16:0), palmitoleic acid (C16:1), and eicosanoic acid (C20) complexes.

The third crystal form grows in the presence of cadmium and/or cobalt ions in the $C121$ space group and contains only monomers. Structures of the apo-state and the complexes formed with linoleic acid (C18:2) and AA (C20:4) were solved using cobalt SAD methods and refined to a resolution of 1.85 Å. Data collection and refinement statistics for deposited structures are given in Supplementary Table 1.

AeOBP22 is monomeric in solution. Previous studies have proposed that dimerization of OBPs may regulate ligand binding and transport^{19,42–47}. Therefore, we used measurements of ¹⁵N NMR relaxation rates, R_1 and R_2 , to determine if AeOBP22 forms dimers or higher order complexes. For a globular protein there is a direct correlation between the average R_2/R_1 ratio and the molecular weight^{48–51}. For AeOBP22 the predicted molecular weights from multiple measurements of the R_2/R_1 ratios in both apo and fatty acid bound states (Supplementary Fig. 6) were in the range 14.1–15.2 kDa, in agreement with the expected molecular weight of 14.3 kDa, indicating that the protein exists as monomers in the absence of other binding partners. We saw no evidence of any significant dimerization over long periods even at the high concentrations used for NMR chemical shift assignments (600–700 μ M). Therefore, we conclude that the domain swapped dimers are an artifact of the crystallization process, and further discussions below are confined to the structures of the monomeric forms.

A complete analysis of the relaxation data using Relax^{52–54} (Supplementary Fig. 7) revealed slightly elevated exchange contributions to the relaxation rates for residues 111–118 suggesting that this region of the protein may undergo conformational exchange on slower time scales, and this may be important to allow access of fatty acids to the ligand binding pocket. This analysis also confirms that residues 74–78 exhibit increased conformational flexibility in the bound state compared to the apo-state.

Description of structure. The monomeric forms of AeOBP22 are similar to other classical OBPs which consist of six α -helices stabilized by three disulfide bridges surrounding a hydrophobic pocket (Fig. 2a). In the bound state, AeOBP22 is unusual in that it contains a seventh C-terminal α -helix that forms one edge of the ligand-binding pocket. The ligand-bound monomers from all our structures superimpose well with a pairwise RMSD of 0.40 Å (Supplementary Fig. 8). Small differences are seen in the position of helix-7, which rotates out

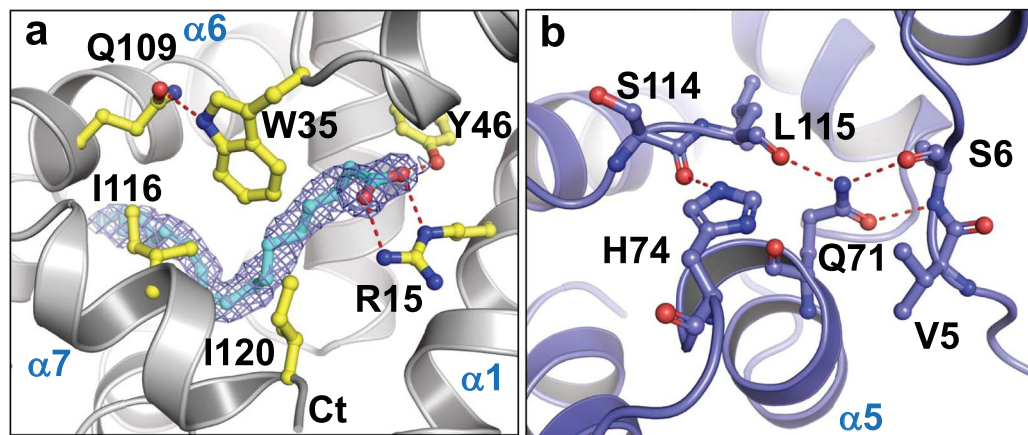


Figure 3. Critical interactions that stabilize the in the bound and apo-states of AeOBP22 (a) Electron density for linoleic acid in a 2Fo-Fc omit map contoured at 1σ with the fatty acid modelled in cyan. Y46 and R15 make specific H-bonds with the fatty acid. Positioning of the head group is reinforced by interactions with I120, I116, and W35, which H-bonds to Q109 in helix-6. (b) In apo-AeOBP22, the position of the C-terminal tail is stabilized by H-bonds between S114 and L115 with the side chains of H74 in the $\alpha 4$ - $\alpha 5$ loop and Q71 in helix-4, which in turn forms hydrogen bonds with S6 in the N-terminal region.

slightly from the binding pocket in the presence of larger fatty acids such that the $C\alpha$ at Ile120 is approximately 2 Å displaced from its position with shorter chain fatty acids. Additionally, we see variations in the position of the $\alpha 4$ - $\alpha 5$ loop (residues 73–78), which may be a result of crystal packing interactions.

The ligand-binding pocket is in the form of a long tunnel approximately 20 Å in length and occupies 144 ± 10 Å³ calculated using CASTp 3.0⁵⁵, and the alkyl chain of the ligand contacts hydrophobic residues that line the pocket (Supplementary Fig. 9). At the opening of the pocket there is an electrostatic patch formed by Arg15, Lys33 and Lys117 (Supplementary Fig. 10). Arg15 in combination with Tyr46 make multiple specific H-bonds to the carboxyl group of the fatty acid (Fig. 3a) and together define the requirement for a negatively charged group at this position. The fatty acid head group also contacts Trp35, which in turn hydrogen bonds to Gln109. All these residues are highly sensitive to ligand binding in NMR experiments.

At the distal end of the pocket from Arg15, three highly ordered water molecules hydrogen bond to Trp100, Ala101, Gly104, Cys88 and Val89. Analogous water molecules are observed in the binding pocket of *Ae. aegypti* Juvenile Hormone Binding Protein (AeJHBP)⁵⁶. This suggests that AeOBP22 may accommodate ligands with polar groups at this position. To test this, we examined the interactions with 16-hydroxy-hexadecanoic acid (C16-OH) and observed large NMR chemical shift changes relative to the apo-protein that are comparable to the changes observed with C16 (not shown). We observed additional changes for residues 86–92 that are consistent with the interaction of the C16-hydroxyl group at this position (Supplementary Fig. 11). However, residues throughout helices 6 and 7 (103–121) show line broadening and reduced intensity indicative of increased conformational averaging. Additionally, we found that C16-OH does not compete for binding of a fluorescent reporter in ligand binding assays (below) and so we conclude that distal polar groups appear to be unfavorable for binding and destabilize the conformation of the AeOBP22 complex, suggesting that the role of the buried water molecules is more likely to be as a structural component.

Structure of the apo-state of AeOBP22. The structure of apo-AeOBP22 confirms that residues 118–123 are disordered and 112–117 adopt an extended structure compared to the α -helix observed in the bound state (Fig. 2b and Supplementary Fig. 8b). In this extended conformation, Leu115 and Phe108 insert into the core of the protein and restrict the size of the binding pocket (Fig. 2b). This extended structure is stabilized by the formation of hydrogen bonds between Ser114 and Leu115, with His74 and Gln71 in the $\alpha 4$ - $\alpha 5$ loop respectively (Fig. 3b). In turn, Gln71 hydrogen bonds with Ser6 in the N-terminus. This network of interactions links conformational changes in the C-terminus to the N-terminus through the $\alpha 4$ - $\alpha 5$ loop. This explains our NMR relaxation data (Fig. 1d and Supplementary Fig. 6 and 7), which shows increased conformational mobility of the $\alpha 4$ - $\alpha 5$ loop on binding fatty acids because these interactions are disrupted.

NMR Solution Structure of AeOBP22 with Arachidonic Acid. In parallel with our crystallographic studies we determined the NMR solution structure of the AeOBP22-AA complex. AA was chosen because the vinylic protons of this fatty acid each have a unique chemical shift and are in a region of the spectrum (4.6–5.6 ppm in ¹H) that has minimal overlap with resonances from the protein⁴⁰. This increased chemical shift dispersion greatly facilitated the assignment of intermolecular NOEs between the protein and the ligand (Fig. 4a). The ensemble of 30 lowest energy structures from the final round of calculations superimpose with a backbone RMSD of 0.32 Å from the mean structure, while the carbon atoms in the ligand superimpose with an overall RMSD of 0.31 Å (Fig. 4b). The list of structural restraints and refinement statistics are given in Supplementary Table 2.

The NMR structure of the AeOBP22-AA complex superimposes with the crystal structure with an RMSD of 0.67 Å for the backbone atoms of residues 7–120 and confirms that the monomeric structure observed in the

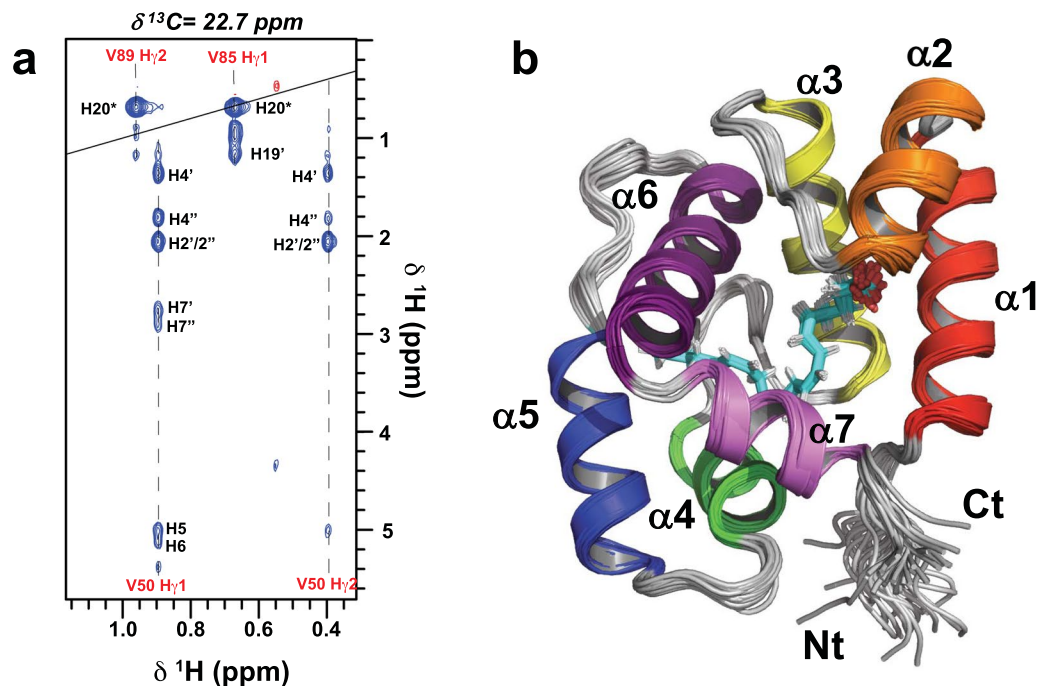


Figure 4. NMR Structure of the AeOBP22-AA Complex **(a)** A slice through a ^{12}C -edited/ ^{13}C -filtered intermolecular NOESY spectrum at $\delta^{13}\text{C} = 22.7$ ppm showing NOEs between resonances from the protein (labeled in red) and arachidonic acid (black labels). For the lipid the hydrogens are numbered according to the attached carbon in the alkyl chain. **(b)** Superposition of the 30 lowest energy structures from the final iteration of NMR calculations (RMSD = 0.32 Å for res 7–121), helices 1–7 are color coded from red-violet. The arachidonic acid is shown in cyan.

crystal is maintained in solution. The biggest differences between the structure are in the conformation of the loop between helices 4 and 5, which are involved in crystal packing, but which are dynamic in solution. The overall structure of the arachidonic acid ligand is similar to the structure observed in the crystal. However, even in the crystal there are differences in the ligand position in the two monomers in the asymmetric unit.

Impact of Chain Length on the Conformation of the C-terminal tail. In the apo-AeOBP22 structure we observed a small binding pocket located between Arg15 and Leu15 that could potentially accommodate short chain fatty acids (up to C6). When we examined the binding of fatty acids containing up to 8 carbons by NMR, we observed chain length dependent chemical shift changes for residues in the N-terminal half of the protein (Fig. 5a and Supplementary Fig. 12a), that are consistent with binding of the carboxylate group to the basic residues in the vicinity of Arg15, Lys33 and Trp35 (Supplementary Fig. 12b). To ensure that addition of free fatty acids did not result in changes in pH that could perturb the NMR spectrum, we examined chemical shift changes caused by pH alone (Supplementary Fig. 12b). We observed that the patterns and magnitude of chemical shift changes caused by short chain fatty acids are distinct from those caused by changes in pH alone. It was surprising that C8 still binds with relatively little impact on the overall structure suggesting a significant plasticity in the binding pocket. However, the exact mode of binding of short chain fatty acids remains to be definitively established.

In contrast to short chain fatty acids, large chemical shift perturbations are observed throughout the protein in the presence of decanoic acid (C10) (Fig. 5b). After reassigning the spectrum, we determined that C10 binding impacts the C-terminal tail, the α 4- α 5 loop, the end of helix-2 (residues 33–35) and helix-1. Several residues in the C-terminal tail show increased exchange broadening and could not be assigned (shaded in grey in Fig. 5b,c). Increasing the alkyl chain from C10 to C15, leads to additional perturbations in the C-terminus (aa 109–121) and in the α 4- α 5 loop (aa 70–75). However, when the alkyl chain is further increased from C15 to C19 additional chemical-shift perturbations localize predominantly to residues 6–21 in α -helix 1 (Fig. 5d), with limited changes in the C-terminal region. There are no additional changes beyond C20, and fatty acids with more than 20 carbon atoms show a dramatic reduction in their ability to bind (not shown). We conclude that fatty acids with at least 10 carbon atoms are required to stimulate a conformational change in the C-terminus, and this is maximally induced with a chain length of 15–16 carbon atoms. Secondly, a conformational shift in helix-1 is required to accommodate binding of longer chain fatty acids C17–C20. This adaptation differs to that observed in the crystal, where helix-7 shifts in response to longer chain fatty acid. We attribute this to the crystal packing contacts formed by helix-1 which restricts its ability to move. In contrast, helix-7 makes few crystal contacts and has greater ability to adapt its conformation to accommodate different ligands.

Fatty acid chain length and degree of unsaturation impact the binding affinity to AeOBP22.

Next, we asked how the chain length and presence of unsaturation impacts the binding affinity for AeOBP22. Many previous studies have used 1-NPN as a fluorescent reporter to examine ligand binding to OBPs^{57,58},

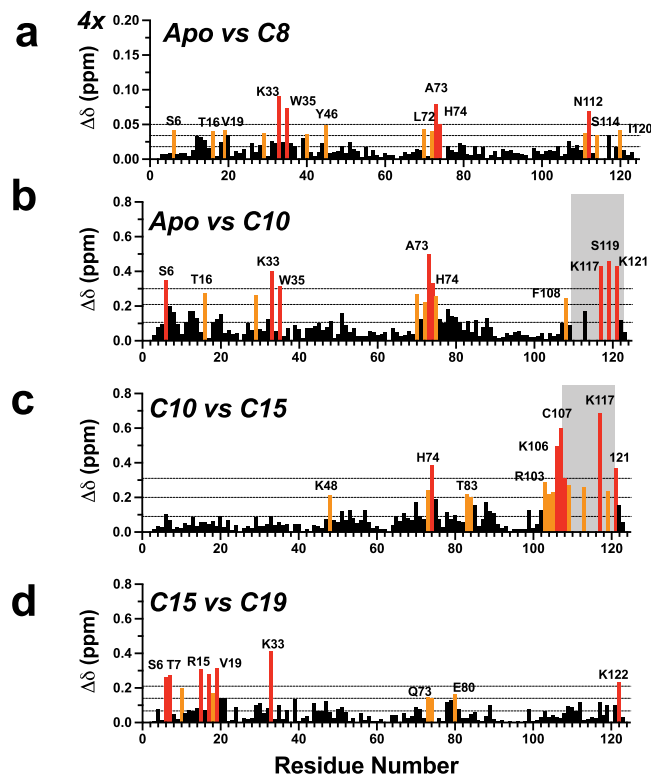


Figure 5. Chemical shift changes as a function of fatty acid chain length. Comparison of chemical shift differences in the ^1H - ^{15}N HSQC spectra recorded at 600 MHz for (a) apo compared to octanoic acid. (b) Apo versus C10. (c) C10 acid vs C15 acid and (d) C15 vs C19 acids. Significant chemical shift perturbations are color coded as in Fig. 1. In panel (a), the vertical scale has been expanded by four compared to the other panels. For C10–C14 conformational averaging in the C-terminal tail limited the assignments for a number of residues in this region (shaded in grey in b and c). In all cases the protein was at $\sim 100\ \mu\text{M}$ and the fatty acid was present at a concentration of $200\ \mu\text{M}$ in sodium phosphate at pH 6.5.

including for AeOBP22^{4,59}. In our hands, we found that 1-NPN was unsuitable for studies of fatty acid binding as it bound with low affinity and showed non-specific interactions that confounded analysis. Further, studies with 1-NPN are often confounded by the ability of ligands to bind simultaneously and quench 1-NPN fluorescence rather than bind in a competitive manner⁶⁰. Therefore, we investigated the use of a fluorescently labeled fatty acid derivative 5-(N-dodecanoyl)-amino-fluorescein (DAF), as a reporter for competition binding assays. DAF shows a strong fluorescence emission with a maximum at 513 nm which is quenched upon addition of AeOBP22 ($\sim 75\%$ reduction from maximal intensity) and with a K_D of $1.19 \pm 0.26\ \mu\text{M}$ ($n = 7$) (Fig. 6a,c). Addition of fatty acids with 14 or more carbon atoms releases DAF from AeOBP22 leading to a recovery of the initial fluorescence (Fig. 6b,d). In addition, NMR spectroscopy shows that DAF produces chemical shift changes comparable to those produced by C11–C14 fatty acids (Supplementary Fig. 13), and conformational line broadening for residues in the C-terminal region. Therefore, we conclude that the alkyl chain of DAF likely binds to the central pocket of AeOBP22 and can displace the C-terminal tail, and that fatty acids bind competitively with DAF.

We used DAF based competition binding assays to determine binding affinities for a series of fatty acids (Table 1 and Supplementary Fig. 14). This revealed that the both chain length and unsaturation impact the binding affinity. Saturated C16 binds weakly with $K_D \sim 690\ \text{nM}$, however, introduction of a single degree of unsaturation at the $\Delta 9$ position (C16:1) leads to an approximate four-fold increase in the binding affinity ($K_D = 175 \pm 41\ \text{nM}$). Increasing the chain length and the introduction of additional unsaturation enhances binding, which appears to be optimal for the fatty acids containing 18 carbon atoms, as increasing the chain length to C20 leads to a two-fold reduction in affinity. However, the extent of unsaturation also impacts the affinity, with the more rigid AA (C20:4) having a lower affinity than the monounsaturated C20:1. These results are consistent with our NMR data that shows binding of longer chain fatty acids requires a conformational change in α -helix 1 (Fig. 5d) and so this appears to be unfavorable for binding.

Changes in pH do not impact the binding of fatty acids. In apo-AeOBP22, His74 makes interactions that maintain the C-terminal residues in a conformation that occludes the ligand binding site. Consequently it is possible that pH may impact the conformation of the C-terminal residues and ligand binding in a manner somewhat analogous to that observed for PBPs^{61–64}. Indeed, previous studies of AeOBP22 have suggested that the helical content of the protein increases at lower pH⁵⁹. To test this, we compared the binding affinity for α -linoleic acid at pH 6.5 and 5.0 but found no significant difference at these two pHs, $52 \pm 8\ \text{nM}$ vs $42 \pm 14\ \text{nM}$ respectively (Fig. 6d,e and Table 1). We could not use DAF below pH 5.0 because its intrinsic fluorescence is severely

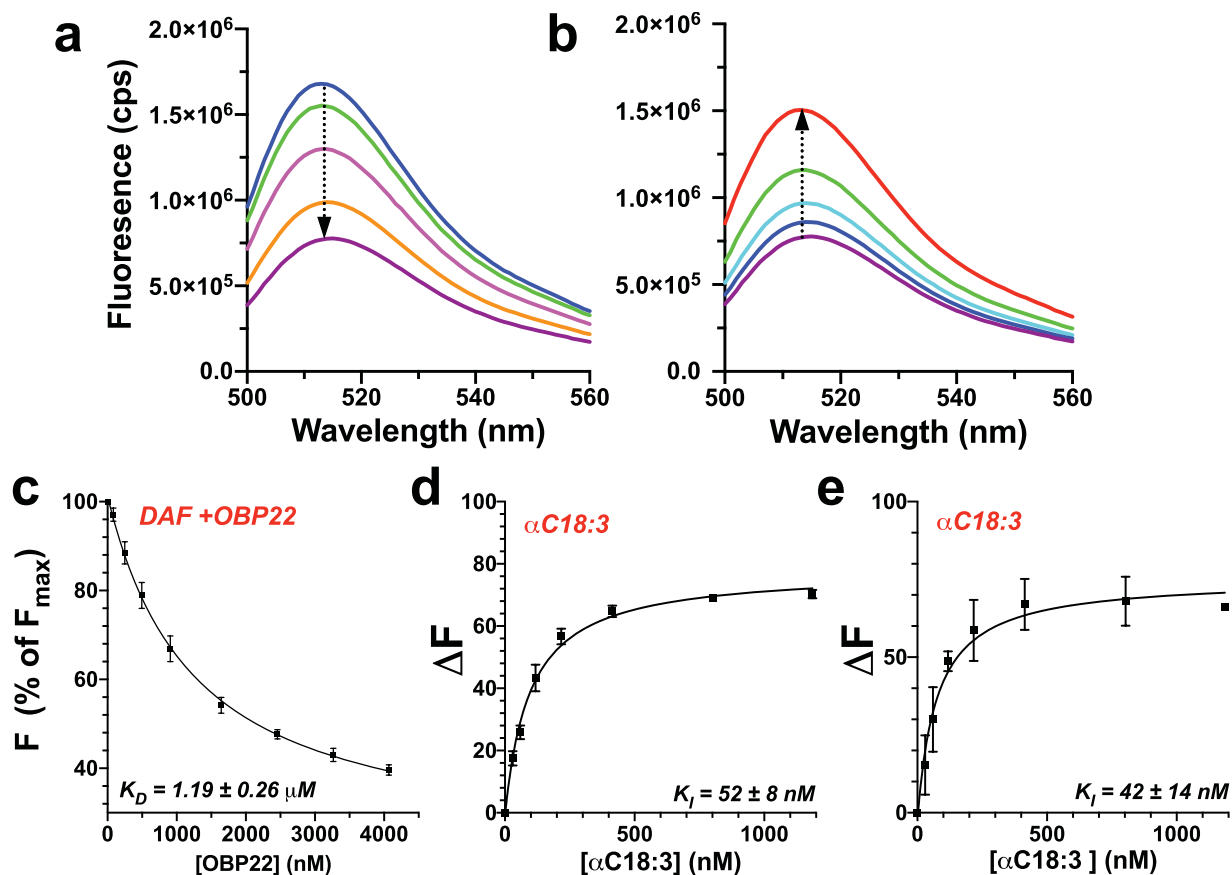


Figure 6. Long chain fatty acids bind with high affinity to AeOBP22. **(a)** Fluorescence emission spectra of DAF (blue) excited at 490 nm, with increasing amounts of AeOBP22 (arrow) shows concentration dependent quenching. **(b)** Titration of the final sample in a, with increasing linoleic acid (arrow) shows that the fatty acid competes for binding of DAF and recovery of initial fluorescence (red). **(c)** Determination of binding constant for AeOBP22 and DAF from 7 replicates of experiments shown in panel (a). DAF concentration was 84 nM. Plot of the intensity at the fluorescence maximum (513 nm) presented as change from initial fluorescence. Data were fit according to Eq. 2. **(d)** Determination of the binding constant of α -linoleic acid at pH 8.0 from multiple replicates ($n = 3$) of experiments shown in panel (b) by fit of the raw intensity using Eqs. 2 and 3 using a K_D for DAF of 1.191 μ M. The DAF concentration was fixed at 100 nM and the protein at 1000 nM. Data are presented as normalized recovery of the quenched fluorescence for ease of viewing. **(e)** The same experiment as in (d) but recorded at pH 5 ($n = 2$). Intrinsic DAF fluorescence is significantly quenched at pH 5.0 leading to increased error in measured points.

Ligand	K_D (nM) ¹
DAF	1191 \pm 261 ²
C16:0	689 \pm 89
C16:1	175 \pm 41
C18:1	84 \pm 5
C18:2	104 \pm 30
α C18:3 pH 8.0	52 \pm 8
pH 5.0	42 \pm 14 ³
γ C18:3	111 \pm 27
C20:1	102 \pm 16
C20:4	286 \pm 29

Table 1. Dissociation constants for binding of fatty acids to AeOBP22. ¹Values reported in nM and \pm the standard deviation, $n = 3$. ² $n = 7$. ³ $n = 2$.

quenched. Therefore, we used NMR to examine the effect of pH on the protein structure and binding. For these studies, samples were prepared in sodium citrate at pH 6.5 and the pH adjusted using hydrochloric acid over the range 6.5–4.5. We found that lowering the pH resulted in chemical shift changes consistent with protonation of His74 and His70 between pH 6.5 and 5.5 (Supplementary Fig. 14b) and sidechain carboxyls between pH 5.5 and

4.5 (not shown). However, the cumulative chemical shift changes (~ 0.3 ppm) over the pH range 6.5–4.5 are 5–6 times smaller than the perturbations induced by the binding of chain fatty acids (Fig. 1), indicating that changes in pH do not induce the conformational rearrangement observed with fatty acids. To validate this, we obtained backbone chemical shift assignments (^1H , ^{15}N , $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$) for both the apo-protein and the α -linoleic acid complex at pH 4.5 and examined differences in the secondary structural propensities and the normalized chemical shift differences between the two samples (Supplementary Fig. 15). These almost exactly matched the SSPs and patterns of chemical shift changes upon ligand binding observed for samples recorded at pH 6.5, and confirms that pH does not impact the structure of the apo-protein or the ability of AeOBP22 to bind to longer chain fatty acids at low pH.

Comparison of AeOBP22 with known structures. The bound state of AeOBP22 contains a seventh α -helix at the C terminus that forms one edge of the ligand-binding pocket. A DALI search^{65,66} shows the highest structural similarity to the N-terminal domains of the insect D7 proteins that contain dual OBP-domains; *Ae. aegypti* juvenile hormone binding protein⁵⁶ (AeJHBP, PDB-ID 5V13) with a DALI Z-score of 15.4, *An. stephensi* D7⁶⁷ (AsteD7, PDB 3NHT, Dali Z = 12.8), and *An. gambiae* D7-Leukotriene E4 complex (AgamD7, PDB 3DZT, Dali Z = 12.2)³¹. The seventh helix of AeOBP22 is in a remarkably similar position to the seventh helix of these D7 proteins (Fig. 7a, only the N-terminal domain of AsteD7 is shown). However, in the D7 proteins the binding pocket is not as deep, and the ligand (magenta in Fig. 7a) extends out of the binding pocket and contacts residues in the C-terminal domain (not shown).

There are three structures of single domain OBPs that have an additional seventh C-terminal helix; *An. gambiae* OBP7 (PDB 3R1O)⁶⁸, *Locusta migratoria* OBP1 (LmigOBP1, PDB 4PT1)⁶⁹ and the Mediterranean fruit fly, *Ceratitis capitata* OBP22 (PDB 6NHE)⁷⁰. Of these, AeOBP22 is most similar to LmigOBP1 (Dali Z = 12.1) (Fig. 7b), however, helix 7 is significantly longer in LmigOBP1 and the angle between helices 6 and 7 differs by $\sim 30^\circ$ (red in Fig. 7b). Additionally, in AeOBP22 helices 4 and 5 are rotated out from the main body of the protein by $\sim 28^\circ$. As a consequence, AeOBP22 has a longer, more extended pocket compared to that in LmigOBP1 (170 \AA^3).

AeOBP22 additionally shows structural conservation with the blowfly, *Phormia regina*, OBP56a (PregOBP56a, PDB 5DIC, Dali Z = 13.7) (Ishida *et al.* not published), which is proposed to transport fatty acids for feeding⁷¹. Helices 1 through 6 of PregOBP56a superimpose with AeOBP22 with an RMSD of 1.2 \AA (Fig. 7c). However, PregOBP56a does not have a seventh helix. Instead, the N-terminal residues 1–5 occupy a similar position to the C-terminus of AeOBP22. The binding pocket of PregOBP56a is significantly larger ($\sim 530 \text{\AA}^3$) and is lined by multiple lysine residues. This allows it to accommodate multiple ligands in a non-specific manner (Fig. 7c). This contrasts with AeOBP22, which has a hydrophobic pocket with Arg15 and Tyr46 positioned to make specific hydrogen bonds with carboxylic acid containing ligands (Fig. 3a).

AeOBP22 also has structural homology to *An. gambiae* OBP22 (AgamOBP22, PDB 3L4L, Dali Z = 12.5) (Zhang and Ren, not published) (Fig. 7d). AgamOBP22 differs in having a significantly shorter N-terminal region but simultaneously a much longer C-terminus, which is only partially observable in the deposited structures. This C-terminal region also forms a seventh α -helix, however the angle formed with helix-6 is more obtuse, $\sim 116^\circ$ compared to $\sim 75^\circ$, seen in AeOBP22 and the binding pocket is relatively shallow and smaller (123 \AA^3) than that of AeOBP22.

Discussion

It is increasingly apparent that insect OBPs have diverse roles that extend beyond those of chemosensory signaling⁷². AeOBP22 is a prime example of this as it is expressed in multiple tissues^{3–7}. Other OBPs expressed in multiple tissues include two related *Helicoverpa* spp., where it is proposed that HeOBP10 functions in both the delivery of an oviposition deterrent to fertilized eggs, and in the subsequent detection of that deterrent in the antenna²⁵. Understanding the native ligands that interact with OBPs can provide insights into the underlying biology. Our results provide evidence that fatty acids are likely natural ligands for AeOBP22 and that the binding of long-chain fatty acids (C15–C20) is enhanced by a conformational change in the C-terminal tail that is critical to generate a high affinity binding site. Previously we showed that burial of a single methylene group from a ligand contributes ~ 1 Kcal mol⁻¹ to the overall stability of the *Drosophila* OBP LUSH⁷³ and so burial of a long alkyl chain can contribute significant binding energy leading to high affinity interactions.

Conformational changes associated with ligand binding have been demonstrated for multiple OBPs and PBPs^{37–39,61,63,64,74–78}. These can stabilize the structure of the protein^{17,45,73,79}, accommodate different ligands^{68,74,75,80,81} or displace ligands^{37,61,64,76–78,82,83}. In *Antheraea polyphemus* PBP1, in the absence of ligand the C-terminal tail forms a long α -helix and occupies the binding site³⁷. It functions to both displace ligands at lower pH but also contributes to high affinity ligand binding at higher pH⁸⁴, even though it is displaced from the binding site, and so the mechanism for how it does this remains uncertain. The conformational change observed for AeOBP22 is more similar to the coil-to-helix transition observed for binding of norepinephrine to the C-terminal domain of *Ae. aegypti* D7 that results in capping of the ligand-binding pocket³¹. In the absence of ligand this C-terminal tail is highly dynamic. In contrast, the C-terminal tail of AeOBP22 is well ordered even in the absence of ligand (Fig. 3b). Therefore, we propose it functions as a selectivity filter for specific ligands of the appropriate chain length.

Previous studies of AeOBP22^{4,59} showed that the length of the ligand was critical for binding, and also branching of the alkyl chain was detrimental to binding indicative of binding to a narrow pocket, consistent with our findings. However, in these studies the “best” ligands were compounds containing two aromatic rings. Such compounds would only partially occupy the ligand binding pocket we observe, and in agreement with this show binding affinities an order of magnitude weaker than long chain fatty acids. We cannot rule out the possibility that ligands other than fatty acids bind to AeOBP22, and indeed, our initial screens identified small molecules

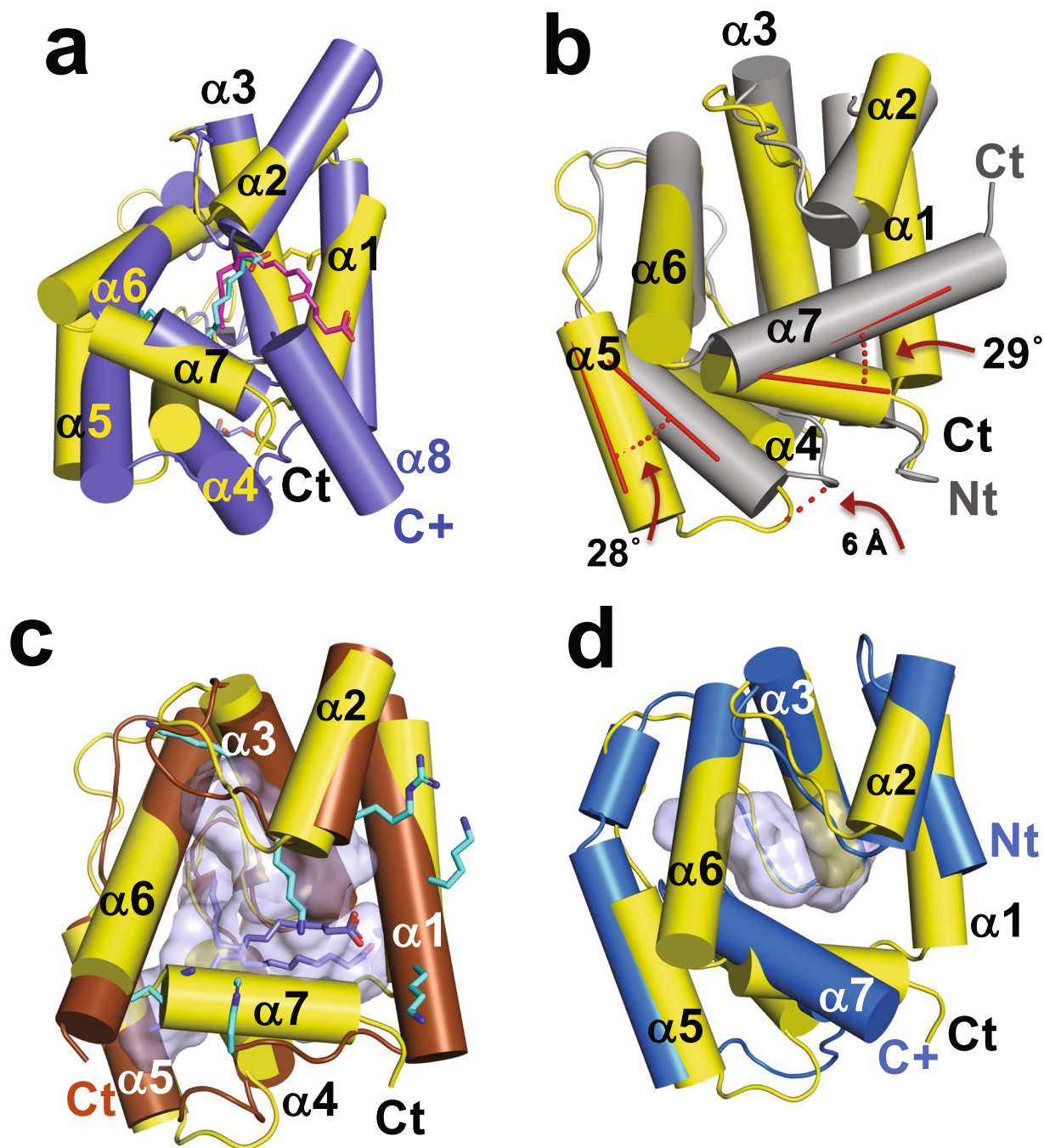


Figure 7. Comparison of AeOBP22 with known structures. Cylinder representation of AeOBP22 in yellow (N and C-termini labeled in black) compared with **(a)** N-terminal OBP domain of *An. stephensi* D7 (blue) bound to Leukotriene C4 magenta (PDB 3NHI)⁶⁷. For clarity, the C-terminal domain, which continues at the position labeled C+ in blue, is not shown. The linoleic acid bound to AeOBP22 is shown in cyan. **(b)** LmigOBP1 (grey) (PDB 4PT1)⁶⁹. The difference in the positions of helices 5 and 7 are shown in red. **(c)** *Phormia regina* OBP56a bound to multiple molecules of palmitic acid (blue) (PDB 5DIC) (Ishida *et al.* not published). The multiple lysines and arginines that surround the pocket (grey) are shown in cyan. **(d)** *An. gambiae* OBP22 (bright blue) (PDB 3L4L) (Zhang and Ren, not published). The C-terminus of AgOBP22 is mostly unstructured (extends from position labeled with blue C+).

that induce conformational changes but only at much higher concentrations (>1 mM) and in a concentration dependent manner, suggesting that these if these compounds bind *in vivo* they likely require multiple ligands to induce a stable conformation of the protein. Our structural studies provide strong evidence that a negatively charged head group is required on the ligand given the number and arrangement of positively charged residues that surround the ligand binding pocket. Our preliminary screens failed to identify lyso-phosphatidic acid (LPA) (C16-LPA) as a ligand, as it did not produce the same stabilization of AeOBP22 we observed with free fatty acids.

We now interpret this as because the overall chain length C16-LPA is too long (23 linear heavy atoms) to be accommodated in the binding pocket and indeed we found that fatty acids longer than C20 chains bind AeOBP22 with dramatically reduced affinities.

Fatty acids components of human sweat and skin and have critical roles in regulating mosquito behaviors. Dodecanoic acid and palmitoleic acid are strong oviposition attractants in *Ae. aegypti*, whilst the saturated C14–C16 and C18 acids are attractants at low concentrations but repellents at high concentrations⁸⁵. In contrast, methyl ester derivatives were found to be oviposition deterrents⁸⁵. Carboxylic acids have also been shown to synergize the effect of lactic acid as an attractant⁸⁶. Carboxylic acids with C1–C3, C5–C8 and C13–C18 alkyl chains all increased attraction compared to lactic acid alone, whereas, C4 and C9–C12 had no effect or showed a decrease in attraction. Whilst other studies have suggested that C9 and C10 acids have a higher stimulatory effect compared to other fatty acids⁸⁷. Our structural studies indicate that AeOBP22 retains the ability to bind to short chain carboxylic acids suggesting that AeOBP22 may have the ability to recognize different ligands in different tissues antennal AeOBP22 may modulate chemosensory responses to volatile short chain fatty acids, whether it functions as a transporter or a buffer of ligand concentration remains unknown²⁰. Within the insect chemosensory system, fatty acids are endogenous components of the sensillar lymph fluid¹⁸, and these fatty acids can interact with both pheromones and PBPs on external sites to regulate pheromone accessibility and interactions between the PBP and the pheromone. In our NMR studies of AeOBP22 even at high fatty acid concentrations, we only observe a single binding site for the fatty acid and no evidence of other alternative sites of interaction.

In the proboscis AeOBP22 could be function to recognize fatty acids. However, its function in the salivary gland and male reproductive tissue is less clear. As a secreted protein in the salivary gland^{6,7} it is likely injected into the host during a blood meal. The closely related protein PregOBP56a from the blow-fly *P. regina* has been proposed to sequester and transport lipids for feeding⁷¹. It is unlikely that AeOBP22 plays such a role in the mosquito given the difference in the mechanisms of feeding. Rather, it seems that it must function to either transport a required ligand from the mosquito to the host or to sequester a signal present in the host. Similarly, given its expression in male reproductive tissue and its transmission to females, it would appear that it must be transporting or sequestering a pheromonal component. Our data suggest that in either case, any compound involved in these processes is likely to be a long chain negatively charged lipid structure, and the ability to undergo a conformational change that enhances binding may be an important evolutionary development that allows it to recognize low abundance compounds with high affinity.

Materials and Methods

Protein expression. The mature form of AeOBP22 lacking the N-terminal signal peptide was expressed and purified as previously described^{88,40}.

NMR spectroscopy. NMR experiments were performed at 25 °C on either a Varian 900 MHz DD2, Varian INOVA 600 MHz or Bruker Avance Neo 600 MHz spectrometer. Samples for NMR were dissolved in sodium phosphate (20 mM, pH 6.5) and 90% H₂O/10% D₂O, with DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) (80 μM) added as an internal reference^{89,90}. For ligand screening, protein concentration was 100 μM and ligands were added to a concentration of 200–500 μM. For chemical shift assignments, relaxation rate measurements and structure calculations, protein concentrations were in the 400–700 μM range. All NMR experiments were collected using non uniform sampling methods⁹¹ using the Poisson-gap sampling schemes implemented by Hyberts *et al.*⁹² with a sampling density of 25–50%. Data were processed using the istHMS package v2111^{93,94} in combination with NmrPipe⁹⁵ and analyzed using Ccpnmr Analysis v 2.4.2⁹⁶.

Chemical shift assignments. Chemical shift assignments for the complex with arachidonic acid and the backbone assignments for the apo-protein were previously reported⁴⁰. Backbone ¹H, ¹⁵N, ¹³Cα and ¹³Cβ assignments for nonadecanoic, linoleic acid (pH 4.5 and pH 6.5), and decanoic acids were made in the same way. All other backbone amide assignments were made by following chemical shift trajectories as a function of ligand concentration or chain length. Normalized chemical shift differences are reported as^{97,98}

$$\Delta\delta = \sqrt{(\Delta\delta H^2 + 0.14 * \Delta\delta N^2)} \quad (1)$$

pH Titrations. Samples were dissolved in sodium citrate (20 mM, pH 6.5) and the pH adjusted using HCl and measured using a Lazar Ultra-M micro pH electrode (Lazar Labs, Los Angeles, CA).

NMR relaxation rates. ¹⁵N R₁, R_{1ρ} and {¹H}-¹⁵N heteronuclear NOEs for the apo-protein and the arachidonic acid complex were measured at 900 and 600 MHz with temperature compensation blocks, and the full set of relaxation delays were acquired prior to incrementation of the ¹⁵N evolution time⁹⁹. Heteronuclear NOEs used a 5 second relaxation delay, with a saturation period of 3 seconds. For R₁ measurements 11 relaxation delays were acquired spanning 0.01–1.2 seconds at 14 T and 0.01–1.8 s at 21 T. For R_{1ρ} measurements, the spin lock field strength at 21 T was 2250 Hz and at 14 T it was 1829 Hz, and relaxation delays were arrayed over 30–210 ms. R_{1ρ} values were converted into R₂ values¹⁰⁰ prior to analysis using the “d’Avergne” protocol in the Relax software suite^{52,53,101–103}.

NMR structure determination. NOE distance restraints were obtained from a simultaneous ¹⁵N/¹³C separated NOESY-HSQC¹⁰⁴ with a 100 ms NOE mixing times. Intermolecular NOEs between the protein and ligand were recorded using a simultaneous ¹³C/¹⁵N F₁-filtered, F₃-edited NOESY-HSQC^{105,106} (150 ms mixing time),

using samples recorded in both 90% H₂O/10% D₂O and a second that was exchanged into 99% D₂O (pD 6.1). Intermolecular restraints were derived from those NOE correlations that were observed uniquely in the intermolecular NOE spectrum.

Backbone dihedral angle restraints were determined using Talos+¹⁰⁷. Any dihedral angles that initially appeared over restrained (clustered at the extreme ends of the restraint) were removed from subsequent calculations. Unambiguous side-chain χ_1 restraints were identified from $^3J_{\text{HAHB}}$ coupling constants measured as previously described¹⁰⁸ in combination with NOE measurements. For structure calculations, the Karplus coefficients (A, B, and C)¹⁰⁹ for $^3J_{\text{HAHB}}$ were parameterized as given in Perez *et al.*¹¹⁰.

Backbone amides involved in hydrogen-bonds were identified from an ¹H-¹⁵N HSQC spectrum after the sample had been exchanged into 99% D₂O based buffer. Acceptor groups for the H-bond restraints were identified after structure calculations had converged and these were included in the final iterations of refinement.

Structure calculations were performed using ARIA 2.3.2^{111,112} in combination with CNS-SOLVE (Version 1.2)¹¹³. NOE restraints lists were generated within CCPNMR analysis and divided into unambiguous and ambiguous restraint sets. Initial iterations used a soft-square well potential for the NOE restraints to identify violations in the restraint lists, and if violated these were subsequently treated as ambiguous restraints. The final rounds of calculations employed the log-harmonic potential in the final cooling stage¹¹⁴ in combination with a modified “soft” force field and updated weighting parameters¹¹⁵. In the final iteration, a total of 100 structures were determined, and the top 50 were refined in water and the top 30 of these selected for analysis based on their total energy as described¹¹⁴. Superposition of structures and calculation of RMSD were performed using the suppose program from Ambertools¹¹⁶.

X-ray crystallography. *Crystallization.* For crystallization trials, AeOBP22 was exchanged into 4-(2-hydroxyethyl)-1-piperazine-ethane-sulfonic acid (HEPES) buffer at pH 7.5. The protein was concentrated to 9–12 mg ml⁻¹ and incubated with ligands at room temperature overnight prior to crystallization trials. Crystals were grown by sitting drop vapor-diffusion.

Initial crystals were obtained in the P3₁21 space group in the presence of benzaldehyde (2 mM), by mixing the protein solution in a 1:1 ratio with a precipitant solution containing sodium-potassium tartrate (0.2 M), sodium citrate (0.1 M) pH 5.0 and ammonium sulfate (2 M). A tantalum bromide derivative was obtained by soaking crystals with a 2 mM Ta₆Br₁₂ solution (Jena Biosciences) for 2–3 days. Crystals of the complexes in the P3₁ space group were obtained at 18 °C through optimization of an initial hit containing ammonium sulfate (1.2–1.4 M), citric acid (0.1 M), and varying the pH over 5.5–5.7, and optimized by addition of sodium chloride to the well solution across a concentration range of 1.0–1.5 M. Crystals in the C121 space group were obtained using PEG3550 (12% w/v) and HEPES (0.1 M) at pH 7.5. For the C18:2 complex the precipitant solution additionally contained cobalt chloride (4 mM), cadmium chloride (4 mM) and magnesium chloride (4 mM). Arachidonic acid complex crystals were obtained with cobalt chloride (6 mM), cadmium chloride (10 mM) and magnesium chloride (6 mM). The apo-protein crystal was obtained with cobalt chloride (6 mM), cadmium chloride (6 mM) and sodium chloride (6 mM).

Data collection. X-ray diffraction data was collected at the Molecular Biology Consortium, Beamline 4.2.2. at the Advanced Light Source, LBNL, Berkeley CA (crystal form 1 and 2) and the University of Colorado School of Medicine Biomolecular X-ray Crystallography Center (crystal form 3) using a Rigaku MicroMax™ 007 HF generator with a copper anode and equipped with a Rigaku PILATUS3 R 200 K detector.

Structure solution and refinement. The initial structure of AeOBP22 was solved using SAD methods of a Ta₆Br₁₂ soaked crystals in the P3₁21 space group collected at a wavelength of 1.254 Å. Data were integrated and scaled using XDS^{117,118} and refined to a resolution of 2.59 Å using Phenix¹¹⁹ to locate heavy atoms sites and generate initial maps. A native data set collected at 1.0 Å on the same crystal was solved using Phenix¹¹⁹ with model rebuilding in Coot¹²⁰ and refined to a resolution of 1.9 Å.

Crystal structures of the complexes in the P3₁ space group were solved by molecular replacement¹²¹ using the structure obtain from the Ta₆Br₁₂ crystal as an initial search model. Data were collected at both the ALS beamline 4.2.2 and on the home-source, which was processed with HKL3000¹²². These crystals are pathologically twinned (~46%), and data were refined using intensity based twin refinement in Refmac¹²³ within CCP4^{124,125}. Structures of the complexes in the C121 space group were solved by cobalt or cadmium SAD methods using data sets collected in house at a wavelength of 1.5418 Å, structures were solved using Phenix and subsequently refined directly against the SAD data in Refmac¹²³.

Ligand binding measurements. *Affinity of AeOBP22 for 5-(N-dodecanoyl)-aminofluorescein (DAF).* All fluorescence experiments were recorded at 25 °C on a Horiba Fluorolog 3–1–1 spectrofluorometer. DAF spectra were acquired using an excitation wavelength of 490 nm and the emission recorded between 500 and 560 nm at 1 nm intervals with an integration time of 1 s and slit widths of 2 nm. Stock solutions of DAF (Thermo Fisher, Batch D-109) were prepared in methanol with potassium hydroxide (0.2 M) and the concentration measured by UV spectroscopy using an extinction coefficient of 88000 M⁻¹ cm⁻¹ at 497 nm provided in the certificate of analysis. A 1:100 dilution was made into HEPES (50 mM, pH 8.1) and the sample vortexed and checked by UV to verify the DAF concentration. A solution of DAF (~80–100 nM) was titrated with increasing AeOBP22 to a final concentration of ~4 μM. The fluorescence intensity (cps) at the emission maxima (513 nm) was plotted against the total concentration of AeOBP22 and the K_D for DAF was determined by fitting the raw fluorescence intensity for each replicate to Eq. (1):

$$F_i = F_0(1 - ([P_T + L_T + K_D] - \sqrt{[P_T + L_T + K_D]^2 - 4P_T L_T})/2L_T) - F_{min} \quad (2)$$

where F_i is the measured fluorescence, F_0 is the initial fluorescence in the absence of any protein, P_T is the total protein, L_T is the total concentration of DAF, and F_{min} is the residual fluorescence after maximal quenching by the protein.

Binding affinity for fatty acids. Binding affinities were determined by monitoring the recovery of DAF fluorescence as a function of fatty acid concentration. Stock solutions of fatty acid were made by dissolving the fatty acid in 100% DMSO, and then making a 1:1000 dilution into HEPES (50 mM pH 8.1). Samples were vortexed prior to each titration point. The protein/DAF solution contained fixed concentrations of DAF (100 nM) and AeOBP22 (1000 nM). Background controls involved titration of the same DAF solution with solvent or fatty acid which showed no significant effect on the fluorescence intensity of DAF.

As there is a single binding site for the fatty acid and the effect of fluorescence quenching can be completely reversed by increasing concentrations of fatty acids, we concluded that the fatty acid is directly competitive with binding of the DAF reporter. Therefore, the term for K_D in Eq. 2 can be replaced with Eq. 3.

$$K_{app} = K_D(1 + I/K_I) \quad (3)$$

where I is the concentration of the competitor and K_I is the dissociation constant of the competitor. The emission maxima were fit using Eq. 2 substituted into Eq. 1, with the concentration of the protein and DAF fixed at their known concentrations.

Data availability

Coordinates for the AeOBP22 and its complexes have been deposited in the Protein Data Bank (<https://www.rcsb.org/>) with the accession numbers: 6OTL – TaBr Complex (SAD); 6P2E – TaBr (Native); 6OMW – palmitoleic acid; 6OPB – eicosanoic acid; 6OG0 – apo state; 6OGH – linoleic acid; 6OII – arachidonic acid. 6NBN – NMR structure with arachidonic acid. NMR chemical shift assignments and structural restraints have been deposited with the Biological Magnetic Resonance Bank (<http://www.bmrb.wisc.edu/>) with accession numbers 27724 and 30550.

Received: 16 August 2019; Accepted: 10 February 2020;

Published online: 24 February 2020

References

- Pelosi, P., Zhou, J. J., Ban, L. P. & Calvello, M. Soluble proteins in insect chemical communication. *Cell. Mol. Life Sci.* **63**, 1658–1676 (2006).
- Leal, W. S. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol.* **58**, 373–391 (2013).
- Sim, S., Ramirez, J. L. & Dimopoulos, G. Dengue Virus Infection of the Aedes aegypti Salivary Gland and Chemosensory Apparatus Induces Genes that Modulate Infection and Blood-Feeding Behavior. *PLoS Path.* **8**, e1002631 (2012).
- Li, S. *et al.* Multiple functions of an odorant-binding protein in the mosquito Aedes aegypti. *Biochem. Biophys. Res. Commun.* **372**, 464–468 (2008).
- Siro, L. K. *et al.* Identity and transfer of male reproductive gland proteins of the dengue vector mosquito, Aedes aegypti: potential tools for control of female feeding and reproduction. *Insect Biochem. Mol. Biol.* **38**, 176–189 (2008).
- Dhawan, R. *et al.* Mosquito-Borne Diseases and Omics: Salivary Gland Proteome of the Female Aedes aegypti Mosquito. *OMICS: J. Integr. Biol.* **21**, 45–54 (2017).
- Dhawan, R. *et al.* Data from salivary gland proteome analysis of female Aedes aegypti Linn. *Data Brief.* **13**, 274–277 (2017).
- Vogt, R. G. & Riddiford, L. M. Pheromone binding and inactivation by moth antennae. *Nat.* **293**, 161–163 (1981).
- Kim, M.-S., Repp, A. & Smith, D. P. LUSH Odorant-binding protein mediates chemosensory responses to alcohols in *Drosophila melanogaster*. *Genet.* **150**, 711–721 (1998).
- Xu, P., Atkinson, R., Jones, D. N. M. & Smith, D. P. *Drosophila* OBP LUSH is Required for Activity of Pheromone-Sensitive Neurons. *Neuron* **45**, 193–200 (2005).
- Krieger, J. & Breer, H. Olfactory reception in invertebrates. *Sci.* **286**, 720–723 (1999).
- Sun, J. S., Xiao, S. & Carlson, J. R. The diverse small proteins called odorant-binding proteins. *Open. Biol.* **8**, 180208 (2018).
- Forstner, M., Breer, H. & Krieger, J. A receptor and binding protein interplay in the detection of a distinct pheromone component in the silkworm *Antheraea polyphemus*. *Int. J. Biol. Sci.* **5**, 745–757 (2009).
- Grosse-Wilde, E., Gohl, T., Bouche, E., Breer, H. & Krieger, J. Candidate pheromone receptors provide the basis for the response of distinct antennal neurons to pheromonal compounds. *Eur. J. Neurosci.* **25**, 2364–2373 (2007).
- Grosse-Wilde, E., Svatos, A. & Krieger, J. A pheromone-binding protein mediates the bombykol-induced activation of a pheromone receptor *in vitro*. *Chem. Senses* **31**, 547–555 (2006).
- Syed, Z., Ishida, Y., Taylor, K., Kimbrell, D. A. & Leal, W. S. Pheromone reception in fruit flies expressing a moth's odorant receptor. *Proc. Natl Acad. Sci. USA* **103**, 16538–16543 (2006).
- Laughlin, J. D., Ha, T. S., Jones, D. N. & Smith, D. P. Activation of pheromone-sensitive neurons is mediated by conformational activation of pheromone-binding protein. *Cell* **133**, 1255–1265 (2008).
- Nardella, J., Terrado, M., Honson, N. S. & Plettner, E. Endogenous fatty acids in olfactory hairs influence pheromone binding protein structure and function in *Lymantria dispar*. *Arch. Biochem. Biophys.* **579**, 73–84 (2015).
- Honson, N., Johnson, M. A., Oliver, J. E., Prestwich, G. D. & Plettner, E. Structure-Activity Studies with Pheromone-binding Proteins of the Gypsy Moth, *Lymantria dispar*. *Chem. Senses* **28**, 479–489 (2003).
- Larter, N. K., Sun, J. S. & Carlson, J. R. Organization and function of *Drosophila* odorant binding proteins. *eLife* **5** <https://doi.org/10.7554/eLife.20242>. (2016).
- Rothmund, S., Liou, Y. C., Davies, P. L., Krause, E. & Sonnichsen, F. D. A new class of hexahelical insect proteins revealed as putative carriers of small hydrophobic ligands. *Structure* **7**, 1325–1332 (1999).
- Takemori, N. & Yamamoto, M. T. Proteome mapping of the *Drosophila melanogaster* male reproductive system. *Proteom.* **9**, 2484–2493 (2009).

23. Sun, Y.-L., Huang, L.-Q., Pelosi, P. & Wang, C.-Z. Expression in antennae and reproductive organs suggests a dual role of an odorant-binding protein in two sibling *Helicoverpa* species. *PLoS One* **7**, e30040 (2012).
24. Scolari, F. *et al.* The Spermatophore in *Glossina morsitans morsitans*: Insights into Male Contributions to Reproduction. *Sci. Rep.* **6**, 20334 (2016).
25. Ameny, D. A. *et al.* Proteomics reveals novel components of the *Anopheles gambiae* eggshell. *J. Insect Physiol.* **56**, 1414–1419 (2010).
26. Costa-da-Silva, A. L., Kojin, B. B., Marinotti, O., James, A. A. & Capurro, M. L. Expression and accumulation of the two-domain odorant-binding protein AegOBP45 in the ovaries of blood-fed *Aedes aegypti*. *Parasit. Vectors* **6**, 364 (2013).
27. Marinotti, O. *et al.* Integrated proteomic and transcriptomic analysis of the *Aedes aegypti* eggshell. *BMC Dev. Biol.* **14**, 15 (2014).
28. Ji, R. *et al.* Comparative transcriptome analysis of salivary glands of two populations of rice brown planthopper, *Nilaparvata lugens*, that differ in virulence. *PLoS One* **8**, e79612 (2013).
29. Celorio-Mancera Mde, L. *et al.* Chemosensory proteins, major salivary factors in caterpillar mandibular glands. *Insect Biochem. Mol. Biol.* **42**, 796–805 (2012).
30. Calvo, E., Mans, B. J., Andersen, J. F. & Ribeiro, J. M. Function and evolution of a mosquito salivary protein family. *J. Biol. Chem.* **281**, 1935–1942 (2006).
31. Calvo, E., Mans, B. J., Ribeiro, J. M. C. & Andersen, J. F. Multifunctionality and mechanism of ligand binding in a mosquito antiinflammatory protein. *Proc. Natl Acad. Sci. USA* **106**, 3728–3733 (2009).
32. Mans, B. J., Calvo, E., Ribeiro, J. M. & Andersen, J. F. The crystal structure of D7r4, a salivary biogenic amine-binding protein from the malaria mosquito *Anopheles gambiae*. *J. Biol. Chem.* **282**, 36626–36633 (2007).
33. Isawa, H., Yuda, M., Orito, Y. & Chinzei, Y. A mosquito salivary protein inhibits activation of the plasma contact system by binding to factor XII and high molecular weight kininogen. *J. Biol. Chem.* **277**, 27651–27658 (2002).
34. Ribeiro, J. M. Blood-feeding in mosquitoes: probing time and salivary gland anti-haemostatic activities in representatives of three genera (*Aedes*, *Anopheles*, *Culex*). *Med. Vet. Entomol.* **14**, 142–148 (2000).
35. Conway, M. J. *et al.* *Aedes aegypti* D7 Saliva Protein Inhibits Dengue Virus Infection. *PLoS Neglected Tropical Dis.* **10**, e0004941 (2016).
36. Conway, M. J. *et al.* Mosquito saliva serine protease enhances dissemination of dengue virus into the mammalian host. *J. Virol.* **88**, 164–175 (2014).
37. Katre, U. V., Mazumder, S., Prusti, R. K. & Mohanty, S. Ligand binding turns moth pheromone-binding protein into a pH sensor: effect on the *Antheraea polyphemus* PBP1 conformation. *J. Biol. Chem.* **284**, 32167–32177 (2009).
38. Zubkov, S., Gronenborn, A. M., Byeon, I. J. & Mohanty, S. Structural consequences of the pH-induced conformational switch in *A. polyphemus* pheromone-binding protein: mechanisms of ligand release. *J. Mol. Biol.* **354**, 1081–1090 (2005).
39. Mohanty, S., Zubkov, S. & Gronenborn, A. M. The solution NMR structure of *Antheraea polyphemus* PBP provides new insight into pheromone recognition by pheromone-binding proteins. *J. Mol. Biol.* **337**, 443–451 (2004).
40. Jones, D. N. M., Wang, J. & Murphy, E. J. Complete NMR chemical shift assignments of odorant binding protein 22 from the yellow fever mosquito, *Aedes aegypti*, bound to arachidonic acid. *Biomol. NMR Assign.* **13**, 187–193 (2019).
41. Marsh, J. A., Singh, V. K., Jia, Z. & Forman-Kay, J. D. Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein Sci.* **15**, 2795–2804 (2006).
42. Plettner, E., Lazar, J., Prestwich, E. G. & Prestwich, G. D. Discrimination of pheromone enantiomers by two pheromone binding proteins from the gypsy moth *Lymantria dispar*. *Biochem. J.* **39**, 8953–8962 (2000).
43. Briand, L., Nespoulous, C., Huet, J. C., Takahashi, M. & Pernollet, J. C. Ligand binding and physico-chemical properties of ASP2, a recombinant odorant-binding protein from honeybee (*Apis mellifera* L.). *Eur. J. Biochem.* **268**, 752–760 (2001).
44. Wogulis, M., Morgan, T., Ishida, Y., Leal, W. S. & Wilson, D. K. The crystal structure of an odorant binding protein from *Anopheles gambiae*: Evidence for a common ligand release mechanism. *Biochem. Biophys. Res. Commun.* **339**, 157–164 (2006).
45. Davrazou, F., Dong, E., Murphy, E. J., Johnson, H. T. & Jones, D. N. New insights into the mechanism of odorant detection by the malaria-transmitting mosquito *Anopheles gambiae*. *J. Biol. Chem.* **286**, 34175–34183 (2011).
46. Tsitsanou, K. E. *et al.* Crystal and solution studies of the “Plus-C” odorant-binding protein 48 from *Anopheles gambiae*: control of binding specificity through three-dimensional domain swapping. *J. Biol. Chem.* **288**, 33427–33438 (2013).
47. Qiao, H. *et al.* Cooperative interactions between odorant-binding proteins of *Anopheles gambiae*. *Cell. Mol. Life Sci.* **68**, 1799–1813 (2011).
48. Kay, L. E., Torchia, D. A. & Bax, A. Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochem. J.* **28**, 8972–8979 (1989).
49. Kroenke, C. D., Loria, J. P., Lee, L. K., Rance, M. & Palmer, A. G. Longitudinal and transverse H-1-N-15 dipolar N-15 chemical shift anisotropy relaxation interference: Unambiguous determination of rotational diffusion tensors and chemical exchange effects in biological macromolecules. *J. Am. Chem. Soc.* **120**, 7905–7915 (1998).
50. Luginbuhl, P. & Wuthrich, K. Semi-classical nuclear spin relaxation theory revisited for use with biological macromolecules. *Prog. Nucl. Magn. Reson. Spectrosc.* **40**, 199–247 (2002).
51. Cavanagh, J., Fairbrother, W. J., Palmer III, A. G., Rance, M. & Skelton, N. *Protein NMR Spectroscopy: Principles and Practice*. 2nd Vol. (ed. Elsevier, 2007).
52. d’Auvergne, E. J. & Gooley, P. R. Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces. *J. Biomol. NMR* **40**, 107–119 (2008).
53. d’Auvergne, E. J. & Gooley, P. R. Optimisation of NMR dynamic models II. A new methodology for the dual optimisation of the model-free parameters and the Brownian rotational diffusion tensor. *J. Biomol. NMR* **40**, 121–133 (2008).
54. Bieri, M., d’Auvergne, E. J. & Gooley, P. R. relaxGUI: a new software for fast and simple NMR relaxation data analysis and calculation of ps-ns and ms motion of proteins. *J. Biomol. NMR* **50**, 147–155 (2011).
55. Tian, W., Chen, C., Lei, X., Zhao, J. & Liang, J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res.* **46**, W363–W367 (2018).
56. Kim, I. H. *et al.* A mosquito hemolymph odorant-binding protein family member specifically binds juvenile hormone. *J. Biol. Chem.* **292**, 15329–15339 (2017).
57. Zhou, J. J. *et al.* Revisiting the odorant-binding protein LUSH of *Drosophila melanogaster*: evidence for odour recognition and discrimination. *FEBS Lett.* **558**, 23–26 (2004).
58. Murphy, E. J., Booth, J. C., Davrazou, F., Port, A. M. & Jones, D. N. Interactions of *Anopheles gambiae* odorant-binding proteins with a human-derived repellent: implications for the mode of action of n,n-diethyl-3-methylbenzamide (DEET). *J. Biol. Chem.* **288**, 4475–4485 (2013).
59. Yang, G., Winberg, G., Ren, H. & Zhang, S. Expression, purification and functional analysis of an odorant binding protein AegOBP22 from *Aedes aegypti*. *Protein Expr. Purif.* **75**, 165–171 (2011).
60. Murphy, E. J., Booth, J. C., Davrazou, F., Port, A. M. & Jones, D. N. Interactions of *Anopheles gambiae* odorant binding proteins with a human-derived repellent: implications for the mode of action of DEET. *J. Biol. Chem.* (2012).
61. Wojtasek, H. & Leal, W. Conformational change in the pheromone-binding protein from *Bombyx mori* induced by pH and by interaction with membranes. *J. Biol. Chem.* **274**, 30950–30956 (1999).
62. Damberger, F. *et al.* NMR characterization of a pH-dependent equilibrium between two folded solution conformations of the pheromone-binding protein from *Bombyx mori*. *Protein Sci.* **9**, 1038–1041 (2000).

63. Sandler, B. H., Nikonova, L., Leal, W. S. & Clardy, J. Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. *Chem. Biol.* **7**, 143–151 (2000).
64. Lautenschlager, C., Leal, W. S. & Clardy, J. Coil-to-helix transition and ligand release of *Bombyx mori* pheromone-binding protein. *Biochem. Biophys. Res. Commun.* **335**, 1044–1050 (2005).
65. Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–549 (2010).
66. Holm, L., Kaariainen, S., Wilton, C. & Plewczynski, D. Using Dali for structural comparison of proteins. *Curr Protoc Bioinformatics* **Chapter 5**, Unit 5 5, (2006).
67. Alvarenga, P. H. *et al.* The function and three-dimensional structure of a thromboxane A₂/cysteinyl leukotriene-binding protein from the saliva of a mosquito vector of the malaria parasite. *PLoS Biol.* **8**, e1000547 (2010).
68. Lagarde, A. *et al.* The crystal structure of odorant binding protein 7 from *Anopheles gambiae* exhibits an outstanding adaptability of its binding site. *J. Mol. Biol.* **414**, 401–412 (2011).
69. Zheng, J. *et al.* Crystal structure of the *Locusta migratoria* odorant binding protein. *Biochem. Biophys. Res. Commun.* **456**, 737–742 (2015).
70. Falchetto, M. *et al.* Structural and biochemical evaluation of *Ceratitis capitata* OBP22 affinity for odorants involved in inter-sex communication. *Insect Mol. Biol.*, imb.1255, (2018).
71. Ishida, Y., Ishibashi, J. & Leal, W. S. Fatty Acid Solubilizer from the Oral Disk of the Blowfly. *PLoS One* **8**, e51779 (2013).
72. Pelosi, P., Iovinella, I., Zhu, J., Wang, G. & Dani, F. R. Beyond chemoreception: diverse tasks of soluble olfactory proteins in insects. *Biol. Rev.* **6**, 1669 (2017).
73. Bucci, B. K., Kruse, S. W., Thode, A. B., Alvarado, S. M. & Jones, D. N. M. Effect of n-alcohols on the structure and stability of the *Drosophila* odorant binding protein LUSH. *Biochem.* **45**, 1693–1701 (2006).
74. Pesenti, M. E. *et al.* Queen bee pheromone binding protein pH-induced domain swapping favors pheromone release. *J. Mol. Biol.* **390**, 981–990 (2009).
75. Pesenti, M. E. *et al.* Structural basis of the honey bee PBP pheromone and pH-induced conformational change. *J. Mol. Biol.* **380**, 158–169 (2008).
76. Xu, X. *et al.* NMR structure of navel orangeworm moth pheromone-binding protein (AtraPBP1): implications for pH-sensitive pheromone detection. *Biochem.* **49**, 1469–1476 (2010).
77. Xu, W., Xu, X., Leal, W. S. & Ames, J. B. Extrusion of the C-terminal helix in navel orangeworm moth pheromone-binding protein (AtraPBP1) controls pheromone binding. *Biochem. Biophys. Res. Commun.* **404**, 335–338 (2011).
78. di Luccio, E., Ishida, Y., Leal, W. S. & Wilson, D. K. Crystallographic observation of pH-induced conformational changes in the Amyeloid transitella pheromone-binding protein AtraPBP1. *PLoS One* **8**, e53840 (2013).
79. Kruse, S. W., Zhao, R., Smith, D. P. & Jones, D. N. Structure of a specific alcohol-binding site defined by the odorant binding protein LUSH from *Drosophila melanogaster*. *Nat. Struct. Biol.* **10**, 694–700 (2003).
80. Lescop, E., Briand, L., Pernollet, J.-C. & Guittet, E. Structural basis of the broad specificity of a general odorant-binding protein from honeybee. *Biochem.* **48**, 2431–2441 (2009).
81. Spinelli, S. *et al.* Crystal structure of *Apis mellifera* OBP14, a C-minus odorant-binding protein, and its complexes with odorant molecules. *Insect Biochem. Mol. Biol.* **42**, 41–50 (2012).
82. Damberger, F. F., Ishida, Y., Leal, W. S. & Wüthrich, K. Structural basis of ligand binding and release in insect pheromone-binding proteins: NMR structure of *Antheraea polyphemus* PBP1 at pH 4.5. *J. Mol. Biol.* **373**, 811–819 (2007).
83. Horst, R. *et al.* NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. *Proc. Natl. Acad. Sci. USA* **98**, 14374–14379 (2001).
84. Katre, U. V., Mazumder, S. & Mohanty, S. Structural Insights into the Ligand Binding and Releasing Mechanism of *Antheraea polyphemus* Pheromone-Binding Protein 1: Role of the C-Terminal Tail. *Biochem.* **52**, 1037–1044 (2013).
85. Ganesan, K., Mendki, M. J., Suryanarayana, M. V. S., Prakash, S. & Malhotra, R. C. Studies of *Aedes aegypti* (Diptera: Culicidae) ovipositional responses to newly identified semiochemicals from conspecific eggs. *Aust. J. Entomol.* **45**, 75–80 (2006).
86. Bosch, O. J., Geier, M. & Boeckh, J. Contribution of fatty acids to olfactory host finding of female *Aedes aegypti*. *Chem. Senses* **25**, 323–330 (2000).
87. Ponnusamy, L. *et al.* Identification of bacteria and bacteria-associated chemical cues that mediate oviposition site preferences by *Aedes aegypti*. *Proc. Natl. Acad. Sci. USA* **105**, 9262–9267 (2008).
88. Prestwich, G. D. Bacterial expression and photoaffinity labeling of a pheromone binding protein. *Prot. Sci.* **2**, 420–428 (1993).
89. Wishart, D. S. *et al.* H-1, C-13 and N-15 Chemical-Shift Referencing in Biomolecular NMR. *J. Biomol. NMR* **6**, 135–140 (1995).
90. Markley, J. L. *et al.* Recommendations for the presentation of NMR structures of proteins and nucleic acids - (IUPAC Recommendations 1998). *Pure Appl. Chem.* **70**, 117–142 (1998).
91. Barna, J. C. J., Laue, E. D., Mayger, M. R., Skilling, J. & Worrall, S. J. P. Exponential sampling, an alternative method for sampling in two-dimensional NMR experiments. *J. Magn. Reson.* **73**, 69–77 (1987).
92. Hyberts, S. G., Takeuchi, K. & Wagner, G. Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data. *J. Am. Chem. Soc.* **132**, 2145–2147 (2010).
93. Hyberts, S. G., Milbradt, A. G., Wagner, A. B., Arthanari, H. & Wagner, G. Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. *J. Biomol. NMR.* **52**, 315–327 (2012).
94. Hyberts, S. G., Arthanari, H., Robson, S. A. & Wagner, G. Perspectives in magnetic resonance: NMR in the post-FFT era. *J. Magn. Reson.* **241**, 60–73 (2014).
95. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR.* **6**, 277–293 (1995).
96. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
97. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1–16 (2013).
98. Schumann, F. H. *et al.* Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions. *J. Biomol. NMR* **39**, 275–289 (2007).
99. Dayie, K. T. & Wagner, G. Relaxation-Rate Measurements for N-15-H-1 Groups with Pulsed-Field Gradients and Preservation of Coherence Pathways. *J. Magn. Reson., Ser. A* **111**, 121–126 (1994).
100. Palmer, A. G. & Massi, F. Characterization of the Dynamics of Biomacromolecules Using Rotating-Frame Spin Relaxation NMR Spectroscopy. *Chem. Rev.* **106**, 1700–1719 (2006).
101. d'Auvergne, E. J. & Gooley, P. R. Set theory formulation of the model-free problem and the diffusion seeded model-free paradigm. *Mol. Biosyst.* **3**, 483–494 (2007).
102. d'Auvergne, E. J. & Gooley, P. R. Model-free model elimination: a new step in the model-free dynamic analysis of NMR relaxation data. *J. Biomol. NMR* **35**, 117–135 (2006).
103. d'Auvergne, E. J. & Gooley, P. R. The use of model selection in the model-free analysis of protein dynamics. *J. Biomol. NMR* **25**, 25–39 (2003).
104. Vögeli, B., Güntert, P. & Riek, R. Multiple-state ensemble structure determination from eNOE spectroscopy. *Mol. Phys.* **111**, 437–454 (2013).

105. Pascal, S. M., Muhandiram, D. R., Yamazaki, T., Formankay, J. D. & Kay, L. E. Simultaneous Acquisition of N-15-Edited and C-13-Edited Noe Spectra of Proteins Dissolved in H₂O. *J. Magn. Reson., Ser. B* **103**, 197–201 (1994).
106. Zwaalen, C. *et al.* Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: Application to a bacteriophage lambda N-peptide/boxB RNA complex. *J. Am. Chem. Soc.* **119**, 6711–6721 (1997).
107. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223 (2009).
108. Born, A. *et al.* Efficient Stereospecific H(beta2/3) NMR Assignment Strategy for Mid-Size Proteins. *Magnetochemistry* **4** (2018).
109. Karplus, M. Vicinal Proton Coupling in Nuclear Magnetic Resonance. *J. Am. Chem. Soc.* **85**, 2870–2871 (1963).
110. Perez, C., Lohr, F., Ruterjans, H. & Schmidt, J. M. Self-consistent Karplus parametrization of 3J couplings depending on the polypeptide side-chain torsion chi1. *J. Am. Chem. Soc.* **123**, 7081–7093 (2001).
111. Linge, J. P., Habeck, M., Rieping, W. & Nilges, M. ARIA: automated NOE assignment and NMR structure calculation. *Bioinforma.* **19**, 315–316 (2003).
112. Linge, J. P., O'Donoghue, S. I. & Nilges, M. Automated assignment of ambiguous nuclear overhauser effects with ARIA. *Methods Enzymol.* **339**, 71–90 (2001).
113. Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
114. Nilges, M. *et al.* Accurate NMR Structures Through Minimization of an Extended Hybrid Energy. *Structure* **16**, 1305–1312 (2008).
115. Mareuil, F., Malliavin, T. E., Nilges, M. & Bardiaux, B. Improved reliability, accuracy and quality in automated NMR structure calculation with ARIA. *J. Biomol. NMR*, 1–14 (2015).
116. AMBER 2018 (University of California, San Francisco, 2018).
117. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D. Biol. Crystallogr* **66**, 133–144 (2010).
118. Kabsch, W. Xds. *Acta Crystallogr. D. Biol. Crystallogr* **66**, 125–132 (2010).
119. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D. Struct. Biol.* **75**, 861–877 (2019).
120. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr* **60**, 2126–2132 (2004).
121. Brünger, A. T. Extension of molecular replacement: A new strategy based on Patterson correlation refinement. *Acta Crystallogr.* **A46**, 472–475 (1990).
122. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr. D. Biol. Crystallogr* **62**, 859–866 (2006).
123. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D.* **53**, 240–255 (1997).
124. Bailey, S. The CCP4 Suite - Programs for Protein Crystallography. *Acta Crystallogr. D.* **50**, 760–763 (1994).
125. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr* **67**, 235–242 (2011).

Acknowledgements

This work was supported by NIH/NIAID grant RO1 AI121253 to DNMJ. The operation of the Structural Biology facilities at CU School of Medicine is supported in part by the University of Colorado Cancer Center Support Grant (NIH 5P30-CA046934). We thank Dr George Dimopoulos for AeOBP22 cDNA and discussions.

Author contributions

D.N.M.J. conceived the experiments; J.W. and E.J.M. performed experiments; J.W., E.J.M., J.C.N. and D.N.M.J. all collected and analyzed X-ray data; J.W., E.J.M. and D.N.M.J. collected and analyzed the N.M.R. data; and J.W. and D.N.M.J. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-60242-9>.

Correspondence and requests for materials should be addressed to D.N.M.J.

Reprints and permissions information is available at 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020