(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 November 2007 (08.11.2007)

PCT

(10) International Publication Number
WO 2007/125371 A2

(51) International Patent Classification:
C07K 7/00 (2006.01)

(21) International Application Number: PCT/GB2007/050225

(22) International Filing Date: 30 April 2007 (30.04.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0608368.7 28 April 2006 (28.04.2006) GB

(71) Applicant *(for all designated States except US)*: **ISIS INNOVATION LTD** [GB/GB]; Ewert House, Ewert Place, Oxford, Oxfordshire OX2 7SG (GB).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **JIANG, Shisong** [CN/GB]; Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford, Oxfordshire OX3 9DS (GB).

(74) Agent: WRIGHT, Gordon; Elkington And Fife LLP, Prospect House, 8 Pembroke Road, Sevenoaks, Kent TN13 1XR (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: PROCESS FOR MAKING OLIGOPEPTIDES

MGGKWSKSSVVGWPAVRERM**IEGR**VGWPAVRER
MRRAEPAADGV**IEGR**RRAEPAADGVGAVSRDLEK
H**IEGR**GAVSRDLEKHGAITSSNTAA**IEGR**GAITSSN
TAATNADCAWLEA**IEGR**TNADCAWLEAQEEEEVG
FPV**IEGR**QEEEEVGFPVTPQVPLRPMT**IEGR**TPQVP
LRPMTYKAAVDLSHF**IEGR**YKAAVDLSHFLKEKG
GLEGL**IEGR**LKEKGGLEGLIHSQRRQDIL**IEGR**IHSQ
RRQDILDLWIYHTQGY**IEGR**DLWIYHTQGYFPDWQ
NYTPE**IEGR**FPDWQNYTPEPGVRYPLTFG**IEGR**PG
VRYPLTFGWCY

(57) Abstract: There is described a method of synthesising a family of oligopeptides of predetermined amino acid sequence, which together make up from 7 amino acids to the complete amino acid sequence of a target protein, which comprises: synthesising a nucleic acid construct which codes for a fusion protein composed of overlapping peptides derived from the desired portion o f the target protein, interspersed with regions which code for protease cleavage sites, expressing the nucleic acid construct in a suitable expression vector, harvesting the fusion protein corresponding to the nucleic acid sequence, and digesting the fusion protein with a protease selective for the cleavage sites to generate the oligopeptides. The oligopeptides may be generated in vitro or in vitro and may be used as vaccines against viral infections and in epitope mapping.

PROCESS FOR MAKING OLIGOPEPTIDES

This invention relates to a process of making recombinant overlapping peptides, in particular recombinant overlapping peptides for use in the preparation of a vaccine for the prevention and treatment of viral infections and tumours.

WO 01/16163 discloses oligopeptide mixtures, comprising a mixture of 10 to 30 amino acids oligopeptides, each with a 5 to 25 amino acid overlap of the adjacent overlapping peptide spanning the amino acid sequence of a viral protein of a virus, for example Hepatitis B. According to this patent application, a peptide mixture composed of seventeen 20 to 23 amino acids oligopeptides spanning the amino acids 1 to 183 of the hepatitis B core antigen (HbcAg) could activate specific T cells regardless of the host MHC/HLA genotype that recognize the native protein processed by professional antigen presenting cells (APCs). The overlapping peptides exemplified are synthesised using standard chemical techniques using an automated synthesizer. These techniques are slow, low yielding and not well adapted to rapid scale up. The use of short peptide fragments, which may be overlapping or contiguous, as vaccines are also described in WO2004/002415.

We have now found a new method for synthesising oligopeptides for use in the vaccination methods described above which overcomes the problems associated with the conventional chemical synthesis and offers the possibility of rapid generation of a family of oligopeptides on a commercial scale.

According to one aspect of the invention we provide a method of synthesising a family of oligopeptides of predetermined amino acid sequence, which together make up from 7 amino acids to the complete amino acid sequence of a protein, which comprises:
synthesising a nucleic acid construct which codes for a fusion protein composed of overlapping peptides derived from the desired portion of the protein, interspersed with regions which code for protease cleavage sites, expressing the nucleic acid sequence in a suitable expression vector, harvesting the protein corresponding to the nucleic acid sequence, and digesting the protein with a protease selective for the cleavage sites to generate the oligopeptides.

According to a second aspect of the invention we provide a method of synthesising a family of oligopeptides of predetermined amino acid sequence, which together make up from 7 amino acids to the complete amino acid sequence of a target protein, which comprises:
synthesising a nucleic acid construct which codes for a fusion protein composed of overlapping peptides derived from the desired portion of the target protein, interspersed with regions which code for protease cleavage sites, introducing the nucleic acid construct into a suitable vector, inducing expression of the fusion protein by a host cell, whereby the expressed fusion protein is cleaved intracellularly by a host protease or a non-host protease encoded by the same or different vector, the protease being selective for the cleavage sites, to generate the oligopeptides, and harvesting the oligopeptides.

The oligopeptides produced by the methods of the invention may be administered as vaccines to humans or animals.

According to a third aspect the invention provides a fusion protein composed of overlapping peptides derived from a desired portion of a target protein interspersed with regions which code for protease cleavage sites. This may be directly administered alone or in combination with a protease to humans or animals as a vaccine.

The oligopeptides making up the desired sequence of the target protein may be either contiguous or overlapping. Preferably the oligopeptides comprise a mixture of 10 to 30 amino acids long oligopeptides each with a 5 to 25 amino acid overlap with the adjacent overlapping peptide spanning the amino acid sequence of the target protein.

The oligopeptides are preferably of equal length, but can differ in length.

The protease cleavage site may be any amino acid sequence which is selectively digested by a protease. The cleavage site may be of any length. Preferably, the cleavage site is less than ten amino acids, more preferably less than eight amino acids, most preferably less than six amino acids in length. We particularly prefer the protease cleavage site to be a Factor Xa digestion site, that is, the sequence Ile-Glu-Gly-Arg-.

This sequence is cleaved after the Arg. Other proteases that may be mentioned include HRV 3 C protease, which cleaves the sequence Leu-Glu-Val-Leu-Phe-Gln/Gly-Pro- between the glutamyl and glycyl residues; the HIV protease; metalloproteinases, tryptases and other proteases such as cathepsins (S, L and B etc) CD13(human aminopeptidase N).

Publications relating to carthepsins and their specificity as proteases include:

Ruckrich, T., J. Brandenburg, A. Cansier, M. Muller, S. Stevanovic, K. Schilling, B. Wiederanders, A. Beck, A. Melms, M. Reich, C. Driessen, and H. Kalbacher. 2006. Specificity of human cathepsin S determined by processing of peptide substrates and MHC class II-associated invariant chain. Biol Chem 387:1503.

Choe, Y., F. Leonetti, D. C. Greenbaum, F. Lecaille, M. Bogyo, D. Bromme, J. A. Ellman, and C. S. Craik. 2006. Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities. J Biol Chem 281:12824.

In addition to the above-mentioned literature, the cathepsin S sequence TVQ/L ("/" represents the cutting site) is particularly preferred.

The target protein may be any protein, which may for example be selected according to the disease to be treated or prevented. Typically, the target protein will be a characteristic of the virus responsible for the disease to be treated, e.g a coat protein of the virus or enzyme characteristic of the virus.

For example, in HIV infections, either the HIV Nef or HIV Gag proteins may be selected as the target protein; in influenza, the target protein could be a hemoglutinin, neuraminidase, nucleoprotein or matrix protein; in colon cancer, the target protein could be carcinoembryonic antigen; in breast cancer, Her2/neu antigen..

Preferably the family of oligopeptides together make up 100% the complete amino acid sequence but could be as few as 7 amino acids of the target protein. In general we prefer the family of oligopeptides to make up at least 50%, more preferably 75% of the complete amino acid sequence of the target protein. The nucleic acid construct which codes for the desired portion of the target protein interspersed with regions which code for protease cleavage sites may be synthesised by commercially available technologies, eg GeneArt^{TM}, GeneMaker^{TM}, GenScript^{TM}.

Suitable expression vehicles for amplifying the construct are preferably prokaryotic expression vector pET 16 when Factor Xa is used for protein digestion. Alternative expression systems could be other prokaryotic expression vectors, such as other pET series vectors, pGEX, pGEMT7, mammalian expression vectors such pEGFP, pcDNA3, pCDM7, or insect expression vectors such as Bac-To-Bac, BaculoDirect.

Suitable vehicles for expressing the target protein coded by the construct will be applied accordingly, e.g. E. coli BL21 for prokaryotic expression, HEK293T cells, HELA cells, etc for mammalian expression, sf9 cells etc for insect expression.

According to the first aspect of the invention the fusion protein may be harvested and purified using standard methods, e.g. in the case of pET16 prokaryotic system, protein expression will be induced by IPTG (isopropyl-beta-D-thiogalactopyranoside), and the protein purified from the cell lysate by affinity chromatography, preferably Ni-NTA column. Following digestion by specific protease, which is normally commercially available, e.g. Factor Xa (Merck Novagen), oligopeptide will be further purified by standard techniques, e.g. high performance liquid chromatography (HPLC), ion exchange chromatography, gel filtration chromatography.

More specifically, the oligopeptides may be prepared *in vitro* as follows:

1. Design and generation of the synthetic gene

A synthetic gene consisting of 20mer overlapping peptides from the original target gene can be designed with protease Factor Xa sites added in between the 20mer peptides. The resultant artificial gene can be optimised for *E. coli.* expression and commercially synthesised before being cloned into prokaryotic expression vector such as pET16.

2. Protein expression, digestion and purification *in vitro*

The construct with the artificial gene can be expressed in *E. Coli.* BL21 as inclusion bodies, which can be purified**,** and the protein completely digested with protease Factor Xa. The resultant 24mer peptides (20mer from the target vaccine gene plus 4mer Factor Xa recognition site) can be purified, e.g.by High Performance Liquid Chromatography (HPLC).

Alternatively, the fusion protein containing protease cleavage sites can be directly administered as a vaccine. It will be cleaved by a protease inside the body or a protease administered separately or together with the fusion protein.

According to the second aspect of the invention the nucleic acid sequence may be expressed in a host cell using any of the expression vectors and systems described above.

Alternatively, the nucleic acid sequence may be expressed by a bacterial or viral vector within a host cell which has been infected by the virus or bacterium. Preferred bacterial vectors are *Listeria*, *Salmonella*, and *Shigella*. Preferred viral vectors are adenovirus, poxvirus, lentivirus and alphavirus, etc. *Listeria* vectors are particularly preferred because these specifically infect antigen presenting cells (i.e. macrophages and dendritic cells). The viral vector may be attenuated or not attenuated. The nucleic acid sequence may be transformed into the viral or bacterial vector via a shuttle vector such as pKSV7 (1).

By “host cell” we mean a cell which is capable of expressing the target protein coded by the construct and in which cleavage of the target protein may occur intracellularly. Consequently, the host cells may be mammalian or insect cells in vitro, such as HEK293T cells, HELA cells, sf9 cells. Alternatively the host cells may be human or animal cells in vivo, which have been infected with a viral or bacterial vector containing a construct encoding the fusion protein, as a vaccine.

The expressed fusion protein is cleaved intracellularly by either a host protease or a protease encoded by the plasmid vector or a protease which is found in the viral or bacterial vector. The protease is selective for the cleavage sites. The vector may also encode an activator for the protease which is selective for the cleavage site. For example, the host protease may be non-active under normal conditions and become active only when the activator is expressed.

According to one embodiment the protease may be pathogen-specific. Accordingly, the nucleic acid may be designed such that protease cleavage sites may only be cleaved by a pathogen-specific protease. Then, the expressed fusion protein is only cleaved when the host cell is infected with a pathogen which produces a protease which is specific for the cleavage sites.

More specifically, the oligopeptides may be prepared *in vivo* as follows:

1. Design and generation of the synthetic gene

A synthetic gene consisting of 20mer overlapping peptides from the original target gene can be designed with protease Factor Xa sites added in between the 20mer peptides. The resultant artificial gene can be optimised for bacterial expression and commercially synthesised before being cloned into prokaryotic expression vector, for example pET16, or a shuttle vector like pKSV7. A gene encoding the protease Factor Xa will also be cloned into the expression vector.

2. Production of vectors which encoding the target gene and/or the genes encoding proteases which can cut the fusion protein.

The construct with the artificial gene and the protease gene will be transformed into a attenuated or not attenuated *Listeria* vector. This may be administered as a vaccine. If the viral vector does not contain a protease gene, the vector may be administered in combination with a protease as a vaccine.

The invention also provides a fusion protein composed of overlapping peptides derived from a desired portion of a target protein interspersed with regions which code for protease cleavage sites and a nucleic acid sequence coding for the fusion protein.

The nucleic acid sequence encoding the fusion protein or the fusion protein itself may be used for the manufacture of a medicament for the prophylactic treatment of viral infections such as HIV, HCV, Herpes etc, or tumours such as melanoma, breast cancer, renal cancer, hepatic cancer, etc. Alternatively the medicament may be used for the treatment of a disease caused by a pathogen which produces a protease specific for the cleavage sites.

The invention also provides a method of treating a disease in a patient comprising:
treating the patient with a nucleic acid sequence encoding a fusion protein composed of overlapping peptides derived from a desired portion of a target protein interspersed with regions which code for protease cleavage sites,
inducing expression of the fusion protein in the patient, whereby the expressed fusion protein is cleaved intracellularly by a protease which is specific for the cleavage sites;
and a method of treating a disease in a patient comprising:
treating the patient with a fusion protein composed of overlapping peptides derived from a desired portion of a target protein interspersed with regions which code for protease cleavage sites, whereby the expressed fusion protein is cleaved intracellularly by a protease which is specific for the cleavage sites.

The disease to be treated may be a viral infection such as HIV, HCV, Herpes etc, or tumours such as melanoma, breast cancer, renal cancer, hepatic cancer, etc. Alternatively, the disease may be caused by a pathogen which produces a protease specific for the cleavage sites.

For example, an HIV fusion protein (e.g. Nef) with HIV protease substrate sites in between overlapping peptide sequences may be administered to a patient infected with

HIV. The HIV protease present in the patient's body will cut the fusion protein into overlapping peptides which then will stimulate the immune system to kill the HIV infected cells.

The invention also provides a method of epitope mapping which comprises synthesising a family of oligopeptides of predetermined amino acid sequence, which together make up from 7 amino acids to the complete amino acid sequence of a target protein, which comprises:

synthesising a nucleic acid construct which codes for a fusion protein composed of overlapping peptides derived from the desired portion of the target protein, interspersed with regions which code for protease cleavage sites, expressing the nucleic acid construct in a suitable expression vector, harvesting the fusion protein corresponding to the nucleic acid sequence, and digesting the fusion protein with a protease selective for the cleavage sites to generate the oligopeptides and using these oligopeptides in an oilgopeptide epitope mapping procedure.

Example 1

The invention is now illustrated by reference to the following embodiment, carried out with HIV Nef protein

Methods:

1. DNA Design

A synthetic gene was designed in such a way that it was composed of a series of 10-amino-acids overlapped 20mer peptides, interspersed with protease Factor Xa cleavage sites (Ile-Glu-Gly-Arg-; I-E-G-R-), covering the full length of the 206 amino acids Nef protein from Lai strain of Human Immunodeficiency Virus (HIV) (Figure 1). The resulting artificial gene was optimised for *E. Coli.* Expression, and restrictive enzyme sites Nde I and Bam were added at the ends for subsequent subcloning and expression in pET16b vector.

2. Gene synthesis

The gene was commercially synthesised by GeneArt, Regensburg, Germany.

3. Subcloning of the target gene into pET16b expression vector

The synthesised target gene, provided in a cloning vector pPCR-Script, was restrictively digested with Nde I and Bam HI. After agarose electrophoresis, the band corresponding to the target gene was sliced under ultraviolet light and the DNA was purified by QIAquick® Gel Extraction kit (Qiagen, Germany). Meanwhile, pET vector was digested with the same restrictive enzymes, ends de-phosphorylated with shrimp alkaline phosphatase (Roche, UK), and gel purified in the same way. The vector DNA and the target gene were ligated with T4 DNA ligase (New England Biolabs, UK) and the ligation product was transformed into DH5α bacteria. Colonies were inoculated from overnight LB agar plate culture and incubated in LB media with 100mg/ml ampicillin. After overnight culture, plasmid DNA was extracted by QIAprep Spin Miniprep Kit (Qiagen, Germany), and sequenced to confirm correct insert.

4. Protein expression

Miniprep DNA was transformed into BL21(DE3) bacteria, and colonies were inoculated and grown in low salt LB media overnight at 37°C or 28°C. The culture were 1:100 diluted in low salt LB media and grown at 37°C under shaking for several hours, until the bacteria reach exponential growth stage, as determined by OD600 = 0.3~0.5. One millimolar IPTG was be added to induce the protein expression. After 3~5 hours culturing with shaking at 37°C or 28°C, bacteria were collected by spinning at 4000rpm, then re-suspended in PBS, and lysed by sonication. The inclusion body, which contains the expressed protein, were collected by centrifugation.

5. Protein purification

The expressed protein was purified by B-per (Pierce,UK) followed by Ni-NTA purification kit (Qiagen, Germany) according to the operation manual. Briefly, the inclusion bodies were re-suspended in denaturing buffer (8M Urea Tris buffer), and passed through the Ni-NTA column, to which the His-tagged protein binds. After extensive washing, protein was eluted by elution buffer, and buffer exchanged to Factor

Xa digestion buffer(Tris buffer with 5mM CaCl2) by passing through PD-10 column (Sephadex G-25).

6. Protease digestion

Protein in digestion buffer was concentrated by Centriprep YM10 (Millipore, USA) and digested completely by incubating the protein with protease Factor Xa (New England Biolabs, UK) at room temperature for around 20 minutes. The protease was removed with benzamidine-agarose (Amersham Biosciences, USA) or simply through a YM10 filter tube, and the free His-tag will be removed with Ni-NTA agarose. The resulting mixture of oligopeptides will be concentrated and buffer-exchanged to PBS for further use.

7. Mass spectrometer analysis

The protease-digested product was analysed by mass spectrometry to confirm the digestion results in the desired overlapping peptides. Samples were carried out on an Applied Biosystems 4700 proteomics analyzer equipped with TOF/ TOF ion optics and a diode pumped Nd:YAG laser with 200 Hz repetition rate. Generally, the width of mass window was set to 50 Da because the mass of precursor ion is less than 5000 Da, MS/MS data was acquired using the instrument default calibration, without applying internal or external calibration.

8. Mice and Immunization.

Inbred mice BALB/c (H-2^d) and C57BL/10 (H-2^b) were immunized with recombinant overlapping pepties at 5 µg of each individual peptide in 100 µl PBS per mouse together with MLP + TDM Adjuvant System. Control mice weregiven only adjuvant. Immunizations were given at weeks 0, 3, and 6. Three weeks after the last immunization, mice were sacrificed for CTL and T-helper cell proliferation assays.

Alternatively, randomly bred, outbred NMRI mice may be immunized subcutaneously (s.c.) with recombinant overlapping peptides at 5 µg of each individual peptide in 100 μl PBS per mouse together with or without MLP + TDM Adjuvant System 3 times at 3-week intervals. Control groups may be only given adjuvant or PBS. Another 3 weeks after the last boost, splenocytes may be collected and IFN-γ-specific ELISPOT assays and intracellular staining for IFN-γ may be performed.

9. Intracytoplasmic Cytokine Staining and Flow Cytometry Analysis.

Mouse splenocytes may be cultured at 5 x 10^6 cells/ml with IL-2 (20 U/ml) with or without 1 μM recombinant overlapping peptides in 24-well culture plates for 6 h. Four hours before harvesting, cells may be treated with Golgistop (BD PharMingen) according to the vendor's protocol. Splenocytes may be then stained with phycoerythrin-(PE)-conjugated monoclonal rat anti-mouse CD8 or CD4 antibody (BD PharMingen) or an immunoglobulin isotype control for 20 min. Splenocytes may be then subjected to intracellular cytokine staining using the Cytofix/Cytoperm kit (BD PharMingen) and FITC-conjugated anti-IFN-γ antibody (20 μg/ml) according to the manufacturer's instructions. Samples may be acquired on an Epics XL flow cytometer (Beckman Coulter, Fullerton, CA), and data may be analyzed using Expo 32 software (Beckman Coulter).

10. ELISPOT Assay.

ELISPOT assays were performed using ELISPOT kits from BD PharminGen. Briefly, splenocytes were restimulated overnight with 1 μM recombinant overlapping peptides in the IFN-γ-precoated plates. Cells were discarded and biotinylated anti-IFN-γ antibodies were added for 1 hour at 37°C followed by another hour of incubation at 37°C of anti-biotin antibody labelled with enzyme. After colour developed, spots were counted under a microscope. Results are expressed as SFU/10^6 cells.

11. Lymphocyte Proliferation Assay.

Splenocytes may be isolated and cultured at 2 x 10^6 / ml in RPMI 1640 plus 15 % FCS plus antibiotics in the presence of either HIV Nef protein (15 μg/ml), recombinant overlapping peptides (3 μg/ml) or OVA (15 μg/ml) for 5 days. Four hours before harvesting, cells may be pulsed with 1 μCi per well of ^{3}H-thymidine. After cells are harvested, ^{3}H-thymidine incorporation may be assessed using a β-counter (Beckman, Fullerton, CA). Results are expressed as stimulation index (SI), i.e., the ratio of cpm of stimulated cells to cpm of cells grown in medium only.

12. Rhesus monkey experiment

Four monkeys may be enrolled initially and different doses of recombinant overlapping peptides will be administered. IFN-γ-ELISPOT and proliferation assays may be performed to find an optimal dose to induce cellular immune responses.

Three groups of rhesus monkeys may be enrolled which will be given respectively; recombinant overlapping peptides (at the optimal dose), recombinant protein from which the overlapping peptides are generated at equivalent dose to the recombinant overlapping peptides, and PBS. IFN-γ-ELISPOT and proliferation assays will be performed to test if overlapping peptides can induce immune responses in the primate model.

Results of Example 1

Brief Description of Drawings

Figure 1 : A recombinant protein composed of 20 overlapping peptides covering Nef and a sort substrate sequence (in bold) of Factor Xa interspeased between each peptide, overlapping amino acid residues underlined;

Figure 2 : Purified (right) and unpurified (left) recombinant overlapping peptides (ROP) fusion protein (55kd):

Figure 3: SDS page and Mass spectrometry profile of ROP peptides after digesting ROP protein by Factor Xa.

(a): titration of Factor Xa. 100 ul ROP protein (2.2mg/ml) incubated with series of two fold diluted Factor Xa. Ln 0: no Factor Xa; Ln 1: 1 μl (2unit/μl); Ln2: ½ μl; Ln 3: 1/4 μl;….Ln9: 1/256 μl;

(b): Mass Spetrometry of digested ROP. The responses were at 4°C 1hour.

Figure 4: ROP peptides and ROP protein taken up by DC2.4 cells, showing that oligopeptides were taken earlier than the protein.

Figure 5: Uptaken ROP peptides (left) and ROP protein (right) by DC2.4 cells ROP peptides and ROP protein were labelled with green fluorescence while lysosomes were stained with LysoTracker-red. Note; in the right picture, the green labelled protein are co-localized with the red stained lysosomes.

Figure 6: Immunization of two different strains of mice with ROP peptides generated immune responses to the peptides, in different strains of mice the immune response target different individual peptides within the pool of overlapping peptides.

Figure 7: Peptide ASNENMDAM TQV/L ASNENMETH was incubated at 37°C for 30 minutes with 5 mU Cathepsin S (Calbiochem). The enzyme reaction was terminated by heating up at 100 °C for 5 minutes. The products were filtered through Millipore Microcon centrifugal filter (YM-10) to separate peptides from the enzyme, which were analysed by mass spectrometry. The peptide was cleaved into two peptides: ASNENMDAM TVQ and LASNENMETH.

Results:

A recombinant protein composed of 20 overlapping peptides covering Nef protein was made (ROP protein) and a short substrate sequence of the protease factor Xa was interspersed between each peptide (Fig. 1).

After the purification (Fig. 2), the ROP protein was cut by Factor Xa into overlapping peptides (ROP peptides) as shown in the mass spectrometry (Fig. 3).

The way that mouse dendritic cells (DC) taking ROP protein and ROP peptides (both labelled with green fluorescence) was compared. It appeared that ROP peptides were taken up by the DC quicker (started from 5 minutes) than ROP protein (Fig. 4) and most of the peptides (labelled with green fluorescence) went directly into the cytoplasma; while the ROP protein (labelled with green fluorescence) went to lysosomes (stained with red colour) (Fig.5).

Immunization of two different strains of mice with ROP peptides generated immune responses to the peptides as well as to the ROP protein in both strains of mice. Nevertheless, in different strains of mice the immune response target different individual peptides within the pool of overlapping peptides (Fig. 6).

Comparison of ROP identified epitopes with database epitopes

	Sequences	**HIV Nef (strains)**	**Position**	**Mouse strains**
Epitopes from database	**CD4**			
	MGGKWSKSSVVGWPTVRERM	Nef (Lai)	1-20	H-2^d
	MGGKWSKSSVIGWPTVRERM	Nef(HXB2)	1-20	H-2^d
	VRERMRRAEPAADGVGAASR	Nef (Lai)	16-35	H-2^d
	GAASRDLEKHGAITSSNTAA	Nef (Lai)	31-50	H-2^d
	SNTAATNAACAWLEAQEEEE	Nef (BRU)	46-65	H-2^d
	SSNTAATNAACAWLEAQEEEEVGFP		46-69	
	QEEEEVGFPVTPQVPLRPMT	Nef (Lai)	61-80	H-2^b
	*LRPMTYKAAVDLSHF*LKEKG	Nef (Lai)	76-95	H-2^b
	VGFPV*TPQVPLRPMTYKAAVDLSHF*LKEKGGL			
	LKEKGGLEGLIHSQRRQDIL	Nef (Lai)	91-110	H-2^b
	RQDILDLWIYHTQGYFPDWQ	Nef (Lai)	106-125	H-2^b
	FPDWQNYTPGPGVRYPLTFG	Nef (Lai)	121-140	H-2^b
	PLTFGWCYKLVPVEPDKVEE	Nef (Lai)	136-155	H-2^d
	DKVEEANKGENTSLLHPVSL	Nef (Lai)	151-170	H-2^d
	HPVSLHGMDDPEREVLEWRF	Nef (Lai)	166-185	H-$2^{b,d}$
	CD8			
	TAATNADCA	Nef (HXB2)	48-56	H-2^b
	GVRYPLTFGWCYKLVP	Nef (Lai)	132-147	H-2^d
	EWRFDSRLAFHHVAREL	Nef(HXB2)	182-198	H-2^d
ROP epitopes	**QEEEEVGFPVTPQVPLRPMT**	Nef Lai	61-80	H-2^b
	TPQVP*LRPMTYKASVDLSHF*	Nef Lai	71-90	H-2^b
	YKAAVDLSHFLKEKGGLEGL	Nef Lai	81-100	H-2^b
	NTRLLHPVSLHGMDDPEREV	Nef Lai	161-180	H-2^b
	FHHVARELHPEY	Nef Lai	191-202	H-2^b
	VGWPAVRERMRRAEPAADGV	Nef Lai	20-40	H-2^d
	TNADCAWLEAQEEEEVGFPV	Nef Lai	51-70	H-2^d

ROP epitopes were compared with epitopes from HIV Molecular Immunology Database(http://hiv-vacdb.lanl.gov/content/immunology/tables/ctl_summary.html). Three of the epitopes from H-2^b mice were either exactly the same or overlap with the known epitopes (showed in bold, italic or underlined).

These results show that immunization with recombinant overlapping peptides (ROP) (containing IEGR), generated immune responses to HIV Nef. In another words, the recombinant overlapping peptides (containing IEGR) immunized cells in the mouse body responded to viral elements containg Nef which has been tested by overlapping synthetic peptides.

Epitopes identified in this way can be the same as the epitope in the HIV Molecular Immunology Database (http://hiv-vacdb.lanl.gov/content/immunology/tables/ctl_summary.html); partially the same as that in the database; or totally different. Acitivity was compared with synthetic peptides (overlapping synthetic peptides covering Nef Lai sequence) supplied by Sigma-Genosys, which did not contain the did not contain IEGR sequence.

It can be concluded that IEGR does not affect the ability of recombinant overlapping peptides (ROP) to stimulate Nef specific immune response.

In further experiments, a cathepsin S cleavage site (TQV/L) was introduced into a peptide sequence. This was successfully cleaved *in vitro* using cathepsin S under conventional conditions and the two fragments characterised by MS (Fig. 7)

Example 2

The invention is now illustrated by reference to the following embodiment, carried out with HIV Nef protein

Methods:

1. DNA Design

A synthetic gene will be designed in such a way that it is composed of a series of 10-amino-acids overlapped 20mer peptides, interspersed with protease trptase or cathepsin S cleavage sites, covering the full length of the 202 amino acids Nef protein from Lai strain of Human Immunodeficiency Virus (HIV).

2. Gene synthesis

The gene will be commercially synthesised by GeneArt, Regensburg, Germany.

3. Making live attenuated vector vaccines

A recombinant live attenuated vector vaccine will be constructed, *Listeria-overlapping Nef*, by a stable modification of its chromosomes using the shuttle vector pKSV7 (1) and a protocol modified from Camilli et al. (2), as described previously (3), to insert HIV-gag into the *sep*A gene of *L. monocytogenes*.

4. Mice and Immunization.

Different strains of inbred and/or outbred mice will be immunized subcutaneously (s.c.) with attenuated Listeria-overlapping Nef 10^6-10^7 pfu per mouse once or twice with 2-4 week's interval. Control groups will be only given Listeria only or PBS. One-two weeks after the last boost, splenocytes will be collected and IFN-γ-specific ELISPOT assays and intracellular staining for IFN-γ will be performed.

5. Intracytoplasmic Cytokine Staining and Flow Cytometry Analysis.

Mouse splenocytes will be cultured at 5 x 10^6 cells/ml with IL-2 (20 U/ml) with or without 1 μM recombinant overlapping peptides in 24-well culture plates for 6 h. Four hours before harvesting, cells will be treated with Golgistop (BD PharMingen) according to the vendor's protocol. Splenocytes will be then stained with phycoerythrin-(PE)-conjugated monoclonal rat anti-mouse CD8 or CD4 antibody (BD PharMingen) or an immunoglobulin isotype control for 20 min. Splenocytes will be then subjected to intracellular cytokine staining using the Cytofix/Cytoperm kit (BD PharMingen) and FITC-conjugated anti-IFN-γ antibody (20 μg/ml) according to the manufacturer's instructions. Samples will be acquired on an Epics XL flow cytometer (Beckman Coulter, Fullerton, CA), and data will be analyzed using Expo 32 software (Beckman Coulter).

6. ELISPOT Assay.

ELISPOT assays will be performed using ELISPOT kits from BD PharminGen. Briefly, splenocytes will be restimulated overnight with 1 μM recombinant overlapping peptides in the IFN-γ-precoated plates. Cells will be discarded and biotinylated anti-IFN-γ. antibodies will be added for 1 hour at 37°C followed by another hour of incubation at 37°C of anti-biotin antibody labelled with enzyme. After colour developed, spots will be counted under a microscope. Results are expressed as SFU/10^6 cells.

7. Lymphocyte Proliferation Assay.

Splenocytes will be isolated and cultured at 2 x 10^6 / ml in RPMI 1640 plus 15 % FCS plus antibiotics in the presence of either HIV Nef protein (15 μg/ml), recombinant overlapping peptides (3 μg/ml) or OVA (15 μg/ml) for 5 days. Four hours before harvesting, cells were pulsed with 1 μCi per well of ^{3}H-thymidine. After cells were harvested, ^{3}H-thymidine incorporation was assessed using a β-counter (Beckman, Fullerton, CA). Results are expressed as stimulation index (SI), i.e., the ratio of cpm of stimulated cells to cpm of cells grown in medium only.

8. Rhesus monkey experiment

Four monkeys will be enrolled initially and different doses of recombinant Listeria-overlapping Nef (Listeria expressing Nef and Listeria only as control) will be administered. IFN-γ-ELISPOT and proliferation assays will be performed to find an optimal dose to induce cellular immune responses.

Three groups of rhesus monkeys will be enrolled which will be given respectively; Listeria-overlapping Nef (at the optimal dose), Listeria-Nef and Listeria only. IFN-γ-ELISPOT and proliferation assays will be performed to test if overlapping peptides can induce immune responses in the primate model.

9. Experiments of dendritic cells (DC) taking up peptides and protein.

Mouse DC cell lines DC2.4 will be used to observe how DC cells taking up the recombinant overlapping peptides (ROP peptides) and its corresponding recombinant protein (ROP protein). Both ROP peptides and protein will be labelled using Invitrogen

LIVE/DEAD® Fixable Dead Cell Stain Kits (green-fluorescent reactive dye - cat no L23101). The green fluorescent dye can be bound to amines of peptides or proteins. Free fluorescence does not permeate into live cells. However the peptides or proteins bound with fluorescence can be taken up by live dendritic cells which will become green. This can be measured by flowcytometry or observed under confocal microscopy.

References

1.Smith, K., and P. Youngman. 1992. Use of a new integrational vector to investigate compartment-specific expression of the Bacillus subtilis spoIIM gene. *Biochimie 74:705.*

2.Camilli, A., L. G. Tilney, and D. A. Portnoy. 1993. Dual roles of plcA in Listeria monocytogenes pathogenesis. *Mol Microbiol 8:143.*

3. Thompson, R. J., H. G. Bouwer, D. A. Portnoy, and F. R. Frankel. 1998. Pathogenicity and immunogenicity of a Listeria monocytogenes strain that requires D-alanine for growth. *Infect Immun 66:3552.*

Claims

1. A method of synthesising a family of oligopeptides of predetermined amino acid sequence, which together make up from 7 amino acids to the complete amino acid sequence of a target protein, which comprises:
synthesising a nucleic acid construct which codes for a fusion protein composed of overlapping peptides derived from the desired portion of the target protein, interspersed with regions which code for protease cleavage sites, expressing the nucleic acid construct in a suitable expression vector, harvesting the fusion protein corresponding to the nucleic acid sequence, and digesting the fusion protein with a protease selective for the cleavage sites to generate the oligopeptides.

2. A method according to claim 1, wherein the oligopeptides are contiguous in making up the desired sequence of the target protein.

3. A method according to claim 1, wherein the oligopeptides overlap in making up the desired sequence of the target protein.

4. A method according to any one of claims 1 to 3, wherein the oligopeptides are of equal length.

5. A method according to any of the preceding claims wherein the protease cleavage site is a Factor Xa digestion site.

6. A fusion protein composed of overlapping peptides derived from a desired portion of a target protein interspersed with regions which code for protease cleavage sites.

7. A nucleic acid sequence coding for the fusion protein according to claim 6.

8. A method of synthesising a family of oligopeptides of predetermined amino acid sequence, which together make up from 7 amino acids to the complete amino acid sequence of a target protein, which comprises:
synthesising a nucleic acid construct which codes for a fusion protein composed of overlapping peptides derived from the desired portion of the target protein, interspersed with regions which code for protease cleavage sites, introducing the nucleic acid construct into a suitable vector, inducing expression of the fusion protein by a host cell, whereby the expressed fusion protein is cleaved intracellularly by a host cell protease or a non-host protease encoded by the vector, the protease being selective for the cleavage sites, to generate the oligopeptides, and harvesting the oligopeptides.

9. A method according to claim 8, wherein the oligopeptides are contiguous in making up the desired sequence of the target protein.

10. A method according to claim 8, wherein the oligopeptides overlap in making up the desired sequence of the target protein.

11. A method according to any one of claims 8 to 10, wherein the oligopeptides are of equal length.

12. A method according to any of claims 8 to 11, wherein the protease cleavage site is a Factor Xa digestion site.

13. A method according to any one of claims 8 to 12, wherein the vector also encodes an activator for the protease which is selective for the cleavage site.

14. A method according to any one of claims 8 to 13, wherein the protease cleavage site is cleavable by a pathogen-specific protease.

15. A method according to any one of claims 8 to 14, wherein the vector is a bacterial vector selected from *Listeria*, *Salmonella*, and *Shigella*.

16. A method according to any one of claims 8 to 15, wherein the vector is a bacterial vector selected from *Listeria*, *Salmonella*, and *Shigella*.

17. Use of a nucleic acid sequence encoding a fusion protein composed of overlapping peptides derived from a desired portion of a target protein interspersed with regions which code for protease cleavage sites, for the manufacture of a medicament for the prophylactic treatment of a viral infection.

18 Use according to claim 17, wherein the viral infection is HIV, HCV, Herpes, or a tumour such as melanoma, breast cancer, renal cancer, hepatic cancer

19. Use of a nucleic acid sequence encoding a fusion protein composed of overlapping peptides derived from a desired portion of a target protein interspersed with regions which code for protease cleavage sites, for the manufacture of a medicament for the treatment of a disease caused by a pathogen which produces a protease specific for the cleavage sites.

20. A method of prophylactically treating a viral infection in a patient comprising:
treating the patient with a nucleic acid sequence encoding a fusion protein composed of overlapping peptides derived from a desired portion of a target protein interspersed with regions which code for protease cleavage sites,
inducing expression of the fusion protein in the patient, whereby the expressed fusion protein is cleaved intracellularly by a protease which is specific for the cleavage sites.

21. A method according to claim 20, in which the viral infection is HIV, HCV, Herpes, or a tumour such as melanoma, breast cancer, renal cancer, hepatic cancer.

22. A method of treating a pathogenic disease in a patient comprising:

treating the patient with a nucleic acid sequence encoding a fusion protein composed of overlapping peptides derived from a desired portion of a target protein interspersed with regions which code for protease cleavage sites,

inducing expression of the fusion protein in the patient, whereby the expressed fusion protein is cleaved intracellularly by a protease produced by the pathogen which is specific for the cleavage sites.

23. Use of a fusion protein composed of overlapping peptides derived from a desired portion of a target protein interspersed with regions which code for protease cleavage sites, for the manufacture of a medicament for the prophylactic treatment of a viral infection.

24. Use according to claim 23, wherein the viral infection is HIV, HCV, Herpes, or a tumour such as melanoma, breast cancer, renal cancer, hepatic cancer

25. Use of a fusion protein composed of overlapping peptides derived from a desired portion of a target protein interspersed with regions which code for protease cleavage sites, for the manufacture of a medicament for the treatment of a disease caused by a pathogen which produces a protease specific for the cleavage sites.

26. A method of prophylactically treating a viral infection in a patient comprising treating the patient with a fusion protein composed of overlapping peptides derived from a desired portion of a target protein interspersed with regions which code for protease cleavage sites, whereby the fusion protein is cleaved intracellularly by a protease which is specific for the cleavage sites.

27. Method according to claim 26, wherein the viral infection is HIV, HCV, Herpes, or a tumour such as melanoma, breast cancer, renal cancer, hepatic cancer

28. Method according to claim 26 wherein the protease is administered together with the fusion protein.

29. A method of treating a pathogenic disease in a patient comprising:

treating the patient with a fusion protein composed of overlapping peptides derived from a desired portion of a target protein interspersed with regions which code for protease cleavage sites, whereby the fusion protein is cleaved intracellularly by a protease produced by the pathogen which is specific for the cleavage sites.

30. A method of epitope mapping which comprises synthesising a family of oligopeptides of predetermined amino acid sequence, which together make up from 7 amino acids to the complete amino acid sequence of a target protein, which comprises: synthesising a nucleic acid construct which codes for a fusion protein composed of overlapping peptides derived from the desired portion of the target protein, interspersed with regions which code for protease cleavage sites, expressing the nucleic acid construct in a suitable expression vector, harvesting the fusion protein corresponding to the nucleic acid sequence, and digesting the fusion protein with a protease selective for the cleavage sites to generate the oligopeptides and using these oligopeptides in an oilgopeptide epitope mapping procedure.

MGGKWSKSSVVGWPAVRERM**IEGR**VGWPAVRERMRRAEPAADGV**IEGR**RRAEPAADGVGAVSRDLEKH**IEGR**GAVSRDLEKHGAITSSNTAA**IEGR**GAITSSNTAATNADCAWLEA**IEGR**TNADCAWLEAQEEEEVGFPV**IEGR**QEEEEVGFPVTPQVPLRPMT**IEGR**TPQVPLRPMTYKAAVDLSHF**IEGR**YKAAVDLSHFLKEKGGLEGL**IEGR**LKEKGGLEGLIHSQRRQDIL**IEGR**IHSQRRQDILDLWIYHTQGY**IEGR**DLWIYHTQGYFPDWQNYTPE**IEGR**FPDWQNYTPEPGVRYPLTFG**IEGR**PGVRYPLTFGWCY

Fig. 1

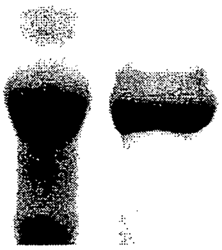


Fig. 2

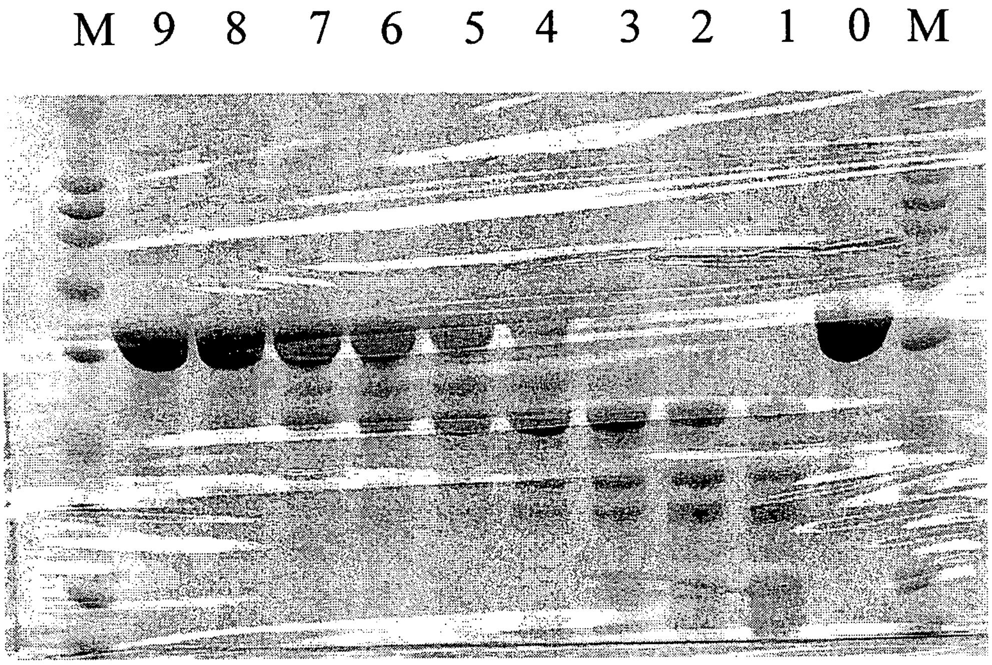


Fig. 3a

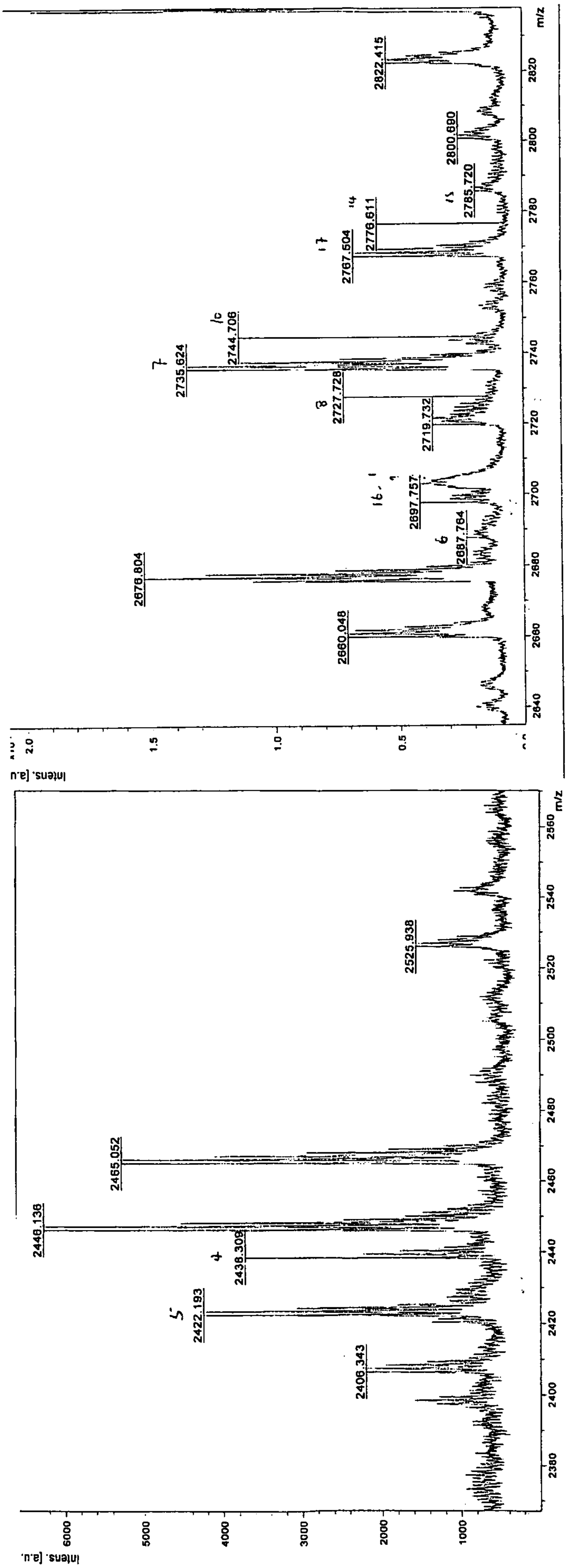
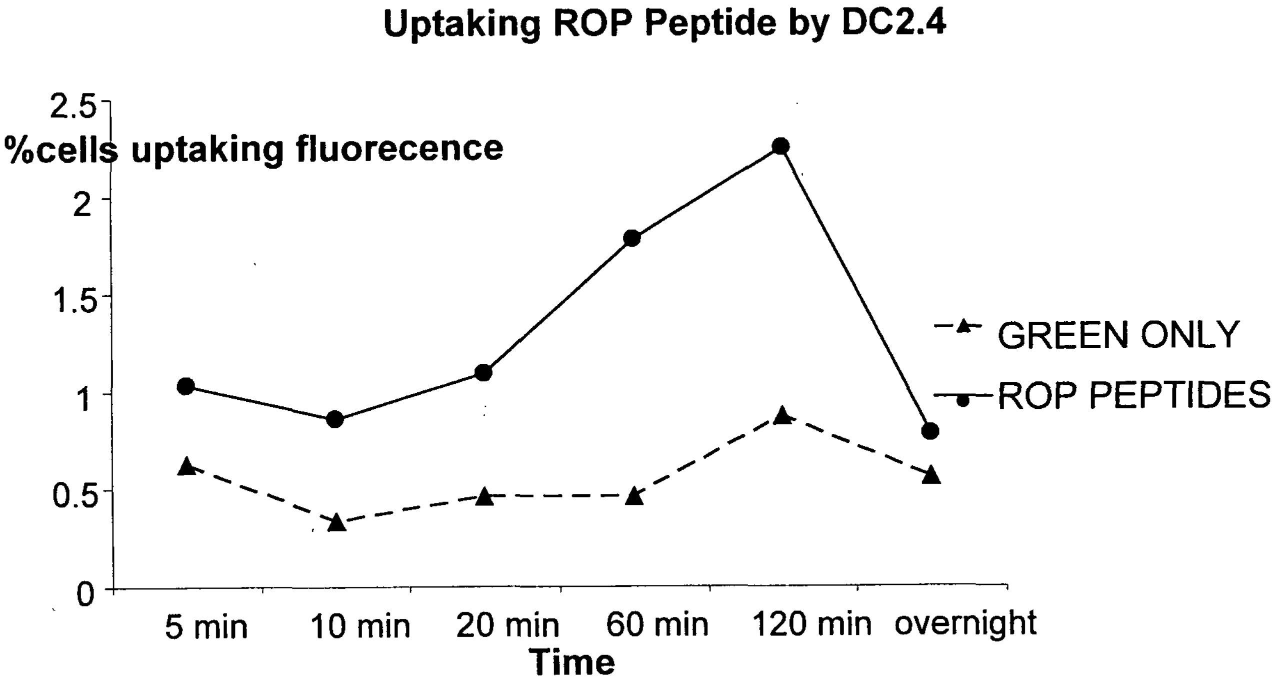

Intens. [a.u.]
6000
5000
4000
3000
2000
1000
2380
2400
2420
2440
2460
2480
2500
2520
2540
2560
m/z
2406.343
2422.193
5
2438.309
4
2446.138
2465.052
2525.938
Intens. [a.u
2.0
1.5
1.0
0.5
2640
2660
2680
2700
2720
2740
2760
2780
2800
2820
m/z
2660.048
2678.804
2687.764
6
2697.757
16
2719.732
2727.728
8
2735.624
7
2744.706
10
2767.504
17
2778.611
14
2785.720
15
2800.690
2822.415

Fig. 3b


Uptaking ROP Peptide by DC2.4
%cells uptaking fluorecence
2.5
2
1.5
1
0.5
0
5 min
10 min
20 min
60 min
120 min
overnight
Time
GREEN ONLY
ROP PEPTIDES

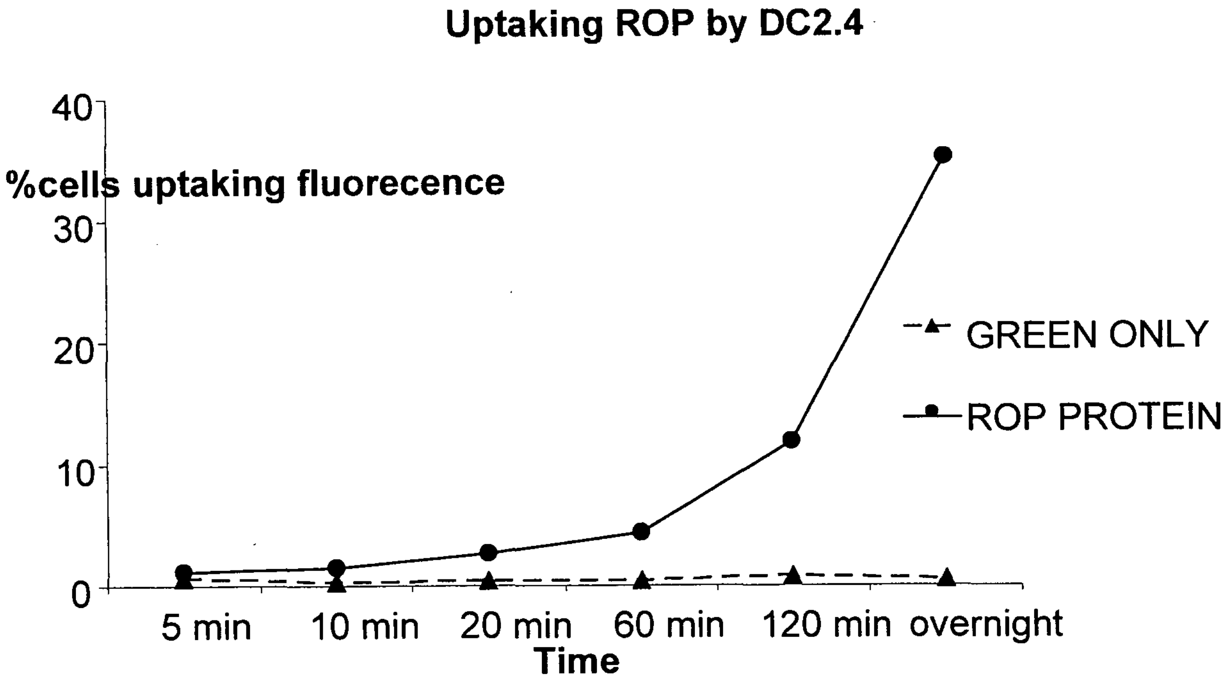

Uptaking ROP by DC2.4
%cells uptaking fluorecence
40
30
20
10
0
5 min
10 min
20 min
60 min
120 min
overnight
Time
GREEN ONLY
ROP PROTEIN

Fig.4

Fig. 5

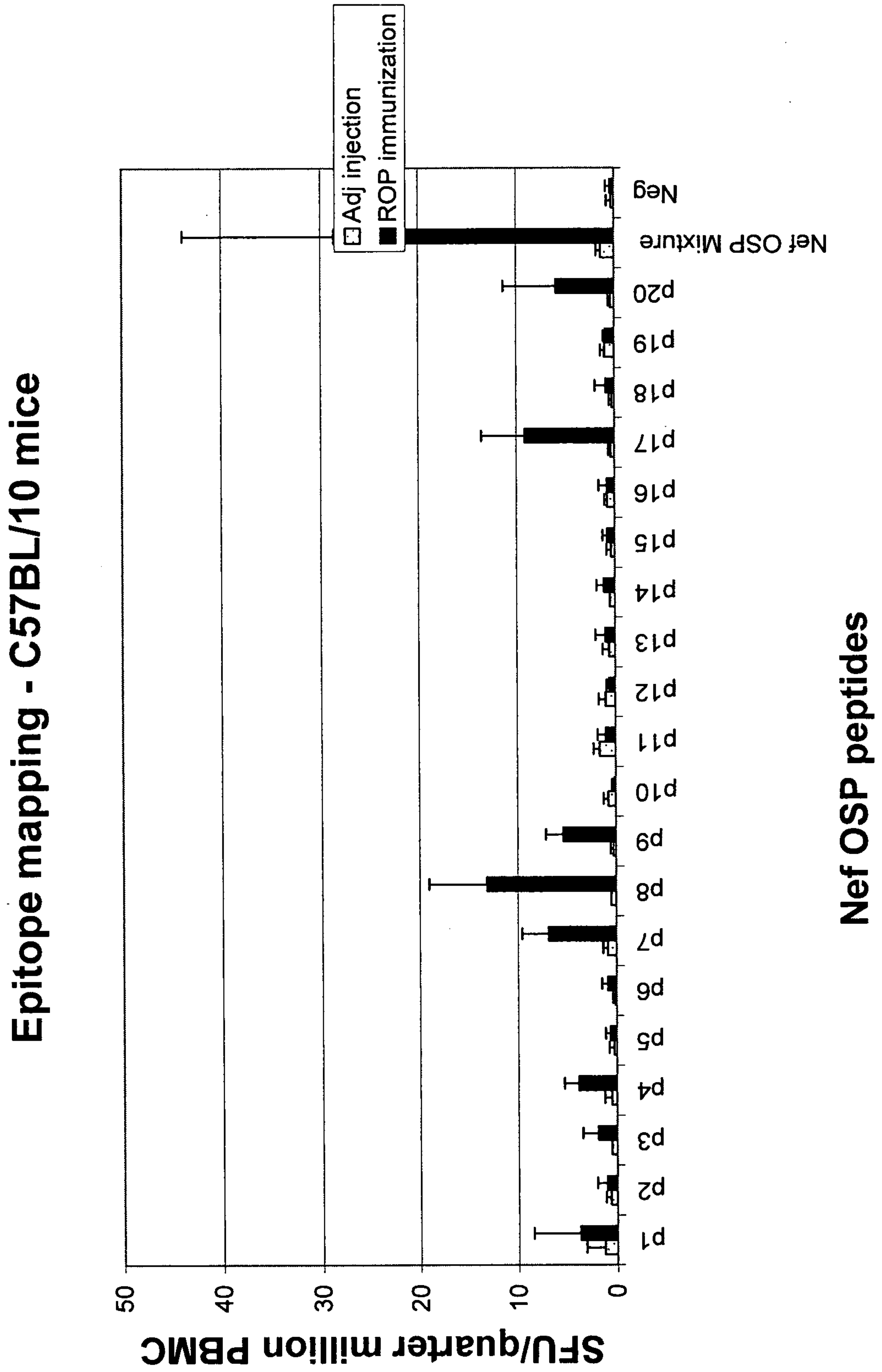

Epitope mapping - C57BL/10 mice
SFU/quarter million PBMC
0
10
20
30
40
50
p1
p2
p3
p4
p5
p6
p7
p8
p9
p10
p11
p12
p13
p14
p15
p16
p17
p18
p19
p20
Nef OSP Mixture
Neg
☐ Adj injection
■ ROP immunization
Nef OSP peptides

Fig. 6

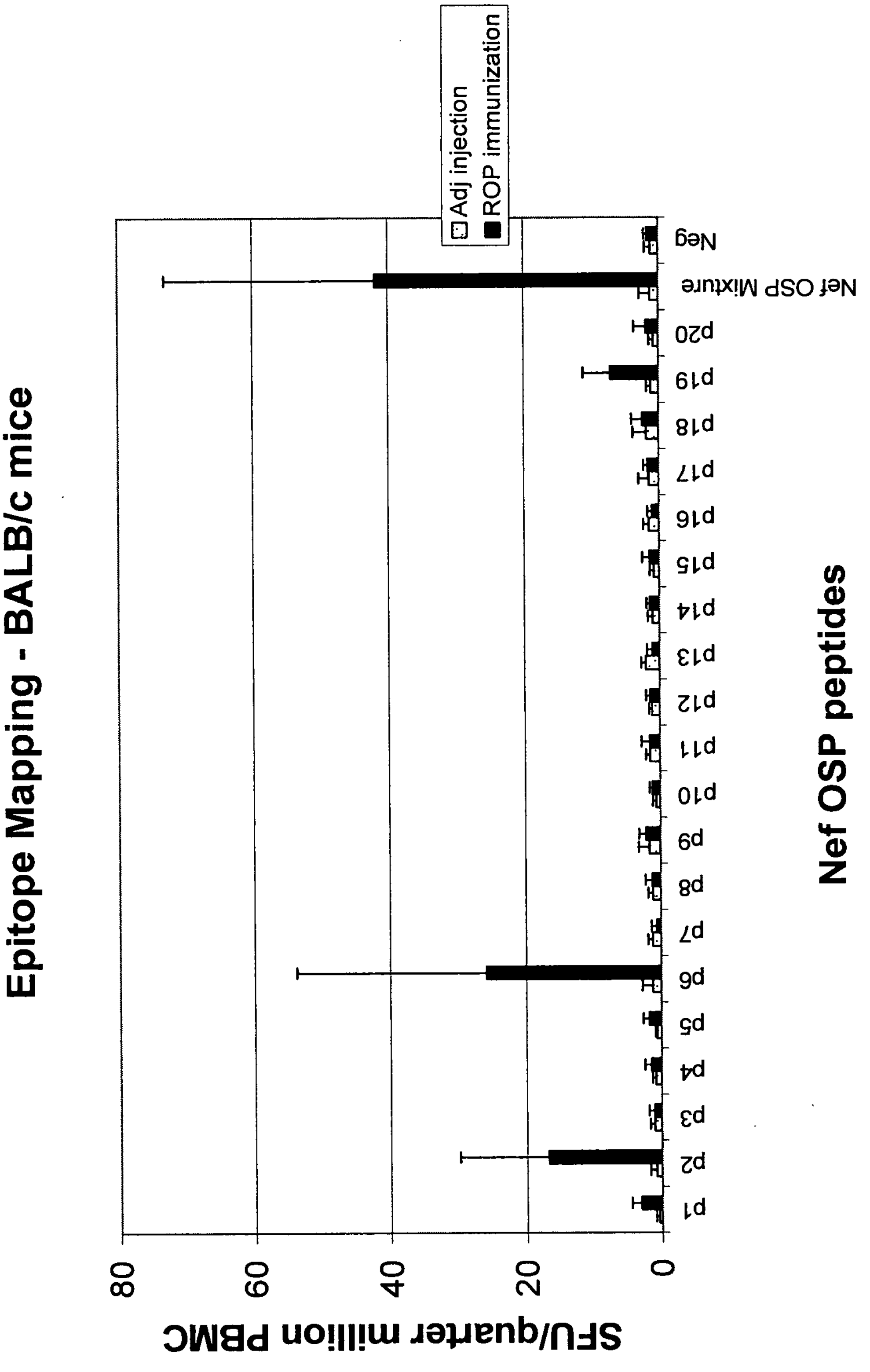


Fig. 6 continued

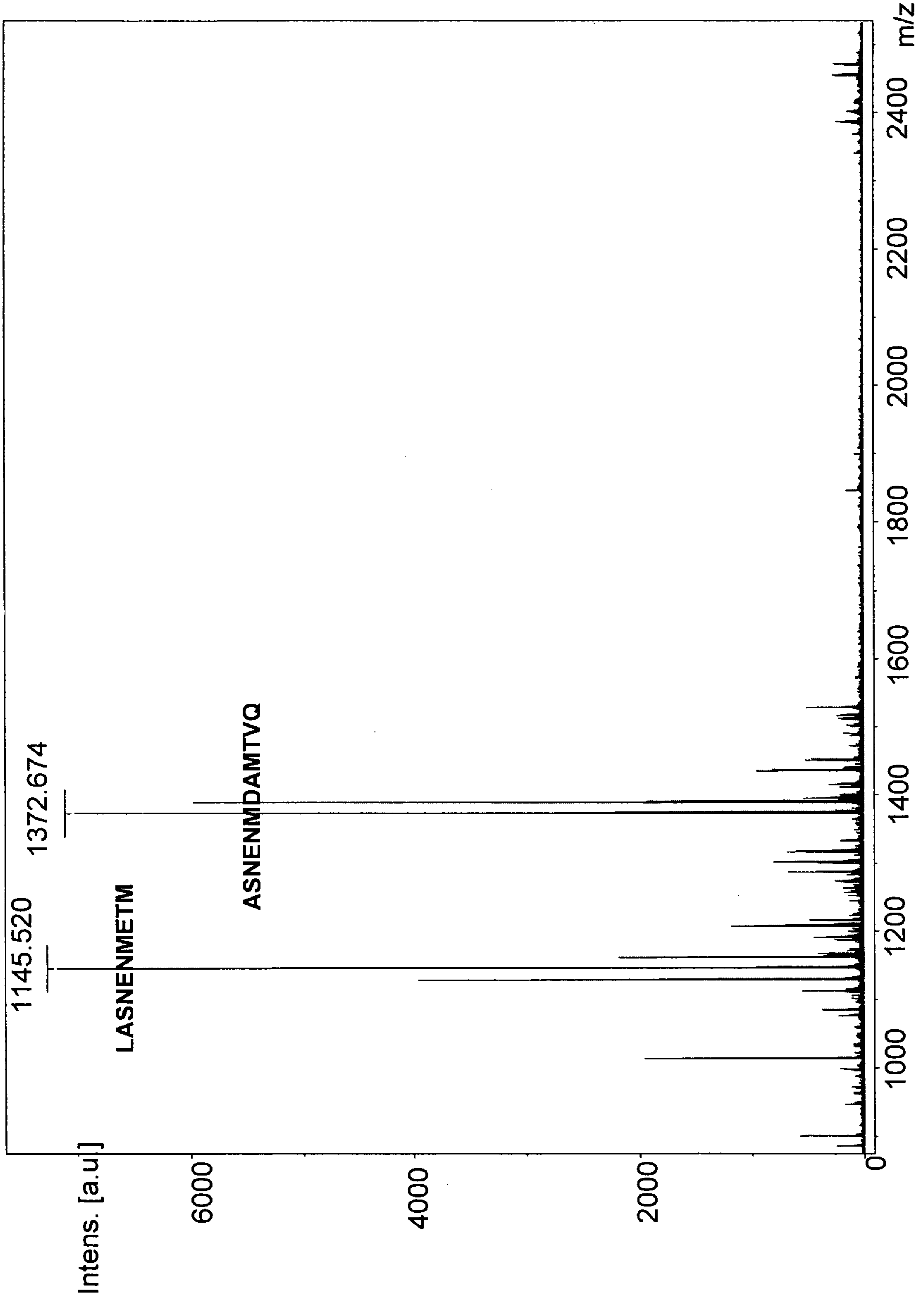

Intens. [a.u.]
6000
4000
2000
0
1145.520
LASNENMETM
1372.674
ASNENMDAMTVQ
1000
1200
1400
1600
1800
2000
2200
2400
m/z

Fig. 7