PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C07K 7/06, A61K 37/02 C12N 15/12, 15/62, C12Q 1/68 C12P 21/08	A1	(11) International Publication Number: WO 90/05142 (43) International Publication Date: 17 May 1990 (17.05.90)

(21) International Application Number: PCT/GB89/01340

(22) International Filing Date: 10 November 1989 (10.11.89)

(30) Priority data:
8826362.9 10 November 1988 (10.11.88) GB
8907660.8 5 April 1989 (05.04.89) GB

(71) Applicant *(for all designated States except US)*: IMPERIAL CANCER RESEARCH TECHNOLOGY LTD. [GB/GB]; Sardinia House, Sardinia Street, London WC2A 3NL (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : TAYLOR-PAPADIMITRIOU, Joyce [GB/GB]; 9 Cedar Road, Berkhampstead, Hertfordshire (GB). GENDLER, Sandra [US/GB]; 20 St James Mansions, West End Lane, West Hampstead, London NW6 (GB). BURCHELL, Joy [GB/GB]; 4 Whites Cottages, Fletching, Uckfield, West Sussex (GB).

(74) Agents: CRESSWELL, Thomas, Anthony et al.; J.A. Kemp & Co., 14 South Square, Gray's Inn, London WC1R 5EU (GB).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), DK, FI, FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), NO, SE (European patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: POLYPEPTIDES

(57) Abstract

Polypeptides including the sequence PDTRP and immunological analogues correspond with an important epitope of the human polymorphic epithelial mucin abnormally expressed in breast cancers and are useful in immunisation and for generating antibodies for diagnosis and therapy. Nucleic acid fragments encoding PDTRP or an immunlogical analogue are useful in probing for diagnosis and producing the polypeptides.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MR	Mauritania
BE	Belgium	GA	Gabon	MW	Malawi
BF	Burkina Fasso	GB	United Kingdom	NL	Netherlands
BG	Bulgaria	HU	Hungary	NO	Norway
BJ	Benin	IT	Italy	RO	Romania
BR	Brazil	JP	Japan	SD	Sudan
CA	Canada	KP	Democratic People's Republic of Korea	SE	Sweden
CF	Central African Republic			SN	Senegal
CG	Congo	KR	Republic of Korea	SU	Soviet Union
CH	Switzerland	LI	Liechtenstein	TD	Chad
CM	Cameroon	LK	Sri Lanka	TG	Togo
DE	Germany, Federal Republic of	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

POLYPEPTIDES

The present invention relates to a polypeptide having a sequence corresponding to an antigenic epitope on human polymorphic epithelial mucin (PEM), to its production and use in diagnosis and therapeutic treatment of cancer.

In WO88/05054 there is described the discovery of a tandem repeat sequence of amino acid residues appearing in the human PEM protein expressed by normal or transformed epithelial cells. At least in the case of breast cancer, malignant cells express an under-glycosylated form of PEM which is detected by the SM3 monoclonal antibody. That antibody does not react with fully processed PEM as produced by normal cells so that the antibody has uses in the diagnosis and therapy of breast cancer.

Further investigations have surprisingly revealed that the epitope recognised by SM3 is a small, continuous amino acid sequence occurring within the tandem repeat sequence previously described. Identification of the epitope permits synthetic preparation of peptides mimicking the epitope which will be useful in generating antibodies with improved specificity for tumour as opposed to normal PEM.

The present invention therefore provides a

(a) polypeptides of formula (I)

$$\{X\}_x\{(A)_mPDTRP(A)_n\}_y\{X\}_z \qquad (I)$$

wherein x and z are the same or different and each is 0 or an integer of one or more;

y is an integer of one or more

each X, where present, is an independently selected amino acid residue;

m and n are the same or different and each is 0 or one and A,P,D,T and R are the internationally recognised one-letter symbols for the residues of the amino acids, alanine, proline, aspartic acid, threonine and arginine respectively;

(b) polypeptides containing a sequence immunologically analogous to $\{(A)_mPDTRP(A)_n\}$;

and (c) polypeptides as defined in (a) or (b) above bearing one or more saccharide moieties;

and pharmaceutically accept salts thereof;

other than a polypeptide specifically disclosed in WO88/05054 and other than beta-glucuronidase, which also contains the sequence PDTRP.

Polypeptides as defined under (a), (b) and (c) above, subject to the same exclusions, will hereafter be referred to as "PDTRP epitopes".

The simplest polypeptides according to the present invention are the pentapeptide PDTRP, immunological

analogues thereof and glycosylated derivatives of the pentapeptide and its analogues. Hexapeptides according to the invention include APDTRP, PDTRPA, immunological analogues thereof and glycosylated derivatives of the hexapeptides and analogues. Heptapeptides according to the invention include APDTRPA and immunological analogues thereof and glycosylated derivatives of the heptapeptide and analogues. Larger polypeptides according to the invention include additional amino acid residues at the amino and/or carboxy terminal of the penta-, hexa- and hepta-peptides, analogues and derivatives. Such polypeptides may comprise, for instance, up to about 200, for instance up to about 50, particularly up to 30, preferably up to 20 and more preferably from 10 to 15 amino acid residues. Maximum and preferred values of x and z are determined accordingly. The polypeptides must include at least one five amino acid residue sequence PDTRP or an immunological analogue thereof and may include several repeats of this sequence, for instance 2 or 3 repeats and perhaps up to 100 or more such repeats. Each such repeat independently may include one or both of the amino and/or carboxy terminal alanine residues, or immunological equivalents thereof and may be extended at either or both the N- and C-termini with other amino acids. Typically each repeat will comprise 19 or less amino acid residues. Where the polypeptide is glycosylated, any repeats of the PDTRP sequence or analogues thereof are independently glycosylated, for instance by attachment of a linkage sugar

such as N-acetyl galactosamine on the threonine residue of PDTRP and/or on threonine or serine residues outside the epitope PDTRP sequence or which are included within the epitope sequence of an immunological equivalent of PDTRP (for instance, wherein the threonine residue has been replaced by a serine residue).

Immunological analogues according to the present inventions are polypeptides having a sequence corresponding to that of formula (I) except that the sequence $(A)_m PDTRP(A)_n$ is modified at at least one amino acid residue by substitution and/or is modified by insertion of one or more additional amino acid residues provided that the modified sequence is capable of acting as an epitope for the generation of antibodies and provided that the antibodies so generated which recognise the modified epitope also recognise at least one of

(a) a polypeptide of formula (I)

(b) a polypeptide comprising the tandem repeat sequence as defined in WO88/05054 and

(c) human PEM core protein

but, in each case have reduced or no reaction and preferably substantially do not react with fully processed human PEM glycoprotein as produced by the normal lactating human mammary gland.

Immunological analogues, also known as mimotopes, may be designed for instance as described in "Peptide Protein and Gene Technology Advances", Issue No. 2, Cambridge Research Biochemicals UK (copy attached) and in

"The delineation of Peptides able to mimic assembled Epitopes", H.M. Geysen et al in "Synthetic Peptides as Antigens" (Ciba Foundation Symposium 119), R. Porter and J. Whelan (Eds), Pitman London 1986 pp 130-149.

Preferably the polypeptides of the present invention are synthetic.

PDTRP epitopes according to the invention may be produced de novo by synthetic methods or by expression of the appropriate DNA fragments described below by recombinant DNA techniques and expressed with or without glycosylation in human or non-human cells.

A peptide may be built up from single amino acids and/or preformed peptides of two or more amino acids in the order of the sequence of the desired peptide and optionally the resultant peptide may be converted into a pharmaceutically acceptable salt if desired.

Solid-phase or solution methods may be employed. In solid-phase synthesis, the amino acid sequence of the desired peptide is built up sequentially from the C-terminal amino acid which is bound to an insoluble resin. When the desired peptide has been produced, it is cleaved from the resin. When solution-phase synthesis is employed, the desired peptide may again be built up from the C-terminal amino acid but the carboxy group of this acid remains blocked throughout by a suitable protecting group, which is only removed at the end of the synthesis. Whichever technique is employed, each amino acid added to the reaction system typically has a protected amino group

and an activated carboxy group. An amino group may be protected by the fluoren-9-ylmethoxycarbonyl (Fmoc) or t-butoxycarbonyl (Boc) group. A carboxy group may be activated as a pentafluorophenyl or 1-oxo-2-hydroxy-dihydrobenzotriazine ester. Functional side-chain groups such as the side chain amino group of lysine and the side chain hydroxy group of threonine may be protected too. Each condensation step may be effected in the presence of dicyclohexylcarbodiimide or 1-hydroxybenzotriazole. After each step in the synthesis, the amino-protecting group is removed. Side-chain functional group protecting groups are generally removed at the end of the synthesis.

To convert the resultant peptide into a pharmaceutically acceptable salt, it may be treated with an organic or inorganic acid to form an acid addition salt. Suitable acids include acetic, succinic and hydrochloric acid. Alternatively, the peptide may be converted into a carboxylic acid salt such as the ammonium salt or an alkali metal salt such as the sodium or potassium salt by use of an appropriate base.

The PDTRP epitopes of the invention may also be prepared by recombinant DNA methodologies; an expression vector is prepared which incorporates a DNA sequence encoding the PDTRP epitope and which is capable of expressing the PDTRP epitope when provided in a suitable host. The DNA sequence is located between translation start and stop signals in the vector. Appropriate transcriptional control elements are also provided, in

particular a promoter for the DNA sequence and a transcriptional termination site. The DNA sequence is provided in the correct reading frame so as to enable expression of the peptide to occur in a host compatible with the vector.

Any appropriate host-vector system may be employed. For instance, the vector may be a plasmid or viral vector and the host may be any prokaryotic or eukaryotic cells such as bacterial or yeast cell or cells of a mammalian cell line.

Alternatively the PDTRP epitopes may be obtained by stripping carbohydrate from native human mucin glycoprotein (which itself may be produced by isolation from samples of human tissue or body fluids or by expression and full processing in a human cell line) (Burchell et al., Cancer Research, 47, 5467-5482, (1987), Gendler et al., P.N.A.S., 84 6060-6064, (1987)), and digesting the core protein.

PDTRP epitopes linked to carrier proteins such as keyhole limpet haemocyanin, albumen or thyroglobulin are also within the invention. Linkage may be by known chemical methods or by expression of the PDTRP epitope and carrier protein as a fusion protein by genetically transformed cells as herein described.

In another aspect the invention provides a nucleic acid fragment encoding a PDTRP epitope.

Fragments according to the invention may be RNA or DNA and DNA fragments may be single or double stranded.

Preferred such nucleic acid fragments are DNA fragments containing a continuous coding sequence of formula (II) which encodes a PDTRP epitope of formula (I) as hereinbefore defined,

$$\{XXX\}_x\{(GCX)_m CCX\ GAPy\ ACX\begin{bmatrix}CGX\\ \text{or}\\ AGPu\end{bmatrix}CCX(GCX)_n\}_y\{XXX\}z \qquad (II)$$

wherein m, n, x, y and z are all as defined in relation to formula (I);

X is any deoxyribonucleotide

XXX is a codon encoding an amino acid residue;

Py is C or T;

Pu is A or G;

and A, C, G and T are the internationally recognised one-letter codes for the four deoxyribonucleotide bases.

Preferred nucleic acid fragments of the invention comprise the sequence;

$(GCC)_m$ CCG GAC ACC AGG CCG $(GCC)_n$

An alternative sequence encoding PDTRP epitopes is to be found in the β-glucuronidase gene of E. coli and contains the coding sequence

CCG GAT ACC CGT CCG

for PDTRP. Fragments containing this sequence and optionally further comprising flanking sequences which may or may not be related to the β-glucuronidase sequence (other than the complete β-glucuronidase gene and fragments specifically disclosed in Proc. Natl. Acad. Sci. USA. 8447-8451 (1986)) form further aspects of the present

invention.

The invention further provides nucleic acid fragments encoding PDTRP epitopes which are immunologically analogous to peptides of formula (I) as hereinbefore defined.

These fragments will be useful both for generating PDTRP epitopes as described below and as probes for detecting DNA encoding PDTRP epitopes. For such purposes it may be convenient to attach detectable labels (such as radioisotopes, fluorescent or enzyme labels) to the fragment or to bind the fragment to a solid support. All of these may be achieved by conventional methods (Matthews et al, Anal. Biochem. 169, 1-25 (1988)).

The nucleic acid fragments of the invention may be produced de novo by conventional nucleic acid synthesis techniques or obtained from human epithelial cells by conventional methods (Huynh et al, "DNA Cloning: A Practical Approach" ed Glover, D.M. (IRL, Oxford), 1985, Vol 1 pp49-48).

In further aspects the present invention provides cloning and expression vectors comprising nucleic acid fragments as hereinbefore defined (the expression vectors having the fragment in open reading frame relation to initiation and termination signals and being associated with appropriate regulatory and promoter sequences so as to form an expressible sequence) and host cells transformed with such cloning or expression vectors.

The nucleic acid fragments of the present

invention may also be used in active immunisation techniques. For use in such methods the fragment, which codes for a PDTRP epitope may be extended at either or both the 5′ and 3′ ends with further coding or non-coding nucleic acid sequences including regulatory and promoter sequences, marker sequences and splicing or ligating sites. In addition to encoding the PDTRP epitope, coding sequences may code for other portions of the mucin protein chain or for other polypeptide chains. The fragment according to the invention, together with any necessary or desirable flanking sequences is inserted, in an appropriate open reading frame register, into a suitable vector such as a plasmid or a viral genome (for instance vaccina virus genome) and is then expressed as a polypeptide product by conventional techniques. In one aspect the polypeptide product including the PDTRP epitope may be produced by culturing appropriate cells transformed with a vector, harvested and used as an immunogen to induce active immunity against the mucin core protein (Tartaglia et al., Tibtech, 6, 43 (1988)). In another aspect the vector, particularly in the form of a virus, may be directly inoculated into a human or animal to be immunised. The vector then directs expression of the PDTRP epitope in vivo and this in turn serves as an immunogen to induce active immunity against the mucin core protein. The DNA sequence may, in one embodiment, be inserted into another gene and expressed as a fusion protein. For example, a nucleotide sequence encoding a PDTRP epitope flanked

by other amino acid sequences (which are the same as or different from the mucin tandem repeat sequence) is expressed in a genetically engineered poliovirus by inserting the oligonucleotide into the gene encoding capsid protein of the Sabin type 1 vaccine strain of poliovirus, as has been done for other protein epitopes (Evans et al., 1989, Nature 339, 385-388). Such an engineered virus may readily be tested for its ability to induce tumour regression in humans or to delay relapse, and ultimately to prevent tumour development.

The invention therefore provides probes, vectors and transformed cells for use in the methods of treatment of the human or animal body by surgery or therapy and in diagnostic methods practised on the human or animal body and for use in the preparation of medicaments for use in such methods. The invention also provides methods for treatment of the human or animal body by surgery or therapy and diagnostic methods practised in vivo as well as ex vivo and in vitro which comprise administering such fragments, probes, vectors or transformed cells in effective non-toxic amount to a human or other mammal in need thereof as well as diagnostic methods practised ex vivo or in vitro comprising contacting a body fluid or tissue sample with the fragments or probes.

The invention further provides a process for producing the PDTRP epitopes comprising culturing such host cells transformed with an expression vector and recovering expressed polypeptides.

In a further aspect the invention provides antibodies other than SM3 and HMFG-2 capable of recognising a PDTRP epitope as hereinbefore defined and preferably having reduced or no reaction with fully processed human PEM glycoprotein as produced by the normal lactating human mammary gland.

As used herein the term "antibody" is intended to include polyclonal and monoclonal antibodies and fragments of antibodies bearing antigen binding sites such as the $F(ab')_2$ fragments as well as such antibodies or fragments thereof which have been modified chemically or genetically in order to vary the amino acid residue sequence of one or more polypeptide chains, to change the species specific and/or isotype specific regions and/or to combine polypeptide chains from different sources. Such antibodies may be obtained by conventional methods (Williams, Tibtech, 6, 36, (1988)) and are useful in diagnostic and therapeutic applications, such as passive immunisation.

Antibodies according to the present invention react with PDTRP epitopes as hereinbefore defined; they also react with HPEM core protein, especially as expressed by colon, lung, ovary and particularly breast carcinomas, but have reduced or no reaction with corresponding fully processed HPEM. In a particular aspect the antibodies react with HPEM core protein but not with fully processed HPEM glycoprotein as produced by the normal lactating human mammary gland. Antibodies according to the present invention preferably have no significant reaction with the

mucin glycoproteins produced by mammary epithelial tissues lactating or pregnant individuals but react with the mucin proteins expressed by mammary epithelial adenocarcinoma cells. These antibodies show a much reduced reaction with benign breast tumours and are therefore useful in diagnosis and localisation of breast cancer as well as in therapeutic methods.

The antibodies may be used for other purposes including screening cell cultures for the polypeptide expression product of the human mammary epithelial mucin gene, or fragments thereof, particularly the nascent expression product. In this case the antibodies may conveniently be polyclonal or monoclonal antibodies.

The invention further provides antibodies linked by known techniques to therapeutically or diagnostically effective moieties or to linking moieties. The linking moieties may be known linking agents for binding therapeutically or diagnostically effective moieties to the antibodies, for instance the linking moiety may be a specific binding partner for a specific binding agent, the therapeutically or diagnosically effective moiety being bound to the specific binding agent. Following separate administration of the antibody:linking moiety (specific binding partner) adduct and of the therapeutically or diagnostically effective moiety: specific binding agent adduct, binding of the specific binding agent to the specific binding partner results in the therapeutically or diagnostically effective moiety being bound to the

antibody. A preferred example of a specific binding agent/ specific binding partner pair is biotin and avidin or streptavidin in which biotin may be the binding agent or the binding partner.

For therapeutic use of the antibodies the therapeutically effective moieties are lethal agents to be delivered to cancerous breast or other tissue in order to incapacitate or kill transformed cells. Lethal agents include toxins, radioisotopes and "direct killing agents" such as components of complement as well as cytotoxic or other drugs or enzymes which activate cytotoxic drugs at the tumour site.

For diagnostic use the diagnostically effective moieties may be moieties such as solid supports and detectable labels such as enzyme labels, chromophores, fluorophores and radioisotopes and other directly or indirectly detectable labels. Preferably monoclonal antibodies are used in diagnosis.

The invention further provides a diagnostic test process or assay method comprising contacting a sample suspected to contain abnormal human mucin glycoproteins with an antibody as defined above. Such methods include tumour localisation involving administration to the patient of the antibody bearing detectable label or administration of an antibody and, separately, simultaneously or sequentially in either order, administering a labelling entity capable of selectively binding the antibody or fragment thereof.

Particular uses of the antibodies include diagnostic assays for detecting and/or assessing the severity of breast, colon, ovary and/or lung cancers.

Diagnostic test kits are provided for use in diagnostic process and assays and comprise antibody and, optionally, suitable labels and other reagents such as buffers, reagents for labelling the antibody and/or for detecting labelled antibody and, especially for use in competitive assays, standard sera.

The invention further provides PDTRP epitopes and antibodies as hereinbefore defined for use in methods of surgery, therapy or diagnosis practised on the human or animal body or for use in the production of medicaments for use in such methods. The invention also provides a method of treatment or diagnosis which comprises administering an effective non-toxic amount of a polypeptide or antibody as hereinbefore described to a human or animal in need thereof. Dosage rates will depend on the age, weight, size, sex and general health of the patient and will be sufficient to achieve appropriate levels of circulating antibodies without undesirable or intolerable side effects.

Administration may be by the oral or parenteral route, for instance by sub-cubtaneous, intravenous or intramuscular injection. Typically a PDTRP epitope is administered to a human adult in an amount up to 1 g per dose, preferably 1 to 200 mg per dose, by either the oral or the parenteral route.

The invention therefore also provides

pharmaceutical compositions comprising a PDTRP epitope of the invention and a suitable carrier, diluent and/or adjuvant, which will, of course, be determined by the route of administration and selected for compatibility with the PDTRP epitope. Suitable carriers, diluents and adjuvants include Freunds incomplete adjuvant (IFA), aluminium hydroxide, saponin, DEAE-dextran, muramyl dipeptide, mineral oils, neutral oils such as miglyol, vegetable oils such as arachis oil, "Iscoms", liposomes, Pluronic (trade mark), polyols or the Ribi adjuvant system (GB-A-2189141). For injection the compositions may be in the form of aqueous sterile, non-pyrogenic solutions optionally containing buffers, antioxidants, biocides such as antibacterials and antifungals and agents to adjust the tonicity. Alternatively the compositions may be provided as dry powders, optionally containing excipients such as those mentioned above, for reconstitution by addition of water for injection.

Antibodies, whether polyclonal or monoclonal, against PDTRP epitope may be used in therapeutic methods for treating humans and animals, for instance in the form of vaccines for passive immunisation. Preferably they are formulated as pharmaceutical compositions such as are described above.

Especially in therapeutic applications it may be appropriate to modify the antibody by coupling the Fab, or complementarity-determining region thereof, to the Fc, or whole framework region, of antibodies derived from the

species to be treated (e.g. such that the Fab region of mouse monoclonal antibodies may be administered with a human Fc region to reduce immune response by a human patient) or in order to vary the isotype of the antibody (see EP-A-0 239 400).

The present invention also provides a process for producing polyclonal antibodies against a PDTRP epitope comprising inoculating a host animal with transfected cells capable of expressing the PDTRP epitope and recovering antibodies from the serum or other body fluid of the animal.

The present invention further provides a process for producing a cell capable of secreting antibodies against a PDTRP epitope comprising inoculating a host animal with PDTRP epitope or with transfected cells capable of expressing the PDTRP epitope and removing antibody-secreting cells from the animal. The cells may be capable of secreting antibodies either in vitro, for instance under appropriate cell culture conditions and/or in vivo, for instance by growth in ascites.

As used herein the term "animal" refers to any animal capable of producing antibodies for instance mammals such as laboratory rodents, and birds and includes humans.

Transfected cells are cells containing a DNA expression vector containing an expressible sequence encoding the PDTRP epitope. The coding sequence may be natural or synthetic and the encoded polypeptide may be expressed directly or as a fusion protein.

For instance, DNA coding for the PDTRP epitope is identified and obtained from a natural source or produced by partial or total synthesis. Techniques for obtaining natural or artificial DNA encoding a particular polypeptide are well known in the field of genetic engineering and are described in Maniatis, T., Fritsch, E and Sambrook J. (1982) Molecular Cloning, a Laboratory Manual, published by Cold Spring Harbor Laboratory Press NY. The vector is constructed by conventional methods to ensure expression in the cells to be transfected and may be provided with marker sequences enabling identification of the DNA and transfectant cells, (Maniatis loc. cit.). Selective markers which are particularly suitable in the present invention include drug resistance markers such as the thymidine kinase (tK) gene for tK^- cells. (Szybalska, E and Szybalski W. (1962) Proc. Natl. Acad Sci., USA, 42, 2026).

Suitably, promoter sequences may be inserted in the vector to increase expression of the PDTRP epitope in the transfectant cells (Subramani, S., Mullingan, R.C. and Berg, P. (1981) Mol. Cell. Biol. 1, 854-864).

The cells to be transfected, i.e. the cells into which the vector is to be inserted, (hereafter "donor cells") are obtained from donor animals, such as rodents, especially rats and mice, and primates, for instance monkeys. The donor cells must be capable, once transfected, of expressing the PDTRP epitope, optionally with post translational processing of the PDTRP epitope.

Preferably the donor cells and the optionally modified DNA are chosen such that the PDTRP epitope appears extracellularly or at the cell surface although it is also possible to apply the present invention to polypeptides which are only expressed intracellularly.

It is preferred, especially for intracellularly expressed polypeptides, that the donor cells are of a type not normally found in the peritoneum or circulating in the bloodstream of the host animals.

Preferably the cells to be transfected are derived from an animal (hereafter the "donor animal") syngeneic with the animal (hereafter the "host animal") which is inoculated with the transfected cells. Without wishing to be bound by a particular theory, it is believed that by using cells from a syngeneic donor animal, the differences between the transfected cells and the host animal's own cells that can be detected and acted upon by the host animal's immune system are minimised. Thus any immune response is more likely to be directed against the foreign polypeptide produced by the transfected cells by expression of the DNA sequence.

To maximise the directing effect on the immune system of the host animal it is desirable that the host and donor animals are substantially genetically identical and usually they will be from a well established inbred strain such as are well known in the art. Use of monozygotic donor and host animals whether or not from an inbred strain, or even cells from the host animal itself, is

contemplated.

Transfection of donor cells is achieved by conventional methods such as using calcium phosphate (Austin, P., Trowsdale, J., Rudd, C., Bodmer, W., Feldmann, M. and Lamb, J. (1985) Nature, 313, 1-4) or by the electroporation technique (Neumann, E., Schaefer- Ridder, M., Wang, Y. and Hofschneider, P.H. (1982) EMBO J, 1, 841-845) or using retroviral vectors (Biotechniques, 6, 608-614 (1988)).

After transfection it may be convenient to screen the cells, or clones or sub-clones thereof, to select those capable of expressing the PDTRP epitope. With cell surface-expressed gene products the use of fluorescent antibodies and a fluorescence activated cell sorter is particularly convenient (Austin loc. cit.) but other techniques are available for screening and selection.

Transfectant cells, optionally after selection, may be treated to enhance the expression of the gene product. Various known enhancement methods can be used including treatment with sodium butyrate (Gorman, C.M. and Howard, B.H. (1983) Nucleic Acids Research, 11, 7631-7648) or interferon (Balkwill, F.R., Stevens, M.H., Griffin, D.B., Thomas, J.A. and Bodmer, J.G. (1987) Eur. J. Cancer Clin. Oncol., 23, 101-106). The enhancement method will be selected as appropriate to the type of donor cells.

The host animal or animals are inoculated with PDTRP epitopes or transfected cells, preferably by intraperitoneal or intravenous injection of transfected

cells in a suitable medium. Preferably the inoculation is repeated at intervals and the animals' antibody titre is monitored to ascertain that the desired immune response occurs. The route chosen for administration and the inoculation regime are preferably selected in order to favour production of the desired class of antibodies, eg production of cells secreting IgM antibodies is favoured by a short interval between inoculation and recovery of antibody-secreting cells, whereas production of cells secreting IgG antibodies is favoured by a long interval between repeated inoculations and eventual recovery of the antibody-secreting cells.

On completion of the inoculation regime antibody -secreting cells eg, spleen cells, are taken from the inoculated host animal(s) and maintained in an appropriate medium. Conventional methods may be used for obtaining and maintaining the host cells. Antibody producing cells may also be obtained from the body fluids or tissues of host animals having carcinoma, particularly lymphocytes from excised lymph nodes of human carcinoma, especially breast or ovarian cancer, patients. (J. Immunol. Methods, 105, 263-273 (1987)).

In another aspect the invention provides a process for producing a cell capable of secreting antibodies against a PDTRP epitope comprising removing antibody-secreting cells from a host animal having carcinoma or inoculated with a PDTRP epitope or with transfected cells capable of expressing the PDTRP epitope.

In a further aspect the invention provides a cell capable of secreting antibodies against a PDTRP epitope which cell has been removed from a host animal having carcinoma or inoculated with a PDTRP epitope or with transfected cells capable of expressing the PDTRP epitope.

It may be convenient at this stage to screen the population of antibody-secreting cells obtained from the host animal in order to select those secreting antibodies against the PDTRP epitope.

The cells are useful particularly for immortalisation in order to form an immortal cell line capable of secreting antibodies (hereafter monoclonal antibodies or MAb's) against the PDTRP epitope. Immortalisation may be by any of the known methods but fusion with an immortal cell to form a hybridoma is perhaps the most convenient method.

Accordingly, the invention also provides a process for producing an immortal cell capable of secreting antibody against a PDTRP epitope comprising immortalising an antibody-secreting animal cell, the antibody-secreting animal cell having been taken from a host animal inoculated with PDTRP epitope or transfected cells capable of expressing the PDTRP epitope.

Immortal cells (which cells are capable of repeated cell division under appropriate culture conditions) suitable for fusion are well known. The immortal cell may be of the same inbred strain as the antibody-secreting cells and usually they will be of the

same species as the antibody-secreting cells from the host animal but this is not essential. Particularly convenient immortal cells are myeloma cells. Fusion is achieved by the method of Kohler and Milstein, Nature, 256, 495-497 (1975) and by variations thereon as described in the literature.

After fusion it may be convenient to screen the resulting population of hybrid cells to select those secreting antibodies against the PDTRP epitope.

The present invention also provides an immortalised cell, preferably a hybridoma, capable of secreting antibodies against a PDTRP epitope, which cell comprises an immortalised, antibody-producing cell capable of secreting antibodies against the PDTRP epitope taken from a host animal inoculated with PDTRP epitope or transfected cells capable of expressing the PDTRP epitope, or is a descendant of such a cell.

Immortal cells (e.g. hybridoma cells) according to the present invention may be used for producing MAb's against the PDTRP epitope.

Accordingly, the present invention provides a process for producing monoclonal antibodies against a PDTRP epitope comprising culturing an immortalised cell as hereinbefore defined capable of secreting antibodies against a PDTRP epitope, or a descendant of such a cell, and recovering the monoclonal antibodies.

The invention further provides monoclonal antibodies against a PDTRP epitope which antibodies are

obtained from growth in tissue culture or ascites of an immortal hybrid cell as hereinbefore defined.

Antibodies produced by the immortalised cells as described above may be screened by conventional techniques to identify those cells producing MAb's against a PDTRP epitope. Such cells may be multiplied and cultured to produce MAb's according to the invention in large quantities. Suitable culturing techniques and ascites growth conditions are well known and are chosen as appropriate to the particular immortalised cell line involved and the MAb's to be produced.

The present invention further provides a fusion protein comprising a PDTRP epitope as hereinbefore defined and a carrier polypeptide or protein bound to the C-terminal or the N-terminal of the PDTRP epitope.

The carrier polypeptide or protein may be any suitable polypeptide or protein. Convenient carrier proteins are thyroglobulin and bovine serum albumin.

The fusion proteins of the invention may be produced by conventional techniques of polypeptide synthesis and protein chemistry from individual amino acid precursors or by coupling smaller polypeptides. Alternatively the fusion proteins may be produced by recombinant DNA techniques involving expression of a gene encoding the fusion protein. These methods are all well known to those skilled in the art.

The invention further provides fusion proteins as hereinbefore defined for use in a method of treatment of

the human or animal body by surgery or therapy or a diagnostic method practised on the human or animal body; the use of fusion proteins in the manufacture of a medicament for use in the therapeutic, surgical or diagnostic treatment of the human or animal body; a pharmaceutical composition comprising the fusion protein and a pharmaceutically acceptable diluent or carrier therefor and, optionally, conventional accessory ingredients, and a method for treating the human or animal body comprising administering an effective, non-toxic amount of a fusion protein to a human or animal in need thereof.

The fusion proteins are particularly useful in the treatment or prophylaxis of cancers such as breast cancer by increasing the patient's immunity to the tumour cells and in the production of polyclonal and monoclonal antibodies for use in diagnosis and therapeutic treatment of cancer.

The invention will now be illustrated by reference to the Figures of the accompanying drawings which

Fig.1 is a graph of the O.D. at 450 nm of a series of octapeptide/SM3 reactions in accordance with Example 1.

Fig. 2 shows the design of overlapping peptide octamers. The starting sequence consisted of a twenty amino acid tandem repeat with four amino acids of adjacent repeats on the amino and carboxyl ends. Overlapping octamers are

indicated [___] and [A] refers to the alanine used as the spacer on the pins. Note, octamers 1 and 21 are identical. The one letter code for the amino acids is used.

Fig. 3 is a graph of the binding of the monoclonal antibody SM-3 to octamer peptides. Monoclonal antibody (50 μg/ml, 100 μl per pin) was reacted with octamers 1-21 (defined in Fig. 2) and the binding visualized using an ELISA assay. The boxed sequences in the insert represent the epitope recognized by SM-3.

Fig. 4 is a series of graphs showing the binding of monoclonal antibodies HMFG-2, LICR-LON-M8 and onc-M15 to octamer peptides. Monoclonal antibody (50 μg/ml, 100 μl per pin for HMFG-2 and onc-M15; 1:500 dilution of LICR-LON-M8 ascites fluid) were reacted with the octamers 1-21 (defined in Fig. 2) and binding visualized using an ELISA. The sequence in brackets indicates the epitope.

Fig. 5 is a diagrammatic representation of epitopes recognised by monoclonal antibodies reactive with the PEM mucin core protein. Potential glycosylation sites are marked with arrows.

Fig. 6 is a series of graphs showing the binding of HMFG-2 to lymphoblastoid and breast carcinoma cell lines detected by fluorescence activated cell sorter. Cells in suspension were incubated with HMFG-2 (-) or TAL-14. (----) an anti

HLA-DR monoclonal antibody, and the binding detected using a fluorescence labelled rabbit anti-mouse. A, lymphoblastoid cell line PGF which expresses HLA-DR2, the β chain contains the sequence Asp-Thr-Arg-Pro (DTRP); B, lymphoblastoid cell line Mann which expresses HLA-DR7, the β chain contains the sequence Asp-Thr-Gln-Pro (DTQP); C, breast carcinoma cell line T47D which expresses the PEM mucin.

Fig. 7 is a graph of the binding of anti-peptide antiserum to octamer peptides. The anti-serum diluted 1:200 was reacted with the peptides as described in Fig. 2. The boxed sequence in the insert represents the dominant epitope recognized by this anti-serum.

Fig. 8 is a pair micrographs showing staining of breast cancer cell lines with antipeptide antiserum.

The invention will further be illustrated by the following Examples which are not intended to limit the invention in any way.

EXAMPLE 1

Determination of the epitope recognised by SM-3

We have previously shown that the monoclonal antibody SM-3 recognises an epitope within the 20 amino acid tandem repeat found in the PEM core protein (Gendler et al., J. Biol. Chem., 263, 12820-12823 (1988)). Using an epitope scanning kit purchased from Cambridge Research Biochemical (Cambridge, England) we have further defined the epitope recognised by SM-3 and have shown that the minimum epitope consists of

Proline - Aspartic Acid - Threonine - Arginine - Proline

The method consists of synthesizing small, overlapping peptides on the end of the polypropylene pins, 96 pins being held in a plastic support in the format and spacing of a microtitre plate. The tips of the pins have been chemically derivatized to produce functional groups to which amino acids may be coupled and the peptides are synthesised on to the tips of the pins by the stepwise elongation of the peptide chain with the addition of one amino acid per day.

Once the desired length of peptide has been

synthesised, the pins can be assayed for their reactivity with antibody in a standard ELISA type assay, dispensing the reagents into 96 well microtitre plates in which the pins are then placed.

We chose to synthesise overlapping octamers moving along the tandem repeat one amino acid with each octamer synthesized, and covering all possible octamers so that the first and last peptides are identical. Thus to design the octamers we used one tandem repeat and 8 amino acids of the next repeat i.e, VTSAPDRTPAPGSTAPPAHGVTSAPDTR. Thus the first octamer was VTSAPDTR, the second TSAPDTRP, the third SAPDTRPA etc. Octamers 2 to 5 are therefore "PDTRP" epitopes as hereinbefore defined. Octamers 1 and 6 to 21, each of which contains only a part or none of the PDTRP sequence and which have little or no reaction with SM-3 are not PDTRP epitopes. The results obtained with such a set of octamers are shown in Fig. 1. indicating that the epitope for SM-3 is PDTRP.

EXAMPLE 2

Many of the monoclonal antibodies which show epithelial specificity and recognise tumour associated antigens react with epitopes found on large molecular weight mucin molecules. Until recently, most of the biochemical characterization of mucins was performed using the large mucins produced, for example, in the gastrointestinal tract, the lung and the cervix (Hounsell and Feizi, 1982; Marshall and Allen, 1978; Slayter et al., 1984; Carlstedt et al.,1983). However, other glands including the mammary, sweat and salivary glands produce mucins and one, which we have termed polymorphic epithelial mucin (PEM), appears to be the target antigen for many monoclonal antibodies recognising epithelial or carcinoma associated determinants (Burchell and Taylor-Papadimitriou, 1989).

The PEM component is of particular interest because, although it is produced by several normal epithelial tissues, and abundantly by the lactating mammary gland, it is also expressed by breast and other carcinomas (Taylor-Papadimitriou et al., 1986; Girling et al., 1989). Moreover, the mucin appears to be glycosylated differently by cancer cells, with the result that novel carbohydrate epitopes appear (Kjeldsen et al., 1988), and core protein epitopes which are masked in the normal mucin are exposed on the cancer-associated component. One such core protein epitope is detected by the antibody SM-3, which was developed using the chemically deglycosylated normal PEM mucin (Burchell et al., 1987). This antibody reacts with more than 90% of breast cancers and with the deglycosylated mucin, but shows little or no reaction with the normal breast or the native PEM mucin purified

from milk (Burchell et al., 1987; Girling et al., 1989). The specificity of the SM-3 antibody makes it a potentially useful tool in the diagnosis and treatment of breast cancer, and it was therefore of great interest to identify the SM-3 epitope.

Recently the amino acid sequence of the immunogenic domain of the PEM mucin (Gendler et al., 1988), was deduced from the nucleotide sequence of cDNA clones coding for this region (Gendler et al., 1987). The domain is made up of repeat units, 20 amino acids in length, and the number of repeats varies from individual to individual, giving rise to the polymorphism seen at the DNA and protein level (Swallow et al., 1987). The reading frame of the tandem repeat was definitively established by showing binding of three core protein reactive antibodies (HMFG-1, HMFG-2 and SM-3) to a synthetic peptide with the corresponding sequence (Gendler et al., 1988). The sequence is in agreement with what might be expected for a mucin and contains serines and threonines (possible O-glycosylation sites) separated by proline rich stretches, suggesting a linear structure. Using overlapping octamers covering the sequence of the twenty amino acid tandem repeat we have now been able to identify a continuous amino acid sequence, Pro-Asp-Thr-Arg-Pro as the core of the SM-3 epitope. Interestingly other core protein epitopes which are found on both the normal and cancer associated mucin, and are recognized by other antibodies, also map to this region, but none contain the proline at the amino end of the SM-3 epitope. The data are consistent with the idea that the exposure of the SM-3 epitope on the cancer associated mucin can be attributed to aberrant processing resulting in shorter carbohydrate side chains. Preliminary studies indicate that it is possible to use

a peptide corresponding to part of the tandem repeat sequence to develop antibodies to the PEM mucin which are reactive with breast cancers.

Materials and Methods

Peptides: The peptides were synthesized by solid phase techniques on a 430A peptide synthesizer (Applied Biosystems), using commercially available t-boc protected amino acids and resins as previously described (Townsend et al., 1986).

Synthesis of overlapping peptide octamers: Overlapping peptide octamers were synthesized onto the end of polypropylene pins using an Epitope Scanning Kit (Cambridge Research Biochemicals, Cambridge, U.K.) which is based on a method originally published by Geysen et al. (1984). The pins are supplied with an alanine attached to the ends which acts as a spacer, thus the amino acid at the carboxyl end of the synthesized peptide is joined by a peptide bond to alanine.

Antibodies: The development and characterization of the monoclonal antibodies HMFG-1, HMFG-2 and SM-3 (Taylor-Papadimitriou et al., 1981; Burchell et al., 1983, 1987; Girling et al., 1989), LICR-LON-M8 (Foster et al., 1982), onc-M15, onc-M23, onc-M27 (Linsley et al., 1988) has been described elsewhere. The monoclonal antibody TAL 14.1 is directed to HLA-DR.

Anti-peptide mouse antiserum was produced by immunizing intraperitoneally with 25 µg of PDTR peptide 1-24 (see Table 2)

coupled to thyroglobulin followed by three injections of 25 µg PDTR 1-14 conjugated to bovine serum albumin.

ELISA: Monoclonal antibodies (50 µg/ml or 1:500 dilution of ascites fluid, 100 µl per pin) and polyclonal serum (1:200) were tested for their reactivity with the octamers in an ELISA assay as previously described (Burchell et al., 1983), using peroxidase conjugated rabbit antimouse antibody (DAKO, Denmark). The octamers were synthesized onto pins conferring to the format of a 96 well microtitre plate and so for the assay pins were inserted into microtitre dishes containing the antibodies or reagents.

Binding of HMFG-2 to cell lines: Cell pellets containing 10^6 cells were washed with RPMI containing 10% foetal calf serum and suspended in 1 ml of hybridoma tissue culture supernatant containing 0.02% sodium azide and incubated at 4°C for 1 hour. After two washes with RPMI containing 10% FCS, 0.02% azide (wash buffer) the cell pellets were suspended in 1 ml of fluorescence conjugated rabbit anti mouse (DAKO, Denmark) which had been diluted 1:20 in wash buffer and filtered prior to use. After 30 minutes at 4°C the pellets were washed once with wash buffer (20 ml) and once with PBSA (20 ml) and then suspended in 1 ml of PBSA. Binding of fluorescence conjugated rabbit anti-mouse was then analyzed in a Spectra-Physics fluorescence cell sorter.

Results

Identification of the epitope recognised by antibody SM-3

The sequence of the tandem repeat unit of the PEM mucin is shown in Figure 2. Arbitrarily the sequence has been written as starting with PDTR since this was the beginning of the first tandem repeat in one of the cDNA clones (Gendler et al., 1987). The antibody SM-3 has been shown previously to react with the PDTR peptide (1-24) which corresponds to one tandem repeat and four amino acids of the next repeat (see Table 1). To further define the epitope recognized by the SM-3 antibody, a series of overlapping peptide octamers, covering a single repeat and eight amino acids of adjacent repeats (Fig. 2) were synthesized onto the tips of polypropylene pins (see Materials and Methods). Figure 3 shows the results obtained when the peptides were reacted with SM-3 in an ELISA. It can be seen that all the peptides reactive with the SM-3 antibody contain the sequence Pro-Asp-Thr-Arg-Pro which appears therefore to represent the minimum epitope.

Immunogenicity of the domain containing the SM-3 epitope

Monoclonal antibodies reactive with the PEM mucin have been developed using a variety of immunogens, including human milk fat globule extracts, which contain the mucin, (Taylor-Papadimitriou et al., 1981; Hilkens et al., 1984; Foster et al., 1982; Ceriani et al., 1983) and carcinoma cells (Bramwell et al., 1983; Price et al., 1985). Some were obtained from mice injected with both the mucin and cultured

Table 1

Binding of antibodies to peptides based on the sequence of the PEM mucin of tandem repeat

Antibody	Immunogen	Reaction with			Ref
		PDTR peptide[1] (a.a. 1-24)	PDTR peptide[2] (1-14)	VTSA peptide[3] (17-20, 1-10)	
HMFG-2	HMFG[4], milk epithelial cells	+	+	+	Burchell et al., 1983
SM-3	Deglycosylated mucin	+	+	+	Burchell et al., 1987
LICR-LON-M8	HMFG	+	+	+	Foster et al., 1982
onc-M15	MCF-7, HMFG	+	+	+	Linsley et al., 1987
onc-M23	MCF-7, HMFG	+	+	+	"
onc-M27	W5-6, HMFG	+	+	+	"

[1]PDTR peptide (1-24) PDTRPAPGSTAPPAHGVTSAPDTR

[2]PDTR peptide (1-14) PDTRPAPGSTAPPAC*

[3]VTSA peptide (17-20, 1-10) VTSAPDTRPAPGST

[4]HMFG Human Milk Fat Globule

*Cysteine included to allow for coupling

cells (Taylor-Papadimitriou et al., 1981; Burchell et al., 1983; see Table 1). We have tested some of these antibodies for their reaction with the peptide coding for the tandem repeat, and found that many show a strong positive reaction with the PDTR peptide (amino acids 1-24, Table 1). These include the antibodies listed in Table 1 as well as antibodies NCRC11 (Price et al., 1985) and DF3 (Sekine et al., 1985). To identify the domain reactive with some of these antibodies in more detail, their reaction with smaller peptides was examined. As shown in Table 1 all of the antibodies tested reacted with two smaller peptides, both of which contained the first 10 amino acids of the PDTR peptide. This region also contains the SM-3 epitope, suggesting that this part of the tandem repeat contains epitopes which are immunodominant in the mouse. The specific epitopes reactive with three of the antibodies shown in Table 1, namely HMFG-2, LICR-LON-M8 and onc-15 were identified in more detail using the overlapping octamers illustrated in Figure 2 and the results are shown in Fig. 4. Antibodies HMFG-2 and LICR-LON-M8 show a remarkably similar pattern of reaction with the octamers and in both cases, the minimum epitope appears to be Asp-Thr-Arg (DTR), while that recognised by onc-15 is Thr-Arg-Pro-Ala (TRPA).

The results of the detailed epitope mapping indicate that the single threonine and flanking amino acids between potential glycosylation sites constitute a highly immunogenic domain of the PEM tandem repeat. The overlap between the different epitopes is shown in Figure 5.

Role of adjacent amino acids

When the PIR protein data base was searched for the epitopes shown in Figure 5 a large number of proteins ranging from bacterial proteins to HLA-DR β chain were shown to contain the sequence Asp-Thr-Arg. However the antibodies HMFG-2 and M8 have been reported to show a restricted reaction with the polymorphic epithelial mucin. To investigate whether one of these antibodies would bind to proteins containing Asp-Thr-Arg with different flanking amino acids, we studied the binding of HMFG-2 to a native protein known to contain the minimum epitope. The β chain of HLA DR2 contains the sequence Asp-Thr-Arg-Pro while the β chain of HLA DR7 contains the sequence Asp-Thr-Gln-Pro (Young et al., 1987), however the adjacent amino acids show no homology with the PEM mucin. The binding of HMFG-2 to lymphoblastoid cell lines expressing HLA DR2 (PGF cells) and HLA DR7 (MANN cells), was analyzed and compared with the binding of HMFG-2 to a breast carcinoma cell line expressing the PEM mucin (T47D). Figure 6 shows that as expected HMFG-2 did not bind to MANN cells expressing DR7, nor did it bind to cells expressing DR2 which contains the sequence Asp-Thr-Arg. However, HMFG-2 did show strong binding to the breast carcinoma cell line T47D. Thus the minimum epitope for HMFG-2 consists of only three amino acids in the context of a conformationally free peptide while in the native protein, which has conformational restraints, the adjacent amino acids have an important role in determining the antigenic site.

Peptides as immunogens for the development of breast cancer reactive antibodies

With the exception of SM-3, the monoclonal antibodies discussed above were developed using glycosylated mucin either isolated from milk or associated with cells, and in these cases the carbohydrate side chains would play a major role in determining which of the core protein epitopes were exposed. We have now examined the response of a mouse injected with the PDTR-peptide (1-24) and the PDTR peptide (1-14) coupled to a carrier protein (see Materials and Methods), by following the reaction of the antiserum with the overlapping octamers. Figure 7 shows that although the overall binding was higher than with the monoclonal antibodies, there appeared to be a peak of binding at peptides 4 and 5. This corresponds to antibodies being present in the serum that recognize epitopes in the sequence P-D-T-R-P-A-P. The antiserum showed a positive reaction with primary breast carcinomas and cell lines (see Figure 8), indicating that some of the antibodies present reacted with epitopes exposed on the cancer associated mucin. This initial experiment suggests that peptides based on the tandem repeat sequence can be used as immunogens to develop antibodies reactive with the cancer associated PEM mucin.

Discussion

The core protein of the polymorphic epithelial mucin is made up of a number of tandem repeats, each of which contains five potential O-linked glycosylation sites (Gendler et al., 1988). These tandem

repeats contain the epitopes recognized by many of the monoclonal antibodies directed to this mucin. Using overlapping octamers the epitope recognized by SM-3, has been defined and has been shown to include the single threonine at position 3. Furthermore, epitopes reactive with other monoclonal antibodies have also been shown to map to this region. It therefore appears that, at least in mice, this sequence of the tandem repeat is highly immunogenic.

Although the epitopes recognized by HMFG-2, LICR-LON-M8, onc 15 and SM-3 are overlapping, the reactivity of these antibodies with the mucin is quite different. All react with the mucin expressed by breast carcinoma cells but the HMFG-2, onc 15 and M8 determinants are also expressed on the normal mucin (Burchell et al., 1983; Linsley et al., 1988; Foster et al., 1982), whereas the SM-3 determinant is not (Burchell et al., 1987; Girling et al., 1989). The fact that the SM-3 monoclonal antibody was raised to the chemically deglycosylated normal milk mucin suggest that carbohydrate may play a role in masking this epitope on the PEM produced by normal cells. The combined epitopes recognized by the three antibodies which react with the normal mucin cover the sequence, DTRPA, which is centrally placed between the flanking glycosylation sites ($T^{18'}$, $S^{19'}$, S^{9}, S^{10} in figure 2., see also figure 5). However, the SM-3 epitope PDTRP contains a proline at the amino end which is only one amino acid away from a glycosylation site (serine 19'). This suggests that in the normal mucin the proline on the amino end of the SM-3 epitope is masked by carbohydrate while in tumours, because of premature termination of the sugar chains on $serine^{19'}$ and $threonine^{18'}$ it is exposed. Recent results lend some support to this idea in that the normal mucin isolated from human milk

has been shown to contain extended and/or branched poly-N-acetyl-lactosamine side chains (Hanisch et al., 1989) while the oligosaccharide side chains of the mucin produced by the breast carcinoma cell line, BT20 consists of only 3 or 4 sugars (Hull et al., 1988).

Since all four of the epitopes mapped contain at their core a threonine, this amino acid in the reactive epitopes cannot be glycosylated. However it should be noted that fewer HMFG-2 specific epitopes are present on the normal mucin found in milk than on the cancer associated mucin (Burchell et al., 1983). This could mean that some of the threonines in the DTR epitope are in fact glycosylated on the normal mucin. An alternative explanation is that due to the heterogeneity in the length of the carbohydrate side chains, some of the DTR epitopes on the normal mucin are masked by long or branched polylactosamine chains, while others are not. Another feature of the HMFG-2 epitope is that the affinity for the antibody appears to vary, depending on the micro environment. Upon comparing the binding of the antibody to the normal mucin and to a cancer associated mucin, it is found that the antibody shows a higher affinity for the normal mucin, suggesting that the normal oligosaccharide side chains are required for maximum affinity. That oligosaccharide side chains may be important is also shown by the fact that both HMFG-2 and SM-3 bind more avidly to the partially deglycosylated mucin than to the mucin stripped of all carbohydrate (Burchell et al., 1987).

Although the oligosaccharide side chains may play a role in determining the conformation of the core protein epitopes recognized by antibodies HMFG-2 and SM-3, it is quite clear that there are

antigenic determinants in the core protein of the PEM mucin which are immunogenic and are expressed by breast cancers. Using a synthetic peptide corresponding to the tandem repeat sequence coupled to a protein carrier we have been able to induce antibodies in the mouse which show reactivity with breast carcinomas. Epitope mapping with the antiserum indicate that dominant epitopes are found in the sequence PDTRPAP between the glycosylation sites at $Thr^{18'}$, $Ser^{19'}$ and Ser 9, Thr 10 (Figure2), again confirming the immunogenic nature of this domain. These results suggest that it should now be possible to take a directed approach to the production of monoclonal antibodies with SM-3-like specificity which react with breast cancers and such work is now in progress. Furthermore, the identification of an epitope which is selectively exposed in carcinoma cells gives a focus for studying the aberrant processing of the mucin by tumour cells and suggests that it may be possible to use peptides or polypeptides containing the tandem repeat as synthetic vaccines for some human cancers.

References for Example 2

BRAMWELL, M.E., BHAVANANDAN, V.P., WISEMAN, G., and HARRIS, H., Structure and function of the Ca antigen. Br. J. Cancer 48, 177-183 (1983).

BURCHELL, J., DURBIN, H. and TAYLOR-PAPADIMITRIOU, J., Complexity of expression of antigenic determinants recognised by monoclonal antibodies HMFG-1 and HMFG-2 in normal and malignant human mammary epithelial cells. J. Immunol. 131, 508-513 (1983).

BURCHELL, J., GENDLER, S., TAYLOR-PAPADIMITRIOU, J., GIRLING, A., LEWIS, A., MILLIS, R., and LAMPORT, D., Development and characterization of breast cancer reactive monoclonal antibodies directed to the core protein of the human milk mucin. Cancer Res. 47, 5476-5482 (1987).

BURCHELL, J. and TAYLOR-PAPADIMITRIOU, J. Antibodies to human milk fat globule molecules. Cancer Invest. (1989) in press.

CARLSTEDT, I., LINDGREN, H., SHEEHAN, J., ULMSTEN, U., and WINGERUP, L., Isolation and characterization of human cervical-mucus glycoproteins. Biochem. J. 211, 13-22 (1983).

CERIANI, R.L., PETERSON, J.A., LEE, J.Y., MONCADA, R., and BLANK, E.W., Characterisation of cell surface antigens of human mammary

epithelial cells with monoclonal antibodies prepared against human milk fat globule. Som. Cell Genet. 9, 415-427 (1983).

FOSTER, C.S., EDWARDS, P.A.W., DINSDALE, E.A. and NEVILLE, A.M., Monoclonal antibodies to the human mammary gland. Virchows Arch. (Path. Anat.) 394, 279-293 (1982).

GENDLER, S.J., BURCHELL, J.M., DUHIG, T., LAMPORT, D., WHITE, R., PARKER, M., and TAYLOR-PAPADIMITRIOU, J., Cloning of partial cDNA encoding differentiation and tumor-associated mucin glycoproteins expressed by human mammary epithelium. Proc. natl. Acad. Sci. USA., 84, 6060-6064 (1987).

GENDLER, S., TAYLOR-PAPADIMITRIOU, J., DUHIG, T., ROTHBARD, J., and BURCHELL, J., A highly immunogenic region of a human polymorphic epithelial mucin expressed by carcinomas is made up of tandem repeats. J. Biol. Chem., 263, 12820-12823 (1988).

GEYSEN, H.M., MELOEN, R., and BARTELING, S., Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. Proc. Natl. Acad. Sci. USA 81, 3998-4002 (1984).

GIRLING, A., BARTKOVA, J., BURCHELL, J., GENDLER, S., GILLETT, C., and TAYLOR-PAPADIMITRIOU, J., A core protein epitope of the PEM mucin detected by the monoclonal antibody SM-3 is selectively exposed in a range of primary carcinomas. Int. J. Cancer (1989) (in press).

HANISCH, F.-G., UHLENBRUCK, G., PETER-KATALINIC, J., EGGE, H., DABROWSKI, J. and DABROWSKI, U., Structures of neutral O-linked polylactosaminoglycans on human skim milk mucins. A novel type of linearly extended poly-N-acetyl-lactosamine backbones with Galβ(1-4)GlcNAcβ(1-6) repeating units. J. Biol. Chem. (in press).

HILKENS, J., BUIJS, F., HILGERS, J., HAGEMAN, P.H., CALAFAT, J., SONNENBERG, A., and VAN DER VALK, M., Monoclonal antibodies against human milk fat globule membranes detecting differentiation antigens of the mammary gland and its tumours. Int. J. Cancer 34, 197-206 (1984).

HOUNSELL, E.F. and FEIZI, E., Gastrointestinal mucins. Med. Biol. 60, 227-236 (1982).

HULL, S.R., BRIGHT, A., CARRAWAY, K.L., ABE, M. and KUFE, D., Oligosaccharides of the DF3 antigen of the BT-20 human breast carcinoma cell line. J. Cell. Biochem., Suppl. 12E, 130, Abstract (1988).

KJELDSEN, T., CLAUSEN, H. HIROHASHI, S., OGAWA, T., IIJIMA, H., and HAKOMORI, S., Preparation and characterization of monoclonal antibodies directed to the tumour associated O-linked sialosyl-2-6α N-acetylgalactosaminyl (Sialosyl-Tn) epitope. Cancer Res. 48, 2214-2220 (1988).

LINSLEY, P., BROWN, J., MAGNANI, J. and HORN, D., Monoclonal antibodies reactive with mucin glycoproteins found in sera from breast

cancer patients. Cancer Res. 48, 2138-2148 (1988).

PRICE, M.R., EDWARDS, S., OWAINATI, A., BULLOCK, J.E., FERRY, B., ROBINS, R.A., and BALDWIN, R.W., Multiple epitopes on a human breast carcinoma-associated antigen. Int. J. Cancer 36, 567-574 (1985).

SEKINE, H., OHNO, T., and KUFE, D., Purification and characterization of a high molecular weight glycoprotein detectable in human milk and breast carcinomas. J. Immunol. 135, 3610-3615 (1985).

SLAYTER, H., LAMBLIN, G., LETREUT, A., GALABERT, C., HOUDRET, N., DEGAND, P., Complex structure of human bronchial mucus glycoprotein. Eur. J. Biochem. 142, 209-218 (1984).

SWALLOW, D.M., GENDLER, S., GRIFFITHS, B., CORNEY, G., TAYLOR-PAPADIMITRIOU, J. and BRAMWELL, M.E., The human tumour-associated epithelial mucins are coded by an expressed hypervariable gene locus PUM. Nature, 328, 82-84 (1987).

TAYLOR-PAPADIMITRIOU, J., PETERSON, J.A., ARKLIE, J., BURCHELL, J., CERIANI, R.L. and BODMER, W.F., Monoclonal antibodies to epithelium specific components of the milk fat globule membrane: production and reactions with cells in culture. Int. J. Cancer 28, 17-21 (1981).

TAYLOR-PAPADIMITRIOU, J., MILLIS, R., BURCHELL, J., NASH, R., PANG, L. and GILBERT, J., Patterns of reaction of monoclonal antibodies HMFG-1 and 2 with benign breast tissues and breast carcinomas. J. Exptl.

Pathol., 2, 247-260 (1986).

TOWNSEND, A.R.M., ROTHBARD, J., GOTCH, F.M., BAHADUR, G., WRAITH, D. and MCMICHAEL, A.J., The epitopes of influenza nucleoprotein recognised by cytotoxic T lymphocytes can be defined with short synthetic peptides. Cell, 44, 959-968 (1986).

YOUNG, J.A., WILKINSON, D., BODMER, W., and TROWSDALE, J., Sequence and evolution of HLA-DR7 and DRW53-associated β-chain genes. Proc. Natl. Acad. Sci. USA. 84, 4929-4933 (1987).

Example 3

Expression of a partial cDNA clone coding for several tandem repeats as a fusion protein in E. coli

Clone 7 which was isolated by Gendler et al. (PNAS 84, 6060-6064, 1987) and is made up of approximately 10 tandem repeat units, was cut out of the Bluescript vector SK with EcoR1 and inserted into the expression vector pRX-1 (Rimm and Pollard, Gene 75, 323-327, 1989). Selection for expression of the mucin fragment in E. coli DH-1 strain was done by probing lifts of the transformants with antibody. DNA from plasmid expressing the mucin fragment was prepared and E. coli strain CAG-456 (see Rimm and Pollard) was transformed for higher expression of the mucin fragment. The fragment containing 18 a.a. of the trpE protein could be detected with antibody on Western blots as a fragment of approximately 30 Kd molecular weight. Coomassie blue staining of the gel-separated bacterial lysate allowed visualization of this band which was increased in cells induced with indolyl acrylic acid, and appeared to make up 10-20% of the bacterial protein.

Example 4

Expression of a partial cDNA clone coding for several tandem repeats in a mammalian cell

Clone 7 was excised with EcoR1 from the Bluescript vector (Stratagene, California) and inserted into the EcoR1 site of a vector based on the pJ3 expression vector so that the correct reading frame of the tandem repeat was maintained. Clones with plasmids containing inserts in the correct orientation were selected using oligonucleotides which overlapped the vector insert sequence. DNA was prepared and purified according to standard methods and transfected into Cos cells. After 48 hours cells were dissolved in sample buffer for acrylamide gel electrophoresis and gels were blotted on to nitrocellulose paper. Probing of the blot with antibody using an ELISA to detect bound antibody showed a band around 30 Kd in the transfected cells.

The procedures of Example 3 & 4 may be used to express and test PDTRP epitope polypeptides or fusion proteins according to the inventon.

CLAIMS

1. A polypeptide of formula (I)

$$\{X\}_x\{(A)_mPDTRP(A)_n\}_y\{X\}_z \qquad (I)$$

wherein x and z are the same or different and each is 0 or an integer of one or more;
y is an integer of one or more
each X, where present, is an independently selected amino acid residue;
m and n are the same or different and each is 0 or one and A,P,D,T and R are the residues of the amino acids, alanine, proline, aspartic acid, threonine and arginine respectively; or containing a sequence immunologically equivalent to the sequence

$$(A)_mPDTRP(A)_n$$

as defined above, optionally bearing one or more saccharide moieties, or a pharmaceutically acceptable salt thereof, other than beta-glycuronidase.

2. A polypeptide selected from

(a) PDTRP

APDTRP

PDTRPA

APDTRPA

(b) polypeptides (a) above bearing one or more saccharide moieties

and (c) pharmaceutically acceptable salts of polypeptides (a) and (b) above.

3. A polypeptide according to claim 1 or claim 2 linked to a carrier protein.

4. A polypeptide according to claim 3 wherein the carrier protein is selected from keyhole limpet haemocyanin, albumen and thyroglobulin.

5. A nucleic acid fragment encoding a polypeptide according to any one of claims 1 to 4.

6. A fragment according to claim 5 containing a continuous coding sequence of formula (II)

$$\{XXX\}_x\{(GCX)_m CCX\ GAPy\ ACX\ RRR\ CCX(GCX)_n\}_y\{XXX\}z \quad (II)$$

wherein RRR is either CGX or AGPu

$$\{X\}_x\{(A)_m PDTRP(A)_n\}_y\{X\}_z$$

x and z are the same or different and each is 0 or an integer of one or more;

y is an integer of one or more

each X, where present, is an independently selected amino

acid residue;

m and n are the same or different and each is 0 or one and A,P,D,T and R are the internationally recognised one-letter symbols for the residues of the amino acids, alanine, proline, aspartic acid, threonine and arginine respectively;

X is any deoxyribonucleotide

XXX is a codon encoding an amino acid residue;

Py is C or T;

Pu is A or G;

and A, C, G and T are the internationally recognised one-letter codes for the four deoxyribonucleotide bases.

7. A fragment according to claim 6 comprising the coding sequence

$(GCC)_m$ CCG GAC ACC AGG CCG $(GCC)_n$.

8. A DNA construct comprising an operational gene, encoding a fusion protein, containing a fragment according to any one of claims 5 to 7 in expressible form.

9. A cloning or expressing vector containing a nucleic acid fragment or construct according to any one of claims 5 to 8.

10. An expression vector consisting of poliovirus genomic DNA containing a coding sequence in expressible form encoding a fragment or construct according to any one of claims 5 to 8.

11. A vector according to claim 10 wherein the

fragment is contained within the poliovirus capsid protein gene.

12. Host cells transformed with a cloning or expression vector according to any one of claims 9 to 11.

13. A nucleic acid probe comprising a fragment according to any one of claims 5 to 7 and a detectable label.

14. An antibody, other than SM-3 and HMFG-2, capable of recognising a polypeptide according to any one of claims 1 to 4.

15. An antibody according to claim 14 linked to a therapeutically or diagnostically effective moiety.

16. A pharmaceutical formulation comprising a carrier or diluent or adjuvant and a polypeptide according to claim 14 or claim 15.

17. A cell, other then HSM-3, capable of secreting antibodies against a polypeptide according to any one of claims 1 to 4.

18. A immortal cell according to claim 17.

19. A polypeptide according to any one of claims 1 to 4, a nucleic acid fragment according to any one of claims 5 to 7, a construct according to claim 8, a vector according to any one of claims 9 to 11, a cell according to claim 12, a probe according to claim 13, an antibody according to claim 14 or claim 15, a pharmaceutical formulation according to claim 16 or a cell according to claim 17 or claim 18 for use in a method for treatment of the human or animal body by surgery or therapy or a

diagnostic method practiced on the human or animal body.

20. Use of a polypeptide according to any one of claims 1 to 4, a nucleic acid fragment according to any one of claims 5 to 7, a construct according to claim 8, a vector according to any one of claims 9 to 11, a cell according to claim 12, a probe according to claim 13, an antibody according to claim 14 or claim 15, a pharmaceutical formulation according to claim 16 or a cell according to claim 17 or claim 18 in preparation of a medicament for use in a method for treatment of the human or animal body by surgery or therapy or a diagnostic method practiced on the human or animal body.

21. A method of treatment or diagnosis comprising administering an effective, non-toxic amount of a polypeptide according to any one of claims 1 to 4, a nucleic acid fragment according to any one of claims 5 to 7, a construct according to claim 8, a vector according to any one of claims 9 to 11, a cell according to claim 12, a probe according to claim 13, an antibody according to claim 14 or claim 15, a pharmaceutical formulation according to claim 16 or a cell according to claim 17 or claim 18 to a human or animal in need thereof.

22. A diagnostic test process comprising contacting a sample suspected to contain abnormal human mucin glycoprotein with an antibody according to claim 14 or claim 15.

23. A diagnostic test kit comprising an antibody according to claim 14 or claim 15.

Fig.1.

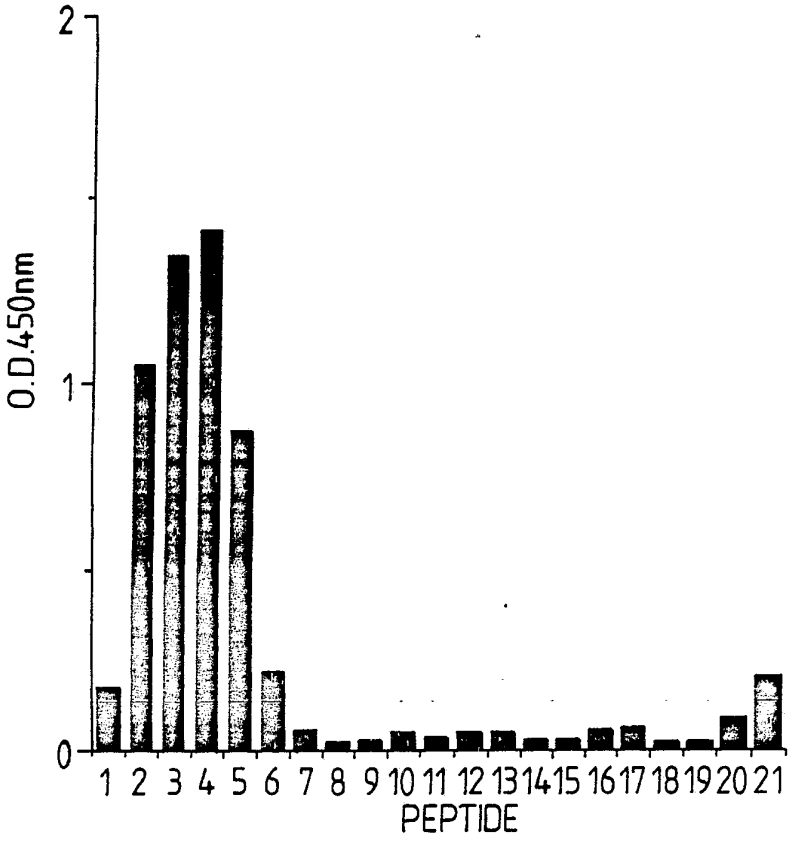

2
1
0
O.D.450nm
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
PEPTIDE

Fig.5.

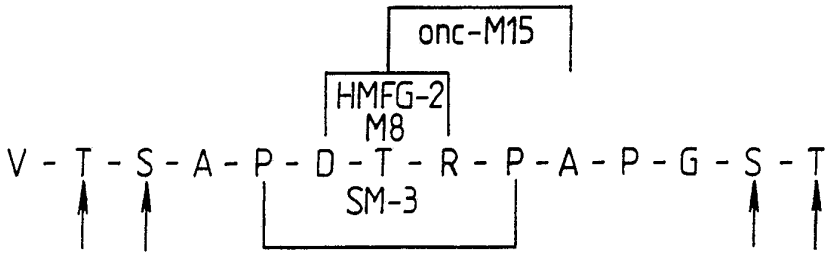

onc-M15
HMFG-2
M8
V - T - S - A - P - D - T - R - P - A - P - G - S - T
SM-3

Fig.2.

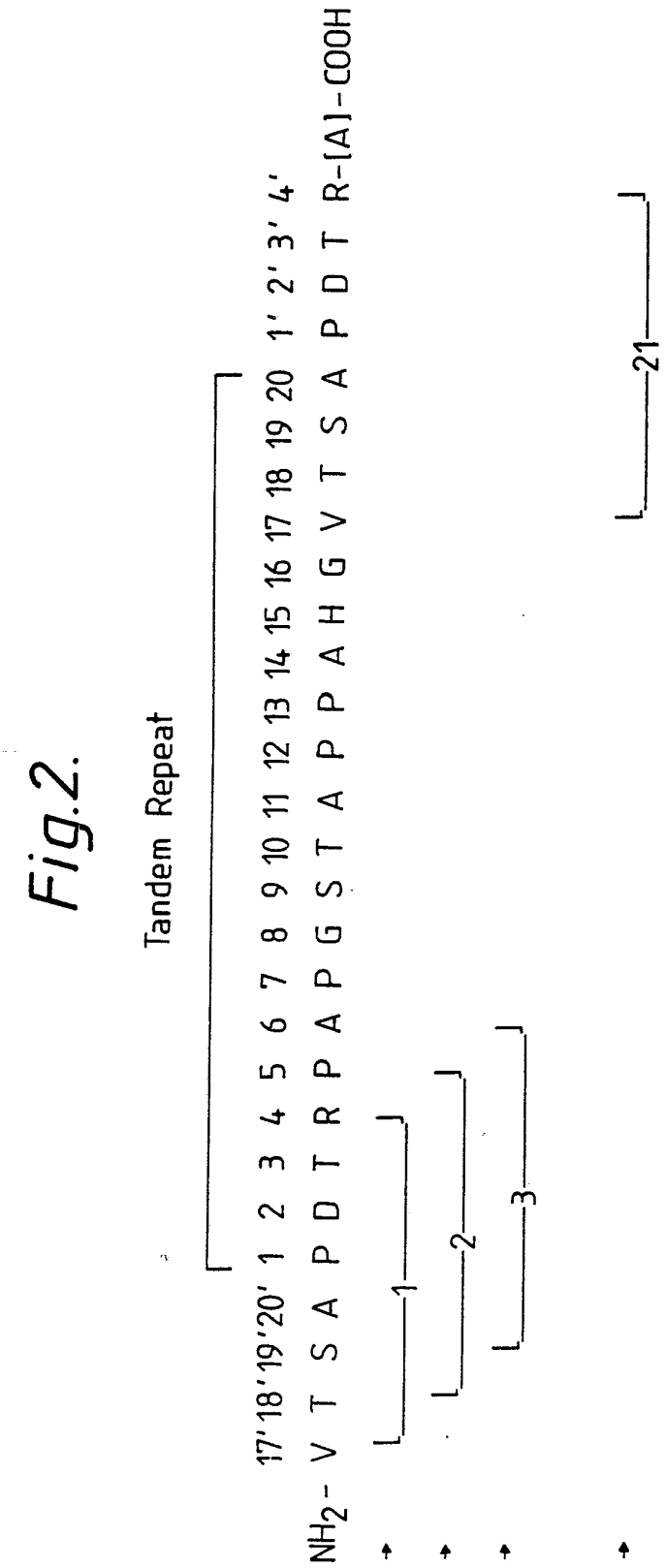


Fig.3.

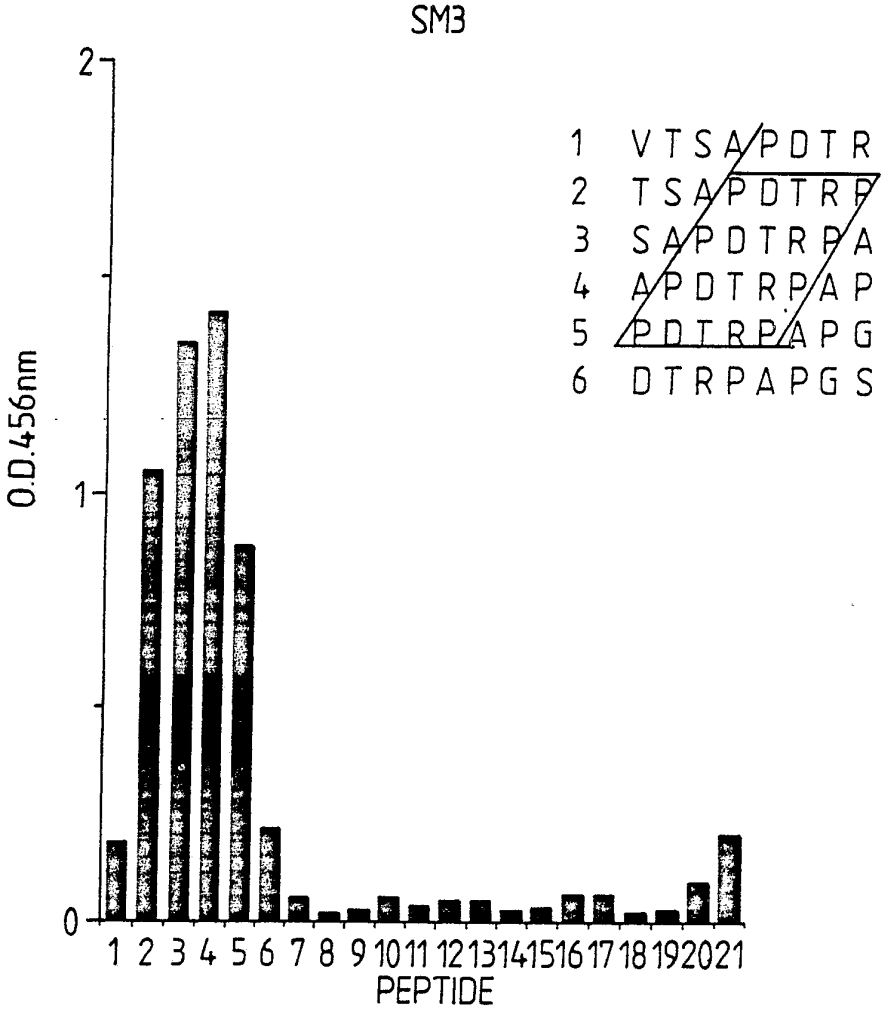
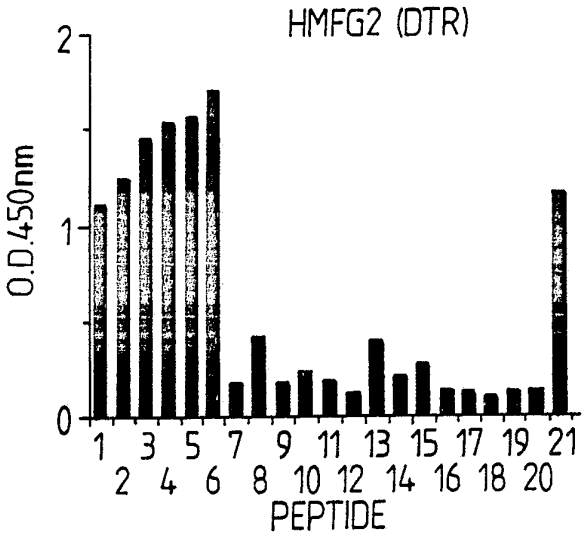
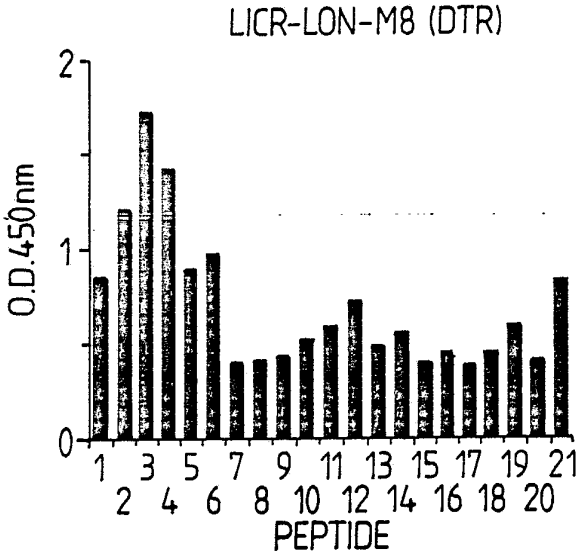

SM3
2
1
0
O.D.456nm
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
PEPTIDE
1 V T S A P D T R
2 T S A P D T R P
3 S A P D T R P A
4 A P D T R P A P
5 P D T R P A P G
6 D T R P A P G S

Fig.4.





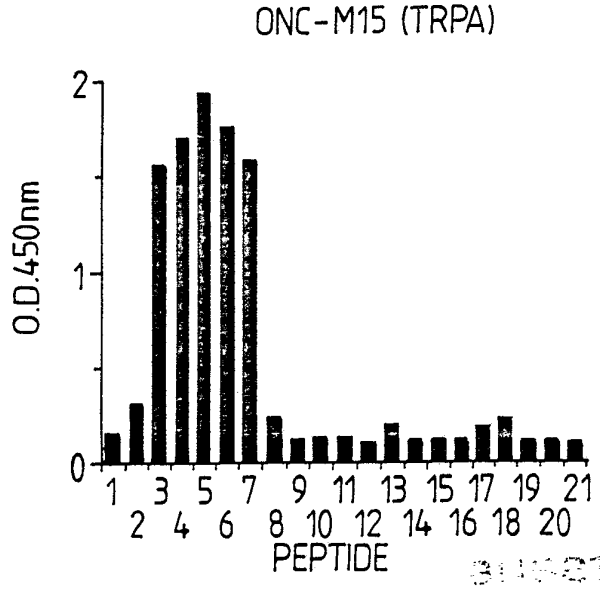


Fig.6.

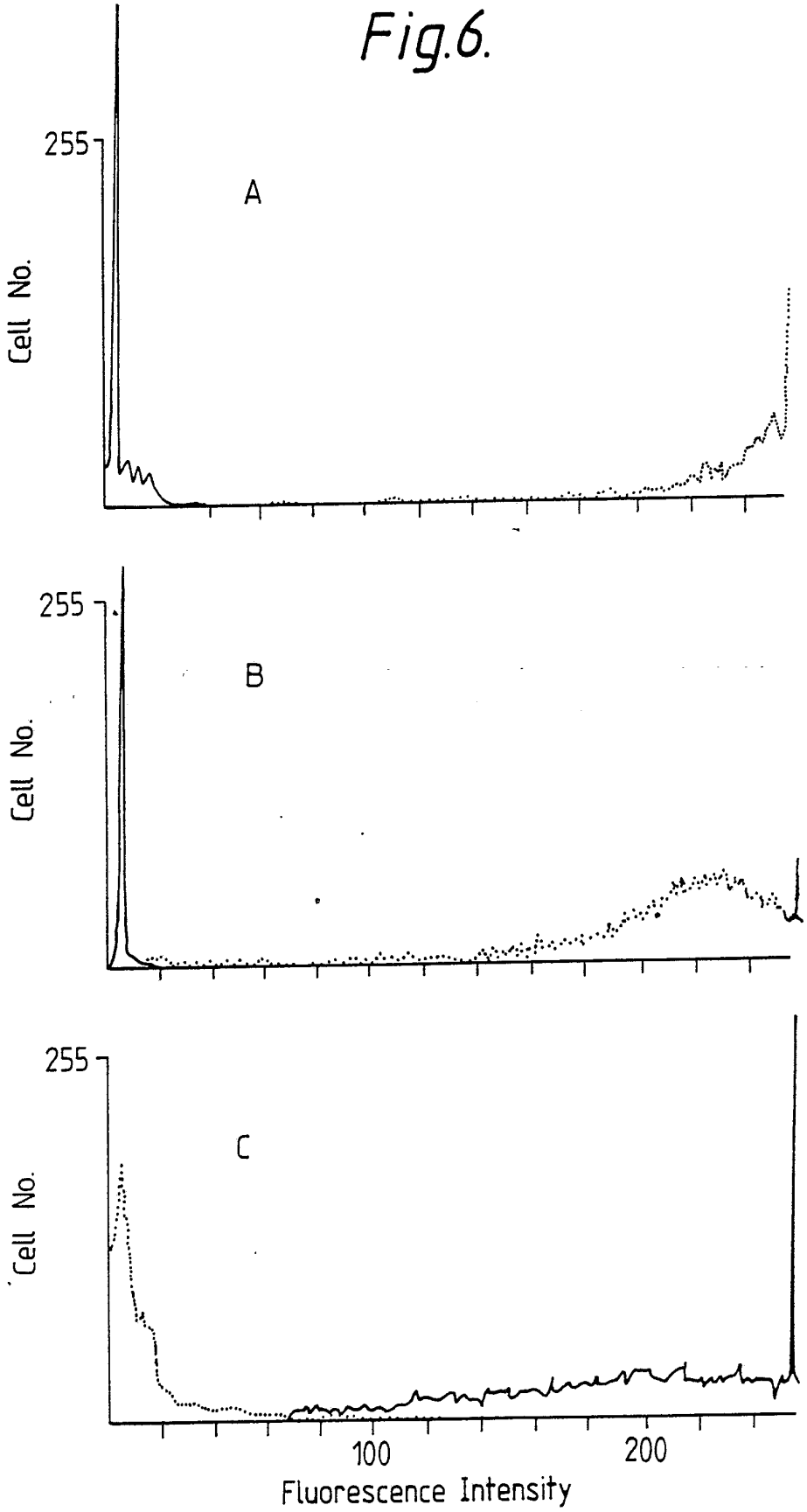

255
A
Cell No.
255
B
Cell No.
255
C
Cell No.
100
200
Fluorescence Intensity

Fig.7.

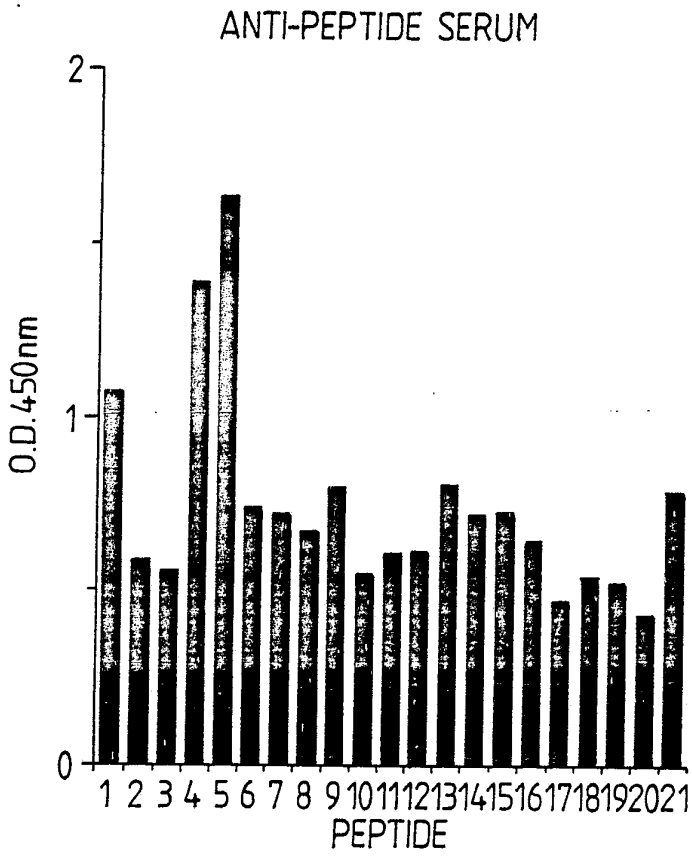
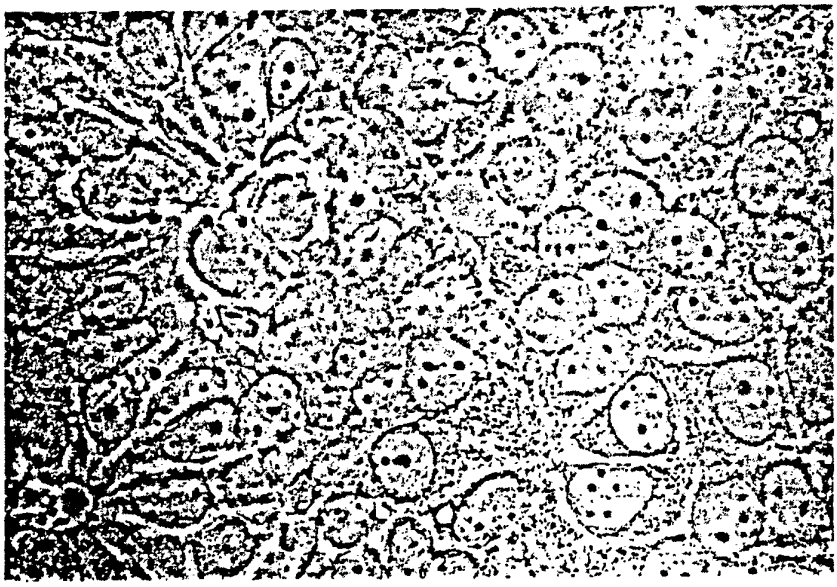


Fig.8A.

Fig.8B.



INTERNATIONAL SEARCH REPORT

International Application No PCT/GB 89/01340

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[5]: C 07 K 7/06, A 61 K 37/02, C 12 N 15/12, C 12 N 15/62, C 12 Q 1/68, C 12 P 21/08

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[5]	C 07 K, A 61 K

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X	WO, A, 88/05054 (IMPERIAL CANCER RES. TECHN. LTD) 14 July 1988, see page 5, paragraph 3 - page 14, paragraph 2; page 27, paragraph 3 - page 30, paragraphe 4; example 2, lines 1-33 (cited in the application)	1,3-6,8-20,. 22,23
Y	--	1-20,22,23
Y	Proc. Natl. Acad. Sci. USA, vol. 81, July 1984 (US) H. Mario Geysen et al.: "Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid", pages 3998-4002, see the whole article (cited in the application)	1-20,22,23
	--	
X	Biochemical and Biophysical Research Communications, vol. 146, no. 2, 31 July 1987, Academic Press, Inc. (US) S. Imajoh et al.: "Calcium-activated	1

./.

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
6th February 1990	13. 03. 90
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	T.K. WILLIS

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

	neutral protease inhibitor from rabbit erythrocytes lacks the N-terminal region of the liver inhibitor but retains three inhibitory units", pages 630-637 see figure 2 --	
P,X	Int. J. Cancer, vol. 44, 1989, Alan R. Liss, Inc. J. Burchell et al.: "A short sequence, within the amino acid tandem repeat of a cancer-associated mucin, contains immunodominant epitopes", pages 691-696, see the whole article ----------	1-20,22,23

V. [X] OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. [X] Claim numbers XX, because they relate to subject matter not required to be searched by this Authority, namely:

Pls. see Rule 39.1 (iv) - PCT:
Method for treatment of the human or animal body by surgery or therapy, as well as diagnostic methods

2. [ ] Claim numbers....., because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claim numbers....., because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. [ ] OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.
2. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:
3. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:
4. [ ] As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

- [ ] The additional search fees were accompanied by applicant's protest.
- [ ] No protest accompanied the payment of additional search fees.

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. GB 8901340 SA 32256

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 06/03/90
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A- 8805054	14-07-88	AU-A- 1103988	27-07-88
		EP-A- 0341252	15-11-89