Identification of Increased Amounts of Eppin Protein Complex Components in Sperm Cells of Diabetic and Obese Individuals by Difference Gel Electrophoresis

Uwe Paasch‡§¶, Falk Heidenreich§¶, Theresia Pursche¶, Eberhard Kuhlisch¶, Karina Kettner§, Sonja Grunewald‡, Jürgen Kratzsch**, Gunnar Dittmar¶¶, Hans-Jürgen Glander‡, Bernard Hoflack¶¶, and Thomas M. Kriegel§

Metabolic disorders like diabetes mellitus and obesity may compromise the fertility of men and women. To unveil disease-associated proteomic changes potentially affecting male fertility, the proteomes of sperm cells from type-1 diabetic, type-2 diabetic, non-diabetic obese and clinically healthy individuals were comparatively analyzed by difference gel electrophoresis. The adaptation of a general protein extraction procedure to the solubilization of proteins from sperm cells allowed for the resolution of 3187 fluorescent spots in the difference gel electrophoresis image of the master gel, which contained the entirety of solubilized sperm proteins. Comparison of the pathological and reference proteomes by applying an average abundance ratio setting of 1.6 and a \( p \leq 0.05 \) criterion resulted in the identification of 79 fluorescent spots containing proteins that were present at significantly changed levels in the sperm cells. Biometric evaluation of the fluorescence data followed by mass spectrometric protein identification revealed altered levels of 12, 71, and 13 protein species in the proteomes of the type-1 diabetic, type-2 diabetic, and non-diabetic obese patients, respectively, with considerably enhanced amounts of the same set of one molecular form of semenogelin-1, one form of clusterin, and two forms of lactotransferrin in each group of pathologic samples. Remarkably, \( \beta \)-galactosidase-1-like protein was the only protein that was detected at decreased levels in all three pathologic situations. The former three proteins are part of the eppin (epididymal proteinase inhibitor) protein complex, which is thought to fulfill fertilization-related functions, such as ejaculate sperm protection, motility regulation and gain of competence for acrosome reaction, whereas the putative role of the latter protein to function as a glycosyl hydrolase during sperm maturation remains to be explored at the protein/enzyme level. The strikingly similar differences detected in the three groups of pathological sperm proteomes reflect a disease-associated enhanced formation of predominantly proteolytically modified forms of three eppin protein complex components, possibly as a response to enduring hyperglycemia and enhanced oxidative stress. Molecular & Cellular Proteomics 10:10.1074/mcp.M110.007187, 1–14, 2011.

Male fertility is compromised by the hormonal and metabolic changes that are associated with type-1 and type-2 diabetes (1, 2), obesity (3), and the metabolic syndrome, the latter disturbance sharing essential pathologic features with the former diseases (4). The deleterious influence of diabetes and obesity on fertility is receiving increasing attention because their prevalence and incidence is escalating worldwide, whereas the age at first diagnosis of both diseases is continuously declining (5, 6). Because of this situation, the fertility of a growing number of individuals is affected before and during their reproductive years (7, 8). Recently, it was shown in a large cohort study that BMI affects fertility potential at the critical age for reproduction, whereas age had a more dominant effect on fertility potential than BMI when looking at patients of all ages (9). Despite a growing body of experimental data, the pathophysiological links between the above metabolic disorders and a reduced semen quality remain to be identified (10–12).

The prevalence of sexual dysfunction in diabetic men has been reported to approach 50% (13). In particular, decreased sperm concentration and motility, abnormal sperm morphology, increased seminal plasma abnormalities, and reduced serum testosterone levels because of impaired Leydig cell
function were detected (13). The hormonal situation is similar in obese males, showing low testosterone and sex hormone-binding globulin levels that are accompanied by elevated estrogen concentrations. Remarkably, the degree of obesity is correlated with reduced levels of inhibin B which, however, are not compensated by increases in follicle stimulating hormone (14). In case of massive obesity, reduced spermatogenesis associated with severe hypotestosteronemia may favor infertility as indicated by a lower total sperm count (12, 15, 16), a higher frequency of oligozoospermia (17) and a negative relationship between BMI and the total number of normal, motile sperm (11). Similar to the consequences of diabetes-associated hyperglycemia, obesity and dyslipidemia provoke increased oxidative stress and sperm DNA damage (4, 11).

In contrast to the existing knowledge on the above alterations of hormonal balance, sperm morphology and sperm overall function, the negative impact of diabetes and obesity on male fertility is poorly understood on a molecular level. Common phenomena of these diseases, however, may be explained by the unifying mechanism causing the vast majority of diabetic complications: excessive production of reactive oxygen species and limitation of the anti-oxidative defense capacity (13, 18, 19). To explore molecular details of sperm physiology and pathology, the experimental tool of comparative proteomics combining protein separation, limited proteolysis, and mass spectrometric peptide identification (20) has been introduced into the analysis of sperm cells (21–23) and epididymal fluid (24). In detail, comparative proteomics has been applied to study sperm maturation (25) and function (23, 26), to analyze pathologic situations in male reproduction (27, 28) and to identify protein targets for contraception (29). The analytical power of two-dimensional gel electrophoresis representing a widely used sperm protein separation method has been demonstrated by the resolution of 3872 single protein spots during the establishment of a high-resolution two-dimensional reference map of human spermatozoal proteins from fertile sperm bank donors (21).

The present study uses the difference gel electrophoresis (DIGE) approach (30) to comparatively analyze the proteomes of sperm cells from type-1 diabetic, type-2 diabetic, non-diabetic obese, and non-obese clinically healthy normozoospermic individuals (reference group). Based on direct labeling of sperm proteins with fluorescent cyanine dyes (CyDyes) before protein separation by isoelectric focusing and SDS-PAGE, the DIGE method allows multiple qualitative and quantitative comparisons when different fluorescent dyes are used to label the reference sample, an internal protein standard and the patients’ proteomes. In accordance with the outcome of a proof-of-principle study (31), this novel experimental approach is expected to function as a general non-invasive tool in the analysis of molecular mechanisms causing male infertility as well as in the identification of disease-associated marker proteins and in the evaluation of fertility-restoring treatment strategies.

**EXPERIMENTAL PROCEDURES**

**Study Design and Approval**—The present study was approved by the authorized ethics commissions at the University of Leipzig (216–2007) and at Dresden University of Technology (EK 36022006). Before sample collection, written informed consent was obtained from all participants. The overall experimental design has already been described elsewhere (31).

**Isolation of Sperm Cells**—Three ejaculates were collected from each of 21 normozoospermic and clinically healthy individuals (reference group), eight type-1 diabetic, seven type-2 diabetic, and 13 non-diabetic obese individuals with periods of sexual abstinence of three to seven days between subsequent donations. The ejaculates were collected by masturbation into sterile plastic dishes and allowed to liquefy for 30 min at 37 °C. Semen analysis was performed according to World Health Organization guidelines (32) by determining the following parameters: volume; sperm concentration, motility and morphology; pH; color; viscosity. To enrich motile sperm cells, ejaculates containing ≥20 × 10⁸ spermatozoa per milliliter were subjected to centrifugation employing an ISolate® two-layer discontinuous density gradient (Irvine Scientific, Santa Ana, CA). The reference group ejaculates subjected to comparative proteome analysis contained a fraction of at least 50% of grade (a+b) progressively motile cells (32), whereas no limitation regarding sperm motility was applied to the pathologic ejaculates. Discontinuous gradients were prepared from equal volumes (2 ml each) of 90% (v/v) lower-layer and 50% (v/v) upper-layer gradient solution, loaded with liquefied semen (2 ml) and centrifuged at 600 × g and 25 °C for 30 min using round-bottom plastic tubes. The cell pellet was re-suspended and washed three times in 3 ml washing buffer (10 mM Tris/HCl, 250 mM sucrose, pH 7.5), applying the centrifugation conditions stated above. Total cell numbers were determined using an improved Neubauer hemocytometer. The cells were finally resuspended in 50 µl lysis buffer (30 mM Tris/HCl, 7 M urea, 2 M thiourea, 4% (w/v) 3-[3-cholamidopropyl]dimethylammonio]propanesulfonate (CHAPS), 10 mM dithiothreitol (DTT), 1× protease inhibitor mix, pH 9.1), immediately transferred into liquid nitrogen and stored at −80 °C until further treatment.

**Cell Disruption and Protein Solubilization**—The frozen samples were thawed, diluted with lysis buffer to give a final concentration of 7–8 × 10⁹ cells per ml, and subsequently incubated at 25 °C for 1 hour with gentle shaking. Lysis was enhanced by nine cycles of sonication (10 s sonication, 10 s cooling) performed on ice using a UP 100H type sonicator (Dr. Hielscher, Teltow, Germany) at 20% output corresponding to 28 µm amplitude. Samples were re-collected after three cycles of sonication by centrifugation (16,000 × g, 4 °C, 30 s). Finally, insoluble material was removed by centrifugation for 10 min under the same conditions. The resulting supernatants containing the solubilized sperm proteins (subsequently referred to as sperm proteomes) were subjected to protein determination employing the RC/DC protein assay (Bio-Rad, Hercules, CA) and were further used for fluorescent labeling and DIGE analysis.

**Difference Gel Electrophoresis**—For multiplex matching of DIGE images, spot normalization and calculation of abundance changes, the internal standard was set up from equal amounts of protein extracted from both the reference and disease groups’ sperm cells. Internal standard, reference and pathological sample proteins were differentially labeled with N-hydroxysuccinimidyl ester derivatives of the cyanine dyes Cy2, Cy3, and Cy5, following the manufacturer’s instructions for minimal labeling (GE Healthcare, Munich, Germany). For analytical gels, 50 µg of protein was labeled with 200 pmol of fluorescent dye derivative in a total volume of 25 µl. Cy2 was exclusively used for labeling the internal standard that was run on each gel in parallel with two reference and/or sample proteomes. To avoid any labeling bias, Cy3 and Cy5 were randomly employed to label the reference and pathologic proteomes. Labeling was stopped by the
addition of 1 μl 10 mM L-lysine. Three differentially labeled proteomes were mixed, reduced with 50 mM DTT and supplemented with 0.75 μl immobilized pH gradient (IPG) buffer pH 3–10 (GE Healthcare, Munich, Germany). Sample rehydration solution (7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 50 mM DTT) was added to give a final volume of 150 μl. Isoelectric focusing was performed using an EttaN IPGPhor 3 unit (GE Healthcare, Munich, Germany). Samples were cup-loaded at the anodic side of the IPG strips (24 cm; pH 3–10, linear) that were previously rehydrated with 450 μl IPG buffer pH 3–10, 7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 1.2% (v/v) DeStreak reagent. For first-dimension separation performed at 20 °C, the following conditions were applied, which correspond to a total of 72 kVh: 150 V/3 h, linear gradient; 300 V/3 h, linear gradient; 1000 V/3 h, linear gradient; 10,000 V/3 h, linear gradient; 10,000 V/55,000 V/h, step. After isoelectric focusing, the strips were treated with equilibration buffer (50 mM Tris/HCl, 6 M urea, 20% (v/v) glycerol, 2% (w/v) SDS, 130 mM DTT, pH 8.8) for 20 min, followed by equilibration for another 20 min in a modified buffer containing 35 mM iodoacetamide for sulphydryl residue alkylation and lacking DTT. Second-dimension denaturing electrophoresis using 12.5% (v/v) SDS-polyacrylamide gels, 20.5 × 25.5 cm, was carried out at 5 mA per gel for the first hour and at 20 W per gel until completion in an Ettan Daltix system (GE Healthcare, Munich, Germany).

Image Acquisition and Data Analysis—The fluorescent two-dimensional gels (in total 74) were scanned with a Typhoon TRIO Variable Mode Imager (GE Healthcare, Munich, Germany) using the excitation/emission wavelengths of 488 nm/520 nm for Cy2, 532 nm/580 nm for Cy3 and 633 nm/670 nm for Cy5. Images (in total 221) were matched and normalized using the DeCyder 2D Software Version 7.0 (GE Healthcare, Munich, Germany). Equivalent protein spots in different gels were identified by employing the fully automated computer assisted alignment module (batch processor). The biological variation analysis program was applied to manually revise the matches. To identify disease-associated proteins, any spot exhibiting an average abundance ratio of ≤ −1.6 or ≥ 1.6 was considered. The latter parameter relates the average amount of any protein in the pathologic proteome to its average amount in the reference proteome (cf. legends to Tables II-IV). Data analysis employed a statistical model for continuous outcome variables in which residuals are considered normally distributed. Disease groups are described by a fixed factor. The variance between individuals and the variability within each individual are modeled as random effects and described by a sum of variance components. A linear mixed model was fitted to the data. The resulting F tests were performed using the SPSS for Windows 16.0.2 release (Chicago, IL, USA) and interpreted as significant when p-values ≤ 0.05 were obtained. Biometric spot evaluation included calculation of a measure which, because of its close relation to the ANOVA effect size parameter, is referred to as “Effect Size Measure” (ESM). ESM values were calculated according to equation (1)

\[
ESM = \frac{|\bar{x}_S - \bar{x}_R|}{s_{SR}} \quad (Eq. 1)
\]

where \(\bar{x}_S\) and \(\bar{x}_R\) stand for the mean log volume of any protein in the sample (S) and reference (R) proteome, respectively, whereas \(s_{SR}\) represents the standard deviation of the log volumes of the protein both in the sample and the reference proteomes. \(s_{SR}\) is calculated from log volume data according to equation (2)

\[
s_{SR} = \sqrt{\frac{s_S^2}{n_S} + \frac{s_R^2}{n_R}} \quad (Eq. 2)
\]

where the numbers of replicated analyses of protein and reference proteomes are indicated by \(n_S\) and \(n_R\), respectively.

Protein Identification—Gel discs containing sperm proteomes at pathologically increased or decreased level were excised from analytical gels after staining with Coomassie Brilliant Blue G-250 when the minimum amount of protein required for limited proteolysis and mass spectrometric peptide analysis was available. Preparative gels typically loaded with 500 μg of protein and stained with Coomassie Brilliant Blue G-250 were manually aligned with the fluorescent images to allow for a correct localization of disease-associated low-abundance protein spots and corresponding gel excision. The excised gel discs were washed twice with deionized water followed by one wash using 50% (v/v) acetonitrile in 25 mM ammonium bicarbonate solution, shrunk by dehydration in 100% acetonitrile and dried in a vacuum centrifuge at 240 × g and 25 °C for 30 min. The dried material was incubated overnight with 15 μl of sequencing grade trypsin (Promega, Mannheim, Germany) solution (5 μg/ml; 5 mM ammonium bicarbonate used as solvent) at pH 7 and 37 °C. For peptide extraction, the proteolyzed samples were mixed with 15 μl of 0.5% (v/v) trifluoroacetic acid in acetonitrile, sonicated for 5 min and finally incubated for another 5 min at 20 °C in a shaker at a frequency of 800 rpm. After centrifugation, supernatant removal and repetition of the extraction procedure, the pooled supernatants were vacuum-dried. The extracted peptides were dissolved in 5 μl 0.1% (v/v) trifluoroacetic acid in water, sonicated for 5 min and incubated at 20 °C for another 5 min. Aliquots of 0.5–1.0 μl of each sample were pipetted onto an AnchorChip™ target plate (Bruker Daltonics, Leipzig, Germany) and allowed to sit for 3 min at 20 °C. Mass spectrometric peptide analysis was performed at an Ultraflex™ Automated High-performance MALDI-TOF/TOF Mass Spectrometry System (Bruker Daltonics, Bremen, Germany) in reflection mode using α-cyano-4-hydroxycinnamic acid as a matrix. Peak lists were generated using the Bruker Daltonics flexAnalysis 2.2 software. Peptide mass fingerprint analysis employed the MASCOT software application 2.2 (Matrix Sciences, London, UK) using the following criteria for searching against the Swiss-Prot 2010.12 non-redundant protein database (523,151 sequences; 184,678,199 residues): taxonomy - human; mass accuracy - 50–90 ppm; fixed modifications - carbamidomethylation of cysteine; variable modifications - methionine oxidation and deamidation of asparagine and/or glutamine; maximum of missed cleavage sites - one. Known masses of contaminants tryptic keratin fragments and autodigestion products of trypsin were excluded from database search. Protein mass fingerprint scores higher than 56 corresponding to p < 0.05 were considered to indicate significant protein identification. For MALDI-TOF/TOF sequencing of trypsic peptides, the mass tolerance for fragment ions was set to 0.8 Da.

To identify low-abundance/low-score molecular species of semenogelin-1 and lactotransferrin, additional information was obtained by MS/MS sequencing. Peptides of corresponding spot proteins obtained by in-gel digestion using sequencing grade trypsin (Promega, Mannheim, Germany) were collected on a C18-column, desalted and prepared for mass spectrometry according to (33). The peptide mix was separated on an in-house packed 15 cm analytical reverse phase column of 75 μm inner diameter containing Reprosil-AQ Pur 3 μm C18-reverse phase beads (Dr. Maisch GmbH, Amberg-Entringen, Germany) by applying a 155 min gradient from 5 to 40% acetonitrile in 0.5% acetic acid using an Eksigent (Dublin, CA, USA) NanoLC HPLC system. For data collection, the effluent was directly electrosprayed into an LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) equipped with an electrospray ion source (Proxeon Biosystems, Odense, Denmark). Data analysis employed the MaxQuant (34) and Mascot (Matrix Science, version 3.2) software packages and the R statistical package (www.r-project.org). For database search, the IPI Human Database Version 3.43 (72,346 entries) was used. The search parameters were set to the default value for MaxQuant searches with fixed modific-
Sperm Eppin Protein Complex: Effect of Diabetes and Obesity

**Table I**

| Parameter Group | REF | DM1 | DM2 | OBE |
|-----------------|-----|-----|-----|-----|
| Age (years)     | 25.8 ± 5.6 (21) | 35.5 ± 10.2 (8)* | 55.9 ± 9.3 (7)* | 41.0 ± 9.8 (13)* |
| BMI (kg/m²)     | 22.5 ± 2.2 (19) | 24.9 ± 1.8 (8)* | 31.7 ± 3.4 (7)* | 34.0 ± 4.7 (13)* |
| Abdominal girth (cm) | 83.1 ± 7.9 (16) | 90.1 ± 3.4 (8)* | 110.9 ± 10.5 (7)* | 115.3 ± 13.8 (13)* |
| Total cholesterol (mmol/l; NR < 5.2) | 4.5 ± 1.1 (21) | 4.7 ± 0.9 (8) | 6.2 ± 3.5 (7) | 5.4 ± 1.2 (13)* |
| LDL cholesterol (mmol/l; NR < 4.1) | 2.7 ± 0.8 (21) | 2.7 ± 0.8 (8) | 2.6 ± 0.6 (6) | 3.1 ± 1.1 (13) |
| HDL cholesterol (mmol/l; NR > 1) | 1.4 ± 0.3 (21) | 1.2 ± 0.3 (8) | 1.1 ± 0.3 (6) | 1.1 ± 0.3 (13)* |
| Triglycerides (mmol/l; NR < 2.3) | 1.3 ± 0.6 (21) | 1.6 ± 1.1 (8) | 6.3 ± 9.6 (7)* | 3.0 ± 1.8 (13)* |
| HbA1c (%) | 4.9 ± 0.3 (20) | 6.6 ± 0.8 (8)* | 6.7 ± 1.4 (7)* | 5.4 ± 0.7 (13)* |
| Systolic blood pressure (mmHg; NR < 130) | 126.9 ± 12.1 (10) | 132.3 ± 10.2 (7) | 130.3 ± 5.1 (7) | 139.6 ± 11.6 (13)* |
| Diastolic blood pressure (mmHg; NR < 85) | 80.2 ± 3.4 (10) | 79.9 ± 5.4 (7) | 84.9 ± 5.4 (7)* | 88.8 ± 4.8 (13)* |
| Sperm concentration (NR) | 20.0 ± 10.0 (10) | 107.8 ± 99.4 (8) | 143.4 ± 74.4 (7) | 64.1 ± 53.2 (13)* |
| Sperm motility WHO (a+b) (NR > 50%) | 52.1 ± 4.4 (21) | 40.8 ± 9.5 (8)* | 41.4 ± 6.3 (7)* | 42.6 ± 8.2 (13)* |
| Sperm morphology (NR > 15%) | 10.8 ± 4.9 (21) | 4.0 ± 3.4 (8)* | 5.7 ± 3.8 (7) | 6.2 ± 3.4 (13)* |
| Sperm DNA fragmentation (%) | 7.5 ± 5.4 (14) | 19.8 ± 17.1 (8)* | 19.0 ± 19.9 (7) | 14.8 ± 13.3 (13) |

Values are expressed as means ± standard deviation. The number of examined individuals is given in parentheses. Data were analyzed by Student’s t-test. Significant differences (p < 0.05) between disease groups and reference group are indicated by an asterisk. NR, normal ratio.

**RESULTS**

Adequacy of Probands and Quality of Ejaculates—The diagnostic parameters of all reference group individuals and patients included in this study are compiled in Table I. All semen donors fulfilled the WHO criteria (32) for sperm concentration, whereas analysis of (a+b) grade progressive sperm motility and morphology according to strict criteria showed a reduction for sperm cells of type-1 diabetic, type-2 diabetic, and obese individuals. DNA-fragmentation was found increased in all three groups of pathologic samples with highest values observed in the type-1 and type-2 diabetic groups. Body mass index, abdominal girth, total cholesterol, LDL and HDL cholesterol, triacylglycerides and HbA1c support—or at least do not contradict—the diagnosed metabolic disorder.

Comparative Proteome Analysis—The application of the DIGE method as described above allowed for the resolution of 3187 fluorescent protein spots in the DIGE images (Fig. 1) compared with ~2000–2500 predicted human sperm proteins (22, 35). Fig. 1A shows the black/white image of the fluorescence pattern of the internal standard proteome obtained by scanning the master gel containing the entirety of sperm proteins expected to be present in the reference and pathological samples. Because of differences in sperm protein abundance and depending on the adjustment of brightness and brilliance, the master gel image may display a number of protein spots that only apparently deviates from the above numbers of identified fluorescent spots and predicted sperm proteins, respectively. Computer-assisted analysis of fluorescence intensities with an average abundance ratio setting of ≥1.6/≤−1.6 and application of a p ≤ 0.05 criterion resulted in the identification of a total of 79 fluorescent spots containing proteins that are present at increased or decreased levels in the proteomes of the diabetic and obese individuals (Fig. 1; supplemental Fig. S1). The outcome of the identification of disease-associated sperm proteins employing in-gel tryptic digestion, mass spectrometric peptide analysis and database search is summarized in Tables II–IV and in supplemental Table SI where the order of proteins is determined by their ESM calculated as described in the Experimental Procedures section.

In case of type-1 diabetes (Table II) and obesity (Table IV), a limited number of 12 and 13 different (forms of) sperm proteins that are present at altered levels were identified, respectively, whereas 71 different protein species were found to be associated with type-2 diabetes (Table III; Table SI). The entirety of observed proteomic changes is visualized in Fig. 1A, whereas the protein spots listed in Tables II–IV are marked in the corresponding DIGE images shown in Figs. 1B–1D. The bars and boundaries in these figures either label the visible protein spots or, in the case of low-abundance proteins, their respective positions in the two-dimensional gel. The dot within each boundary is indi-
cating the center of protein mass, which generally repre-
sents the optimal picking location. For clarity, the complete
set of proteins found to be associated with type-2 diabetes
is marked in the DIGE image shown in Fig. S1 and listed in
Table SI, whereas those proteins that were excluded from
further consideration in the present paper are omitted in
Table III and are not labeled in Fig. 1C. It should be noted
that six additional fluorescent gel spots were identified in
the DIGE images corresponding to the type-2 diabetic (four
spots) and obese (two spots) patients’ proteomes, however,
repeated attempts of mass spectrometric protein identifica-
tion remained unsuccessful.

Table II presents 12 protein species that were detected at
altered levels in the sperm proteomes of the type-1 diabetic
patients. Fluorescent spots containing proteins that occur at pathologically altered concentrations are labeled and annotated in red (type-1 diabetes), blue (type-2 diabetes), and yellow (obesity). Proteins that show very similar abundance changes in all three metabolic disorders are labeled and annotated in white, whereas proteins showing very similar abundance changes in two metabolic disorders are labeled and annotated in green. A, DIGE image of the internal standard comprising all reference and sample proteins (master gel). B–D: Illustration of proteomic changes for type-1 diabetic (B), type-2 diabetic (C) and non-diabetic obese (D) patients in the master gel image. Boundaries surround fluorescent spots exhibiting significantly increased (↑) or decreased (↓) protein volumes or, in case of low-abundance proteins, their respective position in the two-dimensional gel. The dot within each boundary is marking the center of protein mass, which generally represents the optimal picking location for protein identification. In Fig. 1C, the proteins listed in Table III are labeled, whereas the complete set of proteins associated with type-2 diabetes is available in the “Supplemental data” section. Identified proteins are annotated by their gene name according to the Swiss-Prot protein database and by their spot ID used in Tables II–IV and Table SI where full protein names are also given.

FIG. 1. Two-dimensional gel electrophoresis of CyDye-labeled human sperm proteomes (DIGE) and identification of diabetes- and obesity-associated sperm proteins. Fluorescent spots containing proteins that occur at pathologically altered concentrations are labeled and annotated in red (type-1 diabetes), blue (type-2 diabetes), and yellow (obesity). Proteins that show very similar abundance changes in all three metabolic disorders are labeled and annotated in white, whereas proteins showing very similar abundance changes in two metabolic disorders are labeled and annotated in green. A, DIGE image of the internal standard comprising all reference and sample proteins (master gel). B–D: Illustration of proteomic changes for type-1 diabetic (B), type-2 diabetic (C) and non-diabetic obese (D) patients in the master gel image. Boundaries surround fluorescent spots exhibiting significantly increased (↑) or decreased (↓) protein volumes or, in case of low-abundance proteins, their respective position in the two-dimensional gel. The dot within each boundary is marking the center of protein mass, which generally represents the optimal picking location for protein identification. In Fig. 1C, the proteins listed in Table III are labeled, whereas the complete set of proteins associated with type-2 diabetes is available in the “Supplemental data” section. Identified proteins are annotated by their gene name according to the Swiss-Prot protein database and by their spot ID used in Tables II–IV and Table SI where full protein names are also given.

that six additional fluorescent gel spots were identified in
the DIGE images corresponding to the type-2 diabetic (four
spots) and obese (two spots) patients’ proteomes, however,
repeated attempts of mass spectrometric protein identifica-
tion remained unsuccessful.

Table II presents 12 protein species that were detected at
altered levels in the sperm proteomes of the type-1 diabetic
TABLE II

Identification of human sperm proteins associated with type-1 diabetes by difference gel electrophoresis (DIGE)

The terms/abbreviations summarizing the outcome of DIGE analysis as illustrated in Fig. 1B are: Ser. No., numbering according to ESM value; Spot ID, spot number in master image; Protein ID and Gene, data taken from UniProtKB/Swiss-Prot database; Protein Identity, result of mass spectrometric protein identification; Appearance, number of DIGE images in which the respective spot can be matched to the master gel (total number of DIGE images); Average Abundance Ratio, ratio of the average amount of the respective protein in the pathological sample and in the reference proteome for proteins occurring at increased level, negative inverse ratio for proteins occurring at decreased level; p-value, result of ANOVA evaluation of data calculated by SPSS for Windows, Release 16.0.2 (SPSS Inc., Chicago, IL, USA); Effect Size Measure (ESM), biometric parameter calculated as described in Experimental Procedures section; Mr(cal) and Mr(obs), nominal molecular mass of carbamidomethylated protein and molecular mass corresponding to ordinate position in DIGE image, respectively /H20851 Dalton /H20852; pI(cal) and pI(obs), isoelectric point calculated from amino acid sequence and pI value corresponding to abscissa position in DIGE image, respectively; Score (PMF), significance of protein identification from peptide mass finger print (PMF) according to MASCOT software application 2.2 (Matrix Sciences, London, UK) (score value > 56 for p < 0.05); Matched Peptides and Sequence Coverage, number of MS-identified tryptic peptides of the respective protein (related to the total number of tryptic peptides and non-peptide substances resolved in the mass spectrum) and corresponding coverage of the authentic or from DNA translated amino acid sequence.

| Ser. No. | Spot ID | Protein ID | Gene | Protein Identity | Appearance | Average Abundance Ratio | p-value | Effect Size Measure | Mr(cal) | Mr(obs) | pI(cal) | pI(obs) | Score (PMF) | Matched Peptides | Sequence Coverage |
|----------|---------|------------|------|------------------|------------|-------------------------|---------|---------------------|---------|---------|---------|---------|-------------|-----------------|------------------|
| 1        | 1355    | CLUS_HUMAN | CLU  | Clusterin        | 71 (87)    | 1.79                    | 0.001   | 5.32                | 53,031  | 45,600  | 5.89    | 4.77    | 88          | 11 (36)         | 25%              |
| 2        | 1353    | CLUS_HUMAN | CLU  | Clusterin        | 79 (87)    | 2.07                    | 0.005   | 4.78                | 53,031  | 45,345  | 5.89    | 4.87    | 79          | 9 (33)          | 28%              |
| 3        | 247     | TRFL_HUMAN | LTF  | Lactotransferrin | 82 (87)    | 1.98                    | 0.012   | 4.46                | 80,014  | 78,244  | 8.50    | 7.78    | 191         | 25 (69)         | 35%              |
| 4        | 2543    | SPNXB_HUMAN| SPANXC| Sperm protein associated with the nucleus on the X chromosome B/F | 84 (87)    | 2.49                    | 0.019   | 4.10                | 11,876  | 17,449  | 5.92    | 4.77    | 124         | 12 (37)         | 81%              |
| 5        | 1596    | GGH_HUMAN  | GGH  | γ-Glutamyl hydrolase | 84 (87)    | 1.69                    | 0.008   | 3.97                | 36,340  | 39,467  | 6.67    | 7.32    | 128         | 12 (38)         | 33%              |
| 6        | 3042    | SEMG1_HUMAN| SEMG1| Semenogelin-1    | 56 (87)    | 1.72                    | 0.013   | 3.89                | 52,157  | 10,547  | 9.30    | 8.16    | 288         | 16 (68)         | 30%              |
| 7        | 244     | TRFL_HUMAN | LTF  | Lactotransferrin | 79 (87)    | 1.71                    | 0.035   | 3.65                | 80,014  | 78,134  | 8.50    | 7.34    | 133         | 15 (37)         | 23%              |
| 8        | 1945    | SEMG1_HUMAN| SEMG1| Semenogelin-1    | 82 (87)    | 1.6                    | 0.05    | 3.54                | 52,157  | 30,577  | 9.30    | 8.91    | 81          | 10 (47)         | 22%              |
| 9        | 1387    | CLUS_HUMAN | CLU  | Clusterin        | 66 (87)    | 2.06                    | 0.034   | 3.51                | 53,031  | 44,776  | 5.89    | 4.99    | 68          | 5 (20)          | 12%              |
| 10       | 318     | GLB1L_HUMAN| GLB1L| β-Galactosidase-1-like protein | 47 (87)    | 1.82                    | 0.006   | 2.87                | 74,395  | 75,973  | 9.01    | 8.76    | 130         | 19 (34)         | 34%              |
| 11       | 1306    | ZN813_HUMAN| ZNF813| Zinc finger protein 813 | 64 (87)    | 1.63                    | 0.018   | 2.23                | 73,841  | 46,570  | 9.50    | 4.95    | 76          | 18 (53)         | 33%              |

a Evidence at transcript level only.

b Identification involved MALDI-TOF/TOF sequencing of selected tryptic peptides.
### TABLE III
Identification of human sperm proteins associated with type-2 diabetes by difference gel electrophoresis (DIGE)

The terms/abbreviations summarizing the outcome of DIGE analysis as illustrated in Fig. 1C are: Ser. No., numbering according to ESM value; Spot ID, spot number in master image; Protein ID and Gene, data taken from UniProtKB/Swiss-Prot database; Protein Identity, result of mass spectrometric protein identification; Appearance, number of DIGE images in which the respective spot can be matched to the master gel (total number of DIGE images analyzed in this study); Average Abundance Ratio, ratio of the average amount of the respective protein in the pathological sample and in the reference proteome for proteins occurring at increased level, negative inverse ratio for proteins occurring at decreased level; p-value, result of ANOVA evaluation of data calculated by SPSS for Windows, Release 16.0.2 (SPSS Inc., Chicago, IL, USA); Effect Size Measure (ESM), biometric parameter calculated as described in **Experimental procedures** section; \( M_r(\text{cal}) \) and \( M_r(\text{obs}) \), nominal molecular mass of carbamidomethylated protein and molecular mass corresponding to ordinate position in DIGE image, respectively; \( pI(\text{cal}) \) and \( pI(\text{obs}) \), isoelectric point calculated from amino acid sequence and pI value corresponding to abscissa position in DIGE image, respectively; Score (PMF), significance of protein identification from peptide mass fingerprint (PMF) according to MASCOT software application 2.2 (Matrix Sciences, London, UK) (score value \( > 56 \) for \( p < 0.05 \)); Matched Peptides and Sequence Coverage, number of MS-identified tryptic peptides of the respective protein (related to the total number of tryptic peptides and non-peptide substances resolved in the mass spectrum) and corresponding coverage of the authentic or from DNA translated amino acid sequence.

| Ser. No. | Spot ID | Protein ID | Gene       | Protein Identity | Appearance | Average Abundance Ratio | p-value | Effect Size Measure | \( M_r(\text{cal}) \) | \( M_r(\text{obs}) \) | \( pI(\text{cal}) \) | \( pI(\text{obs}) \) | Score (PMF) | Matched Peptides | Sequence Coverage |
|----------|---------|------------|------------|------------------|------------|------------------------|---------|---------------------|----------------|----------------|----------------|----------------|---------------|------------------|------------------|
| 1        | 1355    | CLUS_HUMAN | CLU        | Clusterin        | 68 (84)    | 2.53 \(<0.0005\)       | 8.46    | 53,031              | 45,600         | 5.89           | 4.77           | 88b            | 11 (36)         | 25%              |
| 2        | 2466    | SEMG1_HUMAN| SEMG1      | Semenogelin-1    | 82 (84)    | 2.18 \(<0.0005\)       | 7.81    | 52,157              | 19,303         | 9.3            | 6.41           | 106             | 18 (38)         | 34%              |
| 3        | 2550    | SEMG1_HUMAN| SEMG1      | Semenogelin-1    | 79 (84)    | 2.38 \(<0.0005\)       | 7.49    | 52,157              | 17,134         | 9.3            | 6.56           | 117             | 22 (61)         | 29%              |
| 6        | 2854    | SEMG1_HUMAN| SEMG1      | Semenogelin-1    | 74 (84)    | 2.74 \(<0.0005\)       | 6.85    | 52,157              | 12,763         | 9.3            | 8.18           | 62              | 7 (32)          | 12%              |
| 14       | 3042    | SEMG1_HUMAN| SEMG1      | Semenogelin-1    | 59 (84)    | 2.34 \(<0.0005\)       | 5.79    | 52,157              | 10,547         | 9.3            | 8.16           | 288b            | 16 (68)         | 30%              |
| 16       | 2416    | PGK2_HUMAN | PGK2       | Phosphoglycerate kinase 2 | 76 (84) | 1.68 0.001 | 5.75 | 45,166 | 20,503 | 8.74 | 8.89 | 94b | 10 (89) | 22% |
| 17       | 1353    | CLUS_HUMAN | CLU        | Clusterin        | 76 (84) | 1.88 0.043 | 4.58 | 53,031 | 44,776 | 5.89 | 4.99 | 68b | 5 (20) | 12% |
| 30       | 1387    | CLUS_HUMAN | CLU        | Clusterin        | 67 (84) | 1.67 0.002 | 4.40 | 52,157 | 12,709 | 9.3 | 5.56 | 65 | 24 (89) | 53% |
| 31       | 2859    | SEMG1_HUMAN| SEMG1      | Semenogelin-1    | 74 (84)    | 2.00 0.016 | 4.39 | 52,157 | 20,388 | 9.3 | 8.84 | 68 | 21 (58) | 36% |
| 32       | 2406    | SEMG1_HUMAN| SEMG1      | Semenogelin-1    | 74 (84)    | 2.10 0.002 | 4.38 | 53,031 | 39,578 | 9.3 | 5.46 | 67 | 11 (64) | 29% |
| 33       | 1596    | CLUS_HUMAN | CLU        | Clusterin        | 74 (84)    | 2.14 0.004 | 4.27 | 52,157 | 11,234 | 9.3 | 6.08 | 157b | 15 (59) | 22% |
| 38       | 1282    | CLUS_HUMAN | CLU        | Clusterin        | 47 (84) | 1.7 0.027 | 4.24 | 53,031 | 47,493 | 9.3 | 8.46 | 67b | 7 (31) | 16% |
| 39       | 247     | TRFL_HUMAN | LTF        | Lactotransferrin  | 78 (84) | 1.93 0.031 | 4.20 | 80,014 | 78,244 | 8.5 | 7.78 | 191 | 25 (69) | 35% |
| 41       | 2782    | PIP_HUMAN  | PIP        | Progesterone-inducing protein | 82 (84) | 2.07 0.031 | 4.14 | 16,847 | 15,000 | 8.26 | 5.10 | 130 | 10 (48) | 63% |
| 42       | 1596    | GGH_HUMAN  | GGH        | \( \gamma \)-Gluamyl hydrolase | 81 (84) | 1.67 0.037 | 4.06 | 36,340 | 39,467 | 6.67 | 7.32 | 128 | 12 (38) | 33% |
| 43       | 1584    | CLUS_HUMAN | CLU        | Clusterin        | 78 (84) | 1.86 0.002 | 3.99 | 53,031 | 39,689 | 5.89 | 6.09 | 106 | 9 (29) | 22% |
| 44       | 2853    | SEMG1_HUMAN| SEMG1      | Semenogelin-1    | 82 (84) | 1.78 0.018 | 3.96 | 52,157 | 36,937 | 9.3 | 7.12 | 97 | 24 (69) | 40% |
| 48       | 3051    | SEMG1_HUMAN| SEMG1      | Semenogelin-1    | 82 (84) | 2.00 0.021 | 3.91 | 52,157 | 10,681 | 9.3 | 6.05 | 95 | 10 (62) | 15% |
| 50       | 2581    | SEMG1_HUMAN| SEMG1      | Semenogelin-1    | 82 (84) | 2.04 0.003 | 3.88 | 52,157 | 16,613 | 9.3 | 8.15 | 134 | 11 (27) | 20% |
| 55       | 244     | TRFL_HUMAN | LTF        | Lactotransferrin  | 77 (84) | 1.63 0.034 | 3.66 | 80,014 | 78,134 | 8.5 | 7.34 | 133 | 15 (37) | 23% |
| 56       | 2795    | PIP_HUMAN  | PIP        | Progesterone-inducing protein | 68 (84) | 2.11 0.006 | 3.82 | 16,847 | 13,594 | 8.26 | 4.82 | 134 | 9 (34) | 63% |
| 57       | 318     | GLB1L_HUMAN| GLB1L      | \( \beta \)-Galactosidase-1-like protein | 48 (84) | 1.89 0.004 | 3.58 | 74,395 | 75,973 | 9.01 | 8.76 | 130 | 19 (34) | 34% |
patients. In detail, increased concentrations of multiple molecular forms of clusterin, lactotransferrin, and semenogelin-1 were identified with one form of clusterin (spot ID 1355) exhibiting the highest ESM value, whereas γ-glutamyl hydrolase and β-galactosidase-1-like protein represent the only proteins that were found at reduced level. Limited two-dimensional electrophoretic separation was noticed for spot 1945 containing two different proteins: proteolytically modified semenogelin-1 and outer dense fiber protein 1. Interestingly, the sperm proteome of the obese individuals revealed remarkable similarity by exhibiting increased levels of multiple forms of the semenogelin-1 and lactotransferrin and of one molecular form of clusterin (Table IV). In addition, the β-galactosidase-1-like protein was found similarly reduced.

Compared with the alterations of the sperm proteome of type-1 diabetic and obese individuals, the proteomic changes found to be correlated with type-2 diabetes were more complex with 71 protein species being present at altered levels (Figs. 1 and S1; Tables III and S1). The latter data set coincides with the proteomic characteristics found to be associated with type-1 diabetes and obesity in that increased amounts of different molecular forms of semenogelin-1, clusterin, and lactotransferrin were detected, however, the number of identified protein species was substantially higher for semenogelin-1 (13 species) and clusterin (7 species), and one of the semenogelin-1 species was even decreased (Table III, spot ID 2647). Interestingly, the β-galactosidase-1-like protein was reduced in the type-2 diabetic sperm proteomes as already observed in the two previously described pathological situations.

Taken together, the outcome of the present study (Fig. 1, Tables II–IV) indicates a positive correlation between each of the three metabolic diseases and the concentrations of one semenogelin-1 (spot ID 3042), one clusterin (spot ID 1353) and two lactotransferrin (spot ID 244 and 247) species and a negative correlation with respect to the concentration of the β-galactosidase-1-like protein (spot ID 318) in the pathologic sperm proteomes. Apart from these findings, γ-glutamyl hydrolase (spot ID 1596) and outer dense fiber protein 1 (spot ID 1945 and 2416) occurring at decreased and increased proteomic level, respectively, appear to be specifically associated with type-1 and type-2 diabetes (Tables II and III), whereas prolactin inducible protein (spot ID 2782 and 2795) is found specifically increased in the sperm proteomes of type-2 diabetic and obese patients (Tables III and IV).

**DISCUSSION**

The purpose of the present study was the characterization of alterations of the sperm proteome that are associated with the impairment of male fertility as a consequence of diabetes (1, 2) and obesity (3). In addition to the consideration of disease-specific pathogenic causes, the generally negative impact of the above metabolic disorders on sperm quality has been largely explained by the unifying mechanism causing the
### Table IV

Identification of human sperm proteins associated with obesity by difference gel electrophoresis (DIGE)

The terms/abbreviations summarizing the outcome of DIGE analysis as illustrated in Fig. 1D are: Ser. No., numbering according to ESM value; Spot ID, spot number in master image; Protein ID and Gene, data taken from UniProtKB/Swiss-Prot database; Protein Identity, result of mass spectrometric protein identification; Appearance, number of DIGE images in which the respective spot can be matched to the master gel (total number of DIGE images); Average Abundance Ratio, ratio of the average amount of the respective protein in the pathological sample and in the reference proteome for proteins occurring at increased level, negative inverse ratio for proteins occurring at decreased level; p-value, result of ANOVA evaluation of data calculated by SPSS for Windows, Release 16.0.2 (SPSS Inc., Chicago, IL, USA); Effect Size Measure (ESM), biometric parameter calculated as described in Experimental procedures section; Mr(cal) and Mr(obs), nominal molecular mass of carbamidomethylated protein and molecular mass corresponding to ordinate position in DIGE image, respectively [Dalton]; pI(cal) and pI(obs), isoelectric point calculated from amino acid sequence and pI value corresponding to abscissa position in DIGE image, respectively; Score (PMF), significance of protein identification from peptide mass finger print (PMF) according to MASCOT software application 2.2 (Matrix Sciences, London, UK) (score value > 56 for p < 0.05); Matched Peptides and Sequence Coverage, number of MS-identified tryptic peptides of the respective protein (related to the total number of tryptic peptides and non-peptide substances resolved in the mass spectrum) and corresponding coverage of the authentic or from DNA translated amino acid sequence.

| Ser. No. | Spot ID | Protein ID | Gene | Protein Identity | Appearance | Average Abundance Ratio | p-value | Effect Size Measure | Mr(cal) | Mr(obs) | pI(cal) | pI(obs) | Score (PMF) | Matched Peptides | Sequence Coverage |
|----------|---------|------------|------|------------------|------------|-------------------------|---------|---------------------|---------|---------|---------|---------|-------------|-----------------|-----------------|
| 1        | 247     | TRFL_HUMAN | LTF  | Lactotransferrin  | 92 (102)   | 2.77                    | <0.0005 | 5.16                | 80,014 | 78,244 | 8.50   | 7.78   | 191         | 25 (69)         | 35%             |
| 2        | 250     | TRFL_HUMAN | LTF  | Lactotransferrin  | 91 (102)   | 2.58                    | 0.004   | 4.90                | 80,014 | 78,353 | 8.50   | 8.47   | 313         | 31 (53)         | 47%             |
| 3        | 242     | TRFL_HUMAN | LTF  | Lactotransferrin  | 93 (102)   | 2.64                    | 0.004   | 4.62                | 80,014 | 78,463 | 8.50   | 8.21   | 103         | 17 (35)         | 17%             |
| 4        | 1353    | CLUS_HUMAN | CLU  | Clusterin        | 94 (102)   | 1.79                    | 0.01    | 4.49                | 53,031 | 45,345 | 5.89   | 4.87   | 79          | 9 (33)          | 28%             |
| 5        | 244     | TRFL_HUMAN | LTF  | Lactotransferrin  | 92 (102)   | 1.96                    | 0.003   | 4.17                | 80,014 | 78,154 | 8.50   | 7.34   | 133         | 13 (57)         | 23%             |
| 6        | 3051    | SEMG1_HUMAN | SEMG1 | Semenogelin-1   | 98 (102)   | 2.15                    | 0.01    | 4.13                | 52,157 | 10,681 | 9.30   | 6.05   | 91          | 10 (62)         | 15%             |
| 7        | 2795    | PIP_HUMAN  | PIP  | Prolactin-inducible protein | 84 (102) | 1.68                    | 0.011   | 3.82                | 16,847 | 13,594 | 8.26   | 4.82   | 134         | 9 (34)          | 63%             |

9 318 GLB1L_HUMAN* GLB1L β-Galactosidase-1-like protein 53 (102) 1.83 0.003 3.62 74,395 75,973 9.01 8.76 130 19 (34) 34%

10 1235 ACTB_HUMAN ACTB Actin, cytoplasmic 1c | 95 (102) | 1.73 0.024 3.48 | 42,052 49,119 | 5.29 | 5.55 | 196 | 20 (44) | 49% |

ACTG_HUMAN ACTG1 Actin, cytoplasmic 2p | 95 (102) | 1.73 0.024 3.48 | 42,108 49,119 | 5.31 | 5.55 | 196 | 20 (44) | 49% |

11 271 TRFL_HUMAN LTF Lactotransferrin 95 (102) 1.75 0.017 3.48 | 80,014 77,371 | 8.50 | 8.51 | 430p | 38 (71) | 55% |

12 3091 SEMG1_HUMAN SEMG1 Semenogelin-1 71 (102) 2.06 0.008 3.21 | 52,157 10,183 | 9.30 | 6.52 | 128p | 16 (63) | 26% |

13 3042 SEMG1_HUMAN SEMG1 Semenogelin-1 64 (102) 1.73 0.005 3.16 | 52,157 10,547 | 9.30 | 8.16 | 288p | 16 (68) | 30% |

15 272 TRFL_HUMAN LTF Lactotransferrin 94 (102) 2.04 0.041 2.91 | 80,014 77,588 | 8.50 | 8.32 | 87 | 8 (15) | 13% |

*a* Evidence at transcript level only.

*b* Identification involved MALDI-TOF/TOF sequencing of selected tryptic peptides.

*c* Data not appropriate to distinguish between both protein species.

Spot proteins corresponding to ser. no. 8 (ID 1277) and 14 (ID 647): not identified.
vast majority of diabetic complications: excessive production of reactive oxygen species and limited anti-oxidative defense, both resulting in abnormally enhanced oxidative stress (13, 18, 36). Currently, one of the most promising experimental approaches to the identification of proteomic differences is the combination of protein separation, limited proteolysis and mass spectrometric peptide identification (20). In accordance with this strategy, the DIGE method (30) as specifically adapted to the analysis of limited amounts of sperm proteins (31) not only allows for multiple comparisons of two-dimen-
sional protein patterns, but also permits the detection of covalent protein modifications taking place in transcriptionally inactive (37) but translationally active (38) spermatozoa.

Based on this method, the prominent result of the present work is the detection of substantially increased amounts of the same set of covalently modified molecular forms of semenogelin-1, clusterin, and lactotransferrin in the sperm pro-
tomes of type-1 diabetic, type-2 diabetic, and obese indi-
viduals (Tables II–IV). This novel finding supports the assign-
ment of a central role to the only recently discovered eppin protein complex (EPC) consisting of apparently native forms of clusterin, lactotransferrin, and semenogelin-1, bound to the epididymal proteinase inhibitor eppin. The EPC is present in human seminal plasma and on the surface of human ejaculate spermatozoa where it contributes to ejaculate sperm protection (39–41), motility regulation (42) and acquire-
ment of capacitation-competence (43) and, hence, represents a novel target for male contraception (44, 45).

The subunit composition and the oligomeric structure of the EPC are still largely unknown (39). Basic interactions of the EPC components among each other and with the sperm surface, however, have been identified (39, 40, 46, 47) and are illustra-
ted in Fig. 2 without considering stoichiometric aspects such as eppin oligomer formation (39, 46). Eppin, a cysteine-
rich protein being encoded by the SPINLW1 gene (40, 46), is synthesized in the testis and secreted by epididymal epithelial cells. It coats the surface of human ejaculate sperm (48) by binding to hitherto unidentified receptor molecules. Clusterin (48, 49) and lactotransferrin (24, 50) are also of testicular and epididymal origin, whereas the Zn$^{2+}$-binding protein semenogelin-1 represents a major component of the seminal fluid (51–53). The formation of the sperm-bound EPC is likely to comprise the binding of testicular and/or epididymal eppin to the cell surface, binding of testicular and/or epididymal lactotransferrin and clusterin to surface-bound eppin and, finally, binding of semenogelin-1 to the eppin component of the surface-bound eppin-lactotransferrin-clusterin complex. Direct binding of clusterin to the surface of human sperm is considered unlikely, whereas speculations on lactotransferrin interaction with human sperm lactotransferrin receptors are currently experimentally addressed (39). The interaction of sperm-bound eppin with clusterin, lactotransferrin, and semenogelin-1 may finally result in the formation of a surface protein network exerting antimicrobial and antiproteolytic ac-
tivity as well as regulating sperm capacitation and motility (39, 43). The latter effect is accomplished in part by the activated serine proteinase PSA (prostate-specific antigen) present in the prostatic secretion which, after ejaculation, cleaves the semenogelin-1 component of the EPC. This proteolytic event enables spermatozoa to become motile, to gain capacitation competence and to prepare for acrosome reaction (43, 53, 55). Because PSA activity is initially inhibited by zinc ions which after ejaculation are bound by semenogelin-1 (56), the altered patterns of semenogelin-1 fragmentation observed in the present study might reflect disturbances of zinc homeo-
stasis known to be associated with several diseases including diabetes mellitus (56).

To evaluate the origin and potential pathogenic role of the observed covalent modifications of clusterin, lactotransferrin, and semenogelin-1 corresponding to spot ID 1353, 244/247, and 3042 according to Tables II–IV, the results of mass-
spectrometric protein identification have to be considered in detail. In the case of semenogelin-1, the large difference between calculated (52.157 kDa) and experimentally ob-
served molecular mass for spot protein 3042 (10.547 kDa) indicates extensive proteolytic degradation without excluding additional covalent modifications. The latter semenogelin-1 fragment was shown by peptide mass fingerprinting and par-
tial mass spectrometric sequencing of the tryptic peptides labeled in Fig. 3 to comprise amino acid residues 166–296 (Fig. 3). The corresponding calculated molecular mass and isoelectric point is 14.940 kDa and 9.21, respectively, com-
pared with 10.547 kDa and 8.16 as estimated from spot position 3042 in the master gel (Fig. 1; Tables II–IV). Peptide 234–251 carrying the unique cysteine residue 239 was con-
firmed by MALDI-TOF and LC-ESI-MS/MS analysis. Because cysteine-239 is critical for both eppin binding on the sperm surface (by forming an intermolecular disulfide bridge involving eppin residues 75–133) and inhibition of ejaculate sperm motility (40, 42), the observed pathologic accumulation of semenogelin-1 fragment 166–296 provokes speculations on the detected inhibition of sperm motility in the ejaculates of diabetic and obese patients (Table I).

The second diabetes- and obesity-associated sperm protein identified in the present study is a covalently modified form of clusterin. Native clusterin (also referred to as apolipoprotein J, APOJ) is a highly conserved disulfide-linked heterodimeric protein consisting of α and β chain. Clusterin functions as a chaperone-like lipid carrier and is believed to be involved in fertility-related processes such as the prevention of oxidative damage (57) and complement-mediated sperm lysis (58), agglutination of spermatozoa (60), membrane protection and membrane remodeling during sperm maturation (60). The differences between calculated and experimentally estimated molecular mass and isoelectric point (53.031 kDa/5.89 according to DIGE coordinates) suggest a disease-associated proteolytic degradation of native clusterin without excluding additional covalent modifications as already discussed for semenogelin-1. This consideration is in accordance with the sequence coverage by the set of tryptic peptides used for protein identification, which corresponds to a molecular mass of 45.098 kDa (data not shown). Accumulation of clusterin degradation products as a consequence of the respective metabolic disorder might compromise the structure, protective activity, and regulatory functions of the complex protein network suggested to exist on the sperm surface (39, 43) and, thereby, impair vital sperm functions. The clusterin-related findings reported here together with the observation that sperm cells from asthenozoospermic patients contain an increased amount of clusterin precursor protein (61) support the notion that clusterin plays a critical but ambivalent role during the transition of the ejaculated sperm being immobilized in the coagulum to the progressively motile and capacitation-competent state.

The third protein occurring in modified forms at increased levels in the sperm proteomes of diabetic and obese individuals is lactotransferrin. This protein contributes to the control of free radical formation and lipid peroxidation (63, 64), exerts antibacterial activity (which interestingly is abolished by the binding of advanced glycation end products formed as a consequence of hyperglycemia) (64–66) and modulates immune and inflammatory responses (67). Interestingly, expression of lactotransferrin and also of clusterin (69) and eppin (70) at least in rodents is dependent on estrogens (70). Comparison of calculated and experimentally observed molecular masses and isoelectric points of the lactotransferrin forms being present in protein spots 244 and 247 reveals minor deviations (Tables II–IV). Despite the critically low difference between the corresponding $M_c$(obs) values of 110 Da (Tables II–IV), the detection of two clearly separated protein spots in the DIGE image (Fig. 1) indicates the accumulation of at least one covalently modified form of this iron-chelating protein as a consequence of the respective metabolic disorder. LC-ESI-MS/MS analysis of both protein species revealed the same lack of the 72 N-terminal amino acids and the presence of the C terminus of native lactotransferrin (data not shown). The corresponding calculated molecular mass is 70.012 kDa which deviates by ~8 Da from the masses of spot proteins 244 and 247 as estimated from their ordinate position in the two-dimensional gel. The latter data might reflect the accumulation of two proteolytically modified lactotransferrin forms without explaining their different migration during isoelectric focusing, thus demonstrating the necessity of a comprehensive molecular characterization of these protein species. The limited information on lactotransferrin function(s) during spermatogenesis and ejaculate production and the uncertainty regarding the molecular identity of and differences between spot proteins 244 and 247 interdict speculations on functional consequences of a pathologic accumulation of covalently modified lactotransferrin species on sperm cells. It may be of interest to note, however, that there is evidence at the transcript level for the existence of a protein being highly similar to human lactotransferrin (UniProtKB identifier B7Z4X2_HUMAN for cDNA FLJ58679, highly similar to lactotransferrin (EC 3.4.21.–)). The latter protein is expected to lack the N-terminal 44 amino acids of human lactotransferrin whereas being identical with it in the remaining C-terminal part.

**Fig. 3. Primary structure of human semenogelin-1 (UniProtKB/Swiss-Prot identifier SEMG1_HUMAN, P04279).** Amino acids experimentally verified by peptide mass fingerprinting and by partial MS/MS sequencing of tryptic peptides obtained by in-gel digestion of spot protein 3042 after DIGE analysis are shown in bold red. Determination of molecular mass and isoelectric point employed the ExPASy Compute pi/Mw tool (http://expasy.org/).
The biometrical evaluation of the experimental data discussed above primarily considered ESM values that were determined for the entirety of 3187 protein spots detected in the three groups of pathologic samples. Based on these data, percentages were calculated by relating the number of protein spots exhibiting ESM values that are equal to or higher than the lowest ESM value of the individual proteins of the semenogelin-1/clusterin/lactotransferrin “quadruplet” (spots 3042, 1353, and 244/247) in each pathological situation to the total number of 3187 protein spots. The corresponding histograms in Fig. 4 illustrate the resulting values of 2.5% for type-1 diabetes, 9.7% for type-2 diabetes, and 6.9% for obesity, which justify paramount consideration of the aforementioned semenogelin-1, clusterin, and lactotransferrin species. The latter statement is limited by the numbers of reference individuals and patients included in this study that, for the identified protein quadruplet and for all three pathological situations, correspond to a power of ≥80% and a false discovery rate of ≤0.05. Improvement of these statistical parameters can only be achieved by increasing the number of independent replicates.

The overall outcome of the present study demonstrates the general adequacy of the established experimental approach for the non-invasive detection and identification of proteomic differences between sperm cells from various groups of patients, even when only a critically low amount of biological material is available for analysis. In particular, the proteomic data presented here suggest the involvement of semenogelin-1, clusterin, and lactotransferrin in the pathologic alterations of sperm morphology and functions taking place in diabetic and obese individuals. Considering the stimulation of reactive oxygen species production as a consequence of hyperglycemia-enhanced formation of advanced glycation end products (71), the observed accumulation of clusterin and lactotransferrin species may reflect the adequate cellular response aiming at the prevention of lipid/protein (per)oxidation and DNA fragmentation (72). In addition, the same glycation and oxidation reactions might alter the conformation of sperm proteins and, thereby, promote their proteolytic degradation, resulting in the disease-specific patterns of covalently modified forms of semenogelin-1 and clusterin described above. The apparent failure to detect altered amounts/forms of eppin in the patients’ sperm proteomes does not contradict the pathogenic role of the EPC as discussed here because 1) eppin turnover during spermatogenesis might not be affected by the above metabolic disorders, 2) isolation of the EPC from spermatozoa involves harsh chemical conditions including organic solvent extraction (39) that have not been applied to sperm protein solubilization before DIGE analysis in the present study, and 3) mass spectrometric eppin identification is complicated because eppin is not hydrolyzed by trypsin under the conditions routinely used to prepare samples for MS analysis (39). Evaluation of the remaining large number of sperm proteins listed in Tables II–IV and in Table SI but not considered here including a broad spectrum of additional molecular forms of semenogelin-1, clusterin and lactotransferrin requires additional comprehensive biometric and proteomic analyses that are beyond the scope of the present paper.

**SUMMARY**

Type-1 diabetes, type-2 diabetes, and obesity are accompanied by multiple changes of the sperm proteome.
Comparative proteomic studies employing difference gel electrophoresis (DIGE)\(^1\) indicated an accumulation of modified forms of the eppin (epididymal proteinase inhibitor) protein complex (EPC) components semenogelin-1, clusterin, and lactotransferrin in the sperm proteomes of the above disease groups. The established proteomic approach seems appropriate to identify disease-associated marker proteins and is expected to function as a general non-invasive tool in the analysis of molecular mechanisms causing male infertility as well as in the evaluation of fertility-restoring treatment strategies.

**Acknowledgments**—We thank Mrs. Gabi Kersten and Mrs. Heidrun Janus (University of Leipzig, Medical Faculty, Department of Dermatology, Training Center of the European Academy of Andrology) for technical assistance and secretarial help.

\(^1\) This work was supported by the German Research Council (Deutsche Forschungsgemeinschaft, DFG PA 834/7-1 and GL 199/4-4).

\(^2\) This article contains supplemental Fig. S1 and Table S1.

\(^3\) To whom correspondence should be addressed: University of Leipzig, Medical Faculty, Department of Dermatology, Training Center of the European Academy of Andrology, D-04103 Leipzig, Germany. Tel.: +49-341-9718611; Fax: +49-341-9718619; Email: uwe.paasch@medizin.uni-leipzig.de.

\(^4\) Both authors contributed equally to this work.

**REFERENCES**

1. Bener, A., Al-Ansari, A. A., Zirie, M., and Al-Hamaq, A. O. (2009) Is male fertility associated with type 2 diabetes mellitus? *Int. Urol. Nephrol.* **41**, 777–784

2. Malilidis, C., Agbaje, I. O’Neill, J., and McCuire, N. (2009) The influence of type 1 diabetes mellitus on spermatogenic gene expression. *Fertil. Steril.* **92**, 2085–2087

3. Pauli, E. M., Legro, R. S., Demers, L. M., Kunselman, A. R., Dodson, W. C., and Lee, P. A. (2008) Diminished paternity and gonadal function with type 1 diabetes mellitus on spermatogenic gene expression. *Fertil. Steril.* **90**, 346–351

4. Kasturi, S. S., Tannir, J., and Brannigan, R. E. (2008) The metabolic syndrome and male infertility. *J. Androl.* **29**, 251–259

5. Lavizzo-Mourey, R. (2007) Childhood obesity: what it means for physicians. *Curr. Opin. Chem. Biol.* **4**, 489–494

6. Li, L. W., Fan, L. C., Zhu, W. B., Nien, H. C., Sun, B. L., Luo, K. L., Liao, T. T., Telephoresis, and Lu, G. S. (2007) Establishment of a high-resolution 2-D reference map of human spermatozoal proteins from 12 fertile sperm-bank donors. *Asian J. Androl.* **9**, 321–329

7. Johnston, D. S., Wooters, J., Kopf, G. S., Qiu, Y., and Roberts, K. P. (2005) Analysis of the human sperm proteome. *Ann. N.Y. Acad. Sci.* **1061**, 190–202

8. Martinez-Heredia, J., Estanyol, J. M., Ballesca, J. L., and Oliva, R. (2006) Proteomic identification of human sperm proteins. *Proteomics* **6**, 4356–4369

9. Dacheux, L. J., Belghazi, M., Lanson, Y., and Dacheux, F. (2006) Human epididymal secretome and proteome. *Mol. Cell. Endocrinol.* **250**, 36–42

10. Baker, M. A., Witherden, R., Hetherington, L., Cunningham-Smith, K., and Attken, R. J. (2005) Identification of post-translational modifications that occur during sperm maturation using difference in two-dimensional gel electrophoresis. *Proteomics* **5**, 1003–1012

11. Naaby-Hansen, S., Mandal, A., Wolkowicz, M. J., Sen, B., Westbrook, V. A., Shetty, J., Coonrod, S. A., Klotz, L. L., Kim, Y. H., Bush, L. A., Flickinger, C. J., and Herr, J. C. (2002) CASYR, a novel calcium-binding tyrosine phosphorylation-regulated fibrous sheath protein involved in capacitation. *Dev. Biol.* **242**, 236–254

12. Turner, R. M., Musse, M. P., Mandal, A., Klotz, L., Jayes, F. C., Herr, J. C., Gerton, G. L., Moss, S. B., and Chemes, H. E. (2001) Molecular genetic analysis of two human sperm fibrous sheath proteins, AKAP4 and AKAP116, in men with dysplasia of the fibrous sheath. *J. Androl.* **22**, 302–315

13. Pixton, K. L., Deeks, E. D., Flesch, F. M., Moseley, F. L., Björnmdahl, L., Ashton, P. R., Barratt, C. L., and Brewis, I. A. (2004) Sperm proteome mapping of a patient who experienced failure fertilization at IVF reveals altered expression of at least 20 proteins compared with fertile donors: case report. *Hum. Reprod.* **19**, 1438–1447

14. Shetty, J., Diekmann, A. B., Jayes, F. C., Sherman, N. E., Naaby-Hansen, S., Flickinger, C. J., and Herr, J. C. (2001) Differential extraction and enrichment of human sperm surface proteins in a proteome: identification of immunoblot reactive candidates. *Electrophoresis* **22**, 3053–3066

15. Unlu, M., Morgan, M. E., and Minden, J. S. (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* **18**, 2071–2077

16. Kriegel, T. M., Heidenreich, F., Kettner, K., Pursche, T., Hoflack, B., Grunewald, S., Poenicke, K., Glander, H. J., and Paasch, U. (2009) Identification of diabetes- and obesity-associated proteomic changes in human spermatozoa by difference gel electrophoresis. *Reprod. Biomed. Online* **19**, 660–670

17. World Health Organization. (1999) Laboratory manual for the examination of human semen and sperm - cervical mucus interaction., 4th ed., Cambridge University Press, Cambridge.

18. de Godoy, L. M., Olsen, J. V., Cox, J., Nielsen, M. L., Hubner, N. C., Fröhlich, F., Walther, T. C., and Mann, M. (2008) Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* **455**, 1251–1254

19. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification...
rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26, 1367–1372
35. Baker, M. A., and Atiken, R. J. (2009) Proteomic insights into spermatozoa: critiques, comments and concerns. Expert. Rev. Proteomics. 6, 691–705
36. Sharma, R. K., Pasqualotto, F. F., Nelson, D. R., Thomas, A. J., Jr., and Agarwal, A. (1999) The reactive oxygen species-total antioxidant capacity score is a new measure of oxidative stress to predict male infertility. Hum. Reprod. 14, 2801–2807
37. Grunewald, S., Paasch, U., Glander, H. J., and Anderergg, U. (2005) Mature human spermatozoa do not transcribe novel RNA. Andrologia 37, 69–71
38. Gur, Y., and Breitbart, H. (2006) Mammalian sperm translate nuclear-encoded proteins by mitochondrial-type ribosomes. Genes Dev. 20, 411–416
39. Wang, Z., Widgren, E. E., Richardson, R. T., and O’Rand, M. G. (2007) Characterization of an epfin protein complex from human semen and spermatozoa. Biol. Reprod. 77, 476–484
40. Wang, Z., Widgren, E. E., Sivashanmugam, P., O’Rand, M. G., and Richardson, R. T. (2005) Assocation of epfin with semenogelin on human spermatozoa. Biol. Reprod. 72, 1064–1070
41. Bourgeon, F., Evrard, B., Brillard-Bourdet, M., Colleu, D., Je`gou, B., and Pineau, C. (2004) Involvement of semenogelin-derived peptides in the antibacterial activity of human seminal plasma. Biol. Reprod. 70, 768–774
42. Mitra, A., Richardson, R. T., and O’Rand, M. G. (2010) Analysis of recombinant human semenogelin as an inhibitor of human sperm motility. Biopol. Reprod. 82, 489–496
43. de Lamirande, E., Yoshida, K., Yoshiike, T. M., Iwamoto, T., and Gagnon, C. (2001) Semenogelin, the main protein of semen coagulum, inhibits human sperm capacitation by interfering with the superoxide anion generated during this process. J. Androl. 22, 672–679
44. O’Rand, M. G., Widgren, E. E., Wang, Z., and Richardson, R. T. (2007) Eppin: an epidymal protease inhibitor and a target for male contraception. Soc. Reprod. Fertil. Suppl. 63, 445–453
45. Karande, A. (2004) Eppin: a candidate male contraceptive vaccine? J. Biosci. 29, 373–374
46. Richardson, R. T., Sivashanmugam, P., Hall, S. H., Hamil, K. G., Moore, P. A., Ruben, S. M., French, F. S., and O’Rand, M. (2001) Cloning and sequencing of human Eppin: a novel family of protease inhibitors expressed in the epididymis and testis. Gene 270, 93–102
47. Wang, Z. J., Zhang, W., Feng, N. H., Song, N. H., Wu, H. F., and Sui, Y. G. (2008) Molecular mechanism of epidymal protease inhibitor modulating the liquefaction of human semen. Asian J. Androl. 10, 770–775
48. Sylvester, S. R., Morales, C., Oko, R., and Griswold, M. D. (1991) Localization of sulfated glycoprotein-2 (clusterin) on spermatozoa and in the reproductive tract of the male rat. Biol. Reprod. 45, 195–207
49. Sylvester, S. R., Skinner, M. K., and Griswold, M. D. (1984) A sulfated glycoprotein synthesized by Sertoli cells and by epidydimal cells is a component of the sperm membrane. Biol. Reprod. 31, 1087–1101
50. Goodman, A. M., and Young, L. G. (1981) Immunological identification of lactoferrin as a shared antigen on radioliodated human sperm surface and in radioliodated human seminal plasma. J. Reprod. Immunol. 3, 99–108
51. Jonsson, M., Frohm, B., and Malm, J. (2010) Binding of Semenogelin I to Intact Human Spermatozoo Studied by Flow Cytometry and Surface Plasmon Resonance. J. Androl. 31, 560–565
52. Yoshida, K., Kawano, N., Yoshikke, M., Yoshida, M., Iwamoto, T., and Morisawa, M. (2008) Physiological roles of semenogelin I and zinc in sperm motility and semen coagulation on ejaculation in humans. Mol. Hum. Reprod. 14, 151–156
53. Robert, M., and Gagnon, C. (1999) Semenogelin I: a coagulum forming, multifunctional seminal vesicle protein. Cell Mol. Life Sci. 55, 944–960
54. Zhang, J., Ding, X., Bian, Z., Xia, Y., Lu, C., Wang, S., Song, L., and Wang, X. (2010) The effect of anti-epfin antibodies on ionophore A23187-induced calcium influx and acrosome reaction of human spermatozoa. Hum. Reprod. 25, 29–36
55. Jonsson, M., Linse, S., Frohm, B., Lundwall, A., and Malm, J. (2005) Semenogelins I and II bind zinc and regulate the activity of prostate-specific antigen. Biochem. Biophys. Res. Commun. 327, 447–453
56. Jansen, J., Karges, W., and Rink, L. (2009) Zinc and diabetes–clinical links and molecular mechanisms. J. Nutr. Biochem. 20, 399–417
57. Trougakos, I. P., and Gonsos, E. S. (2002) Clusterin/apolipoprotein J in human aging and cancer. Int. J. Biochem. Cell Biol. 34, 1430–1448
58. Nagahara, A., and Tschopp, J. (1989) Molecular structure and functional characterization of a human complement cytolysis inhibitor found in blood and seminal plasma: identity to sulfated glycoprotein 2, a constituent of rat testis fluid. Proc. Natl. Acad. Sci. U.S.A. 86, 7123–7127
59. Carlsson, L., Ronquist, G., Nilsson, B. O., and Larsson, A. (2004) Dominant prostate specific immunogens for sperm-agglutinating autoantibodies of infertile men. J. Androl. 25, 699–705
60. Griffiths, G. S., Galileo, D. S., Arvindan, R. G., and Martin-DeLeon, P. A. (2009) Clusterin facilitates exchange of glycosyl phosphatidylinositol-linked SPAM1 between reproductive luminal fluids and mouse and human sperm membranes. Biol. Reprod. 81, 562–570
61. Martinez-Heredia, J., de Mateo, S., Vidal-Taboada, J. M., Ballescà, J. L., and Oliva, R. (2008) Identification of proteomic differences in asthenozoospermic sperm samples. Hum. Reprod. 23, 783–791
62. Sanocka, D., and Kurpisz, M. (2004) Reactive oxygen species and sperm cells. Reprod. Biol. Endocrinol. 2, 12
63. Makkere, K., Agarwal, A., and Sharma, R. (2009) Oxidative stress & male infertility. Indian J. Med. Res. 129, 562–570
64. Li, Y. M., Tan, A. X., and Vlassara, H. (1995) Antibacterial activity of lysozyme and lactoferrin is inhibited by binding of advanced glycation-modified proteins to a conserved motif. Nat. Med. 1, 1057–1061
65. Gonzalez-Chavez, S. A., Arevalo-Gallegos, S., and Rascon-Cruz, Q. (2009) Lactoferrin: structure, function and applications. Int. J. Antimicrob. Agents. 33, 301.e1–8
66. Li, Y. M. (1998) Glycation ligand binding motif in lactoferrin. Implications in diabetic infection. Adv. Exp. Med Biol. 443, 57–63
67. Legrand, D., Eliss, E., Carpenter, M., and Mazuirer, J. (2005) Lactoferrin: a modulator of immune and inflammatory responses. Cell Mol. Life Sci. 62, 2549–2559
68. Wünsche, W., Tenniswood, M. P., Schneider, M. R., and Vollmer, G. (1998) Estrogenic regulation of clusterin mRNA in normal and malignant endometrial tissue. Int. J. Cancer 76, 684–688
69. Allan, C. M., Couse, J. F., Simianen, U., Spaliviero, J., Jimenez, M., Rodriguez, K., Korach, K. S., and Handelsman, D. J. (2010) Estradiol induction of spermastogenesis is mediated via an estrogen receptor-(alpha) mechanism involving neuroendocrine activation of follicle-stimulating hormone secretion. Endocrinology 151, 2800–2810
70. Teng, C. T. (2010) Lactoferrin: the path from protein to gene. Biometals 23, 359–364
71. Malik, C., Agbaje, I. M., Rogers, D. A., Glenn, J. V., Atkinson, A. B., Steger, K., Stitt, A. W., and McClure, N. (2009) Advanced glycation end products accumulate in the reproductive tract of men with diabetes. Int. J. Androl. 32, 295–305
72. Mahfouz, R., Sharma, R., Thiyagarajan, A., Kale, V., Gupta, S., Sabanegh, E., and Agarwal, A. (2010) semen characteristics and sperm DNA fragmentation in infertile men with low and high levels of seminal reactive oxygen species. Fertil. Steril. 94, 2141–2146