Biochemical Society TRANSACTIONS

Comparative study of the molecular variation between 'central' and 'peripheral' MUPs and significance for behavioural signalling

Marie M. Phelan*[1], Lynn McLean†, Jane L. Hurst‡, Robert J. Beynon† and Lu-Yun Lian*

*NMR Centre for Structural Biology, Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, U.K.

†Protein Function Group, Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, U.K.

‡Mammalian Behaviour and Evolution Group, Institute of Integrative Biology, University of Liverpool, Leahurst Campus, Neston CH64 7TE, U.K.

Abstract

MUPs (major urinary proteins) play an important role in chemical signalling in rodents and possibly other animals. In the house mouse (*Mus musculus domesticus*) MUPs in urine and other bodily fluids trigger a range of behavioural responses that are only partially understood. There are at least 21 *Mup* genes in the C57BL/6 mouse genome, all located on chromosome 4, encoding sequences of high similarity. Further analysis separates the MUPs into two groups, the 'central' near-identical MUPs with over 97 % sequence identity and the 'peripheral' MUPs with a greater degree of heterogeneity and approximately 20–30 % non-conserved amino acids. This review focuses on differences between the two MUP sub-groups and categorizes these changes in terms of molecular structure and pheromone binding. As small differences in amino acid sequence can result in marked changes in behavioural response to the signal, we explore the potential of single amino acid changes to affect chemical signalling and protein stabilization. Using analysis of existing molecular structures available in the PDB we compare the chemical and physical properties of the ligand cavities between the MUPs. Furthermore, we identify differences on the solvent exposed surfaces of the proteins, which are characteristic of protein–protein interaction sites. Correlations can be seen between molecular heterogeneity and the specialized roles attributed to some MUPs.

Introduction

A variety of MUPs (major urinary proteins), are present in mouse urine at mM and higher concentrations [1,2]. These proteins act as chemical signals in the urine of the mouse, enabling the mice to identify individuals and relatedness from distinct individual patterns [3–7]. One male-specific MUP, darcin, stimulates females to spend time near male scent; it also stimulates strong and rapid associative learning such that females learn the same attraction towards the volatile airborne scent signature of the male [8] and to the remembered location of the pheromone [9]. MUPs are also involved in regulating competitive behaviour between males as a mixture of MUPs applied to the body of a castrated male stimulates aggression from other males [10]. Signalling involves not only the proteins themselves but bound specific volatile ligands [11] and the variation of MUPs influences the pattern of urinary volatiles held and slowly released from drying urine marks [12]. MUP-bound volatile pheromones include the male-specific urinary volatiles 2-*sec*-butyl-4,5-dihydrothiazole, 3,4-dehydro-*exo*-brevicomin and 6-hydroxy-6-methyl-3-heptanone [13–15]. Furthermore, mice can discern differing patterns of MUPs in urine [3,16], indicating that they can detect the differences between MUPs and/or the ligands that they bind. Not all MUPs are excreted in urine; MUP4 is expressed in glands near the nasal cavity and is present in both nasal mucus and the vomeronasal organ where it is thought to play a role in transport of ligands to neuronal receptors [17]. Although mice can easily discriminate between different MUP patterns in urine, the differences between these proteins at a molecular level are minor, with the MUP proteins exhibiting a high degree of sequence similarity [18] and likely to adopt the same three-dimensional fold. This review endeavours to explore the subtle differences between the MUPs at the molecular level and identify variations between the extremely similar 'central', and the more diverse, specialized 'peripheral' MUPs.

Key words: amino acid, darcin, major urinary protein (MUP), pheromone.

Abbreviations: MUP, major urinary protein.

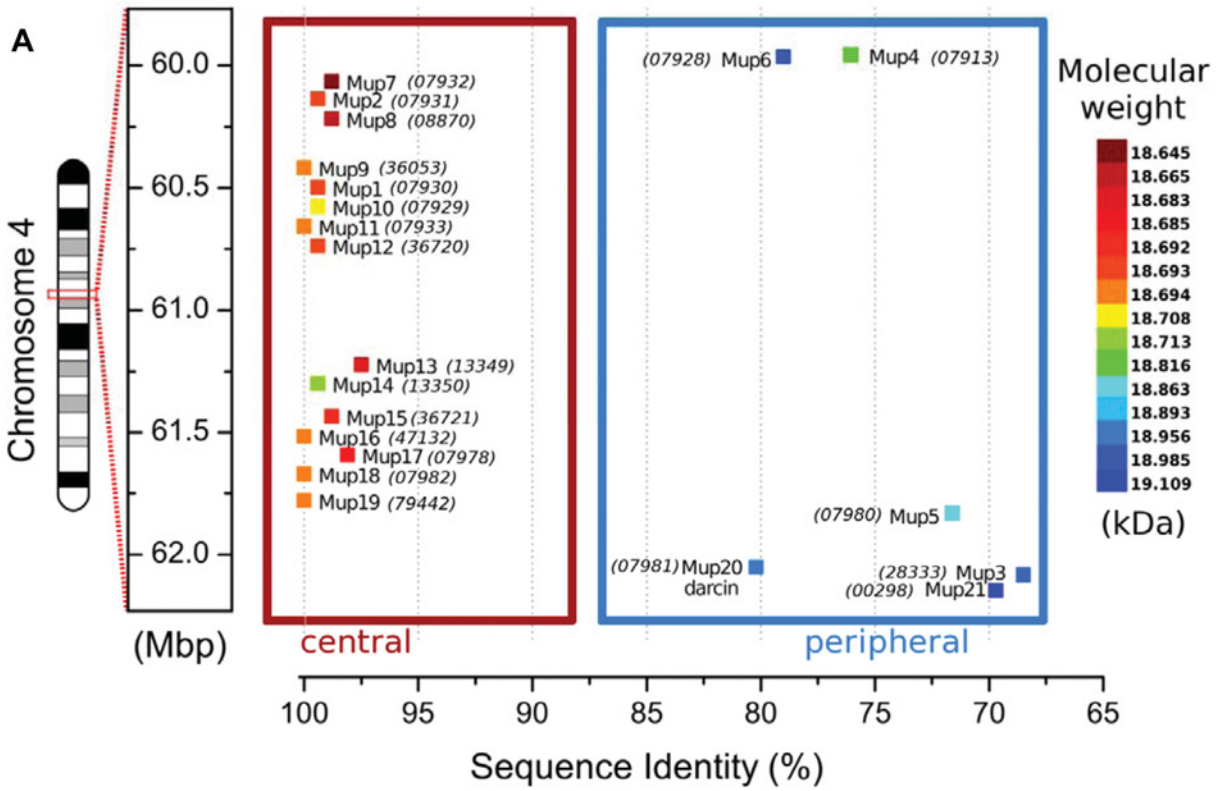
[1]**To whom correspondence should be addressed (email mphelan@liv.ac.uk).**

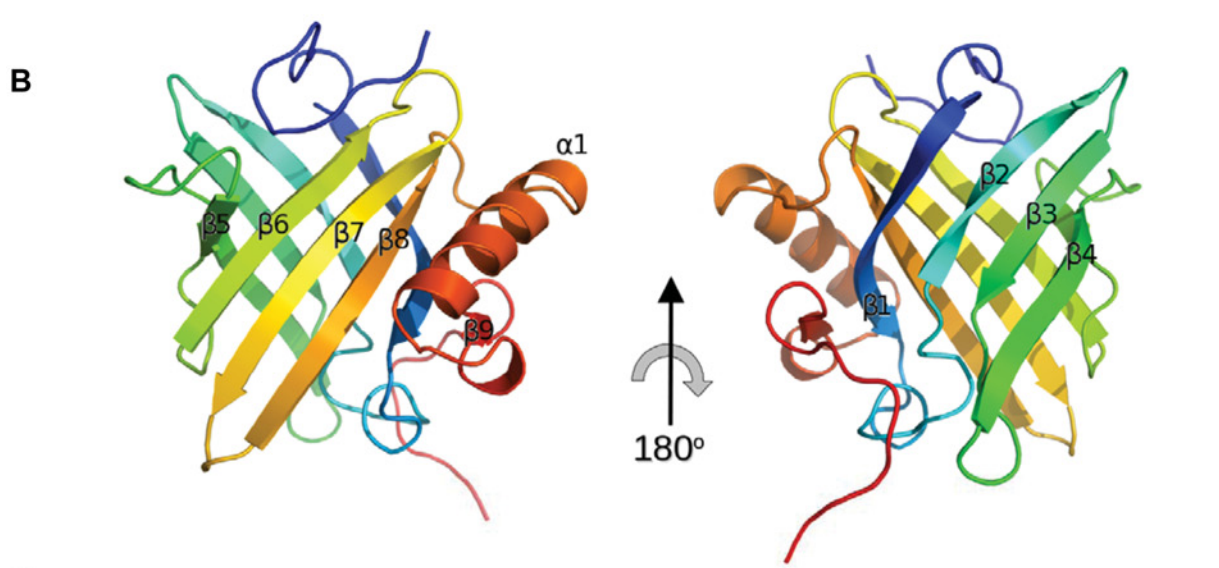
MUP homology

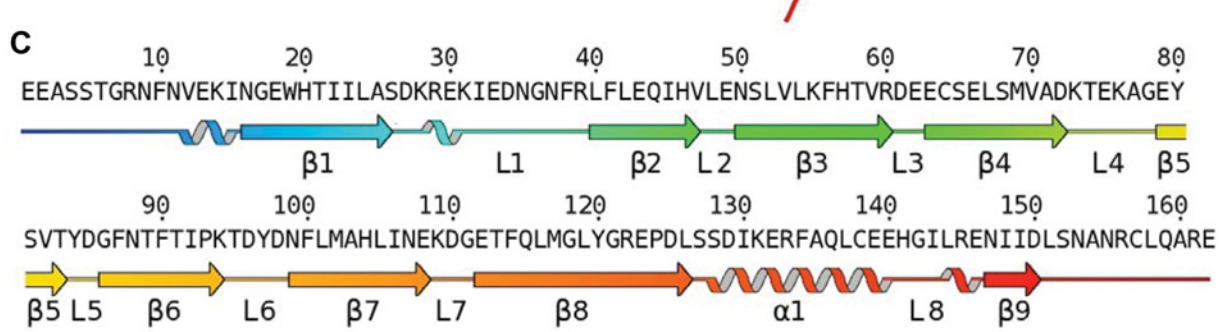
There are 21 *Mup* genes annotated in the mouse genome database, MGI (http://www.informatics.jax.org) [19,20], although several gaps in the genome within this region may harbour additional genes [18,21]. Most *Mup* genes encode proteins of 180 amino acids in length. The N-terminal signal peptide for each MUP is usually 18 amino acids in length and not present in the mature proteins identified in the bodily fluids of the mouse (the majority present in urine) with the mature protein sequence always 162 amino acids in length. Of the mature amino acid sequences there are 16

Figure 1 | **The MUP family of *Mus musculus domesticus***

(**A**) Organization of known mouse *Mup* genes on chromosome 4. The 'central' and 'peripheral' MUPs are coloured according to their mature protein molecular mass; distinction between central MUPs (>97 % mature protein sequence identity) and







peripheral MUPs (<82 % mature protein sequence identity) can clearly be seen. An enlarged view of chromosome 4 B3Q region with base pairs numbered. Chromosome 4 schematic diagram courtesy of Dr David Alder, University of Washington, Seattle. The percentage sequence identity is relative to the most common mature amino acid sequence that is shared by genes *Mup9*, *Mup11*, *Mup16*, *Mup18* and *Mup19*. Numbers in parentheses correspond to the MGI protein identification number. For brevity, the 18-character ID, i.e. OTTMUSP00000007981, is abbreviated to the last five digits only (07981), therefore all numbers are prefixed with OTTMUSP000000. (**B**) Cartoon representation of the typical MUP structure. The cartoon depicts the structure of the most generic protein sequence comprising MUP9, MUP11, MUP16, MUP18 and MUP19 (M.M. Phelan, L. McLean, S.D. Armstrong, J.L. Hurst, R.J. Beynon and L.-Y. Lian, unpublished work) (PDB code 2LB6). (**C**) The annotated amino acid sequence with colour-coding from the N- to the C-terminus from blue to red, to match (**B**).

unique sequences identified with five genes encoding the same mature protein sequence shared by *Mup9*, *Mup11*, *Mup16*, *Mup18* and *Mup19* and a further protein sequence encoded by both *Mup2* and *Mup12*. This high degree of similarity is also made evident by the very limited one, two or three amino acid differences encoded by the genes clustered at the centre of *Mup* region of the chromosome (Figure 1A). The considerable similarity between MUPs encoded within the central region is due to multiple gene duplication events during recent rapid expansion of this region in the house mouse [21]. A further six *Mup* genes are also annotated that flank the central *Mup* region of chromosome 4. For these peripheral *Mup* genes the mature amino acid sequences are more varied, with sequences sharing between 67 and 81 % identity. Behavioural and physiological studies on certain proteins in this group have also identified more specialized roles or locations for several of these more divergent MUPs and thus the nature of these differences at a chemical level are both informative and intriguing ([18,22–24], and M.M. Phelan, L. McLean, J.L. Hurst, R.J. Beynon and L.-Y. Lian, unpublished work). As nomenclature for the MUP protein family has altered between editions of the MGI, and indeed there are often sequence variations available for each gene, we prefer to also annotate the proteins with the MGI protein sequence ID (Figure 1A). For mature protein OTTMUS identification numbers please refer to Beynon et al. [24a] in this issue of *Biochemical Society Transactions*.

MUP structure

There are 25 atomic resolution MUP structures available in the PDB ([24,25–34], and M.M. Phelan, L. McLean, S.D. Armstrong, J.L. Hurst, R.J. Beynon and L.-Y. Lian, unpublished work), comprising 22 X-ray crystal structures and three NMR structures. Despite the large number of structures solved, variation of the primary sequences is limited to four protein sequences corresponding to two central MUPs, MUP9 (identical in sequence with MUP11 MUP16, MUP18 and MUP19) and MUP10, and two peripheral MUPs, darcin (M.M. Phelan, L. McLean, S.D. Armstrong, J.L. Hurst, R.J. Beynon and L.-Y. Lian, unpublished work) and MUP4 [24].

All structures are extremely similar, such that unweighted mean structural alignments of the Cα, N and C backbone atoms of the proteins result in an RMSD of 0.59 Å (1 Å=0.1 nm) for structured regions (i.e. discounting the first and last ten amino acids from the N- and C-termini) calculated using the program uwmn (M.J. Hartshorn and L.S.D. Caves, University of York). The protein fold is that of a typical lipocalin, consisting of an eight-strand β-barrel and a 3.5 turn (approximately 14 amino acid) α-helix with an additional β-strand at the C-terminus (Figure 2). All structures solved to date (both 'central' and 'peripheral') adhere to this fold, with the centre of the β-barrel lined with predominantly hydrophobic residues, forming the binding cavity for the volatile small molecules present in mouse urine [35].

Cavity differences between MUPs

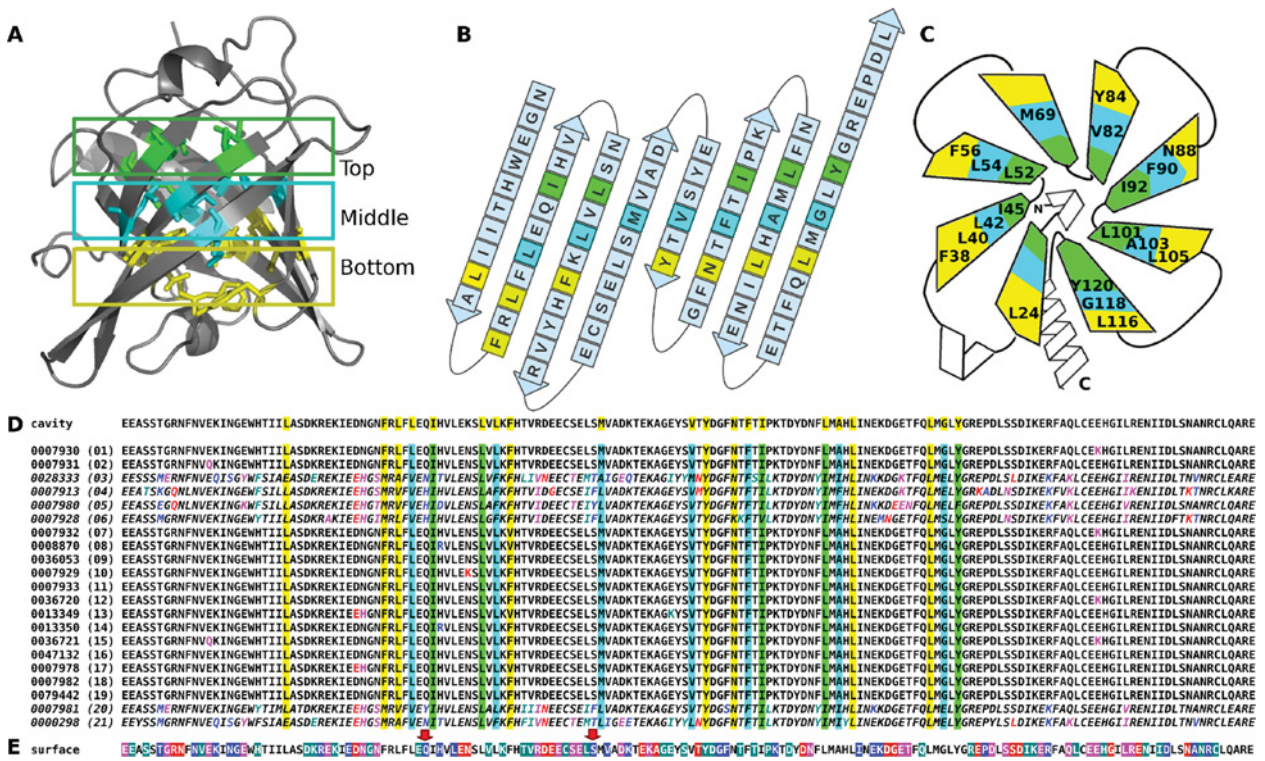
The cavity of the MUPs is well defined by a multitude of structures with ligand partners, some of which may be artefacts of crystallization technique (Table 1). By analysing the eight cyclic ligands and 11 short alcohol or ketone ligands using the program ligplot [36] and PISA [37], a consensus cavity binding motif has been determined (Figures 2A–2D). Furthermore, these cavity residues can be subdivided into their specific location in the β-barrel. If we consider the ends of the barrel as the 'top' and 'bottom' in terms of the location of the N- and C-termini respectively (Figure 2A), there are five residues located at the 'top' and eight residues at the 'bottom' of the cavity; the remaining seven residues form a band around the centre of the cavity where the small molecule volatiles are located (Figures 2B and 2C). The eight residues flanking the bottom of the cavity give rise to a greater distance between the protein backbone on opposing sides of the barrel at the 'bottom' when compared with the 'top'.

A comparison of the cavity residues in the central MUPs provides evidence of the generic nature of the cavity with 19 of the 20 residues conserved between all central MUPs. The only amino acid to vary is located at the bottom of the barrel at residue number 56 that varies between phenylalanine and valine (Figure 2D).

When comparing the cavity residues of peripheral MUPs, both with central and with other peripheral MUPs, a greater degree of variation in amino acids can be observed. Changes can be identified in amino acid properties, e.g. a hydrophobic residue such as leucine at position 29 in

Figure 2 | **Comparison of residues lining the hydrophobic cavity in peripheral and central MUPs**

The top and bottom of the cavity are defined as the ends of the β-barrel nearest the N- and C-terminus respectively. The cavity residues are coloured according to position in three-dimension: top (green) middle (cyan) and bottom (yellow). (**A**) Colour-coded residues of the cavity mapped on to a cartoon representation of central MUP, MUP11. (**B**) Colour-coded schematic representation of β-strand residues: amino acids shown represent those found in the central MUP, MUP11. (**C**) Colour-coded schematic representation of a top-down view into the cavity. (**D**) Alignment of the mature MUP sequences (numbered according to MGI OTTMUSP000 ID and MUP number in parentheses) with peripheral MUPs in italics. The top of the cavity is lined with five residues, i.e. 45, 52, 92, 101 and 120; the centre of the cavity is defined by seven residues, i.e. 42, 54, 69, 82, 90, 103, 118; the bottom of the cavity is lined with eight residues, i.e. 24, 38, 40, 56, 84, 88, 116 and 105. (**E**) Colour-coded schematic representation of surface-exposed residues mapped on to MUP11. Amino acid side-chain exposure is indicated by highlight colour: 100–80 % side chain exposed (red), 79–60 % side chain exposed (magenta), 59–40 % side chain exposed (blue), 39–20 % side chain exposed (blue–green). The percentage surface exposure, determined by mean side-chain exposure of 25 MUP structures, was calculated using the program NACCESS. Largest surface differences in darcin (residues 44 and 68) are indicated by red arrows.



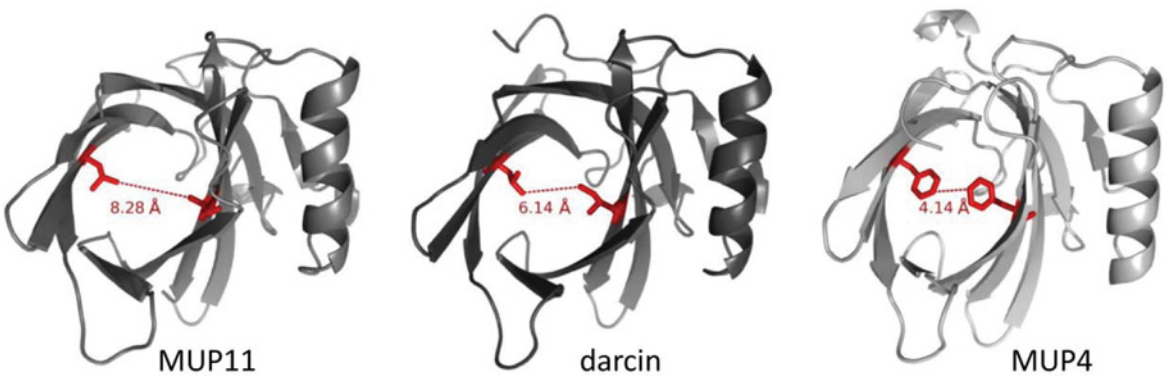
central MUPs compared with an acidic residue glutamate in MUP21. Furthermore, changes in amino acid size such as alanine for isoleucine or phenylalanine at position 103 demonstrate the narrowing of the centre of the cavity in the peripheral MUPS which in turn indicates a more selective cavity (Figure 3). Indeed with the variation of cavity residues between peripheral MUPs one can envisage differences in both selectivity and affinity for small molecule volatiles binding in the cavity. This selectivity is exemplified by the ability of the central MUP, MUP11 to bind the fluorophore NPN (*N*-phenyl-1-naphthylamine), whereas neither darcin nor MUP4 can bind this relatively bulky molecule (M.M. Phelan, L. McLean, S.D. Armstrong, J.L. Hurst, R.J. Beynon and L.-Y. Lian, unpublished work).

Surface differences between MUPs

MUPs are detected directly by V2Rs (Vmn2r putative pheromone receptors) in the basal layer of the vomeronasal organ [10,38]. Most protein–protein interactions involve surface interaction and, from the morphology of the cavity, it is unlikely that a receptor protein would be able to access the protein interior, so a surface interaction is more likely. From the available structures, it is possible to extrapolate the positions of the surface-exposed residues of the group. The surface-exposed side chains (as defined using the program NACCESS, Hubbard and Thornton 1993) can be seen in Figure 2(E). In the central MUPs, these surface variations are limited to seven separate positions; the only positions within the central MUPs of C57BL/6 mice that vary at all are surface

Figure 3 | Comparison of cavity compactness between central and peripheral MUPs

(**A**) Interatomic distance between two equivalent residues in MUP11, darcin and MUP4. (**B**) Amino acids that line the cavity in the different MUPs: central MUP sequence represents all central MUPs as the cavity is invariant excepting position eight. Boxed residues indicate the residues chosen to illustrate the narrowing of the centre of the cavity in peripheral MUPs (darcin and MUP4) compared with central MUPs (MUP11). The interatomic distance is measured between closest non-hydrogen atoms for leucine Cδ1 (MUP11 and darcin) and phenylalanine Cζ (MUP4) at position 54, and alanine Cβ (MUP11), isoleucine Cδ2 (darcin) and phenylalanine Cζ (MUP4) at position 103.



	24	38	40	42	45	52	54	56	69	82	84	88	90	92	101	103	105	116	118	120
central	L	F	L	L	I	L	L	F/V	M	V	Y	N	F	I	L	A	L	L	G	Y
MUP3	E	M	A	V	I	L	F	F	A	M	Y	N	F	I	I	I	L	L	E	Y
MUP4	L	M	V	V	I	L	F	F	L	V	Y	N	F	I	I	F	L	L	E	Y
MUP5	L	M	V	V	I	L	F	F	L	V	Y	N	F	I	I	F	L	L	E	Y
MUP6	L	M	L	V	I	L	F	F	L	V	Y	K	F	V	I	F	L	L	S	F
darcin	L	M	V	V	I	L	L	F	L	V	Y	N	F	I	I	I	L	L	E	Y
MUP21	E	M	A	V	I	L	F	F	L	L	Y	N	F	I	I	I	L	L	E	Y

exposed (Figure 2E). This is not entirely surprising given that a protein recognition site would need to be accessible to a protein receptor and the surface variation allows the possibility for MUP-specific recognition.

By contrast, peripheral MUPs exhibit much greater variation on the protein surface (Figure 3E). Without structures of all MUPs it is difficult to predict whether these residues will form patches of variability on the surface of the protein; homology modelling being of limited use due to the problems in modelling the greater degree of freedom of motion and interaction with the aqueous environment in surface-facing amino acids. However, a previous study we carried out with darcin identified three such clusters that did highlight variation in surface properties by comparison with central MUPs. This change was from a polar hydrophilic surface comprising a glutamine (Glu^{44}) and a serine (Ser^{68}) residue to an aromatic hydrophobic surface comprising a tyrosine (Tyr^{44}) and a phenylalanine (Phe^{68}) residue (Figure 3E). The presence of a hydrophobic patch exposed to an aqueous environment may mediate darcin binding with its receptor.

Other variations between MUPs

Another intriguing piece of the puzzle is the variance in stability between MUPs. As these proteins are detected in drying urine left exposed to the environment over many days, a stable structure is important to maintain the signal for this period [39]. The structure of a β-barrel, being a network of hydrogen bonds, is demonstrably stable. MUPs extend the time domain of volatile signals in urine scent marks [40]. However, what was not anticipated was that stability between MUPs would vary. A chemical denaturation assay using increasing concentrations of urea has shown that darcin maintains its overall fold and function at much higher concentrations of urea than MUP11. The inherent stability of the peripheral MUP darcin, over central MUP11, may be associated with the specific function of darcin; being more stable infers that it has a longer 'life'. An explanation of why this is the case may be related to subtle differences in the cavity, with bulkier side chains in darcin providing a more stable hydrophobic core to the protein.

Conclusion

The MUPs present a substantial challenge in defining the distinctiveness in highly homologous structures that demonstrate different functional effects in mammalian communication. Potential mechanisms for mice to differentiate between MUPs are most likely down to the cavity, surface and inherent stability. The peripheral MUPs with specific roles

Table 1 | **Summary of three-dimensional structures of MUPs**

The ligands are subdivided into (a) larger cyclic molecules {TZL, 2-(*sec*-butyl)thiazole; 25R, 2,5-dimethylpyrazine; ZBT, 2-[(1*S*)-1-methylpropyl]-4,5-dihydro-1,3-thiazole; XBT, 2-[(1R)-1-methylpropyl]-4,5-dihydro-1,3-thiazole; PRZ, 2-isobutyl-3-methoxypyrazine; IPZ, 2-isopropyl-3-methoxypyrazine} and (b) short hydrocarbon chains {GOL, glycerol; 2EH, (2*S*)-2-ethylhexan-1-ol; LTL, 6-hydroxy-6-methyl-heptan-3-one; HTX, heptan-2-one; DE1, decan-1-ol; ODI, octane-1,8-diol; F09, nonan-1-ol; OC9, octan-1-ol; HE4, heptan-1-ol; HE2, hexan-1-ol; PE9, pentan-1-ol}. (c) MUPs without ligands. NMR structures are italicized; X-ray structures are not italicized.

(a) Cyclic ligands

Ligand	PDB code	MUP
TZL	1JV4	MUP10
PRZ	1QY1	MUP10
IPZ	1QY2	MUP10
PRZ	1YP6	MUP10
TZL	1MUP	MUP11
TZL	1I06	MUP11
XBT ZBT	3KFF	MUP4
25R	3KFI	MUP4

(b) Short hydrocarbon chain ligands

Ligand	PDB code	MUP
GOL	1QY0	MUP10
PE9 ×2	1ZND	MUP10
HE9	1ZNE	MUP10
HE4	1ZNG	MUP10
OC9	1ZNH	MUP10
F09	1ZNK	MUP10
DE1	1ZNL	MUP10
ODI	2DM5	MUP10
LTL	1I05	MUP11
2EH	3KFH	MUP4
HTX	3KFG	MUP4

(c) Without ligand

PDB code	MUP
1DF3	*MUP10*
2OZQ	MUP10
1YP7	MUP10
2L9C	*Darcin*
2LB6	*MUP11*
1I04	MUP11

such as darcin (a sexual attraction pheromone) and possibly MUP4 (which is a nasal rather than a urinary MUP) may be afforded function by the higher selectivity of their cavities and individual surface properties. Furthermore, the fact that mice can distinguish between individual profiles of central MUPs with identical small-molecule cavity properties suggests that the ability to discern these MUP patterns may arise from recognition of the MUP itself and not just recognition of the small molecule volatile that is bound.

The MUPs annotated in the MGI are not completely representative of MUPs expressed by wild house mice or indeed other murid species. As studies extend further, examples of central and peripheral MUPs with other specialized functions are likely to come to light. With the structural analysis currently limited to four individual sequences (two central and two peripheral), it is important that structural studies keep abreast of the critical MUPs in the signalling repertoire of the mouse.

Funding

This work was funded by the Biotechnology and Biological Sciences Research Council [sLoLa grant number BB/J002631/1 (to J.L.H., R.J.B. and L.-Y.L.)].

References

1 Beynon, R. and Hurst, J. (2003) Multiple roles of major urinary proteins in the house mouse, *Mus domesticus*. Biochem. Soc. Trans. **32**, 393–396
2 Garratt, M., Stockley, P., Armstrong, S.D., Beynon, R.J. and Hurst, J.L. (2011) The scent of senescence: sexual signalling and female preference in house mice. J. Evol. Biol. **24**, 2398–2409 CrossRef PubMed
3 Hurst, J.L., Payne, C.E., Nevison, C.M., Marie, a D., Humphries, R.E., Robertson, D.H., Cavaggioni, A. and Beynon, R.J. (2001) Individual recognition in mice mediated by major urinary proteins. Nature **414**, 631–634 CrossRef PubMed
4 Beynon, R.J., Veggerby, C., Payne, C.E., Robertson, D.H. L., 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. **28**, 1429–1446 CrossRef PubMed
5 Payne, C.E., Malone, N., Humphries, R., Bradbrook, C., Veggerby, C., Beynon, R.J. and Hurst, J.L. (2001) Heterogeneity of major urinary proteins in house mice: population and sex differences. In Chemical Signals in Vertebrates 9 (Marchlewska-Koj, A., Lepri, J.J. and Müller-Schwarze, D., eds), pp. 233–240, Springer, New York
6 Cheetham, S.A., Thom, M.D., Jury, F., Ollier, W.E. R., Beynon, R.J. and Hurst, J.L. (2007) The genetic basis of individual-recognition signals in the mouse. Curr. Biol. **17**, 1771–1777 CrossRef PubMed
7 Sherborne, A.L., Thom, M.D., Paterson, S., Jury, F., Ollier, W.E. R., Stockley, P., Beynon, R.J. and Hurst, J.L. (2007) The genetic basis of inbreeding avoidance in house mice. Curr. Biol. **17**, 2061–2066 CrossRef PubMed
8 Roberts, S.A., Simpson, D.M., Armstrong, S.D., Davidson, A.J., Robertson, D.H., McLean, L., Beynon, R.J. and Hurst, J.L. (2010) Darcin: a male pheromone that stimulates female memory and sexual attraction to an individual male's odour. BMC Biol. **8**, 75 CrossRef PubMed
9 Roberts, S.A., Davidson, A.J., McLean, L., Beynon, R.J. and Hurst, J.L. (2012) Pheromonal induction of spatial learning in mice. Science **338**, 1462–1465 CrossRef PubMed
10 Chamero, P., Marton, T.F., Logan, D.W., Flanagan, K., Cruz, J.R., Saghatelian, A., Cravatt, B.F. and Stowers, L. (2007) Identification of protein pheromones that promote aggressive behaviour. Nature **450**, 899–902 CrossRef PubMed
11 Darwish Marie, A., Veggerby, C., Robertson, D.H.L., 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 CrossRef PubMed
12 Kwak, J., Grigsby, C.C., Rizki, M.M., Preti, G., Köksal, M., Josue, J., Yamazaki, K. and Beauchamp, G.K. (2012) Differential binding between volatile ligands and major urinary proteins due to genetic variation in mice. Physiol. Behav. **107**, 112–120 CrossRef PubMed

13 Novotny, M.V., Jemiolo, B., Wiesler, D., Ma, W., Harvey, S., Xu, F., Xie, T.-M. and Carmack, M. (1999) A unique urinary constituent, 6-hydroxy-6-methyl-3-heptanone, is a pheromone that accelerates puberty in female mice. Chem. Biol. **6**, 377–383 CrossRef PubMed

14 Bacchini, A., Gaetani, E. and Cavaggioni, A. (1992) Pheromone binding proteins of the mouse, *Mus musculus*. Experientia **48**, 419–421 CrossRef PubMed

15 Robertson, D.H. L., Beynon, R.J. and Evershed, R.P. (1993) Extraction, characterization, and binding analysis of two pheromonally active ligands associated with major urinary protein of house mouse (*Mus musculus*). J. Chem. Ecol. **19**, 1405–1416 CrossRef PubMed

16 Nevison, C.M., Armstrong, S., Beynon, R.J., Humphries, R.E. and Hurst, J.L. (2003) The ownership signature in mouse scent marks is involatile. Proc. Biol. Sci. **270**, 1957–1963 CrossRef PubMed

17 Sharrow, S.D., Vaughn, J.L., Ídek, L.Š.Ž., Novotny, M. V. and Stone, M.J. (2002) Pheromone binding by polymorphic mouse major urinary proteins. Protein Sci. **11**, 2247–2256 CrossRef PubMed

18 Logan, D.W., Marton, T.F. and Stowers, L. (2008) Species specificity in major urinary proteins by parallel evolution. PLoS ONE **3**, e3280 CrossRef PubMed

19 Finger, J.H., Smith, C.M., Hayamizu, T.F., McCright, I.J., Eppig, J.T., Kadin, J.A., Richardson, J.E. and Ringwald, M. (2011) The mouse Gene Expression Database (GXD): 2011 update. Nucleic Acids Res. **39**, D835–D841 CrossRef PubMed

20 Eppig, J.T., Blake, J.A., Bult, C.J., Kadin, J.A. and Richardson, J.E. (2012) The Mouse Genome Database (MGD): comprehensive resource for genetics and genomics of the laboratory mouse. Nucleic Acids Res. **40**, D881–D886 CrossRef PubMed

21 Mudge, J.M., Armstrong, S.D., McLaren, K., Beynon, R.J., Hurst, J.L., Nicholson, C., Robertson, D.H., Wilming, L.G. and Harrow, J.L. (2008) Dynamic instability of the major urinary protein gene family revealed by genomic and phenotypic comparisons between C57 and 129 strain mice. Genome Biol. **9**, R91 CrossRef PubMed

22 Mudge, J.M., Armstrong, S.D., McLaren, K., Beynon, R.J., Hurst, J.L., Nicholson, C., Robertson, D.H., Wilming, L.G. and Harrow, J.L. (2008) Dynamic instability of the major urinary protein gene family revealed by genomic and phenotypic comparisons between C57 and 129 strain mice. Genome Biol. **9**, R91 CrossRef PubMed

23 Sharrow, S.D., Vaughn, J.L., Zídek, L., Novotny, M.V. and Stone, M.J. (2002) Pheromone binding by polymorphic mouse major urinary proteins. Protein Sci. **11**, 2247–2256 CrossRef PubMed

24 Perez-Miller, S., Zou, Q., Novotny, M. V. and Hurley, T.D. (2010) High resolution X-ray structures of mouse major urinary protein nasal isoform in complex with pheromones. Protein Sci. **19**, 1469–1479 CrossRef PubMed

24a Beynon, R.J., Armstrong, S.D., Gómez-Baena, G., Lee, V., Simpson, D., Unsworth, J. and Hurst, J.L. (2014) The complexity of protein semiochemistry in mammals. Biochem. Soc. Trans. **42**, 837–845

25 Rose, P.W., Bi, C., Bluhm, W.F., Christie, C.H., Dimitropoulos, D., Dutta, S., Green, R.K., Goodsell, D.S., Prlic, A., Quesada, M. et al. (2013) The RCSB Protein Data Bank: new resources for research and education. Nucleic Acids Res. **41**, D475–D482 CrossRef PubMed

26 Bingham, R.J., Findlay, J.B. C., Hsieh, S.-Y., Kalverda, A.P., Kjellberg, A., Perazzolo, C., Phillips, S.E.V., Seshadri, K., Trinh, C.H., Turnbull, W.B. et al. (2004) Thermodynamics of binding of 2-methoxy-3-isopropylpyrazine and 2-methoxy-3-isobutylpyrazine to the major urinary protein. J. Am. Chem. Soc. **126**, 1675–1681 CrossRef PubMed

27 Barratt, E., Bingham, R.J., Warner, D.J., Laughton, C.A., Phillips, S.E.V. and Homans, S.W. (2005) van der Waals interactions dominate ligand–protein association in a protein binding site occluded from solvent water. J. Am. Chem. Soc. **127**, 11827–11834 CrossRef PubMed

28 Barratt, E., Bronowska, A., Vondrásek, J., Cerný, J., Bingham, R., Phillips, S. and Homans, S.W. (2006) Thermodynamic penalty arising from burial of a ligand polar group within a hydrophobic pocket of a protein receptor. J. Mol. Biol. **362**, 994–1003 CrossRef PubMed

29 Syme, N.R., Dennis, C., Phillips, S.E.V. and Homans, S.W. (2007) Origin of heat capacity changes in a 'nonclassical' hydrophobic interaction. ChemBioChem **8**, 1509–1511 CrossRef PubMed

30 Malham, R., Johnstone, S., Bingham, R.J., Barratt, E., Phillips, S.E.V., Laughton, C.A. and Homans, S.W. (2005) Strong solute–solute dispersive interactions in a protein–ligand complex. J. Am. Chem. Soc. **127**, 17061–17067 CrossRef PubMed

31 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 CrossRef PubMed

32 Kuser, P.R., Ferrari, E. and Spisni, A. (2001) The X-ray structure of a recombinant major urinary at 1.75 Å resolution: a comparative study of protein at 1.75 Å X-ray and NMR-derived structures. Acta Crystallogr. D Biol. Crystallogr. **57**, 1863–1869 CrossRef PubMed

33 Lücke, C., Franzoni, L., Abbate, F., Löhr, F., Ferrari, E., Sorbi, R.T., Rüterjans, H. and Spisni, A. (1999) Solution structure of a recombinant mouse major urinary protein. Eur. J. Biochem. **266**, 1210–1218 CrossRef PubMed

34 Böcskei, Z., Groom, C.R., Flower, D.R., Wright, C.E., Phillips, S.E., Cavaggioni, A., Findlay, J.B. and North, A.C. (1992) Pheromone binding to two rodent urinary proteins revealed by X-ray crystallography. Nature **360**, 186–188 CrossRef PubMed

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

36 Binkowski, T.A., Naghibzadeh, S. and Liang, J. (2003) CASTp: Computed Atlas of Surface Topography of proteins. Nucleic Acids Res. **31**, 3352–3355 CrossRef PubMed

37 Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. **372**, 774–797 CrossRef PubMed

38 Chamero, P., Katsoulidou, V., Hendrix, P., Bufe, B., Roberts, R., Matsunami, H., Abramowitz, J., Birnbaumer, L., Zufall, F. and Leinders-Zufall, T. (2011) G protein $G\alpha_o$ is essential for vomeronasal function and aggressive behavior in mice. Proc. Natl. Acad. Sci. U.S.A. **108**, 12898–12903 CrossRef PubMed

39 Hurst, J.L. and Beynon, R.J. (2004) Scent wars: the chemobiology of competitive signalling in mice. BioEssays **26**, 1288–1298 CrossRef PubMed

40 Hurst, J., Robertson, D., Tolladay, U. and Beynon, R. (1998) Proteins in urine scent marks of male house mice extend the longevity of olfactory signals. Anim. Behav. **55**, 1289–1297 CrossRef PubMed

Received 23 April 2014
doi:10.1042/BST20140082