Biomol NMR Assign
DOI 10.1007/s12104-010-9253-6

ARTICLE

^{1}H, ^{15}N and ^{13}C resonance assignment of darcin, a mouse major urinary protein

Marie M. Phelan · Lynn McLean · Deborah M Simpson · Jane L Hurst · Robert J Beynon · Lu-Yun Lian

Received: 4 May 2010 / Accepted: 2 August 2010

Abstract Darcin is an important lipocalin of the urinary MUP family. These beta-barrel structures differ subtly in sequence and function and facilitate communication between members of the mouse population via scent marks. Polymorphism within the family has led to the hypothesis that individual MUPs can also contribute to social and physiological information of the scent owner and thus demonstrates the necessity for structural investigation of these variations. Using conventional triple resonance experiments, ^{1}H ^{15}N and ^{13}C assignment of recombinant N terminal hexa-histidine tagged Darcin has been achieved. The corresponding chemical shifts have been deposited in the BioMagResBank; Accession No. 16840.

Keywords Darcin · Major urinary protein (MUP) · Male specific protein · NMR · Resonance assignment

Biological context

In the mouse, the Major Urinary Proteins (MUPs) are a highly polymorphic class of ~20 kDa lipocalins, many of which are synthesised in the liver and released in the urine to be deposited in scent marks. They fulfil multiple roles in chemical signalling between conspecifics, including the binding and delayed release of volatile pheromones (Hurst et al. 1998; Armstrong et al. 2005), the encoding of the ownership of scent marks (Hurst et al. 2001; Nevison et al. 2003; Cheetham et al. 2007), permitting the assessment of genetic heterozygosity (Thom et al. 2008) and driving inbreeding avoidance (Sherborne et al. 2007). Although wild mice demonstrate highly variable individual patterns of MUP expression (Hurst et al. 2001), these patterns are remarkably homogenous among inbred laboratory mouse strains (Cheetham et al. 2009).

Darcin (also known as "Peak IV MUP" (Armstrong et al. 2005)) is an 'atypical' 18,893 Da MUP that differs from other MUPs in primary sequence. It is a peripheral gene in the MUP gene cluster and is an outlier in the MUP phylogeny of C57B/6J mice (Mudge et al. 2008; MGI:3651981). It is the MUP variant that is responsible for the binding and delayed release of the volatile pheromone 2-*sec*-butyl-dihydrothiazole (Armstrong et al. 2005). It differs significantly in nucleotide sequence from the other MUPs, with the central gene MUPS sharing an average nucleotide identity of 99.2%, whereas the peripheral gene MUPs, including darcin, sharing only 88.2% (Mudge et al. 2008). Darcin also exhibits atypical mobility on SDS–PAGE (Armstrong et al. 2005). Expression is strongly male-specific among adult wild house mice (Armstrong et al. 2005; Hurst unpublished data). While it is consistently expressed by male laboratory mice of the C57 lineage, expression is much more variable among strains from the Castle and Swiss lineages (Cheetham et al. 2009). New research on darcin indicates that it is a sexual attraction pheromone that plays a key role in attracting females to the scent of particular males (Hurst unpublished data).

M. M. Phelan · L.-Y. Lian (✉)
NMR Centre for Structural Biology, Biological Sciences, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK
e-mail: Lu-Yun.Lian@liverpool.ac.uk

L. McLean · D. MSimpson · R. JBeynon
Protein Function Group, University of Liverpool, Crown Street, Liverpool L69 7ZJ, UK

J. LHurst
Mammalian Behaviour & Evolution Group, University of Liverpool, Leahurst Campus, Neston CH64 7TE, UK

Published online: 12 August 2010

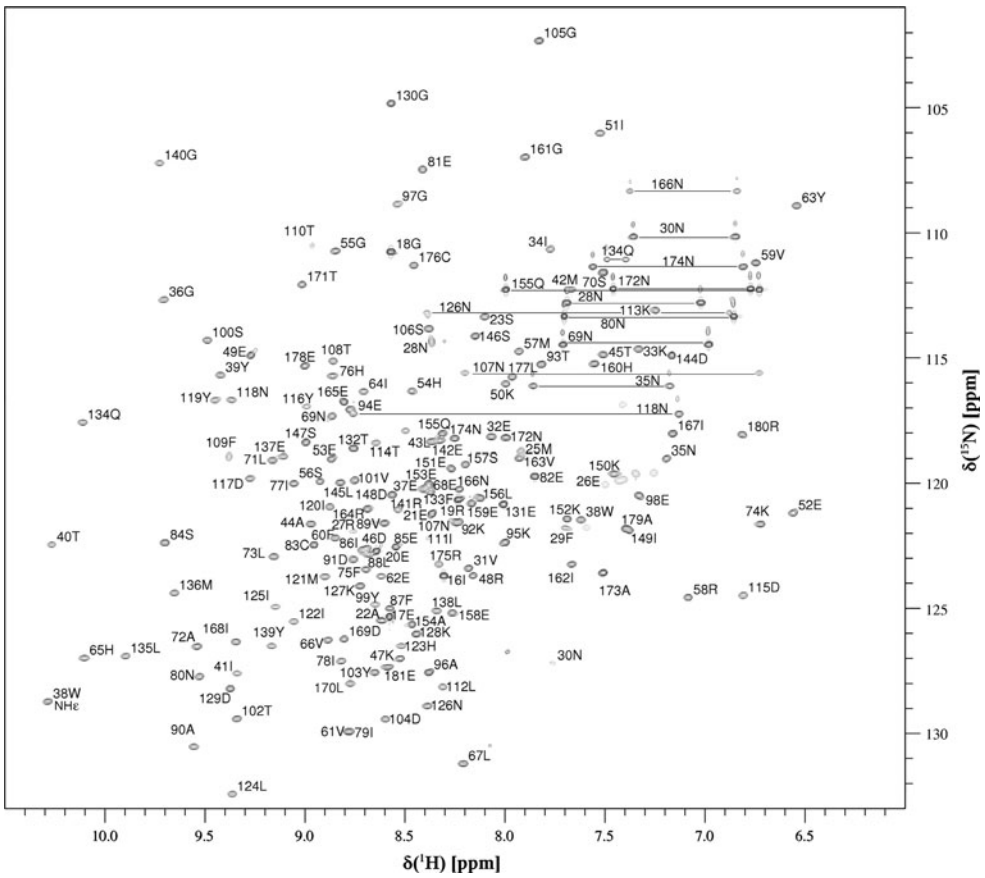

Fig. 1 $^{1}H^{15}N$ HSQC depicting NH assignment for the backbone amides of darcin, folded peaks corresponding to sidechain Arg NH_{ε} are depicted in *light grey* with backbone NH, Gln $_{\delta}NH_2$, Asn $_{\gamma}NH_2$ and Trp NH_{ε} peaks shown in *dark grey*. The Asn and Gln sidechain peaks are lined by *solid lines* with the sidechain of W38 labelled 'NH_{ε}'

Methods and experiments

Native darcin was isolated from mouse urine of the C57 lineage and the darcin sequence was determined by mass spectrometry (Armstrong et al. 2005; Cheetham et al. 2009). This MUP sequence correlates to UniProt sequence Q5FW60 identified from a cDNA library. The isolated native protein lacks nineteen N-terminal amino acid signal peptides identified in the cDNA sequence. We have therefore numbered the native darcin residues from residue twenty (Glu20). A codon-optimised synthetic gene for darcin was expressed in *E. coli* to facilitate the incorporation of ^{15}N and ^{13}C uniform isotopic labelling. The gene encoding darcin was cloned into expression vector pET28b resulting in the first fourteen residues of the recombinant protein (residues numbered 6–19) corresponding to the cloning and hexa-histidine purification tag which are not part of the native sequence. Recombinant expression was carried out in BL21(DE3) *E. coli* cells at 37°C in the presence of 30 μg/mL kanamycin. To facilitate NMR assignment uniform ^{15}N ^{13}C labelling was achieved by expression in M9 minimal media enriched with ^{13}C labelled glucose and ^{15}N labelled ammonium sulphate. Purification was via a Ni^{2+} affinity chromatography using 250 mM imidazole as eluent, prior to desalting by dialysis. Purity was assessed by polyacrylamide gel electrophoresis and mass spectrometry.

The optimal conditions for NMR investigation are 1 mM darcin in 25 mM potassium phosphate, pH 6.8, 300 K, with 0.2% NaN_3 and 10% v/v $^{2}H_2O$. NMR spectra were aquired on Bruker AVANCE 600 MHz and 800 MHz spectrometers equipped with 5 mm triple resonance cryoprobes. Spectra were processed using Topspin2.1 (Bruker) and the Azara processing package provided as part of the CCPN suite with assignment carried out using CCPN Analysis (Vranken et al. 2005). Triple resonance assignment was obtained utilising two dimensional HC and HN HSQCs in conjunction with conventional backbone and side chain experiments.

Assignment and data deposition

Excluding the N-terminal purification tag, 97.3% resonance assignment has been completed from the expected

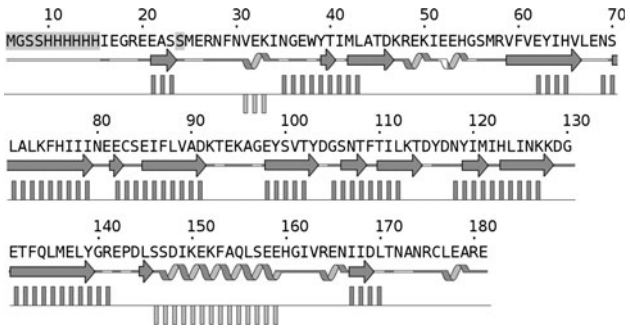

Fig. 2 Sequence of recombinant darcin, amides not assigned are highlighted. Cartoon representation of secondary structure was calculated using the software Dangle (Cheung et al. 2010) based upon assigned $C_{\alpha}C_{\beta}H_{\alpha}C'$ and NH, *unshaded areas* of the cartoon indicate regions of low confidence. The consensus secondary structure calculated using Chemical Shift Index (CSI) (Wishart and Sykes 1994) is also shown. An index of +1 is depicted by *dark grey bars* rising above the median and corresponding to beta strands whereas an index of −1 is depicted by *light grey bars* dropping below the median and corresponding to helical regions

observable ^{1}H ^{13}C and ^{15}N chemical shifts under these conditions. This includes 99.7% backbone NH, CH_{α} and C′ assignment. The HN HSQC (Fig. 1) identifies the primary amide assignments and additionally side chain amides of Asn and Gln marked by horizontal lines together with NH_{ε} of Trp and Arg. Due to spectral folding the Arg NH_{ε} can be seen around 115 ppm, their true $^{15}N_{\varepsilon}$ resonances occur between 81 and 85 ppm. The chemical shift dispersion observed in the HSQC and the chemical shift index indicates the presence of extended beta strands (Fig. 2). These extended beta strands are consistent with an eight stranded beta barrel common to lipocalins with the extended alpha helix predicted between residues Ser146 and Glu159 is also common to the lipocalin family. The corresponding chemical shifts have been deposited in the BioMagResBank; Accession No. 16840.

Acknowledgments This was work supported by BBSRC grant (BBC503897) to JLH and RJB.

References

Armstrong SD, Robertson DHL, Cheetham SA, Hurst JL, Beynon RJ (2005) Structural and functional differences in isoforms of major urinary proteins: a male specific protein that preferentially binds a male pheromone. Biochem J 391:343–350

Cheetham SA, Thom MD, Jury F, Ollier WER, Beynon RJ, Hurst JL (2007) The genetic basis of individual recognition signals in the mouse. Curr Biol 17:1771–1777

Cheetham SA, Smith AL, Armstrong SD, Beynon RJ, Hurst JL (2009) Limited variation in the major urinary proteins of laboratory mice. Physiol Behav 96:253–261

Cheung MS, Maguire ML, Stevens TJ, Broadhurst RW (2010) DANGLE: a Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. J Magn Reson 202(2):223–233

Hurst JL, Robertson DHL, Tolladay U, Beynon RJ (1998) Proteins in urine scent marks of male house mice extend the longevity of olfactory signals. Anim Behav 55:1289–1297

Hurst JL, Payne CE, Nevison CM, Marie AD, Humphries RE, Robertson DHL, Cavaggioni A, Beynon RJ (2001) Individual recognition in mice mediated by major urinary proteins. Nature 414:631–634

Mudge JM, Armstrong SD, McLaren K, Beynon RJ, Hurst JL, Nicholson C, Robertson DHL, Wilming LG, Harrow JL (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. http://genomebiology.com/2008/9/5/R91

Nevison CM, Armstrong SD, Beynon RJ, Humphries RE, Hurst JL (2003) The ownership signal in mouse scents marks is involatile. Proc Royal Soc B 270:1957–1963

Sherborne AL, Thom MD, Paterson S, Jury F, Ollier WER, Stockley P, Beynon RJ, Hurst JL (2007) The genetic basis of inbreeding avoidance in house mice. Curr Biol 17:2061–2066

Thom MD, Stockley P, Jury F, Ollier WER, Beynon RJ, Hurst JL (2008) The direct assessment of genetic heterozygosity through scent in the mouse. Curr Biol 18:619–623

Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 59:687–696

Wishart DS, Sykes BD (1994) The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. J Biomol NMR 4:171–180