The Characterization of the Small Breast Epithelial Mucin (SBEM) Gene and Protein in Human Breast Cancer

By

Mark Mutawe

A thesis submitted to the Faculty of Graduate Studies in partial fulfillment of the requirements for the degree of

Master of Science

Department of Physiology
Faculty of Medicine
University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

The Characterization of the Small Breast Epithelial Mucin (SBEM) Gene and Protein in Human Breast Cancer

BY

Mark Mutawe

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

Master of Science

Abstract

Early detection of breast cancer is an effective means of reducing overall cancer mortality, as well as improving the quality of life for cancer survivors. There is a need to identify genes restricted to the mammary epithelium, which have great potential to report on proliferative changes in the breast. The small breast epithelium mucin (SBEM) is predicted to be a potential candidate for a biomarker, due to its breast-specific expression. To characterize the SBEM gene in breast cancer, normal human breast tissues and breast tumors were systematically examined. Using SBEM-specific primers, a PCR product was detected in all breast tumors and normal breast tissues examined by RT-PCR analysis. Using a ^{32}P-labeled SBEM cDNA probe, an expected 650-700 base transcript was detected in normal breast tissues and breast tumor tissues by Northern blot analysis. Levels of gene expression among tumor samples were variable, suggesting that the expression of the SBEM gene varies in breast cancer. Furthermore, a panel of breast cancer cell lines were examined. A PCR product corresponding to SBEM was detected in the majority of these cells. The BT-20 cell line expressed the greatest level of SBEM mRNA. As heterogeneity exists between between breast cancer cell lines, this differential SBEM mRNA expression was expected.

To evaluate the protein expression of SBEM in breast cancer, immunohistochemical analysis and Western blotting were performed. Immunohistochemical analysis revealed a variable pattern of SBEM protein expression among normal breast tissues and breast tumor tissues. As expected, the SBEM protein was localized to the cytoplasm and

membrane of positively staining ductal epithelial cells. A panel of SBEM antibodies was used to determine the expression levels of SBEM protein by Western blot analysis. No endogenous SBEM protein could be detected by Western blot analysis in any tissues or cell lines examined. This suggests that the level of SBEM protein was low or absent, or that the SBEM-antibodies could not recognize the SBEM antigen. SBEM protein was detected in two SBEM-expressing stably transfected cell lines, HEK-293 and MCF-7. Two sizes of protein were detected in both of these cell lines at 17 and 22kDa, larger than the predicted size of 11kDa. To investigate whether a secreted form of the SBEM protein exists, the culture media of stably transfected cell lines were examined. A broad diffused band ranging in size from 25-33kDa was detected in transfected HEK-293 cell culture media, while a narrow 26kDa band was detected in transfected MCF-7 culture media. The detection of secreted SBEM protein suggests that SBEM may also be secreted into the serum of breast cancer patients. SBEM is predicted to be glycosylated. Differences in glycosylation may give rise to tumor-specific antigens. To address whether SBEM protein was glycosylated, enzymatic deglycosylation assays were performed on the secreted SBEM protein and SBEM protein from stably transfected cell lysates. A 3kDa decrease in size was observed for the SBEM protein secreted from HEK-293 cells. This decrease was much less than expected. No change in size was detected for the SBEM protein secreted from MCF-7 transfected cells. Also, no decrease in size was detected in transfected cell lysates, suggesting that SBEM protein in these lysates were not glycosylated. Taken together, these data suggest that SBEM may be an excellent candidate for a breast cancer marker.

Acknowledgements

I would like to acknowledge my supervisors Dr. Etienne Leygue and Dr. Yvonne Myal for their mentorship and dedication. Without their help and patience, I never would have been able to prove myself worthy of such a high level of education.

I would like to thank my advisory committee members, Dr. Newman Stephens and Dr. Geoff Hicks for their insight and advice on my project. It has been an honour to have two such experienced and knowledgeable individuals on my committee.

I would like to extend special thanks to Mohammad Hamedani for his friendship, selflessness and perseverance. His knowledge and sense of humor made my journey through Graduate studies a memorable one. Many thanks to Shilpa Chooniedass-Kothari for all her help and advice. I am deeply grateful to Florent Hube for all his help with learning techniques. His tireless dedication to the SBEM project will always be remembered.

Last, but definitely not least, I would like to thank my parents for their wisdom, love and support through all the highs and lows. You've always been there for me and I can never thank you enough.

Abbreviations

Nonpolar Amino Acids:

A Ala Alanine
V Val Valine
L Leu Leucine
I Ile Isoleucine
P Pro Proline
F Phe Phenylalanine
W Trp Tryptophan
M Met Methionine

Polar / Uncharged Amino Acids:

G Gly Glycine
S Ser Serine
T Thr Threonine
C Cys Cysteine
Y Tyr Tyrosine
N Asn Asparagine
Q Gln Glutamine

Acidic Amino Acids:

D Asp Aspartic acid [Aspartate]
E Glu Glutamic acid [Glutamate]

Basic Amino Acids:

K Lys Lysine
R Arg Arginine
H His Histidine

α	Alpha
β	Beta
aa	amino acids
Ab	antibody
Ag	antigen

ATAC	Arimidex or Tamoxifen Alone or in Combination Trial
bp	base pairs
BSA	bovine serum albumin
°C	degrees centigrade
CCD	charge-coupled device
cDNA	complementary deoxyribonucleic acid
CM	complete medium
CO_2	carbon dioxide
cpm	counts per minute
dCTP	deoxycytidine triphosphate
ddH_2O	deionized distilled water
DCIS	ductal carcinoma *in situ*
DEPC	Diethyl Pyrocarbonate
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethylsulfoxide
dNTP	Deoxynucleotide Triphosphate
DTT	Dithiotreitol
E2	estradiol
EDTA	ethylene diamine tetraacetic acid
EGF	epidermal growth factor
EGTA	ethylene glycol bis(2-aminoethyl ether)-N,N,N'N'-tetraacetic acid
ER	estrogen receptor
EST	expressed sequence tag
FBS	fetal bovine serum
g / mg / µg / ng	grams/ milli- / micro- / nano-
GRB	gel running buffer
HCl	hydrochloric acid
kb	kilobase
KCl	potassium chloride
kDa	kilodaltons
KH_2PO_4	potassium phosphate
l / ml / µl	liter / milli- / micro-
M / mM / µM	molar / milli- / micro-
M_r	Relative molecular weight
MCID	MicroComputer imaging device
MOPS	morpholinopropanesulfonic acid
MUC1	mucin isoform 1
N_2	nitrogen
NaCl	sodium chloride
NaF	sodium fluoride
$NaHCO_3$	sodium bicarbonate
NaH_2PO_4	sodium phosphate monobasic
Na_3VO_4	sodium orthovanadate
OD	optical density

PAGE	polyacrilamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PR	progesterone receptor
RNA / mRNA	ribonucleic acid / messenger –
RPM	revolutions per minute
SAGE	serial analysis of gene expression
SBEM	small breast epithelial mucin
SDS	sodium dodecyl sulfate
SSC	sodium chloride / sodium citrate
SSPE	sodium chloride / sodium phosphate EDTA
TBE	tris boric acid EDTA
TBS	tris buffered saline
TBST	tris buffered saline + tween
Tris	tris (hydroxymethyl) amino methane
U	unit
V	volt
v / v	volume per volume
w / v	weight per volume
X-FBS	charcoal-stripped fetal bovine serum
1X	one times
10X	ten times
18S/28S	18S/28S ribosomal ribonucleic acid
^{32}P	Phosphorous-isotope 32
^{35}S	Sulphate-isotope 35

List of Figures

List of Tables

Table of Contents

I) Introduction

1) Breast Cancer

1.1. Breast Cancer Statistics

Breast cancer is the second leading cause of cancer-related deaths in women today (after lung cancer) and is the most common cancer among women, excluding non-melanoma skin cancers (www.imaginis.com). According to the World Health Organization, more than 1.2 million people will be diagnosed with breast cancer this year worldwide.

In the United States, approximately 215,990 women will be diagnosed with invasive breast cancer (Stages I-IV), while another 59,390 will be diagnosed with in situ breast cancer (an early form of the disease) (American Cancer Society – www.cancer.org).

In Canada in 2005, an estimated 21,600 women will be diagnosed with breast cancer and 5,300 will die from it (www.cancer.ca). One in 9 women is expected to develop breast cancer during her lifetime. One in 27 will die of it.

Each woman's risk may be higher or lower depending on several factors, including family history, age of menstruation, genetics and other factors. For example, younger women tend to have more aggressive breast cancers compared to older women. Conversely, the older a woman is, the greater her chances are of developing breast cancer. Approximately 77% of breast cancer cases occur in women over the age of 50. Ethnicity appears to be another factor, which dictates occurrence of breast cancer. In the United States, white and African-American women have the highest incidence of

invasive breast cancer, while Korean, American-Indian and Vietnamese woman have the lowest incidence (www.cancer.org).

1.2. Progession of Breast Cancer

Nearly all breast cancers begin in the epitheial cells of the ducts or lobules of the breast (Ma *et al.*, 2003). Tumorigenesis may be initiated when a phenotypically normal luminal epithelial cell enters a premalignant stage of growth called atypical intraductal hyperplasia (ADH) [Figure 1]. Participant ADH cells then progress into a proliferative phase giving rise to the malignant, but preinvasive lesion called ductal carcinoma in situ (DCIS). The growing tumor may then degrade basement membranes and become invasive, moving into adjacent tissues (invasive ductal carcinoma or IDC). At this point, it can metastasize, spreading to other parts of the body through the lymphatic system and bloodstream.

1.3. Screening

Primary prevention through screening, namely breast self-examination, clinical breast examination (CBE) and screening mammography, is believed to be the most promising method for controlling and treating breast cancer at its early stage of onset (Vahabi, 2003). The purpose of screening for breast cancer is to reduce mortality from the disease through the early identification and application of therapies which have a better chance of success than those applied following a clinical diagnosis.

Breast self-examination involves regular examination of the breasts (i.e., monthly) by the individual. It is an inexpensive and non-invasive procedure for early detection of

abnormalities in breasts. However, there are limitations, where the effectiveness depends on the proficiency of the examiner. Even with the best of training, it is difficult for the individual to detect small size tumors (less than 1cm), which are more responsive to available cancer treatments. The average sized tumor detected by breast self-examination is approximately 2.7cm and most often by this point, the tumors have metastasized to the axillary nodes. Though very useful as an initial screening procedure, the majority of tumors detected by breast self-examination are associated with a lower long-term survival.

CBE refers to regular examination (i.e., once a year) of the breasts by a health professional (i.e., physician or nurse). The method is simple, inexpensive and is widely recommended for breast cancer screening. Similar to breast self-examination, the accuracy and effectiveness are dependant on the proficiency of the examiner. Some studies have shown that, CBE, when conducted properly and for an adequate length of time, can detect up to 50% of asymptomatic cancers and may therefore be valuable in reducing breast cancer mortality (Vahabi, 2003). Although, like with breast self-examination, no matter how well trained the examiner may be, CBE is sensitive to tumor size. The average size of tumors detected are 2.1cm in diameter and axillary node metastases are usually already present in about 27% of cases (Vahabi, 2003).

Mammography remains the cornerstone of screening, with technologies such as ultrasonography and magnetic resonance imaging (MRI) having an increasingly defined role. Screening mammography is a low-dose X-ray of the breast and involves pressing the breast firmly between two plastic plates in order to provide a good image of all breast tissue. Mammography is believed to be a more effective method than CBE or breast self-

examination in reducing breast cancer mortality as small breast masses (1.1cm) can be detected. Studies have shown that early detection of breast cancer by mammography reduces breast cancer mortality by 30-40% among women ages 50 and over (Elmore *et al.*, 2005), however, the effectiveness of screening among women younger than 50 years old is controversial. The efficacy of breast screening is based largely on the fact that as a woman's age increases, breast parenchymal density decreases, allowing cancers to be more easily detected in older women's breasts because they are radiolucent (permit the penetration and passage of X-rays or other forms of radiation). Increased breast density is associated with a reduction in the sensitivity of mammography with small or subtle cancers being particularly difficult to detect (Dixon, 2003). As a consequence, cancers could potentially be missed by screening in women with dense breasts, such as pre-menopausal women.

1.4. Breast Cancer Therapies

1.4.1. Hormonal Therapy

Due to increased awareness and numerous screening programs, more and more patients are being diagnosed at earlier stages and therefore are treated with a greater success (Jones and Buzdar, 2004). Advances in the multidisciplinary approach to treatment (chemo/hormonal therapy, surgery and radiation) have enabled practitioners to treat patients better and improve outcomes.

The standard for the past two decades has been the use of the anti-estrogen, tamoxifen, in hormone-sensitive breast cancer patients (Estrogen receptor (ER) and/or Progesterone

receptor (PR) positive). In 1986, tamoxifen (Nolvadex) was the first drug of its kind approved as monotherapy for the treatment of early breast cancer in node-positive, post-menopausal women. Tamoxifen is a molecule that competes with estrogen for binding sites on estrogen receptors. This competition with estrogen limits cell growth in some tissues, making it an effective cancer treatment in some cases. Tamoxifen use is not ideal however, for example among women with ER-positive breast cancer, only 40-50% of patients benefit from treatment, suggesting that a substantial fraction of ER-positive tumors are resistant to this drug (Hayes, 2004). Tamoxifen has also been found to cause common side-effects and life-threatening toxicities such as endometrial cancer, thromboembolic events and uterine sarcoma (Fisher et al., 1996; Wickerham et al., 2002). The undesirable estrogenic activity of tamoxifen has fueled the search for better selective ER modulators (SERMs). Toremifene (Fareston) has been used for the treatment of women with hormone receptor-positive advanced breast cancer (Milla-Santos *et al.*, 2001). Raloxifene has been used for the prevention and treatment of post-menopausal osteoporosis. Raloxifene (Evista) was reported to reduce the risk of invasive breast cancer by 72% and the risk for ER-positive breast cancer by 84% (Cummings et al., 1999). Moreover, raloxifene has not been shown to increase risk of endometrial cancer or vaginal discharge/bleeding.

Another way to prevent estrogen from binding to its receptors in breast tumors is to actually prevent the formation of estrogen. A class of drugs called aromatase inhibitors (AI) block the final step in the conversion of androgen to estrogen and have been extensively used in the treatment of postmenopausal women with advanced breast cancer (Buzdar, 2002). If no estrogen is produced, it cannot bind to the receptors and promote

tumor growth. These AIs can be divided into non-steroidal (anastrazole and letrozole) and steroidal (exemestane) inhibitors. Anastrazole and letrozole have been shown to have a higher efficacy than tamoxifen as first-line therapy for patients with advanced ER-positive tumors (Bonneterre *et al.*, 2001, Mouridsen et al, 2001). Based on the results of the ATAC (Arimidex or Tamoxifen Alone or in Combination) trial, anastrazole appears to be superior to tamoxifen in the adjuvant setting for disease-free survival and time-to-recurrence, particularly in ER-positive patients, and in reducing the incidence of distant metastases and contralateral breast cancer (Howell et al., 2005). The use of letrozole has been shown to give a significant improvement in disease-free survival, a substantial reduction in the rate of distant metstases and a decrease in the rate of death due to breast cancer was shown in postmenopausal women who had been treated with letrozole for early-stage breast cancer, following tamoxifen therapy (for 4.5 to 6 years) [Goss et al., 2003]. Another study showed that first-line letrozole extended the time for tumor progression, compared to tamoxifen alone or in combination with tamoxifen (Long *et al.*, 2004). AI treatment does not appear to benefit patients with ER-negative tumors. In terms of toxicity, side-effects of AIs are associated with estrogen withdrawl and include hot flushes, arthraglia and bone demineralization.

1.4.2. Chemotherapy

There is overwhelming evidence that chemotherapy produces a significant reduction in cancer recurrence and death. Chemotherapy appears to be more effective for women younger than 50 years of age, than for women older than 50 years (Hortobagyi, 2001). Most chemotherapy regimens used in the adjuvant setting contain an alkylating agent,

which kills cells by directly attacking DNA. These drugs produce permanent pre-mature menopause in more than two thirds of pre-menopausal patients (Hortobagyi, 2001). Thus, chemotherapy has both a cytotoxic effect and an endocrine effect in pre-menopausal patients, whereas the endocrine effect would be absent in postmenopausal women (Bonadonna, 1978; Bonadonna and Valagussa, 1986).

Chemotherapy regimens are designed to kill cancer cells throughout the body. It has advantages for nearly every breast cancer patient regardless of whether the cancer is hormone receptor-positive or negative. Regimens currently used are, cyclophosphamide/methotrexate/5-fluorouracil[CMF], cyclophosphamide/doxorubicin/5-fluorouracil [CDF] or similar epirubicin-containing combinations. These regimens are administered for 6 cycles over a 4-6 month period. Two commercially available taxanes, paclitaxel and docetaxel, are effective agents against metastatic breast cancer (Hortobagyi, 2001). These agents are used individually, or in combination with other treatments. Patients who develop metastatic disease (ie, who relapse at distant sites) are generally not curable. Combination therapies, however, are often effective at shrinking tumors and improving quality of life and may even be improving survival rates.

1.4.3. Combination Hormonal Therapy and Chemotherapy

Combination chemotherapy has been found to be significantly more effective than single-agent chemotherapy (Rivkin et al., 1996). Chemotherapy given in combination with tamoxifen was shown to be more effective than tamoxifen alone for patients with node-positive, endocrine-responsive breast cancer (Colleoni *et al.*, 2005). The addition of tamoxifen to adjuvant chemotherapy significantly increases the disease-free and overall

survival rates of patients with hormone receptor-positive tumors. This effect is observed regardless of age or lymph node status. The treatment of choice for these hormone receptor-positive women is the combination of chemotherapy followed by 5 years of tamoxifen therapy (Hartman et al., 2001).

1.4.4. Monoclonal Antibody Therapy

In developing novel anti-cancer agents, the goal is to target specific lesions within tumor cells, leading to improved cure rates and reducing cytotoxicity in normal cells (Esteva and Hortobagyi, 2004). Monoclonal antibody therapy is a form of passive immunotherapy because the antibodies are made in large quantities outside the body (in a lab) rather than by a person's immune system. These treatments do not require the person's immune system to take an "active" role in fighting the cancer. The successful use of trastuzumab/Herceptin, the only approved monoclonal antibody for HER-2 (erbB2) over-expressing metastatic breast cancer, provided the proof of principal that targeting specific receptors results in clinical benefit. HER-2 over-expression is exhibited in 30% of women with breast cancer (Munster and Norton, 2001). Herceptin can be administered multiple times to patients with breast cancer without the risk of resistance developing the potential for immunogenicity is decreased, and the potential for recruiting immune effector mechanisms is increased. Treatment with Herceptin results in clinical benefits, whether given as a single agent or in combination with chemotherapy (Suter *et al.*, 2004). In combination with chemotherapy, improvements in response rate, time to progression, and survival occur in women with HER-2 over-expressing metastatic breast cancer (Keshgegian and Cnaan, 1995; Veronese SM *et al.*, 1993).

1.4.5. Vaccines

In cancer therapy, potential roles for vaccines include passive or adoptive immunotherapy and active specific immunotherapy. The potential advantages of therapeutic vaccines for cancer are that they can augment an established immunogenic response to the tumor, they target specific tumor antigens, they are non-toxic and can be combined with conventional therapies and/or immunotherapies, and they elicit immunogenic memory to prevent re-emergence of the tumor. Active immunization offers multiple theoretical advantages over all other therapies including low toxicities and exquisite specificity (Emens et al., 2005). Several vaccines have been developed and are actively being studied, though it is unclear whether vaccines can actually prolong survival. Examples of currently examined vaccines are MUC-1 peptide-KLH (keyhole limpet hemocyanin), anti-p53 and HER-2 peptide (Reddish et al., 1998, Svane et al., 2004 and Salazar et al., 2003). Additional vaccine platforms are under development, including plasmid DNA-based vaccines, vaccines comprised of recombinant viral vectors or recombinant bacteria incorporating tumor antigens or heat-shock protein-based vaccines.

2) Tumor Markers for Breast Cancer

2.1. Predictive vs. Prognostic Tumor Markers

Tumor markers are substances, usually proteins, that are produced by the body in response to cancer growth or by the cancer tissue itself (Hayes et al., 2001). Detecting

and/or monitoring these changes is expected to assist in evaluating cancer risk, diagnosis, prognosis or response to treatment (Hayes et al., 1996).

A prognostic marker is an indicator of the natural history of the disease, and it is used to help define patients with high and low risks of death that result from the inherent heterogeneity of the disease process. Prognostic markers are associated with risk of metastases, death unrelated to systemic treatment, and are usually markers of cellular proliferation, invasion and angiogenesis (Cianfrocca and Goldstein, 2004). Prognostic markers can also help in identifying women whose prognosis is so poor with conventional approaches as to warrant consideration of more aggressive therapies. These markers are either chronological (indicators of how long the cancer has been present (i.e. tumor size)) or biological (indicators of the metastatic potential behaviour of a tumor (i.e. tumor-grade or axillary node status)). McGuire WL (1991) proposed a set of criteria for a prognostic marker to have clinical relevance; the marker must have biological relevance, be reproducible in different laboratories, be validated in a large series of patients, be confirmed independently by other workers and have cut-off levels, which are optimized.

A predictive marker is an indicator of response to therapy, preferably defining a patient's survival after treatment (Duffy, 2005). Predictive markers are important in oncology as different cancers vary in their response to particular therapies (such as the estrogen receptor (ER) for endocrine therapy or HER-2 protein for Herceptin antibody therapy). These markers may also be epiphenomena that are related to the mechanism of therapy. For example, several studies have suggested that ER status is associated with lower response to chemotherapy (Hayes, 2003). Biologically, it is more likely that ER is

not a direct target of chemotherapeutic agents, but rather, the expression of ER is related to other mechanisms of resistance to chemotherapy.

Markers may also have both prognostic and predictive features. Furthermore, some markers may have diametrically opposing effects. For example, a marker may have a favourable prognostic value, but an unfavourable predictive value, or vice versa. Complicating the issue even further, a marker may be a positive predictive marker for one treatment, but a negative predictive factor for another. The estrogen receptor (ER) is a good example of these concerns. ER is a weak positive prognostic marker and a strong predictive marker for endocrine therapy, but may be a modest negative predictive factor for chemotherapy (Hayes, 2003). Another example is HER-2 (human epidermal growth factor receptor-2), where the over-expression of HER-2 appears to be a weak prognostic marker, a modest negative predictive marker for endocrine therapy, and a positive predictor for anthracyclines (Yamauchi *et al.*, 2001). Thus, trials, wich do not take the precise types of treatments into account, are likely to result in inconsistent conclusions.

2.2. Steroid Receptors

2.2.1. Estrogen Receptor

The estrogen receptor (ER) is a member of the nuclear hormone receptor superfamily that mediates the pleiotropic effects of the steroid hormone estrogen in a diverse range of developmental and physiological processes (Mangelsdorf et al., 1995). The ER is a transcription factor, which modulates transcription of specific genes by directly interacting with estrogen response elements located in the promoter of target genes, as

well as in a variety of additional modes including indirect tethering to DNA via interaction with other transcription factors (Martini and Katzenellenbogen, 2003). Two estrogen receptors have been identified, ERα and ERβ, which are encoded by two independent genes (Mosselman et al., 1996). ERα, which is found on the long arm of chromosome 6, encodes for a 65kDa protein. ERα has been found to be expressed in 15-30% of luminal epithelial cells and not in any of the other cell types within the human breast (Osborne, 1998). ERβ, found on the long arm of chromosome 14, encodes a 477 amino acid 54kDa protein. ERβ has been detected in most luminal epithelial and myoepithelial cells, as well as in fibroblasts and other stromal cells within the normal human breast (Speirs et al., 2002). Structurally, the most conserved domains between the two isoforms are the DNA-binding domain and ligand-binding domain, exhibiting 97% and 60% homology respectively. Although both ERα and ERβ display similar binding affinities for estrogen (E2), they have different roles in the regulation of gene expression.

It has been documented that ERα is expressed in a subset of normal breast epithelial cells, and ERα-containing epithelial cells do not normally proliferate in response to estrogen (Clarke et al., 1997). In contrast to the normal breast, most pre-malignant breast lesions express high levels of ERα. These ERα-expressing breast cancer cells are hormone-dependant and undergo regression when estrogen activity is decreased by aromatase inhibitors or treatment with anti-estrogen, tamoxifen (Fuqua, 2001). Detectable levels of ERα have been observed in 50-70% of invasive breast cancer cases. (Clark et al., 1984, Duffy et al., 2005). ERα expression is increased at the very early stages of ductal hyperplasia and increases further with increasing atypia such that most cells in

atypical ductal hyperplasias and DCIS of low or intermediate grade contain ERα (Allred et al., 2001).

Estrogen-receptor positive cells grow more slowly than receptor negative cells. Women have a better prognosis if their tumors are receptor-positive because these cells grow more slowly than receptor-negative cells and they have more treatment options. Most endocrine therapies for breast cancer are designed to either reduce circulating estrogen levels, which would then reduce the number of bound and activated ERs in a breast tumor, or to bind directly to the available ER in a tumor, rendering them non-functional. The goal of anti-estrogens such as tamoxifen or aromatase inhibitors, is to antagonize ER-mediated transcriptional activation of genes required for tumor growth. Reducing estrogen levels or altering the activity of the receptor has been shown to induce regression of these cancers.

ERα regulation is not limited to direct ligand binding, as it may also be modulated by other pathways. These include the epidermal growth factor (EGF) and insulin-like growth factor-1 (IGF-1) mitogenic pathways and the second messengers like cyclic AMP (cAMP) and dopamine (Driggers et al., 2002). Activation of these pathways influences ER transcriptional activity either by targeting the receptor directly or by regulating the activity of receptor co-regulators. For instance, activation of the epidermal growth factor receptor (EGFR) by EGF leads to phosphorylation of Ser118 residue in the ERα by p44/p42 mitogen-activated protein kinase (MAPK), which in turn recruits the co-activator p68/p72 and activates target gene transcription in a ligand-independent manner (Kato et al., 1995). Thus, there appears to be a cross-talk between ligand-dependant and ligand-independent pathways.

Numerous studies have shown that ERα status in tumors correlates with various tumor characteristics (Osborne, 1998). Nearly 80% of post-menopausal women are ER-positive, while only about 50-60% of pre-menopausal patients are ER-positive (Osborne, 1998). In terms of histological characteristics, low-grade tumors are nearly always ER-positive. In contrast, tumors demonstrating histologic evidence of poor tumor differentiation such as high histological or nuclear grade, extensive necrosis, little or no elastosis, and a prominent lymphoid infiltrate, are frequently ER-negative.

The role of ERβ in breast cancer is not well documented. There are some studies showing that ERβ expression is down-regulated in lesions such as atypical ductal hyperplasia and ductal carcinoma in situ, when compared with that in normal breast epithelium (Roger et al., 2001). ERβ has been found to be expressed in 40-70% of invasive breast cancers (Speirs V et al., 2002). ERβ has also been suggested to negatively modulate the effects of ERα. In terms of predicting response to hormonal therapy, a preliminary report showed that ERβ was produced in higher amounts in tamoxifen-resistant than in tamoxifen-sensitive cancers (Speirs *et al.*, 1999). However, other studies have shown ERβ production to be associated with favourable response to adjuvant tamoxifen therapy (Esslimani-Sahla *et al.*, 2004). The usefulness of ERβ as a clinical prognostic marker remains to be defined.

2.2.1.1. Estrogen Receptor-α as a Prognostic Marker

Over 25 years ago, ER status was first reported to correlate with prognosis of patients with primary breast cancer (Knight et al., 1977). ER-positive patients were found to have longer disease-free and overall survival compared to those with ER-negative status. Other

studies confirmed that patients with ER-negative tumors are more likely to suffer early recurrence than those with ER-positive tumors (Osborne, 1998). However, another report suggests that after the initial few years, in which a patient enjoys a lower relapse rate, the relapse rate in ER-positive patients increases, and thus the prognostic significance of ER disappears (Osborne, 1998). Thus, ER may not be a strong prognostic marker.

2.2.1.2. ER-α as a Predictive Marker

ER status is documented to correlate with response to a variety of hormonal therapies (Osborne et al., 1998). Approximately 60% of previously untreated patients with ER-positive tumors benefit from hormonal therapy, compared to only 5-10% of those with ER-negative tumors. Patients with little or no ER show little or no benefit from adjuvant tamoxifen, while those that express ER have longer disease-free and overall survival time. For example, the Early Breast Cancer Trialists Collaborative group illustrated that 5 years of tamoxifen led to proportional reductions in the risk of recurrence and mortality of 47% and 26%, respectively, of patients with ER-positive tumors (Early Breast Cancer Trialists Collaborative Group, Lancet 1998). Thus, the major clinical utility of measuring ER in breast cancer is to predict response to hormonal therapy. Patients whose tumors are ER-negative are not candidates for hormonal therapy, but are best treated with chemotoxic chemotherapy.

2.2.2. Progesterone Receptor

The progesterone receptor (PR), like the ER, is a member of the nuclear receptor superfamily. The two isoforms of the PR; PR-A and PR-B, are encoded by the same

gene, utilizing two distinct transcriptional start sites and yielding proteins that differ with regards to their amino terminal regions and biological activities. PR-B contains an additional 164 amino acids at its N-terminal region, but otherwise the two proteins are identical (Clarke and Sutherland, 1990). Both isoforms are highly expressed in normal tissues, but PR-B levels are reportedly elevated in breast cancer. Hopp et al. (2004), reported that patients with high PR-A:PR-B ratios in their breast cancers responded poorly to adjuvant therapy. A functional difference between isoforms A and B is that PR-A can act as a dominant repressor of both PR-B and ER in a promoter and cell type-specific manner (Vegeto et al., 1993). Interestingly, more recent findings indicate that unliganded PR can regulate gene transcription, with PR-A the more active isoform (Jacobsen *et al.*, 2003).

PR status, like ERα, is a good predictor of responsiveness to therapy. Nearly 50% of all ER-positive (+) tumors also are reported to be PR (+) and approximately 75% of these ER+/PR+ tumors respond positively to endocrine therapy (Keen and Davidson, 2003). In fact, a recent study showed that the combined measurement of ER and PR is superior to ER alone in predicting benefit from adjuvant anti-estrogen therapy (Bardou *et al.*, 2003). ER+/PR- tumors are reported to be less responsive to anti-estrogen therapy, perhaps suggesting that PR may be necessary for positive therapeutic outcomes with hormone therapy. Alternatively, because ER is a key transcription factor for the activation of PR, lack of PR expression in these ER+/PR- tumors could also suggest that the estrogen response pathway may be non-functional (Graham et al., 1995). Only a small fraction of tumors are ER-/PR+ (< 5%) and they demonstrate and intermediate response to endocrine

therapy (Ravaioli *et al.*, 1998). Overall, the presence of PRs and the ratio of the two isoforms are believed to directly influence tumor differentiation.

2.3. ErbB Receptors

2.3.1. ErbB Receptor Family

The ErbB receptor family includes four members, EGF receptor (also known as ErbB-1/HER1), ErbB-2/Neu/HER2, HER3/ErbB-3 and ErbB-4/HER4, which belong to the type I receptor tyrosine kinase family (Normanno et al., 2003). ErbB receptors and their ligands are frequently expressed in human carcinomas and therefore might play an important role in the disease. EGFR (and other members of the receptor family) consists of an extracellular ligand-binding domain, a hydrophobic transmembrane region, and a highly conserved tyrosine kinase domain. Upon binding of ligands, the receptor undergoes dimerization, leading to the autophosphorylation of the receptor itself and the phosphorylation of cellular substrates. These phosphorylated substrates, which represent the activated, downstream signal transduction pathways, lead to changes in gene expression that may modify proliferation or cell-survival mechanisms (Morris, 2002).

2.3.2.1 HER-1/Epidermal Growth Factor Receptor

The epidermal growth factor receptor (EGFR) was the first receptor protein-kinase to be sequenced (Ullrich *et al.*, 1984). EGFR is expressed by many epithelial tumors, including a proportion of breast cancers, implying the involvement of EGFR signaling is involved (Benz *et al.*, 1993; Pietras *et al.*, 1995). This expression has been associated

with increased proliferation, disease progression, and a poor prognostic outlook (Quenel et al., 1995; Konecny *et al.*, 2003). EGFR has also been correlated with decreased ER expression and increased resistance to endocrine therapy (Knowlden *et al.*, 2003).

EGFR vIII, a constitutively activated variant of EGFR, is the most commonly detected mutant detected in human solid tumors (Mass, 2004). EGFR vIII is highly prevalent in invasive breast cancer, while it is absent in normal breast. However, EGFR vIII expression is extremely rare and is not thought to contribute to the malignant phenotype (Rae et al., 2004).

2.3.2.2. HER-2/*neu* (erbB2)

The human epidermal growth factor receptor 2 (HER-2)/*neu* (c-ERB-2) gene is localized to chromosome 17q and encodes a 185kDa transmembrane tyrosine kinase receptor protein, which is expressed at low levels in epithelial and myoepithelial cells of normal breast tissue (Bundred *et al.*, 2001). While no known ligand for the HER-2 receptor has been identified, it is the preferred dimerization partner of the other family members. The formation of heterodimers and the ensuing activation are temporary and spatially controlled in normal cells and tissue, but dysregulation of this network has been reported in cancer cells, where increased expression of HER-2 is seen (Liu *et al.*, 1995). Indeed, over-expression of HER-2 occurs in about 30% of breast cancers and is associated with a poor prognosis (Munster and Norton, 2001; Suter *et al.*, 2004). High levels of HER-2 are found in high-grade ductal carcinoma in situ (DCIS), but low or undetectable levels are seen in low grade DCIS.

The prognostic value of HER-2 amplification/over-expression in node-positive tumors is widely established, however, there is no consensus on its value in node-negative cases (Volpi *et al.*, 2000; Cooke *et al.*, 2001). HER-2-positivity has been related to endocrine therapy unresponsiveness, even in hormone receptor-positive patients (Elledge *et al.*, 1998). The observation that the level of ER-α is inversely correlated with HER-2 expression, together with clinical data indicating that only high ER-expressing tumors are sensitive to anti-estrogen reagent tamoxifen, likely explain the tamoxifen-resistance of HER-2-positive tumors (Konecny *et al.*, 2003). HER-2 gene amplification and protein over-expression has also been associated consistently with DNA aneuploidy, high cell proliferation rate, increased mitogenesis, p53 mutation, increased cell motility, topoisomerase II alpha amplification and alterations in a variety of other molecular biomarkers of breast cancer invasiveness and metastasis (Masood and Bui, 2002; Eccles, 2001).

A humanized murine monoclonal antibody to HER-2, called Herceptin (Trastuzumab), has been developed, which specifically binds the extracellular portion of HER-2. Since 1998, Herceptin has become an important therapeutic option for patients with HER-2 over-expressing breast cancer. Herceptin is widely used for its approved indication as a treatment for advanced metastatic disease and is also studied as an adjuvant treatment for earlier stage disease and in neo-adjuvant treatment protocols (Hortobagyi, 2001).

2.4. p53

Alterations in the p53 gene, which is found on the short arm of chromosome 17, are the most common genetic changes found in human malignancies, including breast cancer

(Velculescu et al., 1996). The p53 gene encodes for a 393 amino acid nuclear phosphoprotein with a molecular mass of 53kDa. The function of the p53 nuclear phosphoprotein is not fully understood, but there is evidence to suggest that it plays the part of a tumor suppressor gene by controlling transcription, inhibiting cell cycle, promoting apoptosis and inhibiting angiogenesis (Vogelstein *et* al., 2000). Somatic cell mutation in the p53 protein is observed in approximately 20-30% of primary breast carcinoma cases (Sullivan *et al.*, 2002). Although these mutations are scattered throughout the entire gene, the majority of mutations are confined to a 200-amino acid span containing 1 of 4 conserved core domains, which result in decreased DNA-binding affinity and decreased gene transactivation. Tumors with p53 mutations are more likely to be highly invasive, poorly differentiated, high-grade breast tumors. Poorer overall and disease-free survival have also been reported in patients with p53 gene mutations (Lai H et al., 2004). It has been hypothesized that p53 mutations occur late in the progression of breast cancer and tend to correlate with the development of tumors with fully malignant and invasive phenotypes (Bosari et al., 1993). Thus, mutant p53 has been suggested to be a marker predicting risk for subsequent breast carcinogenesis.

Germline p53 mutation also serves as a risk factor for breast carcinoma development as part of the Li-Fraumeni syndrome (LFS). p53 mutations have been identified in nearly 60% of families with this syndrome, suggesting that loss of p53 may be a critical parameter in the development of multiple carcinomas.

2.5. BRCA1 and BRCA 2

Approximately, 5% of breast cancers show a familial pattern of occurrence (Yoshida and Miki, 2004). Numerically, the most important of these genes are BRCA1 and BRCA2 (Eeles, 2000). The BRCA1 gene is located on the long arm of chromosome 17 and encodes a 1863 amino acid-containing protein. The BRCA1 protein appears to have multiple functions, being involved in regulating transcription, inhibiting cell proliferation and repairing DNA (Eeles, 2000). Over 100 different mutations have been found in this gene and most of these mutations result in a truncated protein due to a frameshift or non-sense mutation. BRCA 2 is a larger gene found on chromosome 13, which encodes for a protein of 3418 amino acids (Wooster *et al.*, 1995), appearing to have similar function to BRCA1.

Germline mutations of BRCA1 and BRCA2 are responsible for 80-90% of all hereditary breast cancers (Dumitrescu and Cotarla, 2005). Interestingly, although these genes are important in hereditary breast carcinoma, they have not been found to be associated with development or progression of sporadic cancer. Tumorigenesis in women with BRCA 1 or 2 mutations requires loss or inactivation of the remaining wild-type allele, resulting in expression of a non-functional protein and a loss of cell cycle control and DNA repair mechanisms.

Interestingly, one study has found that 80% of BRCA1 mutated tumors are ER-negative, whereas more than 75% of BRCA2 mutated tumors are ER-positive (King et al., 2001). This suggests that anti-estrogen (eg. tamoxifen) treatment would be beneficial to BRCA2 mutation carriers, but not to those carrying BRCA1 mutations.

Genetic testing is available for BRCA 1 and 2 mutations. The first step in testing involves genetic counseling. In this session, a counselor will discuss the tests, possible risks (such as psychological effects, employer concerns, insurance difficulties), and what a positive or negative result means. During the second session, the blood test and genetic analysis will be performed. At the third session, the results of the test will be discussed with the patient.

2.6. Mammaglobin

Human mammaglobin was first identified in 1996 by Watson and Fleming (Watson and Fleming, 1996), using differential display analysis, as a breast-specific member of the uteroglobin family. Mammaglobin mRNA and protein have been found to be over-expressed in breast cancer and are candidate diagnostic markers for breast cancer (Watson and Fleming, 1996). Using RT-PCR, mammaglobin mRNA was found to be expressed in 60-80% of human breast cancers (O'Brien et al., 2002). Although mammaglobin mRNA has been detected in both normal and malignant breast tissues, relative levels were 10-20 fold higher in the carcinomas than in normal breast tissue (O'Brien et al., 2005). Mammaglobin has been found to be secreted from both, established breast cancer cell lines and primary breast carcinoma cells cultured *in vitro* (Fanger *et al.*, 2002). It has also been detected in sera of patients with breast cancer and may provide a rapid screening test for the diagnosis and management of breast cancer (Fanger *et al*, 2002). A recent study investigating the use of mammaglobin as a vaccine suggested that mammaglobin may have immunotherapeutic potential for the treatment and/or prevention of breast cancer (Narayanan *et al.*, 2004)

A correlation between high mammaglobin levels and the presence of ER and PR, low Ki-67 labeling index (See p26), low nuclear grade and the absence of axillary lymph nodal invasion has been found (Nunez-Villar *et al.*, 2003). Interestingly, these findings indicate that mammaglobin expression in breast cancer is associated with a less aggressive tumor phenotype. In addition, mammaglobin has been successfully used to detect breast cancer cells in axillary lymph nodes (Leygue *et al.*, 1999).

2.7. Prolactin-Inducible Protein

The human prolactin-inducible protein (PIP), also known as gross cystic disease fluid protein-15 (GCDFP-15), is a predominant secretory protein in various body fluids, including saliva, milk and seminal plasma (Haagensen *et al.*, 1979, Tian et al., 2004). PIP has been found to be expressed in more than 90% of human breast cancer biopsies, but not in normal mammary gland (Myal *et al.*, 1998). PIP expression is often conserved in corresponding lymph node metastases.

While PIP expression is found to occur in tumors arising from skin and salivary gland, distinction from breast cancer is rarely an issue and PIP protein has already found application as a marker for the recognition of breast cancer of metastatic cancer (Clark *et al.*, 1999). In human breast cancer cell lines, the gene is up-regulated by a number of hormones including androgen and prolactin (Carsol *et al.*, 2002).

2.8. Plasminogen activators and inhibitors

The urokinase-type plasminogen activator (uPA) is a serine protease, which plays an important role in the invasion and metastasis process through degradation of the

extracellular matrix (Esteva and Hortobagyi, 2004). High levels of tissue uPA and its inhibitors (PAI-1 and PAI-2) have been correlated with poor outcome in node-negative breast cancer patients, independent of HER-2 status (Janicke et al., 1993).

2.9. Ki-67

Tumor proliferation rate is an important prognostic factor in breast cancer, though it is not breast cancer-specific. Ki-67 is a nuclear antigen found in cells in the proliferative phases of the cell cycle, but not cells in the resting phase (G_0 phase). A strong correlation has been shown between the percentage of cells showing Ki-67 staining and the nuclear grade, age and mitotic count (Sahin et al., 1995). Patients whose tumors cells express more than 50% of the Ki-67 staining are at high risk of developing recurrent disease (Veronese et al., 1993). In terms of relation to ER and HER-2 expression in tumors, Ki-67 labeling index has recently been shown to be high in ER(-)/HER-2(-) tumors, whereas ER(+)/HER-2(-) tumors had the lowest labeling index (Umemura *et al.*, 2005).

2.10. Mucins

2.10.1. Mucin Family

Mucins are glycoproteins that are produced by malignant epithelial tumors of pulmonary, gastrointestinal, gynecologic and mammary origin (Diaz et al., 2001). Mucins are broadly defined as proteins containing from 50-90% of their molecular mass as O-linked oligosaccharides. Mucins are complex molecules ranging in size from 400-1,000 kDa. They are a heterogenous group of molecules whose variations in molecular structure

are thought to carry tissue-specific functions (Bartman et al., 1998). All mucins are characterized by a tandemly paired and repetitive central peptide that is rich in serine and threonine residues.

Mucins can be subdivided into secreted mucins and tethered transmembrane mucins. Secreted mucins, which are primarily gel-forming, are the first line of defense for epithelial surfaces, serving as selective physical barriers between the extracellular milieu and the mucosal surface. Five secreted human mucin genes have been identified and well characterized (MUC2, MUC5AC, MUC5B, MUC6 and MUC7). Tethered transmembrane mucins may be the second line of defense, possibly acting as sensors of any disturbance to the environment and signaling this information to the interior of the cell (Gendler, 2001). Six tethered transmembrane mucins have been identifed (MUC1, MUC3, MUC4, MUC11, MUC12 and MUC13). MUC1, also known as epithelial membrane antigen, is the most extensively studied of the mucins.

2.10.2. Structure of MUC1

The human epithelial MUC1 is a heavily O-glycosylated type I transmembrane glycoprotein expressed ubiquitously on the ducts and glands of simple secretory epithelial tissues (Correa et al., 2005). The structure of the core protein of the MUC1 mucin is dominated by the large extracellular domain containing 25-125 tandem repeats (TRs) of amino acids (Taylor-Papadimitriou et al., 2002). TR domains, rich in serines and threonines, have been found to be a common feature of all epithelial mucins, for which they provide scaffold for the attachment of O-glycans. The full-length cDNA encodes a protein with the large extracelluar domain (1000-2200 amino acids), a hydrophobic

membrane-spanning domain, and a phosphorylated cytoplasmic tail of 72 amino acids that contains signaling motifs.

2.10.3. Function of MUC1

Numerous functions have been proposed for MUC1. Mucins trap debris or bacterial or viral pathogens and aid in their removal from mucosal surfaces by normal processes such as muco-ciliary clearance. MUC1 has been shown to bind to bacteria in human breast milk and in tissue culture and to respond with increased levels of tyrosine phosphorylation of the cytoplasmic tail upon bacterial adherence and changes in cellular adhesion (Quin RJ *et al.*, 2000).

MUC1 is thought to be an anti-adhesive protein due to its large, extended conformation (Fontenot *et al.*, 1993) and may block cell-cell interactions by destabilizing cell-cell and cell-matrix connections (Wesseling *et al.*, 1995).

2.10.4. MUC1 Expression in Breast Cancer

In normal breast tissue, MUC1 is expressed on the apical surface of epithelial cells in the ducts and acini from where the molecule is shed via milk globules and in soluble form into the milk (Cheung *et al.*, 2000). In case of tumors, cell polarization is lost and this altered cell surface expression, coupled with the disruption of the normal tissue architecture caused by the growing tumor, allows MUC1 to be shed into the circulation where it can be measured by means of immunoassays (Cheung *et al.*, 2000). The distribution of MUC1 is no longer restricted to apical surfaces of ducts and glands, but rather is found throughout the tumor mass and on the entire surface of tumor cells

(Gendler, 2001). In breast cancer, glycosylation is altered revealing immunodominant peptide sequences in every tandem repeat, which on normal tissues is masked by heavy glycosylation.

MUC1 glycosylation is very complex. Within each tandem repeat, two serines and three threonines represent five potential O-glycosylation sites. In the lactating mammary gland, MUC1 exhibits long polylactosamine chains, which sterically hinder the binding of peptide-reactive antibodies (Burchell and Taylor-Papadimitriou, 1993). In malignancy, MUC1 has shorter, less complex chains, due to changes in glycosyltransferases governing the extension and termination of O-linked glycans.

MUC1 is expressed in approximately 90% of all invasive breast tumors and is found to correlate with high metastatic potential and poor survival (McGuckin et al., 1995; Dong et al., 1997). In fact, statistics have shown that breast cancers expressing MUC1 accounted for about 72% of new cases and for 61% of cancer-related deaths (Gendler, 2001). MUC1 over-expression has also been associated with low tumor grade, smaller tumor size and a higher ER-positive phenotype (Rakha *et al.*, 2005).

2.10.5. MUC1 Detection

MUC1 displays relatively broad expression among epithelial tissues including the colon, breast, pancreas, ovary, prostate, tracheobronchial tree, stomach and uterus. For this reason, MUC-1 derived tumor antigens have relatively poor specificity for individual tumor types and their clinical utility is limited to monitoring the efficacy of cancer therapy and warning of tumor relapse or malignant spread (Cheung et al., 2000). Regardless of this, MUC1 peptides are being investigated for use as cancer vaccines and

have been shown to induce both humoral and cytotoxic T cell-mediated immune responses in animal models, as well as in phase I clinical trials in humans (Cheung et al., 2000). MUC-1 based immunotherapy studies have also been documented. MUC-1 immunotherapy has focused on the tandem repeat domain, based on the assumption that the repetitive structure would provide high levels of protein entering the antigen processing and presenting pathways for cytotoxic T lymphocyte recognition (Correa et al., 2005). However, clinical responses thus far have been disappointing.

Two assays have been developed to detect the MUC1 gene-derived glcyoprotein in patients with breast cancer through the identification of MUC1 antigens, CA15.3 and CA27.29 (Klee and Schreiber, 2004). MUC1, as detected by CA15.3 sandwich capture assay using specific monoclonal antibodies 115D8 and DF3, was the first mucin marker in which sequential changes were reported to correlate with therapeutic response (Hayes et al., 1986). The combination of data from 7 different studies showed that 67% of 352 patients had CA15-3 elevations either before or at the time of recurrence. In 1320 patients without evidence of recurrence, 92% had normal CA15-3 levels (Cheung et al., 2000). Serum CA15.3 levels have also been found to be elevated in 54-80% of patients with metastatic breast cancer.

3) Database Searches as a Tool For Identifying New Tumor Biomarkers

The identification of genes exclusively or abundantly expressed in cancer may yield novel tumor molecular diagnostic markers. Until recently, the approach to understanding the molecular basis of complex biologic processes such as human development and

cancer was to study the behaviour of genes one at a time. The development of technologies that allow a large number of transcripts to be analyzed simultaneously has made it possible to determine the molecular profile of normal and disease cells in a quantitative fashion. Simultaneous determination of the expression pattern of thousands of genes is now possible through the vast number of databases available to researchers. New tools such as, Serial Analysis of Gene Expression (SAGE) (Velculescu *et al.*, 1995), cDNA microarray analysis (Nacht *et al.*, 1999) and subtractive hybridization (Quan and Lu, 2003) are probably the most used.

In cancer research, the identification of new tissue-specific markers has benefited from serial analysis of gene expression (SAGE) databases. SAGE analysis is a powerful genetic profiling technology, which provides qualitative and quantitative assessment of gene transcript population. Several genes involved in breast cancer have been identified by this technology. AMACR (alpha-methylacyl-CoA racemase) which is a new putative tumor marker for several cancers (Zhou et al., 2002); EIT-6 (Estrogen Induced Tag-6), a novel nuclear protein thought to play a role in estrogen induced cell growth (Seth et al., 2002); dermcidin (DCD), a neural survival factor, which is a candidate oncogene for breast cancer (Porter et al., 2003). The concept of SAGE is based on targeting of cDNA sequences: a short nucleotide sequence called "Tag," is used to identify the original transcript. This procedure produces 10-15bp tags, which represent individual genes. Numerous tags are linked and can be sequenced simultaneously and quantified. Thus, the relative amounts of transcripts can be deciphered between different tissue sources. Serial analysis of gene expression (SAGE) profiling was adopted by the Cancer Genome

Anatomy Project (CGAP) in 1998, and over 5 million tags from more than 100 human cell types are posted at NCBI SAGEmap website (http://www.ncbi.nlm.nih.gov/SAGE).

Beginning in 1996, an initiative called the Cancer Genome Anatomy Project (CGAP) [www.cgap.nci.nih.gov/Tissues] set out to identify genes responsible for the establishment of cancer. To do so, libraries of cDNA, produced by reverse transcriptase from mRNA transcripts from normal, pre-cancerous and malignant cell DNA were built. Gene products of normal and malignant cells are then compared to determine which genes may be over-expressed in cancer cells. In the gene-sorting process, CGAP receives genetic information as expressed sequence tags (ESTs), which are placed into the EST database (dbEST). Digital Gene Expression Displayer (DGED) is a tool provided on the CGAP website (http://cgap.nci.nih.gov/Tissues/GXS), which compares gene expression (ESTs) between to pools of libraries. Another tool for finding cancer-related genes is the xProfiler (http://cgap.nci.nih.gov/Tissues/xProfiler). xProfiler compares the cDNA libraries from two different tissue types and searches for genes expressed in one tissue type, but not in the other.

4) The Human Small Breast Epithelial Mucin (SBEM)

4.1. Discovery of SBEM

The discovery of the small breast epithelial mucin (SBEM) gene in 2002 involved an innovative genome mining strategy. Using the cDNA xProfiler tool (http://cgap.nci.nih.gov/Tissues/xProfiler) to search for breast-specific expressed sequence tags (ESTs), 30 ESTs grouped under the UniGene identifier number Hs.348419,

were identified by our laboratory (Miksicek *et al.*, 2002). 15 of these ESTs were ascribed to breast cDNA libraries, 9 were isolated from random activation of gene expression or pooled tissues, 5 were isolated from fetal sources and 1 from a head and neck tumor cDNA library. Alignment of these ESTs led to the construction of a 500bp consensus cDNA sequence, corresponding to what is now known as the SBEM gene sequence. Further investigation using SAGEmap (http://www.ncbi.nlm.nih.gov/projects/SAGE) led to the identification of one SAGE tag, CCTCCTGTGA, which appeared to be over-expressed in 12 of 16 libraries of normal breast and breast cancer cell lines. As of 2005, more than 7500 tags derived from normal breast tissues, breast tumor tissues and breast cancer cell lines corresponded to SBEM mRNA, while only 72 of these were found in tissues other than breast (skin, colon, lung, brain, cartilage, prostate and heart) [SAGE Digital Northern -http://cgap.nci.nih.gov/SAGE/FreqsOfTag]. No tags corresponding to SBEM were detected in cDNA libraries from bone marrow, ovary, kidney, liver, pancreas, stem cells, stomach, testis or thyroid.

4.2. SBEM Gene Expression

The SBEM gene, located on chromosome 12q13, spans a 3.9kb region consisting of 4 exons and 3 introns (Figure 2). Several groups have reported the detection of SBEM mRNA in normal breast tissue (Houghton et al., 2001; Miksicek et al., 2002; Colpitts et al., 2002, Wadle *et al.*, 2005). Dot blot analysis revealed SBEM mRNA expression only in the breast and salivary glands, all other normal tissues such as brain, ovary, uterus, lung or prostate were negative (Miksicek et al., 2002; Colpitts et al., 2002). Using microarray hybridizations, Houghton et al. (2001) found that colon, kidney, heart and

skin were the only other non-breast organs which expressed low, but detectable levels of SBEM mRNA. SBEM mRNA was also reportedly detected in normal lung by RT-PCR, but our laboratory could not replicate the data by Koga *et al.* (2004).

In tumor tissues, variable levels of SBEM mRNA were detected by RT-PCR in more than 90% of breast tumors studied (Miksicek *et al.*, 2002; Colpitts *et al.*, 2002; Wadle *et al.*, 2005). SBEM mRNA was also found to be expressed in axillary lymph nodes containing metastases, but was absent in lymph nodes not expressing metastatic breast cancer cells (Weigelt et al., 2004; Miksicek et al., 2002). SBEM gene expression has been documented to be down-regulated in invasive and metastatic tissues compared with normal breast tissue by SAGE analysis (Zucchi *et al.*, (2004) and Allinen *et al.*, 2004). Recently, an SBEM deletion variant was detected in breast carcinomas (Wadle *et al*, 2005). This variant lacked 24bp (+149 after the translation initiation site) and contained 1 silent gene mutation just upstream of the deletion.

SAGE analysis of normal, early stage ductal carcinoma *in situ* (DCIS) and invasive breast tissues, revealed SBEM mRNA as being restricted primarily to epithelial cells (466 tags total; normal: 370 tags, DCIS: 89 tags, invasive; 7 tags). SBEM tags were also found in normal stromal cells (5 tags) and leukocytes (6 tags) (Allinen et al., 2004). SAGE analysis also showed a high level of SBEM expression in the 'nulliparous' mammary gland, with an almost complete down-regulation found in the normal breast of women with first full-term pregnancies (Verlinden et al., 2005).

4.3. SBEM Protein

The SBEM gene encodes for a protein of 90 amino acids, corresponding to a theoretically predicted size of approximately 9kDa (Figure 3). The first 19 amino acid portion of this sequence has the characteristics of a signal peptide, indicating SBEM is likely a secreted protein (Colpitts et al., 2002; Miksicek et al., 2002; Clark et al., 2003). The amino- and carboxy-termini of the processed 71 aa polypeptide is charged and fairly polar. SBEM is rich in Ala (21%), Thr (18%) and Pro (11%) residues, many of which are organized as a neutral octapeptide core motif (TTAAXTTA) repeated three times. Each of these repeats contains a single amino acid substitution at the fifth position (X) of the consensus sequence; alanine in the first repeat (res 46-53), proline in the second repeat (res 54-61) and serine in the third repeat (res 62-69). These motifs are characteristic of those found in mucins. SBEM protein is predicted to be extensively O-glycosylated due to the presence of 16 Thr residues, all of which are expected to be glycosylated (Figure 4). Lectin binding assays have shown that SBEM was bound to lectins, including peanut lectin (*Arachis hypogaea*), osage orange lectin (*Maclura pomifera*) and jacalin lectin (*Artocarpus integritolia*), which all have recognition elements for galactose or N-acetyl galactose (Colpitts et al., 2002).

An immunohistochemical study performed by Colpitts *et al.*, (2002), showed that the vast majority (70-100%) of early stage breast carcinoma tissues were positive for SBEM protein. SBEM protein was not detected in any other solid tumors by IHC including bladder carcinoma, colon carcinoma, endometrial carcinoma, gastric carcinoma, squamous cell lung carcinoma, ovarian carcinoma, parcreatic and prostate carcinoma (Colpitts *et al.*, 2002). SBEM protein has recently been detected as membrane-bound and

secreted forms in mammalian HEK-293 cells transiently transfected with SBEM cDNA (Wadle *et al.*, 2005). Interestingly, a variant of the SBEM protein (SBEMv) was also detected. This variant lacked 1 of 3 mucin-characteristic tandem repeats (8 amino acids), resulting in an 82 amino acid protein. No difference in the level or frequency of expression was observed between SBEM and SBEMv, as detected by SBEM/SBEMv-specific antibodies isolated from breast cancer patient sera. The serum-reactivity of these SBEM proteins was much stronger (30.2%) using the sera from breast cancer patients compared to the sera donated from healthy female controls (9.5%). This suggests that a greater level of SBEM protein may be expressed in breast cancer patients, or, that the SBEM antigen may be more readily detected by antibodies in these patients.

II) Rationale and Hypothesis

Tumor markers are molecules detected at higher or lower-than-normal levels in the body fluids (blood) or tissues of patients with disease. Detecting and/or monitoring these changes may assist in evaluating cancer risk, diagnosis, prognosis or response to treatment. Many current breast tumor markers are not breast-specific and only help to identify a small subset of patients. There is a need to identify novel genes whose expression is restricted to the mammary epithelium. Our laboratory has identified a novel gene, which we called the small breast epithelial mucin (SBEM). This gene was found to be highly expressed in breast tissues (Miksicek *et al.*, 2002) and thus, is a favorable candidate for a tissue-specific marker.

SBEM protein shares structural similarities with mucin family glycoproteins. The SBEM protein is predicted to contain a signal peptide, which suggests that the SBEM protein may be secreted. If secreted, assays may be developed to detect circulating SBEM protein in the serum of breast cancer patients. SBEM protein is also predicted to be glycosylated. In cancer, proteins are found to be under-glycosylated, as is seen with mucins. A decrease in the glycosylation of the SBEM protein in breast cancer patients, may give rise to tumor-specific antigens, which could be used as potential targets for immunotherapy.

Hypotheses: SBEM is a tissue-specific marker for breast cancer. Different levels of the SBEM gene are expressed between normal breast tissues and breast tumor tissues. SBEM exists as a secreted protein, which is glycosylated. In breast cancer, SBEM protein is under-glycosylated.

III) Research Objectives:

1. **To examine gene expression levels of SBEM in normal breast tissue and breast cancer.**

Specific aim: To determine the expression of the SBEM gene in normal mammary tissue, breast cancer cell lines and breast tumors. SBEM gene expression will be analyzed by RT- PCR and Northern blot analysis.

2. **To determine SBEM protein expression in normal breast tissue and breast cancer.**

Specific aims: a) To investigate the protein expression of SBEM in breast cancer cell lines, normal mammary glands and breast tumors. Endogenous SBEM protein will be detected by Western blot analysis and immunoprecipitation using a panel of anti-SBEM antibodies.

b) To determine SBEM protein expression and localization in normal breast tissues and breast tumors. Normal breast tissue and tumors will be analysed by immunohistochemistry

3. **To investigate the glycosylation status of the SBEM protein in breast cancer cell lines.**

Specific aim: a) To determine whether the SBEM protein is glycosylated. Breast cancer cell line and non-breast cell line total protein extracts will be

enzymatically deglycosylated. These extracts will then be examined by Western blot analysis.

b) To identify the secreted form of the SBEM protein and determine whether it is glycosylated. SBEM protein will be isolated from the culture media of stably transfected MCF-7 and HEK-293 cells and will be and analyzed by Western blot.

IV) Materials and Methods

4.1. Cell Culture

4.1.1. Culture Media

Dulbecco's Modified Eagle Media (DMEM, Gibco BRL, Catalogue #12800-017) was made according to manufacturers' instructions. The pH was adjusted to between 7.2 and 7.3. All media were sterilized using a 0.22μm filter (Millipore) and stored at 4°C. Prior to use, media were supplemented with 5% v/v fetal bovine serum (FBS) (American Type Culture ATCC, Manassas, VA, USA), 5% v/v (stock solution) penicillin-streptomycin (Invitrogen Corporation), 30% glucose (stock solution) and 5% v/v of 200mM L-glutamine (Invitrogen Corporation). This media is referred to as "complete media" (CM).

4.1.2. Cell Lines

Human breast carcinoma cell lines were obtained from the American Type Culture Collection (Manassas, Virginia). Cells (ZR-75, T47-D, MCF-7, MDA-MB-231, MDA-MB-468 and BT-20) were cultured in Dulbecco's Modified Eagle Media (DMEM) and were supplemented with 5% fetal bovine serum (FBS), penicillin-streptomycin (50Units/ml), L-glutamine (200mM) and 30% glucose (Gibco, Grand Island, NY). Pre-malignant MCF10-A1 human epithelial cells were maintained in DMEM supplemented with 10% horse serum, hydrocortisone (0.5mM), choleratoxin (50ug/ml), insulin (2.48mg/ml), EGF (20ug/ml), penicillin-streptomycin (50Units/ml), L-glutamine

(200mM) and 30% glucose. Stably-transfected MCF-7 breast cancer cell lines were cultured in Dulbecco's Modified Eagle Media (DMEM) and were supplemented with 5% fetal bovine serum (FBS), penicillin-streptomycin (50Units/ml), L-glutamine (200mM), 30% glucose and 100μg/ml of Geneticin (G418 - active) (Gibco, Grand Island, NY). SBEM stably-transfected human embryonic kidney 293 (HEK-293) cells (gift from Dr. T. Colpitts, Abbott Laboratories, Abbott Park, Illinois, USA) were cultured in high-glucose DMEM and supplemented with 10% FBS, penicillin-streptomycin (50Units/ml), L-glutamine (200mM), MEM non-essential amino acids, sodium pyruvate and 100μg/ml Geneticin. All cell lines were grown in humidified incubators in the presence of 5% CO_2 at 37°C in 75cm^2 polystyrene culture flasks (Corning Inc., Corning, NY, USA).

4.1.3. Propagation of Cell Lines

For routine cell passage, media was aspirated and the cells were rinsed with 2ml trypsin/EDTA (Invitrogen Corporation). The trypsin/EDTA was removed, 4ml of fresh trypsin/EDTA added, and the flask was incubated at 37°C, 5% CO_2 for 5 minutes. The trypsinized cells were pipetted up and down to loosen any remaining adherent cells. 1ml of cell suspension was added to 15ml of fresh media to continue culture growth.

For long-term storage, cells were trypsinized and were added to cell freezing media (containing; 4ml FBS, 5 ml 5% CM and 1ml DMSO) in cryogenic vials. To ensure gradual freezing, cells were kept at –70°C overnight. The next day, vials were placed into liquid nitrogen for long-term storage.

4.2. Tissue Collection

Normal breast tissue used for RNA and protein analysis was obtained from 4 different pateints who had undergone reduction mammoplasty. Invasive ductal carcinoma cases used for RNA and protein analysis were selected from the Manitoba Breast Tumor Bank by Dr. Peter Watson (Winnipeg, Manitoba, Canada, (Troup et al., 2003), which operates with the approval from the Faculty of Medicine, University of Manitoba, Research Ethics Board. Estrogen receptor and progesterone receptor levels were determined by ligand binding assay ranging from (0-338fmol/mg protein) and (0-729fmol/mg protein) for ER and PR respectively. ER levels of > 3fmol/mg protein and PR levels > 10fmol/mg protein, were considered positive.

Another 10 breast tumor cases were used for RNA analysis and immunohistochemistry. These tumors spanned a wide range of estrogen and progesterone receptor levels (ER, ranging from 2.3 to 180 fmol/mg of protein, and PR, ranging from 4.5 to 105 fmol/mg of protein), and Nottingham grade (5 to 9).

4.3. Isolation of Total RNA

Confluent cells (80-90%) grown in 150mm dishes were collected by 5 minutes incubation with trypsin-EDTA (Gibco). Trypsin was inactivated by adding same volume of 5% CM and cells were collected into 15ml Corning tubes and were briefly centrifuged at 1,000xg for 5 minutes at 4°C. Cell pellets were washed twice with 1X phosphate buffered saline (PBS) (after aspiration of the medium). For tissue RNA extraction, 20μm frozen sections were used. Total RNA was extracted from cell lines and tissues using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's

instructions. Briefly, 1ml of TRIzol reagent was added to cells or tissues and homogenized in 1.5ml Eppendorf microcentrifuge tubes using a Brinkmann Polytron Homogenizer (Brinkmann Instruments, Westbury, NY, USA). The homogenized samples were then incubated at room temperature for 5 minutes and 0.2ml of chloroform was added per 1ml of Trizol reagent. Samples were shaken for 15 seconds and left at room temperature for 2-3 minutes. The homogenate was centrifuged at 12,000xg for 15 minutes in a Thermo IEC Multi RCF table-top centrifuge. The clear upper aqueous phase was removed and placed on ice. To precipitate RNA, 0.5ml of isopropanol (for every 1 ml of TRIzol Reagent) was added to the samples, shaken for 15 seconds and incubated for 10 minutes at room temperature. The samples were then centrifuged at 12,000xg for 10 minutes at 4°C. The supernatant was carefully pipetted off and the RNA pellet washed with 1 ml cold 75% ethanol and spun for 5 minutes at 7,500xg. The ethanol was again carefully removed and the tube centrifuged for an additional 5 minutes to separate out any residual ethanol, which was then removed. The pellet was air-dried and dissolved in 50µl of water. The solution was then heated at 65°C for 5 minutes and then placed on ice for 5 minutes. The purity and yield of the RNA were determined by spectrophometry and only samples with an A260/A280 ratio above 1.6 were kept for further experimentation. The concentration of RNA was determined using the following equation:

Equation 1:

$$\textbf{[Optical Density } A_{260} \textbf{ X 40µg/µl X Dilution Factor] / 1000 = RNA (µg/µl)}$$

The total RNA isolated was stored at -70°C for Northern Blot analysis.

4.4. Mini-Gel Electrophoresis of RNA

The integrity of the extracted RNA was analyzed by gel electrophoresis. A 1.3% (w/v) agarose (Invitrogen Corporation) gel was prepared with 1X Gel Running Buffer (GRB; 1X GRB = 0.04M morpholinopropanesulfonic acid (MOPS) pH 7.0, 10mM sodium acetate, 1mM ethylenedinitroltetraacetic acid (EDTA) pH 8.0) and 2.2M formaldehyde. 2μg of total RNA was denatured at 65°C for 10 minutes in 10μl of a master mix solution containing formamide (Fisher Scientific), 37% formaldehyde (Fisher Scientific) and 5X GRB (0.2M MOPS, 50mM sodium acetate, 5mM EDTA pH 8.0) and 0.25mg/ml ethidium bromide. The samples were immediately placed on ice following incubation to prevent re-annealing and then were briefly centrifuged. Sample loading buffer (50% glycerol, 0.5mM EDTA pH 8.0, 0.25% Bromophenol Blue, 0.25% Xylene Cyanol FF) was added directly to RNA samples and samples were electrophoresed at 100V for 50 minutes in 1X GRB. The gel was illuminated with ultraviolet light and photographed using a CCD camera and the Microcomputer Imaging Device (MCID) M4 software version 2.0 (Imaging Research Inc., St. Catherine's, ON, Canada).

4.5. Northern Blot Analysis

Expression of RNA in normal human tissues and breast cancer cell lines was examined by Northern blot analysis. A 1% agarose gel was prepared with 1X GRB and 2.2M formaldehyde. Total RNA (40μg) isolated from cells and tissues, was prepared in denaturing solution (See Mini Gel Electrophoresis, Section 2.4.) and placed in a water bath at 65°C for 15 minutes. The samples were immediately placed on ice following the incubation time to prevent re-annealing and briefly centrifuged. The RNA samples were

loaded in the gel with sample loading buffer (50% glycerol, 0.5mM EDTA pH 8.0, 0.25% Bromophenol Blue, 0.25% Xylene Cyanol FF and were electrophoresed overnight at 25V in running buffer of 1X GRB. The gel was exposed to ultraviolet light and photographed using a CCD camera and the MCID M4 software (Imaging Research Inc.) The agarose gel was transferred overnight to nitrocellulose membrane in 20X SSC buffer (1X SSC = 0.15M NaCl, 0.015M sodium citrate). The nitrocellulose membranes were baked at 80°C for 2 hours in a drying oven supplied with a vaccuum (Lab-Line Instruments).

Membranes were pre-hybridized in 10ml pre-hybridization solution (50% (v/v) formamide, 5X SSPE (1X SSPE = 0.15M NaCl, 0.01M NaH_2PO_4, 1mM EDTA pH 7.7), 5X Denhardt's Solution, 0.1% sodium dodecyl sulfate (SDS) and 250µg/ml salmon sperm DNA for 2 hours at 42°C in a rotating hybridization incubator (Robbins Scientific, Sunnyvale, CA, USA). A 291bp SBEM PCR product cDNA probe (30-35ng) was prepared with ^{32}P-labeled dCTP isotope and the *Redi*prime II random prime labeling system according to the manufacturer's instructions (Amersham Pharmacia Biotech, Piscataway, NJ, USA). The probe was purified with columns containing Sephadex G-50 DNA-grade. The labeled probe was denatured by boiling for 5 minutes, was immediately cooled and added to hybridization solution containing 10% dextran sulfate. The hybridization solution containing the labeled probe was added to the membranes in the glass tube and allowed to hybridize overnight at 42°C in the rotating oven. The blots were washed once with 2X SSC/0.1% SDS at room temperature for 5 minutes and then washed in 0.1% SSC/0.1% SDS at 65°C for 5 minutes. The membranes were exposed to an intensifying screen for 1 day and were developed using a Molecular Imager FX phospho-

imaging system (Bio-Rad, Hercules, CA, USA) and images were saved (Quantity One Version 4.2 software; Bio-Rad Laboratories).

4.6. Reverse Transcription (RT) Polymerase Chain Reaction (PCR)

2μg of total RNA was reverse-transcribed for 1 hour at 37°C in a mixture containing 300μM of each deoxynucleotide triphosphate, 15mM DTT, 50ng random hexamers, 12 units of RNase Out and 300 units of Moloney Murine Leukemia Virus (M-MLV) Reverse Transcriptase (Life Technologies). A negative RT-PCR was performed without the addition of the MMLV enzyme.

4.7. Polymerase Chain Reaction

PCR was performed as previously described (Leygue *et al*, 1999). Briefly, 2μL of each transcription mixture was amplified in a final volume of 50μL, in the presence of 20mM Tris-HCl (pH 8.4), 50mM KCl, 1.5mM $MgCl_2$, 200μM of each deoxynucleotide triphosphate (dNTP), 200ng of each SBEM primer [SBEM-U (5-atgaagttcttagcagtcctg-3) and SBEM-L (5-ttctaccactgctcgtaaaga-3) (TABLE 1)], and 0.5 units of Taq DNA polymerase (Invitrogen Corp. Burlington ON.) Each PCR consisted of 35 cycles (30s at 94°C, 30s at 52°C, and 30s at 72°C). Primers for the ubiquitously expressed GAPDH gene were GAP-U (5-acccactcctcctccacctttg-3) and GAP-L (5-gggacacactctaccattcg-3) Table 1. To amplify cDNA corresponding to GAPDH, 35 cycles of PCR were used as above for SBEM. PCR products were then separated on a 2% agarose gel, containing 1X TBE and ethidium bromide for 1 hour at 100V. PCRs were performed, visualized with UV irradiation on a GelDoc2000/ChemiDoc System (Bio-Rad), and quantified by

densitometry using QuantityOne software (Version 4.2, Bio-Rad). PhiX174 RF DNA/Hae III DNA ladder was used to size PCR products.

4.8. Protein Extraction

Frozen tissue samples (50mg) were homogenized in a Triton-X extraction buffer containing, 20mM MOPS (pH 7.2), 60mM β-glycerophosphate, 5mM EGTA (pH 8.0), 5mM NaF, 1mM Na_3VO_4, 1 Complete Tablet (Boehringher-Mannheim) and 1% Triton-X (Fisher). Cells (80-90% confluency) were lysed directly using an SDS-buffer containing 1.25mM Tris-HCl (pH 6.8), 20% glycerol, 4% SDS, 0.02% (w/v) Bromophenol Blue. The homogenized samples were sonicated twice for 30 seconds and boiled for 5 minutes. Following centrifugation at 14,500xg for 5 minutes at 20°C, the supernatants were collected and stored at -20°C. Purified SBEM protein from cell culture media of HEK-293 stably transfected cell lines, was a generous gift from Dr. Tracey Colpitts.

4.9. Determination of Protein Concentration

Concentrations of pooled protein were determined by the Lowry assay. 2μl of protein samples in SDS-isolation buffer or Triton-X lysis buffer were diluted in ddH_2O. The BSA standard concentrations used for optical density were 2.5, 5, 10, 20, 40, 60 and 80μg/μl. To each sample, 1ml of Lowry A, B and C (sodium tartrate, copper sulfate, sodium carbonate) solutions were added. 100μl of 1:1 Follins Reagent was then added to each sample. The absorbance was measured using SoftMax Pro software (Molecular Devices) at a wavelength of 650nm. A BSA standard curve was plotted using the GraphPad Prism version 3.02 software. The slope (m) and y-intercept (b) was obtained and used to

calculate the unknown concentrations of the protein samples according to the following equation with a dilution factor of 2.

Equation 2:

[Optical Density – b/m] X Dilution Factor = Protein (μg/μl)

4.10. Antibody Preparation

Rabbit polyclonal SBEM-antibody (Invitrogen) generated against amino acids 70-87 of human SBEM (1:1500) and mouse monoclonal SBEM antibodies (gift from Dr. T. Colpitts) raised against various regions of the SBEM peptide sequence (Figure 5) (1:500) were used for immunodetection on western blots and immunohistochemistry. Goat anti-rabbit and goat anti-mouse horseradish peroxidase (Bio-Rad Laboratories Canada Ltd, Mississauga, ON) (1:5000) and SuperSignal West Pico Chemiluminescent Substrate (Pierce) were used to detect immunoreactive bands on Western blots.

4.11. Western Blot Analysis

Western blot analysis by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) was used to identify proteins in cell and tissue lysates using specific antibodies. Broad range Kaleidoscope prestained standards (Bio-Rad Laboratories, Hercules, CA) were used as molecular weight markers. Each sample (50μg to 100μg) of protein prepared from cell and tissue lysate was separated by electrophoresis at 80V through a 15% SDS-polyacrylamide gel in running buffer (25mM Tris base, 192mM Glycine, 1% SDS). To transfer proteins to a nitrocellulose membrane (Millipore,

ON, Canada) the gel, the membrane and filter papers were soaked in transfer buffer (20% methanol, 0.4M CAPS), layered in a transfer apparatus and set at 100V for 1 hour. The membrane was baked for 15 minutes at 65°C and was then incubated in blocking buffer (5% skim milk in 0.1% TBS) for 1 hour at room temperature. The membrane was incubated overnight at 4°C with the appropriate antibody diluted in blocking buffer. After removal of the primary antibody solution, the membrane was washed for 30 minutes in 0.1% TBST with agitation, changing the buffer every 10 minutes. The membrane was then incubated in blocking buffer containing the secondary antibody for 1 hour. The blot was washed again for 30 minutes in 0.1% TBST with agitation changing every 10 minutes. To determine antibody specificity, duplicate blots were incubated with appropriate antibody pre-mixed with 12μg/ml of the peptide used to raise it. The blot was then developed using SuperSignal West Pico Chemiluminescent Substrate (Pierce). Chemiluminescent signal was captured by video analysis using the Quantity One system (Bio-Rad, Hercules, CA). SDS-PAGE gels were stained with Coomassie Brilliant Blue R-250 (Bio-Rad, Hercules, CA, USA) to assess loading.

4.12. Immunohistochemistry

Immunohistochemistry was performed on 5 μm sections of formalin-fixed paraffin-embedded sections of 4 normal breast tissues and 10 invasive carcinomas. Immunohistochemical staining was performed using a monoclonal SBEM antibody, H218C31 (kindly provided by Dr. Tracey Colpitts). Antibodies were applied using an automated tissue immunostainer (Discovery module, Ventana Medical Systems, Tucson, AZ, USA). iVIEW DAB immunohistochemistry kit and bulk reagents that were supplied

by the manufacturer. Briefly, the Discovery staining protocol was set to 'Standard Cell Conditioning' procedure (according to the manufacturer's instructions), followed by 1 hr incubation with primary antibody and 30 min incubation with secondary antibody. Concentrations of primary antibody initially applied to the Ventana instrument were 1:800 for H218C31, which translates into final concentrations of 1:2400 after a 1:3 dilution with buffer dispensed onto the slide with the primary antibody. A positive and a negative control were used, consisting of samples highly positive and negative for SBEM expression by RT-PCR, respectively. Another negative control, the omission of the primary antibody, was also included (data not shown).

4.13. Immunoprecipitation

Cells were grown to 80% confluency in 100mm culture dishes. Approximately 1 x 10^6 cells were washed twice with ice-cold 1X PBS, scraped off the dishes with a rubber policeman and were pelleted by centrifugation with a Thermo IEC Multi RCF table-top centrifuge at 1000x*g* for 5 minutes. PBS was aspirated off and 1mL of immunoprecipitation lysis buffer (50mM Tris-HCl (pH 7.5), 150mM NaCl, 150mM EDTA (pH 8.0) and 1% Triton-X) was added to the cell pellet. Lysates were incubated for 30 minutes on ice.

As a washing step, 500μL of lysis buffer was added to a 70μL slurry of Protein G Sepharose 4 Fast Flow beads (Amersham), corresponding to a 1:1 mixture of beads and 20% ethanol. This mixture was centrifuged for 30s at 2000x*g* and supernatant was removed. Beads were washed twice more and the final bead volume was resuspended in 1X bead volume with lysis buffer (50% slurry).

As a pre-clearing step, 35μL of pre-washed 50% Protein G bead slurry was added to 1mL of cell lysate, mixture was rotated at 4°C for 45 minutes and was spun down at 16000*g* for 5 min at 4°C. Supernatant was transferred to a new eppendorf and was saved as the pre-cleared lysate. 25μL of 2X SDS lysis buffer was added to the pre-cleared beads, boiled for 5 minutes and was analyzed by Western blot immediately, or stored at -20°C until further use.

According to Sambrook *et al.* (1989), 5-100μL (or 1-5μg) of cultured hybridoma cell supernatant is the ideal quantity for immunoprecipitation. 1.5μg (65μL) of C-terminal monoclonal antibody, H39C51 was added to the pre-cleared lysate and was rotated overnight at 4°C. Next day, the remaining 35μL of pre-washed beads were added to the antigen-antibody immunocomplexes and were rotated for 1hr at 4°C. Beads were pelleted gently for 30s at 1000x*g* and were resuspended in 1mL of lysis buffer. This was repeated 3 times. Finally, supernatant was removed and 20μL of 2X SDS buffer was added directly to the pellet, boiled for 5 minutes and was analyzed by Western blot or was stored at -20°C until further use.

4.14. Cell Fractionation

Cell fractionation was performed on SBEM stably transfected HEK-293 cells, using the FractionPREP Cell Fractionation System kit (BioVision). Cells (4-8 x 10^6) were collected, washed with 1X PBS and were centrifuged at 700xg for 5min. Cytoskeletal, cytosolic, nuclear and membrane fractions were collected as per manufacturer's recommendation. All fractions were stored at -70°C until further use.

4.15. Column Purification

Transfected His-tagged SBEM protein was purified using the His-Select^{TM} HC Nickel Affinity Gel (Sigma, Oakville, ON, Canada). Total protein lysates were collected in equilibration buffer (50mM sodium phosphate, pH 8.0, 0.3M sodium chloride) as per manufacturer's recommendation. Cell extract was added to the affinity gel, was rotated for 15 minutes on a rotating platform and was centrifuged at 5000x*g* for 5 minutes. His-tagged protein was eluted with a series of washes using an elution buffer (50mM sodium phosphate, pH 8.0, sodium chloride) containing varying concentrations of imidazole (50-300mM). Eluted samples were concentrated by filter separation using Microcon YM-10 centrifuge tubes at 14000x*g* for 30min. Concentrated samples were precipitated with 4X acetone and were incubated overnight at -20°C. Next day, precipitates were spun down at 12000x*g* for 20min. Supernatant was removed and pellets were air-dried for 2 hours. Pellets were resuspended in 25µL of 2X SDS-PAGE buffer. Samples were analyzed by SDS-PAGE.

4.16. Enzymatic Deglycosylation

For deglycosylation, cells and tissues were lysed using an extraction buffer containing: 20mM MOPS pH 7.2, 60mM β-glycerophosphate, 5mM EGTA pH 8.0, 5mM NaF, 1mM Na_3VO_4, 1 Complete Tablet (Boehringher-Mannheim) and 1% Triton-X (Fisher). 90µg of total protein lysate was dissolved in 1:1 volume of 2X SDS-PAGE buffer and deionized H_20 to a final volume of 19.2µl. 6.4µl of 5x incubation buffer (0.25M sodium phosphate, pH 7.0) and 1.6µl of denaturation solution (2% SDS and 1M β-mercaptoethanol) were added to the proteins, were gently mixed, and were incubated at

100°C for 5 min. After cooling to room temperature, 1.6µl of detergent solution (15% Nonidet P-40) was added to the proteins. Then 0.64µl each of N-glycanase, O-glycanase, sialidase A, β(1-4)-Galactosidase and β-N-Acetylglucosaminidase or 3.2µl of water (control) was added to the proteins and subsequently incubated for 18 h at 37°C. Samples were diluted with 8µL of 5x Laemmli buffer (final volume of 40µL) and subjected to Western analysis by SDS-PAGE on 15% gels. Fetuin protein was provided with the kit (Glyko, San Leandro, CA, USA), which served as a control ensuring enzymatic deglycosylation was successful.

4.17. Detection of SBEM protein in Cell Culture Media

SBEM stably transfected MCF-7 cell lines were grown to full confluency and were subsequently seeded into new flasks. At 20% confluency, media was aspirated and replaced with serum-free media containing DMEM, penicillin-streptomycin (50Units/ml), L-glutamine (200mM), 30% glucose and 10µg/µL of human transferrin. Media was collected after 48 hours of incubation. Subsequently, media was dialyzed in Spectra/Por dialysis bags (MW cutoff of 3,500Da) for 72 hours. After dialysis, media was lyophilized using a Speed Vac for 24-30hrs. Dried protein was re-constituted in 500µL native buffer (50mM sodium phosphate, pH 8.0, 0.3M sodium chloride). Samples were stored at –20°C until further use or were analyzed immediately by Western blot. Secreted SBEM protein from culture media of stably transfected HEK-293 cells (generously provided by Dr. T. Colpitts) was also examined.

4.18. *In Vitro* Translation

Proteins were generated from 1μg of linearized pcDNA3.1/V5-His-TOPO SBEM plasmid constructs containing 289 and 316 bp PCR inserts (Invitrogen) by *in vitro* translation in the presence of [^{35}S] methionine, using a wheat-germ lysate based coupled transcription/translation system (TNT T7 wheat germ extract protocol; Promega) according to the supplier's recommendations. An empty vector was used as the negative control. The reaction was incubated at 30°C for 2 hours. After reaction, 2.5μl of the final product was mixed with SDS loading buffer and resolved on a 15% SDS-PAGE gel for 2 hours at 100V. Gel was placed into fixative (50% methanol, 10% acetic acid) for 30 minutes and was dried under vaccuum for 30 minutes without heat and then 30 minutes with heat at 70°C and was exposed to an intensifying screen overnight and was developed using a Molecular Imager FX phospho-imaging system (Bio-Rad, Hercules, CA, USA).

V) RESULTS

5.1. SBEM Gene Expression

5.1.1. Analysis of SBEM gene expression in normal human breast tissue

5.1.1.a. RT-PCR analysis

Total RNA was isolated from normal breast tissues obtained from 2 commercial sources (Ambion and BD Biosciences) and 1 reduction mammoplasty specimen. RT-PCR was performed with 2μg of total RNA, followed by subsequent PCR using primers specific for the SBEM sequence as described in Materials and Methods (TABLE 1). The expected PCR product, a 215bp fragment, was observed in all three normal breast tissues examined. (Figure 6A). The same product was detected in the SBEM stably transfected HEK-293 control cells. A 178bp product of the house-keeping gene GAPDH was used as a loading control (Figure 6B) and results were normalized relative to the GAPDH gene expression. The ratio of SBEM product/GAPDH product showed similar levels of expression among all three samples (Figure 6C).

5.1.1.b. Northern blot analysis

An mRNA transcript corresponding to a size of 650-700 bases was detected by Northern blot analysis in normal breast RNA from commercial sources (Ambion and BD Biosciences) (Figure 7A). A detectable level of SBEM mRNA was detected in normal breast taken from reduction mammoplasty (not shown). Human pituitary gland was used as negative control for SBEM expression. On further inspection, it was observed that the integrity of RNA in the salivary gland was poor. Human placenta was negative for SBEM

expression, perhaps because the quantity of RNA was low. The expected 900-950 base transcript was detected in the total RNA of HEK-293 cells stably transfected with the SBEM sequence, another positive control. The ribosomal bands 28S (4.3kb) and 18S (1.8kb) were used as controls for RNA integrity and loading (Figure 7B).

5.1.2. Analysis of SBEM Gene Expression in Breast Cancer Cell Lines

5.1.2.a. RT-PCR analysis

Using RT-PCR, our lab has previously shown that the SBEM gene was detectable in some human breast cancer cell lines (Miksicek *et al.*, 2002). In this study we examined SBEM gene expression in an extensive panel of breast cancer cell lines, including MDA-MB-468 and the non-breast cancer cell line, HEK-293, which were not examined by Miksicek *et al.* (2002). The results show a 215bp PCR product in the BT-20, MCF-7, ZR-75, T47-D and MCF10-A1 breast cancer cell lines (in order of highest to lowest levels expressed) (Figure 8A), obtained with SBEM-specific primers (TABLE 1). The same product was also detected in HEK-293 control cells stably transfected with SBEM. Controls for reverse-transcriptase (RT) and PCR (ddH_20) were negative for SBEM expression. A 178bp PCR product of the house-keeping gene GAPDH was used as a loading control (Figure 8B) and results were normalized relative to the GAPDH gene expression. Various levels of SBEM gene expression were observed amongst the cell lines (Figure 8C). The highest level of SBEM expression was observed in BT-20 by RT-PCR.

5.1.2.b. Northern blot analysis

Northern blot analysis was also used to confirm SBEM gene expression in the breast cancer cell lines. A 650-700 base mRNA transcript was detected in BT-20 cells and ZR-75 by Northern blot, using an SBEM-specific radio-labeled [^{32}P] probe (Figure 9A). The same probe hybridized with a 900-950 base mRNA transcript in SBEM-transfected HEK-293 cells. All other cell lines examined were negative. 28S and 18S ribosomal bands stained with ethidium bromide were used as controls for RNA integrity and loading (Figure 9B). The highest level of SBEM expression was observed in BT-20 by Northern blot.

5.1.3. Analysis of SBEM Gene Expression In Human Breast Cancer tissue

5.1.3.a. RT-PCR

Total RNA derived from breast tumors were analyzed by RT-PCR. The results show that the expected SBEM PCR product, a 215bp band, was apparent in 9 out of 9 tumors (Figure 10A). In addition, the same product was detected in the SBEM stably transfected HEK-293 control cells. A 178bp band was detected for the house-keeping gene GAPDH loading control (Figure 10B), using GAPDH-specific primers (TABLE 1).

5.1.3.b. Northern Blot analysis

Due to the limited quantity of RNA 9 μg of total RNA was used for Northern blot analysis. The expected 650-700 base mRNA transcript was detected with the ^{32}P-SBEM cDNA probe in 7 out of 9 (71%) tumors examined (FIGURE 11A). The level of SBEM

transcript varied between samples from strong (sample #8921, 1962 and 1894) to moderate (sample #9924 and 1897) to weak (sample #1892 and 1896). An additional transcript, approximately 350-375 bases in size, was detected in one sample (sample #1962). A 900-950 base mRNA transcript was observed in HEK-293 transfected control. 28S and 18S ribosomal bands stained with ethidium bromide were used as controls for RNA integrity and loading (Figure 11B). One sample (#1890), which was negative by Northern blot, was positive by RT-PCR.

5.2 SBEM protein expression

5.2.1. Analysis of SBEM protein expression in normal mammary gland tissue

5.2.1.a. Western blot analysis

To investigate SBEM protein expression in normal human breast tissues, 4 reduction mammoplasty cases and a commercially available breast Protein Medley were analyzed by Western blot (see Methods and Materials). One of these 4 reduction mammoplasty samples was previously used for RT-PCR and Northern blot studies. The use of the SBEM polyclonal antibody and monoclonal antibodies, H39C51 and H51C42, resulted in numerous bands being detected (FIGURE 12A-C), all of which were deemed non-specific following peptide neutralization in parallel blots. No endogenous SBEM protein was detected in normal breast tissue by Western blot analysis using these antibodies.

5.2.1.b Immunohistochemistry (IHC)

To determine SBEM protein expression and localization in normal breast tissue, 4 reduction mammoplasty specimens were analyzed by immunohistochemistry. Using the monoclonal antibody, H218C31, only 1 of 4 samples stained positively for SBEM. Positive staining was confined to the cytoplasm and membrane of ductal epithelial cells (FIGURE 13).

5.2.2. Analysis of SBEM Protein Expression in Breast Cancer cell lines

5.2.2.a. Western blot analysis

A polyclonal antibody generated to the C-terminal region of SBEM (70-87) was used as the primary antibody for Western blot analysis of human breast cancer cell lines. Varying intensities of a 33kDa SBEM band was observed in all cell lines (FIGURE 14A). SBEM specificity was verified by neutralization of the antibody with an SBEM peptide (FIGURE 14B).

Further to these studies, a panel of different SBEM monoclonal antibodies were used for Western blotting (TABLE 2). No endogenous SBEM protein was detected in breast cancer cell lines using these monoclonal antibodies (FIGURE 15). The non-breast cancer cell line, HeLa, was negative by Western analysis. Two bands 17kDa and 22kDa were identified in the stably transfected MCF-7 and HEK-293 cell lysates. The 17kDa band was the predominantly expressed band in HEK-293, while the 22kDa band was expressed at higher levels in MCF-7. Both bands were detected in transfected MCF-7 cells using anti-V5 antibody. These bands were also detected in the nickel column purified SBEM stably transfected HEK-293 and MCF-7 total cell lysates with the polyclonal antibody

(FIGURE 16). The untransfected HEK-293 cell line, used as a control for stably transfected HEK-293 cells, was negative by Western blot analysis.

Different sizes of the SBEM protein from transfected MCF-7 and HEK-293 cells were observed using different monoclonal antibodies. Three monoclonal antibodies, H51C42, H76C43 and H218C31 recognized the 17kDa SBEM protein. Two monoclonal antibodies, H39C51 and H24C16 recognized the 22kDa band in MCF-7 cells. The same two antibodies detected both the 17 and 22kDa SBEM bands in HEK-293 cells (Figure 15).

5.2.3. Analysis of SBEM Protein Expression in Breast Tumors

5.2.3.a. Western blot analysis

No endogenous SBEM protein could be detected in breast tumors by Western blot analysis using monoclonal antibodies, H39C51 (Figure 17) or H24C16 (not shown). 17 and 22kDa bands were detected in the HEK-293 positive control cell line. No protein corresponding to SBEM could be detected in the same tumors by Western blot analysis with a polyclonal SBEM antibody (not shown).

5.2.3.b. Immunohistochemistry (IHC)

Immunohistochemical analysis was performed on a small series of 10 breast tumor cases. Nine of 10 tumors stained positively for SBEM protein with the H218C31 monoclonal antibody. The SBEM protein was expressed at different levels in the different tumor specimens (FIGURE 18). Sample (# 8921) exhibited the highest level of staining for SBEM. This very sample was also found to be strongly positive for SBEM

gene expression by Northern blot and RT-PCR. The staining of SBEM was localized to the cytoplasm and membrane of ductal epithelial cells.

Overall, the immunohistochemical data appear to correlate well with the gene expression data as determined by RT-PCR and to a lesser extent with the Northern blot analysis data.

5.3. Immunoprecipitation of Stably Transfected Cells and Breast Cancer Cell Line Protein Extracts

To enhance the sensitivity of detection of endogenous SBEM protein detection by monoclonal antibodies, immunoprecipitation (IP) was performed on BT-20 breast cancer cell lines (see Methods and Materials). To optimize experimental conditions, IP of the HEK-293 transfected cell lysate was first carried out (FIGURE 19A). Under the conditions used, a strong 17kDa band was detected by the H39C51 monoclonal antibody, while a more diffused band was observed at approximately 22kDa. In addition, a strong band, which corresponded to the heavy chain IgG of the H39C51 antibody, was observed at approximately 50kDa. In contrast, the HEK-293 untransfected cell lysate was negative for SBEM protein expression (Figure 19A). A weak band was detected at 13kDa in BT-20 cell lysates, near the theoretically expected size (9kDa) (FIGURE 19B). No detectable SBEM protein was observed in the ZR-75 cell lysates by this method (not shown).

5.4. Sub-localization of SBEM protein in different cell fractions of cell lysate

To examine the sub-cellular localization of SBEM protein, HEK-293 transfected cells were fractionated into cytosolic, membranous, nuclear and cytoskeletal fractions. Western blotting of these fractions was performed using the monoclonal antibody H39C51 and the SBEM polyclonal antibody. Results showed a 17kDa and a 22kDa band in the cytosolic and membranous fractions (Figure 20A and B). The level of the SBEM protein appeared to be highest in the membrane fraction. No detectable level of SBEM protein was observed in the nuclear and cytoskeletal fractions.

5.5. Is the SBEM protein a secreted protein?

SBEM is predicted to be a secreted protein as it possesses a signal peptide in its N-terminal region. To examine whether a secreted SBEM protein exists, culture media derived from 2 SBEM stably transfected cell lines, MCF-7 and HEK-293, were analyzed for the SBEM protein. A 26kDa band was detected by Western blotting with the H39C51 SBEM monoclonal antibody, in media collected from MCF-7 cells, while a broadly diffused band approximately 25-33kDa was observed for HEK-293 transfected cell medium using the same antibody (Figure 21A). Since the construct used to generate the stably transfected MCF-7 cell line contained a V5-His tag at the C-terminal region, a V5 antibody was used to confirm that the 26kDa band was specific (Figure 21B). In the HEK-293 cell line, the SBEM plasmid construct, contained a myc-His tag and lacked a V5-His tag. Thus, the 25-33kDa was not detected by the V5 antibody. In both cell lines, the size of SBEM band appeared to be larger than the band identified in the cell lysate.

5.6. *In Vitro* Translation of SBEM Plasmid Constructs Used to Generate Stably Transfected MCF-7 Cell Lines

To investigate the size of SBEM protein in the absence of post-translational modification, pcDNA 3.1 V5-His-TOPO plasmids (Invitrogen) inserted with SBEM sequences were examined by *in vitro* translation (see Materials and Methods) (Figure 22). The plasmid used to generate a V5-His tag containing SBEM protein, was constructed with an in-frame (inside the start and stop codon) 289bp PCR product. Another SBEM construct was generated with the same pcDNA3.1 vector, however the inserted 316bp PCR product was outside of the coding region (inside the start codon, but downstream of the stop codon), resulting in the production of an un-tagged SBEM protein. The expected translated product size of the 289bp construct is 11kDa, while a 9kDa product is expected for the 316bp construct. Following *in vitro* translation, a 21-22kDa band was detected for the "in-frame" construct, while a 17-18kDa band was observed in the "out-of-frame" construct. *In vitro* translation of empty vector resulted in no protein product. This result agreed with previous Western blot analysis data where 17 and 22 kDa proteins were detected in MCF-7 stably transfected cells.

5.7. Is the SBEM protein glycosylated?

5.7.1. Enzymatic deglycosylation of SBEM protein in cell lysate

The SBEM protein is predicted to be heavily glycosylated due to the presence of numerous threonine residues in the peptide sequence. It is well recognized that glycosylation can sterically mask antibodies from recognizing epitopes, contributing to difficulties in immuno-detection of endogenous proteins (Burchell and Papadimitriou., 1993). To investigate the glycosylation status of the SBEM protein, deglycosylation assays were performed using cell lysates (see Materials and Methods) derived from BT-20, and stably transfected MCF-7 and HEK-293 cell lines by Western blotting. The blot was probed with the SBEM monoclonal antibody, H39C51.

To test the efficiency of the assay, the fetuin control protein was deglycosylated. This also allowed for the optimization of the amount of control protein that should be included in the assay and the conditions required. It was determined that 10μg was the ideal amount to use for Western blotting (Figure 23A). In addition, to ensure that buffers used were not inhibitory to the deglycosylational enzymes, the fetuin protein (10μg) was incubated in 3 alternative buffers in the presence and absence of enzymes (FIGURE 23B). Following deglycosylation, the results show that the fetuin protein decreased in size from 72kDa to 52kDa in the presence of three buffers.

Once it was confirmed that the enzymes were functional under these conditions, deglycosylation experiments were carried out on the stably transfected cell lysates. No change in size was observed following deglycosylation and immuno-detection of HEK-293 and MCF7 stably transfected cells (Figure 24). The results show that 17kDa and

22kDa previously observed (Figure 15) were not modified by deglycosylation. Deglycosylation of BT-20 cell lysate, followed by Western blotting with monoclonal antibody H39C51, did not result in the detection of an endogenous SBEM protein. This antibody was used as it positively detected both the 17 and 22kDa SBEM bands.

5.7.2. Enzymatic Deglycosylation of Secreted SBEM Protein

We next wanted to determine whether the secreted SBEM protein was glycosylated. SBEM isolated from transfected HEK-293 culture media was enzymatically deglycosylated and examined by Western blot. Following deglycosylation, the 25-33kDa SBEM band decreased in size by approximately 3kDa, to 23-30kDa (Figure 25A). Increasing the amount of enzyme 2-fold did not alter the size or degree of smearing of the downshifted band. In addition, deglycosylation was also performed on culture media derived from stably transfected MCF-7 cells (Figure 25B). No change in the SBEM protein size was observed.

VI) DISCUSSION

Breast cancer is the most frequently diagnosed cancer in women and is the second leading cause of cancer-related deaths in developed countries (Cancer Facts and Figures, 2005). Early detection is an important goal in breast cancer treatment, to enable intervention at a potentially curable stage and to maximize the opportunity for breast conservation. The 5-year survival rate is approximately 75% for women with locally advanced breast cancer. However, if the cancer metastasizes, the average survival time decreases to less than 2 years (Ingle *et al.*, 1999). Thus, there is a need to develop strategies, which can detect cancer at its early stages. The utilization of tumor markers has proven to be successful in the management of cancer, as their detection in inappropriate fluids or organs is indicative of disease. Currently, there are only a few tumor markers used to detect breast cancer. For example, the estrogen receptor is believed to herald benefit from tamoxifen-based chemotherapy and the HER-2 gene suggests the presence of aggressive cancer (Ryden *et al.*, 2005, Ross *et al*, 2003). Unfortunately, most current breast tumor markers are not tissue-specific and can only characterize a small subset of patients. Markers restricted to the mammary epithelium would have a significant impact not only to report on proliferative changes in the breast, but also to enhance the detection of micrometastatic disease by identifying cancer cells of breast origin in the lymph nodes.

6.1.1. SBEM Gene Expression

6.1.1. Normal Breast and Tumor Tissue

Normal breast and tumor tissues were examined by RT-PCR and Northern blot analysis in order to characterize SBEM gene expression. Among the normal breast tissues and breast tumors examined in this study, SBEM mRNA was observed in all cases (3 of 3 for normal breast, 9 of 9 for tumors) by RT-PCR. These results agreed with our lab's previous data, where a PCR product corresponding to SBEM was detected in a high proportion of invasive breast tumors (51 of 54). Fewer tumors (7 of 9) were positive for the 650-700 base SBEM transcript in the same tissues, where the levels of SBEM transcript varied between each sample. This variance in the level of gene expression was not surprising as Northern blot analysis is a less sensitive, but more quantitative technique than RT-PCR. Salivary gland shown previously to be positive for SBEM gene expression (Miksicek *et al.*, 2002), was negative in this study. Ethidium bromide staining of 28S and 18S ribosomal RNA bands revealed that the RNA for salivary gland was substantially degraded and of low quantity, likely explaining the absence of a detectable SBEM transcript by Northern blot. Differences in the level of SBEM mRNA detected may be attributed to differences in degradation of individual samples. The degradation rates of individual mRNAs can be regulated and affect mRNA abundance, therefore influencing how much of each protein is produced by translation. However, as a pre-cautionary measure against RNA degradation, all samples were kept on ice, treated with diethyl pyrocarbonate (DEPC)-treated water and all equipment was treated with RNase inhibitors. In this study, no substantial difference was observed between normal breast and breast tumor tissue, in terms of the level of SBEM mRNA expressed. Further

examination is required to delineate whether SBEM mRNA is expressed at a different level in breast cancer, compared with normal tissue.

6.1.2. Breast Cancer Cell Lines

Cancer cells are known to be genetically unstable. They can undergo specific genotypic/phenotypic alterations resulting from long-term storage, or they can differentially evolve and give rise to distinct sub-populations (Lacroix and Leclercq, 2004). Therefore it was necessary to examine cell lines previously documented by our group in order to confirm the presence or absence of SBEM gene expression in these cell lines (Miksicek *et al.*, 2002). In these studies SBEM transcripts were detected in MCF-7, ZR-75, T-47D and MCF10-A1 by RT-PCR, but not in MDA-MB-231 cells. The present study confirmed these results. Additional breast cancer cell lines were analyzed in this study. The breast cancer cell line BT-20 was found to be highly positive for SBEM mRNA expression, whereas both MDA-MB-468 and non-breast cancer HEK-293, were negative. Northern blot analysis was performed on the same panel of cell lines to examine the relative quantity of SBEM mRNA expressed in each cell line. BT-20 cells expressed the highest level of the SBEM gene expression of all the cell lines analyzed, while the ZR-75 cell line exhibited weak expression of the SBEM gene. These results illustrate that the SBEM gene is expressed in some cell lines and not in others, and that some cell lines have strong expression and others weak expression. Heterogeneity in cell line expression of the SBEM gene is not surprising as there is much heterogeneity between the cell lines. Some cell lines are more aggressive and have different hormone receptor (ER/PR) status (positive or negative). For example, BT-20, which expressed the

highest level of SBEM mRNA, are negative for estrogen receptor protein. However, unlike other ER-negative cell lines examined in this study (MDA-MB-231 and MDA-MB-468), BT-20 cells express ER mRNA (deletion in exon 5). All ER-positive breast cancer cell lines (ZR-75, T47-D, MCF-7 and MCF10-A1) were positive for SBEM mRNA expression by RT-PCR. The highly invasive, ER-negative MDA-MB-231 cell line was negative for SBEM mRNA expression. Thus, there could potentially be a link between ER status and SBEM expression.

6.2. SBEM Protein Expression

6.2.1. Immunohistochemistry (IHC) of Normal Breast and Tumor Tissues

To examine the expression and localization of the SBEM protein, immunohistochemical (IHC) analysis was performed on a small series of normal breast tissues and tumors. The H218C31 SBEM monoclonal antibody was used for IHC as it had been previously reported to detect SBEM protein in breast carcinomas (Colpitts *et al,* 2002). 1 of 4 of reduction mammoplasty samples and 9 of 10 breast tumors stained positive for SBEM protein expression. The degree of staining varied between breast tumor samples. Some cases expressed low levels (sample #8674 and 8706), while others expressed high levels (sample #8632 and 8921), which illustrates the amount of heterogeneity among the samples. A correlation was observed between IHC and RT-PCR data for breast tumors, where a high percentage of tumors expressed SBEM (9 of 10 and 9 of 9, respectively). Both assays indicate whether tissues are positive or negative for the particular protein/gene, however they are not quantitative.

IHC is useful in that it is based on the detection and visualization of protein molecules within the cell or on the cell surface. The pattern of SBEM protein staining was shown to be predominantly localized in the cytoplasm and membrane of ductal epithelial cells, with no nuclear staining, as expected. SBEM is not predicted to be localized in the nucleus as it doesn't possess a nuclear localization signal.

6.2.2. Western Blot Analysis of SBEM Protein Expression

6.2.2.1. Normal Breast and Tumor Tissues

The SBEM polypeptide sequence predicts a theoretical molecular mass of 9kDa. A panel of SBEM monoclonal antibodies targeting different regions of the SBEM polypeptide was utilized in Western blot experiments. The objective of Western blotting was to examine the SBEM protein status in normal breast and tumor tissue. In both normal tissues and tumors, no endogenous SBEM protein was detected by any of the monoclonal antibodies studied. These results did not correlate with previously discussed IHC results, where SBEM protein was positively detected in the majority of tumors (9 of 10) and 1 of 4 normal tissue samples. This suggested that SBEM protein was either absent or was too low to be detected by Western blotting using SBEM monoclonal antibodies. It was later determined in our lab that a faint SBEM protein band could be detected at 17kDa in one tumor sample by Western blot analysis (Ionela Gheorgiu, unpublished results). This sample (sample #8921) had the highest amount of SBEM gene expression by Northern blot and the highest degree of staining for SBEM protein by IHC in this study. This result showed that the endogenous SBEM protein was detectable,

albeit weakly, by Western blot analysis. Overall, SBEM protein expression does not reflect gene expression.

A recent study by Wadle *et al.* (2005) identified 2 forms of SBEM protein from a mixture of breast tumor samples using human SBEM-specific polyclonal antibodies, isolated from high titre patient sera. These proteins were identified as being a normal SBEM protein (90 aa) and a variant SBEM protein (82 aa). Perhaps human antibodies directed to the SBEM protein would have identified endogenous SBEM protein in the cohort of breast tissues and tumors examined in this study.

6.2.2.2. Breast Cancer Cell Lines

Western blot analysis was performed with the same breast cancer cell lines used in gene expression studies. A differentially expressed 33kDa band was detected in all breast cancer cell lines. This band was approximately 3.5 times larger than the predicted size of the SBEM protein (9kDa). Surprisingly, a 33kDa band was detected in the MDA-MB-231 cell line, which was negative for SBEM mRNA expression by Northern blot and RT-PCR. Also, the negative control cell line, HeLa, positively expressed this 33kDa band. These observations suggest that the 33kDa band detected by the polyclonal antibody was non-specific.

To further assess SBEM protein expression in breast cancer cell lines, a panel of monoclonal SBEM-antibodies targeting different regions of the SBEM protein sequence was used in Western blot analysis. No endogenous SBEM protein was detected in any breast cancer cell line by any of the different monoclonal antibodies. Surprisingly, even BT-20 cells were negative for SBEM protein expression, which were expected to express

detectable levels of SBEM protein due to their high mRNA expression. This suggests that SBEM protein levels do not reflect gene expression.

There are a number of plausible explanations why the SBEM protein was not detected by Western blotting with monoclonal antibodies. First, the amount of SBEM protein in each cell line may have been too low to be detected by any of the monoclonal antibodies. This appeared to be the case for breast tumors, where only one sample (which had an extremely high amount of SBEM mRNA and stained highly positive for SBEM protein by IHC) expressed detectable levels of SBEM protein. Second, difficulties in detecting endogenous SBEM by Western blotting may be due to epitope masking/shielding of the antigen site, which the antibodies were raised against. As SBEM contains 17 putative O-glycosylation sites, various forms of SBEM protein could exist resulting from different patterns of glycosylation, which could interfere with antibody recognition (Miksicek et al., 2002 and Colpitts et al., 2002). For example, in the MUC1 protein, the reactivity of antibodies that are directed to the peptide backbone are influenced by O-linked glycans, resulting in differential reactivity with MUC1 expressed by normal or malignant breast epithelium (Burchell *et al.*, 2001). In fact, in normal epithelium, antibodies show little to no reactivity for MUC1. A decrease in MUC1 glycosylation is observed in tumors, which exposes hidden epitopes and allows antibodies to recognize previously obscured antigens (Burchell *et al.*, 2001). In this study, monoclonal antibodies H39C51, H24C16 and H218C31, were raised against C-terminal peptides where no glycosylation is predicted to occur. Therefore, the potential for steric hinderance or epitope shielding of the antibodies as a result of glycosylation was expected to be lower. However, glycosylation has been shown to induce conformational changes to protein, which may have induced a masking

of the epitope recognition sites for these antibodies. Third, the stability or half-life of the protein may play a role. If the SBEM protein degrades quickly, it would not be detectable by Western blot. Also, the amount and pattern of glycosylation of the SBEM protein could also modify its half-life. In order to determine the half-life of the SBEM protein, pulse-chase labeling and/or proteasomal inhibition using an agent like cyclohexamide would have to be investigated. Inhibition of the proteasome would also show whether protein degradation was functional. Fourth, monoclonal antibodies were raised against short peptide sequences, as opposed to the full-length protein. Any changes to the protein structure in terms of glycosylation, signal peptide release or potential secondary structures, may restrict antibodies from binding their epitopes. Development of an antibody raised against the full-length peptide may increase detection of SBEM protein by Western blot.

6.2.2.3. Stably Transfected Cell Lines

As there was difficulty in detecting endogenous SBEM protein, it was logical to examine the expression of SBEM protein in cell lines stably transfected with the SBEM sequence. To address this, MCF-7 stably transfected cell lines were generated. HEK-293 stably transfected cells (a generous gift from Dr. T. Colpitts, Abbott Labs, Illinois, USA) were also examined. These cells provided a source of SBEM protein and were used as a positive control for all gene and protein studies. The C-terminal monoclonal antibodies, H39C51 and H24C16 detected a 22kDa band in MCF-7 transfected cell lysates. Using the same antibodies, a weak 17kDa band was observed in some experiments. SBEM protein was not detected in MCF-7 transfected cells by any other antibody. The H39C51

monoclonal antibody detected 17 and 22kDa bands in HEK-293 transfected cell lysates. The other C-terminal antibodies, H24C16 and H218C31 (also use in IHC studies), identified a 17kDa SBEM protein band (at strong and weak levels, respectively). The 22kDa band detected by H39C51 was not detected by the H24C16 antibody, even though both were raised against the same peptide. A 17 and 22kDa band was detected by Western blotting with monoclonal antibodies targeting the internal region of the SBEM protein sequence (H51C42 and H76C43). No endogenous SBEM protein was detected by Western blotting with N-terminal antibodies (H24C29 and H34C78). These data illustrated that not all monoclonal antibodies were able to detect the SBEM protein. Rather, it seemed that each antibody had a different binding affinity for SBEM, even when targeting the same region of the SBEM protein sequence. However, a later study examining HEK-293 transfected cell lysate with all monoclonal antibodies at a 2.5 fold lower dilution (1:200 instead of the 1:500 used in this study), resulted in the detection of a 17kDa SBEM protein in all cases (Ionela Gheorgiu, unpublished data). This suggested that concentration of the antibody used was crucial in detecting SBEM protein. This is result is somewhat surprising since the difference in antibody dilution was only 2.5 fold less (by Gheorgiu). A detectable level of SBEM protein still should have been identifiable, even at 1:500 dilution. Another consideration as to why SBEM protein was undetectable by some antibodies, is that these antibodies were produced by hybridoma cells and were collected in culture supernatant. The amount of actual antibody in the supernatant is very low due to the high volume of the media. Following these studies, it was later determined that purification of the H39C51 monoclonal antibody, prior to

immunodetection, drastically improved the detection of SBEM protein (Ionela Gheorgiu, unpublished results).

Theoretically, SBEM protein produced by stably transfected MCF-7 and HEK-293 cell lines should be 11kDa in size, if one considers the V5-HIS tag (~2kDa) in the MCF-7 transfected cell line or the myc-HIS tag (~2kDa) in the HEK-293 transfected cell line. The visualization of a 17 kDa and a 22kDa protein following Western blot analysis, suggests that post-translational modifications may be partly responsible for the different sizes, which were detected. For example, the SBEM protein was predicted to be highly O-glycosylated. The addition of a number of glycans, would greatly increase the size of the protein. To address whether glycosylation contributed to the 17 and 22kDa bands detected in MCF-7 transfected cell lysates, *in vitro* translation of SBEM plasmid constructs was performed. Any post-translational modifications of the protein (such as phosphorylations or glycosylations) will not take place using the in vitro translation procedure. A 17kDa band was detected for the construct generating a tag-less SBEM protein, while a 22kDa band was identified for the construct generating SBEM protein with the V5-His tag. This study provided evidence that the higher-than-expected SBEM protein size was not due to glycosylation.

As the SBEM protein sequence contained one cysteine residue, and the detectable 22kDa band was double the theoretical size, one can hypothesize that SBEM protein dimerizes through a disulfide linkage. Dimerization is unlikely though, as the reducing agent β-mercaptoethanol was used in all Western blotting experiments. β-mercaptoethanol breaks apart disulfide bridges, reducing dimers to their monomeric units. A later study was performed in our lab, where the cysteine residue was mutated to

eliminate any disulfide bridging (Florent Hube, unpublished results). *In vitro* translation of these cysteine-mutated SBEM constructs also produced the same 17 and 22kDa bands, providing further evidence that the SBEM protein was not dimerized (unpublished results). However, a study by Chalut al. (1999) showed that the penicillin-binding protein was able to dimerize, even after cysteine residues were mutated. Thus, the SBEM protein may still dimerize by strong non-covalent interactions, by a mechanism other than disulfide bridging. Another consideration is that though the SBEM protein is theoretically predicted to be 9kDa (11kDa with tag), it is possible that it simply migrates higher than theoretically expected.

A recent study has documented SBEM protein expression in HEK-293 cells transiently transfected with the SBEM sequence (Wadle *et al.*, 2005). Using Western blot analysis, this group detected SBEM protein at 22kDa and a variant SBEM protein detected at 20kDa. This result confirms the 22kDa SBEM protein observed in this study. The 20kDa variant protein was shown to lack 8 amino-acids (Ala, Thr, Thr, Ala, Thr, Thr, Ala, Ala), a sequence corresponding to a portion of the tandem repeat motif. The detection of an SBEM variant is encouraging, as it may explain why a 17kDa band was also detected by this study's Western blot results. Though SBEM protein was detected at larger than predicted molecular weights, Wadle *et al.* (2005) did not discuss any potential factors contributing to this size.

6.3. Immunoprecipitation (IP) of Breast Cancer Cell Lines and Stably Transfected Cell Lines

In an attempt to specifically bind endogenous SBEM protein, immunoprecipitation (IP) of selected breast cancer cell lines was performed using SBEM monoclonal antibodies. BT-20 and ZR-75 cells were examined as they exhibited the highest level of SBEM mRNA expression (Northern blot data). A faint band was detected at 13kDa by the H39C51 monoclonal antibody in BT-20 cell lysate, close to the theoretical size of 9kDa. Though peptide neutralization suggests that this band is non-specific, the neutralization experiment was performed on the same blot. To address this, a peptide neutralization experiment should be performed on a separate blot. When IP was performed on BT-20 cells using a different antibody (H24C16), this 13kDa band was not visible. Further investigation needs to be performed to determine whether endogenous SBEM protein can be detected in BT-20 cells. No endogenous SBEM protein could be detected by the H39C51 or H24C16 monoclonal antibodies (C-terminal) in ZR-75 protein lysates. These results were in agreement with Western blot data.

Immunoprecipitation of HEK-293 transfected cells resulted in the detection of a strong 17kDa band and a diffused 22kDa band using the H39C51 monoclonal antibody. This confirmed Western blot analysis data of the same cell line. It is likely that the 22kDa band was poorly resolved due to its proximity to the light-chain IgG band at 25kDa. This resolution did not improve even when samples were run for a longer duration.

6.4. Sub-localization of SBEM Protein in Different Cell Fractions of Stably Transfected Cell Lysate

HEK-293 stably transfected cells were fractionated to address the sub-cellular localization of the SBEM protein. 17 and 22kDa SBEM proteins were detected in the cytoplasmic and membranous fractions with H39C51 and H24C16 antibodies. These results correlated with IHC examination of normal breast and tumor tissue, where SBEM protein was also identified in the cytoplasm and membrane of ductal epithelial cells. The H39C51 and H24C16 antibodies have also been shown to successfully detect SBEM protein in tumor tissues by IHC (Colpitts *et al.*, 2002). Of note, a greater level of SBEM protein was detected in the membranous fraction of these cells. This is in concordance with Wadle et al. (2005), who stated that SBEM was detected as a membrane-bound protein in HEK-293 cells.

Proteins have been shown to be differentially localized at different phases of the cell cycle. For example, the tumor suppressor, p53, has been shown to be expressed in the nucleus in G1 and in the cytoplasm during S and G2 (Moll et al., 1996). Synchronization of breast cancer cells at a particular point, by various methods (isoleucine deprivation [G1], double thymidine synchronization [G1/S]) could help determine where the SBEM protein is localized in relation to a specific cell cycle point. However, uniform synchronization may be difficult to achieve since there is much heterogeneity within breast cancer cell populations.

6.5. Secretion of the SBEM Protein

The SBEM protein sequence has a signal peptide in the first 19 N-terminal amino acid residues, indicating the protein is secreted. To address secretion, stably transfected cell line culture media was analyzed for SBEM protein. Using monoclonal antibodies, H39C51 and H24C16, Western blot analysis revealed a secreted 25-33kDa broad diffused band in HEK-293 media and a narrow 26kDa band in MCF-7 media. These results are supported by a recent study examining secreted SBEM protein expression in cell culture supernatant (Wadle *et al.*, 2005). Using HEK-293 cells transiently transfected with cDNA coding for SBEM and SBEMv (a variant of SBEM lacking 24bp [+149 after the translation initiation site] and harboring one silent gene mutation just upstream of the deletion), secreted SBEM proteins were detected by Western blot analysis as broad diffused bands at 25/27kDa (SBEMv) and 33/35kDa (SBEM).

By identifying SBEM as a secreted protein, it is plausible to suggest that breast cancer patients positively expressing SBEM protein in their tissue (tumor or normal) may also secrete SBEM protein into their blood. In this respect, SBEM may resemble the prostate-specific antigen (PSA) (Ward *et al.*, 2001). PSA is a product of the normal prostate gland, and its expression and serum concentration increase in conditions such as prostatic hyperplasia and prostate cancer accounting for its diagnostic value in these diseases (Linton *et al.*, 2003; Ward *et al.*, 2001). The development of enzyme-linked immunosorbent assays (ELISAs) will likely help in providing evidence as to whether this is the case.

6.6. Glycosylation Status of the SBEM Protein

There are structural similarities between mucin family glycoproteins and the SBEM protein. Mucins are characterized by a tandemly paired and repetitive central peptide rich in threonine and serine residues. Mucins are also predicted to be highly O-glycosylated and they contain signal peptides, giving rise to secreted proteins. Mucin glycoproteins have been shown in patient serum when tissues become cancerous. These mucins are heavily O-glycosylated at serine and threonine residues, and these sugars organize themselves into a complex network of branching structures (Gendler, 2001). In malignancy, where the MUC1 isoform is over-expressed, glycosylation is perturbed and sugar moieties not normally expressed are synthesized (Brockhausen *et al.,* 1995). To assess the glycosylation status of the SBEM protein, the secreted proteins obtained from MCF-7 and HEK-293 transfected culture media were isolated and enzymatically deglycosylated. A downshift of approximately 3kDa was observed following deglycosylation of the 25-33kDa protein from HEK-293 media. However, no difference was observed for secreted SBEM protein from MCF-7 transfected cells. The 3kDa downshift after deglycosylation suggests that the size of the SBEM protein is partly due to glycosylation. However, as the SBEM protein is predicted to be highly glycosylated, it was expected that a more pronounced downshift in size would be seen after deglycosylation.

Secreted SBEM protein was obtained from 2 different sources, a breast cancer cell line and a non-breast cancer cell line. Thus, it seems reasonable to assume that the glycosylation pattern between these proteins is different. One way to address this would be to culture these 2 cell lines in media containing a glycosylation inhibitor and then

examine resultant "sugar-less" native secreted proteins. Tunicamycin and benzyl-ON-acetyl-D-galactosamine are examples of agents, which inhibit glycosylation. Tunicamycin prevents N-linked glycosylation, while benzyl-ON-acetyl-D-galactosamine prevents O-linked glycosylation.

The size of the 17 and 22kDa SBEM protein bands detected in transfected cell line lysates did not change following enzymatic deglycosylation. Indeed, this result is in agreement with *in vitro* translation data, which suggests that glycosylation is not a contributing factor to the size of SBEM protein in cell lysates.

6.7. Conclusions

In conclusion, this study showed that levels of the SBEM gene were variable in breast cancer. Different levels of the SBEM gene are expressed among breast cancer cell lines. Differences in the level SBEM gene expression between normal breast tissues and breast tumors could not be determined. Though an endogenous SBEM protein could not be detected in breast cancer tissues and breast cancer cell lines by Western blotting, SBEM protein was detectable in cell lines stably transfected with SBEM. Corresponding SBEM SBEM proteins were found to be over-sized, compared with the theoretically predicted size. A secreted SBEM protein was identified. The identification of a secreted SBEM protein suggests that SBEM may also be secreted into the serum of breast cancer patients. The secreted SBEM protein appeared to be partly glycosylated in HEK-293 transfected cells. Though differences in glycosylation patterns between normal patients and breast cancer patients have yet to be elucidated, differential glycosylation patterns of the SBEM

protein may yield tumor-specific antigens. Thus, SBEM is a strong candidate for a marker, which may be detected in the serum of breast cancer patients.

6.8. Summary

- Differential patterns of SBEM gene expression exist. Levels of the SBEM transcript vary among breast carcinomas and breast cancer cell lines.
- The BT-20 breast cancer cell line expressed the highest level of the SBEM gene.
- Immunohistochemical analysis showed SBEM protein expression was restricted to the cytoplasm and membrane of ductal epithelial cells in normal breast tissues and tumors.
- SBEM protein was detected at 17 and 22kDa in stably transfected cell lines HEK-293 and MCF-7, larger than the predicted size of 11kDa.
- Secreted SBEM protein was identified in stably transfected cell lines. The secreted SBEM proteins were detected at 26kDa (MCF-7) and between 25-33kDa (HEK-293), larger than the 17 and 22kDa SBEM proteins observed in cell lysates.
- The size of the secreted SBEM protein (25-33kDa) from HEK-293 cells, was partly due to the presence of glycosylation.

VII) FUTURE DIRECTIONS

1) No endogenous SBEM protein was detected by Western blot analysis or immunoprecipitation. Investigation of breast cancer cell lines, normal breast and tumor tissues with an antibody generated against the full-length SBEM peptide may enhance detection by these same methods, where shorter peptides may be masked or blocked from binding to their epitope. Also, antibodies derived from human serum have been documented to successfully detect SBEM protein (Wadle *et al.*, 2005).

2) Differences in the level of SBEM gene expression between normal breast tissues and tumor tissues have yet to be determined. A real-time PCR study examining breast tumors and their matching normal tissues will allow us to determine whether the quantity of SBEM gene expression is altered in breast cancer.

3) To date, the perfect circulating tumor marker for breast cancer cells has been elusive. Some proteins, like MUC-1 and the HER-2 extra-cellular domain are shed in the serum and plasma and the measurement of their circulating levels are used to monitor presence and progression of the disease. This study has proven SBEM protein to be secreted. Secretion of SBEM protein by cancer cells may lead to an increased concentration in blood, making SBEM a functional tumor marker. This could be determined by quantifying serum SBEM protein in the sera of breast cancer patients by enzyme-linked immunosorbent assay (ELISA), using SBEM-specific antibodies isolated from breast cancer patient serum.

Figure 1 – Breast cancer progression. The pathologically-defined stages of breast cancer progression are depicted (adapted from Ma et al., 2003). (A) Normal luminal epithelium. (B) Atypical ductal hyperplasia, (C) Ductal carcinoma *in situ*, (D) Invasive intraductal carcinoma.

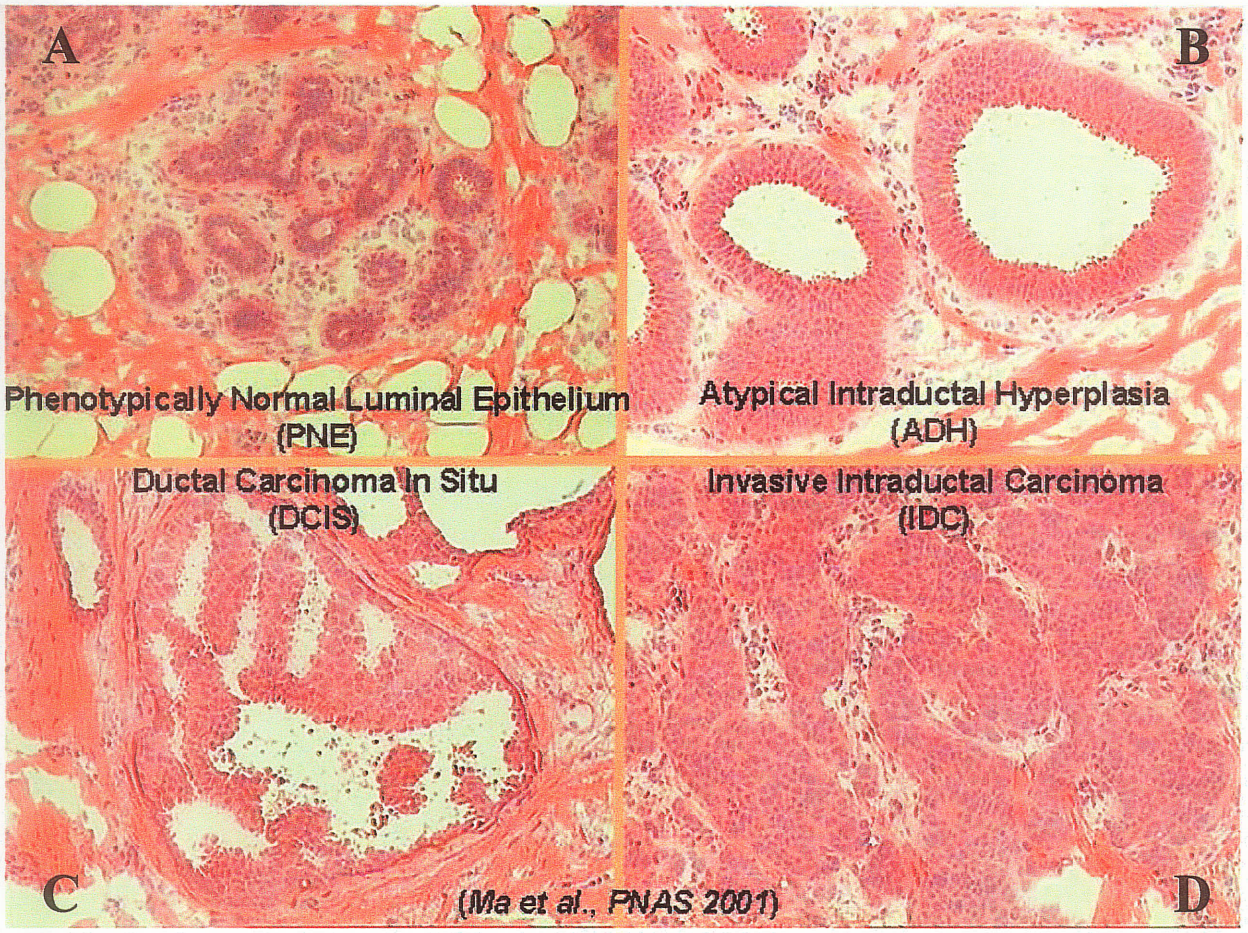
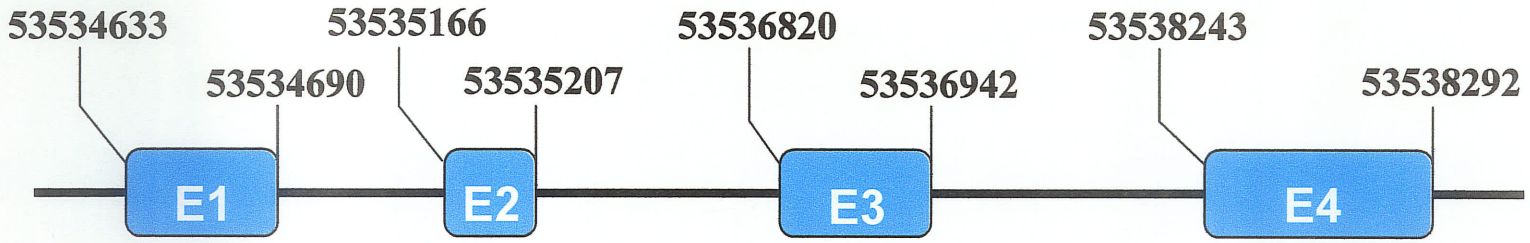

A
Phenotypically Normal Luminal Epithelium
(PNE)
B
Atypical Intraductal Hyperplasia
(ADH)
Ductal Carcinoma In Situ
(DCIS)
Invasive Intraductal Carcinoma
(IDC)
C
(Ma et al., PNAS 2001)
D

Figure 2 - Schematic representation of the SBEM gene. The genomic SBEM DNA sequence is found on chromosome 12 (A). The gene is transcribed into an RNA transcript containing 4 exons and 3 introns with a size of 3.9kb. Splicing of introns gives the resulting mRNA transcript (B). Start codon (ATG) for translation is found in exon 1, while the stop codon (TGA) is found in exon 4, resulting in a 270 base coding sequence (CDS).

A


53534633
53534690
53535166
53535207
53536820
53536942
53538243
53538292
E1
E2
E3
E4

Chr 12q

B

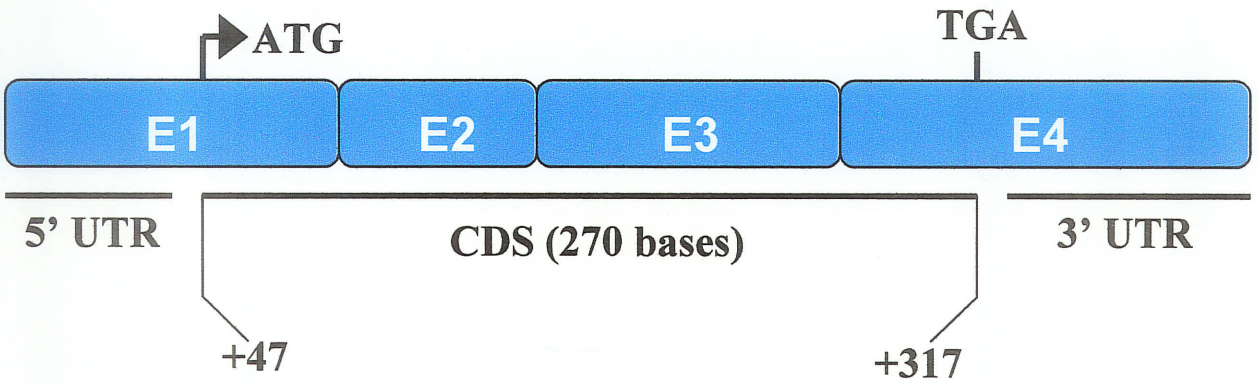

ATG
TGA
E1
E2
E3
E4
5' UTR
CDS (270 bases)
3' UTR
+47
+317

Figure 3 – Predicted polypeptide based on the SBEM cDNA sequence. The predicted hydrophobic signal peptide is depicted by line over residues 1-19, threonine (T) and serine (S) residues predicted to be O-glycosylated are shown in blue, and the octapeptide, the tandem core repeat sequence is boxed.


Signal Peptide
MKFLAVLVLLGVSIFLVSAQNPTTAAPADTYPATGPADDEAPDAE
TTAAATTATTAAPTTATTAASTTA
RKDIPVLPKWVGDLPNGRVCP
Tandem Repeat Motif

Figure 4 - Prediction of O-linked glycosylation of SBEM protein. A graphical representation of predicted O-linked glycosylation at threonine and serine residues of the SBEM polypeptide. O-glycosylation probabilities greater than 50% (0.5) are considered to be positive for O-glycosylation. This graph was generated from the Net-O-Glyc Version 3.1 website: www.cbs.dtu.dk/services/NetOGlyc. 16 of 16 threonine residues and 1 of 3 serine residues are predicted to be O-glycosylated.

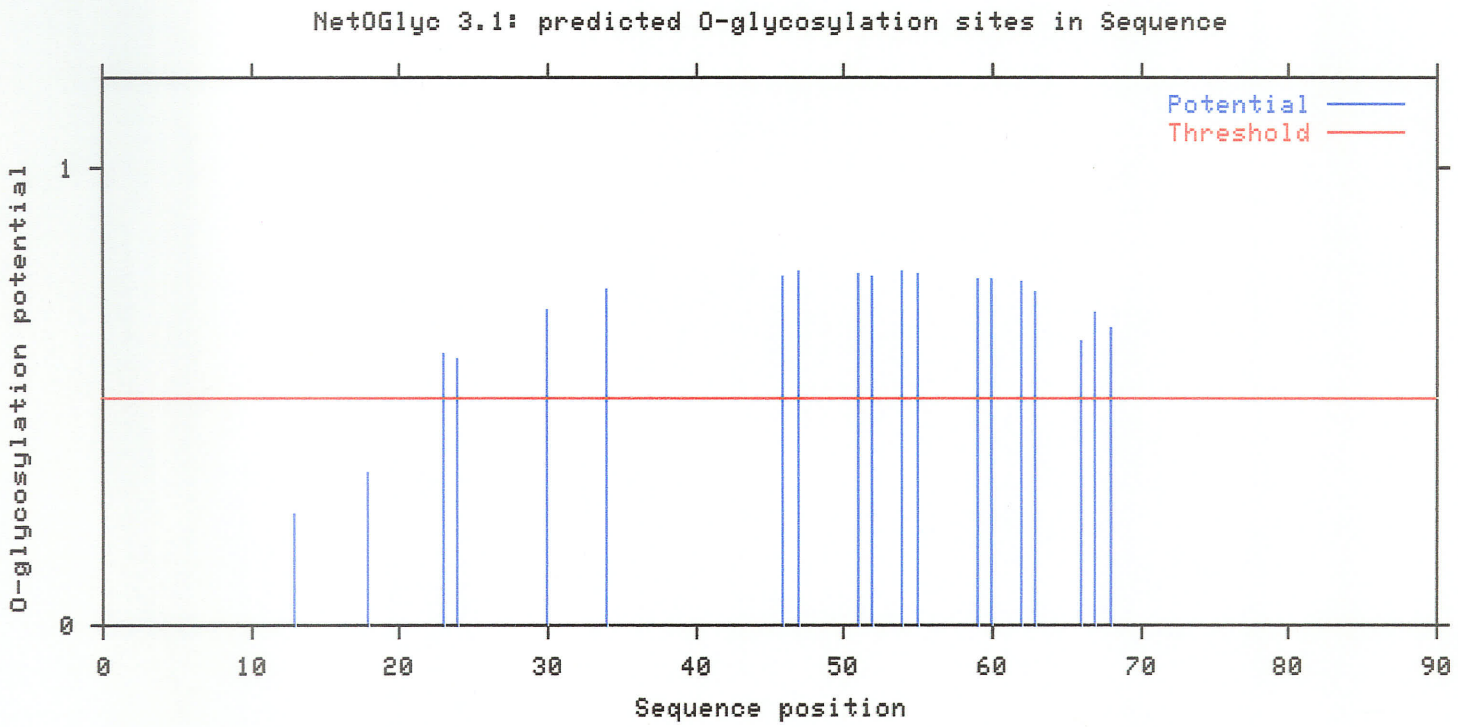

NetOGlyc 3.1: predicted O-glycosylation sites in Sequence
Potential
Threshold
O-glycosylation potential
1
0
0
10
20
30
40
50
60
70
80
90
Sequence position

Figure 5 – Peptide mapping of various antibodies used for Western blot and immunohistochemical analysis of the SBEM polypeptide. A schematic representation of the numerous antibodies utilized in our experiments. Antibodies were raised against different regions of the SBEM peptide sequence to alleviate potential antibody recognition problems due to the unknown structural conformation of the SBEM protein. Polyclonal anti-SBEM was raised in rabbit, against the 70-87 residue C-terminal sequence. All monoclonal antibodies used were a generous gift from Dr. Tracey Colpitts (Abbott Laboratories, Abbott Park, Illinois USA). The C-terminal antibody, H218C31, was raised against residues 69-89, while H39C51 and H24C16 were both raised against residues 69-90, with a cysteine added to alanine 69. Internal sequence antibodies H76C43 and H51C42 were described as being raised against residues 20-57, with a cysteine bound to alanine 57 and to another peptide spanning residues 25-45. N-terminal antibodies H24C29 and H34C78, were raised against residues 20-57. Of note, no antibody was raised against the signal peptide region, amino acids 1-19, as this region is expected to be cleaved off.

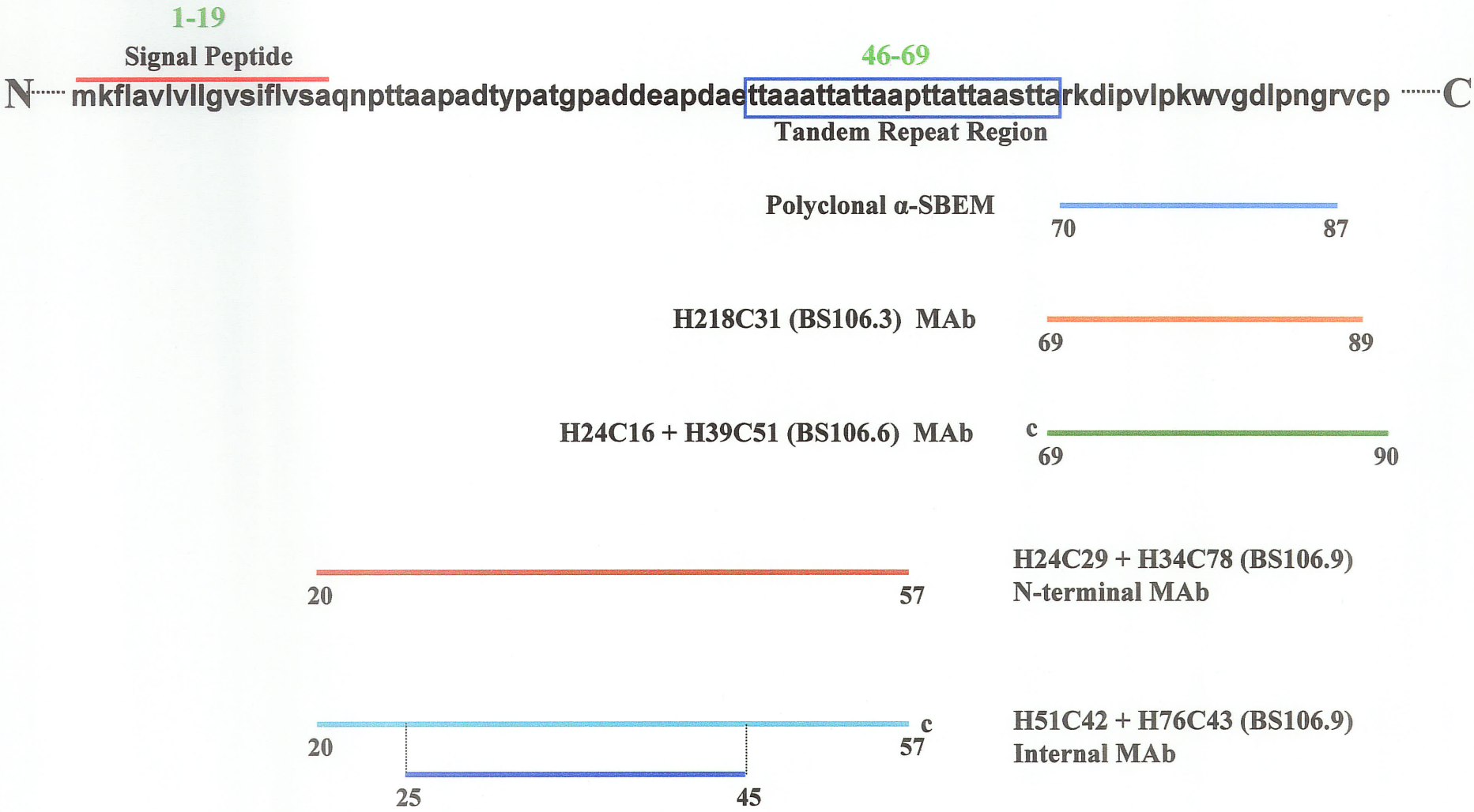

1-19
Signal Peptide
46-69
N
mkflavlvllgvsiflvsaqnpttaapadtypatgpaddeapdaettaaattattaapttattaasttarkdipvlpkwvgdlpngrvcp
C
Tandem Repeat Region
Polyclonal α-SBEM
70
87
H218C31 (BS106.3) MAb
69
89
H24C16 + H39C51 (BS106.6) MAb
c
69
90
20
57
H24C29 + H34C78 (BS106.9)
N-terminal MAb
20
c
57
25
45
H51C42 + H76C43 (BS106.9)
Internal MAb

Figure 6 - RT-PCR analysis of SBEM gene expression in the normal human mammary gland

Total RNA was extracted from different normal human mammary glands and RT-PCR was preformed to amplify the SBEM cDNA. PCR products were visualized on a 2% agarose gel, stained with ethidium bromide. (**A**) A 215bp product was detected in all 3 specimens examined. Normal breast (A); commercial sample (Ambion), normal breast ; (B); commercial sample (BD Biosciences), normal breast (C); reduction mammoplasty specimen. A 215bp product was detected in the positive control, stably-transfected HEK-293 cells. Water (H_20) was used as a negative control for the PCR reaction. PhiX174 RF DNA/Hae III DNA ladder was used for sizing of PCR products. To confirm equal loading, the GAPDH gene was amplified and was detected at 178bp for each sample (**B**). The volumetric analysis values of SBEM were normalized against GAPDH and are illustrated as the ratio of SBEM/GAPDH (**C**). SBEM/GAPDH ratios were then normalized against the SBEM/GAPDH ratio of HEK-293 stably transfected cells, given a value of 1. Experiments were performed in triplicate. Standard error bars are indicated.

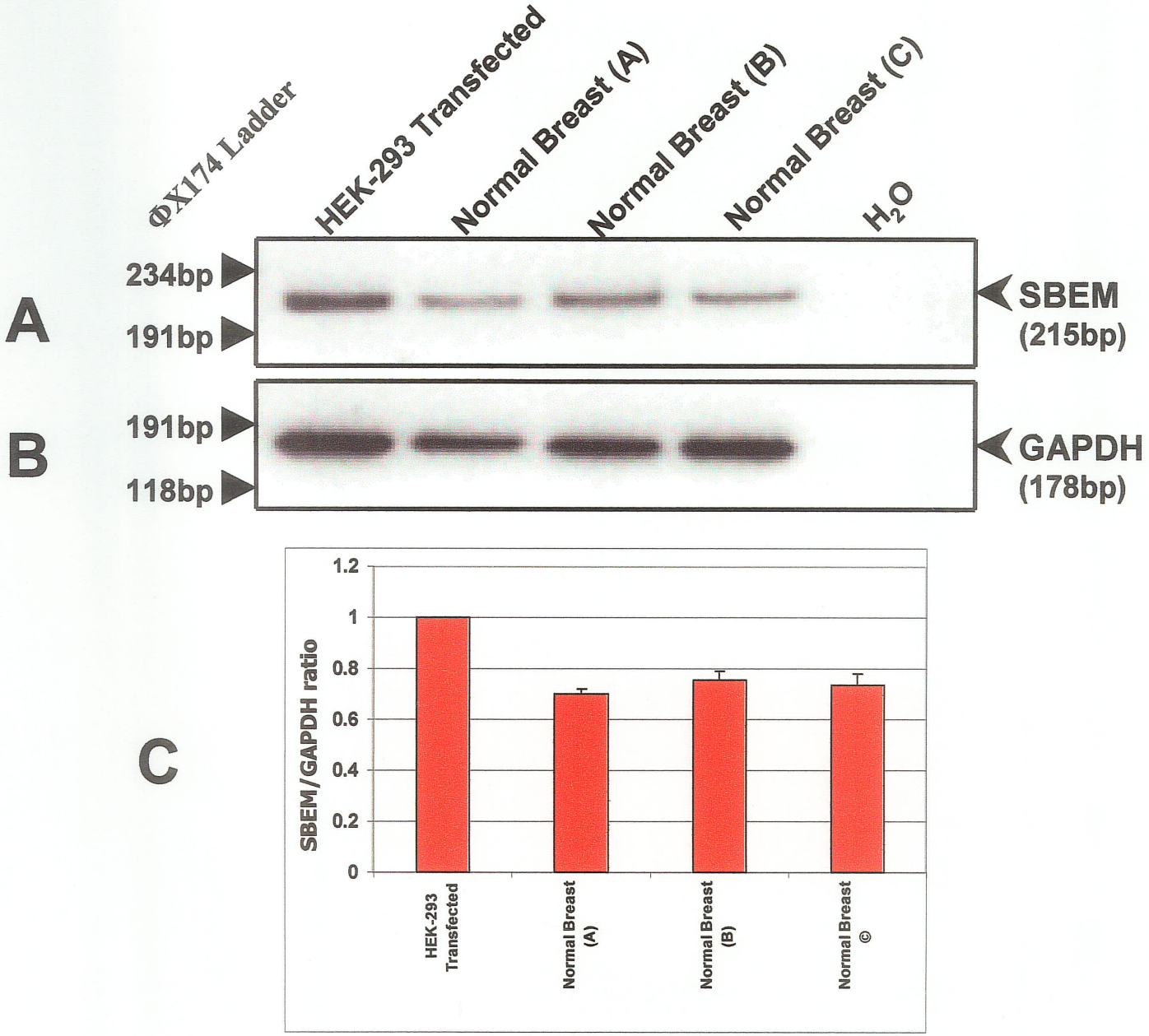

ΦX174 Ladder
HEK-293 Transfected
Normal Breast (A)
Normal Breast (B)
Normal Breast (C)
H2O
A
234bp
191bp
SBEM
(215bp)
B
191bp
118bp
GAPDH
(178bp)
C
SBEM/GAPDH ratio
1.2
1
0.8
0.6
0.4
0.2
0
HEK-293 Transfected
Normal Breast (A)
Normal Breast (B)
Normal Breast ©

Figure 7 - Northern blot analysis of SBEM gene expression in the normal human mammary gland

Total RNA (40μg) extracted from various normal tissues were electrophoresed on a 1% agarose gel denaturating agarose-formaldehyde gel, transferred and hybridized with a ^{32}P –labelled 291bp PCR product SBEM cDNA probe (**A**). A 650-700 base mRNA transcript was observed in both samples of normal breast tissues (Ambion and BD Biosciences). A 900-950 base mRNA transcript was detected in the stably transfected positive control cell line, HEK-293. Human pituitary gland was used as a negative control. The salivary gland, which was positive for SBEM in a previous study (Miksicek *et al.*, 2002), was also negative. To determine RNA loading and quality, the 28S (4.3kb) and 18S (1.8kb) ribosomal RNA bands were examined (**B**).

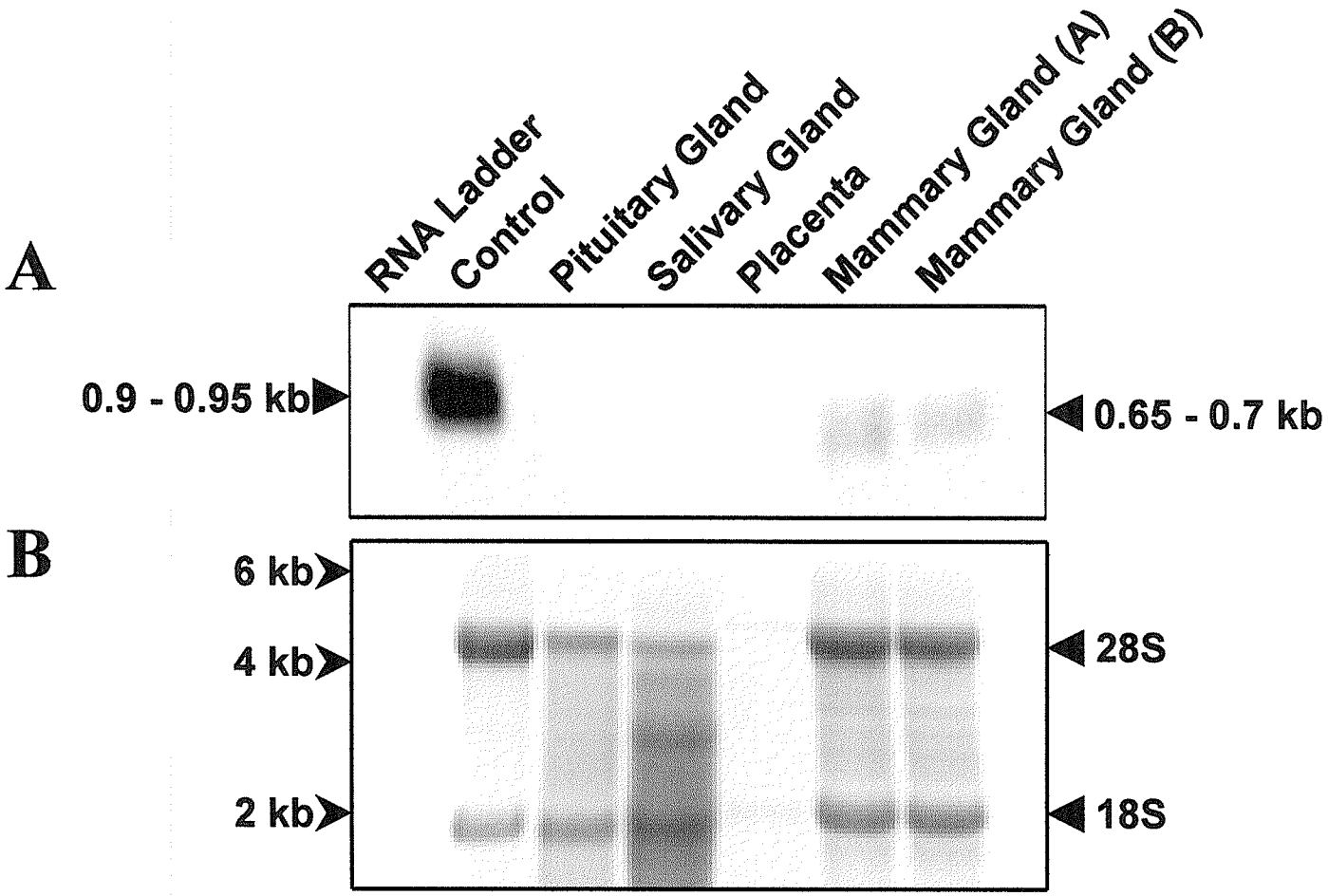

A
RNA Ladder
Control
Pituitary Gland
Salivary Gland
Placenta
Mammary Gland (A)
Mammary Gland (B)
0.9 - 0.95 kb
0.65 - 0.7 kb
B
6 kb
4 kb
2 kb
28S
18S

Figure 8 - RT-PCR analysis of SBEM gene expression in breast cancer cell lines

Total RNA was extracted from cell lines and RT-PCR was performed to amplify SBEM cDNA. PCR products were visualized on a 2% agarose gel stained with ethidium bromide. (**A**) A 215bp product was detected in BT-20, MCF-7, ZR-75, T47-D and MCF10-A1 cells, with BT-20 expressing the highest level. A 215bp product was detected in the stably-transfected HEK-293 cell line positive control. PhiX174 RF DNA/Hae III DNA ladder was used for sizing of PCR products. To confirm equal loading, the GAPDH gene was amplified and was detected at 178bp for each sample (**B**). Cervical cancer cell line, HeLa, was used as a negative control for SBEM expression. HEK-293 untransfected cells were used as a control for HEK-293 transfected cells. H_20 was used as a negative control for the PCR reaction. The volumetric analysis values of SBEM were normalized against GAPDH and are illustrated as the ratio of SBEM/GAPDH (**C**). SBEM/GAPDH ratios were then normalized against the SBEM/GAPDH ratio of HEK-293 stably transfected cells, given a value of 1. Experiments were performed in triplicate. Standard error bars are indicated.

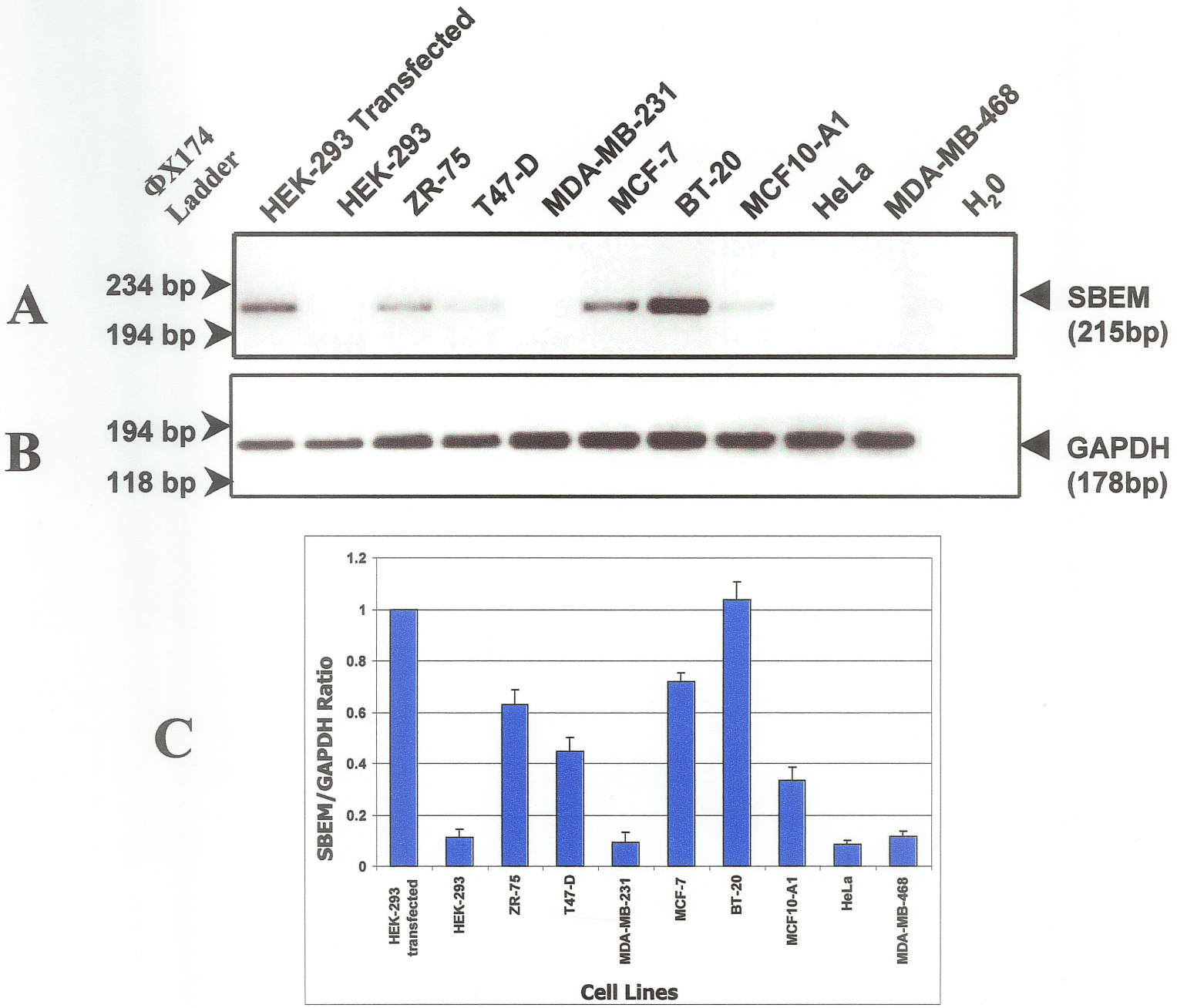

ΦX174 Ladder
HEK-293 Transfected
HEK-293
ZR-75
T47-D
MDA-MB-231
MCF-7
BT-20
MCF10-A1
HeLa
MDA-MB-468
H2O
A
234 bp
194 bp
SBEM (215bp)
B
194 bp
118 bp
GAPDH (178bp)
C
SBEM/GAPDH Ratio
1.2
1
0.8
0.6
0.4
0.2
0
HEK-293 transfected
HEK-293
ZR-75
T47-D
MDA-MB-231
MCF-7
BT-20
MCF10-A1
HeLa
MDA-MB-468
Cell Lines

Figure 9 – Northern blot analysis of SBEM gene expression in breast cancer cell lines

Total RNA (40μg) extracted from a panel of breast cancer cell lines were electrophoresed on a 1% agarose gel denaturating agarose-formaldehyde gel, transferred and hybridized with a ^{32}P –labelled 291bp PCR product SBEM cDNA probe (**A**). A 650-700 base mRNA transcript was observed in BT-20 and ZR-75 (weakly) cells. A 900-950 base mRNA transcript was detected in stably transfected positive control cell line, HEK-293. The cervical cancer cell line, HeLa, was used as a negative control. To determine RNA loading and quality, the 28S (4.3kb) and 18S (1.8kb) ribosomal RNA bands were visualized (**B**).

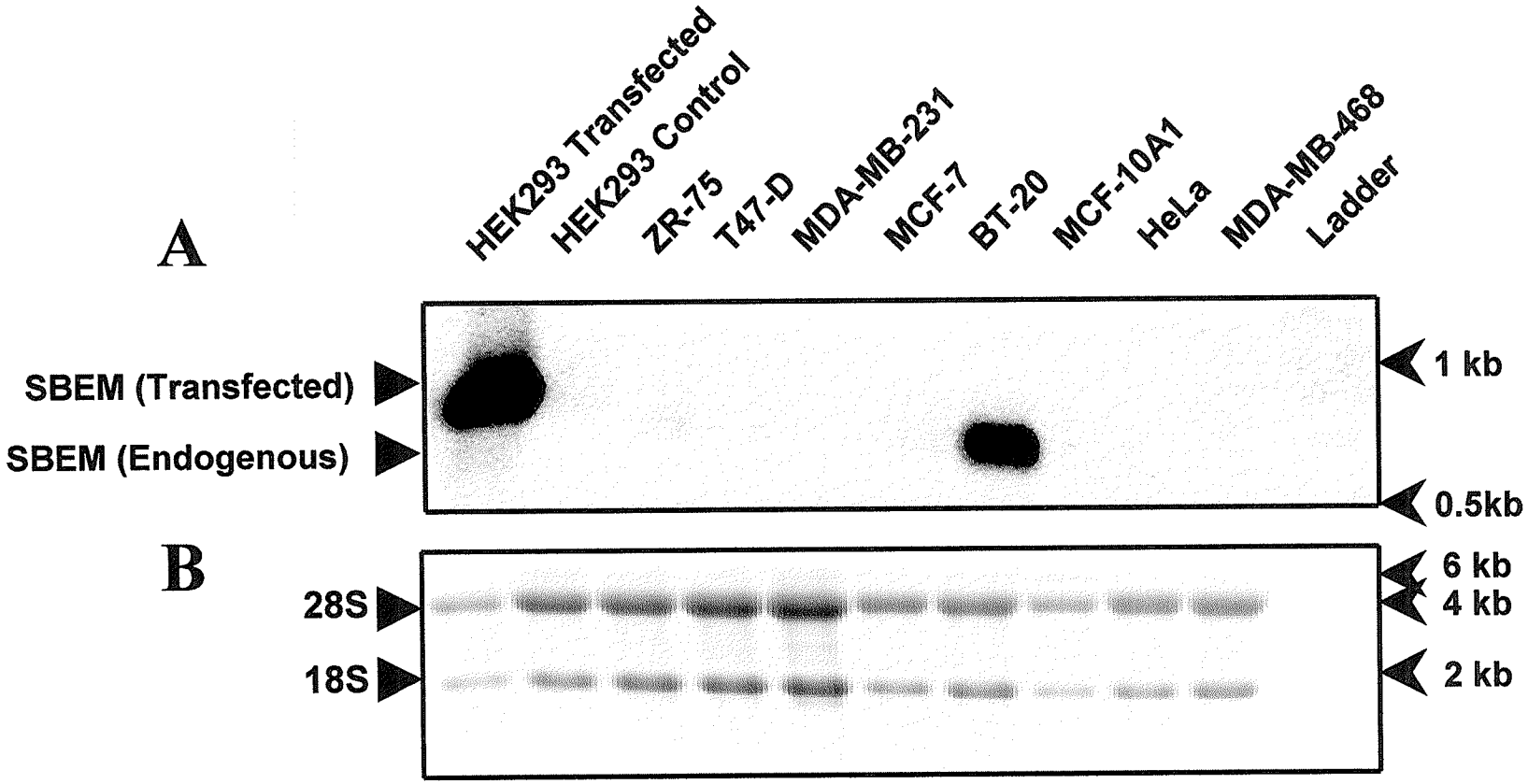

A
HEK293 Transfected
HEK293 Control
ZR-75
T47-D
MDA-MB-231
MCF-7
BT-20
MCF-10A1
HeLa
MDA-MB-468
Ladder
SBEM (Transfected)
SBEM (Endogenous)
1 kb
0.5kb
B
28S
18S
6 kb
4 kb
2 kb

Figure 10 - RT-PCR analysis of SBEM gene expression in breast tumors

Total RNA was extracted from 9 tumor specimens and RT-PCR was preformed to amplify the SBEM cDNA. PCR products were visualized on a 2% agarose gel and stained with ethidium bromide. (**A**) A 215bp product was detected in all samples examined. Stably-transfected HEK-293 cells were used as a positive control. PhiX174 RF DNA/Hae III DNA ladder was used for sizing of PCR products. To confirm equal loading, the GAPDH gene was amplified and was detected at 178bp for each sample (**B**). Glioma cell line, U87, was used as a negative control for SBEM expression. H_20 was used as a negative control for the PCR reaction.

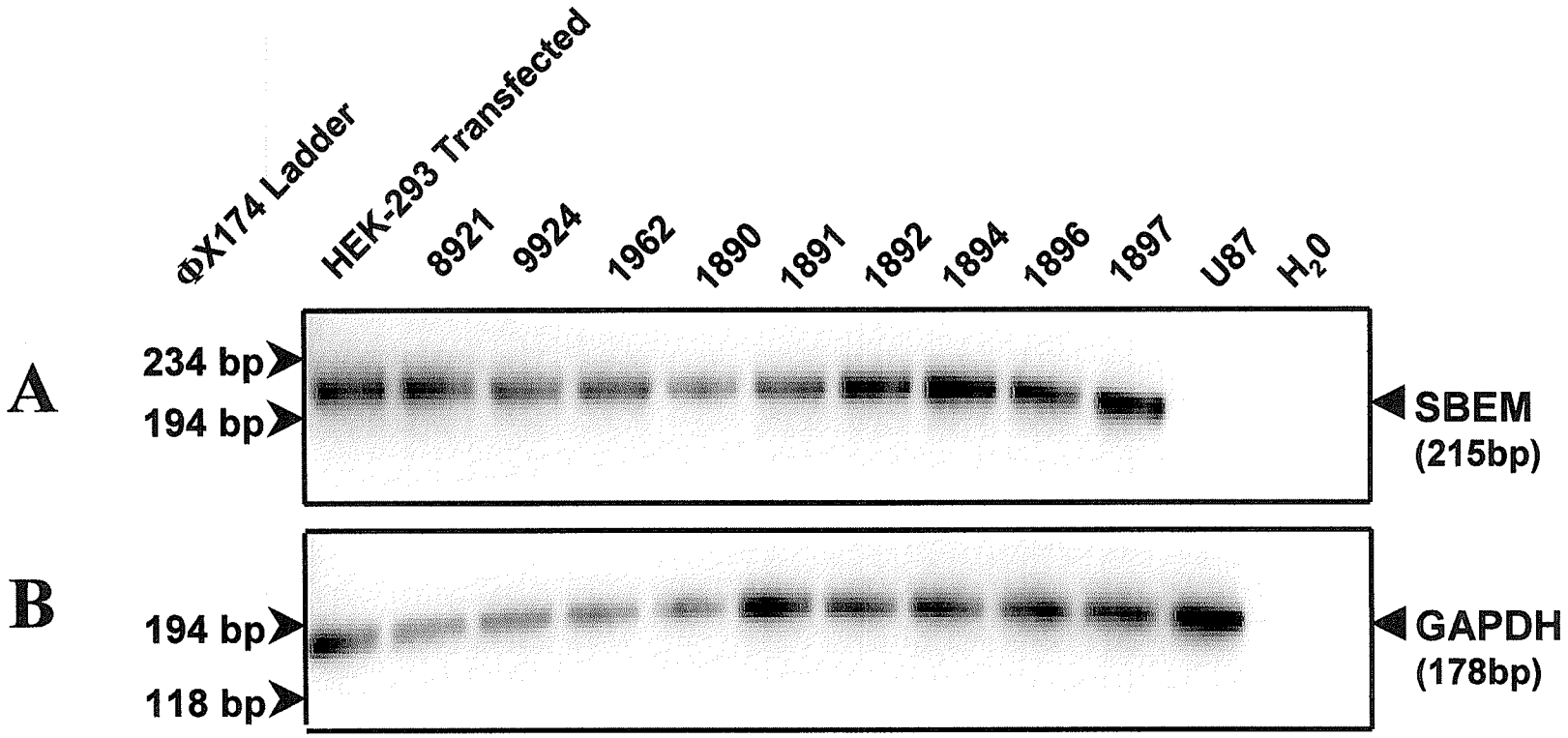

ΦX174 Ladder
HEK-293 Transfected
8921
9924
1962
1890
1891
1892
1894
1896
1897
U87
H20
A
234 bp
194 bp
SBEM
(215bp)
B
194 bp
118 bp
GAPDH
(178bp)

Figure 11 - Northern blot analysis of SBEM gene expression in breast tumors

Total RNA (9μg) extracted from 9 breast tumors specimens were electrophoresed on a 1% agarose gel denaturating agarose-formaldehyde gel, transferred and hybridized with a ^{32}P –labelled 291bp SBEM PCR product cDNA probe (**A**). A 650-700 base transcript was observed in 7 of the 9 samples. An approximately 375 base transcript was detected in sample 1962, in addition to the 650-700bp mRNA transcript. A 900-950 base mRNA transcript was detected in stably transfected positive control cell line, HEK-293. U87 glioma cells were used as a negative control. To determine RNA loading and quality, the 28S (4.3kb) and 18S (1.8kb) ribosomal RNA bands were examined (**B**).

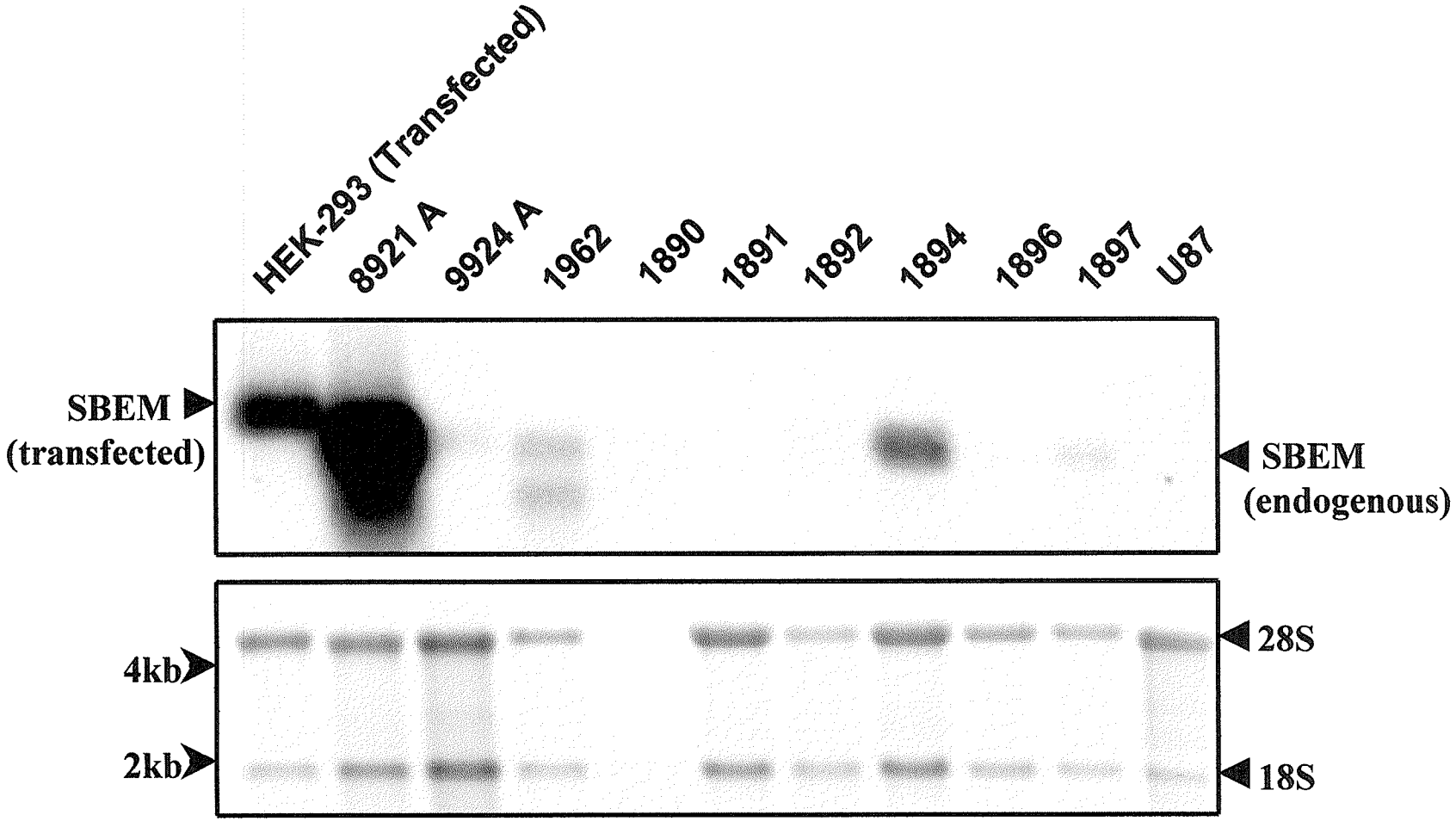

HEK-293 (Transfected)
8921 A
9924 A
1962
1890
1891
1892
1894
1896
1897
U87
SBEM
(transfected)
SBEM
(endogenous)
4kb
2kb
28S
18S

Figure 12 - Examination of SBEM protein expression in normal breast tissues by Western blot analysis.

50μg of total protein extracts taken from 4 frozen normal breast tissues (NB) and one commercially available normal breast protein Medley (BD Biosciences) were analyzed by Western blot. Blots were examined by polyclonal antibody (**A**), H39C51 monoclonal SBEM antibody (**B**) and H51C42 monoclonal SBEM antibody (**C**). Specificity of detected bands was examined by peptide neutralization of the primary antibody with its corresponding peptide (+ Peptide). HEK-293 transfected cell lines were used as the positive control. SBEM protein is indicated by thick black arrow heads. Non-specific binding is indicated by thin arrows. Mr: protein marker.

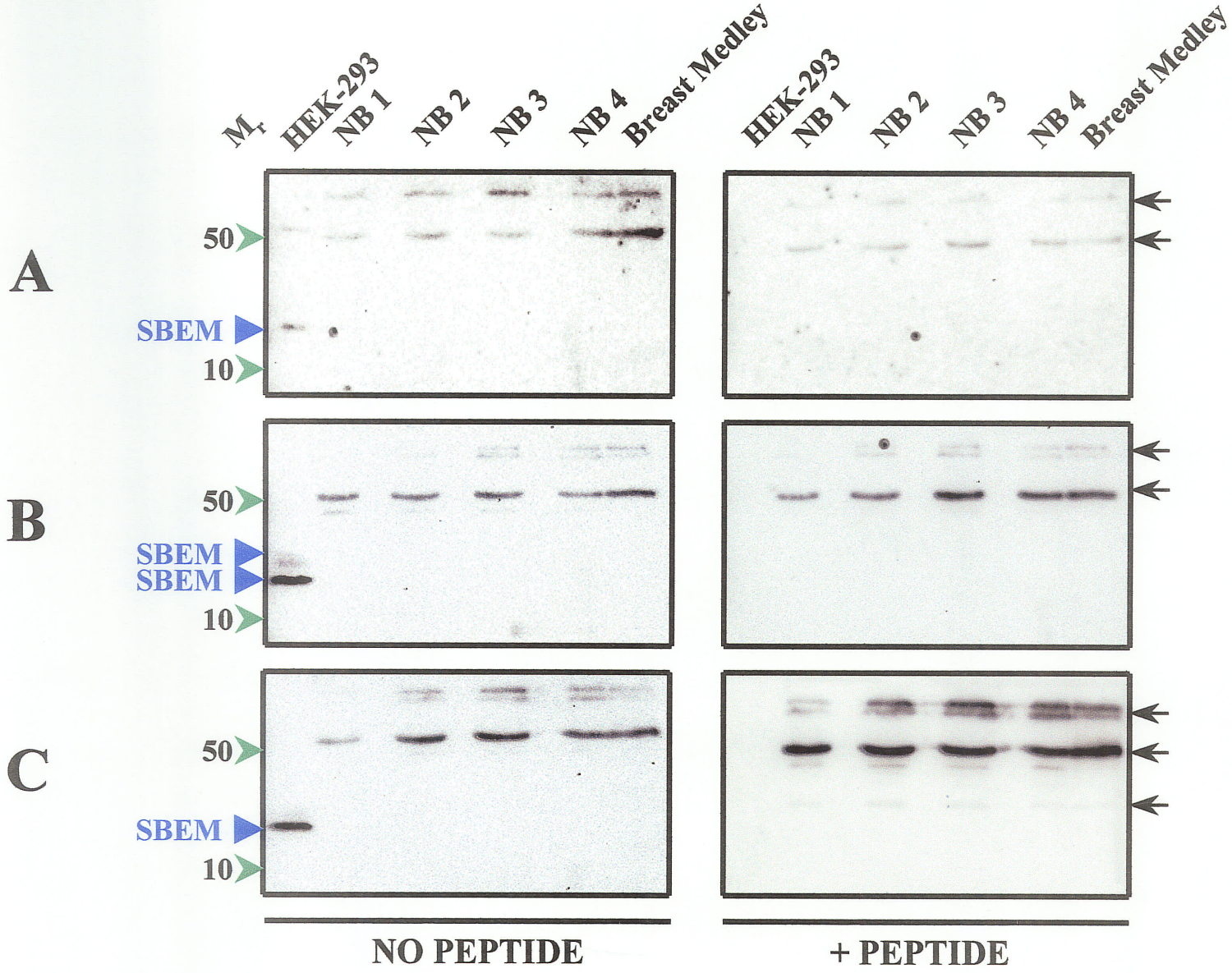

M$_r$
HEK-293
NB 1
NB 2
NB 3
NB 4
Breast Medley
HEK-293
NB 1
NB 2
NB 3
NB 4
Breast Medley
A
B
C
50
SBEM
10
50
SBEM
SBEM
10
50
SBEM
10
NO PEPTIDE
+ PEPTIDE

Figure 13 - Immunohistochemical analysis of normal human mammary tissue.

4 different normal human breast tissues, obtained from mammoplasty reduction surgery, were paraffin-embedded and were examined by immunohistochemical analysis. Samples were analyzed using monoclonal anti-SBEM antibody, H218C31, at a dilution of 1:2400. Specimen sample numbers are given at the bottom right of each slide. Brown DAB-staining indicates positivity for SBEM protein (red arrows), while blue staining indicates the absence of detectable protein (black arrows).

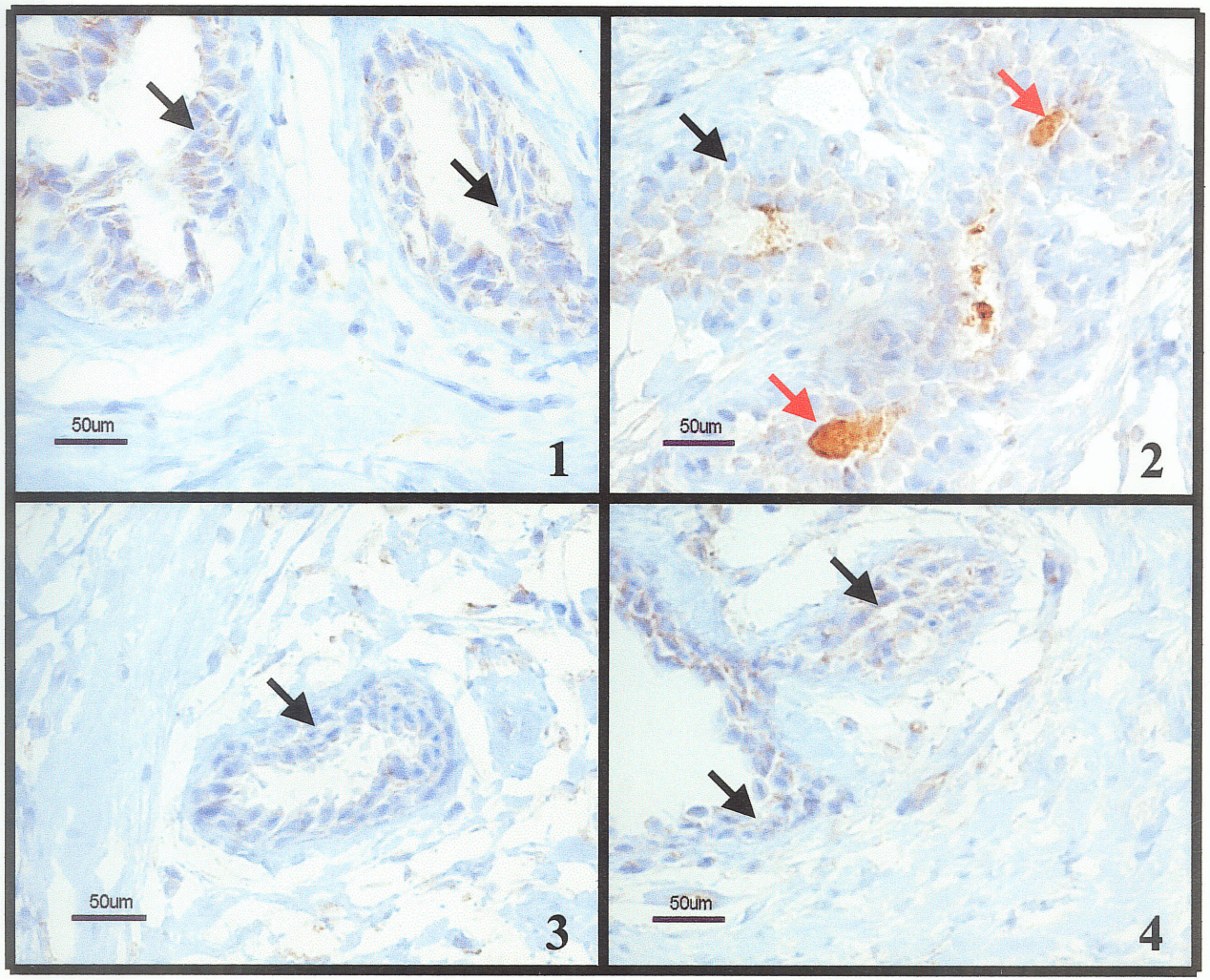

50um
1
50um
2
50um
3
50um
4

Figure 14 - Western blot analysis of breast cancer cell lines using a polyclonal anti-SBEM antibody.

80μg of total protein extracts from various breast cancer cell lines were analysed by Western blot using a polyclonal anti-SBEM antibody alone (**A**), or neutralized with the corresponding peptide (**B**). HEK-293 stably transfected cells were used as positive control (lane 5). Loading was examined by immuno-detection of β-actin (**C**). Mr: protein marker.

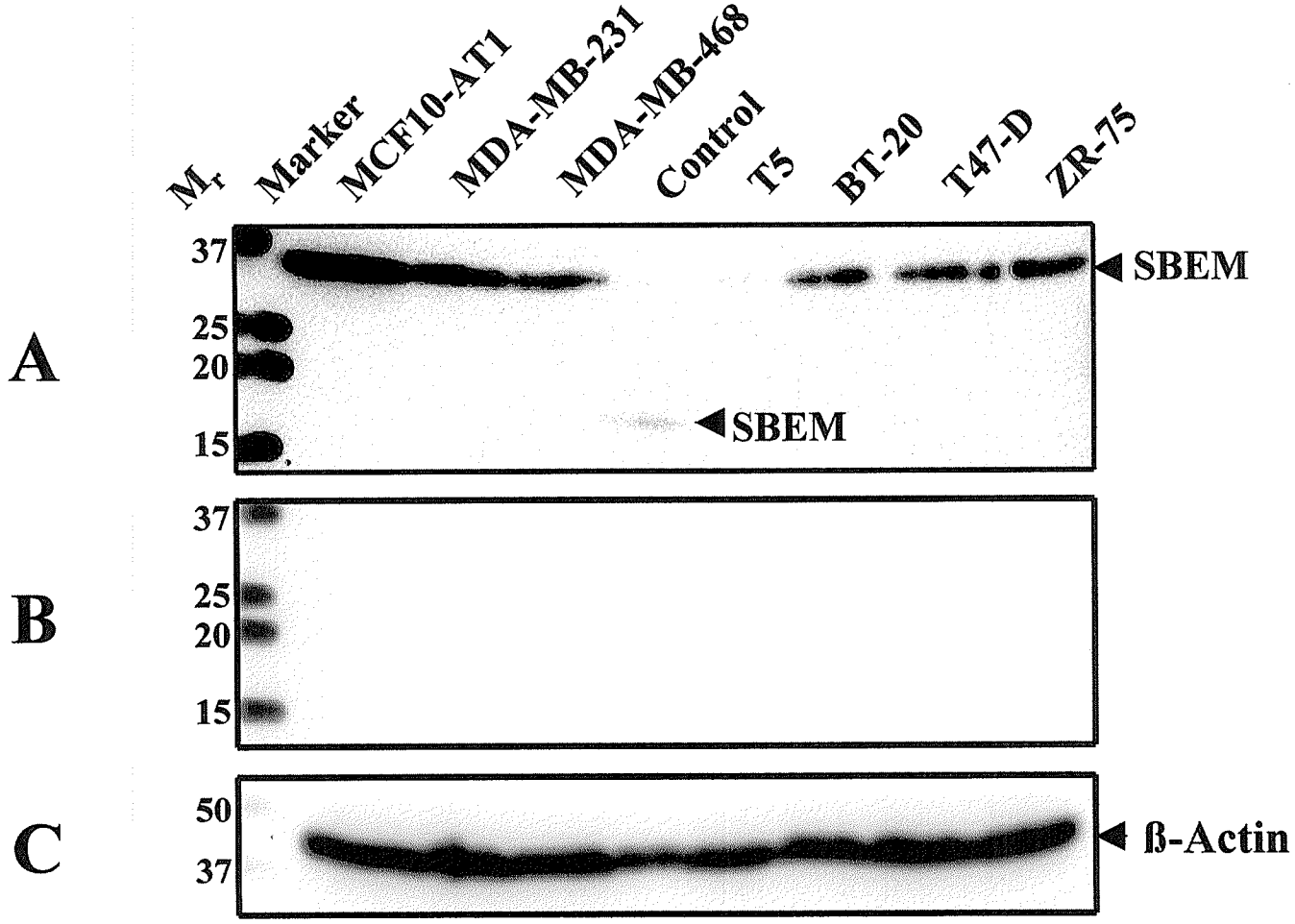

M,
Marker
MCF10-AT1
MDA-MB-231
MDA-MB-468
Control
T5
BT-20
T47-D
ZR-75
A
37
25
20
15
SBEM
SBEM
B
37
25
20
15
C
50
37
ß-Actin

Figure 15 – Western blot analysis of breast cancer cell lines SBEM protein using a panel of monoclonal antibodies

80μg of total protein lysates from various breast cancer cell lines were examined by Western blotting with 7 different monoclonal antibodies (A: H24C16, B: H39C51, C: H218C31, D: H51C42, E: H76C43, F: H24C29, G: H34C78). Stably transfected HEK-293 and MCF-7 cell lines were used as positive controls. No endogenous SBEM protein was detected in any cell lines examined, including negative control cell line; HeLa. In HEK-293 transfected cells, a 17kDa transfected protein was detected with monoclonals A-E, while 17 and 22kDa transfected proteins were observed with monoclonals B, D, E. In MCF-7 transfected cells, a 22kDa band was only observed with monoclonals A and B. All monoclonal antibodies were diluted 1:500. Mr: Molecular mass marker.

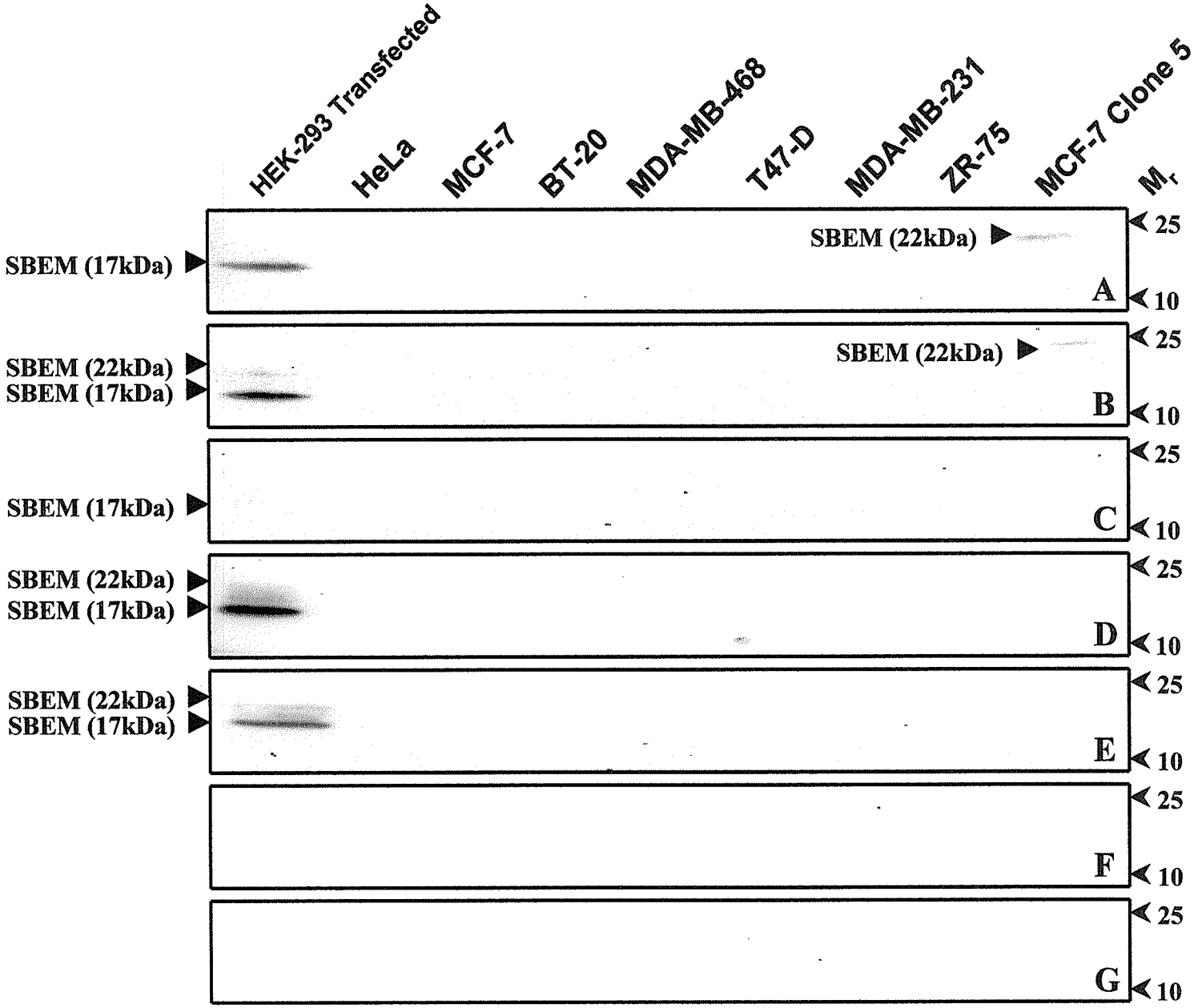

HEK-293 Transfected
HeLa
MCF-7
BT-20
MDA-MB-468
T47-D
MDA-MB-231
ZR-75
MCF-7 Clone 5
M,
SBEM (17kDa)
SBEM (22kDa)
A
25
10
SBEM (22kDa)
SBEM (22kDa)
SBEM (17kDa)
B
25
10
SBEM (17kDa)
C
25
10
SBEM (22kDa)
SBEM (17kDa)
D
25
10
SBEM (22kDa)
SBEM (17kDa)
E
25
10
F
25
10
G
25
10

Figure 16 - Purification of SBEM protein from MCF-7 and HEK-293 stably transfected cell lines

Total lysates of stably transfected cell lines MCF-7 (A) and HEK-293 (B) were purified by nickel column purification. His-tagged protein bound to nickel resin was eluted from columns by increasing concentrations of imidazole (50,100,150,200,250 and 300mM). Eluted fractions were analyzed by Western blot using the C-terminal monoclonal antibody, H39C51 (1:500). SBEM protein was detected at 17 and 22kDa in both transfected cell lines.

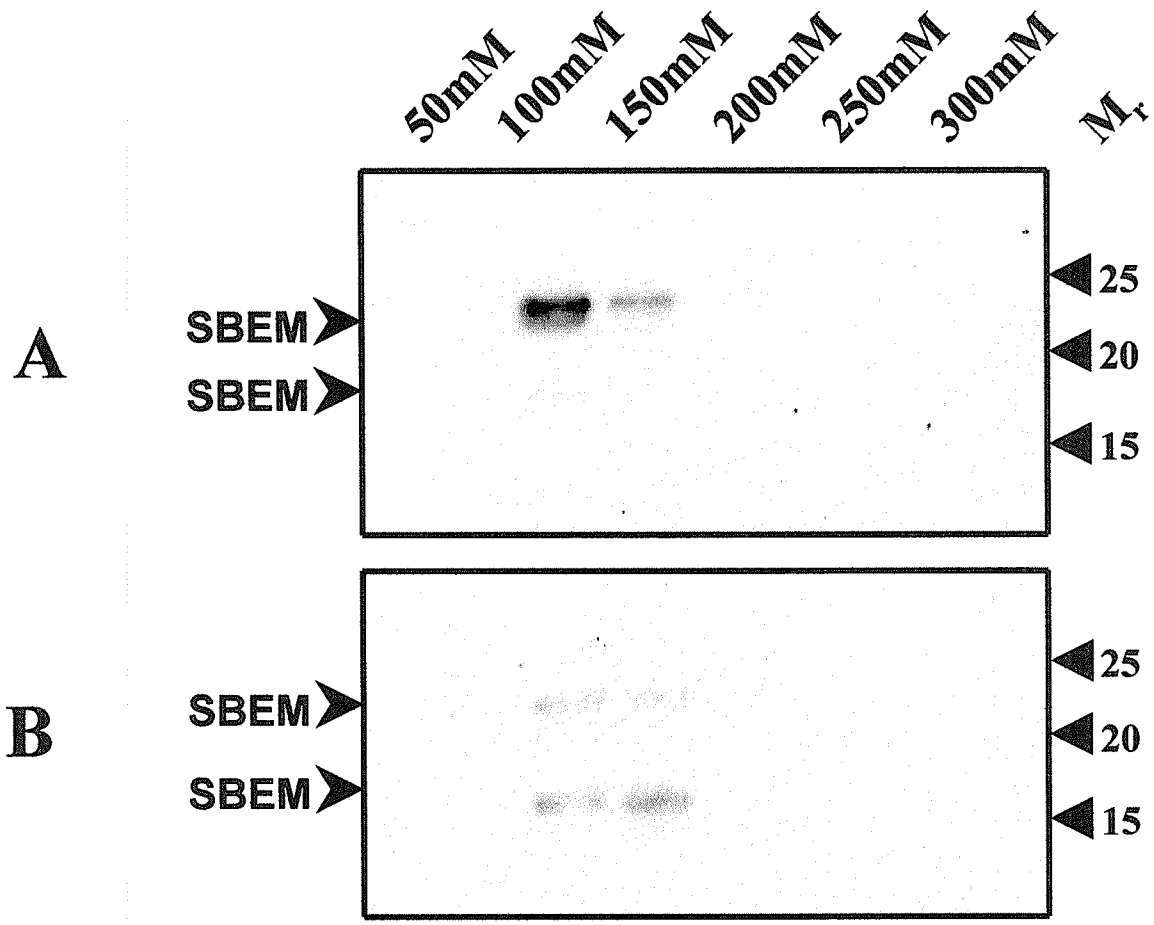

50mM
100mM
150mM
200mM
250mM
300mM
M_r
A
SBEM
SBEM
25
20
15
B
SBEM
SBEM
25
20
15

Figure 17 – Examination of SBEM protein expression in breast tumors by Western blot analysis. 50μg of total protein extracts from various breast tumors were analyzed by Western blot using the H39C51 monoclonal SBEM antibody alone (A), or neutralized with the corresponding peptide (B). C: HEK-293 stably transfected cell line positive control. T: Tumor sample. M_r: Molecular mass marker.

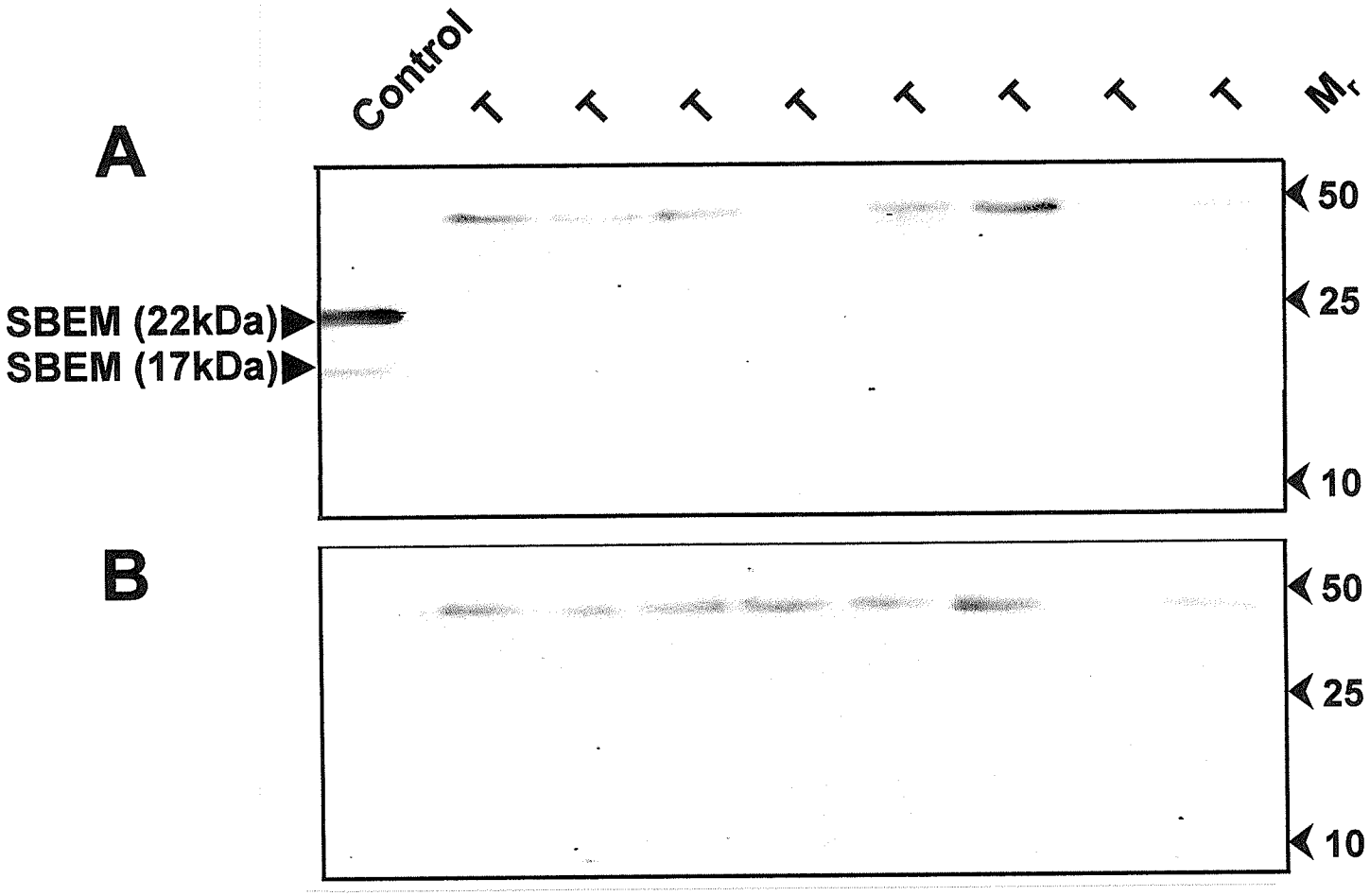

A
Control
T
T
T
T
T
T
T
T
M_r
SBEM (22kDa)
SBEM (17kDa)
50
25
10
B
50
25
10

Figure 18 - Immunohistochemical analysis of a cohort of invasive breast carcinomas.

A subset of 10 paraffin-imbedded invasive ductal carcinomas were examined by immunohistochemical analysis using H218C31 monoclonal SBEM antibody (1:2400 final dilution). Patient identification is indicated on the top left of each picture. Brown DAB-staining indicates positivity for SBEM protein (red arrows), while blue staining indicates the absence of detectable protein (black arrows). Pos: Positive control tissue, positive for SBEM expression by previous RT-PCR. Neg: Negative control tissue, negative for SBEM expression by previous RT-PCR.

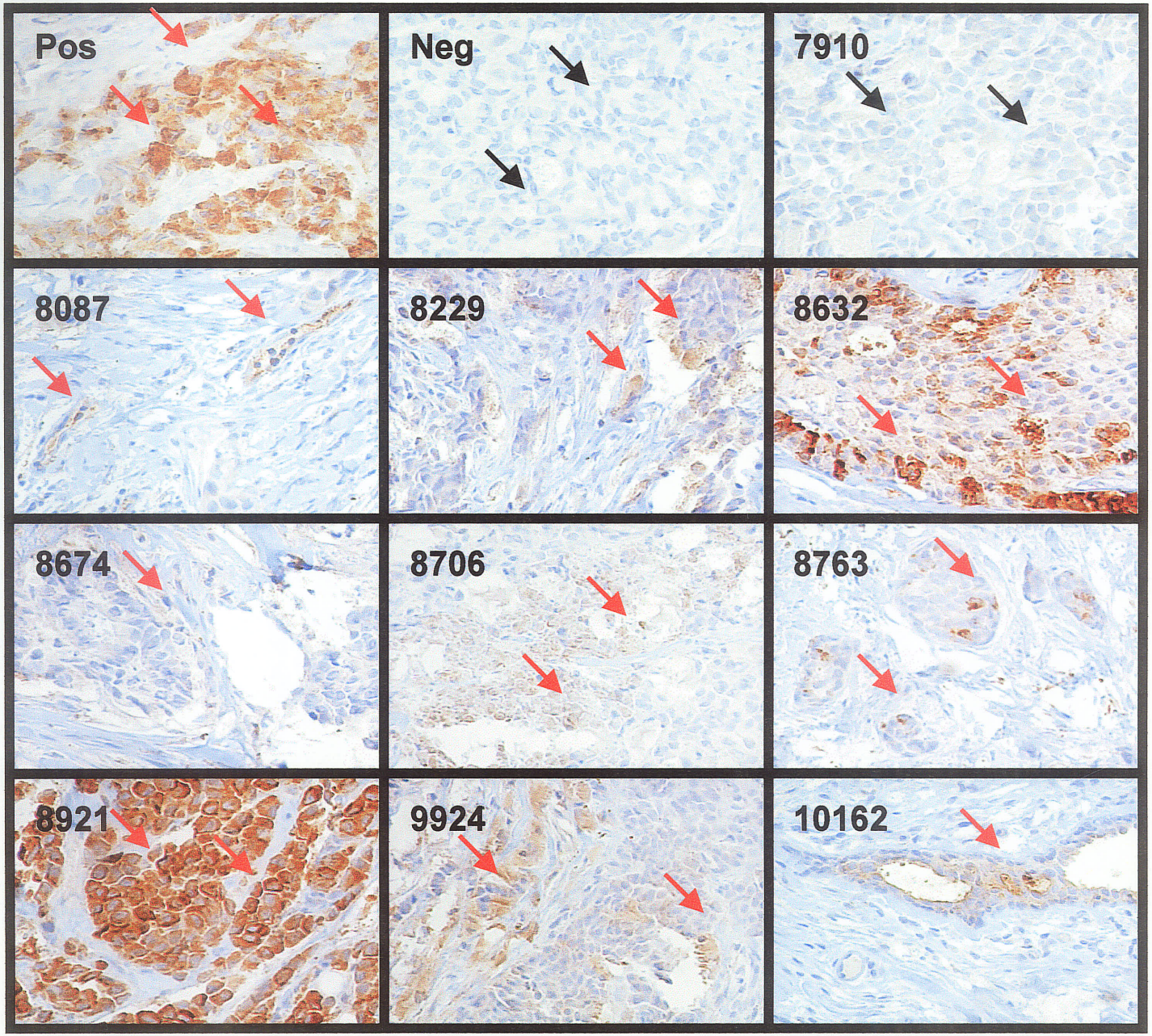

Pos
Neg
7910
8087
8229
8632
8674
8706
8763
8921
9924
10162

Figure 19 – SBEM protein analysis in cell lines by immunoprecipitation

1 x 10^6 cells were extracted in an immunoprecipitation lysis buffer (see Materials and Methods, p.54). Cell lysates were incubated with 1.5μg of anti-SBEM monoclonal antibody, H39C51. Antigen-antibody (Ag-Ab) complexes were immunoprecipitated using Protein G Sepharose (Roche). Immunoprecipitates were electrophoresed on 15% SDS-PAGE gels and were analyzed by Western blot. (**A**) A 17kDa band and a non-distinct, approximately 22kDa band was detected in HEK-293 stably-transfected cells. 50 and 25kDa bands were detected, corresponding to heavy and light chains of antibodies used to form Ag-Ab complexes. 80μg of HEK-293 transfected and BT-20 lysates were used as controls. (**B**) A 17kDa and streaked 22kDa band were detected in HEK-293 transfected cells. Untransfected HEK-293 cells were used as a control for HEK-293 transfected cell SBEM protein expression. No SBEM protein was detected in these cells. Broad bands were identified at 50 and 25kDa, corresponding to heavy and light-chain IgGs, respectively. Mr: Molecular mass marker.

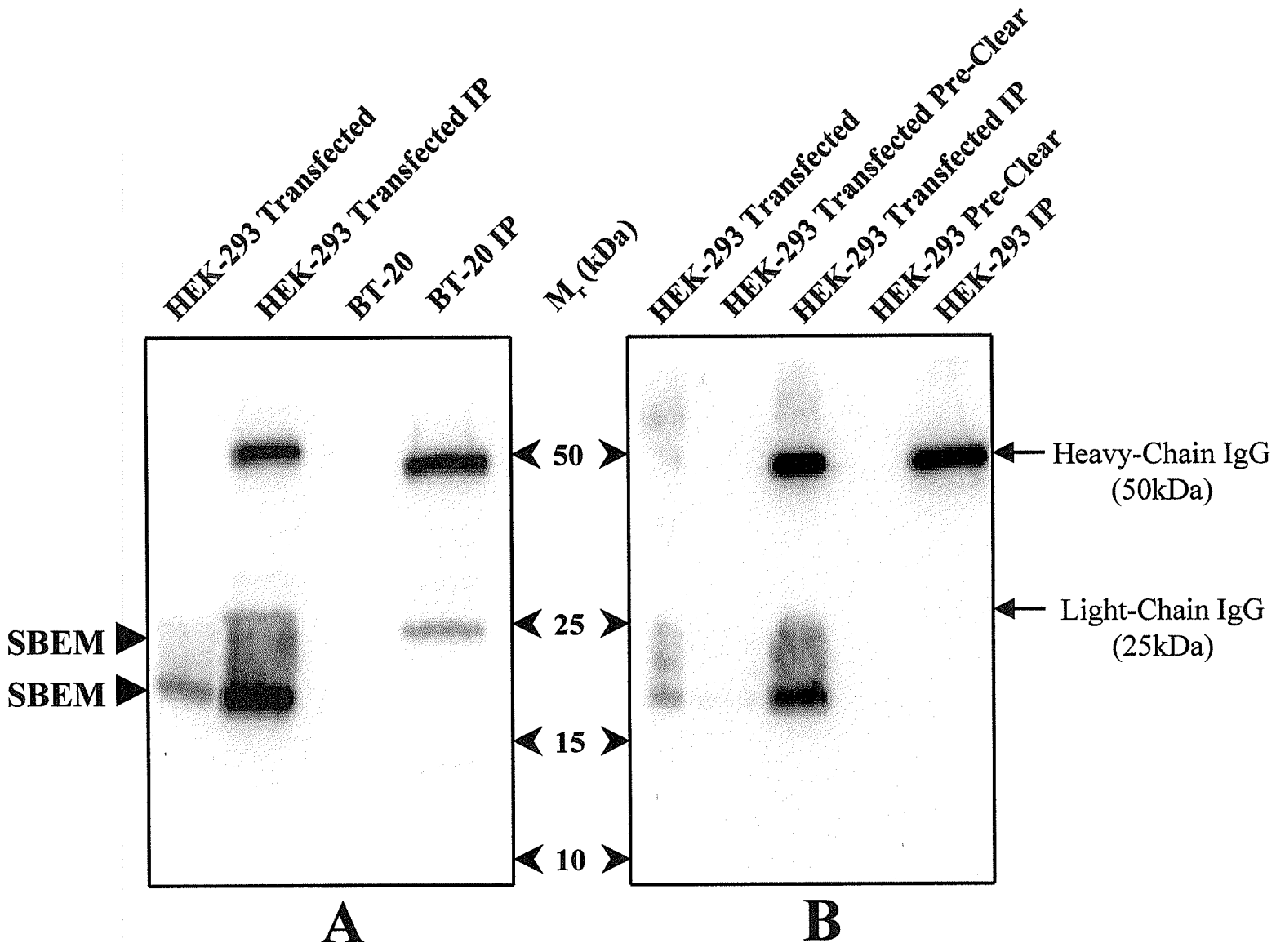

HEK-293 Transfected
HEK-293 Transfected IP
BT-20
BT-20 IP
M, (kDa)
HEK-293 Transfected
HEK-293 Transfected Pre-Clear
HEK-293 Transfected IP
HEK-293 Pre-Clear
HEK-293 IP
50
25
15
10
SBEM
SBEM
Heavy-Chain IgG
(50kDa)
Light-Chain IgG
(25kDa)
A
B

Figure 20 - Cell fractionation of stably transfected HEK-293 cell lines

Following HEK-293 stably transfected cell line cell fractionation into cytosolic, membranous, nuclear and cytoskeletal fractions (BioVision), SBEM protein expression was examined by Western blot analysis. Membranes were probed with 1:1500 polyclonal antibody (**A**) and 1:500 H39C51 monoclonal antibody (**B**). Characteristic 17kDa and 22kDa (weaker) bands were detected in the membranous and cytosolic fractions (arrows). Unfractionated HEK-293 cells were used as a positive control. Equality of loading was determined by staining gels with Coomassie Brilliant blue stain (**C**). Mr: Protein size marker.

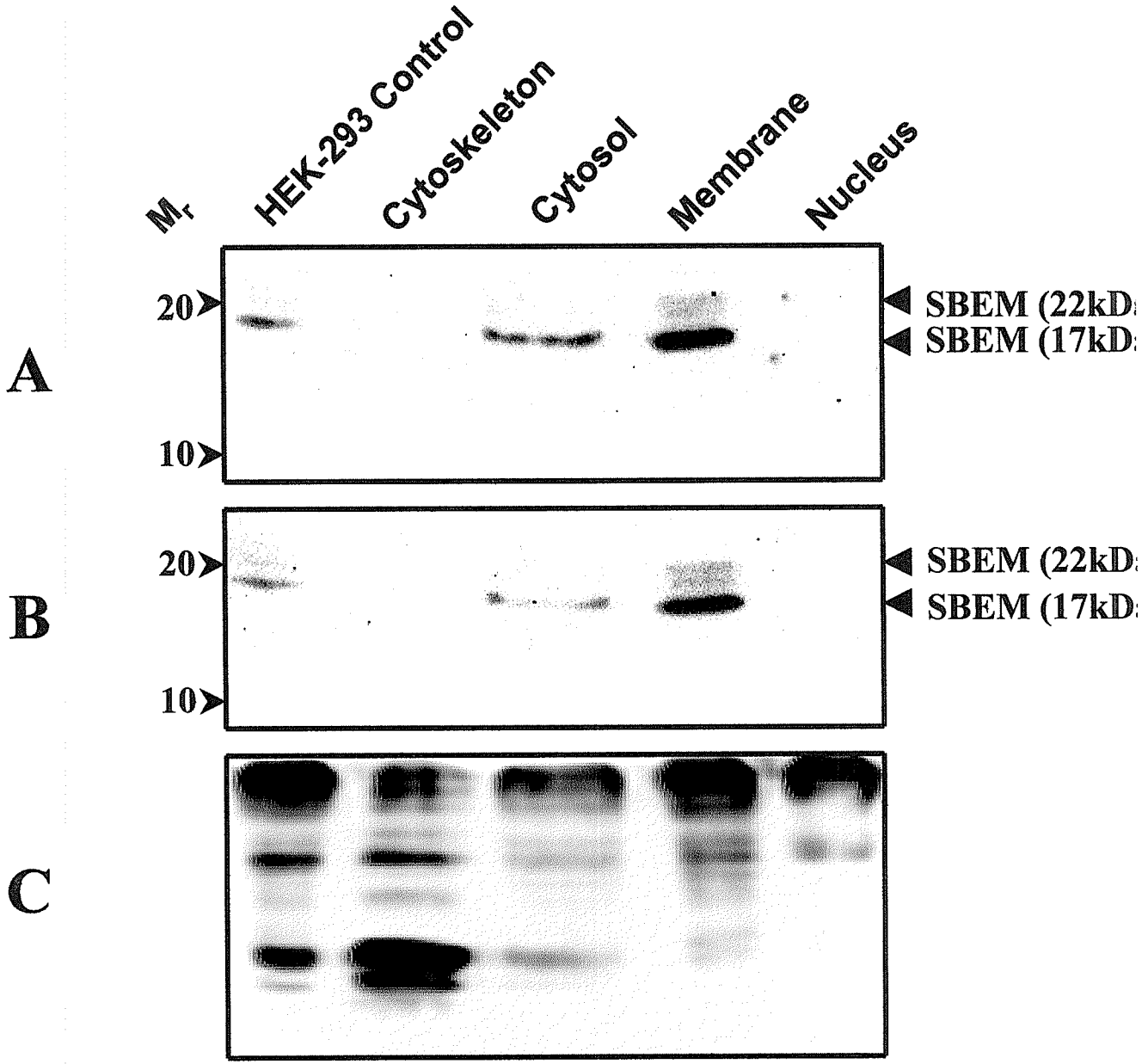

M,
HEK-293 Control
Cytoskeleton
Cytosol
Membrane
Nucleus
A
20
10
SBEM (22kD
SBEM (17kD
B
20
10
SBEM (22kD
SBEM (17kD
C

Figure 21 – SBEM protein is a secreted protein.

Aliquots of stably transfected HEK-293 (HEK-293 (T)) and MCF-7 (MCF-7 (T)) cell line culture media were analyzed by Western blot. 80µl total protein lysates of each transfected cell line were included as a control for corresponding media fractions. (**A**) Using a monoclonal anti-SBEM H39C51 antibody, SBEM protein was detected at 26kDa in MCF-7 transfected media (red arrow), while a smeared band, ranging in size from 25-33kDa, was detected in HEK-293 transfected cells (red arrow). (**B**) To determine specificity, an anti-V5 antibody was used, as the SBEM construct used to generate MCF-7 (T) cells contains a V5 tag. HEK-293 transfected cells do not contain a V5 tag, and thus, were negative. Black arrows: SBEM protein detected in HEK-293 (T) [17kDa] and MCF-7 (T) [22kDa] cell lysates.

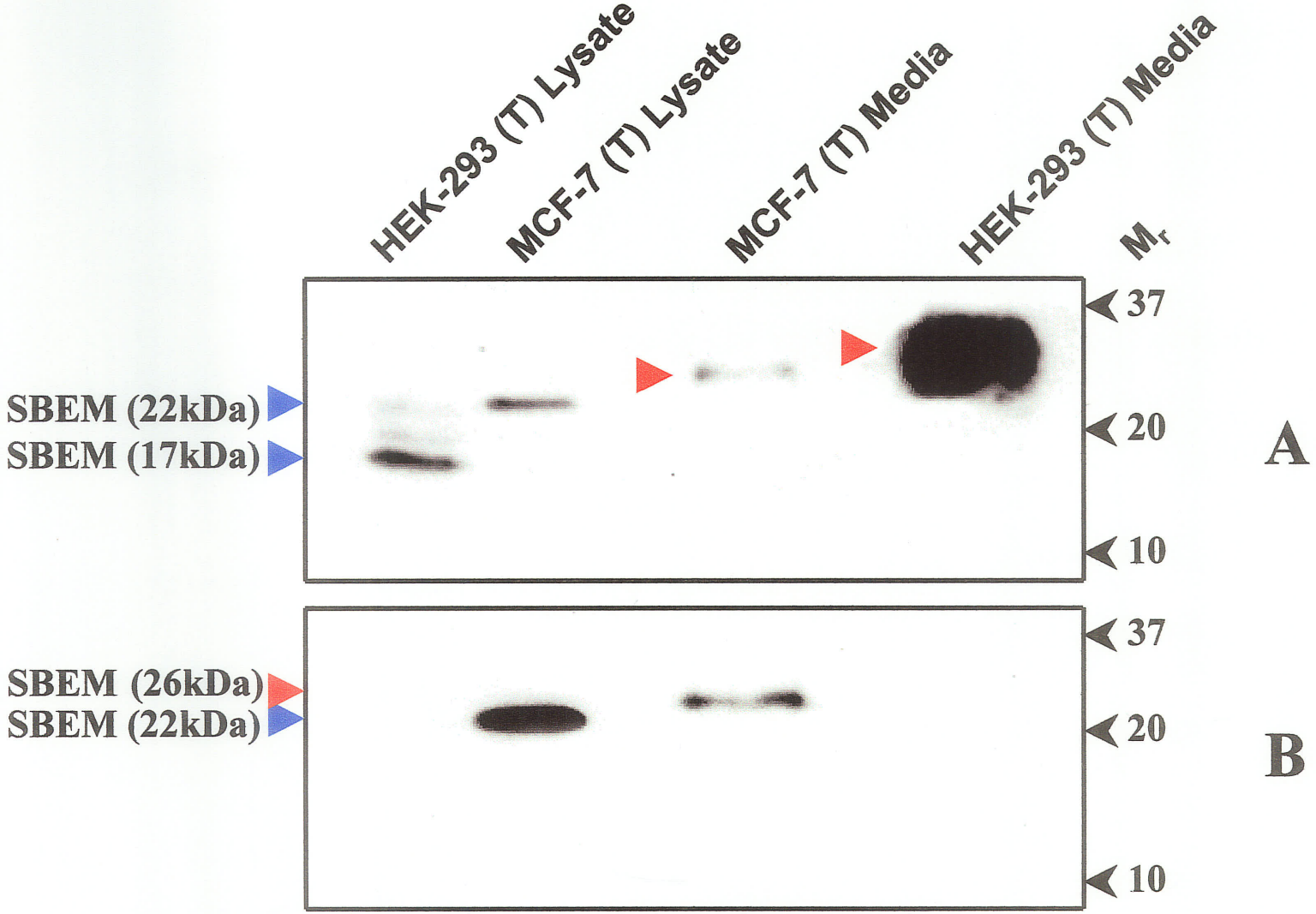

HEK-293 (T) Lysate
MCF-7 (T) Lysate
MCF-7 (T) Media
HEK-293 (T) Media
Mr
37
20
10
SBEM (22kDa)
SBEM (17kDa)
A
37
20
10
SBEM (26kDa)
SBEM (22kDa)
B

Figure 22 - *In vitro* translation of constructs used to generate MCF-7 stably transfected cells.

To confirm the expression of two protein products in stably transfected MCF-7 breast cancer cells, vector (pcDNA3.1/V5-His-TOPO, Invitrogen) constructs used to generate stably transfected cell lines, were examined by *in vitro* translation. A 22kDa band was observed in the SBEM 289 construct (in frame – V5-His tag present), while a size of 17kDa was identified in the SBEM 316 construct (out of frame – no V5-His tag present). Empty vector was the negative control for SBEM expression.

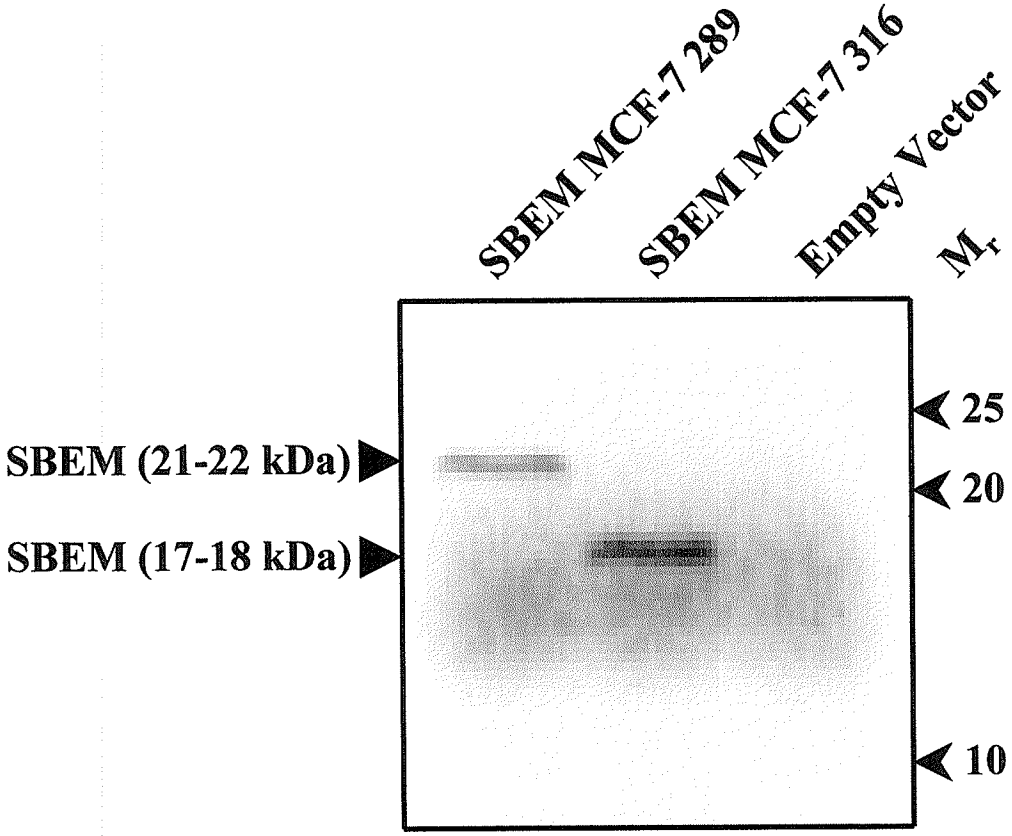

SBEM MCF-7 289
SBEM MCF-7 316
Empty Vector
M_r
SBEM (21-22 kDa)
SBEM (17-18 kDa)
25
20
10

Figure 23 - Optimization of deglycosylation conditions. Increasing amounts of the fetuin protein (0.1μg - 10μg), a positive control provided with the deglycosylation kit (Glyko), were loaded and electrophoresed on a 15% SDS-PAGE gel. (**A**) The gel was stained with Coomassie Brilliant Blue staining and was visualized with a CCD camera and the MCID M4 software (Imaging Research Inc.). The results show the presence of a strong band at approximately 72kDa band. (**B**) Fetuin was resuspended in three different types of lysis buffers (2X SDS buffer, tissue extraction buffer, sodium phosphate pH 8.0 buffer). Fetuin was either treated with ddH_20 (- Enzymes) or deglycosylation enzymes (+ Enzymes). Samples were electrophoresed on SDS-PAGE gels and were examined by Coomassie Brilliant Blue stain.

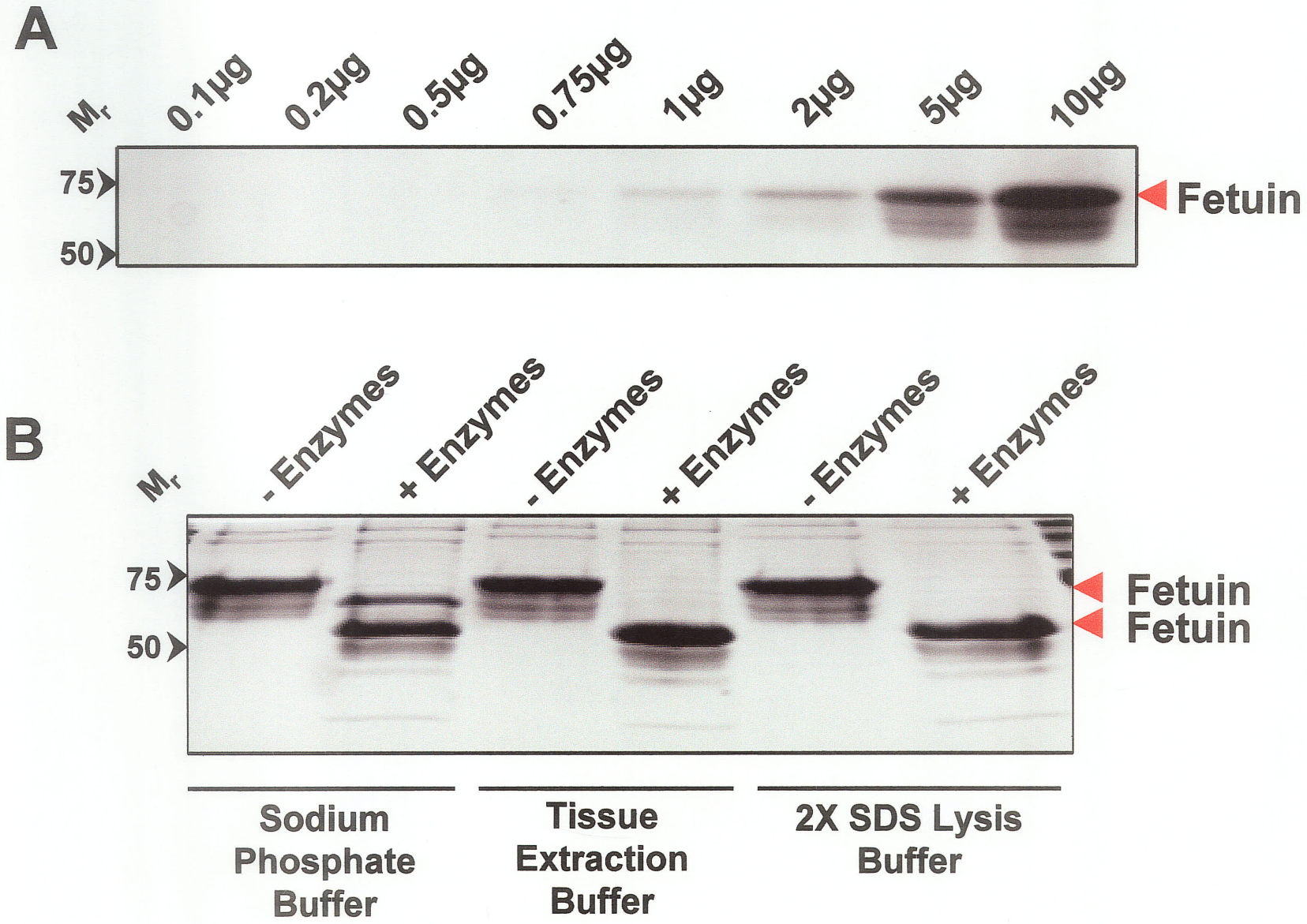

A
Mr
0.1µg
0.2µg
0.5µg
0.75µg
1µg
2µg
5µg
10µg
75
50
Fetuin
B
Mr
- Enzymes
+ Enzymes
- Enzymes
+ Enzymes
- Enzymes
+ Enzymes
75
50
Fetuin
Fetuin
Sodium Phosphate Buffer
Tissue Extraction Buffer
2X SDS Lysis Buffer

Figure 24 – Deglycosylation of SBEM protein in breast cancer cell lines lysate. To examine the glycosylation status of SBEM in breast cancer cell lines, enzymatic deglycosylation was performed on stably transfected HEK-293 (HEK-293 (T)) and MCF-7 (MCF-7 (T)) cells and BT-20 cell lysates. 80μg of total cell lysates were either treated with ddH_20 (- Enzymes) or deglycosylation enzymes (+ Enzymes) and were analyzed by Western blot using an anti-SBEM monoclonal antibody, H39C51. Bands were detected at 17 and 22kDa for HEK-293 (T) and MCF-7 (T) lysates (blue arrows) No shift in size of the SBEM protein in either transfected cell line was observed. BT-20 cells were negative for SBEM protein expression.

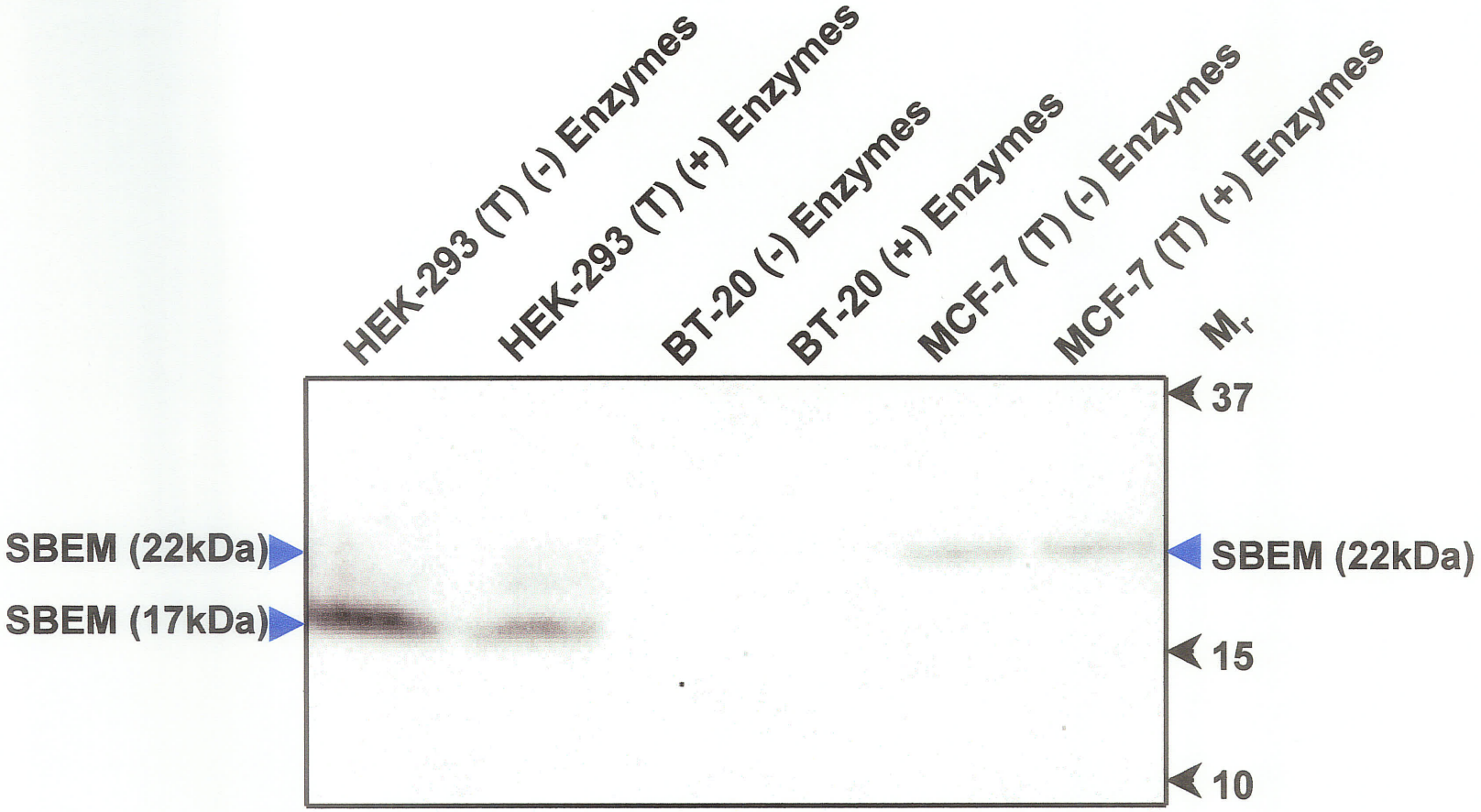

HEK-293 (T) (-) Enzymes
HEK-293 (T) (+) Enzymes
BT-20 (-) Enzymes
BT-20 (+) Enzymes
MCF-7 (T) (-) Enzymes
MCF-7 (T) (+) Enzymes
M_r
37
SBEM (22kDa)
SBEM (17kDa)
SBEM (22kDa)
15
10

Figure 25 - Deglycosylation of the secreted SBEM protein. **(A)** HEK-293 transfected cell culture media was treated with either, ddH_20 (- Enzymes) or deglycosylation enzymes (+ Enzymes) and was analyzed by Western blot, using H39C51 monoclonal SBEM antibody. A broad diffused 25-33kDa SBEM band was detected by Western blot analysis prior to deglycosylation (red arrow). Following enzyme treatment, slightly smaller 23-30kDa diffused band was identified (red arrow). A down-shift of approximately 3kDa seen after enzymatic deglycosylation. Lane 1: Control: HEK-293 transfected cell lysate, Lane 3: 1X deglycosylation enzymes, Lane 5: 2X deglycosylation enzymes. 17 and 22kDa SBEM protein bands were detected in HEK-293 transfected cell lysate (blue arrows). (**B**) Culture media collected from MCF-7 transfected cell lines, was treated with either ddH20 (- Enzymes) or deglycosylational enzymes (+ Enzymes). Western blot analysis reveals a distinct 26kDa band before and after deglycosylation (red arrow). Control: MCF-7 transfected cell total lysate. SBEM detected at 22kDa in total cell lysate (black arrow).

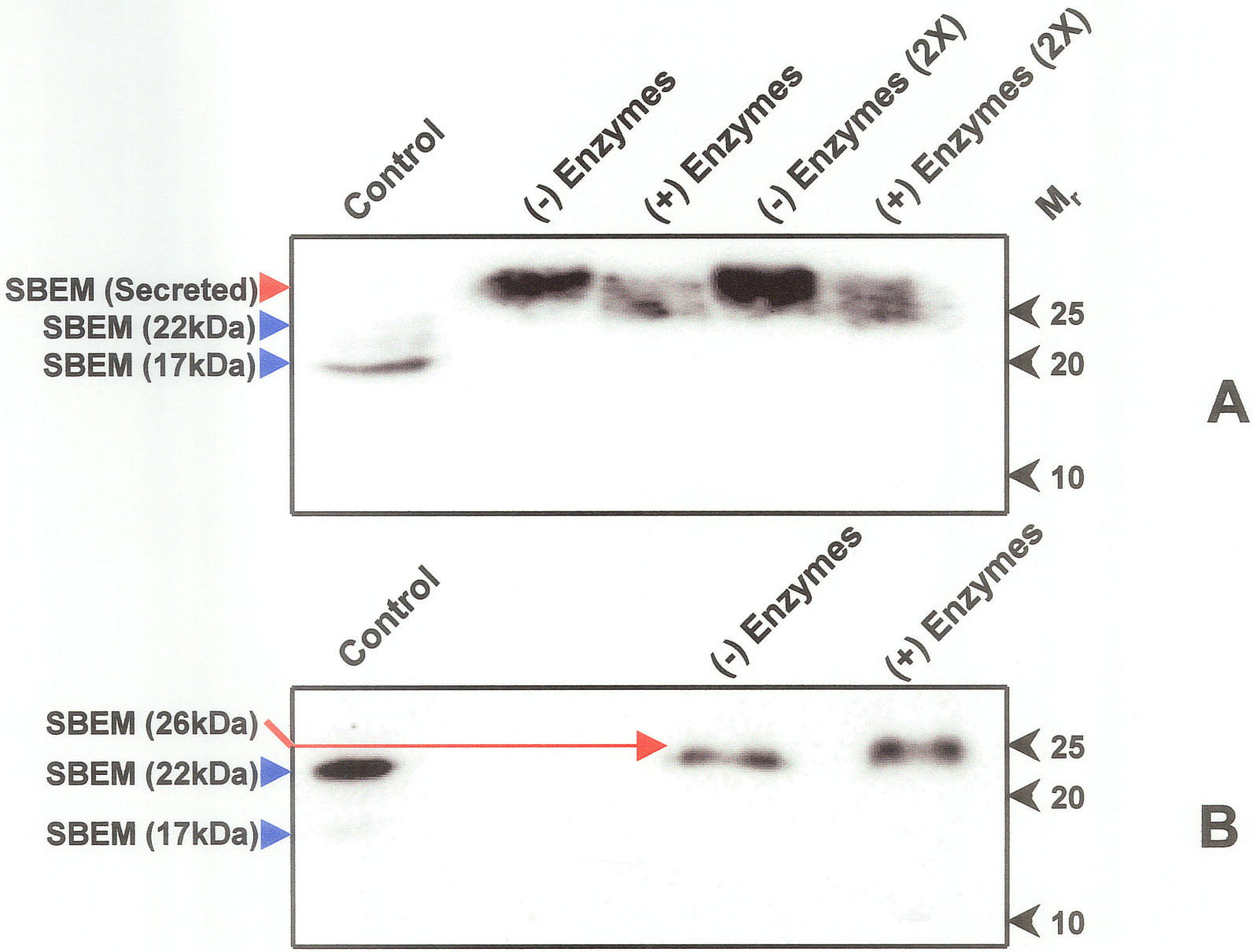

Control
(-) Enzymes
(+) Enzymes
(-) Enzymes (2X)
(+) Enzymes (2X)
Mr
SBEM (Secreted)
SBEM (22kDa)
SBEM (17kDa)
25
20
10
A
Control
(-) Enzymes
(+) Enzymes
SBEM (26kDa)
SBEM (22kDa)
SBEM (17kDa)
25
20
10
B

Table 1 - The primer sets for PCR amplification of cDNA derived from breast cancer cell lines, normal breast tissues and tumors.

Primer Name	Primer Sequence
SBEM upper - YU1	5` - ATGAAGTTCTTAGCAGTCCTG – 3`
SBEM lower – YL1	5` - TTCTACCACTGCTCGTAAAGA – 3`
GAPDH upper - GAP-U	5` - ACCCACTCCTCCACCTTTG – 3`
GAPDH lower - GAP-L	5` - CTCTTGTGCTCTTGCTGGG – 3`

Table 2 - Antibodies used for Western blot and immunohistochemical analysis of the SBEM polypeptide.

Target Protein	Primary Antibody (1°Ab)	Dilution of 1°Ab Used	Secondary Antibody (2°Ab)	Dilution of 2°Ab Used
SBEM C-terminus (aa 70-87)	Polyclonal Rabbit anti-human	1:1500	Goat anti-rabbit	1:5000
SBEM C-terminus (aa c69-90)	H39C51 Monoclonal Mouse anti-human	1:500	Goat anti-mouse	1:5000
SBEM C-terminus (aa c69-90)	H24C16 Monoclonal Mouse anti-human	1:500	Goat anti-mouse	1:5000
SBEM C-terminus (aa 69-89)	H218C31 Monoclonal Mouse anti-human	1:500	Goat anti-mouse	1:5000
SBEM Internal Peptide (aa 20-57, 25-45)	H51C42 Monoclonal Mouse anti-human	1:500	Goat anti-mouse	1:5000
SBEM Internal Peptide (aa 20-57, 25-45)	H76C43 Monoclonal Mouse anti-human	1:500	Goat anti-mouse	1:5000
SBEM N-terminus (aa 20-57)	H24C29 Monoclonal Mouse anti-human	1:500	Goat anti-mouse	1:5000
SBEM N-terminus (aa 20-57)	H34C78 Monoclonal Mouse anti-human	1:500	Goat anti-mouse	1:5000
Progesterone Receptor	Monclonal Mouse anti-human	1:1000	Goat anti-mouse	1:5000
β-Actin	Monclonal Mouse anti-human	1:10000	Goat anti-mouse	1:5000
V5 tag	Monclonal Mouse anti-human	1:5000	Goat anti-mouse	1:5000
His tag	Monclonal Mouse anti-human	1:500	Goat anti-mouse	1:5000

References

Allegra, J. C., Lippman, M. E., Simon, R., Thompson, E. B., Barlock, A., Green, L., Huff, K. K., Do, H. M., Aitken, S. C., and Warren, R. (1979). Association between steroid hormone receptor status and disease-free interval in breast cancer. Cancer Treat Rep *63*, 1271-1277.

Allred, D. C., Mohsin, S. K., and Fuqua, S. A. (2001). Histological and biological evolution of human premalignant breast disease. Endocr Relat Cancer *8*, 47-61.

Bardou, V. J., Arpino, G., Elledge, R. M., Osborne, C. K., and Clark, G. M. (2003). Progesterone receptor status significantly improves outcome prediction over estrogen receptor status alone for adjuvant endocrine therapy in two large breast cancer databases. J Clin Oncol *21*, 1973-1979.

Bartman, A. E., Buisine, M. P., Aubert, J. P., Niehans, G. A., Toribara, N. W., Kim, Y. S., Kelly, E. J., Crabtree, J. E., and Ho, S. B. (1998). The MUC6 secretory mucin gene is expressed in a wide variety of epithelial tissues. J Pathol *186*, 398-405.

Benz, C. C., Scott, G. K., Sarup, J. C., Johnson, R. M., Tripathy, D., Coronado, E., Shepard, H. M., and Osborne, C. K. (1993). Estrogen-dependent, tamoxifen-resistant tumorigenic growth of MCF-7 cells transfected with HER2/neu. Breast Cancer Res Treat *24*, 85-95.

Bobrow, L. G., Happerfield, L. C., Gregory, W. M., and Millis, R. R. (1995). Ductal carcinoma in situ: assessment of necrosis and nuclear morphology and their association with biological markers. J Pathol *176*, 333-341.

Bonneterre, J., Buzdar, A., Nabholtz, J. M., Robertson, J. F., Thurlimann, B., von Euler, M., Sahmoud, T., Webster, A., and Steinberg, M. (2001). Anastrozole is superior to tamoxifen as first-line therapy in hormone receptor positive advanced breast carcinoma. Cancer *92*, 2247-2258.

Bosari, S., Roncalli, M., Viale, G., Bossi, P., and Coggi, G. (1993). p53 immunoreactivity in inflammatory and neoplastic diseases of the uterine cervix. J Pathol *169*, 425-430.

Brockhausen, I., Yang, J. M., Burchell, J., Whitehouse, C., and Taylor-Papadimitriou, J. (1995). Mechanisms underlying aberrant glycosylation of MUC1 mucin in breast cancer cells. Eur J Biochem *233*, 607-617.

Bundred, N. J. (2001). Prognostic and predictive factors in breast cancer. Cancer Treat Rev *27*, 137-142.

Bundred, N. J., Chan, K., and Anderson, N. G. (2001). Studies of epidermal growth factor receptor inhibition in breast cancer. Endocr Relat Cancer *8*, 183-189.

Burchell, J., and Taylor-Papadimitriou, J. (1993). Effect of modification of carbohydrate side chains on the reactivity of antibodies with core-protein epitopes of the MUC1 gene product. Epithelial Cell Biol *2*, 155-162.

Burchell, J. M., Mungul, A., and Taylor-Papadimitriou, J. (2001). O-linked glycosylation in the mammary gland: changes that occur during malignancy. J Mammary Gland Biol Neoplasia *6*, 355-364.

Buzdar, A. U. (2002). New generation aromatase inhibitors--from the advanced to the adjuvant setting. Breast Cancer Res Treat *75 Suppl 1*, S13-17; discussion S33-15.

Carsol, J. L., Gingras, S., and Simard, J. (2002). Synergistic action of prolactin (PRL) and androgen on PRL-inducible protein gene expression in human breast cancer cells: a unique model for functional cooperation between signal transducer and activator of transcription-5 and androgen receptor. Mol Endocrinol *16*, 1696-1710.

Chalut, C., Remy, M. H., and Masson, J. M. (1999). Disulfide bridges are not involved in penicillin-binding protein 1b dimerization in Escherichia coli. J Bacteriol *181*, 2970-2972.

Cheung, K. L., Graves, C. R., and Robertson, J. F. (2000). Tumour marker measurements in the diagnosis and monitoring of breast cancer. Cancer Treat Rev *26*, 91-102.

Cianfrocca, M., and Goldstein, L. J. (2004). Prognostic and predictive factors in early-stage breast cancer. Oncologist *9*, 606-616.

Clark, G. M., Osborne, C. K., and McGuire, W. L. (1984). Correlations between estrogen receptor, progesterone receptor, and patient characteristics in human breast cancer. J Clin Oncol *2*, 1102-1109.

Clark, H. F., Gurney, A. L., Abaya, E., Baker, K., Baldwin, D., Brush, J., Chen, J., Chow, B., Chui, C., Crowley, C., *et al.* (2003). The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment. Genome Res *13*, 2265-2270.

Clark, J. W., Snell, L., Shiu, R. P., Orr, F. W., Maitre, N., Vary, C. P., Cole, D. J., and Watson, P. H. (1999). The potential role for prolactin-inducible protein (PIP) as a marker of human breast cancer micrometastasis. Br J Cancer *81*, 1002-1008.

Clarke, C. L., and Sutherland, R. L. (1990). Progestin regulation of cellular proliferation. Endocr Rev *11*, 266-301.

Clarke, R. B., Howell, A., and Anderson, E. (1997). Type I insulin-like growth factor receptor gene expression in normal human breast tissue treated with oestrogen and progesterone. Br J Cancer *75*, 251-257.

Colleoni, M., Li, S., Gelber, R. D., Coates, A. S., Castiglione-Gertsch, M., Price, K. N., Lindtner, J., Rudenstam, C. M., Crivellari, D., Collins, J., *et al.* (2005). Timing of CMF chemotherapy in combination with tamoxifen in postmenopausal women with breast cancer: role of endocrine responsiveness of the tumor. Ann Oncol *16*, 716-725.

Colpitts, T. L., Billing, P., Granados, E., Hayden, M., Hodges, S., Roberts, L., Russell, J., Friedman, P., and Stroupe, S. (2002). Identification and immunohistochemical characterization of a mucin-like glycoprotein expressed in early stage breast carcinoma. Tumour Biol *23*, 263-278.

Cooke, T., Reeves, J., Lannigan, A., and Stanton, P. (2001). The value of the human epidermal growth factor receptor-2 (HER2) as a prognostic marker. Eur J Cancer *37 Suppl 1*, 3-10.

Correa, I., Plunkett, T., Coleman, J., Galani, E., Windmill, E., Burchell, J. M., and Taylor-Papdimitriou, J. (2005). Responses of human T cells to peptides flanking the tandem repeat and overlapping the signal sequence of MUC1. Int J Cancer *115*, 760-768.

Cummings, S. R., Eckert, S., Krueger, K. A., Grady, D., Powles, T. J., Cauley, J. A., Norton, L., Nickelsen, T., Bjarnason, N. H., Morrow, M., *et al.* (1999). The effect of raloxifene on risk of breast cancer in postmenopausal women: results from the MORE randomized trial. Multiple Outcomes of Raloxifene Evaluation. Jama *281*, 2189-2197.

Diaz, L. K., Wiley, E. L., and Morrow, M. (2001). Expression of epithelial mucins Muc1, Muc2, and Muc3 in ductal carcinoma in situ of the breast. Breast J *7*, 40-45.

Dixon, J. M. (2003). Hormone replacement therapy and the breast. Surg Oncol *12*, 251-263.

Dixon, S. C., Knopf, K. B., and Figg, W. D. (2001). The control of prostate-specific antigen expression and gene regulation by pharmacological agents. Pharmacol Rev *53*, 73-91.

Dong, Y., Walsh, M. D., Cummings, M. C., Wright, R. G., Khoo, S. K., Parsons, P. G., and McGuckin, M. A. (1997). Expression of MUC1 and MUC2 mucins in epithelial ovarian tumours. J Pathol *183*, 311-317.

Driggers, P. H., and Segars, J. H. (2002). Estrogen action and cytoplasmic signaling pathways. Part II: the role of growth factors and phosphorylation in estrogen signaling. Trends Endocrinol Metab *13*, 422-427.

Duffy, M. J. (1999). CA 15-3 and related mucins as circulating markers in breast cancer. Ann Clin Biochem *36 (Pt 5)*, 579-586.

Duffy, M. J. (2005). Predictive markers in breast and other cancers: a review. Clin Chem *51*, 494-503.

Duggan, D. J., Bittner, M., Chen, Y., Meltzer, P., and Trent, J. M. (1999). Expression profiling using cDNA microarrays. Nat Genet *21*, 10-14.

Dumitrescu, R. G., and Cotarla, I. (2005). Understanding breast cancer risk -- where do we stand in 2005? J Cell Mol Med *9*, 208-221.

Eccles, S. A. (2001). The role of c-erbB-2/HER2/neu in breast cancer progression and metastasis. J Mammary Gland Biol Neoplasia *6*, 393-406.

Eeles, R. A. (2000). Future possibilities in the prevention of breast cancer: intervention strategies in BRCA1 and BRCA2 mutation carriers. Breast Cancer Res *2*, 283-290.

Elledge, R. M., Green, S., Ciocca, D., Pugh, R., Allred, D. C., Clark, G. M., Hill, J., Ravdin, P., O'Sullivan, J., Martino, S., and Osborne, C. K. (1998). HER-2 expression and response to tamoxifen in estrogen receptor-positive breast cancer: a Southwest Oncology Group Study. Clin Cancer Res *4*, 7-12.

Elmore, J. G., Armstrong, K., Lehman, C. D., and Fletcher, S. W. (2005). Screening for breast cancer. Jama *293*, 1245-1256.

Emens, L. A., Reilly, R. T., and Jaffee, E. M. (2005). Breast cancer vaccines: maximizing cancer treatment by tapping into host immunity. Endocr Relat Cancer *12*, 1-17.

Esslimani-Sahla, M., Simony-Lafontaine, J., Kramar, A., Lavaill, R., Mollevi, C., Warner, M., Gustafsson, J. A., and Rochefort, H. (2004). Estrogen receptor beta (ER beta) level but not its ER beta cx variant helps to predict tamoxifen resistance in breast cancer. Clin Cancer Res *10*, 5769-5776.

Esteva, F. J., and Hortobagyi, G. N. (2004). Prognostic molecular markers in early breast cancer. Breast Cancer Res *6*, 109-118.

Fanger, G. R., Houghton, R. L., Retter, M. W., Hendrickson, R. C., Babcook, J., Dillon, D. C., Durham, M. D., Reynolds, L. D., Johnson, J. C., Carter, D., *et al.* (2002). Detection of mammaglobin in the sera of patients with breast cancer. Tumour Biol *23*, 212-221.

Fisher, B. (1996). Commentary on endometrial cancer deaths in tamoxifen-treated breast cancer patients. J Clin Oncol *14*, 1027-1039.

Fontenot, J. D., Tjandra, N., Bu, D., Ho, C., Montelaro, R. C., and Finn, O. J. (1993). Biophysical characterization of one-, two-, and three-tandem repeats of human mucin (muc-1) protein core. Cancer Res *53*, 5386-5394.

Fuqua, S. A. (2001). The role of estrogen receptors in breast cancer metastasis. J Mammary Gland Biol Neoplasia *6*, 407-417.

Gendler, S. J. (2001). MUC1, the renaissance molecule. J Mammary Gland Biol Neoplasia *6*, 339-353.

Goss, P. E., Ingle, J. N., Martino, S., Robert, N. J., Muss, H. B., Piccart, M. J., Castiglione, M., Tu, D., Shepherd, L. E., Pritchard, K. I., *et al.* (2003). A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. N Engl J Med *349*, 1793-1802.

Graham, J. D., Roman, S. D., McGowan, E., Sutherland, R. L., and Clarke, C. L. (1995). Preferential stimulation of human progesterone receptor B expression by estrogen in T-47D human breast cancer cells. J Biol Chem *270*, 30693-30700.

Green, B. B., and Taplin, S. H. (2003). Breast cancer screening controversies. J Am Board Fam Pract *16*, 233-241.

Haagensen, D. E., Jr., Mazoujian, G., Dilley, W. G., Pedersen, C. E., Kister, S. J., and Wells, S. A., Jr. (1979). Breast gross cystic disease fluid analysis. I. Isolation and radioimmunoassay for a major component protein. J Natl Cancer Inst *62*, 239-247.

Hartman, A. R., Fleming, G. F., and Dillon, J. J. (2001). Meta-analysis of adjuvant cyclophosphamide/methotrexate/5-fluorouracil chemotherapy in postmenopausal women with estrogen receptor-positive, node-positive breast cancer. Clin Breast Cancer *2*, 138-143; discussion 144.

Hayes, D. F. (2003). Markers of increased risk for failure of adjuvant therapies. Breast *12*, 543-549.

Hayes, D. F. (2004). Tamoxifen: Dr. Jekyll and Mr. Hyde? J Natl Cancer Inst *96*, 895-897.

Hayes, D. F., Bast, R. C., Desch, C. E., Fritsche, H., Jr., Kemeny, N. E., Jessup, J. M., Locker, G. Y., Macdonald, J. S., Mennel, R. G., Norton, L., *et al.* (1996). Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers. J Natl Cancer Inst *88*, 1456-1466.

Hayes, D. F., Isaacs, C., and Stearns, V. (2001). Prognostic factors in breast cancer: current and new predictors of metastasis. J Mammary Gland Biol Neoplasia *6*, 375-392.

Hayes, D. F., Zurawski, V. R., Jr., and Kufe, D. W. (1986). Comparison of circulating CA15-3 and carcinoembryonic antigen levels in patients with breast cancer. J Clin Oncol *4*, 1542-1550.

Holmes, F. A., Fritsche, H. A., Loewy, J. W., Geitner, A. M., Sutton, R. C., Buzdar, A. U., and Hortobagyi, G. N. (1990). Measurement of estrogen and progesterone receptors in human breast tumors: enzyme immunoassay versus binding assay. J Clin Oncol *8*, 1025-1035.

Hopp, T. A., Weiss, H. L., Hilsenbeck, S. G., Cui, Y., Allred, D. C., Horwitz, K. B., and Fuqua, S. A. (2004). Breast cancer patients with progesterone receptor PR-A-rich tumors have poorer disease-free survival rates. Clin Cancer Res *10*, 2751-2760.

Hortobagyi, G. N. (2001). Overview of treatment results with trastuzumab (Herceptin) in metastatic breast cancer. Semin Oncol *28*, 43-47.

Hortobagyi, G. N. (2001). Progress in systemic chemotherapy of primary breast cancer: an overview. J Natl Cancer Inst Monogr, 72-79.

Houghton, R. L., Dillon, D. C., Molesh, D. A., Zehentner, B. K., Xu, J., Jiang, J., Schmidt, C., Frudakis, A., Repasky, E., Maltez Filho, A., *et al.* (2001). Transcriptional complementarity in breast cancer: application to detection of circulating tumor cells. Mol Diagn *6*, 79-91.

Howell, A., Cuzick, J., Baum, M., Buzdar, A., Dowsett, M., Forbes, J. F., Hoctin-Boes, G., Houghton, J., Locker, G. Y., and Tobias, J. S. (2005). Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. Lancet *365*, 60-62.

Hube, F., Mutawe, M., Leygue, E., and Myal, Y. (2004). Human small breast epithelial mucin: the promise of a new breast tumor biomarker. DNA Cell Biol *23*, 842-849.

Ingle, J. N., Suman, V. J., Jordan, V. C., and Dowsett, M. (1999). Combination hormonal therapy involving aromatase inhibitors in the management of women with breast cancer. Endocr Relat Cancer *6*, 265-269.

Jacobsen, B. M., Richer, J. K., Sartorius, C. A., and Horwitz, K. B. (2003). Expression profiling of human breast cancers and gene regulation by progesterone receptors. J Mammary Gland Biol Neoplasia *8*, 257-268.

Janicke, F., Schmitt, M., Pache, L., Ulm, K., Harbeck, N., Hofler, H., and Graeff, H. (1993). Urokinase (uPA) and its inhibitor PAI-1 are strong and independent prognostic factors in node-negative breast cancer. Breast Cancer Res Treat *24*, 195-208.

Jiang, Y., Harlocker, S. L., Molesh, D. A., Dillon, D. C., Stolk, J. A., Houghton, R. L., Repasky, E. A., Badaro, R., Reed, S. G., and Xu, J. (2002). Discovery of differentially expressed genes in human breast cancer using subtracted cDNA libraries and cDNA microarrays. Oncogene *21*, 2270-2282.

Jones, K. L., and Buzdar, A. U. (2004). A review of adjuvant hormonal therapy in breast cancer. Endocr Relat Cancer *11*, 391-406.

Kato, S., Endoh, H., Masuhiro, Y., Kitamoto, T., Uchiyama, S., Sasaki, H., Masushige, S., Gotoh, Y., Nishida, E., Kawashima, H., *et al.* (1995). Activation of the estrogen

receptor through phosphorylation by mitogen-activated protein kinase. Science *270*, 1491-1494.

Keen, J. C., and Davidson, N. E. (2003). The biology of breast carcinoma. Cancer *97*, 825-833.

Kelsey, J. L. (1993). Breast cancer epidemiology: summary and future directions. Epidemiol Rev *15*, 256-263.

Keshgegian, A. A., and Cnaan, A. (1995). Proliferation markers in breast carcinoma. Mitotic figure count, S-phase fraction, proliferating cell nuclear antigen, Ki-67 and MIB-1. Am J Clin Pathol *104*, 42-49.

King, M. C., Wieand, S., Hale, K., Lee, M., Walsh, T., Owens, K., Tait, J., Ford, L., Dunn, B. K., Costantino, J., *et al.* (2001). Tamoxifen and breast cancer incidence among women with inherited mutations in BRCA1 and BRCA2: National Surgical Adjuvant Breast and Bowel Project (NSABP-P1) Breast Cancer Prevention Trial. Jama *286*, 2251-2256.

Klee, G. G., and Schreiber, W. E. (2004). MUC1 gene-derived glycoprotein assays for monitoring breast cancer (CA 15-3, CA 27.29, BR): are they measuring the same antigen? Arch Pathol Lab Med *128*, 1131-1135.

Knight, W. A., Livingston, R. B., Gregory, E. J., and McGuire, W. L. (1977). Estrogen receptor as an independent prognostic factor for early recurrence in breast cancer. Cancer Res *37*, 4669-4671.

Knowlden, J. M., Hutcheson, I. R., Jones, H. E., Madden, T., Gee, J. M., Harper, M. E., Barrow, D., Wakeling, A. E., and Nicholson, R. I. (2003). Elevated levels of epidermal growth factor receptor/c-erbB2 heterodimers mediate an autocrine growth regulatory pathway in tamoxifen-resistant MCF-7 cells. Endocrinology *144*, 1032-1044.

Koga, T., Horio, Y., Mitsudomi, T., Takahashi, T., and Yatabe, Y. (2004). Identification of MGB1 as a marker in the differential diagnosis of lung tumors in patients with a history of breast cancer by analysis of publicly available SAGE data. J Mol Diagn *6*, 90-95.

Konecny, G., Pauletti, G., Pegram, M., Untch, M., Dandekar, S., Aguilar, Z., Wilson, C., Rong, H. M., Bauerfeind, I., Felber, M., *et al.* (2003). Quantitative association between HER-2/neu and steroid hormone receptors in hormone receptor-positive primary breast cancer. J Natl Cancer Inst *95*, 142-153.

Lacroix, M., and Leclercq, G. (2004). Relevance of breast cancer cell lines as models for breast tumours: an update. Breast Cancer Res Treat *83*, 249-289.

Lai, H., Ma, F., Trapido, E., Meng, L., and Lai, S. (2004). Spectrum of p53 tumor suppressor gene mutations and breast cancer survival. Breast Cancer Res Treat *83*, 57-66.

Leygue, E., Dotzlaw, H., Watson, P. H., and Murphy, L. C. (1999). Expression of the steroid receptor RNA activator in human breast tumors. Cancer Res *59*, 4190-4193.

Leygue, E., Snell, L., Dotzlaw, H., Hole, K., Troup, S., Hiller-Hitchcock, T., Murphy, L. C., and Watson, P. H. (1999). Mammaglobin, a potential marker of breast cancer nodal metastasis. J Pathol *189*, 28-33.

Linton, H. J., Marks, L. S., Millar, L. S., Knott, C. L., Rittenhouse, H. G., and Mikolajczyk, S. D. (2003). Benign prostate-specific antigen (BPSA) in serum is increased in benign prostate disease. Clin Chem *49*, 253-259.

Liu, Y., el-Ashry, D., Chen, D., Ding, I. Y., and Kern, F. G. (1995). MCF-7 breast cancer cells overexpressing transfected c-erbB-2 have an in vitro growth advantage in estrogen-depleted conditions and reduced estrogen-dependence and tamoxifen-sensitivity in vivo. Breast Cancer Res Treat *34*, 97-117.

Long, B. J., Jelovac, D., Handratta, V., Thiantanawat, A., MacPherson, N., Ragaz, J., Goloubeva, O. G., and Brodie, A. M. (2004). Therapeutic strategies using the aromatase inhibitor letrozole and tamoxifen in a breast cancer model. J Natl Cancer Inst *96*, 456-465.

Ma, X. J., Salunga, R., Tuggle, J. T., Gaudet, J., Enright, E., McQuary, P., Payette, T., Pistone, M., Stecker, K., Zhang, B. M., *et al.* (2003). Gene expression profiles of human breast cancer progression. Proc Natl Acad Sci U S A *100*, 5974-5979.

Mackay, A., Jones, C., Dexter, T., Silva, R. L., Bulmer, K., Jones, A., Simpson, P., Harris, R. A., Jat, P. S., Neville, A. M., *et al.* (2003). cDNA microarray analysis of genes associated with ERBB2 (HER2/neu) overexpression in human mammary luminal epithelial cells. Oncogene *22*, 2680-2688.

Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., and Evans, R. M. (1995). The nuclear receptor superfamily: the second decade. Cell *83*, 835-839.

Martini, P. G., and Katzenellenbogen, B. S. (2003). Modulation of estrogen receptor activity by selective coregulators. J Steroid Biochem Mol Biol *85*, 117-122.

Masood, S., and Bui, M. M. (2002). Prognostic and predictive value of HER2/neu oncogene in breast cancer. Microsc Res Tech *59*, 102-108.

Mass, R. D. (2004). The HER receptor family: a rich target for therapeutic development. Int J Radiat Oncol Biol Phys *58*, 932-940.

McGuckin, M. A., Walsh, M. D., Hohn, B. G., Ward, B. G., and Wright, R. G. (1995). Prognostic significance of MUC1 epithelial mucin expression in breast cancer. Hum Pathol *26*, 432-439.

McGuire, W. L. (1991). Breast cancer prognostic factors: evaluation guidelines. J Natl Cancer Inst *83*, 154-155.

Miksicek, R. J., Myal, Y., Watson, P. H., Walker, C., Murphy, L. C., and Leygue, E. (2002). Identification of a novel breast- and salivary gland-specific, mucin-like gene strongly expressed in normal and tumor human mammary epithelium. Cancer Res *62*, 2736-2740.

Milla-Santos, A., Milla, L., Rallo, L., and Solano, V. (2001). Phase III randomized trial of toremifene vs tamoxifen in hormonodependant advanced breast cancer. Breast Cancer Res Treat *65*, 119-124.

Moll, U. M., Ostermeyer, A. G., Haladay, R., Winkfield, B., Frazier, M., and Zambetti, G. (1996). Cytoplasmic sequestration of wild-type p53 protein impairs the G1 checkpoint after DNA damage. Mol Cell Biol *16*, 1126-1137.

Morris, C. (2002). The role of EGFR-directed therapy in the treatment of breast cancer. Breast Cancer Res Treat *75 Suppl 1*, S51-55; discussion S57-59.

Mosselman, S., Polman, J., and Dijkema, R. (1996). ER beta: identification and characterization of a novel human estrogen receptor. FEBS Lett *392*, 49-53.

Mouridsen, H., Gershanovich, M., Sun, Y., Perez-Carrion, R., Boni, C., Monnier, A., Apffelstaedt, J., Smith, R., Sleeboom, H. P., Janicke, F., *et al.* (2001). Superior efficacy of letrozole versus tamoxifen as first-line therapy for postmenopausal women with advanced breast cancer: results of a phase III study of the International Letrozole Breast Cancer Group. J Clin Oncol *19*, 2596-2606.

Munster, P. N., and Norton, L. (2001). Predictive factor for the response to adjuvant therapy with emphasis in breast cancer. Breast Cancer Res *3*, 361-364.

Myal, Y., Iwasiow, B., Cosby, H., Yarmill, A., Blanchard, A., Tsuyuki, D., Fresnoza, A., Duckworth, M. L., and Shiu, R. P. (1998). Analysis of tissue- and hormone-specific regulation of the human prolactin-inducible protein/gross cystic disease fluid protein-15 gene in transgenic mice. J Mol Endocrinol *21*, 217-223.

Nacht, M., Ferguson, A. T., Zhang, W., Petroziello, J. M., Cook, B. P., Gao, Y. H., Maguire, S., Riley, D., Coppola, G., Landes, G. M., *et al.* (1999). Combining serial analysis of gene expression and array technologies to identify genes differentially expressed in breast cancer. Cancer Res *59*, 5464-5470.

Narayanan, K., Jaramillo, A., Benshoff, N. D., Campbell, L. G., Fleming, T. P., Dietz, J. R., and Mohanakumar, T. (2004). Response of established human breast tumors to vaccination with mammaglobin-A cDNA. J Natl Cancer Inst *96*, 1388-1396.

Normanno, N., Bianco, C., De Luca, A., Maiello, M. R., and Salomon, D. S. (2003). Target-based agents against ErbB receptors and their ligands: a novel approach to cancer treatment. Endocr Relat Cancer *10*, 1-21.

Nunez-Villar, M. J., Martinez-Arribas, F., Pollan, M., Lucas, A. R., Sanchez, J., Tejerina, A., and Schneider, J. (2003). Elevated mammaglobin (h-MAM) expression in breast cancer is associated with clinical and biological features defining a less aggressive tumour phenotype. Breast Cancer Res *5*, R65-70.

O'Brien, N., Maguire, T. M., O'Donovan, N., Lynch, N., Hill, A. D., McDermott, E., O'Higgins, N., and Duffy, M. J. (2002). Mammaglobin a: a promising marker for breast cancer. Clin Chem *48*, 1362-1364.

O'Brien, N. A., O'Donovan, N., Ryan, B., Hill, A. D., McDermott, E., O'Higgins, N., and Duffy, M. J. (2005). Mammaglobin a in breast cancer: existence of multiple molecular forms. Int J Cancer *114*, 623-627.

Osborne, C. K. (1998). Steroid hormone receptors in breast cancer management. Breast Cancer Res Treat *51*, 227-238.

Osborne, C. K., Yochmowitz, M. G., Knight, W. A., 3rd, and McGuire, W. L. (1980). The value of estrogen and progesterone receptors in the treatment of breast cancer. Cancer *46*, 2884-2888.

Perey, L., Hayes, D. F., and Kufe, D. (1992). Effects of differentiating agents on cell surface expression of the breast carcinoma-associated DF3-P epitope. Cancer Res *52*, 6365-6370.

Pietras, R. J., Arboleda, J., Reese, D. M., Wongvipat, N., Pegram, M. D., Ramos, L., Gorman, C. M., Parker, M. G., Sliwkowski, M. X., and Slamon, D. J. (1995). HER-2 tyrosine kinase pathway targets estrogen receptor and promotes hormone-independent growth in human breast cancer cells. Oncogene *10*, 2435-2446.

Pons-Anicet, D. M., Krebs, B. P., Mira, R., and Namer, M. (1987). Value of CA 15:3 in the follow-up of breast cancer patients. Br J Cancer *55*, 567-569.

Porter, D., Weremowicz, S., Chin, K., Seth, P., Keshaviah, A., Lahti-Domenici, J., Bae, Y. K., Monitto, C. L., Merlos-Suarez, A., Chan, J., *et al.* (2003). A neural survival factor is a candidate oncogene in breast cancer. Proc Natl Acad Sci U S A *100*, 10931-10936.

Punj, V., Saint-Dic, D., Daghfal, S., and Kanwar, J. R. (2004). Microbial-based therapy of cancer: a new twist to age old practice. Cancer Biol Ther *3*, 708-714.

Quan, C., and Lu, S. J. (2003). Identification of genes preferentially expressed in mammary epithelial cells of Copenhagen rat using subtractive hybridization and microarrays. Carcinogenesis *24*, 1593-1599.

Quenel, N., Wafflart, J., Bonichon, F., de Mascarel, I., Trojani, M., Durand, M., Avril, A., and Coindre, J. M. (1995). The prognostic value of c-erbB2 in primary breast carcinomas: a study on 942 cases. Breast Cancer Res Treat *35*, 283-291.

Quinn, M. M., Punnett, L., Christiani, D. C., Levenstein, C., and Wegman, D. H. (1987). Modernization and trends in occupational health and safety in the People's Republic of China 1981-1986. Am J Ind Med *12*, 499-506.

Rae, J. M., Scheys, J. O., Clark, K. M., Chadwick, R. B., Kiefer, M. C., and Lippman, M. E. (2004). EGFR and EGFRvIII expression in primary breast cancer and cell lines. Breast Cancer Res Treat *87*, 87-95.

Rakha, E. A., Boyce, R. W., Abd El-Rehim, D., Kurien, T., Green, A. R., Paish, E. C., Robertson, J. F., and Ellis, I. O. (2005). Expression of mucins (MUC1, MUC2, MUC3, MUC4, MUC5AC and MUC6) and their prognostic significance in human breast cancer. Mod Pathol.

Ravaioli, A., Bagli, L., Zucchini, A., and Monti, F. (1998). Prognosis and prediction of response in breast cancer: the current role of the main biological markers. Cell Prolif *31*, 113-126.

Reddish, M., MacLean, G. D., Koganty, R. R., Kan-Mitchell, J., Jones, V., Mitchell, M. S., and Longenecker, B. M. (1998). Anti-MUC1 class I restricted CTLs in metastatic breast cancer patients immunized with a synthetic MUC1 peptide. Int J Cancer *76*, 817-823.

Rivkin, S. E., Green, S., O'Sullivan, J., Cruz, A. B., Abeloff, M. D., Jewell, W. R., Costanzi, J. J., Farrar, W. B., and Osborne, C. K. (1996). Adjuvant CMFVP versus adjuvant CMFVP plus ovariectomy for premenopausal, node-positive, and estrogen receptor-positive breast cancer patients: a Southwest Oncology Group study. J Clin Oncol *14*, 46-51.

Robertson, J. F., Pearson, D., Price, M. R., Selby, C., Badley, R. A., Pearson, J., Blamey, R. W., and Howell, A. (1990). Assessment of four monoclonal antibodies as serum markers in breast cancer. Eur J Cancer *26*, 1127-1132.

Roger, P., Sahla, M. E., Makela, S., Gustafsson, J. A., Baldet, P., and Rochefort, H. (2001). Decreased expression of estrogen receptor beta protein in proliferative preinvasive mammary tumors. Cancer Res *61*, 2537-2541.

Ross, J. S., Fletcher, J. A., Linette, G. P., Stec, J., Clark, E., Ayers, M., Symmans, W. F., Pusztai, L., and Bloom, K. J. (2003). The Her-2/neu gene and protein in breast cancer 2003: biomarker and target of therapy. Oncologist *8*, 307-325.

Ryden, L., Jonsson, P. E., Chebil, G., Dufmats, M., Ferno, M., Jirstrom, K., Kallstrom, A. C., Landberg, G., Stal, O., Thorstenson, S., and Nordenskjold, B. (2005). Two years of adjuvant tamoxifen in premenopausal patients with breast cancer: a randomised, controlled trial with long-term follow-up. Eur J Cancer *41*, 256-264.

Sahin, A. A., Ro, J., Ro, J. Y., Blick, M. B., el-Naggar, A. K., Ordonez, N. G., Fritsche, H. A., Smith, T. L., Hortobagyi, G. N., and Ayala, A. G. (1991). Ki-67 immunostaining in node-negative stage I/II breast carcinoma. Significant correlation with prognosis. Cancer *68*, 549-557.

Salazar, L. G., Fikes, J., Southwood, S., Ishioka, G., Knutson, K. L., Gooley, T. A., Schiffman, K., and Disis, M. L. (2003). Immunization of cancer patients with HER-2/neu-derived peptides demonstrating high-affinity binding to multiple class II alleles. Clin Cancer Res *9*, 5559-5565.

Seth, P., Krop, I., Porter, D., and Polyak, K. (2002). Novel estrogen and tamoxifen induced genes identified by SAGE (Serial Analysis of Gene Expression). Oncogene *21*, 836-843.

Speirs, V., Malone, C., Walton, D. S., Kerin, M. J., and Atkin, S. L. (1999). Increased expression of estrogen receptor beta mRNA in tamoxifen-resistant breast cancer patients. Cancer Res *59*, 5421-5424.

Speirs, V., Skliris, G. P., Burdall, S. E., and Carder, P. J. (2002). Distinct expression patterns of ER alpha and ER beta in normal human mammary gland. J Clin Pathol *55*, 371-374.

Sullivan, A., Yuille, M., Repellin, C., Reddy, A., Reelfs, O., Bell, A., Dunne, B., Gusterson, B. A., Osin, P., Farrell, P. J., *et al.* (2002). Concomitant inactivation of p53 and Chk2 in breast cancer. Oncogene *21*, 1316-1324.

Suter, T. M., Cook-Bruns, N., and Barton, C. (2004). Cardiotoxicity associated with trastuzumab (Herceptin) therapy in the treatment of metastatic breast cancer. Breast *13*, 173-183.

Svane, I. M., Pedersen, A. E., Johnsen, H. E., Nielsen, D., Kamby, C., Gaarsdal, E., Nikolajsen, K., Buus, S., and Claesson, M. H. (2004). Vaccination with p53-peptide-pulsed dendritic cells, of patients with advanced breast cancer: report from a phase I study. Cancer Immunol Immunother *53*, 633-641.

Taylor-Papadimitriou, J., Burchell, J. M., Plunkett, T., Graham, R., Correa, I., Miles, D., and Smith, M. (2002). MUC1 and the immunobiology of cancer. J Mammary Gland Biol Neoplasia *7*, 209-221.

Tian, W., Osawa, M., Horiuchi, H., and Tomita, Y. (2004). Expression of the prolactin-inducible protein (PIP/GCDFP15) gene in benign epithelium and adenocarcinoma of the prostate. Cancer Sci *95*, 491-495.

Troup, S., Njue, C., Kliewer, E. V., Parisien, M., Roskelley, C., Chakravarti, S., Roughley, P. J., Murphy, L. C., and Watson, P. H. (2003). Reduced expression of the small leucine-rich proteoglycans, lumican, and decorin is associated with poor outcome in node-negative invasive breast cancer. Clin Cancer Res *9*, 207-214.

Ullrich, A., Coussens, L., Hayflick, J. S., Dull, T. J., Gray, A., Tam, A. W., Lee, J., Yarden, Y., Libermann, T. A., Schlessinger, J., and et al. (1984). Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature *309*, 418-425.

Umemura, S., Takekoshi, S., Suzuki, Y., Saitoh, Y., Tokuda, Y., and Osamura, R. Y. (2005). Estrogen receptor-negative and human epidermal growth factor receptor 2-negative breast cancer tissue have the highest Ki-67 labeling index and EGFR expression: gene amplification does not contribute to EGFR expression. Oncol Rep *14*, 337-343.

Vahabi, M. (2003). Breast cancer screening methods: a review of the evidence. Health Care Women Int *24*, 773-793.

Vegeto, E., Bonincontro, C., Pollio, G., Sala, A., Viappiani, S., Nardi, F., Brusadelli, A., Viviani, B., Ciana, P., and Maggi, A. (2001). Estrogen prevents the lipopolysaccharide-induced inflammatory response in microglia. J Neurosci *21*, 1809-1818.

Velculescu, V. E., and El-Deiry, W. S. (1996). Biological and clinical importance of the p53 tumor suppressor gene. Clin Chem *42*, 858-868.

Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W. (1995). Serial analysis of gene expression. Science *270*, 484-487.

Veronese, S. M., Gambacorta, M., Gottardi, O., Scanzi, F., Ferrari, M., and Lampertico, P. (1993). Proliferation index as a prognostic marker in breast cancer. Cancer *71*, 3926-3931.

Vogelstein, B., Lane, D., and Levine, A. J. (2000). Surfing the p53 network. Nature *408*, 307-310.

Volpi, A., De Paola, F., Nanni, O., Granato, A. M., Bajorko, P., Becciolini, A., Scarpi, E., Riccobon, A., Balzi, M., and Amadori, D. (2000). Prognostic significance of biologic

markers in node-negative breast cancer patients: a prospective study. Breast Cancer Res Treat *63*, 181-192.

Wadle, A., Mischo, A., Imig, J., Wullner, B., Hensel, D., Watzig, K., Neumann, F., Kubuschok, B., Schmidt, W., Old, L. J., *et al.* (2005). Serological identification of breast cancer-related antigens from a Saccharomyces cerevisiae surface display library. Int J Cancer.

Ward, A. M., Catto, J. W., and Hamdy, F. C. (2001). Prostate specific antigen: biology, biochemistry and available commercial assays. Ann Clin Biochem *38*, 633-651.

Watson, M. A., and Fleming, T. P. (1996). Mammaglobin, a mammary-specific member of the uteroglobin gene family, is overexpressed in human breast cancer. Cancer Res *56*, 860-865.

Weigelt, B., Verduijn, P., Bosma, A. J., Rutgers, E. J., Peterse, H. L., and van't Veer, L. J. (2004). Detection of metastases in sentinel lymph nodes of breast cancer patients by multiple mRNA markers. Br J Cancer *90*, 1531-1537.

Wesseling, J., van der Valk, S. W., Vos, H. L., Sonnenberg, A., and Hilkens, J. (1995). Episialin (MUC1) overexpression inhibits integrin-mediated cell adhesion to extracellular matrix components. J Cell Biol *129*, 255-265.

Westerberg, H., Gustafson, S. A., Nordenskjold, B., Silfversward, C., and Wallgren, A. (1980). Estrogen receptor level and other factors in early recurrence of breast cancer. Int J Cancer *26*, 429-433.

Wickerham, D. L., Fisher, B., Wolmark, N., Bryant, J., Costantino, J., Bernstein, L., and Runowicz, C. D. (2002). Association of tamoxifen and uterine sarcoma. J Clin Oncol *20*, 2758-2760.

Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., and Micklem, G. (1995). Identification of the breast cancer susceptibility gene BRCA2. Nature *378*, 789-792.

Yamauchi, H., Stearns, V., and Hayes, D. F. (2001). The Role of c-erbB-2 as a predictive factor in breast cancer. Breast Cancer *8*, 171-183.

Yoshida, K., and Miki, Y. (2004). Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. Cancer Sci *95*, 866-871.

Zhou, M., Chinnaiyan, A. M., Kleer, C. G., Lucas, P. C., and Rubin, M. A. (2002). Alpha-Methylacyl-CoA racemase: a novel tumor marker over-expressed in several human cancers and their precursor lesions. Am J Surg Pathol *26*, 926-931.