Eukaryotic expression, purification and structure/function analysis of native, recombinant CRISP3 from human and mouse

Marianna Volpert, Jonathan E. Mangum, Duangporn Jamsai, Rebecca D'Sylva, Moira K. O'Bryan & Peter McIntyre

1Dept of Pharmacology, University of Melbourne, Parkville, VIC, Australia, 2Dept of Anatomy and Developmental Biology, Monash University, Clayton, VIC, Australia, 3Health Innovations Research Institute, RMIT University, Bundoora, VIC, Australia.

While the Cysteine-Rich Secretory Proteins (CRISPs) have been broadly proposed as regulators of reproduction and immunity, physiological roles have yet to be established for individual members of this family. Past efforts to investigate their functions have been limited by the difficulty of purifying correctly folded CRISPs from bacterial expression systems, which yield low quantities of correctly folded protein containing the eight disulfide bonds that define the CRISP family. Here we report the expression and purification of native, glycosylated CRISP3 from human and mouse, expressed in HEK 293 cells and isolated using ion exchange and size exclusion chromatography. Functional authenticity was verified by substrate-affinity, native glycosylation characteristics and quaternary structure (monomer in solution). Validated protein was used in comparative structure/function studies to characterise sites and patterns of N-glycosylation in CRISP3, revealing interesting inter-species differences.

Expression of the Cysteine-Rich Secretory Proteins (CRISPs) occurs primarily in the male reproductive tissues and salivary glands; in mammals, expression is biased to the reproductive tract, while in many poisonous non-mammalian vertebrates such as snakes and lizards, CRISPs are secreted into venom where they exert toxicity through ion channel inhibition. The mechanisms by which CRISPs function in the mammalian reproductive tract are less clear, however individual roles in the processes of decapacitation, acrosome reaction, sperm-oocyte fusion, flagellar motility and immunosuppression have recently been proposed. The latter function has been attributed to equine CRISP3, which is present in high concentrations in seminal plasma and has been shown to inhibit sperm-neutrophil interaction in horses. Whether this function of equine CRISP3 extends to its mouse and human orthologs remains to be determined.

Structurally, CRISPs constitute one of eleven mammalian subfamilies of the CAP (CRISP, Antigen 5 and Pathogenesis-Related 1 proteins) superfamily, to which they belong by virtue of their N-terminal CAP domain. They are characterised by the presence of a C-terminal CRISP domain, a collective term encompassing an ion channel regulator (ICR) domain and the ‘hinge’ which links the ICR and CAP domains (Figure 1). As its name suggests, the CRISP domain comprises an unusually cysteine-rich sequence in which ten cysteine residues form five disulfide bonds; the CAP domain of the protein encodes a further six cysteine residues which are also exclusively disulfide bonded. These eight disulfide bonds confer to CRISPs their high degree of structural complexity. Unsurprisingly therefore, recombinant expression of functional CRISPs has proven difficult, with a number of previous attempts resulting in the production of misfolded, insoluble protein aggregates by bacterial systems; this material must then be subjected to an involved and often inefficient refolding process. This, along with the problem of often inadequate post-translational modification of mammalian proteins by bacteria, creates a need for an improved approach to CRISP expression and purification.

Here we used a mammalian protein expression system to generate useful quantities of soluble, correctly folded and post-translationally modified human CRISP3, as validated by interaction with the known serum binding partner alpha-1-B glycoprotein. We also show that this validated CRISP3 preparation can provide informative results in structure/function analysis. In particular, differences in glycosylation between human and mouse CRISP3 were characterised.
Expression of soluble human CRISP3

Results

Expression of soluble human CRISP3. Stable, inducible HEK 293 cell lines expressing human or mouse CRISP3 were generated. Following 96–120 hours of induction, human His-tagged CRISP3 constituted approximately 1.8% of total protein in the cell media (1.5 μg/ml by comparison with a CRISP3 standard). Protein stability and lack of toxicity was indicated by a proportional increase in CRISP3 concentration with induction time (Figure 2A).

Purification and validation of human CRISP3. We found that the nickel/cobalt coated resins used for His-tag purification lacked specificity with conditioned HEK cell medium, such that many secreted proteins bound to the column despite optimisation attempts. As an alternative enrichment method, we employed carboxymethyl (CM) cellulose cation exchange to achieve a fifteen-fold purification and 75% recovery. Subsequent size-exclusion chromatography achieved a further 2.5-fold enrichment with near-quantitative recovery and 67% purity (Figure 3B and Table 1). Immunoreactivity, SDS-PAGE analysis and glycosylation profile data were supported by mass spectrometry analysis of tryptic peptides from the predominant 26-kDa bands, which confirmed their identity as mouse and human CRISP3 (three high-quality sequence tags for each, not shown).

A single major contaminating protein in both human and mouse preparations was identified by mass spectrometry as alpha-enolase, a 48 kDa glycolytic enzyme, with a Mascot score of 159 (6 peptides, 19% sequence coverage). Secreted alpha enolase binds CM cellulose25–27.

Recombinant human CRISP3 binds A1BG. In the absence of a standard CRISP3 activity assay, we tested the ability of purified human CRISP3 to perform its only reported biological function, that of binding the serum protein alpha-1-B glycoprotein (A1BG), itself a protein of uncertain function28–30. Silver staining of CRISP3-binding human and bovine serum proteins revealed bands corresponding to A1BG in CRISP3-treated samples but not in untreated controls (Figure 4A). Mass spectrometry analysis confirmed their identity as human and bovine A1BG with Mascot scores of 343 (10 peptides, 26% sequence coverage) and 170 (4 peptides, 12% sequence coverage) respectively. Reduction and alkylation of recombinant CRISP3 practically abolished the CRISP3-A1BG interaction, indicating that the recombinant protein contains tertiary structure elements required for A1BG recognition (Figure 4B).

Identification of mouse and human CRISP3 N-glycosylation sites. Purification of mouse CRISP3 (Figure 3C) revealed that its putative glycosylated form possessed lower electrophoretic mobility than glycosylated human CRISP3 (Figure 5A). To identify the site(s) of N-glycosylation in mouse CRISP3, site-directed mutagenesis was performed at three NXS/T motifs; N99, N113 and N156, where amino acid number reflects post-cleavage position. These were mutated individually and in combination to produce the mutants N99Q, N113Q and N156Q. Western blot analysis showed that N99Q and N113Q (but not N156Q) mutations affected protein size, indicating glycosylation at both sites in the wild type protein (Figure 5B, C). Interestingly, we observed that two mutants (N99Q + N113Q and N99Q + N113Q + N156Q) incapable of undergoing glycosylation were of relatively low abundance in the cell media (Figure 5C).

N-glycosylation of human CRISP3 at N220 only was confirmed, with no observable glycosylation of hCRISP3-N220Q (Figure 6B, last lane), and no impairment in protein secretion.

Partial characterisation of mouse CRISP3 N-glycosylation types. To investigate the nature of mouse CRISP3 glycosylation, the N99Q and N113Q mutants were exposed to the deglycosylase Endoglycosidase H (Endo H). Endo H leaves complex oligosaccharides intact while cleaving high mannose and some hybrid N-linked oligosaccharides. Under standard reaction conditions, partial digestion of the mouse wild type, N99Q and N113Q proteins was observed following Endo H treatment (Figure 6A); none of a range of time and temperature conditions employed improved sensitivity to the enzyme. Human CRISP3 was resistant to Endo H, as expected30. Western blot comparison of mouse and human glycans revealed a size difference consistent with complex and high mannose modifications respectively (Figure 6B, lanes 1 and 2).

Finally, in a comparison of the relative mobility of salivary gland CRISP3 with three post-translationally modified variants of the recombinant protein, only the unglycosylated form of mouse CRISP3 could be observed in whole salivary gland (Figure 7).
Discussion

The structural complexity conferred by the high cysteine content of CRISPs has repeatedly caused purification attempts from bacterial cells to result in the production of insoluble inclusion bodies even under conditions intended to favour disulfide bond formation\textsuperscript{12,13}. Alternative purification methods from neutrophil extracts have employed cation exchange and affinity chromatography, but these approaches are limited by tissue availability\textsuperscript{18,19}. Here we demonstrate that recombinant human CRISP3 expression and two-step purification from HEK 293 cells yields milligram quantities of soluble protein suitable for structure/functional analyses. Indications of functional authenticity included efficient secretion, post-translational modification and confirmed interaction with the known binding partner A1BG, which was abolished with denaturation of CRISP3. The recombinant protein reproduced the native CRISP3 glycosylation profile, with site-directed mutagenesis confirming N220 as the exclusive site of N-glycosylation. A published report of Endo H resistance suggesting a hybrid or complex carbohydrate modification was confirmed\textsuperscript{19}.

We present the first report of mouse CRISP3 immunodetection by western blotting using a mouse CRISP3-specific antibody of our own production, as well as the first evidence of N-glycosylation at two sites (N99 and N113). While incomplete digestion by Endo H of both glycosylation sites may reflect either imperfect enzyme access to cleavage sites or a mixture of glycosylation types, both N99 and N113 likely represent sites of high mannose or hybrid glycosylation to some degree. A relationship between glycosylation state and protein function (affecting protein stability or intracellular retention) was suggested by the relatively low abundance of unglycosylated mouse CRISP3. We speculate that glycosylation of mouse, but not human CRISP3, is required for escape from the ER quality control system\textsuperscript{25}.

A gel comparison of salivary gland CRISP3 to three post-translationally modified variants of the recombinant protein indicated that

---

Table 1 | CRISP3 purification from HEK 293 conditioned media

| Step          | Volume (ml) | Protein (mg)\textsuperscript{*} | hCRISP3 (µg)\textsuperscript{†} | Purification (fold) | Recovery (%) |
|---------------|-------------|---------------------------------|-------------------------------|---------------------|-------------|
| 1. Media      | 500         | 40                              | 720                           | 7                   | 100         |
| 2. CM cellulose| 10          | 2.2                             | 540                           | 15                  | 75          |
| 3. S75        | 10          | 0.75                            | 500                           | 37                  | 70          |

\textsuperscript{*}Determined by Bradford assay.
\textsuperscript{†}Determined by western blot comparison with commercial HIS-tagged CRISP3 standard and densitometry of silver-stained SDS PAGE (steps 1 and 2–3 respectively).
in whole salivary gland at least, mouse CRISP3 is wholly or predo-
mimantly unglycosylated. Clarification of glycosylation patterns in
other tissues is required; as we have demonstrated a link between
glycosylation and secretion efficiency, it will be of particular interest
to compare the glycosylation profile of CRISP3 from whole tissues
with that obtained from endocrine secretions. Finally, we note that
while human CRISP3 is glycosylated in the CRISP domain only
(N220), both mouse glycosylation sites N99 and N113 are found in
the N-terminal CAP domain (Figure 1B); this difference in glycosy-
lation pattern may indicate distinct functions. This glycosylation site
distribution appears to reflect a broader species-specific pattern in
which human CRISPs 1 and 3 are N-glycosylated in the ICR domain
with no N-glycosylation motifs in the CAP domain, while all four
mouse CRISPs have potential N-glycosylation sites in their CAP but
not CRISP domains. The major form of seminal plasma human
CRISP1 contains a 4 kDa glycan, consistent with complex modifica-
tion. N-glycosylation motif sequences are absent in human CRISP2
and equine CRISP3, to which a neutrophil-suppressive function has
been attributed.

Sequence differences between pseudochetoxin and pseudecin, two
cyclic nucleotide-gated ion channel blocking venom CRISPs with a
15–30-fold difference in affinity, suggest that most of the critical
functional residues responsible for contact with target ion channels
are concentrated on the concave surface of the proteins spanning the
CAP-CRISP domain interface. That this cleft formed by the
bridge between CRISP and CAP domains is involved in substrate
binding is supported by evidence that the ICR domain of CRISPs
alone provides inadequate or suboptimal ion-channel inhibition, so
that the CAP domain is believed to provide additional specificity to
the ion channel interaction. Indeed, sequence homology studies
have identified a collection of residues within snake venom CRISPs
which form a continuous line along the protein backbone and have
been subject to high frequency evolutionary variation. Sequence
variability in these regions is therefore likely to confer selectivity

---

**Figure 4** | Recombinant human CRISP3 binds human and bovine A1BG. (A) Gel analysis of proteins purified on recombinant CRISP3-coated cyanogen bromide resin showed CRISP3 interaction with A1BG from both human (lane 3, ~85 kDa band) and bovine (lane 3, ~65 kDa band) serum, as confirmed by mass spectrometry. Elution control samples were loaded as follows; lane 1: uncoated beads incubated with serum, lane 2: CRISP3-coated beads incubated with PBS, lane 3: CRISP3-coated beads incubated with serum. In the case of human serum, all eluted proteins in lane 3 other than A1BG bound non-specifically to the beads (compare with lane 1). (B) Denatured CRISP3 (lane 3) was unable to bind A1BG. Full length gels are available under supplementary data; gels were run under the same conditions.

**Figure 5** | Recombinant mouse CRISP3 is glycosylated at N99 and N113. Western blotting with His tag antibody showed glycosylated human and mouse CRISP3-His to be 29 and 33 kDa respectively (A). (B) Peptide pre-absorption of mouse CRISP3 antibody. (C) Mouse CRISP3 N99Q and N113Q mutant proteins were smaller than wild type (WT) mouse CRISP3 (PNGase F untreated samples); N156Q was unaffected. Two mutants unable to undergo glycosylation (N99Q + N113Q and N99Q + N113Q + N156Q) were of relatively low abundance. ‘NT’ represents non-transfected HEK cell media. Full length gels are available under supplementary data; gels were run under the same conditions.
and affinity for binding partners, accounting for the diverse range of CRISP ion channel targets. A comparison of CRISP and CAP proteins from multiple species similarly identified three hypervariable regions in the CAP domain which were postulated to confer substrate specificity. The mouse CRISP3 glycosylation site N113 lies within one of these hypervariable regions, corresponding to G109-V117, in the CAP domain. Due to the high degree of rotational freedom of the CAP and ICR domains and the lack of an ICR domain in many CAP proteins, a separate, non-ion channel binding function has often been postulated for the CAP domain. If the G109-V117 sequence, confers specificity to either an ion-channel interaction or distinct CAP domain function, then carbohydrate addition within this region may have significant functional implications. 25 and 12 amino acids separate the critical 5-amino acid variable ion-channel binding sequence in the ICR domain from the human CRISP3 and CRISP1 N-glycosylation sites, respectively. Despite the addition of a large, complex sugar moity in close proximity to the CRISP3 binding cleft, Udby et al. have previously shown that interaction with at least one binding partner, A1BG, occurs independently of CRISP3 glycosylation status.

In conclusion, this study describes a new method of producing CRISP3 of sufficient abundance and purity for structure/function analysis, as demonstrated by elucidation of the sites of glycosylation of mouse CRISP3, about which very little is currently known.

**Methods**

Cloning, transfection and cell culture. Human and mouse CRISP3 cDNA were amplified from total RNA extracts of the LNCaP prostate cancer cell line and wild type C57BL/6 mouse salivary gland respectively. cDNAs were cloned following the third post-cleavage amino acid. For human CRISP3, these primers were: forward primer 5’ ggacagagctgtaaagaagcgcctctactacgcatctacttacctcgactttgcgggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttacggtgtagctgagcatcttgcgggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttacggtgtagctgagcatcttgcgggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttacggtgtagctgagcatcttgcgggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggttggt tg
were reduced (2.5 mM DTT, 30 minutes, 60°C), alkylated (10 mM iodoacetamide, 30 minutes, room temperature in the dark), and digested (0.5 μg trypsin, 20 mM ammonium bicarbonate, 37°C for 14 hours). Tryptic peptides were analysed by LC-MS/MS using an HCT ULTRA ion trap mass spectrometer (Bruker Daltonics, Bremen, Germany) coupled to a nano HPLC (Ultimate 3000, Dionex Corporation, SunnyBrook, CA, USA), largely as done previously. Identity searches were conducted using the MASCOT MS/MS Ion Search engine (www.matrixscience.com) against the NCBIRef and Swiss-Prot databases with mammalian taxonomy selected. Search parameters were: fixed modifications, carboxymethyl; variable modifications, oxidation (Met), up to one missed cleavage allowed.

Preparation of human tissues. The collection and use of tissues described in this study was approved by the Human Research and Ethics Committee of Southern Health (Monash Medical Centre, Monash Day Surgery and Monash University). Human seminal plasma was obtained with informed consent from Monash IVF (Clayton, Australia) and prepared for western blotting as described elsewhere. Human blood was provided with informed consent by the Monash Medical Centre (Clayton Australia).

Preparation of mouse tissues. Animal procedures were performed in accordance with the Australian NHMRC Guidelines on Ethics in Animal Experimentation and approved by the Monash University Animal Experimentation Ethics Committees. Whole mouse salivary gland was obtained from 12 week old C57BL/6 mice, frozen in liquid nitrogen, powdered with a mortar and pestle and resuspended in tissue lysis buffer (1% SDS, 50 mM Tris, 150 mM NaCl, 1% Triton-X, pH8) containing protease inhibitors (Calbiochem, 359345). Samples were centrifuged at maximum speed (20 minutes, 4°C) and 100 μg of protein was prepared for western blotting.

Homology modelling. A mouse CRISP3 homology model based on the crystal structure template of Natrin, a snake venom CRISP from Taiwan Cobra (pdb 1XTA), was generated using the prime module from the Schrodinger molecular modelling software suite with missing residues added (Prime version 3.8 Schrodinger, Inc, New York, NY 2010). The model was further refined and visualized with VMD software. A 50 nanosecond NAMD molecular dynamics simulation of the CRISP model was produced. Temperature was controlled using Langevin dynamics. The representative glycan models used were complex (tetraantennary with terminal sialic acid residues) and high mannose (mannose-9) for human and mouse CRISP3, respectively.

The structure was superimposed with the PDB model of mouse CRISP3 (4MAK) and refined by energy minimization using the Schrodinger software suite. The model was also visualized using VMD software. The refined model was then used for the generation of the surface electrostatic potential.

Acknowledgments

We thank staff at the Monash Biomedical Proteomics Facility at Monash University, Australia, for support with proteomic analysis, and Monash IVF, Australia, for providing access to human seminal plasma samples. Homology modelling was carried out by Dr.
Michael Kuiper, Victorian Life Science Computation Initiative (VLSCI). This work was funded by NHMRC grant 606563 to MO’B and PMc.

**Author contributions**
M.V. performed the protein purification and glycosylation experiments. J.M. designed the protein purification strategy. P.M. and M.O. supervised the project. D.I. performed cDNA isolation, and both D.I. and R.D. provided technical assistance. M.V., J.M. and P.M. wrote the main manuscript text, and M.V. prepared all figures. All authors reviewed the manuscript.

**Additional information**

**Supplementary information** accompanies this paper at http://www.nature.com/scientificreports

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Volpert, M. et al. Eukaryotic expression, purification and structure/function analysis of native, recombinant CRISP3 from human and mouse. Sci. Rep. 4, 4217; DOI:10.1038/srep04217 (2014).