Ida Pucci-Minafra
Simona Fontana
Patrizia Cancemi
Loredana Basiricò*
Silvana Caricato
Salvatore Minafra

Center of Experimental Oncobiology and Department of Experimental Oncology and Clinical Applications, University of Palermo, Italy

A contribution to breast cancer cell proteomics: Detection of new sequences

Ductal infiltrating carcinoma (DIC) of the breast is the most common and potentially aggressive form of cancer. Knowledge of proteomic profiles, attained both *in vivo* and *in vitro*, is fundamental to acquire as much information as possible on the proteins expressed in these pathologic conditions. We used the breast cancer cell line 8701-BC, established from a primary DIC, with the aim of contributing to the databases on mammary cancer cells, which in turn will be very useful for the identification of differentially expressed proteins in normal and neoplastic cells. Within an analysis window comprising about 1750 discernible spots, we have at present catalogued 84 protein spots. The proteins for which an identity was assigned were identified essentially using gel comparison, *N*-terminal (Nt) microseqencing and immune detection. Among the protein spots Nt-microsequenced, sixteen corresponded to known proteins, four resulted as modified, relative to matching sequences deposited on databases, and seven were unknown. These modified or novel sequences are thus of potential interest to the knowledge of breast cancer proteomics and its applications.

Keyword: Breast cancer PRO 0214

1 Introduction

Breast cancer is the most common cancer among Western women. However, breast cancer is not a single disease, but includes several different forms that can be grouped into invasive and noninvasive histotypes. Among the invasive, ductal infiltrating carcinoma (DIC) is the most common and aggressive form [1]. This kind of carcinoma usually has a poor prognosis, and represents the standard histotype with which the other less frequent subtypes are compared. Carcinomas are potentially malignant tumors of epithelial origin, which, once genetically initiated, grow, progress, invade and spread from one organ or tissue to another with a common phenotype, regardless of their individual derivation. The major cellular changes involved in the conversion of normal to malignant breast are the progressive loss of the stationary epithelial phenotype and the acquisition of a mesenchimal-like phenotype, correlated with the ability to migrate and invade surrounding and distant tissues, even under conditions of low oxygen supply. These features are related to defects in a multiplicity of genes and proteins, whose number, identity and functions are still far from understood. The proteomic approach is therefore considered at present to be the most promising to investigate protein expression and function and to monitor cancer related defects of large numbers of proteins within a cell or tissue. However, some limitations to this approach are intrinsic to the experimental system: on the one hand primary tumor biopsies may contain several different cell types other than carcinoma cells, and laser-capture microdissected sections [2] may include a large amount of proteins resident in the extracellular interstitium. On the other hand neoplastic cells in culture, being free of any tissue components, may not represent their exact *in vivo* counterparts. In spite of this limitation, cancer cells in culture represent a powerful substitute for the study of molecular and cell biology. Therefore, several laboratories, including our own, have developed appropriate *in vitro* models of normal and neoplastic cells for biological and molecular assays including the recent approach of cDNA microarrays [3] and proteomics [4–7].

Among the breast cancer cell lines described so far in the literature, 8701-BC [8–10] deserves particular attention, since it was established from a primary DIC, that is, before the clonal selection of the metastatic process. This cell line maintains a number of properties in culture that are characteristic of mammary tumor cells *e.g.* expression of cytokeratin 8 and 18 (typical of the mammary lumenal phenotype), loss of polarised morphology, accelerated growth rate, release of matrix metalloproteases, loss of contact inhibition and formation of multilayered colonies (domes), low dependence on serum and growth factor

Correspondence: Dr. Ida Pucci-Minafra, Centro di OncoBiologia Sperimentale, Viale delle Scienze, 90128 Palermo, Italy
E-mail: idapucci@unipa.it
Fax: +39-91-424903

Abbreviations: DIC, ductal infiltrating carcinoma; **Nt**, *N*-terminal; **UBCS**, unknown breast cancer secreted

* Current address: University della Tuscia, Viterbo, Italy

 1615-9853/02/0707–919 $17.50+.50/0

and the ability to invade matrigel and nude mice. These properties are clearly distinguished from normal mammary cells in culture (*i.e.* the HB2 derived from lumenal epithelial cells [11]), which maintain the polarised phenotype, establish correct cell-cell interactions, do not invade *in vitro* and *in vivo* and are dependent on serum. We have thus used this cell line model for the analysis of protein expression profiles, with the aim of contributing to the databases on mammary cancer cells present in the literature and on the internet, which in turn will be very useful for the identification of differentially expressed proteins in normal and neoplastic cells. Within an analysis window comprising about 1750 discernible spots, we have at present recorded 84 protein spots, utilising following methods: *N*-terminal (Nt) microsequencing (31), immune detection (23) and gel matching (30) with available maps at the 2-DE databases (http://www.expasy.ch/ch2d/2d-index.html). Among the protein spots Nt-microsequenced, sixteen corresponded to known proteins, four were modified relative to matching sequences deposited on databases, seven were unknown and four were *N*-blocked.

2 Materials and methods

2.1 Cell culture

For the experiments described, we used the neoplastic cell line 8701-BC under the conditions previously described [8]. Briefly, the cells were seeded at a concentration of 1–2 $\times$ 10^5/cm^2 in RPMI 1640 culture medium (Gibco, Paisley, UK), supplemented with 10% foetal calf serum (Gibco) and antibiotics (100 units/mL penicillin and 100 μg/mL streptomycin), and grown in a humidified incubator with 3% CO_2 in air at 37°C. As nontumoral control we used the HB2 human mammary epithelial cell line, kindly given by Dr. J. Taylor-Papadimitriou (Imperial Cancer Research Fund, London, UK) and currently cultivated in our lab under the conditions described by Bartek and coworkers [11].

2.2 Sample preparations

Cells were grown to confluence and then incubated with three changes of serum-free medium over 24–48 h. After washing with ice-cold PBS, they were carefully scraped and incubated on ice for 30 min with RIPA buffer (50 mM Tris pH 7.5, 0.1% Nonidet P-40, 0.1% deoxycholate, 150 mM NaCl, 4 mM EDTA and a mixture of protease inhibitors, 0.01% aprotinin, 10 mM sodium pyrophosphate, 2 mM sodium orthovanadate, 1 mM PMSF). The total cellular lysate was centrifuged at 14 000 rpm for 8 min to clear cell debris, and the supernatant dialysed against ultrapure distilled water, lyophilised and stored at –80°C until analysis. Culture media were collected from confluent cells deprived of serum, extensively dialysed against ultrapure distilled water; lyophilised and stored at –80°C until used. Protein concentration in the cellular extracts and culture media was determined using the Bradford method [12].

2.3 Two-dimensional gel electrophoresis

Aliquots of the dried cell lysate were solubilised in a buffer containing 4% CHAPS, 40 mM Tris, 65 mM DTE (1,4-dithioerythritol) and a trace amount of bromophenol blue in 8 M urea. The first dimensional separation was performed at 20°C on commercial sigmoidal immobilised pH gradient strips (IPGs), 18 cm long with pH range 3.5–10, (Amersham Pharmacia Biotech, Uppsala, Sweden), essentially as described by Görg *et al.* [13] and Bjellqvist *et al.* [14]. Strips were rehydrated in 8 M urea, 2% CHAPS, 10 mM DTE and 0.5% carrier ampholytes (Resolyte 3.5–10; BDH, Merck, Milan, Italy). Aliquots of 45 μg (analytical gels) or 1.5 mg (preparative gels) of total proteins were applied to the gel strip. IEF was carried out by linearly increasing the voltage from 200 to 3500 V during the first 3 h, after which focusing was continued at 8000 V for 8 h. After the run the IPG strips were equilibrated with a solution containing 6 M urea, 30% glycerol, 2% SDS, 0.05 M Tris-HCl pH 6.8 and 2% DTE for 12 min, in order to resolubilise proteins and reduce disulfur bonds. The –SH groups were then blocked by substituting the DTE with 2.5% iodoacetamide in the equilibration buffer. The focused proteins were then separated on 9–16% linear gradient polyacrylamide gels (SDS-PAGE) with a constant current of 40 mA/gel at 10°C. Gels were stained with ammoniacal silver nitrate as described by Hochstrasser *et al.* [15].

2.4 Image acquisition and data analysis

Silver stained gels were digitised using a computing densitometer and processed with the Melanie 3 computer system (GeneBio, Geneva, Switzerland). As internal standard for gel calibration we used an array of 15 features of the cell lysate, whose p*I* and M_r were formerly estimated by interpolation with serum proteins that had comigrated with the whole cell lysate. Background subtraction was performed by considering only the common spots of three different gels of the same sample matched together. For figure presentations, images were transferred to Adobe Photoshop and PowerPoint programs.

2.5 Protein identification

The proteins for which an identity was assigned were identified essentially using gel comparison [16], Nt microseqencing and immune detection (see Sections 2.6 and 2.7). For protein identification by computer gel matching we used the liver map (http://www.expasy.ch/cgi-bin/map2/def?LIVER_HUMAN), the colorectal adenocarcinoma DL-1 cell line map (http://www.expasy.ch/cgi-bin/map2/def?DLD1_HUMAN), the gastrocnemius muscle map (http://www.expasy.ch/cgi-bin/map2/def?MUSCLE_MOUSE) deposited in the SWISS-2DPAGE, and the human breast ductal carcinoma map (http://www.biomol.unisi.it/cgi-bin/map2/def?BREAST_DC_HUMAN) deposited in the SIENA-2DPAGE database.

2.6 Protein microsequencing

N-terminal microsequencing was performed by automated Edman degradation in a protein sequenator (Procise 491, Applied Biosystems, Monza, Italy) on protein samples electrotransferred onto PVDF membranes (Millipore, Immobilon-PTM; Sigma, Milan, Italy) at 50 V for 30 min at 4°C. The SWISS-PROT (ScanProsite) and EMBL (Bic 2) databases were used for on-line sequence similarity search at http://www.expasy.ch/tools/#similarity.

2.7 Western blotting

For immune detection the 2-D IPG gels were electrotransferred onto nitrocellulose membrane (HyBond ECL, Amersham) at 50 V for 1 h at 4°C. After electrotransfer, the membranes were stained with Ponceau S (Sigma) and the separation pattern recorded in order to perform subsequent gel matching. The membranes were then probed with one of the following monoclonal antibodies: anti-cytokeratin-8 and 18 (Chemicon International, Milan, Italy), antivimentin (Novocastra, DBA Italia, Milan, Italy), and anti-alpha-enolase (courtesy of Prof. S. Feo and colleagues, [17]). The reaction was revealed by the ECL detection system, using high performance films (Hyperfilm ECL, Amersham).

3 Results

In Fig. 1A to C phase contrast micrographs of the 8701-BC cell culture are shown at 48 h, 1 week and 10 d from seeding, compared with HB2 cells derived from normal mammary epithelium at the same growth time (Fig. 1D–F). As can be observed from these micrographs, the neoplastic cells display an apolarised morphology with irregular profiles; they spread over the substrate with emission of spikes and once confluence is reached they overlap, forming distinctive domes (Fig. 1C). This growth modality is typical of neoplastic cells and is dramatically evident when compared with nontumoral cells, which maintain *in vitro* the polygonal morphology of epithelial cells, establish extensive cell-cell contact forming orderly islands and stop growing once confluent, being responsive to contact inhibition signals.

Since cell morphology is the first manifestation of cell phenotype and the phenotype of a given cell type is, in turn, determined by the activation status of its genome and by the composition of its protein complement in a given moment of cell life, the proteomic pattern represents the molecular support for phenotyping cells and for describing disease associated alterations, which may involve an unpredictable number of known and unknown proteins, with structural and functional roles in the cell. With this aim, we compared the proteomic profiles of sets of proteins from 8701-BC cells *vs.* HB2 nontumoral mammary cells reported elsewhere [18]. In the present report we extend previous data with validation and additional identification of new sets of proteins.

Figure 2 shows the proteomic map of the 8701-BC cell lysate in which, among the 1750 spots detected by the Melanie system after background subtraction, we have identified and catalogued 84 protein spots. The proteins to which identities were assigned are marked with labels corresponding to the accession number of the SWISS-PROT database, which also includes the major isoelectric variants of assigned proteins. The unknown proteins are labelled with the abbreviations UBC (unknown breast cancer) or UBCS (unknown breast cancer secreted), followed by the approximate molecular mass (kDa). The last category includes three proteins found in the 2-D IPG map of the conditioned medium of the 8701-BC cells, which was matched with the cell lysate map. Figure 3 shows the 2-D gel section of the conditioned medium, which includes the Nt-microsequenced spots, corresponding to the secreted proteins with unknown identity.

Table I summarises the catalogue of the protein spots described in Fig. 2 with the following indications: AC and Entry name, corresponding respectively to the accession numbers and the abbreviations used in the SWISS-PROT database, protein name, experimental p*I* and experimental M_r, method of protein identification (GM: gel matching (30), Nt-Mi: *N*-terminal microsequencing (31), WB: Western blot (23)) and *N*-terminal sequences, along with substitutions (when present) relative to similar sequences deposited in the protein databases. Within the proteins *N*-terminal sequenced, sixteen corresponded to proteins

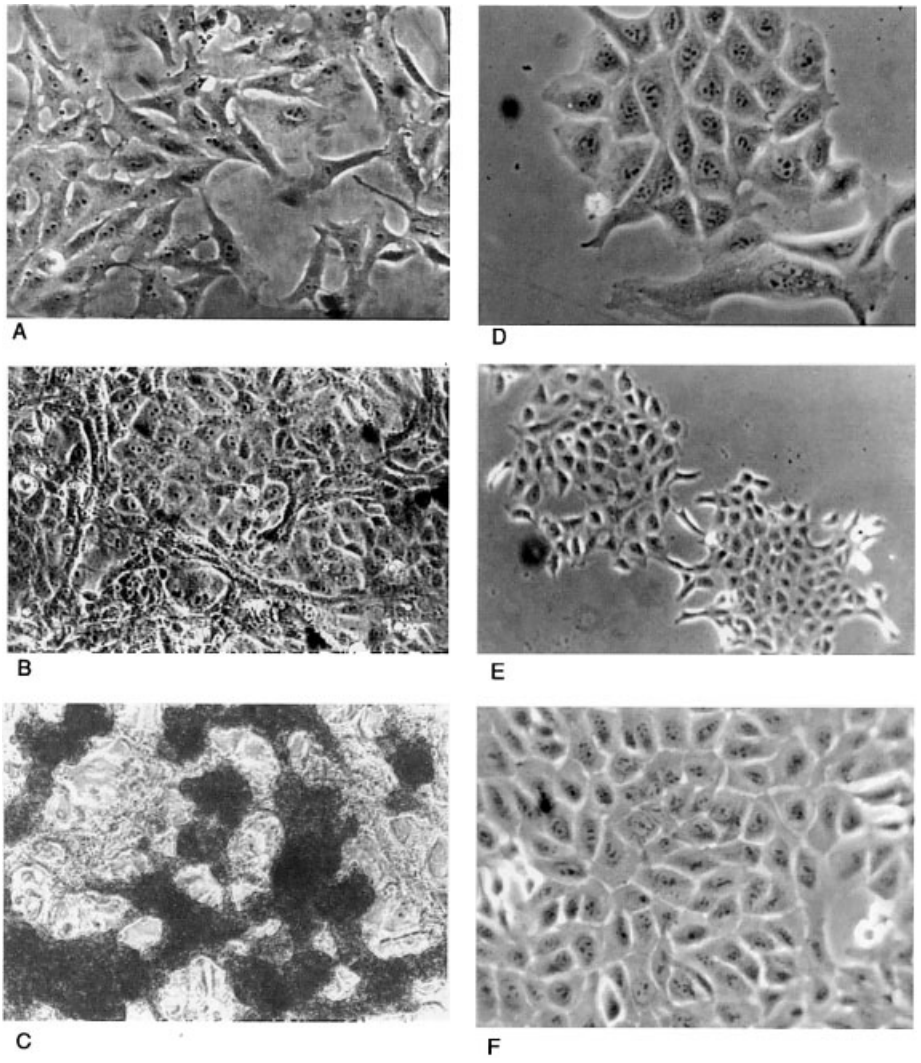

Figure 1. Phase contrast micrographs illustrating the growth modality of the breast cancer cell line (8701-BC: A–C) used for the present research, in comparison with nontumoral mammary cells (HB2: D–F) at 48 h, 1 week and 10 d from seeding. Micrographs were taken from flask culture with an inverted microscope. Objective magnification 20x (A, C, D, F) and 10x (B, E).

already present in the databases, four appeared to be modified relative to the matching proteins, seven resulted unknown and four were *N*-blocked.

Among the proteins present in the databases, two have not been described previously in breast cancer, *i.e.* the ubiquitin thiolesterase L1 and the glycoprotein GC1QBP (MA32). The four proteins showing alterations concerned thioredoxin, in which the threonine in position 8 is substituted with alanine; a truncated and slightly more basic form of annexin I (p*I*/M_r 8.04/25752, starting from residue 26) displaying two substitutions, namely serine 27 and glycine 30, both replaced by isoleucine; a truncated form of alpha-enolase (starting from residue 57) and a shortened form of glyceraldehyde 3-phosphate dehydrogenase starting with regular *N*-terminal residues. In addition, seven protein sequences, some of which showed ambiguities, did not match any known sequence deposited in the databases. The protein spots released into the medium by the 8701-BC cells included two unknown sequences (UBCS) and a small peptide (p*I*/M_r 5.97/17100) matching residues 4199-4207 of the perlecan core protein (PGMB).

Finally, the protein identified by immune detection regarded an isoelectric series corresponding to alpha-enolase, coincident with the silver stain pattern, and a group of intermediate filament proteins. These included seven spots that bound to anticytokeratin-8, five to anticytokeratin-18 and eight to antivimentin (Fig. 4A–C). In parallel, the immune detection for intermediate filaments was also performed on the HB2 cell lysate. In this case the reaction was negative for the antivimentin assay, and positive for both anticytokeratins (Fig. 4D–F). However, as shown in Fig. 4, the number and the isoelectric variants of both cytokeratins were dissimilar from those observed in the 8701-BC cell map. As the majority of the cytoskeletal proteins were almost indiscernible in the silver stained

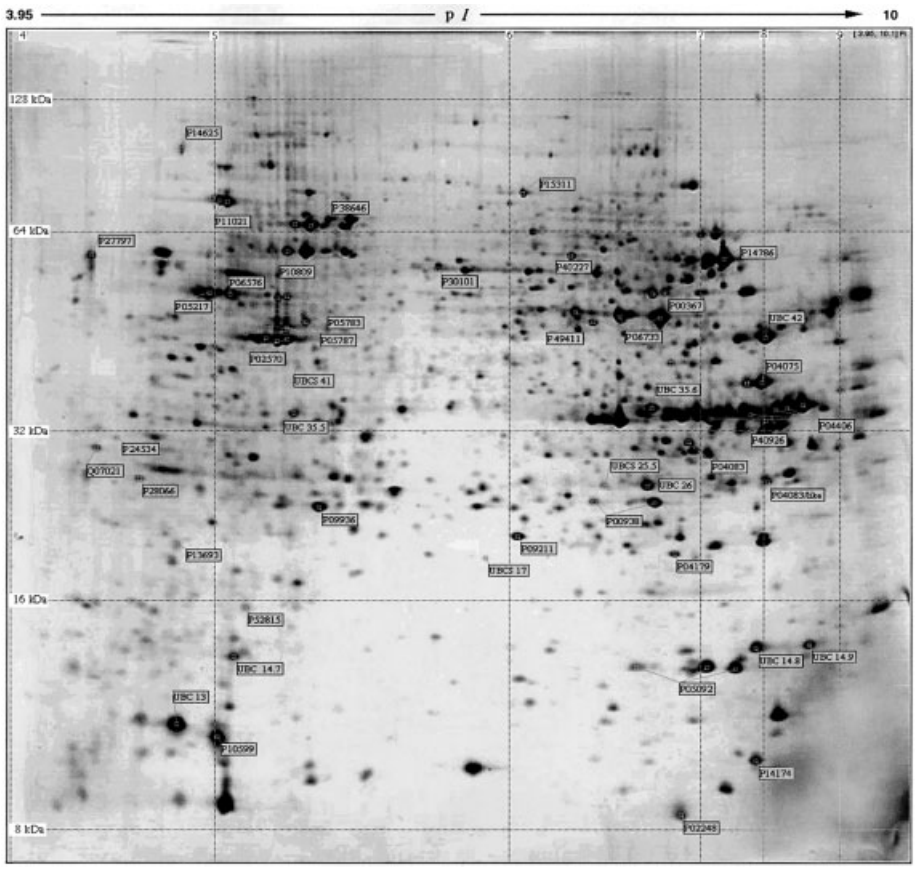

Figure 2. 2-D IPG of the total protein extract from 8701-BC cells at seven days from seeding (confluence). Forty-five μg of proteins was loaded on an IPG gel (18 cm, 3.5–10 NL). The second dimension was performed on a vertical linear-gradient slab gel (9–16%T). The gel was silver stained, digitised and analysed with Melanie 3 software. Protein spots of known identity are labelled with the accession number of the SWISS-PROT database. The unknown proteins are labelled with the abbreviations UBC (unknown breast cancer) or UBCS (unknown breast cancer secreted) followed by the approximate mass (kDa).

gels, their attribution to the proteomic map presented in Fig. 2 was achieved by gel matching with the corresponding pattern on the ECL film.

4 Discussion

Cancer is a somatic gene disease, but at present it is not known how many genes are directly or indirectly involved in the transition from normal to malignant cells. Abnormalities are expected at several levels: gene alterations, qualitative/quantitative protein expression, primary and post-translational modifications of protein structure or cytoplasmic delocalization, any of which may ultimately affect cell function and behavior. Therefore the first goal is to obtain the proteomic profiles of normal and neoplastic cells and tissues and to identify differentially expressed proteins as putative contributors to specific malignancies. Hence the proteomic approach, by allowing the study of variations in gene expression under a large-scale of magnitude, now represents a powerful tool to ascertain which protein combination may play a key role in transformation and tumor progression and apply the obtained information to new clinical strategies.

In the present work we aim to provide a contribution to the knowledge of proteomic profiles of breast cancer using the well-characterised 8701-BC cell line. As stated in Section 1, this cell line is derived from a primary DIC and maintains the phenotypic features of tumorigenic cells in culture, *i.e.* the ability to overlap and form distinctive domes, the loss of polarity and acquisition of motile behavior and an invasive capacity both *in vitro* and *in vivo*. Here we demonstrate for the first time that these cells co-express several forms of cytokeratins (–8 and –18) and vimentin, which are respectively cytoskeletal markers of lumenal mammary epithelium [19] and of mesenchimal cells. These proteins were identified by immune detection on 2-D IPG gels as multiple forms, many of which are not included in current databases. The co-expression of cytokeratins and vimentin by the 8701-BC cells is a further demonstration that many features related to cancer cells (*i.e.* loss of polarity, defective cell-cell contacts, increased motility) reflect an epithelial-mesenchymal transition, as postulated for many carcinomas of different origin [20]. This epithelial-mesenchymal transition is not due to the *in vitro* culture, since HB2 cells used as control were negative for the antivimentin reaction. The proteomic profile of 8701-BC cells, at present a collection of 84 protein spots, including known proteins, and their isoforms and

Table 1. Catalogue of the protein spots identified in the 2-D IPG map of the 8701-BC reported in Fig. 2. The following are given: AC and Entry name, corresponding respectively to the accession numbers and the abbreviations used in the SWISS-PROT database, protein name, isoelectric point and molecular weight, method of protein identification, *N*-terminal sequences along with substitutions (when present) relative to similar sequences deposited in the databases.

SWISS-PROT AC	Entry name	Protein name	p*I*-M_r	Identification method	Nt-Sequence substitutions
P00367	GDH	Glutamate dehydrogenase 1	6.80 49913	GM (Li, Br)	
P00367	GDH	Glutamate dehydrogenase 1	6.86 49913	GM (Li, Br)	
P00938	TPIS	Triosephosphate isomerase	6.50 25030	GM (Li, Br)	
P00938	TIPIS	Triosephosphate isomerase	6.81 25000	Nt-Mi	APSRKFFVGG Res. 1–10
P02248	UBIQ	Ubiquitin	6.89 8700	GM (Li, Br)	
P02570	ACTB	Beta-actin	5.16 42000	GM (Li, Br)	
O02570	ACTB	Beta-actin	5.2042000	GM (Li, Br)	
P02570	ACTB	Beta-actin	5.23 42000	GM (Li, Br)	
P04075	ALFA	Fructose-biphosphate aldolase A	7.60 38021	GM (Br)	
P04075	ALFA	Fructose-biphosphate aldolase A	7.99 38100	GM (Br)	
P04083	ANX1	Annexin I	6.94 30530	Nt-Mi	XIENEEQEYV Res. 12–21
P04179	SODM	Superoxide dismutase [Mn]	6.88 18084	GM (Li, Br)	
P04406	G3P2	Glyceraldehyde 3-phosphate dehydrogenase	7.15 34246	Nt-Mi	GKVKVX Res. 1–6
P04406	G3P2	Glyceraldehyde 3-phosphate dehydrogenase	7.61 35500	Nt-Mi	GKVKVGVNGF Res. 1–10
P04406	G3P2	Glyceraldehyde 3-phosphate dehydrogenase	8.07 35500	Nt-Mi	GKVKVGVNGFGRIGRLV TRA Res. 1–20
P04406	G3P2	Glyceraldehyde 3-phosphate dehydrogenase	8.23 35652	Nt-Mi	GKVKVGVNGF Res. 1–10
P04406	G3P2	Glyceraldehyde 3-phosphate dehydrogenase	8.47 35959	Nt-Mi	GKVKVGVNGFGRIGRLV TRAAFNSGKVDIV Res. 1–30
P04406	G3P2 tr	Glyceraldehyde 3-phosphate dehydrogenase tr	8.46 34490	Nt-Mi	GKVKVGVNGF Res. 1–10
P05092	CYPH	Peptidyl-prolyl cis-trans isomerase A	6.66 14500	GM (Li, Br)	
P05092	CYPH	Peptidyl-prolyl cis-trans isomerase A	7.07 14500	GM (Li, Br	
P05092	CYPH	Peptidyl-prolyl cis-trans isomerase A	7.51 14500	GM (Li, Br)	
P05217	TBB2	Tubulin beta-2 chain	4.98 49900	GM (Br)	
P05783	CK18	Cytokeratin 18	5.05 47074	WB	
P05783	CK18	Cytokeratin 18	5.12 48000	WB	
P05783	CK18	Cytokeratin 18	5.21 49200	WB	
P05783	CK18	Cytokeratin 18	5.23 49300	WB	
P05783	CK18	Cytokeratin 18	5.27 49500	WB	
P05787	CK18	Cytokeratin 8	5.10 42338	WB	
P05787	CK18	Cytokeratin 8	5.14 42600	WB	
P05787	CK8	Cytokeratin 8	5.17 43000	WB	
P05787	CK8	Cytokeratin 8	5.23 42800	WB	
P05787	CK8	Cytokeratin 8	5.25 47000	WB	
P05787	CK8	Cytokeratin 8	5.27 43300	WB	
P05787	CK8	Cytokeratin 8	5.40 47176	WB	
P06576	ATPB	ATP synthase beta chain	5.05 49900	GM (DL-1 cells)	
P06733	ENOA	Alpha enolase	6.41 47074	WB	
P06733	ENOA	Alpha enolase	6.59 47074	WB	
P06733	ENOA	Alpha enolase	6.84 47014	WB	
P06733	ENOA tr	Alpha enolase tr	6.87 41563	Nt-Mi	MGKGVSKAVE Res. 57–66
P08670	VIME	Vimentin	5.05 52945	WB	
P08670	VIME	Vimentin	5.13 53300	WB	
P08670	VIME	Vimentin	5.01 51055	WB	
P08670	VIME	Vimentin	5.14 51029	WB	
P08670	VIME	Vimentin	5.11 51029	WB	
P08670	VIME	Vimentin	4.93 49900	WB	

Table 1. Continued

SWISS-PROT AC	Entry name	Protein name	p*I*-M_r	Identification method	Nt-Sequence substitutions
P08670	VIME	Vimentin	4.82 47278	WB	
P08670	VIME	Vimentin	4.83 42750	WB	
P09211	GSTP	Glutathione S-transferase P	6.01 20635	Nt-Mi	PPYTVVYFPV Res. 1–10
P09936	UBL 1	Ubiquitin thiolesterase L 1	5.34 23940	Nt-Mi	MQLKPMEINPEMLV Res. 1–14
P10599	THIO	Thioredoxin	5.01 12579		XKQIESKAAF Res. 1–10 Substitutions: T 8 –>A
P10809	HSP 60	60 kDa heat shock protein	5.19 59500	GM (Li, Br)	
P10809	HSP 60	60 kDa heat shock protein	5.2359500	Nt-Mi	GADAXA Res. 33–38
P10809	HSP 60	60 kDa heat shock protein	5.27 59500	Nt-Mi	GADAXA
P11021	GRP 78	78 kDa glucose-regulated protein	4.99 7310	GM (Br)	
P11021	GRP 78	78 kDa glucose-regulated protein	5.02 73100	GM (Br)	
P11021	GRP 78	78 kDa glucose-regulated protein	5.04 73100	GM (Br)	
P13693	TCTP	Translationally controlled tumor protein	4.90 19258	GM (Li, Br)	
P14174	MIF	Macrophage migration inhibitory factor	7.62 11900	GM (Br)	
P14625	ENPL	Tumor rejection antigen 1	4.90 91200	GM (Br)	
P14786	KPY 2	Pyruvate kinase M2	7.28 57820	GM (Muscle Mouse)	
P15311	EZRI	Ezrin, Cytovillin	6.02 75000	GM (DL-1 cells)	
P24534	EF1b	Elongation factor 1-beta	4.65 30700	GM (Li, Br)	
P27797	CRTC	Calreticulin	4.52 59000	GM (Br)	
P28066	PSA5	Proteasome subuniut alpha type 5	4.77 25900	Nt-Mi	TRSEYDRGVN Res. 4–13
P30101	ERP60	Disulfide isomerase ER-60	5.88 53600	Nt-Mi	SDVLELTDDNFE Res. 25–36
P38646	GRP 75	Mitochondrial stress-70 protein, Mortalin	5.24 68300	GM (Li, Br)	
P38646	GRP 75	Mitochondrial stress-70 protein, Mortalin	5.28 68300	GM (Li, Br)	
P40227	TCPZ	T-complex protein 1, zeta subunit	6.39 59000	GM (DL-1 cells)	
P40926	MDHM	Malate dehydrogenase	8.03 33835	Nt-Mi	AKVAVLGAXG Res. 25–34
P49411	EFTU	Elongation factor Tu	6.50 43300	Nt-Mi	AVEAKKTYV Res. 44–52
P52815	RM12	Mitochondrial 60S ribosomal protein L7/L12	5.10 15820	GM (DL-1 cells)	
Q07021	MA32	Glycoprotein GC1QBP	4.54 30387	Nt-Mi	LHTEGDKAFV Res. 74–84
	ANX tr	Annexin I-like	8.04 25752	Nt-Mi	XIKGIPGSAV Res. 26–35 Substitutions: S 27–>I; G 30–>I
	UBC 13	Unknown	4.90 12924	Nt-Mi	N-blocked
	UBC 14.7	Unknown	5.06 14687	Nt-Mi	N-blocked
	UBC 14.8	Unknown	7.62 14829	Nt-Mi	AYVADLDQ
	UBC 14.9	Unknown	8.54 14900	Nt-Mi	N-blocked
	UBC 26	Unknown	6.76 25606	Nt-Mi	N-blocked
	UBC 35.5	Unknown	5.25 35500	Nt-Mi	ALFQRAR(D/G)
	UBC 35.6	Unknown	6.80 35600	Nt-Mi	QRQDAQPXQ
	UBC 42	Unknown	8.03 42000	Nt-Mi	G(HAS)L(LQ)E(IE)TNXPG
	UBCS 17	Unknown (perlecan fragment)	5.97 17100	Nt-Mi	DAPGQYGAYF Res. 4199–4207
	UBCS 25.6	Unknown	6.54 25635	Nt-Mi	XRQXTQIT(Q/F)
	UBCS 41	Unknown	5.25 41563	Nt-Mi	SXMERAPXRL

GM = gel matching with reference gels present in the SWISS-2DPAGE and SIENA-2DPAGE; Li = Liver, Br = Human breast ductal carcinoma; DL-1 cells = colorectal adenocarcinoma DL-1 cell line; Muscle mouse = gastrocnemius muscle; WB = Western blot; Nt-Mi = *N*-terminal microsequencing; Res = residue.

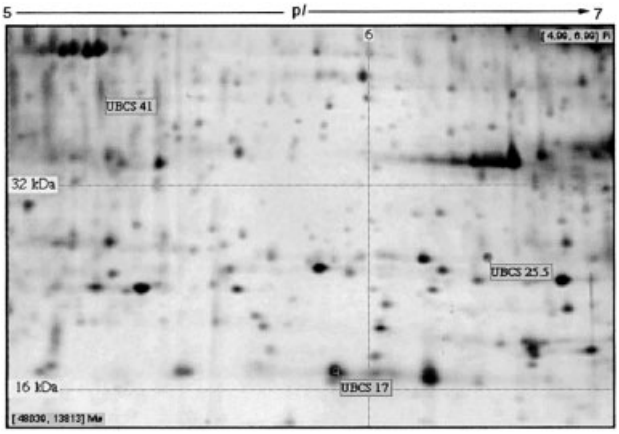

Figure 3. 2-D gel section of the 8701-BC cell conditioned medium, which includes the microsequenced spots labelled with the abbreviation UBCS as in Fig. 2.

unknown or modified proteins, were identified by means of *N*-terminal microsequencing, immune detection and gel matching with reference gels.

Two of the proteins with known identities have not been identified before in breast cancer proteomics: the first is the glycoprotein GC1QBP (MA32), known to bind the globular heads of C1Q, thus inhibiting C1 activation, but with unknown function *in vivo*; the second is UBL1 (ubiquitin thiolesterase L1), a thiol protease that removes ubiquitin from ubiquitinated proteins and prevents them from targeted degradation by proteasomes, first described as a neurospecific marker [21, 22]. A third protein spot, identified by microsequencing as GST, an enzyme that conjugates reduced glutathione to a wide number of exogenous and endogenous hydrophobic electrophiles, was not found to be coincident with the corresponding protein spot described in breast cancer tissue [23]. Indeed the latter migrated with a p*I*/M_r value of 5.40/23325, while on our map the value was 6.01/20635. This may indicate that the form identified on our map represents a variant of that previously described in breast cancer tissue, since genetic polymorphisms for GSTP1 in breast cancer have recently been demonstrated [24]. On the other hand, some authors have reported that the glutathione related

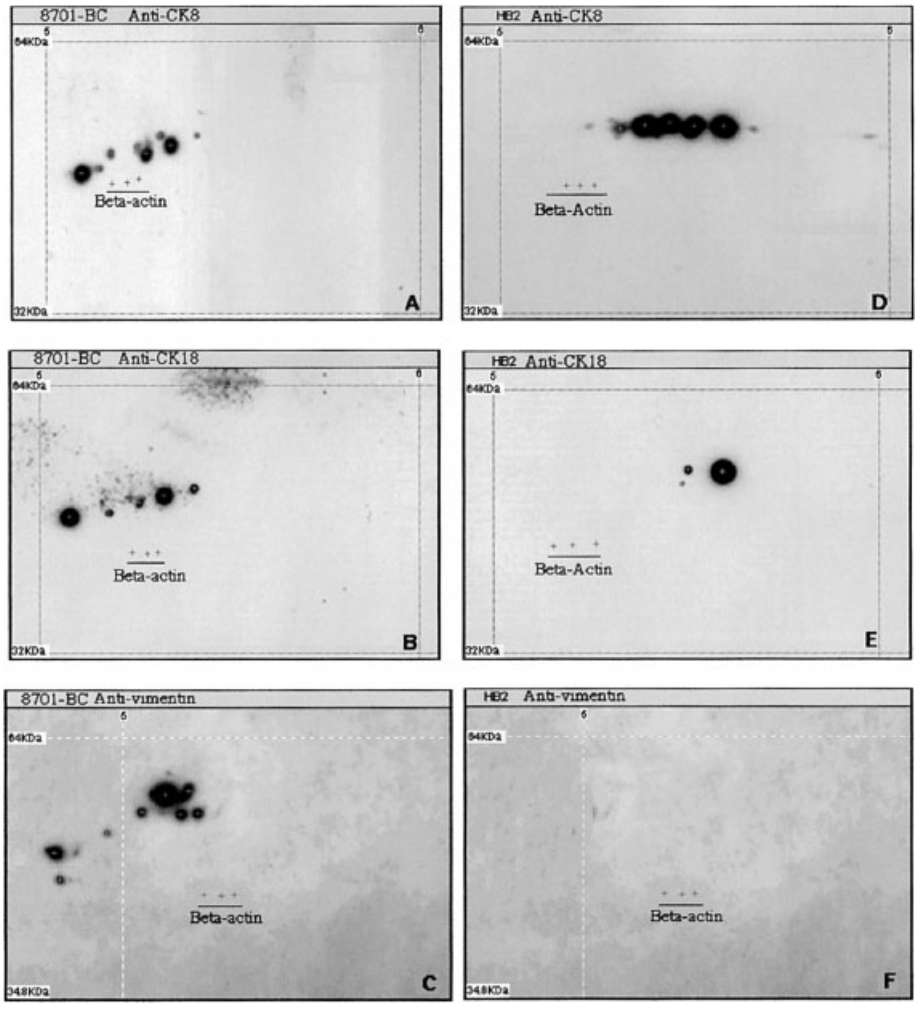

Figure 4. A panel of Western blots for the immune detection of cytokeratin-8, cytokeratin-18 and vimentin on 8701-BC cells (A, B, C) and on HB2 cells (D, E, F). The reaction to antivimentin was negative for the HB2 cells, under present conditions. Ninety μg of proteins was loaded on the IPG gels. The reaction was revealed by the method of ECL: the exposed films were digitised and then processed by Melanie 3. Each image of the panel represents the experimental windows of the western blot matched with the corresponding silver stained gels for the attribution of p*I*/M_r coordinates. The position of beta-actin is reported for reference.

detoxification system increases in human breast cancer in correlation with clinical and histopathological features [25], and therefore GSTP1 is considered as a potential marker for breast tumors.

Among the proteins with substitutions we found a sequence similar to thioredoxin, where the threonine in position 8 is substituted with alanine. Human thioredoxin is a small redox protein that is overexpressed in a number of human primary tumors, where it is associated with rapid cell proliferation and inhibits apoptosis. Recently the possible association of thioredoxin with p53-dependent function, including DNA repair in breast cancer, has been suggested [26].

In addition, two forms of annexin were also identified: one tagging with the full-length annexin I and the second starting from Res. 26 (p*I*/M_r 8.04 25752) and bearing two substitutions: Ser 27->Ile and Gly 30->Ile. As is known, annexin is a calcium/phospholipid-binding protein which promotes membrane fusion and is involved in exocytosis [27]. It is also known that the ability of annexin I to promote membrane aggregation is highly dependent on the structure of the *N*-terminal domain [28] which includes the major phosphorylation sites (namely Tyr-20, a site for phosphorylation by the epidermal growth factor receptor kinase and Thr-23/Ser-27, the sites for phosphorylation by protein kinase C *in vitro*). Therefore it can be deduced that this truncated form of annexin has a different function, which remains to be elucidated.

Finally, a truncated form of alpha-enolase (starting from residue 57, p*I*/M_r of 6.87/41563) and a shortened form of glyceraldehyde 3-phosphate dehydrogenase (p*I*/M_r 8.46/34490) starting with regular *N*-terminal residues have been identified. These forms have not been described before and, due to the multifunctional role of both enzymes [29, 30], the appearance of their derivative peptides in cancer cells is an additional finding of interest.

The protein spots released by the 8701-BC cells into the medium included two unknown sequences (UBCS) and a sequence matching residues 4199–4207 of the perlecan core protein (PGMB). A proportion of the expressed proteins in the proteomic map of 8701-BC appear to be the products of novel genes that are absent in the database, and therefore their inclusion in the database is a further contribution to the knowledge of breast cancer proteomics.

This work was supported by AIRC (to IPM 1999) and MURST (Cofin, 1999 to SM Prot. 9905208811-001). The authors wish to thank Prof. S. Feo and colleagues for providing anti-alpha-enolase antibodies and Dr. J. Taylor-Papadimitriou (Imperial Cancer Research Fund, London, UK) for providing the HB2 cells. Dr. S. Fontana and Dr. P. Cancemi are PhD students at the Department of Cell Biology and Development, University of Palermo.

Received November 8, 2001

5 References

[1] Donegan, W. L., *Cancer J. Clin.* 1997, *47*, 28–51.
[2] Emmert-Buck, M. R., Gillespie, J. W., Paweletz, C. P., Ornstein, D. K. *et al.*, *Mol. Carcinog.* 2000, *27*, 158–165.
[3] Perou, C. M., Jeffrey, S. S., van de Rijn, M., Rees, C. A. *et al.*, *Proc. Natl. Acad. Sci. USA* 1999, *96*, 9212–9217.
[4] Giometti, C. S., Williams, K. T., Tollaksen, S. L., *Electrophoresis* 1997, *18*, 573–581.
[5] Williams, K., Chubb, C., Huberman, E., Giometti, C. S., *Electrophoresis* 1998, *19*, 333–343.
[6] Rasmussen, R. K., Ji, H., Eddes, J. S., Zugaro, L. M. *et al.*, *Electrophoresis* 1998, *19*, 818–825.
[7] Le Naour, F., Misek, D. E., Melissa, C., Krause, M. C. *et al.*, *Clin. Cancer Res.* 2001, *7*, 3328–3335.
[8] Minafra, S., Morello, V., Glorioso, F., La Fiura, A. M. *et al.*, *Br. J. Cancer* 1989, *60*, 185–192.
[9] Minafra, S., Giambelluca, C., Andriolo, M., Pucci-Minafra, I., *Int. J. Cancer* 1995, *62*, 1–7.
[10] Luparello, C., Noel, A., Pucci-Minafra, I., *DNA Cell Biol.* 1997, *16*, 1231–1236.
[11] Bartek, J., Bartkova, J., B., Kyprianou, N., Lalani, E.N. *et al.*, *Proc. Natl. Acad. Sci. USA* 1991, *88*, 3520–3524.
[12] Bradford, M. M., *Anal. Biochem.* 1976, *72*, 248–254.
[13] Görg, A., Postel, W., Gunther, S., *Electrophoresis* 1988, *9*, 531–546.
[14] Bjellqvist, B., Pasquali, C., Ravier, F., Sanchez, J. C., Hochstrasser, D. F., *Electrophoresis* 1993, *14*, 1357–1365.
[15] Hochstrasser, D. F., Harrington, M. G., Hochstrasser, A. C., Miller, M. J., Merril, C. R., *Anal. Biochem.* 1988, *173*, 424–435.
[16] Hochstrasser, D. F., Frutiger, S., Paquet, N., Bairoch, A. *et al.*, *Electrophoresis* 1992, *13*, 992–1001.
[17] Moscato, S., Pratiesi, F., Sabbatini, A., Chimenti, D. *et al.*, *Eur. J. Immunol.* 2000, *30*, 3875–3884.
[18] Pucci-Minafra, I., Fontana, S., Cancemi, P., Alaimo, G., Minafra, S. *Ann. NY Acad. Sci.* in press.
[19] Taylor-Papadimitriou, J., Stampfer, M., Bartek, J., Lewis, A. *et al.*, J. Cell Sci. 1989, *94*, 403–413.
[20] Savagner, P., *Bioessays* 2001, *23*, 912–923.
[21] Day, I. N., Thompson, R. J., *FEBS Lett.* 1987, *210*, 157–160.
[22] Day, I. N., Hinks, L. J., Thompson, R. J., *Biochem. J.* 1990, *268*, 521–524.
[23] Bini, L., Magi, B., Marzocchi, B., Arcuri, F.S. *et al.*, *Electrophoresis* 1997, *18*, 2832–2841.
[24] Dialyna, I., Arvanitis, D. A., Spandidos, D. A., *Int. J. Mol. Med.* 2001, *8*, 79–87.
[25] Perquin, M., Oster, T., Maul, A., Froment, N. *et al.*, *J. Cancer Res. Clin. Oncol.* 2001, *127*, 368–374.
[26] Ueno, M., Matsutani, Y., Nakamura, H., Masutani H. *et al.*, *Immunol. Lett.* 2000, *75*, 15–20.
[27] Rothhut, B., *Cell Mol. Life Sci.* 1997, *53*, 522–526.
[28] Kusumawati, A., Liautard, J. P., Sri Widada, J., *Cell Biol. Int.* 2001, *25*, 809–813.
[29] Pancholi, V., *Cell Mol. Life Sci.* 2001, *58*, 902–920.
[30] Sirover, M. A., *Biochim. Biophys. Acta* 1999, *1432*, 159–184.