Chem. Senses 23: 689–698, 1998

Amino Acid Sequence, Post-translational Modifications, Binding and Labelling of Porcine Odorant-binding Protein

Sara Paolini, Andrea Scaloni[1], Angela Amoresano[2], Silvana Marchese, Elio Napolitano[3] and Paolo Pelosi

Dipartimento di Chimica e Biotecnologie Agrarie, University of Pisa, via S. Michele 4, I-56124 Pisa, [1]Centro Internazionale Servizi di Spettrometria di Massa—IABBAM, National Research Council, via Pansini 5, I-80131 Napoli, [2]CEINGE Biotecnologie Avanzate s.r.c.l., via Pansini 5, I-80131 Napoli and [3]Scuola Normale Superiore, P.za dei Cavalieri 7, I-56100 Pisa, Italy

Correspondence to be sent to: Prof. Paolo Pelosi, Dipartimento di Chimica e Biotecnologie Agrarie, via S. Michele 4, I-56124 Pisa, Italy. e-mail: ppelosi@agr.unipi.it

Abstract

An odorant-binding protein, migrating in SDS–PAGE with an apparent molecular weight of 22 kDa and an isoelectric point of 4.2, has been purified from pig nasal mucosa. Its complete amino acid sequence was determined by a combination of mass spectrometry and Edman degradation procedures. The protein consists of a single polypeptide chain of 157 amino acids, presenting at the N-terminus a pyroglutamic acid residue. The two cysteine residues, occurring in the primary structure at positions 63 and 155, are involved in an intramolecular disulphide bridge. Sequence comparison with other lipocalins revealed a good similarity with bovine odorant-binding protein, the only member of this class which does not contain disulphide bonds and of which the three-dimensional structure recently has been resolved. Nine out of the 16 residues lining the binding pocket in bovine OBP are conserved in the porcine protein, suggesting structural similarities in this region of the molecule. The synthesis of a fluorescent photoaffinity labelling agent and of two tin-containing thymol analogues is also described. These compounds together with other ligands were able to bind the protein as revealed by competitive binding experiments.

Introduction

Odour perception and recognition seems to be mediated, at the peripheral level, by soluble odorant-binding proteins (OBPs), which are very abundant in the nasal mucus of vertebrates and in the sensillar lymph of insects (Vogt and Riddiford, 1981; Pelosi *et al.*, 1982; Pelosi and Maida, 1990; Vogt *et al.*, 1990; Breer *et al.*, 1992; Pelosi, 1994, 1996). The exact physiological role of these proteins is still elusive, although these are the only macromolecules of the olfactory system for which reversible binding of odorants and pheromones has been measured.

The structural characterization of OBPs, combined with their ligand-binding properties, provides the first step towards understanding their physiological role. At present, complete amino acid sequences have been determined for eight members of this family, one in the cow (Tirindelli *et al.*, 1989), two in the rat (Pevsner *et al.*, 1988; Dear *et al.*, 1991), four in the mouse (Miyawaki *et al.*, 1994; Pes *et al.*, 1998) and one in the frog (Lee *et al.*, 1987). The bovine OBP is the only one to have been directly sequenced; in all other cases the primary structures have been derived from their corresponding nucleotide sequences. In addition, N-terminal sequences are available for a number of other odorant-binding proteins (Pes and Pelosi, 1995; Ganni *et al.*, 1997; Garibotti *et al.*, 1997). The expression of several OBPs in the same animal species and the large differences observed in their primary structure seem to indicate different binding specificities towards odorants and pheromones. This fact, when properly documented, could support a ligand-discriminating role for these proteins.

The three-dimensional structure has been resolved only for bovine OBP (Tegoni *et al.*, 1996; Bianchet *et al.*, 1996). It shares the typical β-barrel motif of lipocalins, the superfamily of carrier proteins to which all known vertebrate OBPs belong (Sansom *et al.*, 1994; Flower, 1996). Its binding pocket is located, as in other lipocalins, in the hydrophobic core of the protein. An interesting feature observed in the structure of dimeric bovine OBP is the so-called 'domain swapping', which probably contributes to stabilizing the structure of this protein, which does not contain any disulphide bridge.

Ligand-binding data, obtained for porcine OBP with several odorant molecules (Dal Monte *et al.*, 1993), indicate a strong similarity with the bovine protein (Pelosi and Tirindelli, 1989; Pevsner *et al.*, 1990; Hérent *et al.*, 1994),

suggesting structural analogies between the two proteins. However, the monomeric nature of porcine OBP indicates that the presence of different structural elements is well worth investigating.

In this paper we report the complete amino acid sequence of porcine OBP determined through direct sequencing and its post-translational modifications. The synthesis of different odorants containing fluorescent or metal moieties and their binding to pig OBP are also described.

Materials and methods

Materials

Pig nasal mucosa was obtained from the local abattoir. The tissue was collected within 20 min after animal death and soon utilized for the extraction of OBP.

Tritiated 2-isobutyl-3-methoxypyrazine was prepared at the Amersham Laboratories (UK) by catalytic hydrogenation of the unsaturated precursor 2-(2-methyl-1-propenyl)-3-methoxypyrazine with tritium gas. This pyrazine derivative was synthesized at our laboratory and the experimental conditions for its hydrogenation were determined with hydrogen gas (P. Pelosi, unpublished results). The pure tritiated compound had a sp. act. of ~42 Ci/mmol. For binding experiments, the labelled ligand was diluted 10-fold with unlabelled 2-isobutyl-3-methoxypyrazine. Mono-Q and Superose-12 resins were from Pharmacia-LKB (Uppsala, Sweden). Solvents and reagents for Edman degradation were 'sequencing grade', obtained from the manufacturer of the sequencer. Proteolytic enzymes and other chemical reagents were purchased from Sigma and Aldrich.

Purification of the protein

OBP was purified in milligram amounts from fresh porcine nasal tissue by anion-exchange chromatography on a DE-52 resin (Whatman), according to the procedure described by Dal Monte *et al.* (1991). Highly purified samples of the protein were obtained through a final preparative isoelectric focusing in agarose gel, using a 3.5–9.5 Ampholine gradient and a Pharmacia-LKB cell (Uppsala, Sweden), according to the manufacturer's instructions. This procedure afforded on average 10–15 mg of purified OBP from a single pig's head. Samples for electrospray mass spectrometry (ESMS) analysis were desalted by reversed-phase HPLC on a Vydac 214TP54 column and eluted with a linear gradient of acetonitrile from 2 to 60% (v/v) in 0.1% (v/v) aqueous trifluoroacetic acid (TFA), over 60 min, at a flow rate of 1 ml/min; a single elution peak appeared in the chromatogram.

Reduction and carboxymethylation of pig OBP

The protein (2 mg) was dissolved in 1 ml of 6 M guanidine, 0.3 M Tris–HCl buffer, pH 8.5. One millilitre of 8 mM dithiothreitol was added and the mixture was left for 90 min at room temperature, in the dark. Carboxymethylation was accomplished by treatment with 0.1 ml of 1.4 M iodoacetic acid in 0.3 M Tris–HCl, pH 8.5, for 60 min at room temperature, in the dark. After dialysis against 50 mM ammonium bicarbonate, the carboxymethylated protein was used for selective digestions.

Chemical and enzymatic hydrolyses

In the case of cyanogen bromide hydrolysis, the carboxymethylated protein (200 μg) was dissolved in 150 μl of 70% (v/v) aqueous TFA and treated with 1 mg of reagent. After standing for 22 h under nitrogen in the dark, at room temperature, the reaction was quenched by 10-fold dilution with water and the product dried under reduced pressure.

In the case of 3-bromo-3-methyl-2-[2-nitrophenylmercapto]-3H-indole (BNPS-skatole) hydrolysis, porcine carboxymethylated OBP (200 μg) was dissolved in 70 μl of water, added to 140 μl of acetic acid and reacted with 5 mg of reagent solved in 70 μl of ethanol. After 48 h in the dark, under nitrogen at room temperature, the reaction mixture was diluted with an equal volume of water and lyophilized.

Endoproteinase Glu-C digestion was performed on both the native and the carboxymethylated protein. In the first case, 100 μg of purified OBP was incubated with the proteolytic enzyme in 200 μl of 50 mM ammonium bicarbonate, pH 8.5, for 48 h at 37°C (enzyme:substrate 1:20 w/w). In another experiment, the carboxymethylated protein (100 μg) was incubated for 10 min with the protease (enzyme:substrate 1:10 w/w) under the same experimental conditions.

Trypsin digestion was performed on both the carboxymethylated and the native protein. In the first case, 100 μg of the carboxymethylated protein was incubated with the proteolytic enzyme in 200 μl of 50 mM ammonium bicarbonate, pH 6.5, at 37°C for 15 min or overnight (enzyme:substrate 1:20 w/w). In the disulphide bridge assignment experiment, the native protein (100 μg) was incubated overnight with the protease under the same experimental conditions.

Gel electrophoresis

Electrophoresis of the hydrolysis mixtures was run on a tricine–SDS gel, according to Schägger and von Jagow (1987). SDS–PAGE was performed following the procedure of Laemmli (1970). Protein samples were directly electroblotted on a PVDF membrane (Matsuidara, 1987) and the excised bands submitted to automated Edman degradation, or electroeluted from stained electrophoretic gel bands and utilized for sequence determination.

Mass spectrometry analyses

Intact protein or individual peptide fractions were submitted to ESMS analysis, using a BIO-Q triple quadrupole mass spectrometer (Micromass, Manchester, UK). Samples were dissolved in 1% (v/v) acetic acid and 50% (v/v)

aqueous acetonitrile, and 2–10 µl was injected into the mass spectrometer at a flow rate of 10 µl/min. The quadrupole was scanned from *m/z* 500 to 1800 at 10 s/scan and the spectra were acquired and elaborated using the MASSLYNX software. Calibration was performed by the multiply charged ions from a separate injection of myoglobin (mol. wt 16,951.5 Da). All mass values are reported as average masses.

Matrix-assisted laser desorption ionizationmass spectra (MALDIMS) were recorded using a Voyager DE MALDI-TOF mass spectrometer (Perseptive Biosystem). A mixture of analyte solution, α-cyano-4-hydroxycinnamic acid and bovine insulin was applied to the sample plate and dried *in vacuo*. Mass calibration was performed using the molecular ions from the bovine insulin at 5734.54 Da and the matrix at 379.06 Da as internal standards. Raw data were analysed by using computer software provided by the manufacturer and are reported as average masses.

Amino acid sequence analysis

The amino acid sequence was determined by direct Edman degradation on a Milligen 6600 apparatus or a Perkin-Elmer Applied Biosystems 477A instrument. In the first case samples of electroeluted peptides (200–500 pmol) in ammonium bicarbonate were dried, solubilized in 5% (v/v) *N*-methylmorpholine and 50% (v/v) aqueous isopropanol, immobilized on 1,4-phenylene diisothyocyanate (DITC)-activated membranes and subjected to automatic sequential degradation. DITC membranes were prepared as previously described (Garibotti *et al.*, 1982). In the second case electroblotted samples were directly analysed from PVDF membranes, according to Palmieri *et al.* (1996).

Synthesis of 1-azidoanthracene

1-Aminoanthracene (1.93 g, 10 mmol) was dissolved in concentrated hydrochloric acid (10 ml) and converted into its diazonium salt at 0–5°C by the slow addition of 0.76 g (11 mmol) of $NaNO_2$ dissolved in 2 ml of water. This solution was then treated with 0.72 g (11 mmol) of sodium azide in 2 ml of water and left at room temperature for 1 h. Extraction with diethyl ether gave the crude product as brown-grey crystals (0.77 g, 3.5 mmol, yield: 35%). The 1-azidoanthracene was purified by recrystallization from ethanol, yielding long yellowish needles.

Synthesis of 2-trimethylstannyl-5-methylanisole and 2-trimethylstannyl-5-methoxymethylanisole

Tin-containing compounds were prepared from the corresponding bromoaromatic analogues by the lithium–bromine exchange reaction with BuLi at –78°C, followed by quenching of the intermediate lithiated derivatives with trimethyltin chloride. The compounds were purified by reversed-phase HPLC (C18 column) using mixtures of methylene chloride and acetonitrile as eluants. Attempts to obtain the free phenol from either compound gave extensive destannilation.

Ligand-binding assays

Binding to tritiated 2-isobutyl-3-methoxypyrazine was measured using the filtration assay of Bruns *et al.* (1983), as previously described (Dal Monte *et al.*, 1991). Non-specific binding was measured in the presence of a 1000-fold excess of the cold ligand and accounted for <5% of total binding.

Competitive binding assays were performed using tritiated 2-isobutyl-3-methoxypyrazine and protein, both at a concentration of 1 µM, and competitor ligand in the non-labelled form at concentrations of 1, 10 and 100 µM.

Photoaffinity labelling

Purified OBP at a concentration of 10 mM was incubated with equimolar amounts of 1-azidoanthracene in 20 mM Tris–HCl buffer, pH 7.4, for 1 h at 4°C. The ligand was previously dissolved in ethanol and slowly added to the protein under stirring; the ethanol concentration of the final mixture was 1% (v/v). This solution was then illuminated with an ultraviolet lamp at a distance of 5 cm, at 4°C for periods ranging from 1 to 10 min. The product was then chromatographed through a Sephadex G-15 column to separate unreacted ligand from the protein. Covalent binding of the azido compound to porcine OBP was checked by UV visualization after SDS–PAGE analysis.

Results

Structural characterization of porcine OBP

The protein was purified from pig nasal mucosa by anion-exchange chromatography on a DE-52 resin, as previously reported (Dal Monte *et al.*, 1991), with the addition of a final step of preparative isoelectric focusing (IEF) in agarose gel, in order to obtain ultra-pure samples for amino acid sequencing.

The purified protein, which we call OBP-I, being by far the most abundant OBP in the pig though not the only one expressed (P. Pelosi, unpublished results), showed a single band on electrophoretic gel in denaturing conditions (12% SDS–PAGE), migrating with an apparent mass of 22 kDa. Its isoelectric point, as measured by IEF in a gradient of ampholines, was 4.2, as previously reported (Dal Monte *et al.*, 1991).

ESMS analysis of a reversed-phase HPLC purified sample of porcine OBP-I is reported in Figure 1. A single molecular species exhibiting a molecular mass of 17 689.6 ± 1.8 Da was detected in the spectrum. This protein proved refractory to direct sequencing by Edman degradation. However, after treatment with pyroglutamate aminopeptidase we could determine a short segment of its N-terminal sequence, thus suggesting that the first amino acid was a pyroglutamic acid residue.

Aliquots of the native or carboxymethylated protein were

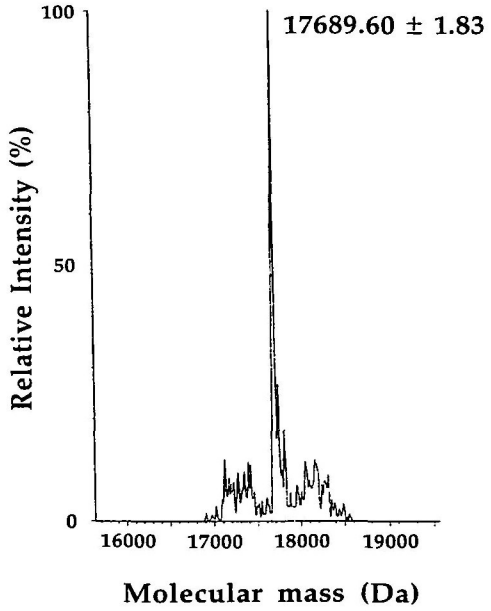

Figure 1 ES/MS analysis of pig OBP-I. Mass spectrum transformed on a real mass scale; the measured value is reported.

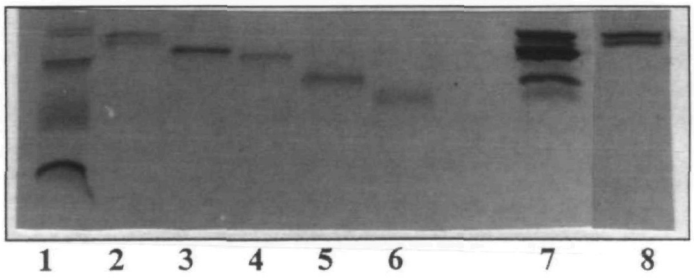

Figure 2 Tricine–SDS–PAGE of the peptides obtained by chemical hydrolysis of pig OBP-I. **Lane 1**: Molecular weight markers (from the top: soybean trypsin inhibitor, 20 kDa; α-lactalbumin, 14 kDa; monellin, 10.7 kDa; insulin, 3.4 kD); **lane 2**: purified pig OBP-I; **lanes 3–6**: peptides obtained by CNBr hydrolysis and purified by electroelution; **lane 7**: crude CNBr hydrolysis product; **lane 8**: crude BNPS-skatole hydrolysis product.

subjected to selective digestions, using enzymes such as endoproteinase Glu-C or trypsin, and chemical reagents such as BNPS-skatole or cyanogen bromide (CNBr). Pig OBP digests were directly analysed by tricine–SDS–PAGE, and the separated polypeptides electroeluted or electroblotted on PVDF membranes and directly sequenced. As an example, Figure 2 reports the electrophoretic separations of the peptides obtained by chemical hydrolyses.

Endoproteinase Glu-C, when used with the native protein, yielded a large peptide of molecular mass slightly lower than the intact protein suitable for sequencing, indicating that pig OBP-I presents a compact structure similar to the bovine protein (Tirindelli *et al.*, 1989). In denaturing conditions, the same enzyme degraded it into a series of smaller fragments. Two of these peptides, with apparent molecular masses of 14 and 11 kDa respectively, provided additional sequence information on internal regions of the protein.

BNPS-skatole hydrolysis afforded a single large peptide of ~16 kDa (lane 8 in Figure 2), suggesting the presence of a single tryptophane residue close to the N-terminus of the protein—a feature characteristic of all lipocalins, including OBPs. These findings were in agreement with the shape of the UV absorption spectrum (data not shown).

Trypsin and CNBr hydrolysed the protein quantitatively, affording several peptides in good yield, suitable for sequencing of the remaining internal regions of pig OBP-I (lanes 3–6 in Figure 2). Figure 3 summarizes the data obtained by direct sequencing of the peptides separated by electrophoresis and the assembly of the partial sequences into the complete OBP primary structure.

Post-translational modifications analysis

In order to ascertain the presence of post-translational modifications, porcine OBP-I digests were characterized by MALDIMS. The tryptic mixture of carboxymethylated OBP-I was directly submitted to mass spectrometry analysis and the spectrum obtained is shown in Figure 4. Each signal

Cleavage	Sequence
pyro-Glu	1——14
Glu-C	7————40
BNPS-skatole	17————46
CNBr	40————84
Glu-C	65————89
Trypsin	88————116
Glu-C	96————118
CNBr	115————157
Trypsin	121————157

Figure 3 Peptides covering the complete amino acid sequence of porcine odorant-binding protein as used for Edman degradation. Peptides were numbered retrospectively according to their location starting from the N-terminus. CNBr, BNPS-skatole, endoproteinase Glu-C, trypsin and pyroglutamate-aminopeptidase hydrolysis products are shown.

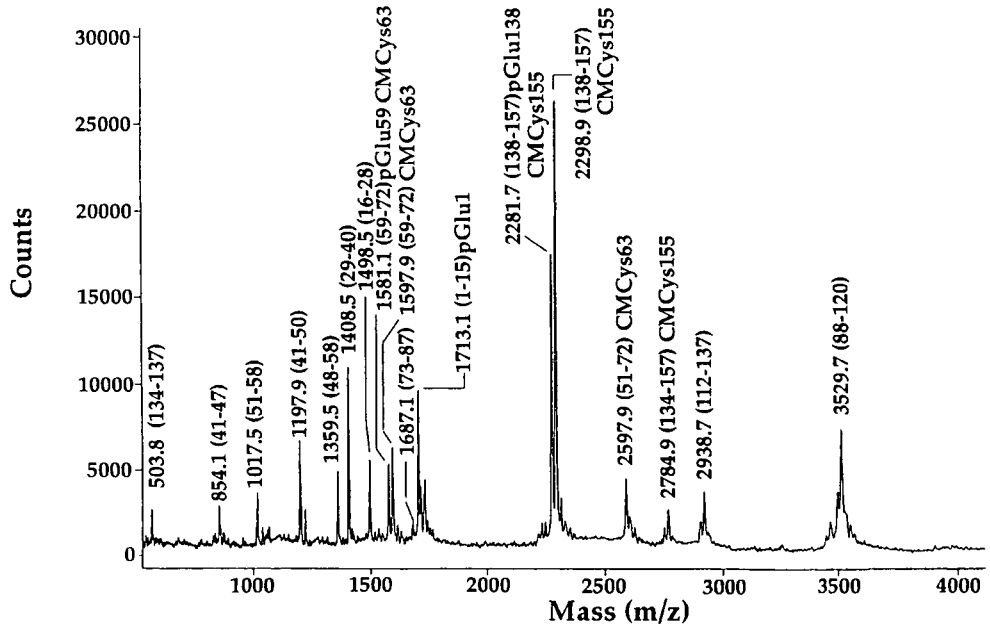

Figure 4 MALDIMS spectrum of reduced and carboxymethylated pig OBP-I following digestion with trypsin. The signals recorded in the spectrum are assigned to the corresponding peptides within the protein sequence on the basis of their molecular masses.

recorded was associated with the corresponding peptide along the amino acid sequence on the basis of its mass value and of trypsin specificity.

The tryptic map allowed the verification of the entire primary structure and the identification of the presence of pyroglutamic acid at the N-terminus of the protein. In fact, the signal at *m/z* 1713.1 was interpreted as corresponding to the peptide 1–15, presenting a pyroglutamic acid at position 1 instead of a Glu residue; on the other hand, the signal at *m/z* 1730.3, expected for the peptide 1–15, was missing, thus explaining the complete reluctance of the native protein to undergo Edman degradation. In addition to the expected signals, the occurrence of satellite peaks 17 Da lower, due to sequence-specific partial cyclization of Glu59 and Glu138 to pyroglutamic acid, was also detected.

Consistent results were obtained in the case of the peptide mixture produced by CNBr hydrolysis. Table 1 summarizes the MALDIMS data generated from this experiment.

In order to determine the redox state of the two cysteine residues present at positions 63 and 155 in the primary structure of porcine OBP-I, an aliquot of the native protein was digested with trypsin and the peptide mixture generated was directly analysed by MALDIMS. In addition to the signals expected on the basis of trypsin specificity, a clear peak at *m/z* 3779.5 was present in the spectrum (reported in Figure 5), assigned to peptide pair (59–72) + (138–157) linked by a disulphide bridge, involving Cys63 and Cys155. This interpretation was confirmed by the presence of a satellite peak 17 Da lower, due to the partial cyclization of Glu138 to 5-oxoproline occurring during the enzymatic hydrolysis. A similar doublet was observed for the signal at

Table 1 MALDIMS analysis of the peptide mixture generated by CNBr hydrolysis of reduced and carboxymethylated porcine odorant-binding

Measured value (*m/z*)	Theoretical value (*m/z*)	Peptide
4398.3	4398.8	(1–39)
4380.5	4380.8	(1–39)-lactone
8513.1	8513.4	(40–114)-carboxymethyl
8495.1	8495.3	(40–114)-carboxymethyl-lactone
4874.0	4874.4	(115–157)-carboxymethyl

m/z 2058.7, corresponding to the peptide pair (59–72) + (153–157) linked by a disulphide bridge between the same Cys residues. Both signals disappeared following dithiothreitol incubation. On the contrary, the peaks corresponding to the reduced peptides (59–72) and (138–157) were not present in the spectrum of the native protein, clearly confirming the presence of a disulphide bridge between Cys63 and Cys155.

Based on these results, the calculated molecular mass of the protein was 17,689.48 Da, in excellent agreement with the value of 17 689.60 ± 1.83 Da, measured by ESMS (Figure 1). The isoelectric point of porcine OBP-I, as calculated on the basis of its amino acid composition, was 3.87, in accord with the value (4.2) determined experimentally. The calculated molar and specific extinction coefficients at 280 nm were 12 330 and 0.696 respectively.

Sequence comparison with other lipocalins

The amino acid sequence of pig OBP-I showed significant

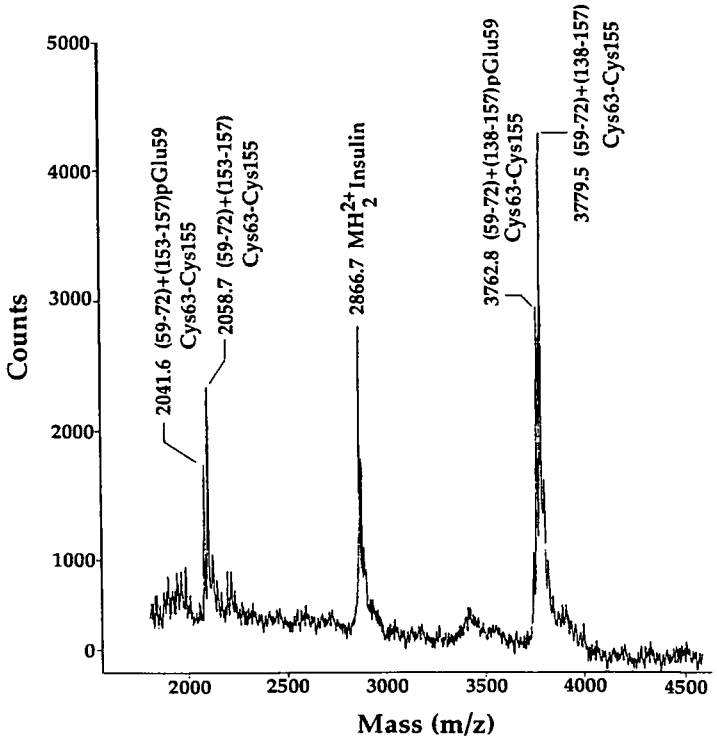

Figure 5 Partial MALDI spectrum of native porcine OBP-I following digestion with trypsin. Signals reported correspond to disulphide-bridged peptides; each signal is assigned to the corresponding peptide pairs and the two cysteine residues involved in the S–S bridge are indicated.

similarity with other members of the same class. Figure 6 shows the alignment of this protein with other, closely related OBPs, as well as with a representative member of the pheromone carrier proteins, hamster aphrodisin (Henzel *et al.*, 1988; Magert *et al.*, 1995). The best similarity (42% of identical amino acids) was found with bovine OBP, but was still significant with mouse (31%) and rat OBP-I (25%). This value dropped to 18% with rat OBP-II (Dear *et al.*, 1991) and to ~10% with frog OBP (Lee *et al.*, 1987) and rat VEG (Schmale *et al.*, 1990)—a lipocalin initially assumed to be involved in taste transduction but later found also to be expressed in the lachrymal glands (Redl *et al.*, 1992; Garibotti *et al.*, 1995) and in the prostate (Holzfeind *et al.*, 1996). Among the lipocalins acting as pheromone carriers, urinary and salivary proteins of mouse and rat (Dinh *et al.*, 1965; Finlayson *et al.*, 1965; Shaw *et al.*, 1983; Shahan *et al.*, 1987) shared only ~20% of their amino acids with pig OBP-I, while hamster aphrodisin (Henzel *et al.*, 1988; Magert *et al.*, 1995), secreted in the vaginal discharge and endowed with pheromonal activity, was significantly more similar to pig OBP-I (34% identity). This protein also shows a relatively high degree of similarity with both subunits of mouse OBP-I (Pes *et al.*, 1998).

Ligand binding experiments

New ligands for the pig OBP have been designed and synthesized. 1-Azidoanthracene is a photoaffinity reagent for covalently labelling the binding site of the protein upon

Name	Sequence	no.
pig OBP-I	EEPQPEQDPFELSGKWITSYIGSSDLEKIGENAPFQVFMRSIEFDDKESKVYLNFF	56
bov OBP	AQEEEAEQNLSELSGPWRTVYIGSTNPEKIQENGPFRTYFRELVFDDEKGTVDFYFS	57
mus OBP-Ia	AMEGPWKTVAIAADRVDKIERGGELRIYCRSLTCEKECKEMKVTFY	46
mus OBP-Ib	HDHPELQGQWKTTAIMADNIDKIETSGPLELFVREITCDEGCQKMKVTFY	50
rat OBP-I	AHHENLDISPSEVNGDWRTLYIVADNVEKVAEGGSLRAYFQHMECGDECQELKIIFN	57
ham APHR	QDFAELQGKWYTIVIAADNLEKIEEGGPLRFYFRHIDCYKNCSEMEITFY	50
	* * * * * *	
pig OBP-I	SKENGICEEFSLIGTKQE-GNTYDVNYAGNNKFVVSYASETALIISNINVDEEGDKTIMT	115
bov OBP	VKRDGKWKNVHVKATKQDDG-TYVADYEGQNVFKIVSLSRTHLVAHNINVDKHGQTTELT	116
mus OBP-Ia	VNENGQCSLTTITGYLQEDGKTYKTQFQGNNRYKLVDESPENLTFYSENVDRADRKTKLL	106
mus OBP-Ib	VKQNGQCSLTTVTGYKQEDGKTFKNQYEGENNYKLLKATSENLVFYDENVDRASRKTKLL	110
rat OBP-I	VKLDSECQTHTVVGQKHEDGR-YTTDYSGRNYFHVLKKTDDIIFFHNVNVDESGRRQCDL	116
ham APHR	VITNNQCSKTTVIGYLKGNG-TYQTQFEGNNIFQPLYITSDKIFFTNKNMDRAGQETNMI	109
	* * * * * * *	
pig OBP-I	GLLGKGTDIEDQDLEKFKEVTRENGIPEENIVNIIERDDCPA	157
bov OBP	GLFVKLN-VEDEDLEKFWKLTEDKGIDKKNVVNFLENEDHPHPE	159
mus OBP-Ia	FILGHGP--LTSEKEKFAELAEEKGIPAGNIREVLITDYCPE	147
mus OBP-Ib	YILGKGEALTHEQKERLTELATQKGIPAGNLRELAHEDTCPE	152
rat OBP-I	VA-GKREDLNKAQKQELRKLAEEYNIPNENTQHLVPTDTCNQ	157
ham APHR	VVAGKGNALTPEENEILVQFAHEKKIPVENILNILATDTCPE	151
	*	

Figure 6 Amino acid sequence alignment of pig OBP-I with most similar members of the lipocalin family. Amino acid residues identical with pig OBP-I are indicated in bold. The asterisk marks residues lining the ligand binding site in bovine OBP; those conserved in most of the proteins of this figure are indicated with a bold asterisk. Hamster aphrodisin (ham APHR) shares 34% of its amino acids with pig OBP-I. Identity with the other proteins are: 42% with bovine OBP (8), 25% with rat OBP-I (9) and 31% with both subunits of mouse OBP-I (12) respectively.

generation of a nitrene intermediate by irradiation with UV light. It was prepared from 1-aminoanthracene (Figure 7), which proved to be a good ligand for porcine OBP-I (compound 2 in Figure 8). The newly synthesized fluorescent derivative bound the protein with high affinity (compound 3 in Figure 8). The fluorescence measured for the azido derivative was much weaker, but the reaction of the nitrene intermediate generated an amino group and was accompanied by a strong increase in the fluorescence yield (data not shown). This behaviour can be efficiently exploited to monitor the covalent attachment of the molecular probe to the protein. 2-Trimethylstannyl-5-methylanisole and 2-trimethyl-stannyl-5-methymethoxylanisole (compounds 4 and 5 in Figure 8) are tin-containing analogues of a good ligand, thymol (Dal Monte *et al.*, 1993), which is used in structural analysis by X-ray diffraction spectroscopy. Their synthesis was also easily accomplished from the corresponding bromo derivatives as described in Figure 7. Other special ligands include 2-amino-4-butyl-5-propyl-selenazole, already employed in the identification of the

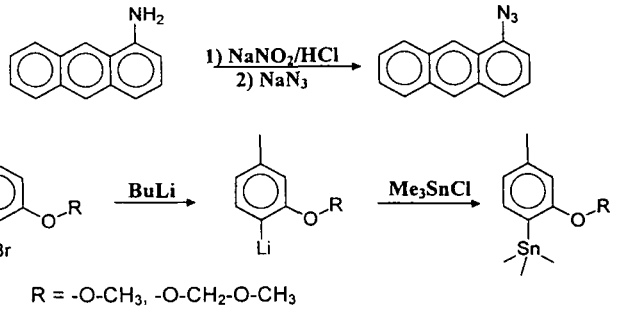

Figure 7 Synthesis schemes. The photoaffinity labelling agent 1-azido-anthracene was synthesized from 1-aminoanthracene via the intermediate diazonium salt (upper panel). The two tin-containing ligands were obtained from their corresponding bromo derivatives (lower panel) as reported.

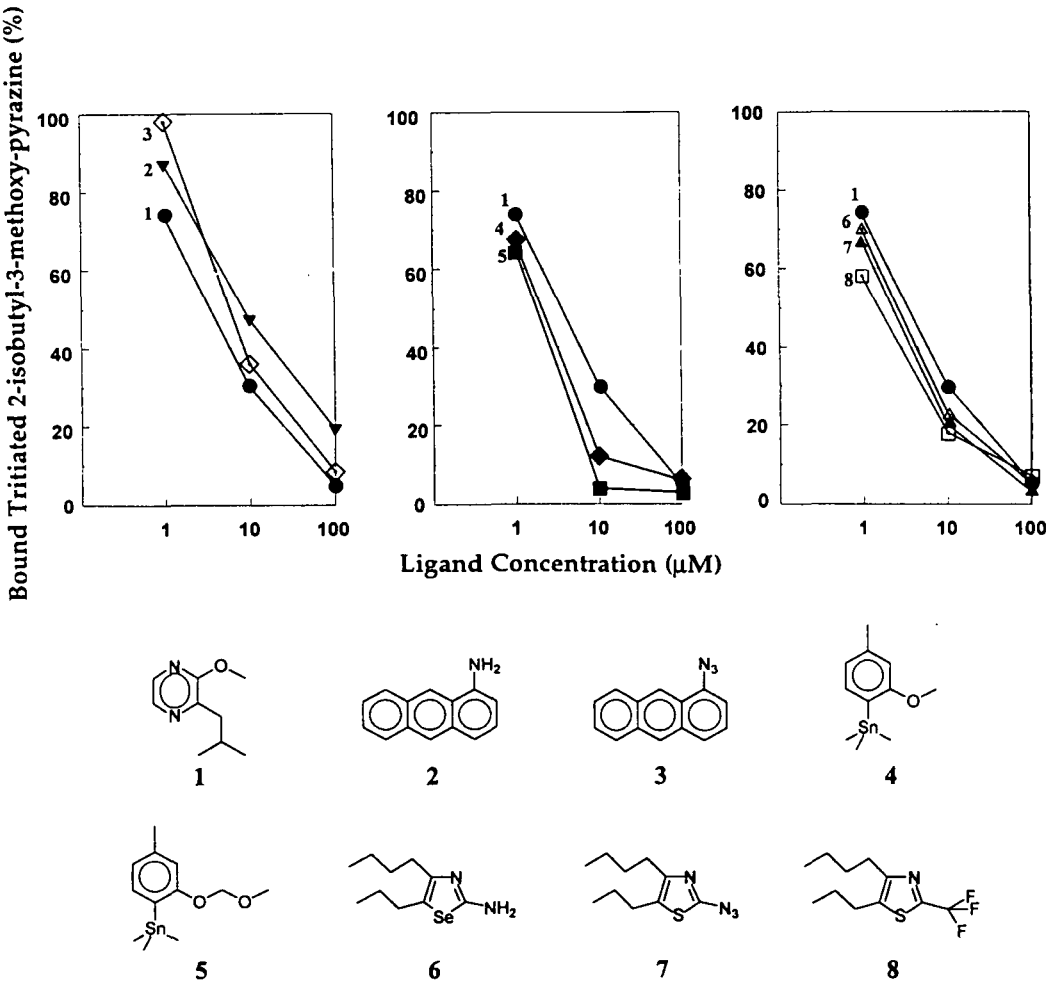

Figure 8 Competitive binding assays of different ligands with purified porcine OBP-I. Tritiated 2-isobutyl-3-methoxypyrazine was used as the radioactive ligand, at 1 μM concentration, while the displacing ligand was assayed at concentrations of 1, 10 and 100 μM respectively. The competition by unlabelled 2-isobutyl-3-methoxypyrazine is also reported as a positive control. All the compounds assayed proved to be good ligands for pig OBP-I, thus making their use feasible as specific probes in fluorescence binding assay (**2**), photoaffinity labelling experiments (**3** and **7**), X-ray diffraction studies (**4**, **5** and **6**) and NMR spectroscopy (**8**).

ligand-binding site of bovine OBP (Bianchet *et al.*, 1996), the photoaffinity label 2-azido-4-butyl-5-propylthiazole and the fluorinated NMR probe 2-trifluoromethyl-4-butyl-5-propylthiazole (compounds 6, 7 and 8 respectively in Figure 8). The synthesis of the three last compounds has been described previously (Napolitano and Pelosi, 1992).

All the above compounds proved to be good ligands for the protein in competitive binding assays, as compared with the reference ligand 2-isobutyl-3-methoxypyrazine (compound 1 in Figure 8). The 4-butyl-5-propylthiazole and thymol analogues described here presented the highest values of affinity toward porcine odorant-binding protein.

Photoaffinity labelling

In order to covalently label the protein in its ligand-binding site, pig OBP-I was incubated with the photoaffinity labelling agent 1-azidoanthracene and irradiated with UV light. The mixture, was purified by gel filtration chromatography with the aim to eliminate most of the excess ligand and then analysed by SDS–PAGE. Figure 9 shows the sharp fluorescent band clearly visible at an apparent mol. wt just above 20 kDa, coincident with the protein band of OBP-I. The fluorescence measured for the purified sample was in agreement with the stoichiometric value expected.

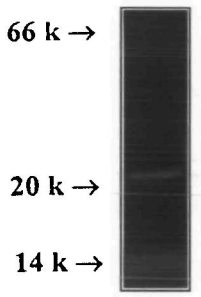

Figure 9 SDS–PAGE analysis of porcine OBP-I derivatized with photoaffinity labelling agent 1-azidoanthracene. A sharp fluorescent band migrated with an apparent molecular weight of ~22 kDa, as the native protein. The band at the bottom of the gel is due to unreacted ligand.

Discussion

The study of the physiological role of OBPs in olfactory transduction requires detailed information on the structure of these proteins, particularly relative to the ligand-binding site. Complete amino acid sequences have been reported only for a limited number of vertebrate odorant-binding proteins, compared with the much wider information available for insect OBPs. This is certainly one of the reasons why in vertebrates the classification of these proteins into subfamilies is not as clear as in insects. The available information, however, including N-terminal sequences of a number of OBPs, often indicates lower similarity between OBPs of the same species than between species, suggesting that a classification could be attempted on the basis of their amino acid sequence.

Thus, a first subclass would include pig OBP-I, bovine OBP, rat OBP-I, both subunits of mouse OBP-I and hamster aphrodisin. In a second one we might find rat OBP-II (Dear *et al.*, 1991), two OBP-like proteins present in the vomeronasal organ of the mouse, VNSP1 and VNSP2 (Miyawaki *et al.*, 1994), frog BG protein (Lee *et al.*, 1987) and the VEG proteins (Schmale *et al.*, 1990; Redl *et al.*, 1992; Garibotti *et al.*, 1995). Finally, a third subclass would include OBPs, such as mouse OBPs III and IV (Pes and Pelosi, 1995), rabbit OBP III (Garibotti *et al.*, 1997) and porcupine OBP I (Ganni *et al.*, 1997), that are very similar to urinary and salivary proteins of mouse and rat, as well to the salivary pheromone carrier protein SAL of the boar (Marchese *et al.*, 1998). Comparison of the respective hydropathy profiles, calculated according to Kyte and Doolittle (1982) (data not shown), provides further support to such a classification.

The number of cysteine residues does not appear to be a conserved motif among vertebrate OBPs. They range from complete absence in the bovine protein up to five in rat OBP-I and in the 'a' subunit of mouse OBP-I; the 'b' subunit of this latter protein instead contains four residues (Pes *et al.*, 1998), as does aphrodisin. The positions of the cysteines present, however, are rather conserved, as can be seen from Figure 6. The porcine protein only contains two cysteines, at positions 63 and 155, connected by an intramolecular disulphide bridge. Both residues are well conserved in all the proteins described in the figure, except for the bovine OBP. It is noteworthy that the same two cysteines are also conserved in other members of OBPs that bear poor similarity to those here reported: namely rat OBP-II, the above-mentioned VNSP1 and VNSP2 from mouse vomeronasal organ, and frog BG. In this context, the bovine OBP seems to represent the exception, with its complete lack of cysteines.

The sequence similarity between the porcine and bovine proteins also matches the close similarity in their binding specificity towards odorous molecules, suggesting the presence of chemically and stereochemically related binding sites in the two proteins.

So far, the bovine OBP is the only protein of this family whose three-dimensional structure has been resolved (Tegoni *et al.*, 1996; Bianchet *et al.*, 1996). The precise location of the ligand-binding site has been identified with a selenium-containing odorant, 2-amino-4-butyl-5-propyl-selenazole (Napolitano and Pelosi, 1992), a strong ligand for the protein. The amino acid residues lining the binding pocket are indicated in Figure 6 and appear to belong to different regions of the polypeptide chain. It is interesting that some of these amino acids are well conserved in the sequences reported in the figure, although much less than in those of other OBPs or OBP-like proteins, such as the urinary and salivary proteins of rat and mouse (Dinh *et al.*, 1965; Finlayson *et al.*, 1965; Shaw *et al.*, 1983; Shahan *et al.*, 1987). The conserved residues are Ile21, Phe55, Tyr/Phe79,

Asn83, Phe/Tyr85 and Ile/Leu95; these amino acids could also line the binding pockets of pig OBP-I and of the other proteins reported in the figure.

Such a highly conserved motif at the level of the ligand-binding site is apparently obtained with different molecular architectures in the bovine and porcine OBPs. In the first protein cysteine residues are completely absent, while the second contains two cysteines involved in an intramolecular disulphide bridge. Moreover, while the bovine OBP is a homodimer in native conditions, the pig OBP-I is a monomer. It is possible that, given the absence of cysteines, the bovine protein achieves a compact structure by assembling two monomers, interacting in the unusual system of 'domain swapping', while a certain rigidity in the monomeric porcine protein is provided by the presence of the mentioned S–S bridge.

The experiments with 1-azidoanthracene have indicated that this ligand can be efficiently used in covalently photolabelling the binding site. The fluorescence of the amino precursor, which is restored in the modified azido compound only after the photoaffinity reaction, provides evidence for the chemical reaction having occurred. The labelled protein could be selectively hydrolysed with the same chemical and enzymatic reagents used in this work, to afford peptides to be analysed by MALDIMS. A comparison of the spectra thus obtained with the native and the labelled protein would indicate the regions of the protein chain covalently bound to the 1-aminoanthracene residue (A. Scaloni and P. Pelosi, in preparation).

The location of the binding site will be studied also by the use of ligands containing unusual atoms, such as the selenium- or tin-containing derivatives here reported, in the X-ray diffraction maps obtained for the crystals of OBP-I grown in the presence of these compounds.

Acknowledgement

The pig nasal tissue was obtained with the kind cooperation of Dr Della Longa at the abattoir of Pontedera (Pisa).

References

Bianchet, M.A., Bains, G., Pelosi, P., Pevsner, J., Snyder, S.H., Monaco, H.L. and **Amzel, L.M.** (1996) *The three dimensional structure of bovine odorant-binding protein and its mechanism of odor recognition*. Nature Struct. Biol., 3, 934–939.

Breer H., Boekhoff I., Krieger J., Raming K., Strotmann J. and **Tareilus, E.** (1992) *Molecular mechanism of olfactory signal transduction*. In Corey, D.P. and Roper, S.D. (eds), Sensory Transduction. The Rockefeller University Press, New York, pp. 94–108.

Bruns, R.F., Lawson-Wendling, K. and **Pugsley, T.A.** (1983) *A rapid filtration assay for soluble receptors using polyethylenimine-treated filters*. Anal. Biochem., 132, 74–81.

Dal Monte, M., Andreini, I., Revoltella, R. and **Pelosi, P.** (1991) *Purification and characterization of two odorant binding proteins from nasal tissue of rabbit and pig*. Comp. Biochem. Physiol., 998, 445–451.

Dal Monte, M., Centini, M., Anselmi, C. and **Pelosi, P.** (1993) *Binding of selected odorants to bovine and porcine odorant-binding proteins*. Chem. Senses, 18, 713–721.

Dear, T.N., Campbell, K. and **Rabbitts, T.H.** (1991) *Molecular cloning of putative odorant-binding and odorant-metabolizing proteins*. Biochemistry, 30, 10376–10382.

Dinh, B.L, Tremblay, A. and **Dufour, D.** (1965) *Immunochemical study of rat urinary proteins, their relation to serum and kidney proteins*. J. Immunol., 95, 574–582.

Finlayson, J.S., Asofsky, R., Potter, M. and **Runner, C.C.** (1965) *Major urinary protein complex of normal mice: origin*. Science, 149, 981–982.

Flower, D.R. (1996) *The lipocalin protein family: structure and function*. Biochem. J., 318, 1–14.

Ganni M., Garibotti, M., Scaloni, A., Pucci, P. and **Pelosi, P.** (1997) *Microheterogeneity of odorant-binding proteins in the porcupine revealed by N-terminal sequencing and mass spectrometry*. Comp. Biochem. Physiol., 117B, 287–291.

Garibotti, M., Christiansen, H., Schmale, H. and **Pelosi, P.** (1995) *Porcine VEG proteins and tear prealbumins*. Chem. Senses, 20, 69–76.

Garibotti, M., Navarrini, A., Pisanelli, A.M. and **Pelosi, P.** (1997) *Three odorant-binding proteins from rabbit nasal mucosa*. Chem. Senses, 22, 383–390.

Henzel, W.J., Rodriguez, H., Singer, A.G., Stults, J.T., Macrides, F., Agosta, W.C. and **Niall, H.** (1988) *The primary structure of aphrodisin*. J. Biol. Chem., 263, 16682–16687.

Hérent, M.F., Collin, S. and **Pelosi, P.** (1994). *Affinities of nutty and green-smelling compounds to odorant-binding proteins*. Chem. Senses, 20, 601–610.

Holzfeind, P., Merschak, P., Rogatsch, H. Culig, Z., Feichtinger, H. Klocker, H. and **Redl, B.** (1996) *Expression of the gene for tear lipocalin/von Ebner's gland protein in human prostate*. FEBS Lett., 395, 95–98.

Kyte, J. and **Doolittle, R.F.** (1982) *A simple method for displaying the hydropathic character of a protein*. J. Mol. Biol., 157, 105–132.

Laemmli, U.K. (1970) *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*. Nature, 227, 680–685.

Lee, H.K., Wells, R.G. and **Reed, R.R.** (1987) *Isolation of an olfactory cDNA: similarity to retinol binding protein suggests a role in olfaction*. Science, 253, 1053–1056.

Magert, H.J., Hadrys, T., Cieslak, A., Groger, A., Feller, S. and **Forssmann, W.G.** (1995) *cDNA sequence and expression pattern of the putative pheromone carrier aphrodisin*. Proc. Natl Acad. Sci USA, 92, 2091–2095.

Marchese, S., Pes, D., Scaloni, A., Carbone, V. and **Pelosi, P.** (1998) *Lipocalins of boar salivary glands binding odours and pheromones*. Eur. J. Biochem., 252, 563–568.

Matsuidara, P. (1987) *Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes*. J. Biol. Chem., 262, 10035–10038.

Miyawaki, A., Matsushita, F., Ryo, Y. and **Mikoshiba, K.** (1994) *Possible pheromone-carrier function of two lipocalin proteins in the vomeronasal organ*. EMBO J., 13, 5835–5842.

Napolitano, E. and **Pelosi, P.** (1992) *Synthesis of thiazole and selenazole derivatives with affinity for the odorant-binding protein*. Bioorg. Med. Chem. Lett., 2, 1603–1606.

Palmieri, G., Di Palo, M., Scaloni, A., Orru', S., Marino, G. and **Sannia,**

G. (1996) *Glutamate 1-semialdehyde aminotransferase from* Sulfolobus solfataricus. Biochem. J., 320, 541–545.

Pelosi, P. (1994) *Odorant-binding proteins*. Crit. Rev. Biochem. Mol. Biol., 29, 199–228.

Pelosi, P. (1996) *Perireceptor events in olfaction*. J. Neurobiol., 30, 3–19.

Pelosi, P. and **Maida, R.** (1990) *Odorant binding proteins in vertebrates and insects: similarities and possible common function*. Chem. Senses, 15, 205–215.

Pelosi, P. and **Tirindelli, R.** (1989) *Structure/activity studies and characterization of an odorant-binding protein*. In Brand, J.G., Teeter, J.H., Cagan, R.H. and Kare, M.R. (eds), Chemical Senses. Vol. 1. Receptor Events and Transduction in Taste and Olfaction. Marcel Dekker, New York, pp. 207–226.

Pelosi, P., Baldaccini, N.E. and **Pisanelli, A.M.** (1982). *Identification of a specific olfactory receptor for 2-isobutyl-3-methoxypyrazine*. Biochem. J., 201, 245–248.

Pes, D. and **Pelosi, P.** (1995). *Odorant-binding proteins of the mouse*. Comp. Biochem. Physiol., 112B, 471–479.

Pes, D., Mameli, M., Andreini, I., Krieger, J., Weber, M., Breer, H. and **Pelosi, P.** (1998) *Cloning and expression of odorant-binding proteins Ia and Ib from mouse nasal tissue*. Gene, 212, 49–55.

Pevsner, J., Reed, R.R., Feinstein, P.G. and **Snyder, S.H.** (1988) *Molecular cloning of odorant-binding protein: member of a ligand carrier family*. Science 241, 336–339.

Pevsner, J., Hou, V., Snowman, A.M. and **Snyder, S.H.** (1990) *Odorant-binding protein: characterization of ligand binding*. J. Biol. Chem, 265, 6118–6125.

Redl, B., Holzfeind, P. and **Lottspeich, F.** (1992) *cDNA cloning and sequencing reveals human tear prealbumin to be a member of the lipophilic-ligand carrier protein superfamily*. J. Biol. Chem., 267, 20282–20287.

Sansom C.E., North A.C.T. and **Sawyer L.** (1994) *Structural analysis and classification of lipocalins and related proteins using a profile-search method*. Biochim. Biophys. Acta, 1208, 247–255.

Schägger, H. and **von Jagow, G.** (1987) *Tricine–sodium dodecylsulfate–polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa*. Anal. Biochem., 166, 368–379.

Schmale, H., Holtgreve-Grez, H. and **Christianse, H.** (1990) *Possible role for salivary gland protein in taste reception indicated by homology to lipophilic-ligand carrier proteins*. Nature, 343, 366–369.

Shahan, K.M., Denaro, M., Gilmartin, M., Shi, Y. and **Derman, E.** (1987) *Expression of six mouse major urinary protein genes in the mammary, parotid, sublingual, submaxillary and lachrymal glands and in the liver*. Mol. Cell. Biol., 7, 1947–1954.

Shaw, P.H., Held, W.A. and **Hastie, N.D.** (1983) *The gene family for major urinary proteins, expression in several secretory tissues of the mouse*. Cell, 32, 755–761.

Tegoni, M., Ramoni, R., Bignetti, E., Spinelli, S. and **Cambillau, C.** (1996) *Domain swapping creates a third putative combining site in bovine odorant binding protein dimer*. Nature Struct. Biol., 3, 863–867.

Tirindelli, R., Keen, J.N., Cavaggioni, A., Eliopoulos, E.E. and **Findlay, J.B.C.** (1989) *Complete amino acid sequence of pyrazine-binding protein from cow nasal mucosa*. Eur. J. Biochem., 185, 569–572.

Vogt R.G. and **Riddiford L.M.** (1981) *Pheromone binding and inactivation by moth antennae*. Nature, 293, 161–163.

Vogt R.G., Rybczynski R. and **Lerner M.R.** (1990) *The biochemistry of odorant reception and transduction*. In Schild, D. (ed.), NATO ASI Series H, Vol. 39, Chemosensory Information Processing. Springer-Verlag, Berlin, pp. 33–76.

Accepted July 17, 1998