The 'omics revolution and our understanding of sperm cell biology

Mark A Baker

The foundations of proteomics are to study gene products and their regulatory roles within cells. Paradoxically, the only evidence that sperm cells make new proteins is through mitochondrial protein synthesis. Yet despite this, spermatozoa are the perfect candidates for mass spectrometry and hence, proteomic analysis. These enterprising cells use a plethora of post-translational modifications in order to gain functionality following their production within the testis. By using a combination of two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), and more recently liquid chromatography–mass spectrometry (LC–MS)/MS, recent advances in sperm cell biology, through the use of proteomics, is making unparalleled progress. The protein inventory lists being generated have shed light on transmembrane proteins, kinases and chaperones never previously recognized. In addition, the ability to isolate either phosphopeptides or glycopeptides and quantify the differences between cells of two different populations make proteomic analysis of spermatozoa a real chance to finally answer some age old questions.

Keywords: capacitation; DIGE; epididymis; LC–MS; proteomics; quantitation; sperm maturation; spermatozoa

SPERMATOZOA, A ‘SILENT’ ACHIEVER

Truth can often be found only after studying both extremes of an argument. The use of different 'omics approaches to study sperm cell biology, whether it be gene- or protein-based, will depend on the fundamental belief of when protein translation is, or is not, turned off. The thesis often put forward, is that sperm cells are transcriptionally silent. The antithesis held by others is that sperm cells are not only capable, but do require transcription for successful fertilization.1 In support of the former, it is recognized that following spermatogenesis, there is a twofold reduction in ploidy, which precedes chromatin remodeling through the sequential replacement of histones with transition proteins, and transition proteins with protamines.2–5 The latter packages DNA so tightly that it is unlikely to allow binding of transcription factors, leading to silencing of the DNA.6 Indeed experimentally, incubation of rabbit spermatozoa with radiolabeled leucine demonstrated no uptake, leading the authors to conclude that no protein synthesis was taking place within spermatozoa of this species.7 However, in what seems to be a contradiction, by adding a mixture of [3H]-amino acids to washed Balb/c mouse spermatozoa, incorporation of activity appeared to increase linearly over a 3-h time period.8 This activity was shown to be sensitive to the mitochondrial production of proteins with spermatid protein synthesis inhibitor chloramphenicol, but not the cytoplasmic protein synthesis inhibitor, cyclohexamine, suggesting mitochondrial involvement.9 Supporting this argument, autoradiographs of mouse sperm following incubation clearly showed that the radiolabeled amino acids localized to the sperm midpiece. Furthermore, SDS–polyacrylamide gel electrophoresis (PAGE) demonstrated that [3H]-amino acids were incorporated into full length proteins, and therefore, the activity was not a result of free amino acids storage with the sperm cells. The observation that spermatozoa produce new proteins using mitochondrial ribosomes has also been shown in other species. For example, Premkumar and Bharagava reported protein synthesis by bovine sperm, which again was inhibited by chloramphenicol and not cycloheximide.9 Furthermore, Mujica10 have shown evidence for human sperm protein biosynthesis and finally Gur and Breitbart have recently demonstrated incorporation of [35S]-methionine and [35S]-cysteine into newly synthesized proteins of sperm during ‘capacitation’ of either human, mouse, bovine or rat spermatozoa.11 In this latter report, incorporation of radioactivity was suppressed by chloramphenicol, tetracycline and gentamycin (all mitochondrial ribosomal inhibitors), but not cycloheximide. So in summary, a number of independent laboratories have clearly demonstrated mitochondrial production of proteins with spermatozoa of several species. The canonical view of the latter, is that mitochondrial DNA encodes for 13 proteins (seven from complex 1, one from complex 3, three from complex 4 and two from complex 5), all of which are known.12 The controversy at this present time, is not whether sperm mitochondrion produce new proteins (and theoretically, given that the mitochondrial DNA is not protonated, there appears to be little reason why either transcription or translation within the mitochondria does not continue as per a somatic cell), but which proteins are being produced. Gur and Breitbart have argued that the addition of chloramphenicol stops the production of 22 proteins,13 many of which have been shown by others to have a critical involvement in sperm capacitation, including the sperm-egg fusion receptor Izumo. What’s novel and controversial about these...
22 proteins, is that none are the 13 canonical ones derived from mitochondrial ribosomes. Making this concept even more challenging, it is difficult to conceptualize how some of the ‘new’ proteins being produced are handled by the cell. For example, how does the structural protein, A-kinase-anchoring protein 4 (AKAP4), which is normally insoluble, be transported through the mitochondria to the sperm flagella? Furthermore, if AKAP4 is produced, then why do others not see the radiolabeled signal within the sperm flagellum? Unfortunately, the authors failed to produce the raw data necessary to judge their interpretation. For these reasons, and the ability of pharmacology intervention, including kinase inhibitors to block sperm function,14–17 it is still the view of many today, that further maturation of the sperm cell, following elongation, is dominated through post-translational modifications. It is to this end, that proteomics, and not transcriptomics, stands alone as the dominant force to study sperm maturation.

THE MOST PROFOUND STATEMENTS ARE OFTEN SPOKEN IN SILENCE

Although the testicular environment develops a morphological sperm cell, paradoxically, this cell is incapable of movement, recognizing the egg and undergoing the acrosome reaction. Terminally differentiated and in an autonomous state, the male gamete can no longer rely on further protein biosynthesis as a means for maturation. Thus, the pivotal question raised to reproductive biologist is how to study a cell that essentially does not speak.

The different fields of ‘omics have indeed tried to answer this question. Remarkably, several laboratories have investigated the human sperm transcriptome, through both microarray-based experiments and serial analysis of gene expression analysis. Both have concluded that the mRNA species present within sperm must be either from past events; the gene expression of spermatozogenesis, or for providing the zygote with unique suite of paternal mRNAs.14–17 It is still the view of many today, that further maturation of the sperm cell, following elongation, is dominated through post-translational modifications. It is to this end, that proteomics, and not transcriptomics, stands alone as the dominant force to study sperm maturation.

THE SECRETS OF A SILENT CELL

Initial attempts to identify those proteins present within sperm began at the two-dimensional (2D)-PAGE level.20 The dynamic range of this technology lies between 10^3 and 10^6 orders of magnitude, depending on whether fluorescent dyes are used for protein detection. In practice, this means that even when using narrow pH ranges, only a small percentage of the proteome is visualized.21 Nonetheless, the earliest attempts at defining the human proteome found 1000 protein spots, of which, 98 unique proteins were identified by matrix-assisted laser desorption/ionization–time of flight (MALDI–ToF–ToF)–MS.22,23 Greater coverage of the sperm proteome was demonstrated in the human, when newer technologies, including 2D-liquid chromatography (LC), directly coupled to an ion-trap mass spectrometer were employed.24 Here, pooled human sperm ejaculates were purified using Percoll gradients, then subject to 1% Triton X-100 solubilization.24 This has the advantage that the soluble proteins would not be saturated with the higher abundant proteins including AKAP4, tubulin and outer dense fibers, which typically require a harsher lyses approach. Thus, the sample complexity and dynamic range is reduced.

Furthermore, the insoluble pellet was subject to SDS–PAGE extraction, then both the Triton-X 100 soluble and insoluble proteins were run on an SDS gel. Gel strips were taken every 2 mm. Due to the limitations in mass scanning range (typically around 50–6000 Da), peptides are often used for what is known as ‘bottom-up’ sequencing approach. For this cause, trypsin digestion of the gel strips is performed, and the peptides are then injected onto a reversed phase column and eluted directly into the MS. The latest state-of-the-art mass spectrometers can handle up to 10^5 orders of magnitude, again, well below 10^8 needed for a whole-system approach. Nevertheless, in the human, 1056 proteins were positively identified. The identification of these proteins presented some interesting facts, which either reinforced current concepts or introduced novel ideas into reproductive biology not previously considered. For example, 437 proteins were shown to belong to various metabolic pathways within spermatozoa. Not only does this demonstrate a huge reliance on glycolysis, but also these data suggest that glycolytic proteins in particular are amongst the highest expressed proteins with human spermatozoa. Before the sequencing of the human sperm proteome was undertaken, very little was known about the surface receptors with sperm cells. However, at least six 7-transmembrane protein receptors and six tyrosine kinase receptors were detected. One receptor that is of prolactin has recently been studied in further detail.25 This protein was found to be localized to the postacrosomal region of the sperm head, neck midpiece and principal piece of the tail. Interestingly, addition of prolactin was shown to be a prosurvival factor for human spermatozoa. Hence, in the presence of 25–1000 ng ml^-1 prolactin, up to 50% of human sperm cells continued to be motile in a basic salt solution, compared to only 5% of the vehicle control after 19 h.25 Interestingly, direct measurement of the level of prolactin in epididymal plasma recorded a level of 650 ng ml^-1,25 suggesting that prolactin plays a survival role within the epididymal lumen.

The amount of proteins identified involved in post-translational modification or protein (re)folding events was astonishing. Included in this inventory were heat shock proteins (HSPs) 10, 40, 70, 72, 75 and 90.26 Furthermore, HSP60 has been detected in other species.19,26,27 Why do sperm carry around so many HSPs within their proteome? And, as is the case for HSP70 and HSP90, why are they expressed at such high levels? The answers may be as intriguing as the question. Several HSPs, including HSP60, HSP90 and HSP70, undergo phosphorylation during capacitation.28–34 Recent evidence has shown that HSP60 and HSP90 (ERp99 and Grp94) migrate to the surface of the sperm head during capacitation, in concert with the ability of sperm to bind to the egg. In addition, although circumstantial, it was interesting to note that the only tyrosine-phosphorylated sperm heads were able to bind to the zona pellucida. Since the addition of anti-HSP60 or
-HSP90 antibodies did not inhibit sperm–zona pellucida binding, the authors have speculated that HSPs on the sperm surface are involved in formation of a complex, which is ultimately responsible for zona pellucida interaction.33

Apart from the human, other large scale studies utilizing LC–MS have been performed on rat mouse,26 bull36 and fruit fly sperm cells. An analysis of the ‘common’ proteins present within these species (common being found in at least two species) has been performed.37 From these lists, several classes of proteins/enzymes were found. For example, no less than 18 different kinases were found to be common between the human, rat and mouse sperm proteomes. One remarkable facet was the different isoforms of protein kinase A (PKA) present within sperm cells. This enzyme has been reported to play a role in just about every aspect of sperm function that has been studied.12,13,38–42. PKA is a tetramer, consisting of both regulatory and catalytic subunits. By using a molecular approach, deletion of the regulatory type IÎ± regulatory subunit delocalized PKA in the mouse sperm, but did not affect motility or fertility.43 Interestingly, sperm proteomic analyses have suggested that two other types of PKA-regulatory subunits are present, including RÎ±z and RÎ±ż. A haploinsufficiency of the latter gives rise to fertility defects in male mice and men.44 Thus, it has been suggested that there is some redundancy toward the catalytic subunit.40 Genetic approaches investigating the PKA-catalytic subunit lists have demonstrated the importance of this enzyme. By removing the PKAcÎ± gene (and abolishing both PKAcÎ±1- and PKAcÎ±2-catalytic subunits), a small percentage of knockout mice survived to adulthood.38 These spermatozoa had normal spermatogenesis, but defects in forward motility, with a near-total loss in progressive movement. By measuring enzyme activity, the authors could demonstrate that the PKAcÎ±1 knockout mice had little to no PKA activity in sperm, suggesting that PKAcÎ± was indeed the major kinase(s) present within these cells. To sharpen this approach, individual catalytic subunits of PKA have since been targeted. Deletion of the PKCÎ±II produced mice whose spermatozoa swim spontaneously in vitro, suggesting that this particular catalytic subunit has no role in the formation of a functional flagellum of the initiation of motility. However, the protein was necessary to speed the flagella beat, and for the tyrosine phosphorylation of proteins normally seen during capacitation.45 To my knowledge, no report exists on sperm PKAÎ±1 knockout mice; however, a chemical-genetic approach has been applied to sperm PKAc. Herein, the ATP-binding socket of the gene products (Î±Î± and Î±Î±1) are made susceptible to the chemical 4-amino-5-(4-methylphenyl)-7-(t-butyl)pyrazolo[3,4-d]-pyrimidine (PP1) through an M120A substitution. Addition of PP1 did not stop forward progressive motility per se; however, it did stop an increase in forward progressive motility usually seen upon addition of HCO3– (see Figure 3 of Ref. 46). This raises the question of whether either isoform of PKAc is responsible for the initiation of motility as suggested by the PKAc Î± knockout or could there be evidence for the involvement of another PKA isoform. The answer may lie in the proteomic analysis of the rat sperm cells.74 Herein, a non-redundant peptide hit (i.e., a peptide that only matches to one isoform), was found for the catalytic subunit of PKAb. Hence, it is tempting to speculate that while the PKAc Î± global knockout produced sperm that were essentially immotile, this is more likely a developmental defect, rather than the involvement of any of the PKAc Î± isoforms in motility activation per se. Hence, PKAb may be responsible for the initial onset of motility, whilst PKAc Î± then gives a ‘boost’ to the forward progressive motility.46 Finally, PKAcÎ±II appears to be responsible for the hyper-activated movement necessary for sperm fertilization.

**MAKING SPERMATOZOA SING**

The creation of proteomic lists within spermatozoa gives us a ‘snapshot’ or inventory of proteins present within the cell. However, functionally, it is still difficult to ascertain which of these proteins are, for example, modified during epididymal transit to allow sperm–egg binding, or initiate sperm motility. To answer these questions, comparative proteomic analysis has to be performed. In difference to simply identifying a protein, quantitative proteomic analysis requires a higher degree of expertise.

Initial comparisons of sperm proteomes were again performed at the level of 2D-PAGE, with 47 infertile patients with a cohort of fertile donors compared. From this study, 67 spots were found to undergo a change between the groups.34 However, the significant variation between the gels has led to criticism of the analysis.35 A second attempt at 2D-PAGE profiling comparing normal sperm with that of a patient who experienced in vitro fertilization failure found 20 differential spots, of which, actin-binding protein and outer dense fiber 2 could be identified.48 To overcome the inherent difficulties and reproducibility problems of 2D-PAGE, a less complex sample is desirable. Thus, 2D-PAGE applied to the nucleus of spermatozoa demonstrated 12 spots corresponding to protamines P1 and P2.49 A quantitative analysis of these spots demonstrated a higher P1 to P2 level in two of the 184 men undergoing male infertility diagnosis.50,51

A second way to overcome the inherent difficulties of 2D-PAGE is to use internally labeled fluorescent dyes, in the process now known as ‘difference in gel electrophoresis’ or DIGE. This technology allows three samples to be run simultaneously within a 2D gel. One sample is comprised of a pooled internal standard, and if run correctly, generates a standard for every protein spot detected, from which to accurately quantify protein levels. By comparing rat spermatozoa from the cauda and caput epididymis, 60 protein spots were shown to exhibit significant changes using the DIGE technology.52 Of these eight proteins were identified by MALDI–TOF–MS, including one protein undergoing serine phosphorylation as demonstrated by 2D western blotting.

The power of quantification is not limited to 2D-PAGE, but can also be applied to LC–MS technology, in perhaps a more versatile manner. Following tryptic digestion of proteins to produce peptides, several techniques are available to enrich for phosphopeptides, including the use of immobilized metal affinity chromatography. In a recent study, both non-capacitated and capacitated sperm digested were subject to phosphopeptide enrichment. Then, peptides were labeled using d0 and d3 methyl esters. This results in an increase in parent mass of the d3-labeled peptide by 3 Da. Thus, for example, if a peptide had an original mass of 1000 Da, the d3 peptide would remain at this mass, whilst the d0-labeled peptide increases to 1003 Da. The samples were then combined and run through reversed phase, nanoflow chromatography, coupled to an ion-trap Fourier transform instrument. One major advantage of such labeling strategies is that the retention time of the labeled peptides under reversed phase conditions does not change. Hence, two co-eluting peptides will emerge from the c-18 column, with a difference of 3 Da. By integrating the peptide counts, it is very easy to perform a quantitative measurement and locate kinase-/phosphatase-regulated proteins. Using this approach, the laboratory of Visconti found 55 unique phosphorylation sites, of which 42 were said to change during capacitation.53 From this study, it is clear that phosphoregulation plays a major role in capacitation. For example, the regulatory binding domain within AKAP, underwent major phosphorylation changes, such that the R-Helix involved in binding PKArII demonstrated charge–density shifts. This is likely to affect the binding movement of this regulatory domain to AKAP4.
The use of such 'label' strategy for sperm proteomics is not limited to this technology. Indeed, other labels can be used, albeit they work in slightly different ways. For instance, $^{18}$H$_2$O can be used to label a sample, which results in an increased mass of the parent peptide by either 2 or 4 Da. Labeling occurs in the presence of trypsin. Again, the samples can then be mixed (multiplexed) and run through reversed phase, without affecting the retention times. Heavy water labeling quantitation is also based upon the integrated area of the parent mass peptide. Other labels amenable to sperm cell biology include isotopically-coded affinity tag or isobaric tags for relative and absolute quantitation. The principal of increasing the mass of one peptide is the same; however, for both isotope-coded affinity tag and isobaric tags for relative and absolute quantitation, quantitation is performed at the MS/MS level and can only be used on high resolution, time of flight instruments including a Qq-ToF or MALDI-ToF-ToF.

The advantage of a global, quantitative proteomics approach is that the method is unbiased and capable of generating large amounts of meaningful data in a relatively short time frame. Coupled with techniques, including phosphopeptide enrichment through either immo-lized metal affinity chromatography of titanium dioxide$^{54,55}$ or indeed glycopeptides enrichment through the use of hydrazine,$^{56}$ LC–MS-based proteomic analysis is ensured to be a technique that will be foundational in answering biological question related to sperm cells and their ability to undergo post-testicular maturation for successful fertilization.

**COMPETING FINANCIAL INTERESTS**

The authors declare no competing financial interests.


54 Thingholm TE, Jensen ON, Robinson PJ, Larsen MR. SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. Mol Cell Proteomics 2008; 7: 661–71.
