J Chem Ecol (2007) 33:613–630
DOI 10.1007/s10886-006-9247-0

Characterization and Comparison of Major Urinary Proteins from the House Mouse, *Mus musculus domesticus*, and the Aboriginal Mouse, *Mus macedonicus*

Duncan H. L. Robertson · Jane L. Hurst · Jeremy B. Searle · İslam Gündüz · Robert J. Beynon

Received: 15 September 2006 / Accepted: 18 December 2006 /
Published online: 1 February 2007

Abstract Urine from the house mouse, *Mus musculus domesticus*, contains a high concentration of major urinary proteins (MUPs), which convey olfactory information between conspecifics. In wild populations, each individual expresses a different pattern of around 8 to 14 electrophoretically separable MUP isoforms. To examine whether other *Mus* species express MUPs and exhibit a similar level of individual heterogeneity, we characterized urinary proteins in urine samples from an aboriginal species, *Mus macedonicus*, captured from different sites in Turkey. Anion exchange chromatography and electrospray ionization mass spectrometry demonstrated that *M. macedonicus* urine contained a single major peak of mass 18,742 Da, and in contrast to *M. m. domesticus*, all individuals were the same. The *M. macedonicus* masses were not predicted from any known MUP gene sequence. Endoproteinase Lys-C (Lys-C) digestion of the purified *M. macedonicus* urinary protein followed by matrix assisted laser desorption time of flight (MALDI-TOF) mass spectrometry demonstrated that it shared considerable, but not complete, sequence homogeneity with *M. m. domesticus* MUPs. Three *M. macedonicus* Lys-C peptides differed in mass from their *M. m. domesticus* counterparts. These three peptides were further characterized by tandem mass spectrometry. The complete sequences

D. H. L. Robertson · R. J. Beynon (✉)
Protein Function Group, Faculty of Veterinary Science,
University of Liverpool, PO Box 147, Liverpool, L69 3BX, UK
e-mail: r.beynon@liverpool.ac.uk

D. H. L. Robertson
e-mail: dhlr@liverpool.ac.uk

J. L. Hurst
Mammalian Behaviour & Evolution Group Faculty of Veterinary Science,
University of Liverpool, Leahurst, Neston, CH64 7TE, UK
e-mail: jhurst@liverpool.ac.uk

J. B. Searle · İ. Gündüz
Department of Biology (area 2), University of York, PO Box 373, York, YO10 5YW, UK

İ. Gündüz
Department of Biology, Faculty of Arts and Sciences, University of Ondokuz Mayis, Samsun, Turkey

of two were determined, and in conjunction with methyl esterification, the amino acid composition of the third was inferred, and the sequence narrowed down to three permutations. The complete *M. macedonicus* sequence contained a maximum of seven amino acid substitutions, discernible by tandem mass spectrometry, relative to a reference *M. m. domesticus* sequence. Six of these were on the surface of the molecule. Molecular modeling of the *M. macedonicus* sequence demonstrated that the amino acid substitutions had little effect on the tertiary structure. The differences in the level of heterogeneity between the two species are discussed in relation to their environment and behavior. In addition, the differences in protein structure allow speculation into molecular mechanisms of MUP function.

Keywords Aboriginal mouse · *De novo* protein sequencing · House mouse · Major urinary proteins · Peptide mass spectrometry · Protein mass spectrometry · Protein sequence heterogeneity

Introduction

House mice, *Mus musculus domesticus*, have evolved intricate and sophisticated methods of chemical communication (Hurst, 1987; Hurst and Beynon, 2004). The major urinary proteins (MUPs) are key components of this signaling system (Beynon and Hurst, 2004). MUPs are eight stranded beta-barrel pheromone binding proteins (Adams and Sawyer, 1990; Flower et al., 1993), encoded by a multigene family, expressed in the liver, and secreted into the urine via the kidneys (Finlayson et al., 1965). A number of small, volatile molecules have been identified as endogenous MUP ligands, most of which are reproductive priming pheromones (Bacchini et al., 1992; Robertson et al., 1993; Novotny et al., 1999) that also influence behavior (Malone et al., 2001). MUPs prolong the release of the volatile ligands from scent marks and may also protect them from oxidation (Hurst et al., 1998). However, this does not explain the extreme polymorphism that is a feature of these proteins in *M. m. domesticus*. Each individual expresses around 8 to 14 electrophoretically separable MUP isoforms, with only very closely related individuals (Hurst et al., 2001) or inbred laboratory mice (Robertson et al., 1996) expressing the same pattern. We have demonstrated that this extreme polymorphism provides a genetically stable method of communicating the individual ownership of scent marks (Hurst et al., 2001, 2005).

Knowledge of the structure and function of MUPs has mostly been derived from *M. m. domesticus*. These mice are thought to have become commensal some 10,000 yrs ago in the Fertile Crescent, utilizing some of the earliest human settlements there (Cucchi et al., 2005). Therefore, the history of *M. m. domesticus* has become inextricably intertwined with that of humans. Their innate agility and flexibility has proved ideal for the continued exploitation of human populations as a ready source of food and other resources, which has facilitated their spread from the Fertile Crescent to all parts of the world (Silver, 1995). Where food resources are abundant (for example around grain stores and livestock housing), population densities typically reach high levels (Berry, 1981; Bronson, 1979), and mice live in territorial social groups in which the ranges of many individuals overlap (Hurst, 1987; Barnard et al., 1991). *Mus m. domesticus* is one of at least three subspecies of house mice that have parapatric distributions and that interbreed where they make contact (Boursot et al., 1993). There are also three other *Mus* species (*Mus macedonicus*, *Mus spretus*, and *Mus spicilegus*) that are closely related to and occur sympatrically with *M. musculus* in Europe

and the Middle East (Boursot et al., 1993; Suzuki et al., 2004). These species are termed aboriginal mice as, unlike the commensal *M. musculus*, they live independently of humans. Commensal mouse populations differ in a number of aspects of their behavior compared to free living aboriginal species (Patris and Baudoin, 1998; Ivantcheva and Cassaing, 1999). Here, we studied the aboriginal mouse, *M. macedonicus*, a short-tailed species that is found in Greece, Turkey, and elsewhere in the Middle East. The aim of this study was to characterize MUPs from *M. macedonicus* and compare them to the well-characterized MUPs from *M. m. domesticus*.

Methods and Materials

Animals and Sampling

Adult male *M. macedonicus* and *M. m. domesticus* urine donors were caught during May–July 2001 in Sherman traps from four locations dispersed around Turkey (Fig. 1). Urine voided by the animals in response to handling (typically 20–50 μl) was collected into 1.5 ml capped polypropylene test tubes and frozen within 5 min. All samples were subsequently stored at −20°C before analysis. Reference MUPs were obtained from pooled urine samples from adult male C57BL/6 mice. These were obtained by bladder massage, frozen immediately, and stored at −20°C.

Urine Sample Preparation

Urine samples were desalted before further analysis with Vivaspin 5,000 Da molecular weight cutoff centrifugal concentrators (Sartorius, Epsom, UK). Typically, a 50 μl urine sample was placed in a prerinsed concentrator and made up to a final volume of 500 μl with distilled water. The sample was then centrifuged at 10,500×*g* for 15 min, which was sufficient to reduce it to 50 μl. This was then expanded to 500 μl with distilled water, and the

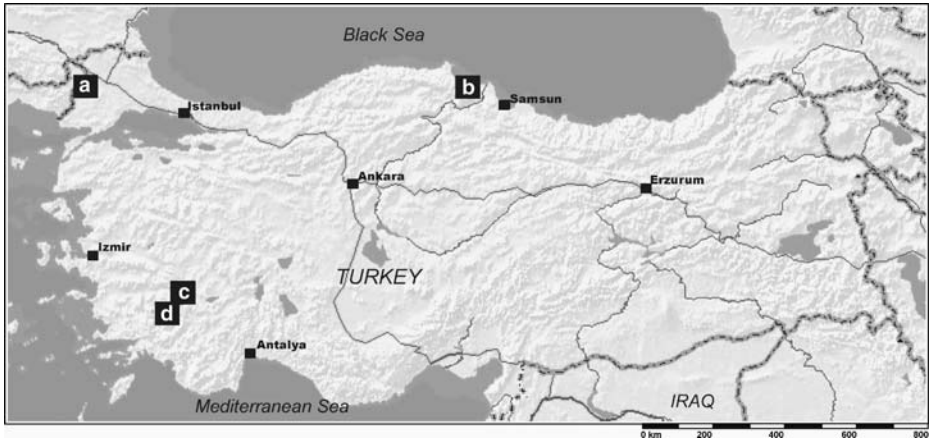


Fig. 1 Map of *Mus macedonicus and M. musculus domesticus* collection sites. Wild specimens of *M. macedonicus* and *M. m. domesticus* were collected for urine samples from four locations in Turkey, labeled (a–d). *Mus macedonicus* were caught in sites: (a), Eskiköy, Uzunkoprü, Edirne, 41°20′ N 26°37′ E (sample 123) and (c), Korucuk, Denizli, 37°49′ N 29°09′ E (samples 129 and 131), whereas *M. m. domesticus* were from (b), Kozköy, Yakakent, Samsun, 41°38′ N 35°31′ E (sample 171) and (d), Denizli, 37°47′ N 28°59′ E (samples 141 and 145)

process was repeated. The sample was then removed from the apparatus and stored at −20°C until required. An identical process was used to desalt fractions collected from anion exchange chromatography before analysis by electrospray ionization mass spectrometry (ESI-MS)

Ion Exchange Chromatography

Anion exchange chromatography was performed on a Dionex Bio-LC platform fitted with a Dionex ProPac SAX column (2×250 mm) and a ProPac SAX guard column (2×50 mm). In all cases, the column flow rate was 0.2 ml/min. After equilibration with 20 mM Tris, pH 8.5, desalted urine samples, (typically 1–5 μl) were loaded onto the column. Samples were eluted from the column with a linear NaCl gradient of 0–500 mM in 30 min. The eluent from the column was monitored at 214 nm in a flow cell of 9 mm path length. Where applicable, fractions were collected by hand directly after passage through the flow cell. All aspects of data acquisition and processing were controlled through the Dionex Chromeleon software.

Proteolysis of MUPs

Anion exchange purified MUPs were proteolyzed by endoproteinase Lys-C (Lys-C), endoproteinase Glu-C (Glu-C), or trypsin (Roche Diagnostics, Lewes, UK). An aliquot (typically 100 μl) of the MUP fraction was reduced with 10 mM 2-mercaptoethanol for 1 hr at room temperature (RT) or with 10 mM dithiothreitol (DTT) at 55°C for 1 hr. In some instances, cysteine residues were carbamidomethylated by incubating the reduced protein preparation with a 55-mM final concentration of iodoacetamide for 1 hr at RT. Removal of reduction and carbamylation reagents was achieved by protein precipitation or centrifugal filtration in a Vivaspin apparatus as described previously. Protein precipitation was achieved by the addition of an equal volume of 20% (w/v) tricarboxylic acid (TCA). After centrifugation at 14,000×*g* for 10 min, the supernatant was discarded, and the precipitate was washed twice with diethyl ether. Residual ether was removed from the sample by incubation at 40°C for 10 min, and the precipitate was resuspended in its starting volume of either 100 mM Tris, pH 8.5 (Lys-C and Glu-C) or 100 mM Tris/2 mM $CaCl_2$, pH 8.5 (trypsin). Protease solution (2 μl of 0.1 mg/ml) was added to the resuspended precipitate, which was then incubated overnight at 37°C. The reaction was subsequently stopped by addition of a 5-μl aliquot of formic acid. Samples prepared by Vivaspin desalting were digested under identical conditions.

Esterification of MUP Lys-C Peptides

Esterification of carboxyl groups within MUP Lys-C peptides was achieved according to Shevchenko et al. (2003). Lys-C peptides from digests of *M. macedonicus* and *M. m. domesticus* MUPs were initially desalted into a solution of 0.1% (v/v) trifluoro-acetic acid (TFA)/50% (v/v) acetonitrile with Zip-Tips (Millipore, Billerica, MA, USA). This solution was subsequently reduced to dryness in a vacuum centrifuge. A 1 ml aliquot of ethanol was placed in a 1.5-ml test tube and placed at −20°C for 15 min, after which, a 150 μl aliquot of acetyl chloride was added, and the mixture was incubated at room temperature for 10 min. The dried peptides were immediately treated with 15 μl of this mixture and incubated at room temperature for 45 min before drying in a vacuum centrifuge. The dried, esterified peptides were finally dissolved in MALDI matrix solution (see below) before matrix assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry.

Matrix Assisted Laser Desorption Ionization-Time of Flight (MALDI-TOF) Mass Spectrometry

MALDI-TOF mass spectrometry of MUP digests was performed on a Waters-Micromass M@LDI instrument. Peptide mixtures from proteolytic digests of MUPs were mixed in a 1:1 ratio with a matrix solution consisting of saturated α-cyano-4-hydroxy-cinnaminic acid (Sigma Chemical, Poole, UK) in 50% (v/v) acetonitrile/0.2% (v/v) TFA. A 1 μl aliquot of this preparation was deposited on the MALDI target and allowed to dry at RT. Spectra were subsequently acquired between 1,000 and 4,000 Th with the laser energy optimized to give the best signal to noise ratio for each sample. The laser firing rate was 5 Hz and 10 spectra (collected over 2 sec) were combined. The final mass spectrum was a combination of 10–15 such combined data sets, representing 100–150 individual laser shots. All aspects of data acquisition, processing, and machine management were controlled through the MassLynx software suite (versions 3.5 and 4.0).

Electrospray Ionization (ESI) Mass Spectrometry

Electrospray ionization mass spectrometry (ESI-MS) and tandem mass spectrometry (ESI-MSMS) were performed on a Micromass Q-ToF Micro instrument, fitted with a nanospray source. The electrospray was created from a silver-coated glass capillary with a 10-μm orifice (New Objective, Woburn, MA, USA), held at a potential of +2,000 V relative to the sample cone. For measurement of the mass of intact MUPs, a desalted sample was introduced into the mass spectrometer by syringe pump infusion (Harvard Instruments Ltd, Edenbridge, UK) at a rate of 0.5 μl/min. In this case, the instrument was operated in TOF only mode, with the quadrupole analyzer operating in R_f only mode to allow transmission of all ions. Raw data were gathered between 700 and 1,400 Th at a scan/interscan speed of 2.4/0.1 sec. These raw data were subsequently de-convoluted using the MaxEnt 1 module contained within the MassLynx 3.5 software. To create the MaxEnt damage model, peak width and resolution parameters of 0.75 and 1 Da/channel were used, respectively, and data were processed over the mass range 18,400–19,000 Da.

De novo Sequencing

For tandem mass spectrometry on peptides derived from MUP proteolytic digests, samples were introduced into the mass spectrometer at 0.2 μl/min following reversed phase high pressure liquid chromatography (RP-HPLC) (see below). In this case, the initial quadrupole analyser was set to allow passage only of selected precursor ions to a gas cell, where they were fragmented by collision with argon. The mass of the resulting fragment ions was then measured by the TOF analyzer. Selection of precursors and fragmentation energy were controlled automatically by using the data dependent acquisition facility within the MassLynx software. Precursor spectra were acquired between m/z 400 and 1,500 at a scan/interscan speed of 2.4/0.1 sec. Product ion spectra were acquired between m/z 100 and 2,000 at a scan/interscan speed of 1.0/0.1 sec. Raw product ion spectra were de-convoluted using the MaxEnt 3 algorithm within the MassLynx software. The charge state of the parent peptide was determined from the isotope envelope in the precursor spectrum. Interpretation of product ion spectra and the determination of peptide sequences were facilitated by the PepSeq module within MassLynx.

Reverse Phase-HPLC

Before ESI-MSMS, MUP peptides were separated by RP-HPLC on a Dionex Ultimate system. The system was fitted with a PepMap C18 column (LC Packings, Camberley, UK), 15 cm×75 μm, bead size 3 μm and pore size 100 Å. Before separation, aliquots (2–5 μl) of MUP peptides were desalted in-line using a Dionex Switchos apparatus, fitted with a 1 mm×300 μm, C18 precolumn. The precolumn was initially equilibrated in 0.2% (v/v) formic acid at 30 μl/min. Peptides were then loaded and washed for 3 min at the same flow rate and then eluted with 90% acetonitrile/0.2% formic acid, introduced as a linear gradient of 0–50% in 30 min at 0.2 μl/min. The column eluent was monitored by UV absorbance at 214 nm, before delivery to the mass spectrometer.

Molecular Modeling

Protein sequences derived from the *M. macedonicus* MUPs were used to direct homology models. Briefly, protein structures were modeled in the structure of MUP1 1I04.PDB (Timm et al., 2001). After alignment of the two sequences (no insertions or deletions were necessary), the models were built to the highest quality possible using the Modeler package of the Discovery Studio package (Accelrys, Cambridge, UK, ver 1.5). Because the four C-terminal most residues of MUP1 cannot be defined because of main chain flexibility (Krizova et al., 2004), these residues were not modeled.

Results and Discussion

We analyzed urine from three samples each of *M. macedonicus* and *M. m. domesticus*, obtained from mice caught from geographically distinct locations (Fig. 1). The two species demonstrated different patterns of MUP expression. When the urinary proteins were resolved by ion exchange chromatography, the elution profile of *M. macedonicus* urine consisted of a single major peak and two barely detectable smaller peaks (Fig. 2a). The retention time of the major peak expressed by each individual was similar, (maximum absorbance at 21.69, 21.83, and 21.86 min), suggesting little difference in their MUP profiles. By contrast, the *M. m. domesticus* samples were noticeably more complex and varied among individuals in the number of major and minor peaks (Fig. 2b). Furthermore, the retention times of the major peaks differed by more than 0.25 min among samples, contrasting strongly with the invariable elution times of the major peak in the *M. macedonicus* samples. It is probable that such differences in elution time reflect different proteins (Robertson et al., 1996). Each of the three *M. m. domesticus* samples also differed in the relative proportions of the peaks, which added to the overall level of heterogeneity. For *M. m. domesticus*, such heterogeneity is not surprising and is similar to that observed previously in wild *M. m. domesticus* populations (Robertson et al., 1997; Pes et al., 1999; Veggerby et al., 2002; Beynon and Hurst, 2004). Of greater interest was the relative simplicity and the lack of comparable heterogeneity in the *M. macedonicus* urine samples, which we have never observed in *M. m. domesticus* samples caught from many different locations.

We previously demonstrated the utility of mass spectrometric analysis of intact urinary MUPs (Evershed et al., 1993; Robertson et al., 1996, 1997; Beynon et al., 2002). All six samples generated complex and overlapping mass spectra corresponding to multiple charged protein variants that were readily transformed into true mass spectra (Fig. 3). The

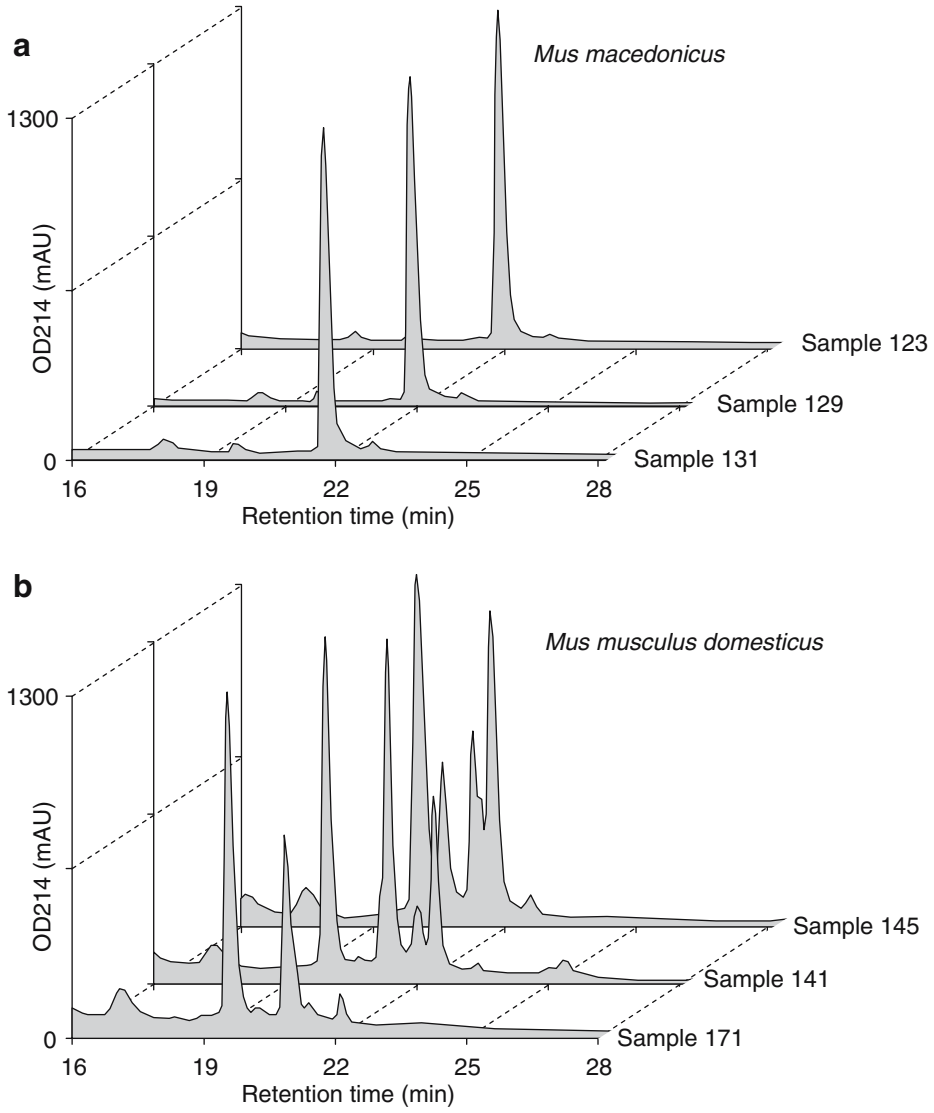


Fig. 2 Anion exchange chromatography profiles of MUPs from three urine samples of (a) *Mus macedonicus* (123, 129, and 131) and (b) *M. m. domesticus* (141, 145, and 171). MUPs were desalted with a centrifugal concentrator (see Methods and Materials section for details). Anion exchange chromatography was performed on a 2×250 mm Propac SAX anion exchange column mounted on a Bio LC chromatography platform (Dionex, Camberly, UK). The column was equilibrated with 20 mM Tris buffer, pH 8.5 at 0.2 ml/min, before loading aliquots (5 μl) of desalted MUPs from the same *M. macedonicus* and *M. m. domesticus* samples as described previously. After loading, each sample was washed for 5 min, after which, the sample was eluted with a linear NaCl gradient of 0–500 mM in 30 min. Proteins eluting from the column were detected by monitoring the absorbance at 214 nm of the column eluate, in a flow cell of 1 cm path length

three *M. macedonicus* samples were characterized by a dominant mass peak at 18,742 Da (Fig. 3a). The close agreement between these three masses (0.4 Da) is within the expected instrument error (approximately ± 2 Da) and is consistent with all three samples containing the same, single protein species. The spectra all contained a second smaller peak at 18,723,

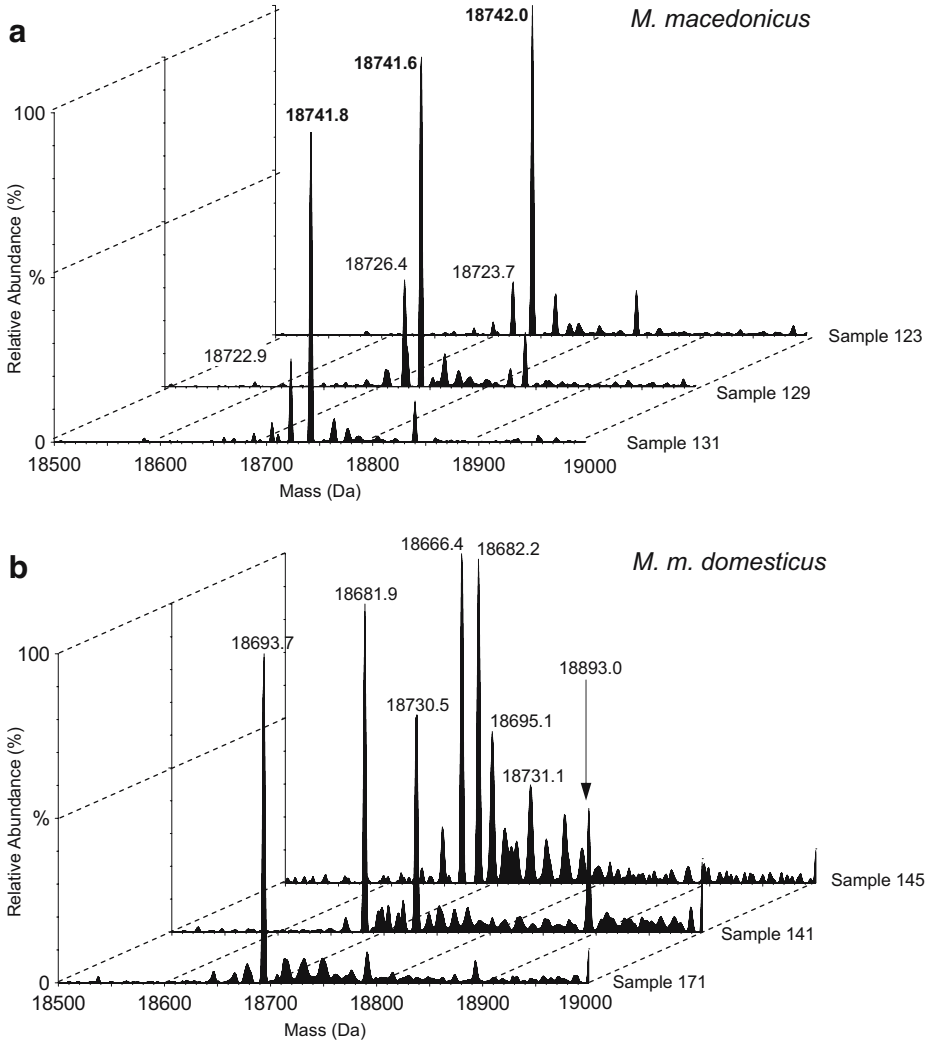


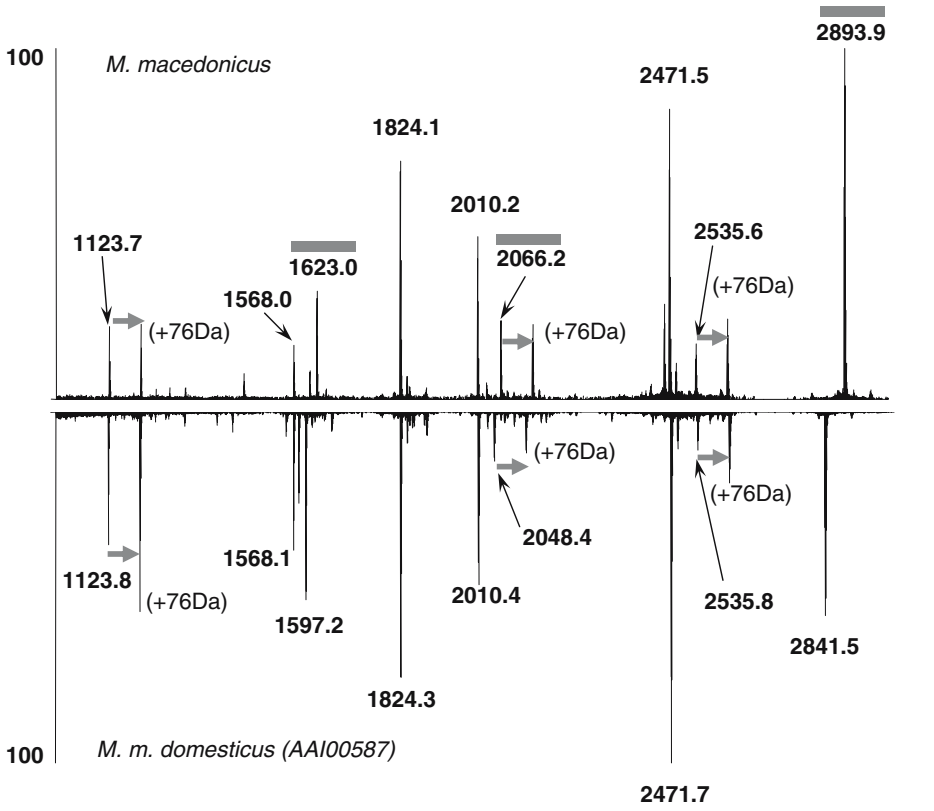
Fig. 3 ESI Mass Spectra of Unfractionated MUPs from three urine samples of (a) *Mus macedonicus* (123, 129, and 131) and (b) *M. m.domesticus* (141, 145, and 171). MUPs were separated (Fig. 2) and diluted 1:50 with a solution of 50% (v/v) acetonitrile/0.2% formic acid. This was introduced into a Q-tof Micro mass spectrometer (Micromass plc, Manchester UK) at 0.5 μl/min from a syringe infuser (Harvard Instruments, Holliston, MA, USA). The mass spectrometer was operated in TOF only mode, and raw data were gathered between m/z 700 and 1,400 with a scan/interscan time of 2.4/0.1 sec, respectively. Finally, raw data were deconvoluted and transformed to a true mass scale using MaxEnt 1 software

18,726, and 18,724 Da that probably also corresponds to a single protein. Two other features were also common to these mass spectra; a small peak at 18,763 Da and a slightly more abundant peak at 18,840 Da. The former, 21 Da higher than the base peak, is consistent with formation of a sodium adduct, whereas the latter is 98 Da higher than the base peak, consistent with the formation of a phosphate or sulfate adduct. We have commonly observed these artifacts in previous analyses, and it is unlikely that these reflect additional proteins.

The mass spectra obtained from urine samples of *M. m. domesticus* were more complex and heterogenous among individuals (Fig. 3b) than found for *M. macedonicus*, even though the two species were both collected from the same geographical region (Fig. 1). The number of major peaks (>20% base peak intensity) differed greatly. Sample 171 was the simplest, and unusually, expressed a single major peak at 18,694 Da, whereas sample 141 contained three peaks at 18,682, 18,730, and 18,893 Da. Sample 145 was most complex with peaks at 18,666, 18,682, 18,695, and 18,731 Da. Some of these masses have previously been observed in other wild mice or in inbred mouse strains, notably the protein at 18,694 Da, in C57BL/6, BALB/c, and wild mice (Robertson et al., 1996, 1997); 18,682 Da in wild mouse populations (Robertson et al., 1997), and 18,893 Da in inbred C57BL/6 mice (Armstrong et al., 2005).

Urine samples from *M. macedonicus* exhibited a single peak on anion exchange chromatography comprising a protein of mass 18,742 Da. To confirm that this protein was a MUP, and to investigate its relationship to known *M. m. domesticus* MUPs, the primary structure of the protein was analyzed by mass spectrometry. The protein was digested by endopeptidase Lys-C (Lys-C) in parallel with a MUP from urine of the inbred strain C57BL/6. Lys-C was selected over other proteases for this analysis as it was known from previous work to produce a small number of large peptides that were readily analyzed by MALDI-TOF mass spectrometry. This has previously allowed confirmation of >95% of the amino acid sequence from *M. m. domesticus* MUPs (Robertson et al., 1996). The amino acid sequence of the C57BL/6 reference MUP (GenBank accession number AAI00587, molecular mass 18,644 Da) had previously been inferred by comparison of intact protein molecular mass to the theoretical mass predicted by the cDNA sequence (Robertson et al., 1996) and comparison of Lys-C peptide masses to those predicted by the cDNA sequence (unpublished observation). This particular MUP was chosen on experimental grounds, as it was readily purified from C57BL/6 urine. The resultant peptides were surveyed by MALDI-TOF mass spectrometry. The mass spectrum of C57BL/6 MUP Lys-C peptides (Fig. 4), displayed nine masses corresponding to those predicted by the cDNA sequence (AAI00587). This analysis confirmed the identity of 156 of the 162 amino acids in the sequence (>96% coverage). The remaining unassigned amino acids were two tripeptides, the masses of which were too low for MALDI-TOF analysis. Three peptides (1123.8, 2048.4, and 2535.8 Th) were associated with additional "shadow" peaks at +76 Da. The predicted sequence for each of these peptides contains a cysteine residue that was maintained in the reduced state during proteolytic digestion by the addition of 2-mercaptoethanol. The 76-Da mass increase is consistent in mass with the formation of a β-mercaptoethanol adduct with those cysteine residues, providing good confirmation of the presence of cysteine residues in these peptides.

The corresponding mass spectrum of *M. macedonicus* Lys-C peptides revealed six peaks which shared masses with peptides in the C57BL/6 reference spectrum (Fig. 4). This was presumptive evidence for the identity of 100 amino acid residues and that the 18,742-Da *M. macedonicus* protein was indeed a MUP. The amino acid coverage in this instance was 61% (assuming a total of 162 residues). Three of the Lys-C peptides predicted from the cDNA sequence and present in the *M. m. domesticus* sample were absent from the *M. macedonicus* MUP (LysC peptides L2, L4, and L5), whereas the *M. macedonicus* sample contained three different peptides (m/z 1,623.0, 2,066.2, and 2,893.9). These peptides, likely to be the equivalent peptides to L2, L4, and L5, were selected for further MS analysis. The presence of a +76-Da adjacent peak (putatively a 2-mercaptoethanol adduct) to the peptide at m/z 2,066.2, suggested that in common with the equivalent *M. m. domesticus* peptide L5, it contained a cysteine residue.



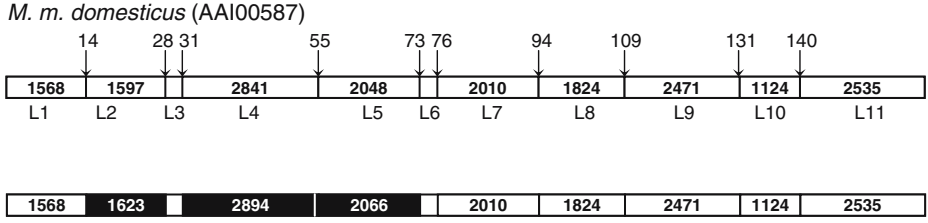


Fig. 4 MALDI-TOF peptide mapping of endoproteinase Lys-C peptides from a *Mus macedonicus* and *M. musculus domesticus* (C57BL/6) MUP. Anion exchange purified MUPs were digested with endoproteinase Lys-C (see Methods and Materials section for details). The mass of the resulting peptides was determined using MALDI-TOF mass spectrometry. Peptides were identified in the *M. m. domesticus* sample by comparison of the MALDI masses to a set of predicted masses (*L1–L11*) generated from a theoretical digest of a translated cDNA sequence (accession number AAI00587). Comparison of *M. macedonicus* peptide masses to those of *M. m. domesticus* allowed identification of both identical peptides and those where amino acid substitutions had occurred. Peptides thought to contain amino acid substitutions in the *M. macedonicus* sample are *labeled with a horizontal bar*. Additional ions, attributed to reaction of cysteine residues with β-mercaptoethanol, were observed 76 Da higher than certain peptides. Such peaks are labeled (+76 Da)

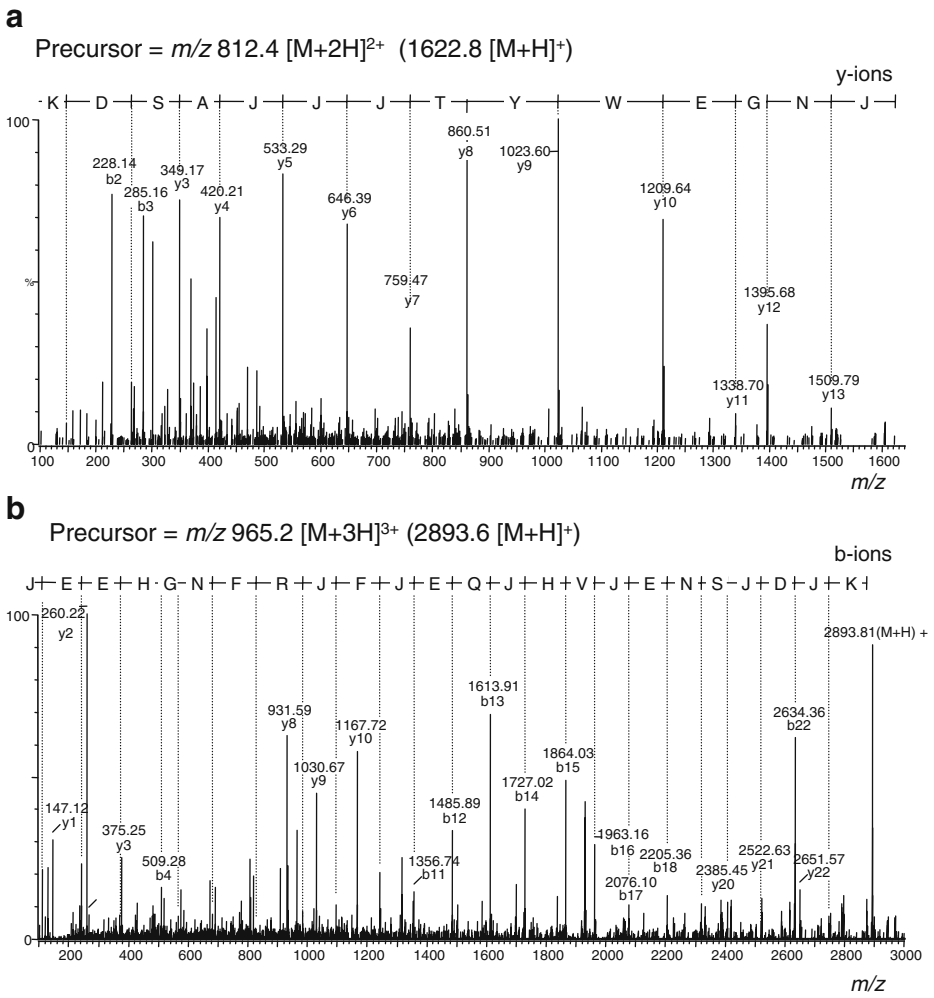


Fig. 5 Product ion mass spectra of (a) m/z 812.4 ($[M+2H]^{2+}$, 1,623.0 Da) and (b) m/z 965.2 ($[M+3H]^{3+}$, 2,893.9 Da) *Mus macedonicus* peptides. Tandem LC-MSMS experiments were performed on a Micromass Q-tof Micro mass spectrometer and a Dionex Ultimate RP-HPLC system (see Methods and Materials section for details). A Lys-C digestion of an anion exchange purified *M. macedonicus* MUP (18,742 Da) was introduced into the system, and ions of interest (m/z 812.4, 2+ and 965.2, 3+) were selected for tandem mass spectrometry using the data dependent switching feature within the MassLynx software. Fragmentation energy was similarly automatically selected using this feature. Product ion spectra were combined and de-convoluted using MaxEnt 3 software, with the charge state parameter determined from the isotope envelope of the precursor ion scan. De-convoluted spectra were interpreted with the MassSeq software included in the MassLynx package. Sequence information for the m/z 812.4 peptide was gained from a y-ion series, whereas that of the m/z 965.2 peptide was determined from a b-ion series. In this latter case, the identity of the 8 C-terminal residues was confirmed by a similar experiment on a trypsin digest (data not shown)

The three novel *M. macedonicus* peptides were sequenced by electrospray ionization tandem mass spectrometry. The first of the three unique *M. macedonicus* peptides (1,623.0 Th, $[M+H]^+$) was observed as a doubly charged ion at m/z 812.4 under electrospray ionization conditions. This fragmented readily to yield an abundant series of y ions (Fig. 5a). From these data, the amino acid sequence of the peptide was determined to be

*J*NGEWYT*JJJ*ASDK, where *J* is used to represent either leucine or isoleucine, which are isobaric and unable to be differentiated. This sequence is virtually identical to peptide L2 from *M. m. domesticus* AAI00587, ($_{15}$INGEWHTIILASDK$_{28}$), assuming correspondence of the leucine and isoleucine residues. The only difference between these peptides was the substitution of a tyrosine for a histidine residue in the *M. macedonicus* peptide. The mass difference of a His/Tyr substitution is +26 Da. The difference between the measured m/z values of the AAI00587 L2 peptide and the 1,623.0 *M. macedonicus* peptide is +25.8 Da, which confirms this as the sole change. In addition to characterizing this substitution, this experiment confirmed the identities of 14 additional amino acids in the *M. macedonicus* MUP, with the caveat of the residual ambiguity due to the isobaric pair of leucine and isoleucine.

The second unique *M. macedonicus* peptide (m/z 2,893.9 Th, $[M+H]^+$) was observed as a triply charged ion of m/z 965.2 in the precursor spectrum, implying an internal basic residue. During tandem mass spectrometry, this peptide did not generate a full series of y ions but fragmented to produce a complete set of b ions (Fig. 5b), from which the sequence of the peptide was determined as JEEHGNFRJFJEQJHVJENSJDJK. The sequence of the N-terminal eight residues of this peptide (JEEHGNFR) was confirmed by tandem mass spectrometry of a tryptic digest of the same preparation (data not shown). Although this sequence shared significant homology with peptide L4 from AAI00587, ($_{31}$IEDNGNFRLFLEQIHVLENSLVLK$_{55}$), it differed at three positions. The observed differences (from *M. m. domesticus* to *M. macedonicus*) were $D_{33}E$, $N_{34}H$, and $V_{53}D$. The arginine residue at position 40 was consistent with the triply charged precursor ion. Thus, an additional 24 residues of the *M. macedonicus* MUP were identified, and again, the interpretation was confirmed from the mass difference between the *M. m. domesticus* L4 peptide (52.4 Da) and the proposed *M. macedonicus* sequence (53 Da).

The final *M. macedonicus* peptide identified as a target for tandem mass spectrometry sequencing (2,066.2 Th, $[M+H]^+$) was also observed as a triply charged ion of m/z 689.4 under electrospray ionization conditions, consistent with an internal basic residue. Although this peptide did not generate a complete set of y or b ions (Fig. 6), the analysis yielded partial sequences from both y and b ion series that overlapped. The y ion series yielded the partial sequence of FHJ.....JSMVADK. The b ions gave a sequence of ...ECSEJSMV. Amalgamating these two sequences gave an overall amino acid sequence of FHJ...ECSEJSMVADK. When aligned to peptide L5 from AAI00587 ($_{56}$VHTVRDEECSELSMVADK$_{73}$), a high level of similarity was apparent. Two amino acid substitutions were observed in the *M. macedonicus* peptide; these were $V_{56}F$ and $T_{58}L$. The substitution of F/V at position 56 has been previously observed in other MUPs, and the V_{56} residue in the C57BL/6 sequence used in this study (AAI00587) is only observed in a minority of cases. This residue is located within the central, ligand-binding calyx and modulates the ability to bind specific ligands (Darwish Marie et al., 2001).

Four residues in the middle of the peptide remained unassigned. To establish the identity of these residues in L5, a Lys-C digest of the *M. macedonicus* protein was treated with an ethanolic solution of acetyl chloride to convert free carboxyl groups within the peptides to the corresponding methyl esters. Peptides, thus modified, increase in mass by 14 Da for each methyl group added, quantifying the number of acidic residues within the peptide. MALDI-TOF mass spectrometry of a reduced, carbamidomethylated and esterified Lys-C digestion of the *M. macedonicus* MUP (Fig. 7) yielded ions of masses consistent with the predicted number of carboxyl groups from previous assignments. For peptide L5, the mass difference between the carbamido peptide and the methylated carbamido peptide was 84 Da (6×14), which identified five acidic residues plus the C terminus. Previous tandem mass

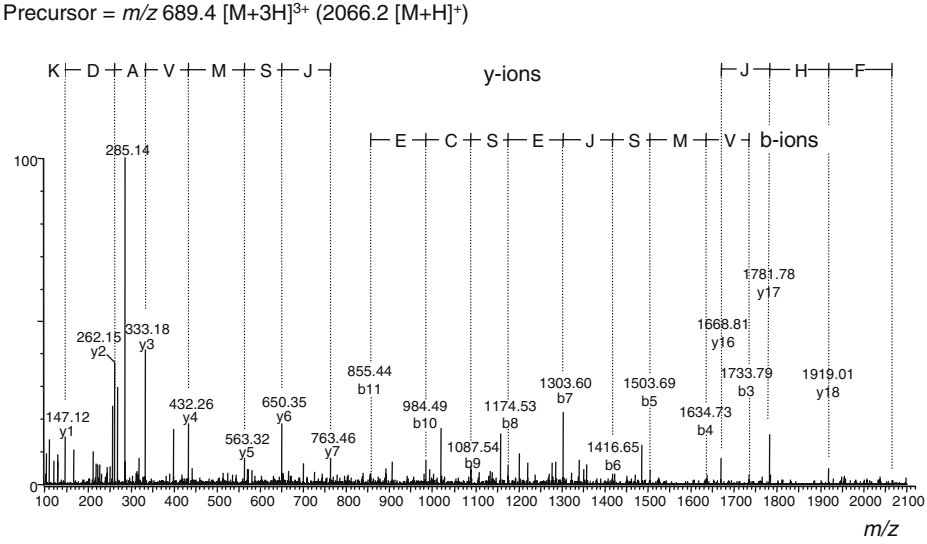


Fig. 6 Product ion mass spectrum of m/z 689.4 ($M+H^+$ 2,066.2 Da) *Mus macedonicus* peptide. A tandem LC-MSMS experiment was performed on a *M. macedonicus* Lys-C peptide of m/z 689.4, under identical conditions to those described in the Fig. 5 legend. Once again, this peptide was isolated from a *M. macedonicus* MUP Lys-C digest, using data dependent switching. The MaxEnt 3 de-convoluted product ion spectrum was interpreted using PepSeq software. In this instance, a composite partial sequence was inferred from both b and y ions

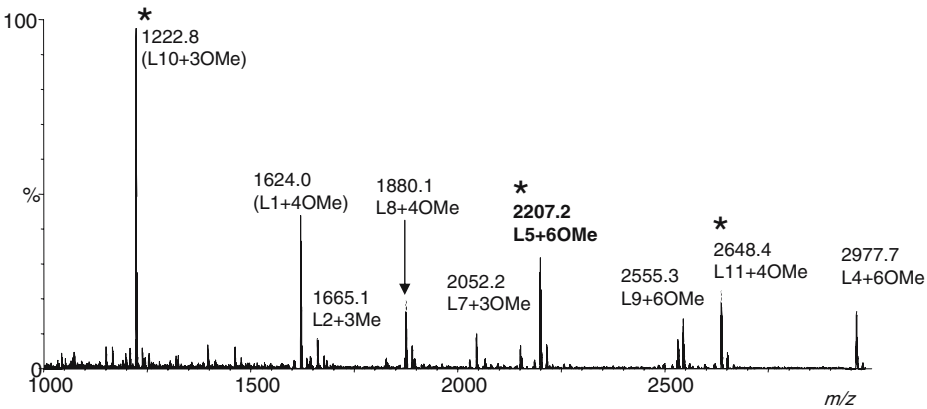


Fig. 7 MALDI-TOF mass spectrum of a methyl esterified Lys-C digest of a carbamidomethylated *M. macedonicus* MUP. Quantification of acidic residues within Lys-C peptides from a *M. macedonicus* MUP was achieved by methyl esterification. Before digestion, an anion exchange purified *M. macedonicus* MUP preparation was reduced, carbamidomethylated, and digested with Lys-C (see Methods and Materials). Peptides were then extracted from the solution using reverse phase chromatography media (Zip Tips), washed with 0.1% (v/v) trifluoroacetic acid (TFA), and finally desorbed and collected in a solution of 50% (v/v) acetonitrile/0.1% (v/v) TFA. Solvents and TFA were then removed from this preparation using a vacuum centrifuge. A 1-ml aliquot of methanol was incubated at −20°C for 15 min before addition of acetyl chloride (150 µl). A 20-µl aliquot of this solution was then added to the dried Lys-C peptides and incubated at RT for 45 min. Esterified Lys-C peptides unique to *M. macedonicus* are marked with an asterisk (*)

spectrometric analysis of this peptide identified the positions of three of these residues, leading to the conclusion that two of the four unidentified residues were also acidic, namely, aspartate (D) or glutamate (E).

The identification of two acidic residues in the unassigned portion of L5 allows inferences to be drawn as to the identities of the remaining two unassigned residues. On ESI-MS, peptide L5 is triply charged, implying that it contains an internal basic residue. Although the histidine residue at position 57 is weakly basic, a more likely explanation of this charge state is that the arginine residue at position 60 in the *M. m. domesticus* sequence is conserved, as observed in all the urinary MUPs previously characterized in our laboratories (Robertson et al., 1996). The L5 peptide from *M. m. domesticus* is 42 Da heavier than the equivalent peptide from *M. macedonicus* (allowing for residues assigned by tandem mass spectrometry). The assumption that R_{60} is conserved and the knowledge that two of the residues are acidic allows for only two possibilities for the final assignment of the sequence. In the first scenario, the two acidic residues at positions 61 and 62 are identical to the *M. m. domesticus* sequence, in which case a V/G substitution at position 59 in the *M. macedonicus* sequence would account for the mass difference. The second scenario allows for the possibility of a D/E substitution at position 61 and a V/A substitution at position 59 to give a net loss of 42 Da in the *M. macedonicus* peptide. On the basis of this analysis, the remaining four amino acid residues in L5 are assigned either as [GR(DE)] or [AR(DD)]. The order of these amino acids is not apparent from the above analysis and has been assigned on the basis of the overall level of conservation observed between the two MUPs (Fig. 8).

Endopeptidase Lys-C digestion of AAI00587 creates two tripeptides (L3 and L6), each of which has a predicted mass of less than 400 Da. Such peptides are difficult to analyze with MALDI-TOF mass spectrometry because of background signal from ions in the chemical matrix. We have, however, observed in previous experiments that Lys-C has a

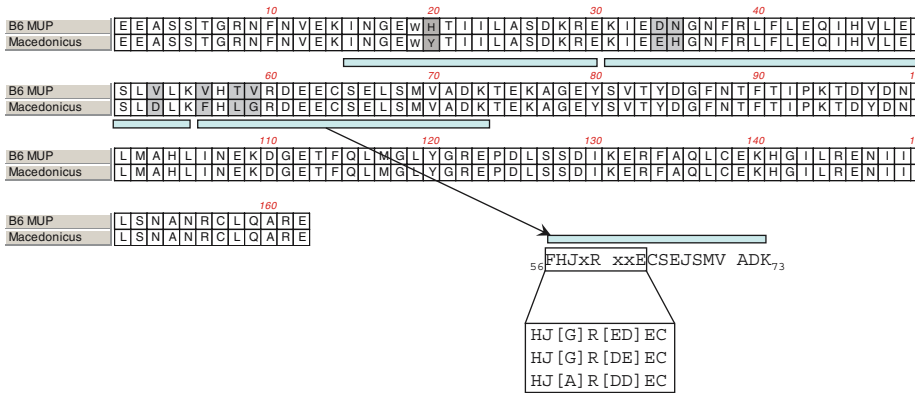


Fig. 8 Comparison of *M. macedonicus* and *M. m. domesticus* MUP sequences. Inferred data about the sequence of the *M. macedonicus* MUP from MALDI-TOF, LC-MS, and LC-MSMS analysis were amalgamated into a single composite sequence and compared to the known sequence of a *M. m. domesticus* MUP (accession number AAI00587, labeled "B6 MUP"). Positions of amino acid substitutions in the two sequences are shaded. Full sequence identification was not achieved between residues 59 and 62, but was narrowed down to three permutations following methyl esterification (see Fig. 7). An expanded view of this area is inset and the three possible permutations included. Amino acid identities within the three permutation sequences inferred from methyl esterification are shown in brackets. The arginine residue at position 60 is assigned provisionally, pending further proof, on the basis of overall peptide charge state (3+) and conservation between known MUP sequences

tendency to omit cleavages in the peptide chain in areas where lysine residues are in close proximity. To assess whether the two tripeptides in *M. macedonicus* possessed the same sequence as those in AAI00587, an additional LC-MS experiment was run on the Lys-C peptides. In this instance, the m/z values for L2+L3 (1,018.54, $[M+2H]^{2+}$) and L5+L6 (820.70, $[M+3H]^{3+}$ with an oxidized methionine residue) were readily observed in extracted ion chromatograms (data not shown). The two tripeptides were therefore assigned to the *M. macedonicus* sequence as identical to the *M. m. domesticus* sequence.

The assembled *M. macedonicus* sequence was aligned to that of AAI00587 (Fig. 8). Of the 162 amino acids in the *M. macedonicus* sequence, 152 are identical to AAI00587, seven residues are known to differ, and the identity of a further three is unclear, although their chemical class has been determined. This has allowed the modeling of the variant residues into the three dimensional structure using the Modeler facility within the Discovery software package and crystallographic coordinates (Kuser et al., 2001) (Fig. 9), lodged in the National Centre for Biotechnology Information (NCBI) structural database (http://www.ncbi.nlm.nih.gov).

In constructing the model *M. macedonicus* MUP structure, we had to consider two types of uncertainty in the sequences derived from proteomics experiments. First, the amino acids between residues 58 and 61 could not be unambiguously determined by tandem mass spectrometry, and the three possibilities that existed (GRED, GRDE, ARDD) were modeled. Secondly, it is not possible to discriminate between the isobaric residues leucine and isoleucine. To assess the importance of these residues, we constructed two artificial sequences based on AAI00587, in which all Ile/Leu residues were converted to either Leu or Ile. Models built on the parent 1I04 structure were compared, and the RMSD of the alpha carbon atoms of the two structures was 0.14 Å. The overall main chain trajectories were virtually identical. We conclude that Leu/Ile substitution would not have a major effect on the model structures. Accordingly, we built *M. macedonicus* models assuming that the identity of Leu/Ile residues were the same as for AAI00587. For the six variant structures based on the unresolved pentapeptide (see above), all yielded high quality models with alpha carbon RMSD of less than 0.25 Å and similar main chain trajectories (Fig. 9). All proteins passed the Protein Health checks built into the Modeler package. Other than

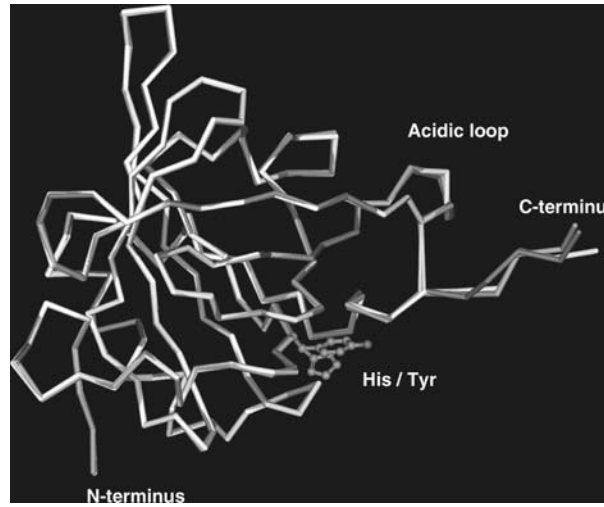


Fig. 9 Homology model building of *M. macedonicus* sequences. The three variants of the possible *M. macedonicus* sequence, varying at the acidic loop (GRED, GRDE, and ARDD) were used to construct homology models based on the MUP structure 1I04.PDB. All three models were aligned structurally to the 1I04 structure, with an RMSD in each instance of less than 0.16 Å. The uncertain acidic loop is highlighted, as is the His/Tyr substitution that is in close proximity to the GxW lipocalin motif. Because the structures are so similar, no labels have been included to identify them individually

the ambiguous acidic loop region of the *M. macedonicus* protein sequences, there are a number of amino acid substitutions relative to MUPs from *M. m. domesticus*. These changes ($H_{20}Y$, $D_{34}E$, $N_{35}H$, $V_{53}D$, $T_{58}L$, $V_{59}G$) do not influence the ligand binding cavity and are solvent exposed.

To assess the level of similarity between the *M. macedonicus* sequence and those from *M. m. domesticus*, we compared it to all the known *M. m domesticus* sequences in the NCBI database. The amino acid sequence of the *M. macedonicus* MUP shares most similarity with a group of MUPs for which only incomplete sequences are available. These, in turn, are most similar to the male-specific 18,893 Da MUP discovered in C57BL/6 mouse, but also widely present in wild caught *M. m. domesticus*. This protein has a high capacity for binding of the male specific ligand 2-*sec*-butyl-4,5-dihydrothiazole (thiazole) that accords with a function in male-specific signaling (Armstrong et al., 2005).

The simple MUP pattern in *M. macedonicus*, together with the lack of individual variability, suggests that MUPs do not have sufficient polymorphism to provide an individual ownership signal in scent marks in this species, in contrast to *M. m. domesticus*. This may reflect the difference in the population ecology of this aboriginal grassland species compared to the commensal house mice and possibly reflects the more ancestral form. While the field ecology of *M. macedonicus* is not well known, in line with the other aboriginal species of *Mus* in Europe and the Middle East, *M. macedonicus* in Turkey can be presumed to have large territories, and individuals are highly agonistic to each other (reviewed by Frynta et al., 2005). Both male and female *M. macedonicus* are much more aggressive than *M. m. domesticus* (Frynta and Čiháková, 1996), and individuals may be largely nonoverlapping. By contrast, in commensal *M. m. domesticus* populations, multiple males and females live within territorial social groups, and there may be extensive spatial overlap between neighbors when borders are not easily defended (Hurst, 1987; Barnard et al., 1991). This results in an unusually high level of aggregation and contact between individuals in this species. As male *M. m. domesticus* advertise their territorial dominance through scent marks (Hurst and Beynon, 2004), it is possible that this may have been a strong driver for the evolution of individual-specific MUP patterns in *M. m. domesticus* but not in the aboriginal species. Given their dispersed distribution, there is likely to be much less requirement for advertising individual scent ownership in *M. macedonicus* and other widely dispersed rodent species. This suggests the intriguing hypothesis that the extreme polymorphism and individual variability in MUP patterns seen in *M. m. domesticus* is a species-specific adaptation for signaling individuality in a complex social system where individuals vary in social status. *M. macedonicus* is one of a number of aboriginal mouse species that have evolved independently of *M. m. domesticus*, inhabit different environments, and display different forms of behavior. Information about the MUPs and MUP genes in these animals remains scarce (Sampsell and Held, 1985), yet they provide a valuable opportunity to investigate the function of both MUPs and their ligands and the evolution of a family of proteins used for scent signaling. To this end, a wider survey of MUPs from aboriginal mice of both sexes, which includes both primary structure and ligand status, has the possibility to produce greater insights into the wider context of MUP polymorphism and function.

Acknowledgments We are grateful for the assistance of Coskun Tez and support from the Genetics Society toward the fieldwork. The work was carried out under research grants from the Biotechnology and Biological Sciences Research Council, UK to J.L.H. and R.J.B.
In memory of J.E. Robertson, 1928–2003.

References

ADAMS, P. and SAWYER, L. 1990. Structure of rodent urinary proteins. *Biochem. Soc. Trans.* 18:936–937.

ARMSTRONG, S. D., ROBERTSON, D. H. L., CHEETHAM, S. A., HURST J. L., and BEYNON, R. J. 2005. Structural and functional differences in isoforms of mouse major urinary proteins: a male-specific protein that preferentially binds a male pheromone. *Biochem. J.* 391:343–350.

BACCHINI, A., GAETANI, E., and CAVAGGIONI, A. 1992. Pheromone binding proteins in the house mouse *Mus musculus*. *Experientia* 48:419–421.

BARNARD, C. J., HURST, J. L., and ALDHOUS, P. 1991. Of mice and kin: the functional significance of kin bias in social behaviour. *Biol. Rev.* 66:379–430.

BERRY, R. J. 1981. Population dynamics of the house mouse. *Symp. Zool. Soc. Lond.* 47:395–425.

BEYNON, R. J. and HURST, J. L. 2004. Urinary proteins and the modulation of chemical scents in mice and rats. *Peptides* 25:1553–1563.

BEYNON, R. J., VEGGERBY, C., PAYNE, C. E., ROBERTSON, D. H., GASKELL, S. J., HUMPHRIES, R. E., and HURST, J. L. 2002. Polymorphism in major urinary proteins: molecular heterogeneity in a wild mouse population. *J. Chem. Ecol.* 7:1429–1446.

BOURSOT, P., AUFFRAY, J.-C., BRITTON-DAVIDIAN, J., and BONHOMME, F. 1993. The evolution of house mice. *Annu. Rev. Ecol. Syst.* 24:119–152.

BRONSON, F. H. 1979. The reproductive ecology of the house mouse. *Q. Rev. Biol.* 54:265–299.

CUCCHI, T., VIGNE, J.-D., and AUFFRAY, J.-C. 2005. First occurrence of the house mouse (*Mus musculus domesticus* Schwarz & Schwarz), 1943 in the Western Mediterranean: a zooarchelogical revision of subfossil occurrences. *Biol. J. Linn. Soc.* 84:429–445.

DARWISH MARIE, A., VEGGERBY, C., ROBERTSON, D. H., GASKELL, S. J., HUBBARD, S. J., MARTINSEN, L., HURST, J. L., and BEYNON, R. J. 2001. Effect of polymorphisms on ligand binding by mouse major urinary proteins. *Protein Sci.* 10:411–417.

EVERSHED, R. P., ROBERTSON, D. H., BEYNON, R. J., and GREEN, B. N. 1993. Application of electrospray ionization mass spectrometry with maximum-entropy analysis to allelic 'fingerprinting' of major urinary proteins. *Rapid Commun. Mass Spectrom.* 10:882–886.

FINLAYSON, J. S., ASOFSKY, R., POTTER, M., and RUNNER, C. C. 1965. Major urinary protein complex of mice: origin. *Science* 149:981–982.

FLOWER, D. R., NORTH, A. C. T., and ATTWOOD, T. K. 1993. Structure and sequence relationships in the lipocalins and related proteins. *Protein Sci.* 2:753–761.

FRYNTA, D. and ČIHÁKOVÁ, J. 1996. Neutral cage interactions in *Mus macedonicus* (Rodentia: Muridae): an aggressive mouse? *Acta Soc. Zool. Bohem.* 60:97–102.

FRYNTA, D., SLÁBOVÁ, M., VÁCHOVÁ, H., VOLFOVÁ, R., and MUNCLINGER, P. 2005. Aggression and commensalisms in house mouse: a comparative study across Europe and the Near East. *Aggress. Behav.* 31:283–293.

HURST, J. L. 1987. The functions of urine marking in a free living population of house mice. *Anim. Behav.* 35:1433–1422.

HURST, J. L. and BEYNON, R. J. 2004. Scent wars: the chemobiology of competitive signalling in mice. *Bioessays* 26:1288–1298.

HURST, J. L., ROBERTSON, D. H. L., TOLLADAY, U., and BEYNON, R. J. 1998. Lipocalins in mouse urine provide a slow release mechanism for olfactory signals. *Anim. Behav.* 55:1289–1297.

Hurst, J. L., Payne, C. E., Nevison, C. M., Marie, A. D., HUMPHRIES, R. E., ROBERTSON, D. H. L., CAVAGGIONI, A., and BEYNON, R. J. 2001. Individual recognition in mice mediated by Major Urinary Proteins. *Nature* 414:631–634.

HURST, J. L., THOM, M. D., NEVISON, C. M., HUMPHRIES, R. E., and BEYNON, R. J. 2005. MHC odours are not required or sufficient for recognition of individual scent owners. *Proc. R. Soc. Lond. B Biol. Sci.* 272:715–724.

IVANTCHEVA, A. and CASSAING, J. 1999. Male–female interactions and socio-sexual isolation of *Mus spicilegus* toward two other species of east-European mice, *M. macedonicus* and *M. m. musculus. C. R. Acad. Sci. III* 322:597–605.

KRIZOVA, H., ZIDEK, L., STONE, M. J., NOVOTNY, M. V., and SKLENAR, V. 2004. Temperature-dependent spectral density analysis applied to monitoring backbone dynamics of major urinary protein-I complexed with the pheromone 2- *sec*-butyl-4,5-dihydrothiazole. *J. Biomol. NMR* 4:369–384.

KUSER, P. R., FRANZONI, L., FERRARI, E., SPISNI, A., and POLIKARPOV, I. 2001. The X-ray structure of a recombinant major urinary protein at 1.75 A resolution. A comparative study of X-ray and NMR-derived structures. *Acta. Crystallogr. D. Biol. Crystallogr.* 57:1863–1869.

MALONE, N., PAYNE, C. E., BEYNON, R. J., and HURST, J. L. 2001. Social status, odour communication and mate choice in wild house mice, pp. 217–224, *in* A. Marchlewska-Koj, D. Muller-Schwarze, and J. Lepri (eds.). Chemical Signals in Vertebrates. Plenum Press, New York.

NOVOTNY, M. V., MA, W., WIESLER, D., and ZIDEK, L. 1999. Positive identification of the puberty-accelerating pheromone of the house mouse: the volatile ligands associating with the major urinary protein. *Proc. R. Soc. Lond. B Biol. Sci.* 266:2017–2022.

PATRIS, B. and BAUDOIN, C. 1998. Female sexual preferences differ in *Mus spicilegus* and *Mus musculus domesticus*: the role of familiarization and sexual experience. *Anim. Behav.* 56:1465–1470.

PES, D., ROBERTSON, D. H. L., HURST, J. L., GASKELL, S., and BEYNON, R. J. 1999. How many major urinary proteins are produced by the house mouse *Mus domesticus*?, pp. 149–162, *in* R. E. Johnston, D. Müller-Schwarze, and P. W. Sorensen (eds.). Advances in Chemical Signals in Vertebrates 9. Plenum Press, New York.

ROBERTSON, D. H. L., BEYNON, R. J., and EVERSHED, R. P. 1993. Extraction, characterisation and binding analysis of two pheromonally active ligands associated with major urinary protein of house mouse. *J. Chem. Ecol.* 19:1405–1416.

ROBERTSON, D. H. L., EVERSHED, R. P., COX, K., GASKELL, S., and BEYNON, R. J. 1996. Characterisation of molecular heterogeneity in the major urinary proteins of the house mouse *Mus musculus*. *Biochem. J.* 315:461–466.

ROBERTSON, D. H., HURST, J. L., BOLGAR, M. S., GASKELL, S. J., and BEYNON, R. J. 1997. Molecular heterogeneity of urinary proteins in wild house mouse populations. *Rapid Commun. Mass Spectrom.* 11:786–790.

SAMPSELL, B. M. and HELD, W. A. 1985. Variation in the major urinary protein multigene family in wild-derived mice. *Genetics* 109:549–568.

SHEVCHENKO, A., SUNYAEV, S., LISKA, A., BORK, P., and SHEVCHENKO, A. 2003. Nanoelectrospray tandem mass spectrometry and sequence similarity searching for identification of proteins from organisms with unknown genomes. *Methods Mol. Biol.* 211:221–234.

SILVER, L. M. 1995. Mouse Genetics, Concepts and Applications. Oxford University Press, Oxford, UK.

SUZUKI, H., SHIMADA, T., TERASHIMA, M., TSUCHIYA, K., and APLIN, K. 2004. Temporal, spatial, and ecological modes of evolution of Eurasian *Mus* based on mitochondrial and nuclear gene sequences. *Mol. Phylogenet. Evol.* 33:626–646.

TIMM, D. E., BAKER, L. J., MUELLER, H., ZIDEK, L., and NOVOTNY, M. V. 2001. Structural basis of pheromone binding to mouse major urinary protein (MUP-I). *Protein Sci.* 10:997–1004.

VEGGERBY, C., PAYNE, C. E., ROBERTSON, D. H. L., GASKELL, S. J., HUMPHRIES, R. E., HURST, J. L., and BEYNON, R. J. 2002. Polymorphism in major urinary proteins: heterogeneity in a wild mouse population. *J. Chem. Ecol.* 28:1425–1440.