Search for Cancer Markers from Endometrial Tissues Using Differentially Labeled Tags iTRAQ and cICAT with Multidimensional Liquid Chromatography and Tandem Mass Spectrometry


Department of Chemistry and Centre for Research in Mass Spectrometry, York University, Toronto, Ontario, Canada, Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario, Canada, Biochemistry, Toronto General Hospital, Toronto, Ontario, Canada, and Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada

Received October 4, 2004

A total of nine potential markers for endometrial cancer (EmCa) have been discovered and identified from endometrial tissue homogenates using a combination of differentially labeled tags, iTRAQ and cICAT, with multidimensional liquid chromatography and tandem mass spectrometry. The tissues were snap frozen in liquid nitrogen within 15–20 min after devitalization. Samples for proteomic analysis were treated with protease inhibitors before processing. Marker proteins that were overexpressed in EmCa are chaperonin 10, pyruvate kinase M1 or M2 isozyme, calgizzarin, heterogeneous nuclear ribonucleoprotein D0, macrophage migratory inhibitory factor, and polymeric immunoglobulin receptor precursor; those that were underexpressed are alpha-1-antitrypsin precursor, creatine kinase B, and transgelin. The chaperonin 10 result confirms our earlier observation of overexpression in EmCa tissues using surface-enhanced laser desorption/ionization mass spectrometry, verified by Western analysis and immunohistochemistry [Yang, E. C. C. et al. J. Proteome Res. 2004, 3, 636–643]. Pyruvate kinase was observed to be overexpressed using both iTRAQ and cICAT labeling. All nine markers have been found to be associated with various forms of cancer. A panel of these plus other markers may confer sufficient selectivity for diagnosing and screening of EmCa. The use of cICAT led to identification of a higher proportion of lower-abundance signaling proteins; conversely, iTRAQ resulted in a higher percentage of the more abundant ribosomal proteins and transcription factors.

Keywords: endometrial cancer markers • iTRAQ • cICAT • liquid chromatography • tandem mass spectrometry

Introduction

Endometrial carcinoma (EmCa) is a common malignancy among women in North America, with the fourth highest age-standardized incidence rate for female cancers in Canada as recently as 2003.¹ No accepted screening methods for this cancer or its precursor lesions exist. Current methods for definitive diagnosis require invasive tissue-sampling procedures, and are usually performed in women without EmCa, but who have presented with symptoms possibly stemming from this disease. No tumor markers are available for the clinical follow-up of known EmCa patients. Thus there is a real need for markers that will permit screening, early and accurate diagnosis, and disease monitoring. If differential protein expression between normal and malignant endometrial tissues could be identified through tissue proteomic analysis, then such differential expression could serve as a basis for the development of new serum or direct tissue testing.

The search for such protein markers in EmCa is a particular challenge due to inherent physiologic variation in protein expression levels as a result of normal menstrual cycling. This physiologic variation is unlike that in many other tissue types which have stable phenotypes. The endometrial cycle is broadly classified into two phases: the proliferative phase begins on the first day after menstruation (day 5) and ends at ovulation (day 14). During the proliferative phase and under increasing levels of estrogen, the tubular glands increase in tortuosity and length, and the stromal amount increases. Shortly after ovulation, the glands and stroma move through an orderly process of maturation in a secretory phase, prompted and supported by the appearance of progesterone. In the absence of any implanted fertilized ovum, menses occurs two weeks following ovulation.² Differences in protein expression between the proliferative and secretory phases have been noted.² The cyclic nature of the endometrium requires that any assessment of
relative abundances of individual proteins in malignant and normal endometria must include both phases of the normal
physiologic endometrium.

Our previous work using solid-phase extraction followed by matrix-assisted laser desorption/ionization mass spectrometry
(MALDI MS) as well as selective surface binding and surface-enhanced laser desorption/ionization (SELDI MS, Ciphergen
Biosystems Inc, CA) has succeeded in identifying individual markers that show significant enhanced expression in EmCa
tissues.5,6 The MALDI/SELDI MS strategy relies on side-by-side markers that show selective surface binding and surface-
matrix-assisted laser desorption/ionization mass spectrometry relative abundances of individual proteins in malignant and
research articles
DeSouza et al.

differentially expressed proteins involves differential tagging of proteins from samples that are being compared using isotopic dilution mass spectrometry experi-
mament.6 This strategy has recently been applied to discover differentially expressed proteins between the proliferative and
secretory phases of the human endometrium1 using a cleavable, second-generation ICAT reagent (cICAT, Applied Biosystems
Inc, CA). Recently, a variation of the ICAT technology, iTRAQ (also from the Applied Biosystems Inc, CA), has been intro-
duced. Both cICAT and iTRAQ tagging permit online identification of multiple markers and relative quantification of these
proteins. Although similar in their basic concepts, the tagging reagents and methodologies differ in significant areas.
The cICAT method relies on tagging cysteine residues and isolating peptides containing these tagged residues by affinity chromatography. The net result is a reduction in the complexity of peptide pools generated by digestion with proteases including trypsin.6 In the case of the new iTRAQ method, tagging is
on primary amines. This difference in labeling strategy elimi-
nates the dependence on relatively nonabundant cysteine containing peptides intrinsic to ICAT-based methods, thus
potentially allowing the tagging of most tryptic peptides. Other noteworthy features of the iTRAQ technology are that relative
quantification is performed via MS/MS and that there are four possible tags, which permit multiplexing of up to four samples
tissue (tissue states) in a single experiment. Quantification is performed via the differences in abundances of four product ions, 114, 115, 116, and 117 Th that are each cleaved from one of the four possible tags. The tags have an identical mass, a result of differences in other parts of the iTRAQ tag structure, with the consequence that an identical peptide in the four samples will have an identical mass and LC retention time after tagging. This strategy simplifies analysis and will potentially increase analytical accuracy and precision. The multi-sample capability of the iTRAQ technology is ideally suited for our study, as it now provides us with a means to perform a proteomic analysis of both the major phases of the normal endometrium, while simultaneously comparing them against cancer samples.

This report describes a feasibility study that compares protein expression profiles between normal and cancerous endometria using iTRAQ as well as contrasts it against a similar feasibility study using the cICAT technology. As this is one of the first studies using iTRAQ, contrasting the results obtained from both methods would be useful. It will be shown that both iTRAQ and cICAT labeling afforded discovery of a number of differentially expressed proteins that are potential cancer markers (PCMs). There is little overlap in the PCMs discovered

and identified, thus pointing to the complementary nature of the two technologies. It is noteworthy that application of the iTRAQ methodology permits us to confirm the overexpression of chaperonin 10 in EmCa tissues;4 chaperonin 10 is an approximately 10 kDa heat shock protein that does not contain the cysteine residue, which is the tagging site for cICAT. Thus, in addition to our primary interest in the discovery of markers for endometrial carcinoma, these experiments have also helped to illustrate the relative strengths and differences of the two tagging techniques, when applied to studies of clinical samples.

Materials and Methods

Sample Preparation. Endometrial tissue was retrieved from an in-house dedicated, research endometrial tissue bank. All
tissues were snap frozen in liquid nitrogen within 15–20 min of devitalization at the time of hysterectomy, and were obtained with patient consent. The patient consent forms and tissue-
banking procedures were approved by the Research Ethics Boards of York University, Mount Sinai Hospital, University
Health Network, and North York General Hospital. In each case, the endometrium was classified by a pathologist (TJC). Histological classification was verified by examination of a section from the frozen research tissue. Tissue for proteomic analysis was taken from the mirror-face of the residual block. In the case of samples used for iTRAQ analysis, 0.5 mL phosphate buffered saline (PBS) containing protease inhibitors (1mM AEBSF, 10 µM leupeptin, 1 µg/mL aprotinin and 1 µg/mL pep-
statin) was added. The tissue was then mechanically homogenized at 30 000 rpm using a Polytron PT 1300D handheld homogenizer (Brinkmann, Westbury, USA). For samples used for cICAT analysis, tissues were similarly homogenized in 1 mL Hanks’ Balanced Salt Solution with the same concentration of protease inhibitors as listed above. The samples were stored in aliquots at −70 °C until used for further processing. These whole tissue homogenates contained not only endometrial epithelium, but supportive stroma and vessels, as well as secretions. The iTRAQ analysis involved one normal proliferative, one normal secretory, and two cancer homogenates, while the cICAT analysis combined one normal proliferative homogenate with three different cancer homogenates in pairwise comparisons.

Chemicals. Reagent grade chemicals were purchased from Sigma Aldrich (Oakville, ON, Canada), or Fisher Scientific
(Nepean, ON, Canada). All iTRAQ and cICAT reagents and buffers were obtained from Applied Biosystems (Foster City, CA).

iTRAQ Sample Preparation Procedure. Cell debris from each of the homogenates was removed by centrifugation in a microfuge at 4 °C for 30 min at 14 000 rpm. The clarified supernatant was then transferred to fresh microfuge tubes and the total protein content determined using a commercial Bradford assay reagent (Bio-Rad, Mississauga, ON, Canada). A standard curve for the Bradford assay was made using γ-glob-
ulin as a control. 100 µg of each sample was then denatured and the cysteines blocked as described in the iTRAQ protocol
(Applied Biosystems, Foster City, CA). Each sample was then digested with 0.2 mL of a 50 µg/mL trypsin (Promega) solution at 37 °C overnight and labeled with the iTRAQ tags as follows: normal proliferative endometrium, iTRAQ114; normal secretory endometrium, iTRAQ115; and the two EmCa samples, iTRAQ116 and iTRAQ117. The labeled samples were then pooled and acidified by mixing with Eluent A (see later) to a total volume of 2.0 mL for strong cation exchange (SCX) chromatography.
1.5 mL of this acidified labeled sample was injected into an HP1050 LC system (Agilent, Palo Alto, CA) with a 1.5 mL injection loop and a 2.1 mm inner diameter (ID) × 10 mm length PolyLC Polysulfoethyl A column packed with 5 µm beads with 300 Å pores (The Nest Group, Southborough, MA). A 2.1 mm ID × 10 mm length guard column of the same material was plumbed upstream from the analytical column. Fractionation was effected by a binary mobile-phase gradient at a total flow rate of 0.2 mL/min. Eluent A consisted of a 10 mM KH₂PO₄ solution in 25% acetonitrile and 75% deionized water acidified to a pH of 3.0 with phosphoric acid. Eluent B consisted of a 10 mM KH₂PO₄ and 350 mM KCl solution in 25% acetonitrile and 75% deionized water acidified to a pH of 3.0 with phosphoric acid. Initially, the gradient comprised 100% Eluent A. At the 2nd minute, the % Eluent B was changed linearly from 0 to 100% at the 56th min. The run was terminated at the 60th min.

A total of 30 fractions were collected for the sample, one every two minutes, using an SF-2120 Super Fraction Collector (Advantec MFS, Dublin, CA). Following fractionation, the samples were dried by speed vacuuming and stored at −20 °C. Prior to reverse phase nanobore liquid chromatography-tandem mass spectrometric (nanoLC MS/MS) analysis, these fractions were redissolved in an aqueous buffer containing 5% acetonitrile and 0.1% formic acid for nanoLC MS/MS.

**cICAT Sample Preparation Procedure.** Samples were clarified and their total protein content determined as described above. cICAT sample preparation procedure was carried out according to the cleavable ICAT protocol (Applied Biosystems, Foster City, CA). Three 100 µg aliquots of the clarified normal proliferative homogenate were paired with 100 µg each of the three clarified cancer homogenates separately. After denaturing and reduction, the normal homogenates were labeled with the light reagent while the EmCa homogenates were labeled with the heavy reagent. A light labeled sample was then mixed with one of the heavy labeled samples to form in total three ICAT sample pairs: A, B, and C. The final volume of each sample was 0.2 mL. The sample pairs were then digested by incubating each pair with 0.2 mL of a 50 µg/mL trypsin solution at 37 °C overnight. Afterward each sample pair was mixed with 2.0 mL of Eluent A and fractionated into 30 fractions using SCX chromatography as described above. The fractions were screened by MALDI MS analysis. Those showing the signature ICAT peak pairs separated by 9 Da were further processed by affinity purification using the avidin cartridge provided with the ICAT kit. The affinity-purified sample was then dried, treated with the cleavage reagent to eliminate biotin, and dried again by speed vacuuming. The resulting solids were redissolved in an aqueous buffer containing 5% acetonitrile and 0.1% formic acid for nanoLC MS/MS.

**Nanobore LC MS/MS.** The nanobore LC system was from LC Packings (Amsterdam, The Netherlands) and consisted of a Famos autosampler and an Ultimate Nano LC system. It was interfaced to an API QSTAR Pulsar QqTOF mass spectrometer (Applied Biosystems/MDS Sciex, Foster City, CA) using a Protana NanoES ion source (Protana Engineering A/S, Odense, Denmark). The spray capillary was a PicoTip SilicaTip emitter with a 10 µm ID tip (New Objective, Woburn, MA). The nanobore LC column was a 75 µm ID × 150 mm length reversed-phase PepMap C18 nano capillary column (LC Packings, Amsterdam, The Netherlands) packed with 3 µm beads with 100 Å pores. One µL of sample was injected via the full-loop mode. Separation was performed using a binary mobile-phase gradient at a total flow rate of 200 nL/min. Eluent A consisted of 94.9% deionized water, 5.0% acetonitrile and 0.1% formic acid (pH ≈ 3). Eluent B consisted of 5.0% deionized water, 94.9% acetonitrile and 0.1% formic acid. The following binary gradient was used for the iTRAQ experiments:

<table>
<thead>
<tr>
<th>time (min)</th>
<th>0</th>
<th>5</th>
<th>10</th>
<th>125</th>
<th>135</th>
<th>155</th>
<th>160</th>
<th>162</th>
<th>188</th>
</tr>
</thead>
<tbody>
<tr>
<td>% eluent A</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>35</td>
<td>60</td>
<td>80</td>
<td>80</td>
<td>5</td>
<td>stop</td>
</tr>
</tbody>
</table>

While the binary gradient used for cICAT experiments was:

<table>
<thead>
<tr>
<th>time (min)</th>
<th>0</th>
<th>5</th>
<th>125</th>
<th>135</th>
<th>157</th>
<th>160</th>
<th>163</th>
<th>190</th>
</tr>
</thead>
<tbody>
<tr>
<td>% eluent B</td>
<td>5</td>
<td>5</td>
<td>30</td>
<td>60</td>
<td>80</td>
<td>80</td>
<td>5</td>
<td>stop</td>
</tr>
</tbody>
</table>

For nanospray, the source conditions were a curtain-gas setting of 20 and an ionspray voltage in the range of 1800−3000 V that was optimized daily. In the Q0 region, the instrument parameters were a declustering potential (DP) of 65 V and a focusing potential (FP) of 265 V. Nitrogen was used as the collision gas at a setting of CAD = 5 for both TOF MS and MS/MS scans. All nanoLC MS/MS data were acquired in information-dependent acquisition (IDA) mode in Analyst QS SP8 with Bioanalyist Extension 1.1 (Applied Biosystems/MDS Sciex). We performed two sets of runs for the iTRAQ fractions. For the first set, MS cycles comprised a TOF MS survey scan with an m/z range of 400−1500 Th for 2 s, followed by three product ion scans with an m/z range of 70−2000 Th for 10 s each. For the second, we used a 1 s TOF MS survey scan, followed by three product ion scans of 3 s each. For cICAT experiments, the MS cycles consisted of a TOF MS survey scan for 1 s, followed by two product ion scans of 2 s each. The ranges for the TOF MS and product ion scans were the same as those of the iTRAQ experiments. Collision energy (CE) was automatically controlled by the IDA CE Parameters script. Switching criteria were set to ions greater than m/z = 400 Th and smaller than m/z = 1500 Th with a charge state of 2 to 5 and an abundance of ≥10 counts/s. Former target ions were excluded for 60 s and ions within a 4 Th window were ignored. In addition, for iTRAQ runs, the IDA Extensions II script was set to no repetitions before dynamic exclusion for the first run, and one repetition for the second. While for the iCAT experiments, the IDA Extensions II script was set to two repetitions before dynamic exclusion. In all three experiments, the script was set to select a precursor ion nearest to a threshold of 15 count/s every 4 cycles. These settings ensured examination of not only high abundance ions, but low abundance ones as well.

**Data Analysis.** Data analysis for the iTRAQ experiments was performed with ProQUANT 1.0, while that for cICAT experiments was with ProICAT 1.0 SP3. The cut off for the confidence settings for both analyses was at 75; that for the score was at 20. The tolerance set for peptide identification in ProQUANT searches were 0.15 Da for MS and 0.1 Da for MS/MS, whereas those for ProICAT searches were 0.2 and 0.1 Da, respectively. All identifications were manually inspected for correctness. Relative quantification of proteins in the case of iTRAQ was performed on the MS/MS scans and is the ratio of the areas between the pairs of light and heavy label series of peaks. In both cases, the results of the quantification were normalized using the overall ratio obtained for all tagged peptide pairs in the sample. On the basis of our previous studies utilizing cICAT...
Results

iTRAQ runs 1 and 2 led to sequencing and identification of 645 and 1026 peptides, respectively. Our previous studies showed a confidence setting of 75 and a score setting of 20 were optimal. Manual inspection of hundreds MS/MS spectra concluded that while these conditions were sufficiently stringent in that most identifications were correct, they also did not exclude too many peptides that could have been identified based on spectral quality. The MS/MS spectra of all peptides that scored above the cut offs were manually inspected to verify proper identifications. The shorter MS and MS/MS cycles used for run 2 resulted in a higher number of peptides identified. A significant number of these peptides were identified more than once, thus the number of unique peptides dropped to 292 and 312 for runs 1 and 2, respectively. Many of the more abundant proteins were identified by several peptides (Figure 1); the numbers of proteins identified were 101 and 126, respectively. 63 proteins were identified in both runs. The numbers are modest, which were likely due to the small amount of starting materials. The list of proteins and their abundance ratios are shown in Table 1s in the Supporting Information section. Five proteins have abundance ratios that show more than a 2-fold change (≥2.0 or ≤0.5) in both cancer samples relative to the proliferative as well as the secretory endometrium, and meet the criteria of differential expression. These are shown in Table 1.

Relative quantification is expressed as three pairwise ratios against the proliferative endometrium (iTRAQ114) for the secretory endometrium (iTRAQ115) and the two EmCa samples (iTRAQ116 and 117). To account for small differences in protein loadings across the samples, these ratios have been normalized using the overall ratios for all proteins in the samples, as recommended by Applied Biosystems. The rationale for this choice is based on the assumption that the relative abundances for the majority of proteins are close to one. This assumption is exemplified in this study in terms of the abundance ratios in secretory versus proliferative endometrium for the following abundant proteins: cytoplasmic actin, 0.98; alpha enolase, 0.91; alpha filamin, 1.04; serum albumin precursor, 1.56; and tropomyosin alpha 4 chain, 0.99. Figure 2a shows the MS/MS spectrum for the doubly charged cytoplasmic actin peptide at 725.4 Th in one of the runs. The cluster of peaks around 115 Th is better shown in the mass spectral window 110–120 Th in Figure 2b, demonstrating near identical abundances in the four samples: proliferative (iTRAQ114), secretory (iTRAQ115), EmCa 1 (iTRAQ116), and EmCa 2 (iTRAQ117).

Typical MS/MS windows (110–120 Th) for tryptic peptides in four differentially expressed proteins, (A) chaperonin 10, (B) alpha-1-antitrypsin precursor, (C) creatine kinase B, and (D) transgelin, are shown in Figure 3. Chaperonin 10 is overexpressed in both EmCa samples, whereas alpha-1-antitrypsin precursor, creatine kinase B, and transgelin are underexpressed. The remaining differentially expressed protein, pyruvate kinase M1 or M2 isozyme, is also overexpressed in the two EmCa samples.

The cICAT experiments led to identification and quantification of 68 proteins, all were manually verified. These are listed in Table 2s in the Supporting Information section. Again, the modest number stemmed from the small sample size. Fewer proteins were identified by multiple peptides relative to iTRAQ (Figure 1), in accordance with the more selective nature of cICAT labeling on only cysteine residues. Five proteins that meet the two-fold differential expression criterion in all three EmCa versus proliferative endometrium pairs are: calgizzarin, heterogeneous nuclear ribonucleoprotein (hnRNP) D0, macrophage migration inhibitory factor (MIF), polymeric immunoglobulin receptor (PIGR) precursor, and pyruvate kinase M1 or M2 isozyme. These results are summarized in Table 2.
five proteins are overexpressed in EmCa tissues; pyruvate kinase is also shown to be overexpressed with iTRAQ. Figure 4 shows an example of overexpressed protein, calgizzarin A. Again, the relative abundance ratios are normalized to the overall ratio of all proteins in a given sample pair to account for small differences in protein loadings in the two samples. Cytoplasmic actin exhibits a heavy/light label ratio of 0.95 (0.24 (standard deviation of three samples).

Discussion

The combination of iTRAQ and cICAT labeling results in the discovery and identification of nine differentially expressed proteins that are potential cancer markers (PCMs) for EmCa. Six of the nine PCMs are overexpressed, whereas three are underexpressed in EmCa.

PCMs Overexpressed in EmCa. Chaperonin 10 is a heat shock protein that we have previously discovered as a PCM using MALDI/SELDI MS and identified by offline separation, preconcentration, trypsinization and MS/MS. Overexpression of chaperonin 10 in EmCa tissues was verified independently by Western analysis; chaperonin 10 was localized to the cancerous epithelium by means of immunohistochemistry (IHC). Elevation of levels of chaperonin 10 has been associated with large bowel and cervical carcinomas. The level of chaperonin 10 in serum was demonstrated to be an indicator of trophoblastic tumor. The observation of chaperonin 10 overexpression in EmCa samples in this study with iTRAQ confirms our previous finding using MALDI/SELDI MS, Western analysis and IHC. The possibility of detecting chaperonin 10 in serum as a marker protein for EmCa is noteworthy.

Overexpression of pyruvate kinase M1 or M2 isozyme was demonstrated by both the iTRAQ and cICAT applications. Pyruvate kinase was found to be expressed at elevated levels in both plasma and fecal samples in patients of gastrointestinal

Table 1. Differentially Expressed Proteins Quantified by iTRAQ

<table>
<thead>
<tr>
<th>accession_id</th>
<th>name</th>
<th>S:P (± SD)</th>
<th>C1:P (± SD)</th>
<th>C2:P (± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi/461730</td>
<td>10 kDa heat shock protein, mitochondrial (Hsp10) (10 kDa chaperonin) (CPN10)</td>
<td>1.06 ± 0.12</td>
<td>2.71 ± 0.52</td>
<td>2.23 ± 0.29</td>
</tr>
<tr>
<td>gi/1703025</td>
<td>alpha-1-antitrypsin precursor</td>
<td>1.25 ± 0.39</td>
<td>0.34 ± 0.07</td>
<td>0.44 ± 0.16</td>
</tr>
<tr>
<td>gi/125294</td>
<td>Creatine kinase, B chain (B–CK)</td>
<td>0.96 ± 0.04</td>
<td>0.38 ± 0.16</td>
<td>0.52 ± 0.06</td>
</tr>
<tr>
<td>gi/20178296 or gi/125604</td>
<td>Pyruvate kinase, M1 or M2 isozyme</td>
<td>1.03 ± 0.11</td>
<td>2.75 ± 0.03</td>
<td>2.02 ± 0.08</td>
</tr>
<tr>
<td>gi/3123283</td>
<td>Transgelin (Smooth muscle protein 22-alpha)</td>
<td>1.46 ± 0.26</td>
<td>0.26 ± 0.04</td>
<td>0.45 ± 0.05</td>
</tr>
</tbody>
</table>

*S:P is the ratio of secretory phase relative to the proliferative phase, C1:P and C2:P are the ratios of cancer samples 1 and 2 relative to the proliferative phase, respectively.

Figure 2. Cytoplasmic actin sequencing and quantification using iTRAQ. A. MS/MS spectrum of the doubly protonated peptide, EITALAPSTMK, at 725.4 Th. Residues in black are sequenced from the b-ion series, while those in purple from the y-ion series. Residue 'J' in the y-ion series is the iTRAQ-modified lysine residue. B. Expanded view of the low-m/z end of the MS/MS spectrum in 'A', showing relative abundances of the signature iTRAQ ions at 114.1, 115.1, 116.1, and 117.1 Th.
The ability to detect pyruvate kinase in plasma samples is of particular interest as this might indicate that screening plasma for elevated pyruvate kinase levels in endometrial cancer cases may also be possible. An investigation into the activity of 12 enzymes related to the glycolytic pathway in cervical and endometrial cancers found that only two, pyruvate kinase and phosphoglucone isomerase, were significantly higher in both cancers. Our observations of pyruvate kinase overexpression in both iTRAQ and cICAT experiments are in accordance with this earlier result regarding the same protein.

Calgizzarin was found to be overexpressed in EmCa tissues using cICAT labeling. Calgizzarin was one of two proteins that exhibited significant upregulation in colorectal and lung carcinoma cell lines over normal colorectal mucosal cells. Calgizzarin was one of three proteins that have been identified as tumor markers in mouse colon cancer. The other two tumor markers were calgranulins A and B, the former of which we recently discovered also as a protein marker in human endometrial carcinoma using SELDI MS. Identification was made possible by size-exclusion chromatography, trypsinization and online nanoLC MS/MS; confirmation of calgranulin A overexpression in EmCa was rendered by IHC in a tissue microarray format.

HnRNP D0, also known as AU-rich element RNA-binding protein 1, binds to the AU rich 3' UTR of many protooncogenes. One study has shown that this protein is more abundant in murine neoplastic lung epithelial cell lines and that this abundance decreases when nontumorigenic cells reach confluence or growth arrest. Conversely, it was also found that the abundance was unaffected in spontaneous

cancer. The ability to detect pyruvate kinase in plasma samples is of particular interest as this might indicate that screening plasma for elevated pyruvate kinase levels in endometrial cancer cases may also be possible. An investigation into the activity of 12 enzymes related to the glycolytic pathway in cervical and endometrial cancers found that only two, pyruvate kinase and phosphoglucone isomerase, were significantly higher in both cancers. Our observations of pyruvate kinase overexpression in both iTRAQ and cICAT experiments are in accordance with this earlier result regarding the same protein.

Calgizzarin was found to be overexpressed in EmCa tissues using cICAT labeling. Calgizzarin was one of two proteins that exhibited significant upregulation in colorectal and lung carcinoma cell lines over normal colorectal mucosal cells. Calgizzarin was one of three proteins that have been identified as tumor markers in mouse colon cancer. The other two tumor markers were calgranulins A and B, the former of which we recently discovered also as a protein marker in human endometrial carcinoma using SELDI MS. Identification was made possible by size-exclusion chromatography, trypsinization and online nanoLC MS/MS; confirmation of calgranulin A overexpression in EmCa was rendered by IHC in a tissue microarray format.

HnRNP D0, also known as AU-rich element RNA-binding protein 1, binds to the AU rich 3' UTR of many protooncogenes. One study has shown that this protein is more abundant in murine neoplastic lung epithelial cell lines and that this abundance decreases when nontumorigenic cells reach confluence or growth arrest. Conversely, it was also found that the abundance was unaffected in spontaneous

---

**Figure 3.** Relative abundances of differentially expressed proteins. The ion assignments are as follows: 114 Th, proliferative endometrium; 115 Th, secretory endometrium; 116 Th, first endometrial carcinoma; and 117 Th, second endometrial carcinoma. A. Tryptic peptide from chaperonin 10, showing higher abundances in the 116 and 117 Th ions relative to the 114 and 115 Th ions; B. tryptic peptide from alpha-1-antitrypsin precursor, showing lower abundances in the 116 and 117 Th ions relative to the 114 and 115 Th ions; C. tryptic peptide from creatine kinase B, showing lower abundances in the 116 and 117 Th ions relative to the 114 and 115 Th ions; D. tryptic peptide from transgelin, showing once more lower abundances in the 116 and 117 Th ions relative to the 114 and 115 Th ions.

**Table 2.** Differentially Expressed Proteins Quantified Using cICAT

<table>
<thead>
<tr>
<th>accession no.</th>
<th>name</th>
<th>avg. C:N ratio (± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S111_HUMAN</td>
<td>Calgizzarin (S100C protein) (MLN 70).</td>
<td>Singleton H&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>(P31949)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROD_HUMAN</td>
<td>Heterogeneous nuclear ribonucleoprotein D0</td>
<td>4.62 (± 0.39)</td>
</tr>
<tr>
<td>(Q14103)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIF_HUMAN</td>
<td>Macrophage migration inhibitory factor (MIF)</td>
<td>2.74 (± 0.92)</td>
</tr>
<tr>
<td>(P14174)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEBP_HUMAN</td>
<td>Phosphatidylethanolamine-binding protein (PEBP)</td>
<td>3.72 (± 1.03)</td>
</tr>
<tr>
<td>(P30086)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KPY1_HUMAN</td>
<td>Pyruvate kinase, M1 isozyme (EC 2.7.1.40)</td>
<td>4.61 (± 1.40)</td>
</tr>
<tr>
<td>(P14618)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIGR_HUMAN</td>
<td>Polymeric-immunoglobulin receptor presursor</td>
<td>9.83 (± 6.60)&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>(P01833)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Singleton H signifies extreme overexpression in the cancer sample relative to the normal proliferative control. Standard deviations were calculated from all three sample pairs. <sup>b</sup> Calculation of SD is not possible. <sup>c</sup> Large SD is due to a very high ratio, 17.45, with the other two being 5.74 and 6.31.
transformants from this cell line. Another recent study using transgenic mice showed that overexpression in an isoform of this protein altered mRNA levels of several oncogenes, including c-myc, c-jun, c-fos, and TNF-alpha. The mouse line with the highest level of this isoform developed sarcomas, suggesting that hnRNP D0 overexpression leads to tumorigenesis. Our result showing an overexpression of hnRNP D0 in EmCa is consistent with this conclusion.

Macrophage Migratory Inhibitory Factor is another protein that has been well documented as being involved with cancer. This ranges from hepatocellular carcinomas, to non small cell lung cancer, to brain tumors and gliblastomas. A recent clinical study, however, showed that MIF is in fact elevated in the peritoneal fluid of women with endometriosis. This nonspecificity suggests a decrease in the utility of MIF as a cancer marker.

PIGR has previously been detected in serum from patients with lung cancer and was shown to be significantly upregulated in the secretory component by an ELISA study involving 45 lung cancer patients compared with 45 control subjects. Another study has also demonstrated a possible linkage between PIGR and bladder carcinoma with protein levels in serum being significantly increased in patients with transitional cell carcinoma.

PCMs Underexpressed in EmCa. Alpha-1-antitrypsin precursor is one of the three proteins that are underexpressed in EmCa tissues. Recent studies have shown downregulation of alpha-1-antitrypsin to be associated with malignant lymphoma as well as liver, lung, stomach, bladder, and gall bladder cancer. There is also evidence for upregulation contributing to enhanced cell migration and metastases of human colon cells in a rat model. Such a contradiction might be explained by the effect of alpha-1-antitrypsin being tissue specific. If this were proven to be the case, this would prove useful for distinguishing between forms of cancer.

Creatine kinase B shown here as being underexpressed in EmCa has likewise been demonstrated as being downregulated in cervical, colon, and lung cancers but not in hepatocarcinoma. Transgelin is another protein observed to be underexpressed in this study. It has also been implicated as being downregulated in breast and colon carcinomas as well as in a lung epithelial cell line. Transgelin was isolated and identified as an antigen from renal cell carcinoma; subsequent in-situ hybridization experiments, however, found that the malignant cells were negative with respect to transgelin and that the transgelin source was from the mesenchymal cells of the stroma. Thus, downregulation of transgelin appears to be true for renal cell carcinoma as well.

Proteins Showing Possible Differential Expression. There are four proteins that showed differential expression by cICAT, but a smaller than critical (two-fold) change by iTRAQ labeling. One such protein is cyclophilin A, which was observed as being overexpressed by approximately four-fold in the ICAT analysis; however, the iTRAQ experiments showed a smaller overexpression of 1.47 ± 0.29 across both cancer samples in both runs compared with the normal proliferative sample, which is not considered as significant according to our criterion of a two-fold change. Cyclophilin A has recently been reported as overexpressed in lung cancer. Further investigations into the relative abundances of cyclophilin A in EmCa and normal epithelia may be warranted.

Triosephosphate isomerase is another protein showing overexpression by cICAT tagging. However, the relative standard
deviation for this protein at ±72% was also large, suggesting a variable expression level. There is a recent study that suggests triosephosphate isomerase is overregulated by two-fold in lung cancer. Triosephosphate isomerase level was found to be highly variable in renal cell carcinoma. ITRAQ labeling shows an overexpression of 1.55 ± 0.28. Together these probably suggest that triosephosphate isomerase is unsuitable as a cancer marker.

Superoxide dismutase [Cu–Zn] is another protein that showed a large variation, ±50%, in the cICAT results. This protein has been implicated in pancreatic adenocarcinoma. ITRAQ results at 1.46 ± 0.08 is below the criterion of overexpression. Phosphatidylethanolamine binding protein showed a cICAT overexpression level of 3.7 ± 1.0 and an iTRAQ overexpression level of 1.39 ± 0.13. There is evidence for upregulation in rat hepatoma cell lines.

The use of isotope-coded affinity tags and nanoLC MS/MS afforded examination of a large number of proteins for differential expression. This method is a lot more efficient and powerful in terms of the number of proteins that can be examined in a given experiment than SELDI MS. However, complex data analysis and the need for manual examination of MS/MS data limit the number of samples that can be examined within a given period of time. As a consequence, iTRAQ tagging involved two normal endometrial tissues versus two EmCa tissues, whereas cICAT pairing involved only one normal endometrial versus three EmCa tissues. By contrast, PCM discovery using SELDI MS involved in excess of 40 samples. A contribution to the uncertainties of the aforementioned four proteins may simply be individual variations in protein expression, both within the normal group as well as the EmCa group. Our experience with the results of chaperonin 10 and calgranulin A shows that, while there is overexpression in EmCa tissues, there are considerable variations in protein abundances, possibly reflecting variability in the cellular subpopulations of the whole tissue homogenates or in the type or nature of the EmCa. In addition, there is variation in the abundances of these proteins across individual normal endometrial samples. Thus, uncertainties regarding the status of some proteins are to be expected because of the limited sample size in the feasibility studies. Differential expression of the proteins identified in this endeavor should be verified with a large data set and/or confirmed with an independent method, e.g., IHC or Western analysis. Furthermore, as these differentially expressed proteins were discovered in whole tissue homogenates, the precise cellular origin(s) is unknown. The exception is chaperonin 10, which has already been discovered and identified to be overexpressed in EmCa by means of SELDI MS, and verified by IHC and Western analysis. The use of IHC permits explicit protein localization. This is potentially a more viable method to localize protein markers than proteomic analysis on laser capture microdissection (LCM)-procured cancer cells. Li et al. reported the use of 100 µg proteins per sample from LCM-procured cancer and normal liver cells for cICAT analysis, and dissection of 50 000–100 000 cells, which was very laborious and may not be possible for other types of cancer.

Selectivity of PCMs. Literature data show that the nine differentially expressed proteins discovered and identified here are associated with cancers and are, indeed, potential cancer markers. Individually, these nine PCMs are nonspecific for endometrial carcinoma, as each has been linked with cancers other than EmCa. As the principal concern in screening, diagnosis, and monitoring of EmCa is the exclusion or omission of any malignant endometrial disease, the fact that these PCMs have been noted in other cancers would rarely be a potential clinical drawback. In other words, for envisaged clinical use, the sensitivity is of far more concern than specificity for EmCa. It is probable that different cancers will share similar pathways in tumorigenesis, which will induce over- and underexpression of similar proteins. Nevertheless, it is unlikely that the differential expression pattern involving many proteins will be identical for all or some primary tumor sites. Their collective use as a panel may be more effective in indicating the site of origin of a cancer, despite the association of individual markers with a variety of primary sites. Markers that are secreted or shed into blood or body fluids in detectable concentrations may constitute the basis of a less-invasive diagnostic test. Chaperonin 10 and pyruvate kinase have been detected in blood, which raises the possibility that a non-tissue based diagnostic test may be feasible. Judging from current results based on serum analyses by mass spectrometry, the sensitivity of this technology will need to be dramatically raised before it will be adopted as the means of detection in such a diagnostic test.

The absolute expression level of PCMs will be an important issue. Knowing protein abundances in normal and diseased states will allow establishment of threshold levels beyond which EmCa is signaled. ITRAQ permits simultaneous investigation of up to four samples, thus facilitating the inclusion of synthetic tryptic peptides of known amounts in absolute isotope-dilution experiments. It stands to reason that, in the near future, absolute measurements of these PCMs in EmCa and normal
tissues should be made, which will permit establishment of ranges of PCM concentrations in normal and EmCa tissues.

**iTRAQ versus cICAT.** Our results obtained by iTRAQ and ICAT analyses suggest that the information generated by the two methods is complementary. There are, however, a few aspects on which each of these methods has advantages over the other. Quantification by cICAT can be compromised by overlapping peaks in the MS spectrum; this complication is resolved in iTRAQ as quantification is performed on the MS/MS spectrum. iTRAQ on the other hand requires processing samples separately until after the tryptic digestion. This increases the potential for errors introduced, as a result of sample handling or different extents of tryptic digestion. Another aspect on which iTRAQ differs is complexity because of the relatively nonspecific nature of labeling. Many more proteins are identified by multiple peptides. This is not necessarily a disadvantage, as it permits quantification of multiple peptides, thereby increasing the confidence in the ratios report. After classifying identified proteins into broad categories (Figure 5), it is apparent that there is a higher proportion of signaling proteins identified by the cICAT method. Conversely, iTRAQ analysis identified a larger percentage of the more abundant ribosomal proteins and transcription factors. This observation is inline with the expectation that age of the more abundant ribosomal proteins and transcription factors. This observation is inline with the expectation that

**Conclusion.** While the results obtained from both the methods used in this study have provided us with exciting possibilities, they are only early indicators and need to be further verified by a larger study or by other techniques, such as IHC or Western analysis. Further studies with iTRAQ will profit from absolute quantification of the potential markers from this study. This will involve the use of known amounts of synthetic peptides from these protein markers being labeled, replacing the duplicate cancer sample used in this study. Multiple marker techniques such as iTRAQ and cICAT analyses may turn out to be better suited for initial marker discovery, the results of which would require validation by other more "traditional" methods that are more amenable to large-scale screening.

**Acknowledgment.** The authors gratefully acknowledge the support and cooperation of their surgical and clinical laboratory collaborators, including Drs. K. Joan Murphy (University Health Network), Wusun Paek (Mount Sinai Hospital), and Titus Owalabi, Ali Qizilbash and Dennis MacDonald (North York General Hospital). This research was supported by the National Cancer Institute of Canada (NCIC), and the Ontario Genomics Institute (OGI) and Genome Canada. Infrastructure support was provided by the Canadian Foundation for Innovation (CFI) and the Ontario Innovation Trust (OIT), and the Ontario Research and Development Challenge Fund (ORDCF).

**Supporting Information Available:** Table 1s: Proteins identified and abundance ratios in the iTRAQ experiments. Table 2s: Proteins identified and abundance ratios in the cICAT experiments. These materials are available free of charge via the Internet at http://pubs.acs.org.

**References**


research articles


