科技资料

Breast Epithelial Antigens

Molecular Biology to Clinical Applications

BREAST EPITHELIAL ANTIGENS

Molecular Biology to Clinical Applications

Edited by

Roberto L. Ceriani

John Muir Cancer and Aging Research Institute
Walnut Creek, California

PLENUM PRESS • NEW YORK AND LONDON

Library of Congress Cataloging in Publication Data

International Workshop on Monoclonal Antibodies and Breast Cancer (4th: 1990: San Francisco, Calif.)

Breast epithelial antigens: molecular biology to clinical applications / edited by Roberto L. Ceriani.

p. cm.

"Proceedings of the Fourth International Workshop on Monoclonal Antibodies and Breast Cancer, held November 5–6, 1990, in San Francisco, California."—T.p. verso.

Includes bibliographical references and index.

ISBN 0-306-44009-1

1. Breast—Cancer—Immunodiagnosis—Congresses. 2. Breast—Cancer—Immunotherapy—Congresses. 3. Tumor antigens—Congresses. I. Ceriani, Roberto L. II. Title.

[DNLM: 1. Antibodies, Monoclonal—diagnostic use—congresses. 2. Breast Neoplasms—diagnosis—congresses. 3. Breast Neoplasms—therapy—congresses. WP 870 I614b 1990]

RC280.B8I58 1990

616.99'4490756—dc20

DNLM/DLC 91-24095

for Library of Congress CIP

Proceedings of the Fourth International Workshop on
Monoclonal Antibodies and Breast Cancer,
held November 5–6, 1990, in San Francisco, California

ISBN 0-306-44009-1

A Division of Plenum Publishing Corporation
233 Spring Street, New York, N.Y. 10013

Printed in the United States of America

4TH INTERNATIONAL WORKSHOP ON MONOCLONAL ANTIBODIES AND BREAST CANCER

San Francisco, California
November 5-6, 1990

Organized by the John Muir Cancer & Aging Research Institute, with the cooperation of the International Association for Breast Cancer Research

WORKSHOP CHAIRPERSON

Dr. Roberto L. Ceriani
John Muir Cancer & Aging Research Institute

ORGANIZING COMMITTEE

Chairperson: Dr. Roberto L. Ceriani

Members: Dr. Jerry A. Peterson
Ms. Carolyn Klinepeter

ACKNOWLEDGEMENTS

The Organizing Committee for the 4th International Workshop on Monoclonal Antibodies and Breast Cancer, together with the John Muir Cancer & Aging Research Institute gratefully acknowledge the support of the following in making the Workshop possible:

SPONSORS

COULTER IMMUNOLOGY, HIALEAH, FLORIDA
ABBOTT LABORATORIES
CYTOGEN CORPORATION
TRITON BIOSCIENCES
Burroughs Wellcome Company
Bio-Rad Laboratories
Toys 2 Go (Discovery Toys) - Judy Corley
In Memory of Fred Vinson
American Speedy Printing - Walnut Creek
John Muir Medical Center
Heublein Fine Wine Group consisting of:
  Beaulieu Vineyards
  Inglenook - Napa Valley
  Christian Brothers - Greystone Cellars
  Quail Ridge Cellars
  Gustave - Niebaum Collections
  Rutherford Estate Cellars

PREFACE

The wealth of research results in the area of breast cancer diagnosis and therapy with monoclonal antibodies presented in previous workshops is now complemented with a new rendition of the proceedings of the 4th Workshop on Monoclonal Antibodies and Breast Cancer, held in San Francisco on November 5-6, 1990. Basic science findings reported in previous workshops have now percolated to the clinical level and have become immunoassays and imaging and therapy reagents, as the program shows us. Thus, the latest discoveries in immunology, biochemistry and molecular biology of breast epithelial antigens and their corresponding antibodies have produced newer diagnostic tests and therapeutic approaches that are altering and improving the way we attack breast cancer.

The recent spectacular and rapid advancements in the molecular biology of several of the breast epithelial antigens are presented in this volume. The way in which the final assembly of different components of the breast antigens is achieved and their functions are now within our grasp as a result of new understanding of molecular structure of these breast antigens. In addition, newer immunoassays aiming at the earliest detection of the disease are also described that integrate with promising attempts at imaging and radioimmunotherapy to set the stage for new oncological possibilities in breast cancer treatment. All these areas of intense involvement of scientists with diverse specialties are presented in this volume, which proves the need for multidisciplinary approaches to increase our chances for success in this field of medical research.

The publication of these collected papers represents the cutting edge of an area of rapid scientific and clinical development. They are grouped into 4 chapters that include the molecular biology of breast epithelial antigens, their cell biology and immunology, the use of anti-breast epithelial antibodies in serum tests, and imaging and therapeutic application of these novel drugs in breast cancer.

R.L. Ceriani

CONTENTS

SESSION I

SESSION II

SESSION III

SESSION IV

MOLECULAR ANALYSIS OF H23 EPITHELIAL TUMOR ANTIGEN - DIFFERENTIALLY SPLICED FULL LENGTH cDNAs AND GENE

D. H. Wreschner, I. Tsarfaty, M. Hareuveni, J. Zaretsky, N. I. Smorodinsky, M. Weiss[1], S. Zrihan[2], M. Burstein[3], J. Horev, P. Kotkes[4], R. Lathe[2], C.A. Hart[3], K. McCarthy[3], C. Williams[4], A. Dion[4] and I. Keydar

Dept. Microbiology/Cell Biology, Tel Aviv University, Ramat Aviv, Israel 69978; [1]Dept. Medicine, Tel Hashomer, Israel; [2]AFRC, University of Edinburgh, King's Bldgs, Edinburgh, U.K.; [3]Medical Microbiology, University of Liverpool, Prescott St., Liverpool, U.K.; [4]Center for Molecular Medicine and Immunology, Newark, New Jersey, U.S.A.

ABSTRACT

The isolation and characterization of the complementary DNAs (cDNAs) and gene which code for an epithelial tumor antigen (H23-ETA), aberrantly expressed in human breast tumor tissue, are described here. A diversity of H23-ETA protein forms is generated by a series of alternative splicing events that occur in regions located upstream and downstream to a central tandem 20 amino acid (aa) repeat array (TRA) that is rich in proline, serine and threonine residues. The upstream region shows that differential usage of alternative splice acceptor sites generates two protein forms containing putative signal peptides of varying hydrophobicities located at the NH_2 terminus. The region downstream to the tandem repeat array indicates that one mRNA transcript is collinear with the gene and defines a 160 aa open reading frame (secreted or sec form). A second cDNA correlates with a mRNA that is generated by a series of splicing events and codes for 149 aa downstream to the TRA, identical with the aa sequence of the unspliced cDNA, after which it diverges and continues for an additional 179 aa. This sequence (transmembrane or tm form) contains a highly hydrophobic transmembrane domain of 28 aa followed by a hydrophilic "transfer-stop signal" (Arg Arg Lys) and a cytoplasmic domain of 72 aa. The various protein forms (alternative signal sequences, secreted and transmembrane) are likely routed to different cytoplasmic, cell membrane and extracellular compartments. Reverse PCR indicates that the relative ratios of the alternatively spliced forms vary in different epithelial tissues. To identify the individual protein species, monoclonal antibodies (mAb) are being generated against synthetic peptides unique to each form. The H23-ETA gene was also isolated and sequenced, demonstrating a putative promoter region that includes a 'TATA' box, Sp1 binding elements and an upstream putative hormone responsive element. Commensurate with these findings, H23-ETA expression was increased following hormonal treatment of BT549 breast tumor cells. These molecular studies have unravelled novel H23-ETA protein and gene structures, and facilitate future investigations that will focus on H23-ETA function and interaction with other cellular proteins.

Breast Epithelial Antigens, Edited by R.L. Ceriani
Plenum Press, New York, 1991

INTRODUCTION

A number of monoclonal antibodies (mAbs) have been developed that recognize an antigen of epithelial cell origin that is expressed primarily in breast tumor tissue and, to lesser extents, in other malignant epithelial tissues (1-11). This antigen is likely a constituent of normal epithelial cells which, in malignant breast tissue, undergoes quantitative and/or qualitative changes. The H23 mAb detects the breast cancer associated antigen (H23-ETA) in both cytoplasm of human breast cancer cells and body fluids of breast cancer patients (11,12). Since H23-ETA serum levels have clinical significance and correlate with the severity of disease (12), an analysis of H23-ETA at the molecular level was undertaken.

The cDNA coding for the epitope recognized by H23 is composed of tandem 60 base pair (bp) repeating units that code for a 20-amino-acid (aa) repeat motif rich in proline, serine, threonine and alanine (13,14). Other mAbs (DF3, HMFG1 and HMFG2, MAM6) that are also directed against a breast cancer antigen, were used to isolate cDNAs that are composed of almost identical 60-bp tandem repeats (15-19). Southern blot analyses indicated that the H23-ETA gene is highly polymorphic and 70-80% of individuals are heterozygous at this locus (13,14,20). The different allelic sizes result from variation in the number of repeating 60-bp units present within the genomic repeat array and the polymorphism detected in the H23-ETA protein products correlates with the various allelic sizes (14).

In addition to its presence in serum, the cellular localization of the breast tumor associated antigen has been variously designated as apical, membranous, intracytoplasmic or focal (1-11). It appears therefore that varying forms of H23-ETA may localize to different cellular and extracellular compartments. In order to understand the postulated different H23-ETA protein forms and to clarify its overexpression in breast cancer, both full length cDNAs (13,21) and the H23-ETA gene were isolated and characterized (21-24). These molecular studies, presented here, have unravelled novel protein and gene structures, that may elucidate H23-ETA function in tumor progression, as well as the regulatory mechanisms responsible for its overexpression in breast cancer tissue.

RESULTS AND DISCUSSION

H23-ETA amino acid sequence - NH_2 terminal to tandem repeats

A number of cDNAs correlating with the region 5'upstream to the central tandem 20 aa (60 bp) repeat array (TRA) as well as with the 3' downstream region were isolated, sequenced and characterized (Figs. 1, 2 and 3). The NH_2 terminal aa sequence deduced from the pSe 2 cDNA (Figs. 1, 2 and 3) open reading frame (ORF), demonstrates the initiating methionine followed by a highly hydrophobic 13-aa peptide that includes 5 tandem leucine residues. Because of its size, hydrophobicity and proximity to the amino terminus it seems likely that this domain represents a signal sequence. H23-ETA secretion or its insertion into the plasma cell or/and cytoplasmic membranes may be mediated by this classical signal sequence. Interestingly, the ORF determined by a second 5' cDNA (pSe 4, Fig. 1) indicates diversity in the signal sequence domain (Fig. 2A and B). Alternative usage of two splice acceptor sites generates the variability between the pSe-2 and pSe-4 cDNA sequences. The splice event generating the pSe-2 cDNA sequence determines a Thr-Val-Val peptide. The ORF of the pSe-4 cDNA sequence, however, interrupts this tripeptide and determines an in-frame insertion of nine aa (Fig. 2A and B). As the residues introduced by this alternative splice change the hydrophobicity of the signal peptide region, cellular routing and targeting of H23-ETA may also be affected.

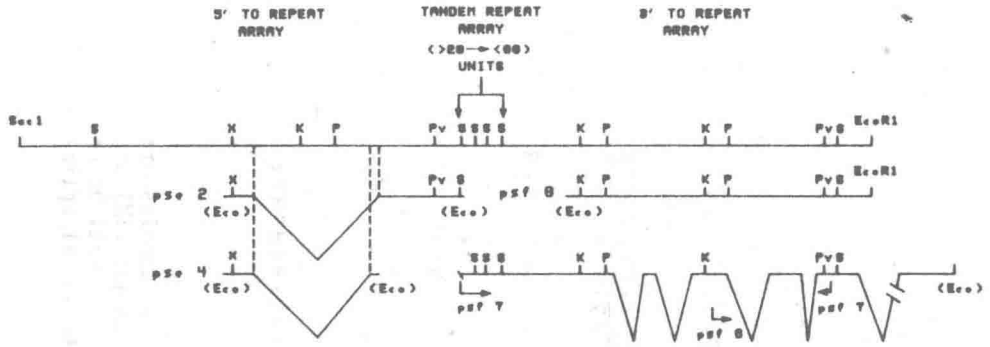

Fig. 1 Restriction map of genomic fragment coding for H23-ETA and localization of isolated cDNA inserts. The top line represents the restriction map of a 5.3-kb SacI-EcoRI genomic fragment part of which codes for H23 ETA. The various cDNA inserts obtained are shown below the genomic restriction map. For convenience the tandem array is depicted as three repeat units when, in fact, it ranges over 20-80 repeats. The restriction enzymes KpnI, PstI, PvuII, SmaI and XmnI are represented by K, P, PV, S and X respectively, and the synthetic EcoRI sites are shown in parentheses. The genomic fragments used to reprobe the cDNA library are the SmaI-PstI and PstI-PstI, PstI-EcoRI fragments located 5' and 3' to the repeat array, and are designated pse and psf, psg respectively. Pse 2 and 4 and psf6, 7 and 8 are the cDNA inserts isolated.

H23-ETA amino acid sequence - tandem repeats

The ORF determined by the in-frame initiation methionine extends pSe 2 into the 60-bp repeat unit which codes for a highly conserved proline-rich 20-aa repeat motif that also contains 3 threonine, 2 serine, 2 glycine and 2-5 alanine residues (Fig. 2 and 3).

The epitope in H23-ETA recognized by H23-ETA mAb is situated within this repeat motif. Synthetic peptides correlating with different parts of the 20 aa repeat motif, determined by the ORF, were synthesized and analyzed for immune reactivity with H23 mAb. The Pro-Asp-Thr-Arg sequence was found to be essential for recognition by H23 mAb, although flanking residues on both the NH_2 and COOH termini are also required for maximal immune reactivity. It is notable that this tetrapeptide represents the hydrophilic segment and possibly, therefore, the most antigenic part of the 20 aa repeat motif.

H23-ETA amino acid sequence - COOH terminal to tandem repeats

The region downstream to the TRA (Fig. 1, psF8, psF7 and psF6) indicates that one mRNA transcript (cDNA psF8) is unspliced and collinear with the gene and defines a 160 aa ORF (secreted or sec Form). This H23-ETA protein form contains a short 12 amino acid hydrophobic region bounded by potential N-glycosylation sites at positions 289 and 315 (Fig. 3B). The small size of this hydrophobic region is insufficient to support transmembrane localization, but it may function as a signal for glycolipid mediated membrane anchorage. The aa sequence of the sec form, on the carboxyl side of the TRA, is serine rich and contains four potential N-linked glycosylation sites (NxS/T).

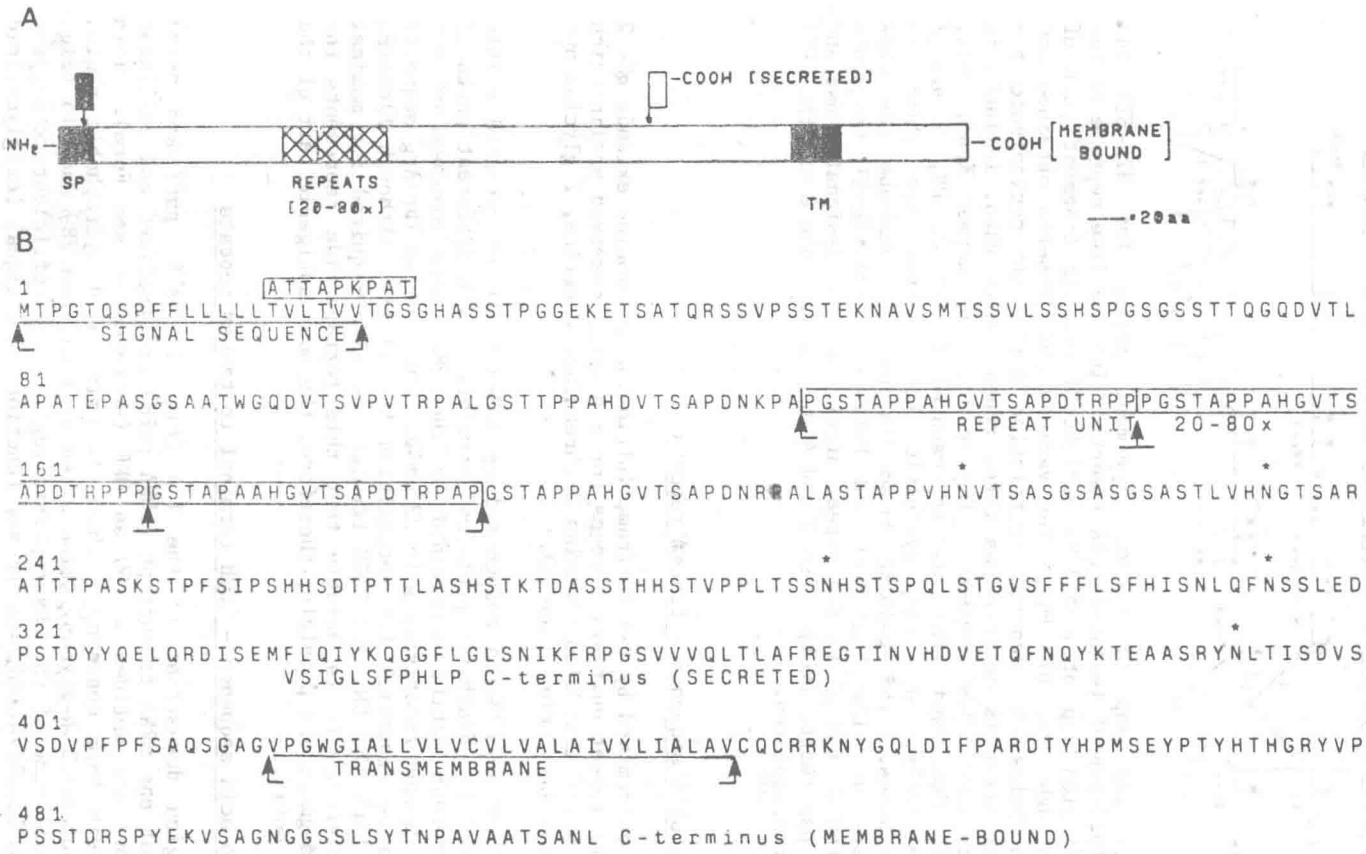

Fig. 2 Scheme demonstrating the structure and sequence of various ETA forms. A) Different domains are presented schematically: the hydrophobic signal peptide (SP) and transmembrane domain (TM) are in black and the highly conserved 20 amino acid repeat units are crosshatched. The small boxes above the membrane bound form represent the protein variants resulting from differential splicing. B) The complete ETA amino acid sequences.

A second cDNA form (psF7 and psF6, Fig. 1) codes for 149 aa downstream to the TRA that are identical with the aa sequence of the unspliced cDNA, after which it diverges and continues for an additional 179 aa. This sequence (transmembrane or tm form) contains a highly hydrophobic transmembrane domain of 28 aa followed by a cytoplasmic domain of 72 aa (Figs. 2, 3 and 5). Viable intact breast tumor cells can be immunofluorescently stained with H23 mAb, indicating that the tandem 20 aa repeats are located in the extracellular domain of the H23-ETA tm form. The presence of 5 potential N-linked glycosylation sites (AsnxSer/Thr) in this region lends credence to this hypothesis.

That the H23-ETA tm form may be involved in signal transmission and serve as a receptor for an as yet unidentified ligand, is supported by the fact that the cytoplasmic domain comprises a long 72 aa tail. Were the H23-ETA tm form solely a structural transmembrane protein, a significantly smaller cytoplasmic domain would be sufficient to anchor the protein securely in the cell membrane. It therefore seems plausible that the cytoplasmic domain may interact with cellular proteins and elicit changes in cell behaviour by transmitting signals from the cell exterior. Whether ligand binding to the extracellular domain or direct interaction of the ETA-tm form with extracellular matrix, substratum or adjacent cells induces such signals, remains to be determined.

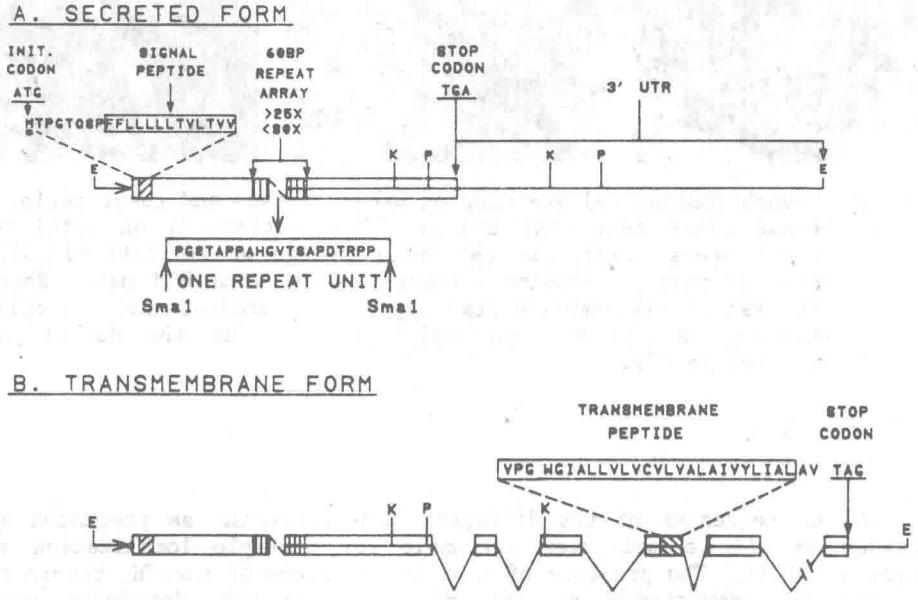

Fig. 3 Scheme of differential splicing that may generate secreted and membrane-bound forms of the epithelial tumor antigen. The initiation codon (ATG) and stop codons (TGA and TAG, secreted and transmembrane forms respectively) are indicated. The restriction sites for KpnI and PstI are designated by K and P. The one-letter code for amino acids is used and the region bracketed by 3'UTR represents the 3' untranslated region.

Multiple H23-ETA forms - multiple localization sites

As H23-ETA can be detected in the serum of patients, it is obviously secreted into the peripheral circulation. Furthermore, immunohistochemical analyses of breast tumor sections stained with H23 mAbs demonstrated primarily intracytoplasmic staining and an immunofluorescence study indicated that H23-ETA is also membrane bound (11). H23-ETA localization to different sites is presented in Fig. 4. Immunohistochemical staining of T47D breast tumor cells and a breast tumor tissue paraffin section (Figs. 4A and B) demonstrates both membrane and cytoplasmic staining. On the other hand, staining of a benign breast fibroadenoma paraffin section (C) shows extracellular glycocalyx staining at the ductal apical surfaces.

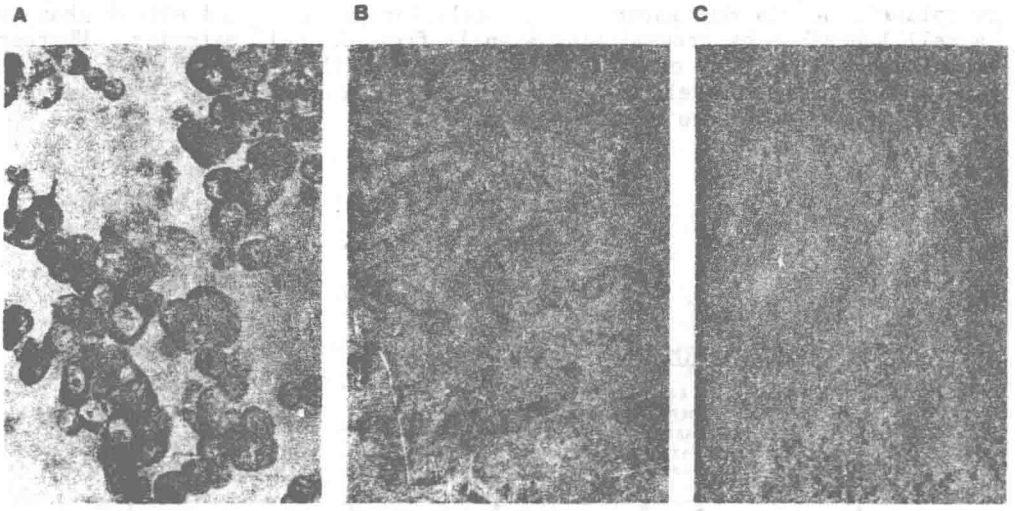

Fig. 4 Immunohistochemical staining of breast tissue and tumor cells. T47D breast tumor cells (A) and paraffin sections of an infiltrating ductal breast carcinoma (B) and a benign breast fibroadenoma (C) were indirectly immunoperoxidase stained with H23 mAb. Note the cytoplasmic and membrane staining in (A), predominantly cytoplasmic staining in (B) and glycocalyx staining at the ductal apical surfaces in (C).

The aa sequences of the different H23-ETA forms, as presented here, provides us with a molecular rationale for multiple localization sites (Figs. 3 and 5). The presence of a signal sequence at the NH_2 terminus and sec sequence downstream to the repeat array may determine protein secretion. On the other hand, presence of the NH_2 signal sequence and transmembrane sequence downstream to the repeat array could target the protein for cell membrane localization. The effect of nonapeptide insertion into the signal sequence on post-translational processing and cellular routing are, at present, not known. It may well be, however, that insertion of the nine residue peptide disrupts signal sequence cleavage, rendering the uncleaved hydrophobic NH_2 terminus, a lipophilic endoplasmic reticulum membrane anchor, thereby explaining H23-ETA cytoplasmic localization.

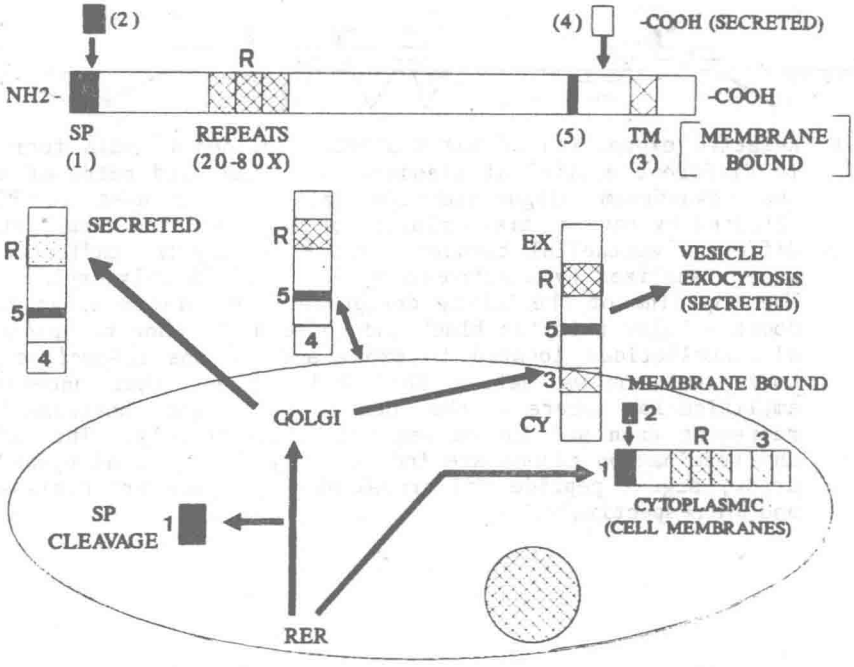

Fig. 5 Scheme demonstrating possible localization sites of the various H23-ETA forms. The H23-ETA protein is illustrated at the top of the figure with NH_2 and COOH termini at the left and right borders respectively. The signal peptide (SP) is indicated by the left black box (1) near the NH_2 terminus and the 9 aa insertion in the signal sequence, determined by a 5' alternative splice, is designated by the black box (2). The 3 densely hatched boxes (R) in the middle represent tandem 20 aa repeats which may occur in H23-ETA from 20 to 80 times. The large 28 aa highly hydrophobic domain (TM) present in the H23-ETA transmembrane form is represented by the sparsely cross-hatched box (3) close to the COOH terminus. The 11 aa sequence, unique to the putative secreted H23-ETA form and representing its COOH terminus (COOH secreted), is illustrated by the open box (4). The shorter 12 aa hydrophobic region present both in the transmembrane and secreted H23-ETA forms is designated by the slender black box (5). Attachment of the secreted form to the cell membrane may be mediated by this short hydrophobic region (5). The extracellular and cytoplasmic domains of the transmembrane form are indicated by Ex and Cy respectively. The diagram is schematic and is not drawn to scale.

Relative expression of alternatively spliced H23-ETA forms varies in different epithelial tissues

It is of obvious interest to know whether all H23-ETA forms are equally expressed in epithelial tissues or, alternatively, whether different epithelial tissues preferentially express certain H23-ETA forms. This was investigated by using reverse PCR technology in which the mRNA is reverse transcribed into cDNA which is then amplified with a chosen pair of downstream and upstream primers (Fig. 6). The sizes of the PCR products

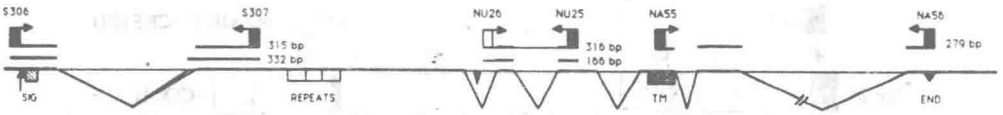

Fig. 6 Relative expression of alternatively spliced H23-ETA forms varies in different epithelial tissues. The indicated pairs of upstream and downstream oligonucleotide primers were used to PCR cDNA obtained by reverse transcription of mRNA isolated from a number of different epithelial tissues. The PCR products (indicated in bp) were visualized by electrophoresis through 8% polyacrylamide gels. The top line of the scheme designates the various oligonucleotide codes - below this the black and white boxes show the positions of oligonucleotides located in exons and introns respectively. The horizontal arrows define the cDNA segment that underwent PCR amplification whereas the heavy and light horizontal lines represent exon and intron sequences respectively. The initiation and termination codons are indicated by the vertical upward facing arrow; signal peptide and transmembrane region are designated SIG and TM respectively.

reflect the alternatively spliced forms of mRNA (see Fig. 6). The analysis performed with RNA from a variety of different epithelial tissues confirmed that the relative ratios of expression of the alternatively spliced forms vary in different epithelial tissues.
For example, the pair of primers designated S306 and S307 will lead to 315 bp and/or 332 bp reverse PCR products depending on the alternative splice taken within the signal sequence region. An analysis with this primer pair demonstrated that whereas some epithelial tissues express both alternatively spliced forms, other tissues exclusively expressed one or the other species. It is clear, therefore, that certain tissues may preferentially express specific H23-ETA forms.

Evidence supporting the existence of a secreted H23-ETA form

As previously noted a cDNA, psF 8 (Fig. 1), located in the region downstream to the TRA was isolated, and nucleotide sequencing indicated that the mRNA transcript from which it is derived is unspliced and collinear with the gene (Fig. 1). Retention of the intron immediately downstream to the TRA introduces an in-frame termination codon that will result in a protein product that, as lacking the transmembrane domain, could be secreted from the cell (Figs. 2 and 3).

Precedent exists for intron inclusion in mature mRNA in one tissue, whereas the same intron is removed by splicing in other tissues, resulting in a mRNA with a longer open reading frame that produces a larger protein product (26-28) (for review see 25). For example, P element transposition in Drosophila depends on germ-line pre-mRNA splicing of the P element third intron, producing a germline mRNA that is translated into an active 87Kd transposase protein. In somatic tissues, however, this same P element third intron is not spliced out resulting in a mature mRNA that contains an in-frame termination codon within the unspliced intron and is translated into a truncated 66 Kd repressor protein that functions as a negative regulator of transposition (27,28).

It is of course possible that the isolated psF 8 cDNA described here represents a pre mRNA transcript present within the population of

polyadenylated RNAs used to prepare the cDNA library, thereby explaining gene collinearity and intron presence. Alternatively, it may represent a bona fide mature mRNA transcript coding for a secreted H23-ETA form. Support for this latter possibility is derived from several lines of evidence.

Firstly, Northern blot analysis of RNA extracted from both breast tumor cell lines and primary breast tumor tissue demonstrated hybridization to sequences within the introns of the transmembrane mRNA (29).

Secondly, reverse PCR was performed on RNA extracted from a variety of epithelial tissues using an upstream oligonucleotide NU26, that is located within the first intron 3' to the TRA and a downstream oligonucleotide Nu25 (Fig. 6). Two PCR products were observed - one sized at 316 bp that is collinear with the gene and a second smaller 166 bp product that corresponds to a mRNA that is spliced in the second intron downstream to the TRA (Fig. 6). As the upstream NU26 primer is located in the first intron downstream to the TRA, both the 316 bp and 116 bp reverse PCR products obviously contain this intron downstream to the TRA (Fig.6)

Thirdly, mAbs have been recently generated against the amino acid sequence SIGLSFPMLP that represents the C-terminal 10 aa, unique to the putative secreted H23-ETA form (Fig. 2). These sec specific mAbs, when analyzed in an ELISA assay, not only recognized the synthetic peptide with which the mice were immunized, but also bound H23-ETA secreted into culture medium by T47D breast tumor cells. Furthermore, the binding of these sec specific mAbs to H23-ETA was competed out by preincubation of the mAbs with the synthetic sec specific peptide.

These data support the hypothesis that a secreted H23-ETA form exists, and work on its further characterization is presently in progress.

Isolation and characterization of the H23-ETA gene - Identification of a hormonal responsive element

To understand mechanisms that regulate expression of the H23-ETA gene and to elucidate its overexpression in breast tumor tissue, a genomic library prepared with DNA from the MCF7 breast tumor cell line was used to isolate and subsequently characterize the H23-ETA gene (22). The H23-ETA EcoRI-EcoRI 7.5 Kb genomic fragment isolated was designated I7.5. It contained a tandem repeat array of 2.3 Kb and unique sequences both 5' upstream and 3' downstream to the TRA. A SacI-EcoRI 5.3 Kb fragment of I7.5 was sequenced and comparison with cDNA nt sequences identified different domains in the H23-ETA gene. As in the cDNA sequence, the genomic sequence contains the same consensus initiation sequences, putative signal sequence divided by an intron, TRA and unique sequences 3' to the TRA.

Comparison of the cDNA and gene sequences revealed one intron located 5' to the TRA that divides the Thr and Val residues towards the end of the signal peptide. The 499 bp intron contains a highly purine rich region (78% AG) over the first 419 bp, followed by a pyrimidime rich sequence (75% CT) for the remaining 80 bp. A nucleotide homology search revealed a putative enhancer sequence situated within the purine rich region of the intron that showed 86% identity with a 28 bp murine cellular enhancer of retroviral gene expression.

The genomic sequence upstream from the 5' terminal cDNA sequences revealed a putative promoter region consisting of TATA sequences (Hogness box) flanked on the 5' and 3' sides by G+C rich regions that include several Sp1 binding sites.