

SCIENTIFIC REPORTS



OPEN

Effects of point mutations in the binding pocket of the mouse major urinary protein MUP20 on ligand affinity and specificity

Jimena Ricatti^{1,2}, Laura Acquasaliente³, Giovanni Ribaudò³, Vincenzo De Filippis³, Marino Bellini³, Ramiro Esteban Llovera⁴, Susi Barollo⁵, Raffaele Pezzani⁵, Giuseppe Zagotto³, Krishna C. Persaud⁶ & Carla Mucignat-Caretta^{1,7}

The mouse Major Urinary Proteins (MUPs) contain a conserved β -barrel structure with a characteristic central hydrophobic pocket that binds a variety of volatile compounds. After release of urine, these molecules are slowly emitted in the environment where they play an important role in chemical communication. MUPs are highly polymorphic and conformationally stable. They may be of interest in the construction of biosensor arrays capable of detection of a broad range of analytes. In this work, 14 critical amino acids in the binding pocket involved in ligand interactions were identified in MUP20 using *in silico* techniques and 7 MUP20 mutants were synthesised and characterised to produce a set of proteins with diverse ligand binding profiles to structurally different ligands. A single amino acid substitution in the binding pocket can dramatically change the MUPs binding affinity and ligand specificity. These results have great potential for the design of new biosensor and gas-sensor recognition elements.

Major urinary proteins (MUPs) in the mouse belong to the lipocalin family and contain a conserved β -barrel structure with a characteristic central hydrophobic pocket^{1,2}. They are mainly synthesised in the liver and excreted into the urine³. The tertiary structure of MUPs consists of eight β -strands, arranged in an anti-parallel β -barrel open on one side, with α -helices at both ends, encompassing a cup-like binding pocket (Fig. 1). They bind volatile pheromones or other lipophilic molecules^{4,5}, are transported across the kidney, and excreted in the urine⁶. Mice deliberately deposit numerous small urine drops, that act as scent marks. Once deposited in urine marks, MUPs release bound volatile compounds slowly into the air. These are highly important for chemical communication^{7,8} and indeed mice recognise scent marks by their molecular composition⁹.

The stability of lipocalins, like vertebrate Odorant Binding Proteins (OBP), against proteolysis, temperature, and solvents make them attractive for development of biosensors¹⁰. Recent work has shown that OBPs can be successfully integrated into a variety of transducers, including quartz crystal microbalances, surface acoustic wave devices, organic field effect transistors^{11–15}, to produce gas sensors. If the structures of the binding sites can be tailored to produce proteins capable of sensitive and selective detection of ligands for specific applications, then this would open a new generation of gas sensing technology. MUPs are interesting candidates for this purpose. They are highly polymorphic, and the relative abundance of different isoforms in urine is thought to function as individual identity signatures of mice^{16,17}. However, the relationship between extreme polymorphism and chemical communication is not clear¹⁸. There are 21 MUPs isoforms characterized by small differences in amino acid sequences, highly conserved biochemical properties and tertiary folding features. The sites of variation are found

¹Department of Molecular Medicine, University of Padua, Padua, Italy. ²Cell Biology and Neuroscience Institute, University of Buenos Aires—National Scientific and Technical Council (UBA-CONICET), Buenos Aires, Argentina.

³Department of Pharmaceutical and Pharmacological Sciences, University of Padua, Padua, Italy. ⁴Multidisciplinary Institute of Cell Biology, National Scientific and Technical Council (CONICET) and Department of Science and Technology, National University of Quilmes, Buenos Aires, Argentina. ⁵Department of Medicine, University of Padua, Padua, Italy. ⁶School of Chemical Engineering and Analytical Science, University of Manchester, Manchester, UK.

⁷National Institute of Biostructures and Biosystems, Rome, Italy. Correspondence and requests for materials should be addressed to C.M.-C. (email: carla.mucignat@unipd.it)

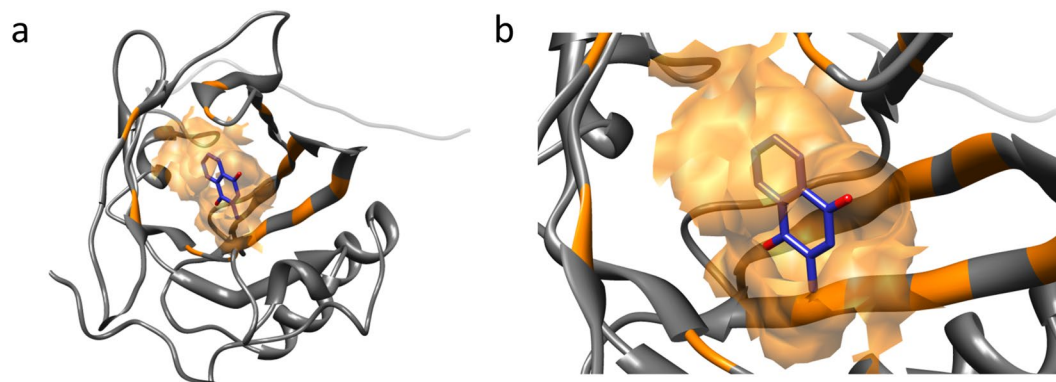


Figure 1. Three-dimensional structure of murine MUP20 with menadione inside the binding cavity. **(a)** Inside the binding pocket, the ligand structure is depicted in blue with oxygen in red; the binding cavity surface and the interacting residues are highlighted in orange. The ligand is buried within the binding cavity and appears partially covered by the binding surface. **(b)** A closer view of the binding pocket reveals that the aromatic ring and the methyl group of menadione are inside the cavity. Docking was performed with Autodock/Vina^{42,43}, starting from the first conformer of the NMR structure of MUP20 (2L9C.pdb).

Physico-chemical properties of selected MUP ligands.				
Ligand	Molecular mass (a.m.u.)	Molecular Volume (Å ³) ^a	LogP ^a	ZINC Accession Number ^b
L-Adrenaline (L-ADR)	183.20	171	−0.82	ZINC00039089
Methylnaphthalene-1,4-dione (Menadione, MEN)	172.18	155	1.91	ZINC00001677
2,4-Dimethylphenol (2,4-DMP)	122.16	125	2.37	ZINC01672873
2-Butyl-1-octanol (OCT)	186.33	222	5.36	ZINC1640892
Linalool (LIN)	154.25	176	2.68	ZINC01529819
2-Isobutyl-3-methoxypyrazine (PYR)	166.22	168	2.11	ZINC00156517

Table 1. List of ligands used for docking simulations with MUP20 and their physico-chemical properties. The molecular mass values, the molecular volumes and the octanol/water partition coefficients (LogP) are reported along with the ZINC accession. ^aValues of molecular volume and LogP were determined using the online Molinspiration software version 2011.06 (www.molinspiration.com) and ALGOPS2.1 program⁷¹. ^bAccession number of the listed molecules in the ZINC data base⁶⁹.

only on a restricted segment of the polypeptide chain, which projects to a patch on the surface of the protein¹⁶. Each isoform displays a unique spectrum of binding specificities across a variety of ligands¹⁹. These proteins and their volatile ligands are detected by the main and accessory olfactory systems of the mouse and trigger adaptive behavioural responses and/or physiological processes^{8,20}. Apart from the release of volatiles into the environment, there is evidence that the proteins themselves may also function as chemosignalling molecules^{21–23}, supporting the segregation of the *Mus* genus species²⁴.

Of the MUP proteins, MUP20 is also known as Darcin for its peculiar expression in male mice and for its primary role in social communication^{25,26}, but there is limited information available about the critical amino acid residues of the binding pocket involved in interactions with the ligand. We utilised this protein for our work. The approach taken here was to use known structures of MUPs to identify the critical amino acids in the binding pocket that interact with ligands. A subset of possible mutant proteins was identified, that could form an array of proteins with diverse ligand binding characteristics. From these studies, an interesting outlook of artificial sensing systems for biotechnological applications of MUPs is now emerging^{10,27}.

Results

Selection of MUP20 ligands. While the lock-and-key model may be inadequate, bioisosteric substitutions may be relevant for exploring odorant-protein interactions²⁸. Here, six chemically unrelated ligands, some of which display odorant/pheromone characteristics^{9,25,29}, were selected for their different physico-chemical properties, i.e. molecular size, chemical structure, hydrophilic/hydrophobic balance, and hydrogen bonding properties (Table 1). In particular, 2-isobutyl-3-methoxypyrazine (PYR) is a low-threshold odor^{30,31}, used in MUPs/ligand binding studies^{4,7}, methylnaphthalene-1,4-dione (menadione, MEN), a naphthoquinone analogue of vitamin K is also found in MUP scent marks³² while 2,4-dimethylphenol (2,4-DMP), and 2-butyl-1-octanol (OCT), along with linalool (3,7-dimethylocta-1,6-dien-3-ol) (LIN), contribute to urine scent marks in male mice^{9,33}. L-Adrenaline (L-ADR) was also tested as a precursor of vanilloid ligands, which may bind to Transient Receptor Potential Vanilloid (TRPV) receptors widely found in the olfactory epithelium³⁴.

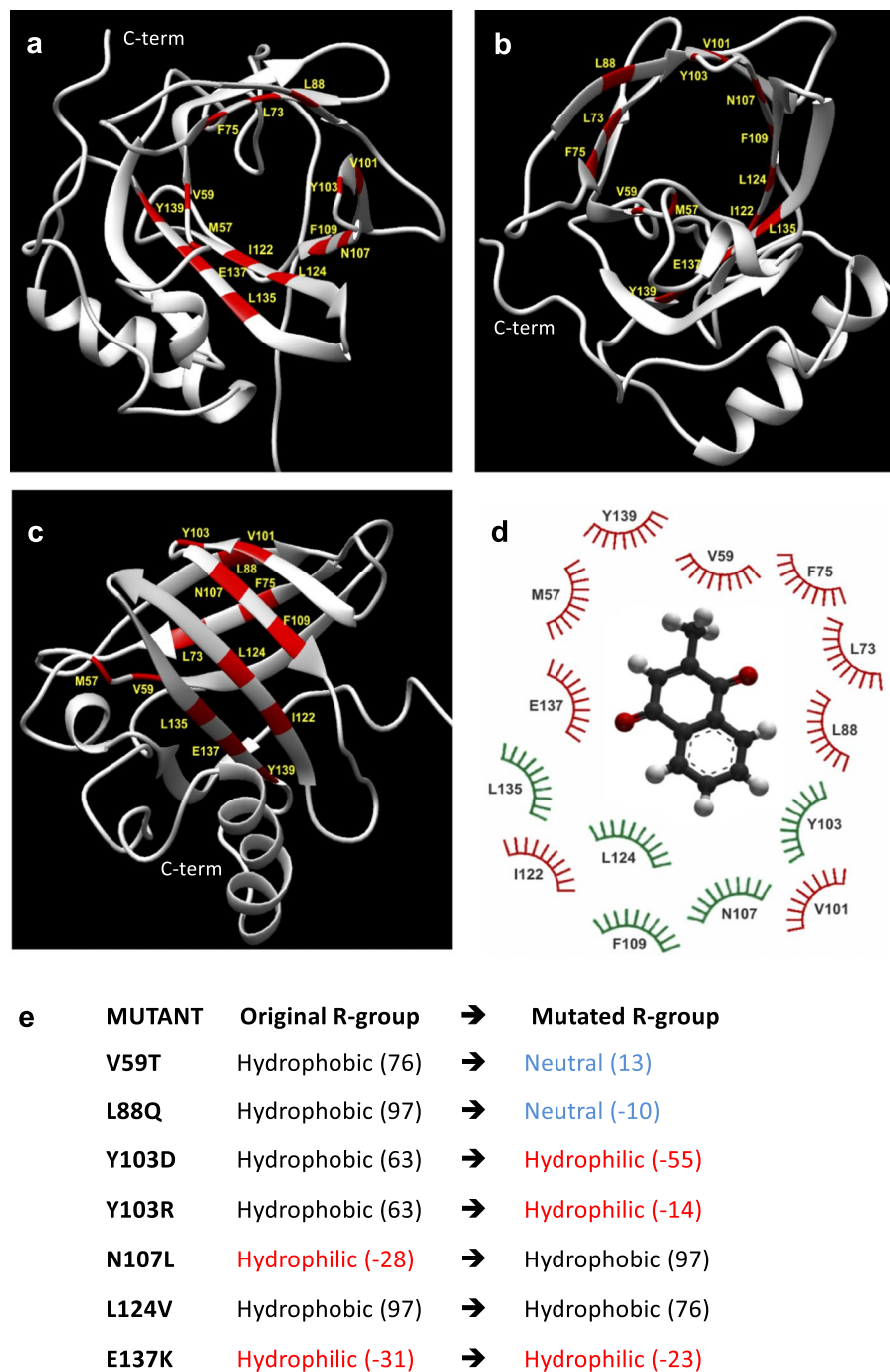


Figure 2. (a–c) The structure of wild type MUP20 (PDB ID: 2L9C, 6xHis tag is not represented) in different orientations. The 3D representation shows in red the amino acids belonging to the hydrophobic pocket involved in the binding process, according to the LPC/CSU analysis. C-term indicates C-terminus. (d) Relevant residues involved in the ligand-MUP20 interaction. Those in green are conserved across the natural MUP isoforms, while the variable residues are in red. (e) Amino acid mutations in MUP20. The polar/apolar nature of the original and substituted amino acids is reported together with the corresponding hydrophobicity index values (in brackets)³⁶.

Selection of MUP20 mutations. Figure 1 illustrates the tertiary structure of MUP20 (PDB ID: 2L9C.pdb). The ligand-binding pocket in MUP20 is shaped by a region containing mostly hydrophobic amino acids (1 Ile, 2 Phe, 4 Leu, 2 Val, and 1 Met) lining a cavity of approximately 435 Å³ (Fig. 2a–c). Interestingly, the polar –OH groups of two Tyr side-chains (103 and 139) protrude into the apolar cavity surface of MUP20, along with the hydrophilic side chain of Asn107 and the negatively charged Glu137, which in the crystallographic structure of the sequence/structural homologous MUP4 (PDB ID: 3KFF) is partially compensated by two structural water molecules³⁵.

Residue position	Distance ^a	Surface ^b	Hphi-Hphi ^c	A-A ^d	Hpho-Hpho ^e	Hpho-Hphi ^f
M57	5.5 Å	11.7 Å ²	—	—	+	—
V59	4.8 Å	24.0 Å ²	—	—	+	—
L73	3.3 Å	36.8 Å ²	—	—	+	—
F75	5.0 Å	19.1 Å ²	—	—	+	—
L88	3.6 Å	26.7 Å ²	—	—	+	—
V101	4.9 Å	7.2 Å ²	—	—	+	—
Y103	3.4 Å	41.5 Å ²	—	—	+	—
N107	4.7 Å	5.8 Å ²	—	—	—	+
F109	3.7 Å	35.9 Å ²	—	—	+	—
I122	3.9 Å	22.2 Å ²	—	—	+	—
L124	3.4 Å	20.3 Å ²	—	—	+	+
L135	3.8 Å	13.0 Å ²	—	—	+	—
E137	2.5 Å	46.3 Å ²	+	—	—	+
Y139	3.4 Å	14.3 Å ²	+	—	—	—

Table 2. LPC/CSU analysis of the interactions of the amino acids in the ligand-binding pocket of MUP20 with menadione. ^aThe closest measured distance (Å) between atoms of the ligand and the residue. ^bContact surface area (Å²) between the ligand and the residue. ^cHydrophilic-hydrophilic contact. ^dAromatic-aromatic contact. ^eHydrophobic-hydrophobic contact. ^fHydrophobic-hydrophilic contact.

From the computational analysis of interatomic contacts on digital models of MUP20 mutants docked with menadione, a reference ligand for MUPs³², fourteen amino acids localised in the binding pocket were identified as the most relevant for interactions with ligands (Table 2, Fig. 2d). Notably, the residues shown in green are conserved across all natural MUP isoforms. These were ranked on the basis of minimum distance existing between the amino acids in the binding pocket and the ligands, as well as according to the greater contact surface. The calculated ligand-MUP distances obtained from the LPC/CSU analysis ranged from 2.5 to 5.5 Å, while the contact surface area was estimated in the 5.8–46.3 Å² range (Table 2). From these data, six critical residues in the binding region were selected for substitution in order to produce an effective mutation that would change the binding characteristics of MUP20 (Fig. 2e), while retaining the structural stability of the protein. The selected amino acids included strictly conserved residues in MUP20, like Tyr103, Asn107, and Leu124, as well as partially conserved residues, like Val59, Leu88, and Glu137 (Fig. 2d). The substituent amino acids were selected according to their size and hydrophobicity index³⁶ and mutations were designed to mainly alter the hydrophilic/hydrophobic balance at the mutation site, without significantly perturbing the conformational and stability properties of MUP20 (see Supplementary S1 for 3D models and Root-mean-square deviation of atomic positions, RMSD, data).

The amino acid substitutions shown in Fig. 2e were also intended to exaggerate the effect of variations found in natural MUP, within the constraints of also maintaining conformational stability of the protein. Although V59 is not strictly conserved in MUP family, possibly replaced by more hydrophobic or less hydrophobic amino acids like Leu and Ala, respectively, the naturally occurring residues at position 59 are all hydrophobic in nature. Hence, we decided to perturb position 59 by replacing Val with Thr, which retains the conformational properties of Val, whereby both residues are strong β -sheet stabilizers. However, due to the presence of the –OH group, Thr is much less hydrophobic than Val (see Table 1)³⁶. At position 88, Leu, Ala or Met are observed in natural MUP isoforms and even in this case all these residues are highly hydrophobic. Here Leu88 was replaced by Gln, which has a flexible and polar side chain amenable to form stable hydrogen bonds. At position 137 Glu and Gly residues are frequently observed in MUP family, where Glu has a negatively charged and flexible side-chain while Gly is neutral and dramatically increases the conformational flexibility of the backbone³⁷. Here, we decided to alter the charge and conformational properties at position 137 by replacing Glu with Lys, which has a long/flexible and positively charged side chain. Among the naturally conserved amino acids, Leu124 was replaced with Val, where Val-side chain retains the apolar character of Leu but has a smaller size and different conformational properties. Actually, Leu is one of the best helix-stabilizers among the natural amino acids, whereas Val is a helix-breaker and stabilizes the β -sheet secondary structure. Next, drastic chemical variations were introduced at position 103, where the rigid aromatic side-chain of the Tyr-residue in wild-type MUP20 was replaced by Arg, having a long and flexible positively charged side chain. To test the role of the electrostatic charge at position 103, Tyr103 was also replaced by Asp, which is negatively charged. Finally, the neutral and polar side chain of Asn107 was substituted with the apolar side chain of Leu.

The selected ligands have different chemical features such as steric hindrance and hydrophilic/hydrophobic properties (see Table 1).

The theoretical binding profiles of these ligands to the seven selected MUP20 mutants are depicted in Fig. 3. Predicted ligand affinities to MUP mutants are expressed as theoretical free energy change of binding (ΔG_b)³⁸ and are the result of blind docking simulations. Accordingly, the wild type MUP20 (WT) displayed binding selectivity for MEN and to a lesser extent for 2,4-DMP and this binding profile was roughly retained for V59T, L88Q, L124V, N107L and, rather surprisingly, even for the mutant carrying the charge reversal E137K. The mutant L124V showed the broadest specificity profile, compared to other mutants, whereas Y103R and Y103D displayed a strong reduction in the ability to recognize the selected ligands, such that these mutants were predicted to bind

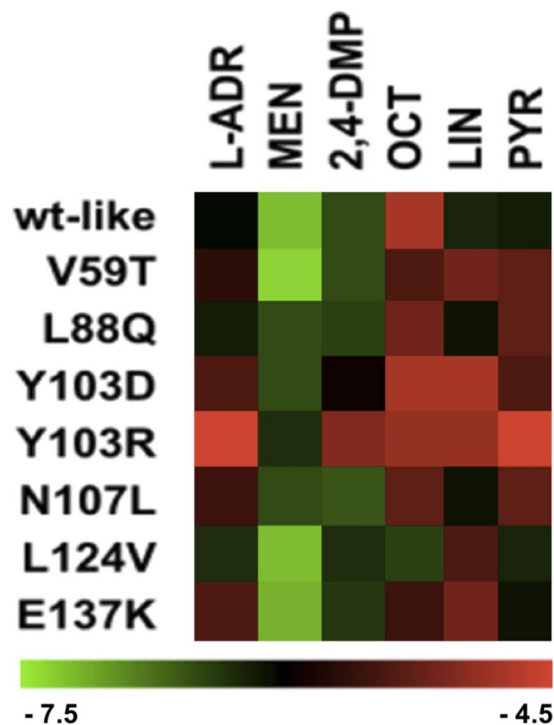


Figure 3. *In silico* binding analysis. Theoretical binding energies (ΔG_b) (kcal/mol) of the selected ligands to wild type (WT) MUP20 and mutant models, plotted as a heatmap. The change of ΔG_b values across mutants, and between different ligands are represented from light green to light red. Refer to the Supplementary File S1 for ΔG_b values.

solely to MEN. According to our *in silico* experiments, while most of the ligands cluster in the internal MUP20 pocket, a limited number of interacting molecules may also bind to an external portion of the protein in some of the models (see Supplementary File S1 for complete 2D interaction patterns).

Production, purification and characterization of MUP20 mutants. To test the working hypotheses that emerged from the *in silico* analysis, wild type MUP20 and the seven selected MUP20 mutants were expressed in *E. coli* as 6xHis-tagged N-terminal fusion proteins, as detailed in the Methods section, and further characterized with respect to their chemical identity, conformational and ligand-binding properties. Recombinant MUPs were purified to homogeneity by IMAC, as documented by reducing SDS-PAGE and Coomassie staining, where MUP proteins migrated as a single band (>85%) with an apparent molecular weight of ~20 kDa (Supplementary Fig. S2b). The homogeneity of recombinant MUPs preparations was further checked by RP-HPLC (Fig. 4a), while high-resolution mass spectrometry analyses yielded mass values in agreement with those deduced from the amino acid composition of MUP analogues within 50 ppm mass accuracy (Supplementary Material S3), thus providing clear-cut evidence for the chemical identity of mutant MUPs. To verify that point mutations did not appreciably affect the conformation of MUP20, circular dichroism CD spectra were recorded in the far-UV region (Fig. 4b), a spectroscopic technique that provides key information on the protein secondary structure content. The CD spectrum of recombinant wild-type MUP20 displays a single minimum centered at about 218 nm, reflecting the high β -sheet content of MUP family members¹. Notably, the spectra of MUP20 mutants are almost superimposable to that of the wild-type protein, thus suggesting that the single point mutations inserted into MUP20 sequence do not alter the overall MUP structure. Importantly, the latter result allows us to interpret ligand binding data solely on the basis of the different physico-chemical properties of the ligands and/or of the mutated amino acid side chain.

Fluorescence and calorimetric ligand binding measurements to MUP20 mutants. The affinity of the six ligands reported in Table 1 for wild-type and seven mutant MUPs was first determined by fluorescence competition experiments, using 1-NPN as a fluorescent MUP ligand^{39,40}. Using this method, rough estimates of ligand affinity for MUPs were obtained and graphically represented in Fig. 5 as radar plot and heatmap. Wild-type MUP20 binds to the selected ligands with a K_d range of 0.64 to 5.36 μ M, with the highest affinity to menadione, followed by PYR and OCT (Supplementary Materials S4 and S5). The heatmap indicates that each mutant had a different selectivity profile, with MUP20 L124V showing less specificity in binding to all the selected ligands, and Y103R showing the highest selectivity. Intriguingly, N107L mutant displayed a binding profile remarkably different from that of the wild-type MUP20.

The binding energetics of the best ligands, previously identified in fluorescence binding measurements (see Fig. 5), i.e. 2-isobutyl-3-methoxy pyrazine (PYR) and menadione (MEN), to wild-type (WT) and seven mutant MUPs was studied by isothermal titration calorimetry (ITC). This gave an independent analysis of binding without the constraints of ligand displacement experiments, i.e. linearity of fluorescence changes and possible

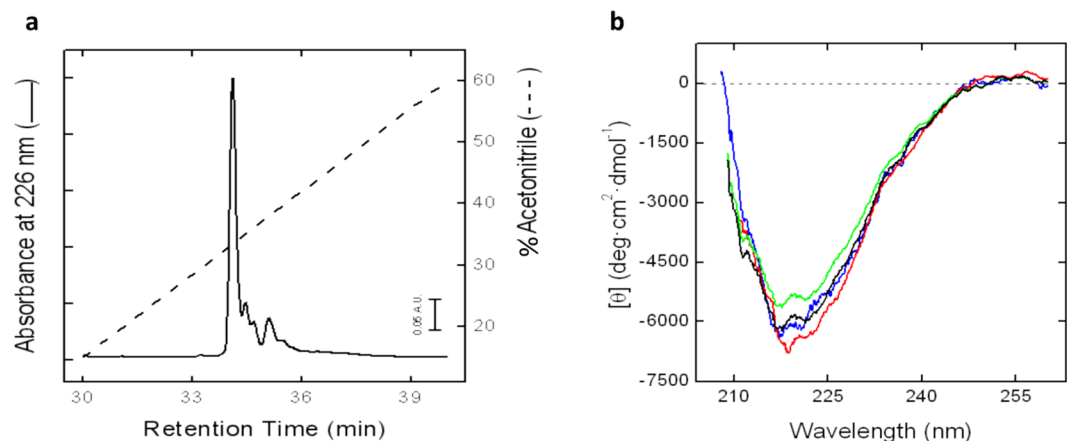


Figure 4. Characterization of recombinant MUP20 proteins. **(a)** RP-HPLC analysis of recombinant wild type MUP. An aliquot (20 μ g) of protein was applied to a Vydac C4 analytical column eluted with a linear acetonitrile-0.1% TFA gradient (—) at a flow rate of 0.8 ml/min. The protein material corresponding to the major peak in the chromatograms was collected and subjected to MS analysis (Supplementary Material S3). **(b)** Far-UV CD spectra of wild-type and mutant MUPs. CD spectra of recombinant wild type MUP (black line), and MUP mutants: L88Q (blue line), L124V (red line) and Y103R (green line). All measurements were carried out at 25 ± 0.1 °C 10 mM Na_2HPO_4 pH 7.4, 0.15 M NaCl, at a protein concentration of 0.1 mg/ml.

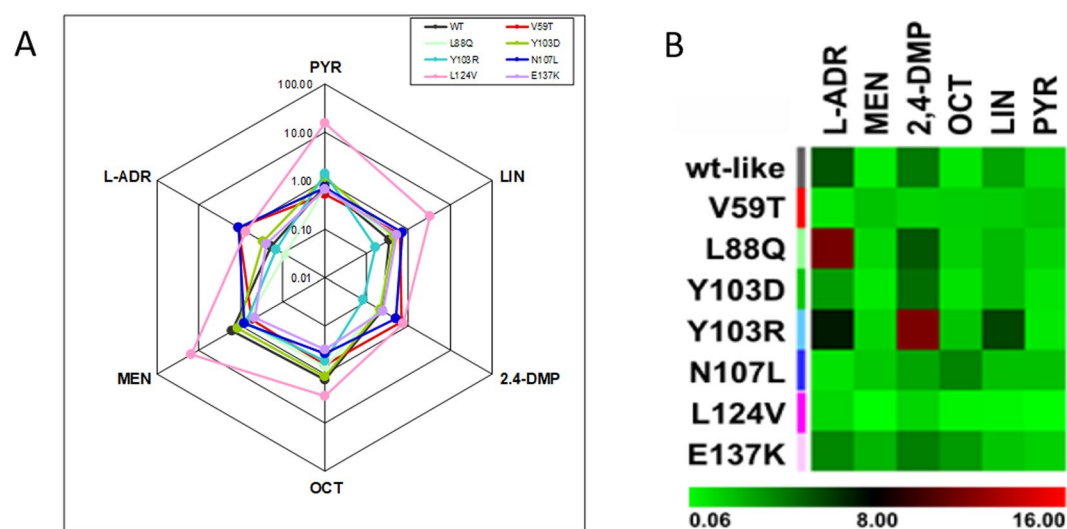


Figure 5. *In vitro* binding analysis. **(A)** Radar plot of the binding profile for displacement of the fluorescent ligand 1-NPN from MUP mutants. The equilibrium association constants, K_a , relative to the binding of each ligand to each mutant MUP are plotted on a logarithmic scale. **(B)** The equilibrium dissociation constants, K_d , are plotted as a heatmap, displaying the variations across mutants, and between different ligands binding to the same mutant. The change of K_d values are represented from light green to light red and the numbers refer to the range of dissociation constants (μ M).

unspecific interactions with 1-NPN³⁹. The titration of wild type MUP20 and mutants L88Q and L124V gave an exothermic reaction with both PYR and MEN (Fig. 6A,B), approaching saturation at high ligand/MUP molar ratio, yielding a ligand:MUP binding stoichiometry close to 1. Notably, the binding strength of the same ligands to Y103R mutant (Fig. 6A,B panel d) was barely detectable.

ITC measurements indicate that PYR binds to the wild type MUP20 with an affinity >40 fold higher than MEN: this may be due to the nature of the pyrazine ring of PYR which can form more hydrogen bonds than the naphthoquinone scaffold of MEN, when inserted into the binding pocket of MUP20. The replacement of L88 with a linear and more polar amino acid such as Gln did not appreciably affect the affinity for PYR, whereas the same L88Q mutation increased the affinity for MEN by about 6-fold. On the other hand, the mutation of L124 with the smaller and hydrophobic valine residue resulted in a drop of the affinity for PYR and MEN by 1300- and 25-fold, respectively, compared to the wild-type MUP20. Notably, the L124V mutant displayed identical affinity, within the limits of the experimental error, for both PYR and MEN.

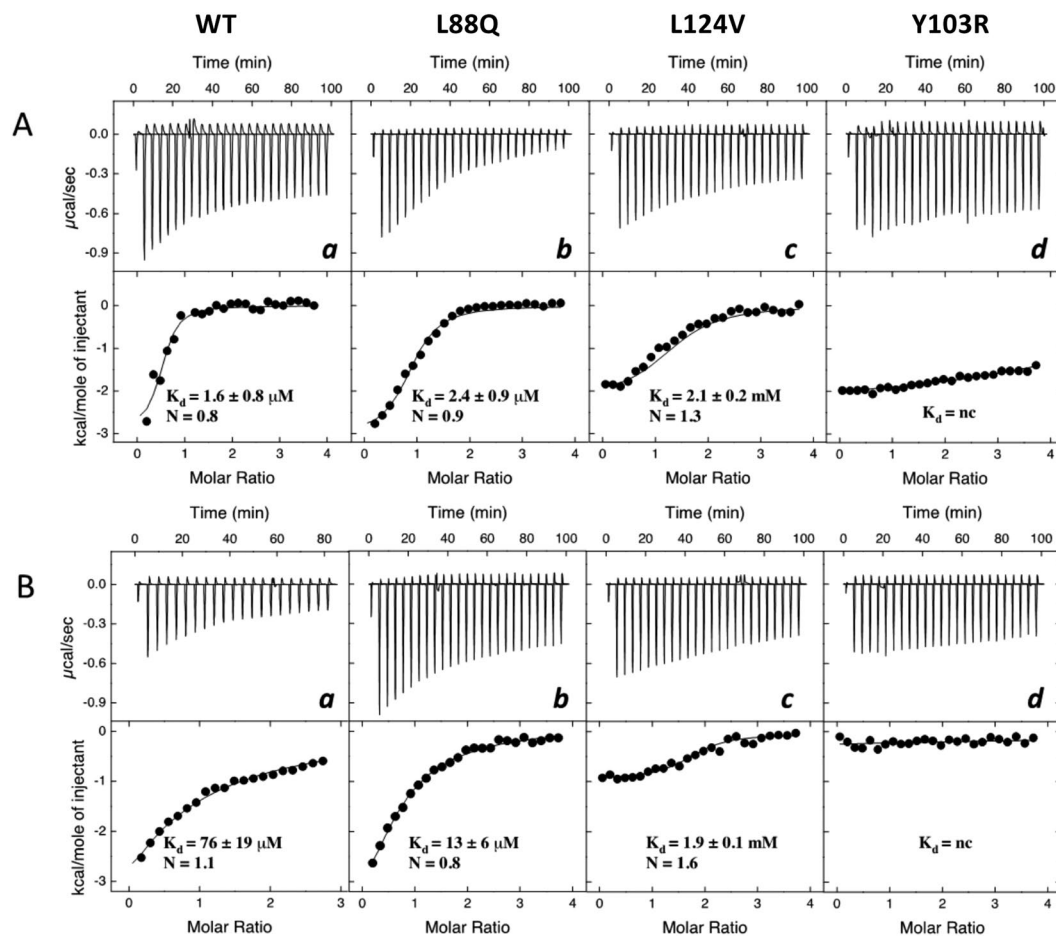


Figure 6. ITC binding measurements. Binding of 2-isobutyl-3-methoxy pyrazine (PYR) and menadione (MEN) to wild type and mutant MUPs. ITC data for PYR (A) and MEN (B) binding to recombinant wild type MUP (WT) (a), and mutants L88Q (b), L124V (c) and Y103R (d). Raw injection heats (expressed as differential power) are shown in the top panels and the corresponding specific binding isotherms, calculated from the integrated injection heats and normalized to the moles of injectant, are shown in the bottom panels. The equilibrium dissociation constants (K_d) and stoichiometry values (N) were determined by fitting the data to a one-site binding model.

Discussion

The phylogenetic framework of chemical communication of mice is complex. Different MUP isoforms may trigger a variety of behavioural and physiological effects. MUP bind a variety of small molecules^{7,41} which may be used for communication among mice, suggesting that they may bind other molecules of interest. However, it is still not clear how variations in amino acid sequence across isoforms that are minimal in many cases, may affect binding of ligands or their release to the environment. Here, we establish the basis for understanding how a substitution of one critical amino acid in the binding pocket can produce a large change in selectivity to volatile ligands that may be bound, as well as their affinities.

In this study, we used the solution structure of MUP20 to digitally dock MEN, as a prototype of MUP ligands (Fig. 1), and identify the potential hot spots on MUP20 structure important for ligand binding. The changes in binding affinity of WT and mutant MUPs for the six chemically unrelated ligands were first estimated by docking simulations, expressed as free energy change of binding (ΔG_b) (Fig. 3). Starting from the results of simulations, seven point-mutants of MUP20 were expressed, purified and thoroughly characterized with respect to their purity, chemical identity, and conformational properties. Thereafter, experimental values of the dissociation constants (K_d) of each mutant MUP for the selected ligands were determined by fluorescence competition measurements using 1-NPN as a fluorescent MUP ligand (Fig. 5). The complex figure that emerges from the affinity data is that each mutant MUP has a different binding selectivity profile and that a given mutation can perturb binding of a specific ligand without substantially affecting the affinity of the others. The comparison of the theoretical and experimental binding data (Figs 3 and 5) highlights the partial predictive power of our theoretical approach. The docking protocol used in this study assumes that (i) ligand-receptor interaction follows a rigid-body mechanism, (ii) the contribution of desolvation is negligible, and (iii) the receptor structure remains unchanged upon mutation^{42,43}. These assumptions can hardly apply to ligand-MUP binding: the high conformational flexibility of MUPs, their structural plasticity in ligand binding, and the key role that solvation water molecules play in

molecular recognition of MUPs are all implicated in the structure-function relationship^{44–50}. All these contributions make difficult to accurately predict the effect of mutations in MUP binding. Notwithstanding, docking simulations have been a useful tool in guiding the selection of potential mutants for expression.

Considered the complex nature of ligand-MUP interaction and the moderate/low expression yields of mutant MUPs, we used ITC to measure the affinity of the best ligands identified in competition experiments (i.e. MEN and PYR) to WT and some selected mutant MUPs (i.e. L88Q, Y103R, and L124V), being ITC the gold standard technique for studying ligand-receptor interaction in a label-free system⁵¹. From the thermograms in Fig. 6 it was possible to extract K_d values for the binding of PYR to WT ($K_d = 1.6 \mu\text{M}$) and L88Q ($2.4 \mu\text{M}$), which were in excellent agreement with the affinities deduced from fluorescence competition experiments, yielding $K_d(\text{WT-PYR}) = 1.3 \mu\text{M}$ and $K_d(\text{L88Q-PYR}) = 1.4 \mu\text{M}$. ITC data, however, indicate that the replacement of Leu124 with Val leads to a drop in the affinity for both PYR and MEN. Likewise, the substitution of the aromatic Tyr103 with the long, positively charged Arg-residue almost abolished binding of both ligands tested. Intriguingly, fluorescence measurements were able to pinpoint this dramatic drop in the affinity of Y103R mutant only for the binding to L-ADR, 2,4-DMP, and LIN. These apparent discrepancies between fluorescence and ITC measurements can be reconciled considering that, contrary to ITC which directly measures the heat exchange of ligand binding, the spectroscopic method used here to estimate ligand-MUP affinities is actually an indirect procedure relying on the decrease of the fluorescence intensity upon displacement of the MUP-bound fluorophore 1-NPN, when increasing concentrations of a competing ligand in the solution. Noteworthy, the results of competition binding measurements are affected by poor linearity of the spectroscopic signal, possibly caused by fluorescence inner filter effects⁵², along with the unique binding selectivity profile of 1-NPN for the mutant MUPs. Furthermore, an increase (instead of a decrease) of fluorescence intensity may be observed with some ligands, as a result of unpredictable unspecific interaction of the added ligand with the apolar 1-NPN fluorophore either in the free or MUP-bound state³⁹.

Beyond comparing the accuracy of fluorescence and calorimetric techniques, the data in Fig. 6 indicate that position 88 in MUP20 structure is quite tolerant to mutation, as the substitution of the apolar Leu with the polar Gln-residue does not alter the affinity of L88Q for PYR and even improved the binding strength for MEN by 6-fold compared to WT. Leu88 is not conserved across the MUP family members (Fig. 2d). Therefore, our affinity data are consistent with the general trend that mutations at evolutionary non-conserved positions usually do not significantly alter neither protein structure nor function, whereas even subtle changes at conserved sites dramatically affect structure, stability and function⁵³. Hence, the replacement of the highly conserved Leu124 with Val dramatically decreased the affinity for both the ligands tested. The lack of the methylene-group, upon Leu \rightarrow Val exchange, may alter the physico-chemical forces driving ligand-MUP interaction and may stabilize the apolar pocket of MUP20 into a conformation unproductive for ligand binding. This interpretation is realistic considering that the energetics of ligand-MUP interaction is intrinsically complex and contributed by several different interactions, including hydrophobic desolvation of both the protein and the ligand, formation of a buried water-mediated hydrogen bond network between the protein and ligand, formation of strong van der Waals interactions, and changes in the structure, dynamics, and/or hydration of the protein upon binding^{44–50}. Moreover, ligand-MUP interactions follow an induced-fit mechanism, whereby the apolar binding cavity is more occluded in the ligand-free protein than in the ligand-bound structure². The interaction pocket in MUPs is optimised for a variety of ligand binding modes, thus accounting for the known promiscuity of MUPs in ligand binding⁵⁴. A further complication in ligand-MUP interaction stems from NMR data showing that, counterintuitively, the backbone flexibility of MUP slightly increases upon ligand binding⁵⁵. The combination of opposing entropic and enthalpic effects make difficult to interpret binding data to MUPs on rigorous structural grounds. In the case of Leu124Val mutation, it is likely that even a small reduction in the amino acid side-chain volume of Leu124 ($\text{vol} = 124 \text{ \AA}^3$), after mutation with Val ($\text{vol} = 105 \text{ \AA}^3$), creates an uncompensated cavity in the binding pocket of MUP20 that induces a rearrangement of the surrounding amino acids to a conformation that disfavours ligand binding. Actually, far-UV circular dichroism data indicate that all the mutant MUPs retain the overall secondary structure of the wild type MUP20 (Fig. 4b). Nevertheless, there may be subtle but functionally important changes in the local structure and dynamics of the binding pocket which cannot be detected by far-UV CD measurements⁵⁶. This would indicate that there is great scope for MUP analogues to have a different binding selectivity profile depending on small, elusive structural changes in the apolar cavity. Lastly, the replacement of the conserved Tyr103 with Arg at the entrance of the binding cavity almost abrogates affinity for the ligand molecules tested (Fig. 6). Previous work from different laboratories have shown that the presence of an aromatic amino acid, like Tyr and Phe, is required to stabilize the binding of natural ligands to MUPs and other lipocalins important for chemical communication in mice and rats^{32,57} and also in other lipocalins like the Odorant Binding Proteins⁵⁸. The dramatic effect of Tyr \rightarrow Arg exchange might be expected considering the drastic structural change associated to the mutation of the aromatic and rigid Tyr-residue with the long/flexible and positive Arg side chain, protruding into and limiting the access to the apolar cavity of MUP20.

The scope for this work widens to the biotechnological sector where MUP may be designed to bind a variety of non-natural ligands that can be used as recognition sites for biosensors or gas sensors¹⁰. Because of their conformational stability, as well as resistance to degradation in the environment, this is opening a new pathway to designing biohybrid sensors capable of detecting a variety of analytes²⁷. It is shown here that a single substitution of an amino acid involved in the binding process can radically change the MUP binding ability and ligand specificity. Our results pave the way to express different MUPs significantly different in sensitivity and selectivity for binding to a variety of volatile ligands, so that large arrays of proteins can be constructed, immobilised on to appropriate transducers to produce artificial sensing systems that are robust and suitable for applications such as environmental monitoring, medical diagnostics and food quality evaluation. Other researchers are actively pursuing the expression of olfactory receptor proteins and interfacing of nanoelectronic devices to produce promising chemical sensors^{59–63}. However, because these proteins are membrane bound proteins, it has been difficult to

retain the conformational stability needed to create robust sensors. The use of soluble and stable biorecognition elements such as MUPs provides a new tool for development of new chemical sensors.

Methods

Protein-Ligand Interactions. The NMR solution structure of MUP20 isoform²⁵ (PDB ID: 2L9C) in the ligand-free state was used as the starting point for preliminary ligand docking simulations, which were run using the LPC/CSU Server – Weizmann AC and SwissDock (<http://www.swissdock.ch>). Prior to docking, the key residues putatively important for ligand binding were identified starting from the three-dimensional structure of the complex of MUP4 with 2-sec-butyl-4,5-dihydrothiazole (PDB ID: 3KFF)³⁵. MUP4 shares high sequence homology (91% homology, 81% identity) and structural similarity with MUP20, especially in the ligand-binding pocket, where 16 out of 18 amino acids are conserved in the two proteins (the two different amino acids being L73 → S in MUP4 and E122 → F in MUP4). This information was transferred to the MUP20 structure. The Ligand-Protein Contacts and Contacts of Structural Units (LPC/CSU)⁶⁴ were obtained for each atom–atom contact and ranked selecting those residues represented by nearest distance (Å) between atoms of the ligand and the MUP amino acids and largest contact surface (Å²)³⁸. Potentially destabilising contacts were also identified. Residues with significant contact were selected for mutation in order to modulate the affinity of MUP for selected ligands. Values of Molecular Volume and LogP were determined using the online Molinspiration software version 2011.06 (www.molinspiration.com) and ALGOPS2.1 software.

Selection of single point mutations in MUP20. A total of six amino acids (V59, L88, Y103, N107, L124, E137) with strong atomic interactions within protein-ligand complex were selected. Of these, 3 are conserved in all natural MUPs (Y103, N107, L124) and 3 are not (V59, L88, E137). They were mutated computationally to generate seven mutants (V59T, L88Q, Y103D, Y103R, N107L, L124V, E137K). The molecular modelling framework software Sirius 1.2 (San Diego Supercomputer Center, USDC) was used to visualize the 2L9C.pdb NMR solution structure of the wild type MUP20, in which the selected mutation sites were identified. After the single amino acid mutation in the structure of the protein, the mutants were evaluated for conformational stability using PyMOL software^{43,65}. Structural comparison with original MUP20 3D model, including sequence alignment, structure alignment and RMSD statistics was carried out using SuperPose Version 1.0 (see Supplementary material S1)⁶⁶.

Ligand selection and in silico docking. Protein and ligands were prepared and then blind docking experiments were performed using Autodock Vina (Molecular Graphics Laboratory, Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA)^{42,43}.

Output data (energies, interaction patterns) were analyzed and scored using UCSF Chimera molecular viewer⁶⁷ and BIOVIA Discovery Studio software (Dassault Systèmes BIOVIA, San Diego), which were also used to produce the artworks⁶⁸.

The ZINC database⁶⁹ provided molecules in 3D format for ready-to-dock virtual screening.

Expression and Characterisation of mutant proteins. Selected mutant proteins were expressed *in vitro*. Salts, solvents and reagents were of analytical grade and purchased from Sigma (St. Louis, MO, USA), unless otherwise specified. Isopropyl-β-D-thiogalactoside and ampicillin LB-Media were from BioChemica, EuroClone, Milan, Italy.

Protein Expression: The wild type 6xHis-tagged MUP20 was designed with the addition of a 6xHisTag N-terminal tail in a pEAX plasmid (see Supplementary Material S2a), the resulting gene 6His-MUP20 was synthesised (Eurofins, Milan, Italy). The plasmid carrying the MUP20 was NdeI/EcoRI digested and the insert genes were then recombined in pET5b vector for heterologous expression in DE3-BL21 E. coli (Supplementary Material S2b). After induction of bacterial expression with 0.4 mM IPTG, the expressed protein was confirmed by SDS-page based on the molecular weight. The same procedure was used for the expression of the mutants MUP20V59T, MUP20L88Q, MUP20L124V, MUP20Y103R, MUP20Y103D, MUP20N107L and MUP20E137K from genes obtained by site specific mutagenesis, validated by Sanger sequencing, in which the signal peptide (19 amino acids: MKLLVLLCLGLTLVCVHA) was deleted from the sequences.

Protein purification: The His tag system (HisTrap™ FF crude, GE Healthcare Bio-Sciences) was used to purify the mutant proteins. A HisTrap column was eluted in 20 mM sodium phosphate buffer, pH 7.4, containing 0.5 M NaCl and 0.5 M imidazole. The eluted protein was delipidated on a Sephadex LH-20 column, eluted with 50 mM acetate buffer (pH 4.5). Delipidated protein were dialyzed overnight against 50 mM Tris-HCl buffer, pH 7.4, containing 0.15 M NaCl, and stored at –80 °C.

Chemical characterization of the mutant MUPs: The homogeneity of MUP mutants was analysed by SDS-PAGE (12% acrylamide), under reducing conditions, after Coomassie staining. RP-HPLC analyses were run on a Grace-Vydac (Hesperia, CA, USA) C4 analytical column (4.6 × 150 mm). The column was eluted with a linear 0.1% (v/v) TFA-acetonitrile gradient from 10 to 60% in 30 minutes at a flow rate of 0.8 ml/min and the absorbance of the effluent was recorded at 226 nm. The chemical identity of the purified mutant MUPs was established by high-resolution mass spectrometry on a Xevo G2-S Q-TOF mass spectrometer (Waters, MA, USA). The proteins were also analysed for possible volatile contaminants via Solid Phase Micro Extraction (SPME) of the headspace followed by GC/MS analysis.

Spectroscopic measurements: The protein concentrations were determined spectrophotometrically, by measuring the absorbance at 280 nm, on a V-630 spectrophotometer (Jasco, Tokyo, Japan) using an absorptivity value of 1.618 mg⁻¹·cm². Circular dichroism (CD) spectra were recorded on a Jasco J-810 spectropolarimeter equipped with a thermostated cell holder and a Peltier PTC-423S temperature control system. Spectra were recorded using 0.1- or 1-cm pathlength cuvettes in the far- and near-UV region, respectively. Each spectrum was the average of four accumulations, after baseline subtraction. Ellipticity data were expressed as the mean residue ellipticity,

$[\Theta] = (\Theta \cdot \text{MRW}) / (10 \cdot l \cdot c)$, where Θ is the measured ellipticity in degrees, MRW is the mean residue weight, l is the cuvette pathlength, and c is the protein concentration in g/ml⁷⁰. All measurements were conducted in PBS buffer (10 mM Na₂HPO₄ pH 7.4, 0.15 M NaCl) at 25 ± 0.1 °C.

Fluorescence competitive binding assays: Fluorescence-competitive binding assays were carried out using the fluorescent probe N-phenyl-1-naphthylamine (1-NPN), as described^{39,40}. A binding curve was constructed and the dissociation constant of the probe (K_D) was calculated. Proteins were diluted to 1 μM in 50 mM Tris-HCl, pH 7.4, then 1.6 μM of the fluorescent probe 1-NPN was added and equilibrated for 1 minute. The solution was excited at 337 nm and the emission spectra were measured at 25 °C in the wavelength range 380–450 nm on a Perkin Elmer LS50 pulsed lamp spectrofluorimeter.

The binding analyses assumed that the protein had 100% activity and that the stoichiometry of binding was 1:1 at saturation. The affinity of various volatile ligands was measured in competitive binding assays.

The change in fluorescence was recorded after adding increasing concentrations of ligand. Competitor concentrations causing a decay of fluorescence to half-maximal intensity were determined as IC50 and the dissociation constant for the test ligands (expressed as K_d) was calculated using the following formula:

$$K_d = \frac{[\text{IC50}]}{1 + \frac{[\text{Probe}]}{K_D}} \quad (1)$$

where [Probe] is the free concentration of the fluorescent probe and K_D is the dissociation constant of the complex Protein/Probe.

Isothermal Titration Calorimetry (ITC): The binding energetics of the best binding ligands, 2-isobutyl-3-methoxypyrazine (PYR) and menadione (MEN), to wild-type (WT) and mutant MUPs were studied. ITC titrations were performed at 25 ± 0.1 °C in 50 mM Tris HCl pH 7.4, using a MicroCal VP-ITC instrument. To a MUP solution (1.7 ml, 25 μM) were sequentially added 25 aliquots (10 μl each) of a ligand stock solution (500 μM), with a delay of 4 min after each injection with continuous stirring of the solution (at 307 r.p.m.) in the sample cell. Before analysis, protein samples were dialyzed overnight in the same buffer, using a Slide-A-Lyzer (3.5-kDa cutoff) from ThermoFisher Scientific (Waltham, MA, USA), and thoroughly degassed. The heat of dilution was determined in control experiments by injecting aliquots (10 μl) of a ligand stock solution (500 μM) into buffer and this was subtracted from the integrated binding isotherm prior to curve fitting. Each injection generated a heat-burst curve (μcal s⁻¹) versus time (min). The area under each peak was determined by integration using Origin 7.5 software (Microcal, Inc.) to give the heat associated with the injection. The heat exchanged per mole of injectant was plotted against the ligand:protein molar ratio. The binding constants (K_a) and stoichiometry (N) parameters of the binding process were obtained by fitting the integrated heats of binding to the one-site binding model using the MicroCal ITC Data Analysis software.

References

- Flower, D. R., North, A. C. T. & Sansom, C. E. The lipocalin protein family: Structural and sequence overview. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1482**, 9–24 (2000).
- Kuser, P. R., Franzoni, L., Ferrari, E., Spisni, A. & Polikarpov, I. The X-ray structure of a recombinant major urinary protein at 175 Å resolution: A comparative study of X-ray and NMR-derived structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **57**, 1863–1869 (2001).
- Cavaggioni, A. & Mucignat-Caretta, C. Major urinary proteins, alpha(2U)-globulins and aphrodisin. *Biochim. Biophys. Acta* **1482**, 218–228 (2000).
- Böcskei, Z. *et al.* Pheromone binding to two rodent urinary proteins revealed by X-ray crystallography. *Nature* **360**, 186–188 (1992).
- Mucignat-Caretta, C. & Caretta, A. Message in a bottle: Major urinary proteins and their multiple roles in mouse intraspecific chemical communication. *Anim. Behav.* **97**, 255–263 (2014).
- Cavaggioni, A., Mucignat, C. & Tirindelli, R. Pheromone signalling in the mouse: Role of urinary proteins and vomeronasal organ. *Arch. Ital. Biol.* **137**, 193–200 (1999).
- Bacchini, A., Gaetani, E. & Cavaggioni, A. Pheromone binding proteins of the mouse, *Mus musculus*. *Experientia* **48**, 419–421 (1992).
- Hurst, J. L. *et al.* Individual recognition in mice mediated by major urinary proteins. *Nature* **414**, 631–634 (2001).
- Cavaggioni, A., Mucignat-Caretta, C. & Redaelli, M. Mice recognize recent urine scent marks by the molecular composition. *Chem. Senses* **33**, 655–663 (2008).
- Pelosi, P., Mastrogiacomo, R., Iovinella, I., Tuccori, E. & Persaud, K. C. Structure and biotechnological applications of odorant-binding proteins. *Appl. Microbiol. Biotechnol.* **98**, 61–70 (2014).
- Persaud, K. C. Biomimetic olfactory sensors. *IEEE Sensors Journal* **12**, 3108–3112 (2012).
- Di Pietrantonio, F. *et al.* A surface acoustic wave bio-electronic nose for detection of volatile odorant molecules. *Biosens. bioelectron.* **67**, 516–523 (2015).
- Mulla, M. Y. *et al.* Capacitance-modulated transistor detects odorant binding protein chiral interactions. *Nature Communications*. **6**, 6010 (2015).
- Kotlowski, C. *et al.* Fine discrimination of volatile compounds by graphene-immobilized odorant-binding proteins. *Sensors and Actuators, B: Chemical*. **256**, 564–572 (2018).
- Larisika, M. *et al.* Electronic Olfactory Sensor Based on A. mellifera Odorant-Binding Protein 14 on a Reduced Graphene Oxide Field-Effect Transistor. *Angewandte Chemie*. **54**, 13245–13248 (2015).
- Beynon, R. J. *et al.* Polymorphism in major urinary proteins: Molecular heterogeneity in a wild mouse population. *J. Chem. Ecol.* **28**, 1429–1446 (2002).
- Darwish Marie, A. *et al.* Effect of polymorphisms on ligand binding by mouse major urinary proteins. *Protein Sci.* **10**, 411–417 (2001).
- Kwak, J. *et al.* Differential binding between volatile ligands and major urinary proteins due to genetic variation in mice. *Physiol. Behav.* **107**, 112–120 (2012).
- Sharrow, S. D., Vaughn, J. L., Zidek, L., Novotny, M. V. & Stone, M. J. Pheromone binding by polymorphic mouse major urinary proteins. *Protein Sci.* **11**, 2247–2256 (2002).
- Yang, L. *et al.* Major urinary protein (MUP) profiles show dynamic changes rather than individual “barcode” signatures. *Front. Ecol. Evol.* **1482**, 1667–1677 (2016).

21. Beynon, R. J. & Hurst, J. L. Urinary proteins and the modulation of chemical scents in mice and rats. *Peptides* **25**, 1553–1563 (2004).
22. Mucignat-Caretta, C., Caretta, A. & Cavaggoni, A. Acceleration of puberty onset in female mice by male urinary proteins. *J. Physiol.* **486**, 517–522 (1995).
23. Papes, F., Logan, D. W. & Stowers, L. The Vomeronasal Organ Mediates Interspecies Defensive Behaviors through Detection of Protein Pheromone Homologs. *Cell* **141**, 692–703 (2010).
24. Sherborne, A. L. *et al.* The genetic basis of inbreeding avoidance in house mice. *Curr. Biol.* **17**, 2061–2066 (2007).
25. Phelan, M. M. *et al.* The structure, stability and pheromone binding of the male mouse protein sex pheromone darcin. *PLoS One* **9**, e108415 (2014).
26. Roberts, S. A. *et al.* Darcin: a male pheromone that stimulates female memory and sexual attraction to an individual male's odour. *BMC Biol.* **8**, 75 (2010).
27. Scorsone, E. *et al.* Major Urinary Proteins on Nanodiamond-Based Resonators Toward Artificial Olfaction. *IEEE Sens. J.* **16**, 6543–6550 (2016).
28. Poivet, E. *et al.* Applying medicinal chemistry strategies to understand odorant discrimination. *Nat. Commun.* **7**, 11157 (2016).
29. Pes, D. & Pelosi, P. Odorant-binding proteins of the mouse. *Comp. Biochem. Physiol.—Part B Biochem.* **112**, 471–479 (1995).
30. Buttery, R. G., Seifert, R. M., Guadagni, D. G. & Ling, L. C. Characterization of some volatile constituents of bell peppers. *J. Agric. Food Chem.* **17**, 1322–1327 (1969).
31. Seifert, R. M., Buttery, R. G., Guadagni, D. G., Black, D. R. & Harris, J. G. Synthesis of Some 2-Methoxy-3-Alkylpyrazines with Strong Bell Pepper-Like Odors. *J. Agric. Food Chem.* **18**, 246 (1970).
32. Robertson, D., Hurst, J. L., Hubbard, S., Gaskell, S. J. & Beynon, R. Ligands of Urinary Lipocalins from the Mouse: Uptake of Environmentally Derived Chemicals. *J. Chem. Ecol.* **24**, 1127–1140 (1998).
33. Redaelli, M., Orsetti, A., Zagotto, G., Cavaggoni, A. & Mucignat-Caretta, C. Airborne molecules released from male mouse urine affect female exploratory behaviour. *Front. Ecol. Evol.* **2**, 28 (2014).
34. Ahmed, M. K., Takumida, M., Ishibashi, T., Hamamoto, T. & Hirakawa, K. Expression of transient receptor potential vanilloid (TRPV) families 1, 2, 3 and 4 in the mouse olfactory epithelium. *Rhinology* **47**, 242–247 (2009).
35. Perez-Miller, S., Zou, Q., Novotny, M. V. & Hurley, T. D. High resolution X-ray structures of mouse major urinary protein nasal isoform in complex with pheromones. *Protein Sci.* **19**, 1469–1479 (2010).
36. Monera, O. D., Sereda, T. J., Zhou, N. E., Kay, C. M. & Hodges, R. S. Relationship of sidechain hydrophobicity and alpha-helical propensity on the stability of the single-stranded amphipathic alpha-helix. *J. Pept. Sci.* **1**, 319–329 (1995).
37. Margarit, I. *et al.* Cumulative stabilizing effects of glycine to alanine substitutions in *Bacillus subtilis* neutral protease. *Protein Eng.* **5**, 543–550 (1992).
38. Hetényi, C. & Van Der Spoel, D. Toward prediction of functional protein pockets using blind docking and pocket search algorithms. *Protein Sci.* **20**, 880–893 (2011).
39. Sun, Y.F. *et al.* Two odorant-binding proteins mediate the behavioural response of aphids to the alarm pheromone (e)- β -farnesene and structural analogues. *PLoS One* **7**, <https://doi.org/10.1371/journal.pone.0032759> (2012).
40. Löbel, D., Marchese, S., Krieger, J., Pelosi, P. & Breer, H. Subtypes of odorant-binding proteins—heterologous expression and ligand binding. *Eur. J. Biochem.* **254**, 318–324 (1998).
41. Cavaggoni, A., Mucignat-Caretta, C., Redaelli, M. & Zagotto, G. The scent of urine spots of male mice, *Mus musculus*: Changes in chemical composition over time. *Rapid Commun. Mass Spectrom.* **20**, 3741–3746 (2006).
42. Trott, O. & Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
43. Seeliger, D. & De Groot, B. L. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J. Comput. Aided. Mol. Des.* **24**, 417–422 (2010).
44. Lücke, C. *et al.* Solution structure of a recombinant mouse major urinary protein. *Eur. J. Biochem.* **266**, 1210–1218 (1999).
45. Barratt, E. *et al.* Van der Waals interactions dominate ligand-protein association in a protein binding site occluded from solvent water. *J. Am. Chem. Soc.* **127**, 11827–11834 (2005).
46. Sharrow, S. D., Novotny, M. V. & Stone, M. J. Thermodynamic analysis of binding between mouse major urinary protein-I and the pheromone 2-sec-butyl-4,5-dihydrothiazole. *Biochemistry* **42**, 6302–6309 (2003).
47. Sharrow, S. D., Edmonds, K., Goodman, M., Novotny, M. V. & Stone, M. J. Thermodynamic consequences of disrupting a water-mediated hydrogen bond network in a protein:pheromone complex. *Protein Sci.* **14**, 249–256 (2005).
48. Nespoulous, C., Briand, L., Delage, M. M., Tran, V. & Pernollet, J. C. Odorant binding and conformational changes of a rat odorant-binding protein. *Chem. Senses.* **29**, 189–198 (2004).
49. Golebiowski, J., Topin, J., Charlier, L. & Briand, L. Interaction between odorants and proteins involved in the perception of smell: the case of odorant-binding proteins probed by molecular modelling and biophysical data. *Flavour Fragr. J.*, <https://doi.org/10.1002/ffj.3121> (2012).
50. Hajjar, E., Perahia, D., Débat, H., Nespoulous, C. & Robert, C. H. Odorant binding and conformational dynamics in the odorant-binding protein. *J. Biol. Chem.* **281**, 29929–29937 (2006).
51. Vecellio Reane, D. *et al.* A MICU-1 splice variant confers high sensitivity to the Ca²⁺ uptake machinery of skeletal muscle. *Mol. Cell* **64**, 760–773 (2016).
52. Pozzi, N. *et al.* Loop electrostatics asymmetry modulates the pre-existing conformational equilibrium in thrombin. *Biochemistry* **55**, 3984–3994 (2016).
53. De Filippis, V., Vangelista, L., Schiavo, G., Tonello, F. & Montecucco, C. Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin. *Eur. J. Biochem.* **229**, 61–69 (1995).
54. Pertinhez, T. A. *et al.* The binding cavity of mouse major urinary protein is optimised for a variety of ligand binding modes. *Biochem. Biophys. Res. Commun.* **390**, 1266–1271 (2009).
55. Macek, P., Novák, P., Krizová, H., Zídek, L. & Sklenár, V. Molecular dynamics study of major urinary protein-pheromone interactions: a structural model for ligand-induced flexibility increase. *FEBS Lett.* **580**, 682–684 (2006).
56. De Filippis, V., P de Laureto, P., Toniutti, N. & Fontana, A. Acid-induced molten globule state of a fully active mutant of human interleukin-6. *Biochemistry* **35**, 11503–11511 (1996).
57. Meillour, P. N. *et al.* Phenylalanine 35 and tyrosine 82 are involved in the uptake and release of ligand by porcine odorant-binding protein. *Biochim. Biophys. Acta.* **1794**, 1142–1150 (2009).
58. Ilayaraja, R. *et al.* Evaluating the binding efficiency of pheromone binding protein with its natural ligand using molecular docking and fluorescence analysis. *Sci. Rep.* **4**, 5201 (2014).
59. Dung, T. T. *et al.* Applications and advances in bioelectronic noses for odour sensing. *Sensors.* **18**, 103 (2018).
60. Wasilewski, T., Gebicki, J. & Kamysz, W. Advances in olfaction-inspired biomaterials applied to bioelectronic noses. *Sens. Actuators, B.* **257**, 511–537 (2018).
61. Guo, Z. *et al.* A novel platform based on immobilized histidine tagged olfactory receptors, for the amperometric detection of an odorant molecule characteristic of boar taint. *Food Chem.* **184**, 1–6, <https://doi.org/10.1016/j.foodchem.2015.03.066> (2015).
62. Akimov, V. *et al.* Nanobiosensors based on individual olfactory receptors. *Analog Integr. Circuits Signal Process.* **57**, 197–203 (2008).
63. Lim, J. H. *et al.* Nanovesicle-Based Bioelectronic Nose for the Diagnosis of Lung Cancer from Human Blood. *Adv. Healthcare Mat.* **3**, 360–366 (2014).

64. Sobolev, V. *et al.* SPACE: A suite of tools for protein structure prediction and analysis based on complementarity and environment. *Nucleic Acids Res.* **33**, W39–43 (2005).
65. DeLano, W. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr.* **40**, 700 (2002).
66. Maiti, R., Van Domselaar, G. H., Zhang, H. & Wishart D. S. SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res.* **32**, W590–594 (2004).
67. Pettersen, E. F. *et al.* UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
68. Elengoe, A., Abu Naser, M. & Hamdan, S. Modeling and docking studies on novel mutants (K71L and T204V) of the ATPase domain of human heat shock 70 kDa protein 1. *Int. J. Mol. Sci.* **15**, 6797–6814 (2014).
69. Irwin, J. J. & Shoichet, B. K. ZINC - A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **45**, 177–182 (2005).
70. Pontarollo, G. *et al.* Non-canonical proteolytic activation of human prothrombin by subtilisin from *Bacillus subtilis* may shift the procoagulant-anticoagulant equilibrium toward thrombosis. *J. Biol. Chem.* **292**, 15161–15179 (2017).
71. Tetko, I. V. *et al.* Virtual computational chemistry laboratory—design and description. *J. Comput. Aided Mol. Des.* **19**, 453–463 (2005).

Acknowledgements

Supported by EU FP7 Project No. 285203 – SNIFFER to CMC, and PRAT-2015 from the University of Padova to VDF. We gratefully acknowledge Marco Redaelli for data collection, Paolo Pelosi and Andrea Cavaggioni for helpful suggestions and advice.

Author Contributions

J.R. and C.M.C. designed the study. J.R., L.A., M.B., R.E.L., S.B., R.P. performed the experiments. J.R., G.R., V.D.F., G.Z., K.C.P. and C.M.C. analyzed the data and wrote the paper. All authors approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-36391-3>.

Competing Interests: The authors declare no competing interests.

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) 2019