This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
Proteomic characterization of the qualitative and quantitative differences in cervical mucus composition during the menstrual cycle


The chemical composition of the cervical mucus (CM), its physical characteristics and the volume secreted change cyclically throughout the menstrual cycle. Aim of this study was to identify the constitutive protein composition of CM of fertile women and the changes in the CM proteome throughout the menstrual cycle. Five fertile women who had a term delivery within 1 year before the study were enrolled. Proteomic analysis was performed by an Ultimate 3000 Nano/Micro-HPLC apparatus equipped with an FLM-3000-Flow manager module and coupled with an LTQ Orbitrap XL hybrid mass spectrometer; bioinformatic software were used for functional and quantitative analysis. 59, 81 and 43 proteins (mean) were respectively identified in the pre-ovulatory, ovulatory and post-ovulatory samples. 38 common proteins were identified. 42, 38 and 17 exclusive proteins were respectively identified in pre-ovulatory, ovulatory and post-ovulatory CM. The main part of CM constituents has a catalytic activity, which is mainly related to hydrolase activity. The label-free quantitative analysis for the common proteins revealed a significant reduction in the protein abundance index for antileukoproteinase, after the ovulation, and a peak of haptoglobin at ovulation. This is the first application of high-resolution MS-based proteomics for the identification of protein constituents of CM. This approach may contribute in identifying putative biomarkes of the female reproductive tract.

Introduction

The secretory cells of the glands in the cervical crypts produce cervical mucus (CM), that facilitates sperm penetration into the uterus during ovulation. The ovulatory CM provides in fact a defense for the sperm against the hostile acidic vaginal secretions, fulfills its energy demands, filters the ejaculate of the morphologically abnormal and dismotile sperm and provides conditions for sperm capacitation (1). Moreover the CM represents an inhibitory microenvironment for sperm propagation during the non-ovulatory phases of the cycle. However the function of cervical mucus is not yet fully understood.

The chemical composition of the CM is important for the fertility, since the addition of mucospissic agents and the presence of antisperm antibodies in human mid-cycle CM may transform CM from a sperm-receptive to a sperm-hostile element (2).
recognize the fertile phase for the purpose of fertility awareness (7).

Due to the very complex chemical and biophysical structure of cervical mucus, the available data are frequently incomplete and sometimes conflicting.

Mucins are the most abundant proteins near ovulation while the abundance of other proteins increases during the luteal phase (8). Wolf et al. proposed that the variable viscoelastic properties of cervical mucus are determined by its carbohydrate composition and/or structure. They also suggested that mucin concentration is dependent on variations in the hormonal milieu (9). Previous studies have shown the characteristics of mucin aggregates by means of atomic force microscopy and demonstrated the switch from fibrous preovulatory to globular ovulatory mucus (10).

In the postgenomic era, proteomic technology has rapidly developed becoming a powerful tool in the research of human physiology, in particular in biological fluid, characterizing the comprehensive proteomic composition and identifying potential novel biomarkers for diagnosis, prognosis and therapy in different clinical aspects, including reproduction (11).

A comprehensive human cervical mucus proteome catalog is still missing, even though the material is generally accepted as a biofluid. The proteome composition of cervical mucus is probably less complex when compared to plasma and urine but how many proteins cervical mucus contains and how the cervical mucus proteome undergoes to cyclical changes remain unclear.

Few studies have in fact investigated the proteic composition of cervical mucus. The presence of mucins and of anti-microbial peptides were previously reported in the cervical mucus plugs during pregnancy (12). More recently, 137 common proteins were identified in a pool of five cervical mucus plugs obtained from women in labor at term, by using LC–MS/MS (13).

Andersch-Björkman et al. have reported proteomic and glycomic analyses of the cervical mucus obtained from 12 non-pregnant women using LC-FT-ICR and MS/MS (14). They performed a pre-fractionation of the samples by one-dimensional SDS-PAGE and SDS-Agarose composite gel electrophoresis for the separation of proteins and mucins with subsequent tryptic analysis; major result was related to an interesting alternations in the mucin glycans. Moreover Panicker et al. applied SELDI-TOF MS in order to study CM composition (15).

The aim of this study was to investigate proteomic changes of CM at different phases of the menstrual cycle by an LTQ-Orbitrap XL mass spectrometer, implemented by bioinformatic tools for functional and quantitative analysis.

**Results and discussion**

In this study modern and more specific methodologies of high resolution mass spectrometry and bioinformatics analysis were utilized in order to determine the in vivo changes of CM proteic composition during the menstrual cycle.

The biochemical analysis of the intact CM is difficult because of its extremely high viscosity. To reduce viscosity cervical mucus was diluted in aqueous TFA solution. The acidic environment generated by the aqueous TFA, in addition, decreased protein degradation induced by proteinases after collection. This treatment facilitated the solubilization of small peptides that could otherwise co-precipitate with high molecular weight proteins, such as mucins, that were excluded by this procedure from the analysis. This approach can be adopted as the standard protocol for future studies on CM proteome.

Protein identification using stringent criteria lead to the characterization of a mean of 59 proteins in the pre-ovulatory samples, 81 proteins in the ovulatory samples and 43 proteins in the post-ovulatory samples.

The increase in the number of proteins at ovulation suggests that the major synthesis of proteins in CM is induced at the moment of ovulation and might be the consequence of the increase in estrogen levels.

Thirty-eight proteins were identified in all the 3 phases of the menstrual cycle. The complete list of these proteins, considered as constitutive of the CM in all the phases is reported in Table 1. For each protein we reported the following informations: UniPROT, Gene name, molecular weight (MW), isoelectric point (pI) and Description.

### Table 1: Constitutive CM proteins

<table>
<thead>
<tr>
<th>UniPROT</th>
<th>Gene name</th>
<th>MW [kDa]</th>
<th>calc. pI</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>P60709</td>
<td>ACTB</td>
<td>11.3</td>
<td>7.24</td>
<td>Actin</td>
</tr>
<tr>
<td>P04217</td>
<td>A1BG</td>
<td>69.3</td>
<td>6.28</td>
<td>Alpha-1B-glycoprotein</td>
</tr>
<tr>
<td>P03973</td>
<td>SLPI</td>
<td>15.2</td>
<td>8.68</td>
<td>Antileukoproteinase</td>
</tr>
<tr>
<td>P02647</td>
<td>APOA1</td>
<td>16.0</td>
<td>7.28</td>
<td>Apolipoprotein A-I</td>
</tr>
<tr>
<td>Q9UBG3</td>
<td>CRNN</td>
<td>11.6</td>
<td>5.87</td>
<td>Cornulin</td>
</tr>
<tr>
<td>P00738</td>
<td>HPT</td>
<td>16.0</td>
<td>8.05</td>
<td>Haptoglobin</td>
</tr>
<tr>
<td>P69905</td>
<td>HBA</td>
<td>13.2</td>
<td>6.13</td>
<td>Hemoglobin subunit alpha</td>
</tr>
<tr>
<td>P06871</td>
<td>HBB</td>
<td>11.2</td>
<td>7.24</td>
<td>Hemoglobin subunit beta</td>
</tr>
<tr>
<td>P02042</td>
<td>HBD</td>
<td>18.1</td>
<td>8.57</td>
<td>Hemoglobin subunit delta</td>
</tr>
<tr>
<td>P02790</td>
<td>HEMO</td>
<td>14.3</td>
<td>8.75</td>
<td>Hemopexin</td>
</tr>
<tr>
<td>P23527</td>
<td>H2B10</td>
<td>16.5</td>
<td>9.16</td>
<td>Histone H2B type I-O</td>
</tr>
<tr>
<td>P01876</td>
<td>IGHA1</td>
<td>37.6</td>
<td>6.51</td>
<td>Ig alpha-1 chain C region</td>
</tr>
<tr>
<td>P01877</td>
<td>IGHA2</td>
<td>77.0</td>
<td>7.12</td>
<td>Ig alpha-2 chain C region</td>
</tr>
<tr>
<td>P01857</td>
<td>IGHG1</td>
<td>36.5</td>
<td>6.10</td>
<td>Ig gamma-1 chain C region</td>
</tr>
<tr>
<td>P01859</td>
<td>IGHG2</td>
<td>13.9</td>
<td>10.32</td>
<td>Ig gamma-2 chain C region</td>
</tr>
<tr>
<td>P01860</td>
<td>IGHG3</td>
<td>36.1</td>
<td>8.19</td>
<td>Ig gamma-3 chain C region</td>
</tr>
<tr>
<td>P01861</td>
<td>IGHG4</td>
<td>10.2</td>
<td>6.99</td>
<td>Ig gamma-4 chain C region</td>
</tr>
<tr>
<td>P01834</td>
<td>IGKC</td>
<td>13.0</td>
<td>4.84</td>
<td>Ig kappa chain C region</td>
</tr>
<tr>
<td>P01620</td>
<td>KV302</td>
<td>35.9</td>
<td>7.59</td>
<td>Ig kappa chain V-III region SIE</td>
</tr>
<tr>
<td>P01717</td>
<td>LV403</td>
<td>15.0</td>
<td>8.27</td>
<td>Ig lambda chain V-IV region Hil</td>
</tr>
<tr>
<td>P0CG05</td>
<td>LAC2</td>
<td>5.0</td>
<td>5.06</td>
<td>Ig lambda-2 chain C regions</td>
</tr>
<tr>
<td>P0CG06</td>
<td>LAC3</td>
<td>22.6</td>
<td>8.91</td>
<td>Ig lambda-3 chain C regions</td>
</tr>
<tr>
<td>P01591</td>
<td>IGJ</td>
<td>78.1</td>
<td>8.12</td>
<td>Immunoglobulin J chain</td>
</tr>
<tr>
<td>O60282</td>
<td>KIF5C</td>
<td>41.3</td>
<td>7.90</td>
<td>Kinesin heavy chain</td>
</tr>
<tr>
<td>P05164</td>
<td>PERM</td>
<td>45.2</td>
<td>6.58</td>
<td>Myeloperoxidase</td>
</tr>
<tr>
<td>P02788</td>
<td>TRFL</td>
<td>35.9</td>
<td>7.36</td>
<td>Lactotransferrin</td>
</tr>
<tr>
<td>P61626</td>
<td>LYSC</td>
<td>41.7</td>
<td>5.48</td>
<td>Lysozyme C</td>
</tr>
<tr>
<td>P59665</td>
<td>DEF1</td>
<td>18.1</td>
<td>5.24</td>
<td>Neutrophil defensin 1</td>
</tr>
</tbody>
</table>
possess, in addition to the antiprotease activity, other properties published data demonstrating that WFDC domains may also
form four disulfide bonds. This motif is present in a number of including eight cysteine residues at defined positions, which
are released from the neutrophils (19).

Human defensins are small cationic antimicrobial polypeptides
and include the α-defensin human neutrophil peptides 1 to 3 (HNP1-3), that are released from the neutrophils (19).

WFDC motifs are composed of approximately 50 amino acids including eight cysteine residues at defined positions, which
form four disulfide bonds. This motif is present in a number of other secreted proteins, including elafin, eppin, and antileukoproteinase, which form the WFDC family of proteins. Initially, it was thought that the biological role of these proteins was to regulate inhibition of a wide spectrum of microbial and leucocyte proteolytic enzymes (20). There are at present several published data demonstrating that WFDC domains may also possess, in addition to the antiprotease activity, other properties such as antibacterial, anti-fungal, antiviral and anti-inflammatory functions (21).

Lactotransferrin, also called lactoferrin, is a mammalian iron scavenging defense protein, constitutively present in secretions that are consistently exposed to microbial flora: milk, tears, mucus.

gastrointestinal fluids, cervical mucus and seminal fluid. The main function of lactoferrin is to scavenge non-protein-bound iron in body fluids (22) and to suppress oxidative stress-mediated damage (23). The secretion of humoral soluble factors in cervical mucus such as lactoferrin, represents a first non immune defence involved in preventing ascending infections.

Further studies are needed in order to understand the interactions between immunoglobulins, defense and binding proteins and other CM proteins, which might open the perspective to the development of vaccines designed to protect the mucosal barriers from sexually transmitted diseases.

Comparative analysis for the identification of phase-specific proteins lead to the identification of 42 specific proteins in the pre-ovulatory phase, 38 specific proteins in the ovulatory phase and 17 specific proteins in the post-ovulatory phase (Table 2).

Table 2: Phase-specific proteins

<table>
<thead>
<tr>
<th>UniPROT</th>
<th>Gene name</th>
<th>MW [kDa]</th>
<th>pI</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q9Y217</td>
<td>FYY1</td>
<td>10.9</td>
<td>5.78</td>
<td>1-phosphatidylinositol 3-phosphate 5-kinase</td>
</tr>
<tr>
<td>P04083</td>
<td>ANXA1</td>
<td>15.8</td>
<td>8.75</td>
<td>Annexin A1</td>
</tr>
<tr>
<td>Q05994</td>
<td>AGR2</td>
<td>9.9</td>
<td>5.06</td>
<td>Anterior gradient protein 2 homolog</td>
</tr>
<tr>
<td>P02652</td>
<td>APOA2</td>
<td>15.9</td>
<td>5.76</td>
<td>Apolipoprotein A-II</td>
</tr>
<tr>
<td>P02749</td>
<td>APOH</td>
<td>15.9</td>
<td>5.76</td>
<td>Apolipoprotein A-II</td>
</tr>
<tr>
<td>P06022</td>
<td>DEFB1</td>
<td>11.2</td>
<td>6.62</td>
<td>Beta-defensin 1</td>
</tr>
<tr>
<td>Q07DL5</td>
<td>BPIB1</td>
<td>8.6</td>
<td>7.25</td>
<td>BPI fold-containing family B member 1</td>
</tr>
<tr>
<td>Q06P11</td>
<td>BRD7</td>
<td>29.8</td>
<td>7.97</td>
<td>Bromodomain-containing protein 7</td>
</tr>
<tr>
<td>Q0N03</td>
<td>CCL28</td>
<td>23.2</td>
<td>8.10</td>
<td>C-C motif chemokine 28</td>
</tr>
<tr>
<td>P08123</td>
<td>CO1A2</td>
<td>17.0</td>
<td>7.24</td>
<td>Collagen alpha-2(I) chain</td>
</tr>
<tr>
<td>P02748</td>
<td>CO9</td>
<td>20.0</td>
<td>9.00</td>
<td>Complement component C9</td>
</tr>
<tr>
<td>P08603</td>
<td>CFAH</td>
<td>21.0</td>
<td>7.53</td>
<td>Complement factor H</td>
</tr>
<tr>
<td>P01034</td>
<td>CYTC</td>
<td>14.6</td>
<td>10.04</td>
<td>Cystatin-C</td>
</tr>
<tr>
<td>Q8NCM8</td>
<td>DHHC2</td>
<td>38.2</td>
<td>6.95</td>
<td>Cytoplasmic dynein 2 light chain 1</td>
</tr>
<tr>
<td>P68104</td>
<td>EF1A1</td>
<td>19.0</td>
<td>7.84</td>
<td>Elongation factor 1-alpha 1</td>
</tr>
<tr>
<td>P15328</td>
<td>FOLR1</td>
<td>52.4</td>
<td>7.23</td>
<td>Folate receptor alpha</td>
</tr>
<tr>
<td>Q03880</td>
<td>LG3BP</td>
<td>18.7</td>
<td>9.09</td>
<td>Galectin-3-binding protein</td>
</tr>
<tr>
<td>P15076</td>
<td>LV302</td>
<td>65.3</td>
<td>5.27</td>
<td>Ig lambda chain V-III region LOI</td>
</tr>
<tr>
<td>Q2692</td>
<td>IBP4</td>
<td>27.9</td>
<td>7.15</td>
<td>Insulin-like growth factor-binding protein 4</td>
</tr>
<tr>
<td>Q08N6S</td>
<td>AN13C</td>
<td>37.1</td>
<td>6.84</td>
<td>Ankyrin repeat domain-containing protein 13C</td>
</tr>
<tr>
<td>Q9Y4D</td>
<td>DAAM1</td>
<td>31.2</td>
<td>6.43</td>
<td>Disheveled-associated activator of morphogenesis 1</td>
</tr>
<tr>
<td>Q16270</td>
<td>IBP7</td>
<td>32.3</td>
<td>6.65</td>
<td>Insulin-like growth factor-binding protein 7</td>
</tr>
<tr>
<td>Q03405</td>
<td>UPAR</td>
<td>14.3</td>
<td>10.20</td>
<td>Urokinase plasminogen activator surface receptor</td>
</tr>
<tr>
<td>P09609</td>
<td>CLUS</td>
<td>50.1</td>
<td>9.01</td>
<td>Clusterin</td>
</tr>
<tr>
<td>P09466</td>
<td>PAEP</td>
<td>38.3</td>
<td>7.97</td>
<td>Glycodelin</td>
</tr>
</tbody>
</table>
OVULATORY SPECIFIC PROTEINS

Q9XR5X HHLA3 11.9 5.08 HERV-II LTR-associating protein 3
Q99996 AKAP9 38.7 7.02 A-kinesin anchor protein 9
Q8N4C8 MINK1 28.8 8.12 Misshapen-like kinase 1
Q9Y6F6 MRVI1 63.1 5.59 Protein MRVI1
P30044 PRDX5 65.8 5.43 Peroxiredoxin-5
P02750 A2GL 57.0 6.87 Leucine-rich alpha-2-glycoprotein
Q38SD2 LRRK1 129.2 8.95 Leucine-rich repeat
Q75556 SG2A1 91.5 6.34 Mammaglobin-B
P10033 TIMP1 191.2 4.89 Metalloproteinase inhibitor 1
Q6UB82 MAP12 74.1 6.39 Methionine aminopeptidase 1D
Q96PE7 MCEE 122.2 7.23 Keratin, type II cytoskeletal 8
Q13232 NDK3 139.0 6.61 Nucleoside diphosphate kinase 3
P30086 PEBP1 143.5 7.12 Phosphatidylinositol-binding protein 1
P10023 A2MG 67.0 6.04 Alpha-2-macroglobulin
P18054 LOX12 75.6 6.21 Arachidonate 12-lipoxygenase, 12S-type
Q5TGY3 AHDC1 54.81 6.61 AT-hook DNA-binding motif-containing protein 1
Q5PSV4 BRM1L 67.5 4.79 Breast cancer metastasis-suppressor 1-like protein
Q86VK4 ZN410 44.2 6.57 Zinc finger protein 410
Q8WXX0 DYH7 57.6 7.36 Dynein heavy chain 7
Q9NYU1 UGGG2 237.0 6.70 UDP-glucose:glycoprotein glucosyltransferase 2

POST-OVULATORY SPECIFIC PROTEINS

Q13164 MK07 56.29 8.37 Mitogen-activated protein kinase 7
Q43148 MCES 108.46 5.43 mRNA cap guanine-N7 methyltransferase
Q13765 NACA 56.90 9.64 Nascent polypeptide-associated complex subunit alpha
Q94964 SOGA1 226.55 6.20 Protein SOGA1
O14795 UN13B 94.61 5.10 Protein unc-13 homolog B
Q9Y3P9 RBGPI 97.88 6.38 Rab GTPase-activating protein 1
P08240 SRPR 106.44 6.48 Signal recognition particle receptor subunit alpha
Q96F8X SLIK1 332.97 7.28 SLIT and NTRK-like protein 1
Q9H3E2 SNX25 243.79 6.89 Sorting nexin-25
Q96T0 TI2CD1 163.19 6.46 TBC1 domain family member 1
P23458 JAK1 113.90 9.38 Tyrosine-protein kinase JAK1
Q9NYU2 UGGG1 73.17 5.69 UDP-glucose:glycoprotein glucosyltransferase 1
Q92614 MY18A 133.00 6.96 Unconventional myosin-VIId
Q6QZQ6 WD8R7 162.93 7.68 WD repeat-containing protein 87
Q86UP3 ZFHX4 125.15 4.88 Zinc finger homeobox protein 4
Q9H412 ZHX3 253.27 6.60 Zinc fingers and homeoboxes protein 2
Q9H412 ZHX3 253.27 6.60 Zinc fingers and homeoboxes protein 3
Q6SF0 PAMD1 56.0 7.58 Atherin
Q8SWXX0 DYH7 57.6 7.36 Dynin heavy chain 7
Q9N5S6 TOPR5 112.3 9.41 E3 ubiquitin-protein ligase Topors
Q14576 ELAV3 39.5 9.28 ELAV-like protein 3
Q5VW36 FOCAD 199.9 6.62 Focaladhesin
Q48637 GSHB 40.3 5.11 Glutathione synthetase
Q6UXS9 CASPC 29.4 6.21 Inactive caspase-12
P01579 IFNG 19.3 9.47 Interferon gamma
P13645 K1C10 58.8 5.21 Keratin, type I cytoskeletal 10
Q5PSV4 BRM1L 67.5 4.79 Breast cancer metastasis-suppressor 1-like protein
Q08493 PDE4C 347.38 6.20 cAMP-specific 3',5'-cyclic phosphodiesterase 4C
Q81YX3 CC116 115.76 5.49 Coiled-coil domain-containing protein 116
Q8N4S0 CCD82 37.61 5.15 Coiled-coil domain-containing protein 82
P12109 C06A1 92.25 6.86 Collagen alpha-1(VI) chain
Q8WTW3 COG1 63.96 5.01 Conserved oligomeric Golgi complex subunit 1
P35462 DRD3 275.48 7.24 D(3) dopamine receptor
P32926 DSG3 40.73 8.54 Desmoglein-3
Q8TEC5 SH3R2 180.56 5.99 E3 ubiquitin-protein ligase SH3RF2
Q992K8 E2AK4 104.59 6.07 Eukaryotic translation initiation factor 2-alpha kinase 4
Q5XX13 FBB10 379.64 8.51 F-box/WD repeat-containing protein 10
P02751 FINC 238.73 6.68 Fibronectin
Q5CZC0 FSIP2 174.87 5.60 Fibrous sheath-interacting protein 2
P08107 HSP71 232.84 5.68 Heat shock 70 kDa protein 1A/1B
P42858 HD 107.47 5.00 Huntingtin
Q9BY13 HYCC1 66.52 9.04 Hyocin
Q8N315 INADL 131.63 6.71 InAD-like protein
P27816 MAP4 168.24 9.04 Microtubule-associated protein 4

These data seem apparently in contrast with the previously data reported by Andersch-Bjorkman et al (15). They reported in fact in 2007 that, when other proteins than mucins were separated by gel electrophoresis, the silver-stained protein profiles from different cycle days revealed a common pattern and no significant differences have been reported. Their major result were related to an interesting alterations in the mucin glycans. On the contrary, we preferred a different procedure excluding mucins from analytical plan by TFA treatment and mass spectrometer of similar sensitivity. Our experimental approach led us to obtain evidence on alterations of protein pattern over the different menstrual phases. We hypothesize...
that the main reason of this discrepancy may rely on the different sample treatment.

In the pre-ovulatory CM, 42 specific proteins were identified including cystatin C, clusterin, glycodelin, metalloproteinase inhibitor 1.

Cystatins are cysteine protease inhibitors, found in a wide range of organisms (24). Cystatin-C was identified for the first time in CM as a marker of the pre-ovulatory phase.

Clusterin is a protein previously identified in different secretions, including milk (25) and seminal plasma (26). It is related to the damaging oxidative reactions (27) and protein precipitation (28). Its role in reproduction is mainly related to the control of the agglutination of abnormal spermatozoa (29).

Glycodelin was reported to be involved in inhibiting sperm capacitation and function (30). The correlation between CM proteins, sperm function and reproduction represents an interesting field for future studies.

There are some literature reports about the inhibitors of metalloproteinases (TIMP) in reproduction, linking the expression of metalloproteinases-TIMPs system in the cervical mucus plug during pregnancy with the proteolytic activity in connection with term and pre-term birth (31). The significance of the exclusive presence in the pre-ovulatory CM of TIMP-1 might be associated to the inhibition of proteolytic activity, which leads to the liquefaction of CM in the ovulatory phase.

In the post-ovulatory phase 17 specific proteins were identified, including glycolytic enzymes such as triosephosphate isomerase (TPI) and structural proteins such as keratin I and II, spectrin and dynein chains which are essential for the maintenance and organization of intermediate filament network.

Sperm movement modifications are depending on external regulatory factors, some of which are secreted by the cervical glands of the cervix, by the oocyte or by the perioocyte layers to create a sperm chemotaxis which would optimize the gametic interaction (32). Previous data demonstrated that TPI is also localized at the acrosomal level in human spermatozoa (33) and that TPI is involved in the sperm-oocyte interaction. TPI is an enzyme that catalyzes the interconversion of dihydroxyacetone phosphatase and glyceraldehyde-3-phosphatase, completing the preparatory phases of glycolysis. TPI moreover contributes to cytoskeleton regulation interacting with Rho in the regulation of intracellular sodium through Na,K-ATPase activation, by providing glycolitic ATP that fuels energy for membrane function (34). TPI is associated with the plasma membrane and binds indirectly to structural proteins such as actin ad microtubules, which are involved in the microfibrillar structure of the CM, which constraints the flagellar wave amplitude of sperm.

Dynein chain might be associated with the exocytosis of the cervical granules, which could contribute in the changes in biochemical and microfibrillar pattern of CM.

Finally, we reported the presence of E3 ubiquitin-protein ligase SH3RF2 in the ovulatory CM and of the E3 ubiquitin-protein ligase Topors in the post-ovulatory phase. Ubiquitin is a 76-residue protein, involved in protein degradation by the ubiquitin-proteasome system. Ubiquitilation is important in processes such as apoptosis and antigen presentation, and represents a central process in oncogenesis, including the development of cervical neoplasia (35). Here we reported the first identification of enzymes involved in ubiquitinilation in CM that can be associated to the selection of the spermatozoa.

The bioinformatic analysis for the molecular functions of the common proteins, as reported in figure 1, revealed that the main part of CM constituents have a catalytic activity (36%), which is mainly related to hydrolase activity (60% of the catalytic enzymes involved in mucin breakdown as proteases, sialidases, other glycosidases and sulphatases. Their activities change mucin structure and hence its physical properties. These enzymes may influence the physical barrier to sperm cells and pathogens or may enhance bacterial adhesion and hence colonization (36) and might be moreover involved in enzymatic processes associated with sperm capacitation.

Binding proteins represent the second group for molecular activity classification (31%), mainly involved in protein-protein interactions (50% of the binding proteins).

The comparative analysis of GO annotations for molecular function for the exclusive proteins identified in pre-ovulatory and ovulatory CM reported significative differences in some important protein classes such as an increase of defense/immunity proteins at the ovulatory phase. This data suggests an associations between estrogens and the protective capacity of cervical mucus. This is consistent with previous studies demonstrating that the expression of factors reducing Candida Albicans binding to vaginal epithelial cells are most prominent in the glandular epithelium of the endocervix during estrus and is stimulated by estradiol treatment in mice (37).
This is the first study in which bioinformatic functional analysis of proteomic data offers additional information regarding the changes in CM functions during the menstrual cycle.

Finally, we applied for the first time a label free quantitative proteomic approach evaluating the differences in term of protein abundance for each identified protein between the different phases. This approach take in account the area of the three most abundant peptides, based on the premise that the MS signals of the most efficiently ionized peptides directly correlate with the corresponding protein amount. This method was originally proposed by Silva et al. (38) for Q-ToF instruments, but has been validated recently also for ion-trap based mass spectrometers running in data-dependent acquisition (DDA) mode (39).

The quantitative analysis for the common proteins revealed a reduction in the protein abundance index for antileukoproteinase, after the ovulation, and a peak of elastase and cathepsinG, contributing to local changes in CM functions during the menstrual cycle.

This is consistent with previous literature data, demonstrating that ALP concentration is significantly higher in the ovulatory phase than in the follicular and luteal ones (40). ALP, also called secretory leukocyte protease inhibitor (SLPI), is a mucosal secreted protein with a low molecular-weight, which inhibits the activities of the neutral serine lysosomal proteinases, elastase and cathepsin-G, contributing to local protection against epithelial cell destruction (41). Potential regulators of pig endometrial SLPI gene expression include the steroid hormones, estrogen (E) and progesterone (P4) (42). Interestingly, E and P4 regulation of SLPI synthesis appears to be specific to the endometrium. In humans, SLPI was demonstrated to be present in the CM during all the menstrual cycle, with a peak at ovulation, as confirmed by our quantitative proteomic analysis.

A peak of haptoglobin (HP) was verified at ovulation. HP is an acute phase alpha-sialoglycoprotein with hemoglobin-binding capacity. Hemoglobin-HP complexes are removed as a consequence of the binding with the CD163 receptor, which is expressed on the surface of monocytes and macrophages. HP also have been demonstrated to modulate several aspects of the innate and adaptive immune response and to have anti-inflammatory activities (43). The quantitative increase of haptoglobin at ovulation as reported by quantitative proteomic analysis is consistent with the observed increase in defence activity, which we observed by functional protein annotation.

Although our present results may provide a basis for a better understanding of the role of CM in reproduction, we must also acknowledge certain limitation of our experimental design. First, we performed our study on a relatively small sample scale. Second, the study was based solely on proteomic data, not confirmed by other techniques.

However this approach has permitted, by high resolution mass spectrometry, the detection of an array of proteins in cervical mucus reflecting both the constitutive composition and the changes during the menstrual cycle.

**Experimental**

The design of the study was approved by the Institutional Review Board of our University.

Five 25-30 year-old fertile women, with no history of fertility problems and who had a term delivery within 1 year before the study were enrolled. The subjects gave informed consent according to the guidelines of the Declaration of Helsinki.

The subjects maintained sexual abstinence during the study. Vaginal and cervical swabs were obtained before sample collection to exclude vaginal and cervical infection. PAP test and colposcopy were also performed before sample collection.

Cervical mucus samples (n=3 per patient) were obtained by gentle aspiration from the cervical canal with a catheter for intrauterine insemination (Gynecetics Medical Products, Achel, Belgium). Samples were collected before (day 7), during (day 12), and after (day 18) ovulation of the same menstrual cycle. Ovulation was assessed by transvaginal sonography and confirmed by measurement of midluteal serum progesterone levels.

The cervical mucus was collected in plastic tubes and mixed 1:1 (v/v) with aqueous trifluoroacetic acid solution (0.2% v/v) and centrifuged at 9200g for 10 minutes. The soluble acidic fraction was stored at -80°C until analysis.

An aliquot of the soluble acidic fraction of each CM sample, corresponding to 0.5 mg of total protein (as measured by Bradford assay), was mixed with 1 M ammonium bicarbonate pH 8.0 and reduced with 200 mM dithiothreitol (DTT 10 mM final, Sigma) for 5 minutes at 100°C and 15 minutes at 50°C, and then alkylated with 200 mM iodoacetamide (IAA 55 mM final, Sigma) in the dark at room temperature for 60 minutes. The samples were left to digest overnight at 37°C by adding 100 mM ammonium bicarbonate (pH 8) with sequencing grade modified porcine trypsin (1:50, trypsin: protein concentration, Pierce Biotechnology). To stop the digestion, the samples were acidified with aqueous trifluoroacetic acid solution (0.2% v/v) and immediately frozen and lyophilized.

The samples were resuspended in 40 µl of aqueous formic acid solution (0.1% v/v) and analyzed by an Ultimate 3000 Nano/Micro-HPLC apparatus (Dionex, Sunnyvale, CA, USA) equipped with an FLM-3000-Flow manager module, and coupled with an LTQ Orbitrap XL hybrid mass spectrometer (ThermoFisher, San Jose, CA). Separations were performed by a Zorbas 300SB-C18 column (3.5 µm particle diameter; column dimension 1mm i.d. x 15 cm) (Agilent Technologies, Santa Clara, CA) using the following eluents: (A) 0.05% (v/v) aqueous formic acid and (B) acetonitrile:water 80:20 with 0.05% (v/v) aqueous formic acid. The applied gradient was linear from 0 to 55% of solvent B in 40 min, at a flow rate of 80 µL/min. The LTQ-Orbitrap mass spectrometer was operated in a data-dependent mode in which each full MS scan (60000 resolving power) was followed by three MS/MS scans where the three most intense multiple-charged ions were dynamically selected and fragmented by collision-induced dissociation (CID) using a normalized collision energy of 35%.

Samples were analyzed individually; proteomic analysis was performed at the same time for all samples, while data analysis was subsequently performed.

Tandem mass spectra were analysed by the Thermo Proteome Discoverer 1.4 software, using SEQUEST cluster (University of Washington, Seattle, WA, licensed by Thermo Electron Corp) as search engine against UniProtKB/Swiss-Prot protein knowledgebase release 2013-02: 20247 Homo Sapiens protein database.
In order to obtain a reliable identification of the peptides, the following stringent filters were used: XCorr versus charge, 1.8 and 2.5 for doubly and triply charged ions respectively; high value peptide confidence; false discovery rate (FDR) <5%. Data were searched with three missed cleavages, fixed carbamidomethylation of cysteines and the oxidation of methionines as variable modification.

In order to identify a proteic pattern for cervical mucus we considered for this study the panel of proteins identified in at least one sample for each menstrual cycle phase (preovulatory, periovulatory and postovulatory) with an X score assigned by the software >1.5. Subsequently, in order to identify proteic markers for each cycle phase, we further considered the proteins exclusively identified in each ovulatory phase. Proteins were reported as “phase-specific” when they have been identified in the samples of one specific menstrual phase but they were not present or under the instrumental limit of sensitivity in the other phases, so that they are anyway more abundant and characterizing each of the menstrual phase.

Proteins identified by SEQUEST were analyzed using the publicly available PANTHER (protein annotation through evolutionary relationship) software (http://www.pantherdb.org/). For the aim of this study, we evaluated the GO annotations for molecular function in the common proteins. The differences between pre-ovulatory and ovulatory exclusive proteins and between ovulatory and post-ovulatory exclusive proteins were also evaluated according to GO classes for molecular function. Statistical significance was given for p <0.05.

The label-free quantification of common proteins was performed using the Peak Area Calculation Quantification during the bioinformatic analysis by Proteome Discoverer software. This quantification method is used to obtain an idea of the relative quantities of all peptides in a sample. The Proteome Discoverer application calculates peptide areas during processing, using them to automatically calculate protein areas for the proteins in the report. It calculates the area of any given protein as the average of the three most abundant distinct peptides identified for the protein. Statistical analysis was performed by the software MiniTab 1.7 (GMSL, Nerbiano (MI), Italia).

Conclusions

In summary our present proteomic approach reveals differences during the menstrual cycle in the expression of several proteins of the CM, that are involved in defense system, inflammation and structural activity. This work constitutes one of the few proteomic studies reported on CM and the first analysis of CM changes during the different phases of the menstrual cycle by high-resolution MS, implemented by quantitative tools.

Further characterization of proteins differentially expressed in CM may contribute to a better understanding of their role in reproduction, their effect on fertility and their biological regulation.

Notes

This journal is © The Royal Society of Chemistry 2012

Author contributions:

G. Grande substantially contributed to the design of the work, through analysis and interpretation of the data and drafting the work;
D. Milardi substantially contributed to the conception of the work and to the drafting of the work;
F. Vincenzoni substantially contributed to the design of the work, through sample proteomic analysis and drafting the work;
G. Pompe substantially contributed to the design of the work, through clinical evaluation and cervical mucus collection. She moreover revised critically the manuscript for important intellectual content.
A. Biscione substantially contributed to the design of the work, through cervical mucus collection and sample preparation. She moreover revised critically the manuscript for important intellectual content.
A.L. Astorri substantially contributed to the design of the work, through clinical evaluation and cervical mucus collection. She moreover revised critically the manuscript for important intellectual content.
E. Fruscella substantially contributed to the design of the work, through clinical and ultrasound evaluation of the patients. She moreover revised critically the manuscript for important intellectual content.
A. De Luca substantially contributed to the design of the work, through sample preparation. She moreover revised critically the manuscript for important intellectual content.
I. Messana substantially contributed to the conception of the work. She moreover revised critically the manuscript for important intellectual content.
M. Castagna substantially contributed to the design of the work, through clinical and ultrasound evaluation of the patients. She moreover revised critically the manuscript for important intellectual content.
R. Marana substantially contributed to the conception of the work. He moreover revised critically the manuscript for important intellectual content.

References:

4. M. Curlin and D. Bursac, Front Biosci Cervical mucus: from biochemical structure to clinical implications, 2013, 5,507;
5. R.J. Fehring, Contraception Accuracy of the peak day of cervical mucus as a biological marker of fertility, 2002, 66(4), 231;


17. K.M. Fahrbach, O. Malykhina, D.J. Steh and T.J. Hope, *PLoS One* Differe binding of IgG and IgA to mucus of the female reproductive tract, 2013, 8(10), e76176;


23. V. Turk, V. Stoka and D. Turk, *Front Biosci Cytostatin: biochemical and structural properties, and medical relevance*, 2008, 13, 5406;


27. S. Poon, S.B. Easterbrook-Smith, M.S. Rybchyn, J.A. Carver and M.R. Wilson, *Biochemistry Clusterin* is an ATP-independent chaperone with very broad substrate specificity that stabilizes stressed proteins in a folding-competent state, 2000, 39(51), 15953;


40. R.C. Thompson and K. Ohlsson, *Proc Natl Acad Sci USA* Isolation, properties, and complete amino acid sequence of human secretory leukocyte protease inhibitor, a potent inhibitor of leukocyte elastase, 1986, 83(18), 6692;

