



(86) Date de dépôt PCT/PCT Filing Date: 2005/02/03
(87) Date publication PCT/PCT Publication Date: 2005/08/25
(45) Date de délivrance/Issue Date: 2013/10/15
(85) Entrée phase nationale/National Entry: 2006/08/01
(86) N° demande PCT/PCT Application No.: EP 2005/001114
(87) N° publication PCT/PCT Publication No.: 2005/077977
(30) Priorités/Priorities: 2004/02/11 (US60/543,649);
2004/12/14 (US60/635,791)

(51) Cl.Int./Int.Cl. *C12N 15/62* (2006.01),
A61K 39/00 (2006.01), *A61P 35/00* (2006.01),
A61P 37/04 (2006.01), *C07K 14/705* (2006.01),
C07K 19/00 (2006.01), *C12N 15/00* (2006.01),
C12N 15/861 (2006.01)

(72) Inventeurs/Inventors:
LA MONICA, NICOLA, IT;
FACCIABENE, ANDREA, IT;
AURISICCHIO, LUIGI, IT;
CILIBERTO, GENNARO, IT

(73) Propriétaire/Owner:
ISTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE
P. ANGELETTI SPA, IT

(74) Agent: NORTON ROSE FULBRIGHT CANADA

(54) Titre : PROTEINES DE FUSION A ANTIGENE CARCINOEMBRYONIQUE ET LEURS UTILISATIONS
(54) Title: CARCINOEMBRYONIC ANTIGEN FUSIONS AND USES THEREOF

(57) **Abrégé/Abstract:**

Polynucleotides encoding carcinoembryonic antigen (CEA) fusion proteins are provided, the CEA fusion proteins comprising a CEA protein, or functional variant thereof, fused to a substantial portion of an immunoenhancing element. The polynucleotides of the present invention can elicit an immune response in a mammal, which, in preferred embodiments, is stronger than the immune response elicited by a wild-type CEA. The gene encoding CEA is commonly associated with the development of human carcinomas. The present invention provides compositions and methods to elicit or enhance immunity to the protein product expressed by the CEA tumor-associated antigen, wherein aberrant CEA expression is associated with a carcinoma or its development. This invention specifically provides adenoviral vector and plasmid constructs carrying polynucleotides encoding CEA fusion proteins and discloses their use in vaccines and pharmaceutical compositions for preventing and treating cancer.



(74) **Agent(suite/continued):** LLP/S.E.N.C.R.L., S.R.L.

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

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
25 August 2005 (25.08.2005)

PCT

(10) International Publication Number
WO 2005/077977 A3

(51) International Patent Classification⁷: **C07K 14/705**,
C12N 15/62, A61K 39/00, A61P 35/00

(21) International Application Number:
PCT/EP2005/001114

(22) International Filing Date: 3 February 2005 (03.02.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/543,649 11 February 2004 (11.02.2004) US
60/635,791 14 December 2004 (14.12.2004) US

(71) Applicant (for all designated States except US): **ISTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P ANGELETTI SPA** [IT/IT]; Via Pontina KM 30.600, I-00040 Pomezia (IT).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **LA MONICA, Nicola** [IT/IT]; IRBM, Via Pontina KM 30.600, I-00040 Pomezia (IT). **FACCIABENE, Andrea** [IT/IT]; IRBM, Via Pontina KM 30.600, I-00040 Pomezia (IT). **AURISICCHIO, Luigi** [IT/IT]; IRBM, Via Pontina KM 30.600, I-00040 Pomezia (IT).

(74) Agent: **BUCHAN, Gavin, MacNicol**; Terlings Park, Eastwick Road, Harlow, Essex CM20 2QR (GB).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, 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, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, 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, MC, 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:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(88) Date of publication of the international search report:
6 October 2005

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: CARCINOEMBRYONIC ANTIGEN FUSIONS PROTEINS AND USES THEREOF

(57) Abstract: Polynucleotides encoding carcinoembryonic antigen (CEA) fusion proteins are provided, the CEA fusion proteins comprising a CEA protein, or functional variant thereof, fused to a substantial portion of an immunoenhancing element. The polynucleotides of the present invention can elicit an immune response in a mammal, which, in preferred embodiments, is stronger than the immune response elicited by a wild-type CEA. The gene encoding CEA is commonly associated with the development of human carcinomas. The present invention provides compositions and methods to elicit or enhance immunity to the protein product expressed by the CEA tumor-associated antigen, wherein aberrant CEA expression is associated with a carcinoma or its development. This invention specifically provides adenoviral vector and plasmid constructs carrying polynucleotides encoding CEA fusion proteins and discloses their use in vaccines and pharmaceutical compositions for preventing and treating cancer.



WO 2005/077977 A3

TITLE OF THE INVENTION

CARCINOEMBRYONIC ANTIGEN FUSIONS AND USES THEREOF

FIELD OF THE INVENTION

5 The present invention relates generally to the therapy of cancer. More specifically, the present invention relates to polynucleotides encoding fusion proteins wherein the fusion proteins comprise at least a portion of the tumor associated polypeptide carcinoembryonic antigen. The present invention also provides recombinant vectors and hosts comprising said polynucleotides, purified fusion proteins and methods for enhancing an immune response against CEA using the
10 compositions and molecules disclosed herein.

BACKGROUND OF THE INVENTION

 The immunoglobulin superfamily (IgSF) consists of numerous genes that code for proteins with diverse functions, one of which is intercellular adhesion. IgSF proteins contain at least
15 one Ig-related domain that is important for maintaining proper intermolecular binding interactions. Because such interactions are necessary to the diverse biological functions of the IgSF members, disruption or aberrant expression of many IgSF adhesion molecules has been correlated with many human diseases.

 The carcinoembryonic antigen (CEA) belongs to a subfamily of the Ig superfamily
20 consisting of cell surface glycoproteins known as CEA-related cell adhesion molecules (CEACAMs). CEACAMs have been shown to act as both homotypic and heterotypic intercellular adhesion molecules (Benchimol et al., *Cell* 57: 327-334 (1989)). In addition to cell adhesion, CEA (also known as CEACAM5) inhibits cell death resulting from detachment of cells from the extracellular matrix and can contribute to cellular transformation associated with certain proto-oncogenes such as
25 *Bcl2* and *C-Myc* (see Berinstein, *J. Clin Oncol.* 20(8): 2197-2207 (2002)). Sequences coding for human CEA have been cloned and characterized (U.S. Patent No. 5,274,087; U.S. Patent No 5,571,710; and U.S. Patent No 5,843,761. See also Beauchemin et al., *Mol. Cell. Biol.* 7:3221-3230 (1987); Zimmerman et al., *Proc. Natl. Acad. Sci. USA* 84:920-924 (1987); Thompson et al. *Proc. Natl. Acad. Sci. USA* 84(9):2965-69 (1987)).

30 Normal expression of CEA has been detected during fetal development and in adult colonic mucosa. CEA overexpression was first detected in human colon tumors over thirty years ago (Gold and Freedman, *J. Exp. Med.* 121:439-462 (1965)) and has since been found in nearly all colorectal tumors. Additionally, CEA overexpression is detectable in a high percentage of adenocarcinomas of the pancreas, liver, breast, ovary, cervix, and lung. Because of its prevalence in

these tumor types and limited normal tissue expression, CEA is considered a self tumor-associated antigen and a target for active and passive immunotherapy. Recent clinical data have established that different vaccine strategies can generate human B and T cells specific for CEA, providing additional evidence that CEA is a target for molecular and immunological intervention for treatment of these cancer types.

Therapeutic approaches targeting CEA include the use of anti-CEA antibodies (*see* Chester et al., *Cancer Chemother. Pharmacol.* 46 (Suppl): S8-S12 (2000)), as well as CEA-based vaccines (for review, *see* Berinstein, *supra*). The development and commercialization of many vaccines have been hindered by difficulties associated with obtaining high expression levels of exogenous genes. Success of DNA-based vaccines has also been hindered by an inability to generate an immune response of sufficient magnitude in treated individuals. Although DNA vaccines targeting various proteins have been developed, the resulting immune responses have been relatively weak compared with conventional vaccines.

The ease of DNA manipulation has offered an opportunity to develop vaccines incorporating gene fusion strategies in which antigens are linked to various immunoenhancing elements. Enhancement of immune response to target antigens has been demonstrated in animal models by vectors encoding antigens fused to heat shock protein (HSP) 70 (Liu et al., *J. Virol.* 74: 2888-94 (2000); Cheng et al. *J. Immunol.* 166: 6218-26 (2001); Chen et al., *Cancer Res.* 60: 1035-42 (2000)), to Fc portion of IgG1 (You et al., *J. Immunol.* 165: 4581-92 (2000)), to lysosome-associated membrane protein (LAMP) (Su et al., *Cancer Res.* 62: 5041-48 (2002)), and universal Th epitope from tetanus toxin (Renard et al., *J. Immunol.* 171:1588-95 (2003); King et al., *Nature Med.* 4: 1281-86 (1998); Lund et al., *Cancer Gene Ther.* 10: 365-76 (2003); Padua et al., *Nature Med.* 9(11): 1413-17 (2003); Savelyeva et al., *Nature Biotechnol.* 19: 760-64 (2001); Wahren et al., WO 2004/092216). The enhancement of immune responses to target antigens is particularly relevant for cancer vaccines in view of the limited immunogenicity of tumor antigens and of the need to overcome tolerance to exert effective antitumor effects.

Therefore, despite the identification of the wild-type nucleotide sequences encoding CEA proteins described above, it would be highly desirable to develop a vaccine which is capable of eliciting an enhanced CEA-specific immune response relative to a wild-type full-length CEA cDNA, when delivered to a mammal. It would also be desirable to develop methods for treating or preventing CEA-associated cancers which utilize nucleic acid molecules or proteins that safely and effectively potentiate a CEA-specific immune response.

SUMMARY OF THE INVENTION

The present invention provides polynucleotides encoding fusion proteins wherein the fusion proteins comprise at least a portion of the tumor associated polypeptide carcinoembryonic antigen, fused to a substantial portion of an immunoenhancing element, such as a bacterial toxin. In preferred embodiments, the CEA portion of the encoded CEA fusion protein is deleted of its C-terminal anchoring domain. In preferred embodiments, the immunoenhancing element is the A or B subunit of the heat labile enterotoxin of *E.coli*, or substantial portion thereof. In other preferred embodiments, the immunoenhancing element is the minimized domain of *tetanus* toxin fragment C (DOM), or substantial portion thereof. The present invention also provides recombinant vectors, including but not limited to, adenovirus and plasmid vectors, comprising said polynucleotides and host cells comprising said recombinant vectors. Also provided herein are purified fusion proteins encoded by invention polynucleotides.

The present invention further provides methods for inhibiting or preventing the development of a cancer in a mammal by eliciting an immune response to the CEA protein by administering a vaccine or pharmaceutical composition comprising the CEA fusions or CEA fusion proteins described herein. In preferred embodiments of the methods herein, the immune response is enhanced relative to the response elicited by a wild-type CEA vaccine.

As used throughout the specification and in the appended claims, the singular forms "a," "an," and "the" include the plural reference unless the context clearly dictates otherwise.

As used throughout the specification and appended claims, the following definitions and abbreviations apply:

The term "promoter" refers to a recognition site on a DNA strand to which the RNA polymerase binds. The promoter forms an initiation complex with RNA polymerase to initiate and drive transcriptional activity. The complex can be modified by activating sequences termed "enhancers" or inhibiting sequences termed "silencers".

The term "cassette" refers to a nucleotide or gene sequence that is to be expressed from a vector, for example, the nucleotide or gene sequence encoding the hCEA-LTB fusion. In general, a cassette comprises a gene sequence that can be inserted into a vector, which in some embodiments, provides regulatory sequences for expressing the nucleotide or gene sequence. In other embodiments, the nucleotide or gene sequence provides the regulatory sequences for its expression. In further embodiments, the vector provides some regulatory sequences and the nucleotide or gene sequence provides other regulatory sequences. For example, the vector can provide a promoter for transcribing the nucleotide or gene sequence and the nucleotide or gene sequence provides a transcription termination sequence. The regulatory sequences that can be provided by the vector

include, but are not limited to, enhancers, transcription termination sequences, splice acceptor and donor sequences, introns, ribosome binding sequences, and poly(A) addition sequences.

The term "vector" refers to some means by which DNA fragments can be introduced into a host organism or host tissue. There are various types of vectors including plasmid, virus
5 (including adenovirus), bacteriophages and cosmids.

The term "first generation," as used in reference to adenoviral vectors, describes adenoviral vectors that are replication-defective. First generation adenovirus vectors typically have a deleted or inactivated E1 gene region, and preferably have a deleted or inactivated E3 gene region.

The abbreviation "DOM" refers generally to the N-terminal domain of fragment C of
10 *tetanus* toxoid.

The abbreviation "LT" refers generally to the heat labile enterotoxin of *E. coli*. "LT" may refer to the complete enterotoxin, comprising subunits A and B or a substantial portion of subunit A, or a substantial portion of subunit B. The abbreviation "LTA" refers to the A subunit of the heat labile enterotoxin of *E. coli*, or substantial portion thereof, including subunits which are
15 truncated on the C-terminal or N-terminal end but maintain biological activity, as well as subunits that contain internal amino acid insertions, deletions, or substitutions but maintain biological activity. The abbreviation "LTB" refers to the B subunit of the heat labile enterotoxin of *E. coli*, or substantial portion thereof, including subunits which are truncated on the C-terminal or N-terminal end but maintain biological activity, as well as subunits that contain internal amino acid insertions, deletions,
20 or substitutions but maintain biological activity.

The designation "pV1J/hCEAopt" refers to a plasmid construct, disclosed herein, comprising the CMV immediate-early (IE) promoter with intron A, a full-length codon-optimized human CEA gene, bovine growth hormone-derived polyadenylation and transcriptional termination sequences, and a minimal pUC backbone (see EXAMPLE 2). The designation "pV1J/hCEA" refers
25 to a construct essentially as described above, except the construct comprises a wild-type human CEA gene instead of a codon-optimized human CEA gene.

The designation "pV1J/hCEA-LTB" refers to a plasmid construct, disclosed herein, comprising the CMV immediate-early (IE) promoter with intron A, a human CEA gene devoid of its GPI anchor coding sequence, fused at its C-terminal end to the B subunit of *E. coli* heat labile
30 enterotoxin, bovine growth hormone-derived polyadenylation and transcriptional termination sequences, and a minimal pUC backbone.

The designation "pV1J/hCEAopt-LTB" refers to a construct essentially as described immediately above, except the construct comprises a codon-optimized human CEA gene devoid of its GPI anchor coding sequence instead of the corresponding portion of the wild-type human CEA gene.

It is provided a nucleic acid molecule comprising a sequence of nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a C-terminally truncated human CEA protein which lacks amino acids 679 - 702 of SEQ ID NO:20, fused to an immunoenhancing element which is the B subunit of the heat labile enterotoxin of E. coli (LTB) optionally truncated of its signal sequence; and wherein the fusion protein is capable of producing an immune response in a mammal.

It is also provided an adenovirus vaccine vector comprising an adenoviral genome with a deletion in the E1 region, and an insert in the E1 region, wherein the insert comprises an expression cassette comprising:

(a) a polynucleotide comprising a sequence of nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a C-terminally truncated human CEA protein which lacks amino acids 679 - 702 of SEQ ID NO:20, fused to an immunoenhancing element which is the B subunit of the heat labile enterotoxin of E. coli (LTB) optionally truncated of its signal sequence; and wherein the fusion protein is capable of producing an immune response in a mammal; and

(b) a promoter operably linked to the polynucleotide.

It is further provided a vaccine plasmid comprising a plasmid portion and an expression cassette portion, the expression cassette portion comprising:

(a) a polynucleotide comprising a sequence of nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a C-terminally truncated human CEA protein which lacks amino acids 679 - 702 of SEQ ID NO:20, fused to an immunoenhancing element which is the B subunit of the heat labile enterotoxin of E. coli (LTB) optionally truncated of its signal sequence; and wherein the fusion protein is capable of producing an immune response in a mammal; and,

(b) a promoter operably linked to the polynucleotide.

The designation "pV1J/hCEAopt-LTBopt" refers to a plasmid construct essentially as described immediately above, except that both the CEA sequences and the LTB sequences are codon-optimized for high level expression in human cells.

5 The designation "pV1J/rhCEAopt-LTBopt" refers to a construct essentially as described above except that the human codon-optimized CEA gene is replaced with a rhesus monkey CEA gene, codon-optimized for high-level expression in human cells.

10 The designation "pV1J/hCEA-LTA" refers to a plasmid construct, disclosed herein, comprising the CMV immediate-early (IE) promoter with intron A, a human CEA gene devoid of the GPI anchor coding sequence, fused at its C-terminal end to the A subunit of *E. coli* heat labile enterotoxin, bovine growth hormone-derived polyadenylation and transcriptional termination sequences, and a minimal pUC backbone. Construction of plasmid vectors comprising various CEA-LT fusions is described in EXAMPLE 2.

15 The designation "pV1J/hCEA-DOM" refers to a plasmid construct, disclosed herein, comprising the CMV immediate-early (IE) promoter with intron A, a human CEA gene devoid of its GPI anchor coding sequence, fused at its C-terminal end to the N-terminal domain of Fragment C of tetanus toxoid (DOM), bovine growth hormone-derived polyadenylation and transcriptional termination sequences, and a minimal pUC backbone (EXAMPLE 2).

20 The designation "pV1J/rhCEAopt-DOMopt" refers to a construct essentially as described above except that the human codon-optimized CEA gene is replaced with a rhesus monkey CEA gene, codon-optimized for high-level expression in human cells.

25 The designation "pV1J/hCEA-FcIgG" refers to a plasmid construct, disclosed herein, comprising the CMV immediate-early (IE) promoter with intron A, a human CEA gene devoid of the GPI anchor coding sequence, fused at its C-terminal end to the heavy fragment of constant chain of immunoglobulin G1, bovine growth hormone-derived polyadenylation and transcriptional termination sequences, and a minimal pUC backbone. (EXAMPLE 2). pV1J/hCEAopt-FcIgGopt refers to a construct essentially as described, except the nucleotide sequences encoding CEA and FcIgG have been codon-optimized for high-level expression in human cells.

30 The designations "Ad5/hCEAopt" and "Ad5/hCEA" refer to two constructs, disclosed herein, which comprise an Ad5 adenoviral genome deleted of the E1 and E3 regions. In the "Ad5/hCEAopt" construct, the E1 region is replaced by a codon-optimized human CEA gene in an E1 parallel orientation under the control of a human CMV promoter without intron A, followed by a bovine growth hormone polyadenylation signal. The "Ad5/hCEA" construct is essentially as described above, except the E1 region of the Ad5 genome is replaced with a wild-type human CEA sequence. The designation "Ad5/hCEAopt-LTB" refers to an Ad5 construct, essentially as described

above, except that the codon-optimized human CEA sequence is devoid of the GPI anchor coding sequence and is fused at its C-terminus to the B subunit of *E. coli* heat labile enterotoxin.

Construction of adenovirus vectors comprising various CEA-LT fusions is described in EXAMPLE 3.

“Immunoenhancing element” refers to a portion of the CEA fusion proteins of the present invention which is capable of stimulating or enhancing the immune response to the associated CEA protein, relative to full-length wild-type CEA. Immunoenhancing elements of the present invention are selected from the group consisting of: heat shock protein (HSP) 70, lysosome-associated membrane protein (LAMP), fragment C of tetanus toxoid (FrC), the N-terminal domain of FrC (DOM), the heavy fragment of constant chain of immune globulin G1 (FcIgG), the vesicular stomatitis virus glycoprotein (VSV-G), cholera toxin (CT) from *Vibrio cholerae*, and heat labile enterotoxin of *E. coli* (LT). The term “immunoenhancing element” is used interchangeably herein with the term “adjuvant.”

As used herein, a “fusion protein” refers to a protein having at least two polypeptides covalently linked in which one polypeptide comes from one protein sequence or domain and the other polypeptide comes from a second protein sequence or domain. The fusion proteins of the present invention comprise a CEA polypeptide or fragment or variant thereof, and a second polypeptide, which comprises a substantial portion of an immunoenhancing element, which, in some cases, is a bacterial toxin. The CEA polypeptide, fragment or variant thereof may be a human CEA or CEA homolog from another species. The polypeptides that comprise the fusion protein are preferably linked N-terminus to C-terminus. The CEA polypeptide and the toxin subunit can be fused in any order. In some embodiments of this invention, the C-terminus of the CEA polypeptide is fused to the N-terminus of the toxin subunit, as exemplified in FIGURE 1A. However, fusion proteins in which the immunoenhancing element is fused to the N-terminus of the CEA polypeptide are also contemplated. The term “CEA fusion protein” is intended to be a general term which refers to a fusion as described above, which comprises a CEA polypeptide or fragment or variant thereof fused to a polypeptide comprising an immunoenhancing element.

The term “CEA-LT fusion” refers to a nucleic acid sequence in which at least a portion of the CEA gene is fused to a substantial portion of either the LTA or the LTB subunit of *E. coli* heat labile enterotoxin. The term “CEA-LT fusion protein” refers to a polypeptide encoded by a CEA-LT fusion as described. The terms “CEA-LT fusion” and “CEA-LT fusion protein” are also understood to refer to fragments thereof, homologs thereof, and functional equivalents thereof (collectively referred to as “variants”), such as those in which one or more amino acids is inserted, deleted or replaced by other amino acid(s). The CEA-LT fusions of the present invention, upon administration to a mammal such as a human being, can stimulate an immune response by helper T

cells or cytotoxic T cells, or stimulate the production of antibodies at least as well as a "wild-type" CEA sequence. In preferred embodiments of the invention, the CEA-LT fusion can enhance the immune response as compared to a wild-type CEA.

5 The term "CEA-DOM fusion" refers to a nucleic acid sequence in which at least a portion of the CEA gene is fused to a substantial portion of the minimized domain of tetanus toxin fragment C, unless the context clearly dictates that said term refers to the protein sequence. The term "CEA-DOM fusion protein" refers to a polypeptide encoded by a CEA-DOM fusion as described. The terms "CEA-DOM fusion" and "CEA-DOM fusion protein" are also understood to refer to fragments thereof, homologs thereof, and functional equivalents thereof (collectively referred to as
10 "variants"), such as those in which one or more amino acids is inserted, deleted or replaced by other amino acid(s). The CEA-DOM fusions of the present invention, upon administration to a mammal such as a human being, can stimulate an immune response by helper T cells or cytotoxic T cells, or stimulate the production of antibodies at least as well as a "wild-type" CEA sequence. In preferred embodiments of the invention, the CEA-DOM fusion can enhance the immune response as compared
15 to a wild-type CEA.

The abbreviation "AD" refers to the anchoring domain of a CEA gene or protein. The anchoring domain of the wild-type human CEA is located from about amino acid 679 to about amino acid 702 of SEQ ID NO:20.

20 The term "treatment" refers to both therapeutic treatment and prophylactic or preventative measures. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in which the disorder is to be prevented.

A "disorder" is any condition that would benefit from treatment with the molecules of the present invention, including the nucleic acid molecules described herein and the fusion proteins that are encoded by said nucleic acid molecules. Encompassed by the term "disorder" are chronic and
25 acute disorders or diseases including those pathological conditions which predispose the mammal to the disorder in question. The molecules of the present invention are intended for use as treatments for disorders or conditions characterized by aberrant cell proliferation, including, but not limited to, breast cancer, colorectal cancer, and lung cancer.

30 The term "effective amount" means sufficient vaccine composition is introduced to produce the adequate levels of the polypeptide, so that an immune response results. One skilled in the art recognizes that this level may vary.

A "conservative amino acid substitution" refers to the replacement of one amino acid residue by another, chemically similar, amino acid residue. Examples of such conservative substitutions are: substitution of one hydrophobic residue (isoleucine, leucine, valine, or methionine)

for another; substitution of one polar residue for another polar residue of the same charge (e.g., arginine for lysine; glutamic acid for aspartic acid).

“hCEA” and “hCEAopt” refer to a human carcinoembryonic antigen and a human codon-optimized carcinoembryonic antigen, respectively.

5 “rhCEA” and “rhCEAopt” refer to a rhesus monkey carcinoembryonic antigen and a rhesus monkey codon-optimized carcinoembryonic antigen, respectively.

“Substantially similar” means that a given nucleic acid or amino acid sequence shares at least 75%, preferably 85%, more preferably 90%, and even more preferably 95% identity with a reference sequence. In the present invention, the reference sequence can be relevant portions of the wild-type human CEA nucleotide or amino acid sequence, or the wild-type nucleotide or amino acid sequence of a bacterial toxin or subunit thereof, such as the LTB or LTA subunits of the *E.coli* heat labile enterotoxin, as dictated by the context of the text. The reference sequence may also be, for example, the wild-type rhesus monkey CEA sequence. Thus, a CEA protein sequence that is “substantially similar” to the wild-type human CEA protein or fragment thereof will share at least 15 75% identity with the relevant fragment of the wild-type human CEA, along the length of the fragment, preferably 85% identity, more preferably 90% identity and even more preferably 95% identity. Whether a given CEA, LTB, or LTA protein or nucleotide sequence is “substantially similar” to a reference sequence can be determined for example, by comparing sequence information using sequence analysis software such as the GAP computer program, version 6.0, available from the 20 University of Wisconsin Genetics Computer Group (UWGCG). The GAP program utilizes the alignment method of Needleman and Wunsch (*J. Mol. Biol.* 48:443, 1970), as revised by Smith and Waterman (*Adv. Appl. Math.* 2:482, 1981).

A “substantial portion” of a gene, variant, fragment, or subunit thereof, means a portion of at least 50%, preferably 75%, more preferably 90%, and even more preferably 95% of a 25 reference sequence.

A “gene” refers to a nucleic acid molecule whose nucleotide sequence codes for a polypeptide molecule. Genes may be uninterrupted sequences of nucleotides or they may include such intervening segments as introns, promoter regions, splicing sites and repetitive sequences. A gene can be either RNA or DNA. A preferred gene is one that encodes the invention peptide.

30 The term “nucleic acid” or “nucleic acid molecule” is intended for ribonucleic acid (RNA) or deoxyribonucleic acid (DNA), probes, oligonucleotides, fragment or portions thereof, and primers. DNA can be either complementary DNA (cDNA) or genomic DNA, e.g. a gene encoding a CEA fusion protein.

“Wild-type CEA” or “wild-type protein” or “wt protein” refers to a protein comprising a naturally occurring sequence of amino acids or variant thereof. The amino acid sequence of wild-type human CEA is shown in FIGURE 7E (SEQ ID NO:20). The amino acid sequence of the wild-type rhesus monkey CEA was previously described (WO 2004/072287, see FIGURES 7A-7B).

“Wild-type CEA gene” refers to a gene comprising a sequence of nucleotides that encodes a naturally occurring CEA protein, including proteins of human origin or proteins obtained from another organism, including, but not limited to, other mammals such as rat, mouse and rhesus monkey. The nucleotide sequence of the human CEA gene is available in the art (*supra*). See also Beauchemin et al., *Mol. Cell. Biol.* 7:3221-3230 (1987); Zimmerman et al., *Proc. Natl. Acad. Sci. USA* 84:920-924 (1987); and Thompson et al. *Proc. Natl. Acad. Sci. USA* 84(9):2965-69 (1987). The nucleotide sequence of the wild-type rhesus monkey gene is shown in FIGURES 7C-7D.

The term “mammalian” refers to any mammal, including a human being.

The abbreviation “Ag” refers to an antigen.

The abbreviations “Ab” and “mAb” refer to an antibody and a monoclonal antibody, respectively.

The abbreviation “ORF” refers to the open reading frame of a gene.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1 shows a schematic representation of the vectors developed in this study. The essential features of the plasmid and Ad vectors encoding the CEA-LTA and CEA-LTB fusions are indicated. The inverted terminal repeats (ITR) of the Ad5 genome are also shown.

FIGURE 2 shows the nucleotide (SEQ ID NO:7, Panel A) and amino acid sequence (SEQ ID NO:8, Panel B) of an exemplary hCEA-LTA fusion. The LTA nucleotide sequence is shown in bold.

FIGURE 3 shows the nucleotide (SEQ ID NO:9, Panel A) and amino acid sequence (SEQ ID NO:10, Panel B) of an exemplary hCEA-LTB fusion. The LTB nucleotide sequence is shown in bold.

FIGURE 4 shows the nucleotide sequence of an exemplary hCEAopt-LTB fusion (SEQ ID NO:11). The LTB nucleotide sequence is shown in bold.

FIGURE 5 shows the nucleotide (SEQ ID NO:12, Panel A) and amino acid sequence (SEQ ID NO:13, Panel B) of an exemplary fully optimized hCEA-LTB fusion, designated herein hCEAopt-LTBopt. The LTB nucleotide and amino acid sequences are shown in bold. Junction

sequences, created by the cloning strategy employed to fuse the CEA and LTB sequences are underlined.

FIGURE 6 shows the nucleotide (SEQ ID NO:14, Panel A) and amino acid sequence (SEQ ID NO:15, Panel B) of a fully optimized rhesus monkey CEA-LTB fusion, designated herein rhCEAoptLTBopt. LTB nucleotide and amino acid sequences are shown in bold. Junction sequences, created by the cloning strategy employed to fuse the CEA and LTB sequences, are underlined.

FIGURE 7 shows nucleotide sequences of wild-type genes encoding rhesus monkey CEA (Panels A and B, SEQ ID NOs:16 and 17) and the amino acid sequences of the corresponding proteins (Panels C and D, SEQ ID NOs:18 and 19), as previously described (U.S.S.N. 60/447,203). Panel E shows the amino acid sequence of wild-type human CEA (SEQ ID NO:20), which was previously described (*see, e.g.*, U.S. Patent No. 5,274,087).

FIGURE 8 shows a comparison of CEA expression efficiency in cells transfected with different CEA constructs. Panel A depicts the expression efficiencies of HeLa cells transfected with 3 µg of plasmids carrying the wild type sequences of hCEA, hCEA-LTA, and hCEA-LTB, in conjunction with 0.2 µg of plasmid pV1J/mEPO as tracer. Panel B shows results from a similar transfection experiment using pV1J/hCEAopt and pV1J/hCEAopt-LTB. Expression efficiency was determined three days post-transfection by measuring the amount of CEA protein present in cell extracts and by normalizing this value for EPO expression. Data shown relates to the average CEA expression values of two independent transfections.

FIGURE 9 shows a comparison of the expression efficiency of different Adenovirus recombinant vectors expressing CEA. HeLa cells were infected at an moi of 100 and 1000 with Ad/hCEAopt and Ad/hCEAopt-LTB. Expression efficiency was determined by measuring three days post infection the amount of CEA protein released in cell extracts. Data shown reflects the average CEA expression values of two independent infections.

FIGURE 10 shows an analysis of the cell mediated immune response elicited by different plasmid vectors encoding human CEA. Three groups of C57BL/6 mice were electroinjected intramuscularly with 50µg of the indicated plasmid (CEA, CEA-LTA fusion or CEA-LTB fusion) at 0 and 3 weeks. A fourth group of mice was immunized with a mixture of 25 µg of pV1J/hCEA-LTA and 25 µg of pV1J/hCEA-LTB. Panel A. Two weeks post boost, the number of IFNγ-secreting T cells specific for CEA was determined by ELISPOT assay on splenocytes from individual mice (empty circles) using peptide pools that encompass the entire protein. Geometric mean values (filled diamonds) are also indicated. Panel B depicts results of IFNγ intracellular staining of pooled

splenocytes from immunized mice using peptide pool D. The nonspecific IFN γ production (DMSO) is shown for each group.

FIGURE 11 shows antibody titers from mice immunized with plasmid DNA vectors encoding CEA. Individual titers against purified human CEA protein were measured by ELISA on
5 serum from individual mice immunized with plasmids pV1J/hCEA, pV1J/hCEA-LTA and pV1J/hCEA-LTB. Average values are also shown (filled diamonds).

FIGURE 12 shows an analysis of the cell mediated immune response elicited by different plasmid vectors encoding CEA. Groups of 4 BALB/c mice were electroinjected with the indicated plasmid as indicated above (FIGURE 4). Two weeks after the last injection, the number of
10 IFN γ secreting T cells specific for CEA was determined by ELISPOT assay on splenocytes from individual mice (empty circles) using peptide pools that encompass the entire protein. Average values (filled diamonds) are also indicated.

FIGURE 13 shows an analysis of the CEA-specific CD8 $^{+}$ T cell response elicited by different plasmid vectors encoding CEA. C57/DR4 mice were electroinjected with the indicated
15 plasmid as described above (see FIGURE 4). Two weeks after the last injection, IFN γ intracellular staining of pooled splenocytes from immunized mice was performed using peptide pool D. The nonspecific IFN γ production (DMSO) is shown for each group.

FIGURE 14 shows an analysis of CEA-specific CD8 $^{+}$ T cell response elicited by different plasmid vectors encoding CEA. HHD mice were electroinjected with the indicated plasmid
20 as described above (see FIGURE 4). Two weeks after the last injection, IFN γ intracellular staining of pooled splenocytes from immunized mice was performed using peptide pools B and D. The nonspecific IFN γ production (DMSO) is shown for each group.

FIGURE 15 shows the cell-mediated and humoral immune response of CEA transgenic mice (N=9) immunized with 5 weekly electroinjections of the indicated plasmids. A total
25 amount of 50 μ g of plasmid DNA was injected i.m. at each vaccination. Panel A. Two weeks after the last injection, the number of IFN γ secreting T cells specific for CEA was determined by intracellular staining on splenocytes from individual mice (circles) using peptide pool D. Geometric mean values (triangles) are also indicated. Panel B. Individual titers against purified human CEA protein were measured by ELISA on each serum from mice immunized with plasmids pV1J/hCEAopt and pV1J/hCEA-LTB. Geometric mean values are also shown (filled diamonds). These data indicate
30 that the CEA-LTB fusion breaks tolerance to CEA in transgenic mice.

FIGURE 16 shows an analysis of the CEA-specific CD8 $^{+}$ T cell response elicited by different Adenovirus vectors encoding CEA. CEA transgenic mice were immunized with different doses of Ad/hCEAopt and Ad/CEAopt-LTB at 0 and 2 weeks. Two weeks after the last injection,

IFN γ intracellular staining of PBMC from each immunized mouse was performed using peptide pool D (filled circles). Geometric mean values are also shown (filled diamonds). The nonspecific IFN γ production (DMSO) of each injected group was less or equal to 0.01%.

FIGURE 17 shows the results of tumor protection studies of immunized CEA transgenic mice challenged with MC38-CEA cells. Groups of 10 CEA transgenic mice were immunized with 5 weekly electroinjections of the indicated plasmid DNA (50 μ g/injection). Two weeks after the last DNA injection, mice were boosted with a single injection of 1×10^{10} vp of the corresponding Ad vector. Fourteen days after the Adenovirus boost, mice were challenged with a subcutaneous injection of 5×10^5 MC38-CEA cells. Panel A shows the percentage of tumor free mice at the indicated timepoint. Panel B reports the average tumor volumes of each immunized group. These data demonstrate that immunization of CEA transgenic mice with CEA-LTB protects mice from tumor development

FIGURE 18. Panel A shows a schematic representation of representative CEA fusion proteins used in this study. Vectors expressing the CEA fusion proteins were derived from plasmid pV1Jns as described in EXAMPLE 2. The constructs comprise a CEA nucleotide sequence from nt 1 to nt 2037 with a net deletion of 64 aa corresponding to the GPI anchoring sequence and express CEA from aa 1 to aa 679. The sequence coordinates of each protein fused to CEA are also indicated. Panel B shows expression of pV1J-derived constructs in transfected cells. HeLa cells were transfected with plasmids pV1J/CEA-VSV-G, pV1J/CEA-FcIgG, pV1J/CEA-DOM, pV1J/CEA-HSP70, pV1J/CEA-LAMP, or pV1J/CEA and processed for Western blot analysis as described in EXAMPLE 5. The specificity of the antibody used for Western blot is indicated. The CEA protein is indicated (black arrow). The positions of molecular size standards (in kilodaltons) are also shown.

FIGURE 19 shows a comparison of expression efficiency of the CEA fusion constructs. HeLa cells were transfected with the indicated plasmids and CEA derived protein present in cell lysates (A) and supernatants (B) was measured by ELISA as described in EXAMPLE 8. Results obtained are representative of two independent experiments.

FIGURE 20 shows a comparison of immunogenicity of different constructs encoding CEA fusion proteins. C57BL/6 mice were electroporated intramuscularly with a 5 or 50 μ g/dose of the indicated plasmids. Injections were carried out at days 0 and 14. Panel A The number of IFN γ -secreting T cells in PBMC in each individual mouse was determined using a pool of peptides covering aa 497-703 (pool D) as described in EXAMPLES 6 and 15. Average number of IFN γ -secreting T cells are also shown (filled circles). SFC values of the pV1J/CEA-DOM, and pV1J/CEA-FcIgG are significantly different from those of pV1J/CEA. Panel B. Antibody titer was measured by ELISA using purified CEA as substrate. Average values of each cohort immunized with 50 μ g dose

of the indicated plasmid are shown. Titers that are significantly different from those of mice injected with pV1J/CEA are indicated with an asterisk.

FIGURE 21 shows the induction of CEA-specific immune responses in CEA transgenic mice. Groups of 12 CEA transgenic mice were immunized with plasmid DNA (50 μ g/dose electroinjected in the quadriceps muscle) or Adenovirus vectors (10⁹ vp/dose) carrying the codon usage optimized cDNA of CEA, CEA-DOM or CEA-FcIgG. CEA-specific CD8⁺ T cells elicited by the DNA/DNA (A) and Ad/Ad (C) immunization regimen were measured by intracellular IFN γ staining on PBMC of each immunized mouse. The average values for each cohort are also shown (filled circle). The CEA-DOM and CEA-FcIgG cohorts immunized with DNA/DNA and Ad/Ad regimens were significantly different from the CEA vaccinated group. CEA-specific antibody titers of each individual mouse vaccinated with the DNA/DNA (B) or Ad/Ad (D) immunization regimen were measured by ELISA. Titers elicited by CEA-DOM and CEA-FcIgG vectors were significantly different from those elicited by CEA.

FIGURE 22 shows the immunogenicity of the DNA/Ad regimen. Groups of 12 CEA transgenic mice were immunized with plasmid DNA (50 μ g/dose) and Adenovirus vectors (10⁹ vp/dose) carrying the codon usage optimized cDNA of CEA, CEA-DOM or CEA-FcIgG. CEA-specific CD8⁺ T cells were measured by intracellular IFN γ staining on PBMC of each immunized mouse (A). The average values for each cohort are also shown (filled circle). The CEA-DOM and CEA-FcIgG cohorts were significantly different from the CEA vaccinated group. CEA-specific antibody titers of each individual mouse were measured by ELISA (B). Titers elicited by CEA-DOM and CEA-FcIgG vectors were significantly different from those elicited by CEA. Average values are shown (filled circles).

FIGURE 23 shows the detection of CD4⁺ T cell response to tetanus toxoid protein. CEA transgenic mice were immunized with pV1J/CEA-DOMopt as described in EXAMPLE 16. IFN γ intracellular staining on pooled PBMC from immunized mice was performed with peptide p30. Whole lymphocyte gating and gating for CD8⁺ (R3) and CD4⁺ T cells (R4) are shown.

FIGURE 24 shows the antitumor effect of vaccination with vectors carrying codon optimized cDNA of CEA, CEA-DOM or CEA-FcIgG. Groups of 10 CEA transgenic mice were immunized with DNA/DNA (A), Ad/Ad (B) and DNA/Ad (C) vaccination regimens using plasmid DNA and Ad vectors carrying the codon usage optimized cDNAs of CEA, CEA-DOM or CEA-FcIgG, as described in EXAMPLE 18. Two weeks after the last injection, mice were challenged with sc inoculation of 5x10⁵ MC38-CEA tumor cells. Percentage of tumor free mice in the vaccinated groups was determined at weekly intervals and compared to that of untreated controls.

Mice vaccinated with CEA-DOM vectors (DNA/Ad modality) was significantly different from control mice (log rank test $p < 0.05$).

FIGURE 25 shows the effect of CD4, CD8, or NK depletion on the induction of anti-tumor effect induced by CEA-DOM DNA/Ad immunization. CEA transgenic mice were immunized with repeated weekly injections of 50 μ g of pV1J/CEA-DOMopt followed by a boost with 1×10^9 vp of Ad-CEA-DOMopt (EXAMPLE 19). One week after the last injection, mice were either not depleted, or were depleted of CD4⁺ T cells, CD8⁺ T cells, or NK cells. Two weeks after the last immunization, mice were challenged with sc inoculation of 5×10^5 MC38-CEA tumor cells. Percentage of tumor free mice in the vaccinated groups was determined at weekly intervals and compared to that of untreated controls. The data indicate that the percentage of tumor-free mice in the vaccinated group was significantly different from untreated controls and depleted cohorts.

FIGURE 26 shows the nucleotide sequence (SEQ ID NO:21) of an exemplary, fully optimized hCEA-DOM fusion, herein designated hCEAoptDOMopt. The amino acid sequence of the encoded protein is also shown (SEQ ID NO:45). The CEA portion of the nucleotide sequence of this particular CEA fusion consists of nucleotides 1 to 2037, which are codon-optimized for high-level expression in a human host cell. The DOM portion of the nucleotide sequence is shown in bold and is also codon-optimized for high-level expression in human cells. Junction sequences, created by the cloning strategy employed to fuse the CEA and LTB sequences are underlined.

FIGURE 27 shows an exemplary nucleotide (SEQ ID NO:25) sequence of a hCEA-FcIgGopt fusion, herein designated hCEAoptFcIgGopt. The sequence of the encoded protein (SEQ ID NO:46) is also shown. The CEA portion of the nucleotide sequence of this particular CEA fusion consists of nucleotides 1 to 2037, which are codon-optimized for high-level expression in a human host cell. The FcIgG portion of the nucleotide sequence, which is also codon-optimized for high-level expression in human cells, is shown in bold. Junction sequences, created by the cloning strategy employed to fuse the CEA and LTB sequences are underlined.

FIGURE 28 shows the nucleotide sequence of a portion of the wild-type human CEA cDNA from nt 1 to nt 2037 (SEQ ID NO:22, Panel A), encoding a portion of the hCEA protein from aa 1 to aa 679 (SEQ ID NO:23, Panel B).

FIGURE 29 shows the non-optimized nucleotide sequence of the minimized domain of tetanus toxin fragment C (DOM) cDNA from nt 1 to nt 825 (SEQ ID NO:47), encoding the DOM protein, also shown (SEQ ID NO:48).

FIGURE 30 shows the non-optimized nucleotide sequence of an exemplary human CEA-DOM fusion (SEQ ID NO:49). The CEA portion of the nucleotide sequence of this particular

CEA fusion consists of nucleotides 1 to 2037. The DOM portion of the nucleotide sequence is shown in bold.

FIGURE 31 shows an exemplary nucleotide sequence (SEQ ID NO:50) of a rhesus monkey CEA-DOM fusion, herein designated rhCEA-DOMopt. The sequence of the encoded fusion protein (SEQ ID NO:51) is also shown. The CEA portion of the nucleotide sequence of this particular CEA fusion consists of nucleotides 1 to 2037, which are codon-optimized for high-level expression in a human host cell. The DOM portion of the nucleotide sequence, which is also codon-optimized for high-level expression in human cells, is shown in bold.

FIGURE 32 shows an exemplary nucleotide sequence (SEQ ID NO:52) of a rhesus monkey CEA- CTB fusion, herein designated rhCEA-CTBopt. The sequence of the encoded fusion protein (SEQ ID NO:53) is also shown. The CEA portion of the nucleotide sequence of this particular CEA fusion consists of nucleotides 1 to 2037, which are codon-optimized for high-level expression in a human host cell. The CTB portion of the nucleotide sequence, which is also codon-optimized for high-level expression in human cells, is shown in bold.

DETAILED DESCRIPTION OF THE INVENTION

Carcinoembryonic antigen (CEA) is commonly associated with the development of adenocarcinomas. The present invention relates to compositions and methods to elicit or enhance immunity to the protein product expressed by the CEA tumor-associated antigen, wherein aberrant CEA expression is associated with the carcinoma or its development. Association of aberrant CEA expression with a carcinoma does not require that the CEA protein be expressed in tumor tissue at all timepoints of its development, as abnormal CEA expression may be present at tumor initiation and not be detectable late into tumor progression or vice-versa.

To this end, the present invention provides polynucleotides, vectors, host cells, and encoded proteins comprising a CEA sequence or variant thereof for use in vaccines and pharmaceutical compositions for the treatment and/or prevention of a cancer. The polynucleotides of the present invention comprise a nucleotide sequence encoding a CEA protein or variant thereof, fused to a nucleotide sequence encoding at least a subunit of an immunoenhancing element, such as a bacterial enterotoxin or substantial portion thereof, which can effectively adjuvant an immune response to the associated CEA.

The CEA nucleotide sequences of the present invention can be of human origin or can be a CEA homolog from another species. The wild-type human CEA nucleotide sequence has been reported (*see, e.g.*, U.S. Patent No. 5,274,087; U.S. Patent No 5,571,710; and U.S. Patent No 5,843,761). The rhesus monkey CEA sequence was recently described (WO 2004/072287). The

CEA portion of the CEA fusion may be full-length, or any variant sufficient to elicit a CEA-specific immune response in a mammal. CEA variants of the present invention include, but are not limited to sequences that are C- or N-terminally truncated, sequences with conservative substitutions, and sequences with internal deletions or insertions.

5 In preferred embodiments of the present invention, the CEA portion of the CEA fusion is human CEA or a functional equivalent thereof. In other preferred embodiments, the CEA portion is a rhesus monkey CEA, or functional equivalent thereof.

Accordingly, the present invention relates to a synthetic polynucleotide comprising a sequence of nucleotides encoding a CEA fusion protein, said fusion protein comprising a CEA
10 protein or a biologically active fragment or mutant form of a CEA protein fused to an immunoenhancing element or subunit thereof, which can effectively enhance the immune response to the CEA protein. Said mutant forms of the CEA protein include, but are not limited to: conservative amino acid substitutions, amino-terminal truncations, carboxy-terminal truncations, deletions, or additions. Any such biologically active fragment and/or mutant will encode either a protein or protein
15 fragment which at least substantially mimics the immunological properties of the CEA protein as set forth in SEQ ID NO:20. The synthetic polynucleotides of the present invention encode mRNA molecules that express a functional CEA fusion protein so as to be useful in the development of a therapeutic or prophylactic cancer vaccine.

In preferred embodiments of the present invention, the CEA portion of the encoded
20 CEA fusion protein is human CEA (SEQ ID NO: 20) or a functional equivalent thereof, for example, a human CEA deleted of its C-terminal anchoring domain (AD) (SEQ ID NO: 23), which is located from about amino acid 679 to about amino acid 702 of the human full-length CEA. While not being bound by theory, deletion of the anchoring domain increases secretion of the CEA fusion protein, thereby enhancing cross priming of the CEA-LTB immune response. In other preferred embodiments,
25 the CEA portion is a rhesus monkey CEA (SEQ ID NOs:18 and 19), or functional equivalent thereof.

The immunoenhancing element portion of the CEA fusion proteins of the present invention are capable of stimulating or enhancing the immune response to the associated CEA protein and are selected from the group consisting of: heat shock protein (HSP) 70, lysosome-associated membrane protein (LAMP), fragment C of tetanus toxoid (FrC), the N-terminal domain of FrC
30 (DOM), the heavy fragment of constant chain of immune globulin G1 (FcIgG), the vesicular stomatitis virus glycoprotein (VSV-G), cholera toxin (CT) from *Vibrio cholerae*, and heat labile enterotoxin of *E.coli* (LT). In preferred embodiments of the present invention, the adjuvant portion of the CEA fusion comprises the N-terminal domain of FrC (DOM), which has been shown to strongly potentiate the immunogenicity of codelivered antigens. In further preferred embodiments, the

adjuvant portion of the CEA fusion comprises a subunit of LT, or substantial portion thereof. In still further preferred embodiments, the adjuvant portion of the CEA fusion is a substantial portion of FcIgG.

5 A CEA fusion comprising a truncated human CEA fused to a single epitope of tetanus toxin (Q830 – L844) has been described (Lund et al. *Cancer Gene Therapy* 10: 365-376 (2003)). Unlike this single-epitope fusion, the CEA fusions of the present invention comprise a substantial portion of an immunoenhancing element or subunit thereof, as described above, which is capable of enhancing the immunogenicity of a CEA protein or variant thereof. A substantial portion of an immunoenhancing element to be used for the compositions and methods described herein does
10 not include portions that are less than 50% of a full-length toxin subunit. The strategy used herein, which utilizes full-length adjuvant subunits or substantial portions thereof, was employed to ensure a greater immune response to the fused CEA sequence. While not being bound by theory, it is believed that if the bacterial toxin chosen as adjuvant comprises greater than one helper epitope, limiting the toxin sequence of the fusion protein to a single epitope would arguably lead to a reduced effect on the
15 immunogenicity of the target protein. Additionally, it is believed that if the adjuvant-mediated enhancement of the immune response is dependent on the interaction of the adjuvant with specific cell receptors and not based on a universal epitope, then the receptor interaction could depend on a specific structural configuration that would require a substantial portion of the immunoenhancing element to exert an adjuvant effect. In such a case, a short adjuvant sequence comprising a single
20 epitope would be insufficient in mediating an increase of the immune response.

Also contemplated for use in the present invention are nucleotide sequences encoding variants or mutants of the immunoenhancing elements described herein, including but not necessarily limited to: nucleotide substitutions, deletions, additions, amino-terminal truncations and carboxy-terminal truncations. In some cases, it may be advantageous to add specific point mutations to the
25 nucleotide sequence encoding the adjuvant or subunit thereof to reduce or eliminate toxicity of the encoded protein. In exemplary embodiments of this aspect of the present invention, an LT subunit is fused to the CEA sequence of the CEA fusion, wherein the LT subunit is truncated of its signal sequence. While not being bound by theory, deletion of the toxin signal sequence, e.g. the LTB signal sequence, ensures that posttranslational processing of the CEA fusion is driven by the CEA
30 signal sequence.

The immunoenhancing element, subunit, or substantial portion thereof may be fused to the amino terminus or the carboxy terminus of the CEA sequence. Further, the immunoenhancing element sequence and the CEA sequence can be fused N-terminus to N-terminus, C-terminus to C-terminus, C-terminus to N-terminus or N-terminus to N-terminus. In preferred embodiments of the

present invention, the C-terminus of the CEA polypeptide is fused to the N-terminus of the immunoenhancing element.

The present invention relates to a synthetic nucleic acid molecule (polynucleotide) comprising a sequence of nucleotides which encodes mRNA that expresses a novel CEA fusion protein; for example, nucleotide sequences encoding the fusion proteins as set forth in SEQ ID
5 NOs:8, 10, 13, 15, 45, 46, 51 and 53. The nucleic acid molecules of the present invention are substantially free from other nucleic acids.

The present invention also relates to recombinant vectors and recombinant host cells, both prokaryotic and eukaryotic, which contain the nucleic acid molecules disclosed throughout this
10 specification. The synthetic DNA molecules, associated vectors, and hosts of the present invention are useful for the development of a cancer vaccine.

Exemplary nucleic acid molecules of the present invention comprise a nucleotide sequence selected from the group consisting of: SEQ ID NOs: 7, 9, 11, 12, 14, 21, 25, 49, 50, and 52, as shown in FIGURES 2-6, 26-27, 30-32, which encode exemplary CEA-LTA, CEA-LTB, CEA-
15 DOM, CEA-FcIgG, and CEA-CTB fusion proteins of the present invention.

The present invention also includes biologically active fragments or mutants of SEQ ID NOs: 7, 9, 11, 12, 14, 21, 25, 49, 50, and 52, which encode mRNA expressing exemplary CEA fusion proteins. Any such biologically active fragment and/or mutant will encode either a protein or protein fragment which at least substantially mimics the pharmacological properties of the hCEA
20 protein, including but not limited to the hCEA protein as set forth in SEQ ID NO:20. Any such polynucleotide includes but is not necessarily limited to: nucleotide substitutions, deletions, additions, amino-terminal truncations and carboxy-terminal truncations. The mutations of the present invention encode mRNA molecules that express a functional CEA fusion protein in a eukaryotic cell so as to be useful in cancer vaccine development.

Also included within the scope of this invention are mutations in the DNA sequence
25 that do not substantially alter the ultimate physical properties of the expressed protein. For example, substitution of valine for leucine, arginine for lysine, or asparagine for glutamine may not cause a change in the functionality of the polypeptide.

As stated above, the present invention further relates to recombinant vectors that
30 comprise the nucleic acid molecules disclosed throughout this specification. These vectors may be comprised of DNA or RNA. For most cloning purposes, DNA vectors are preferred. Typical vectors include plasmids, modified viruses, baculovirus, bacteriophage, cosmids, yeast artificial chromosomes, and other forms of episomal or integrated DNA that can encode a CEA fusion protein.

It is well within the purview of the skilled artisan to determine an appropriate vector for a particular gene transfer or other use.

Also provided by the present invention are purified CEA fusion proteins encoded by the nucleic acids disclosed throughout this specification. In exemplary embodiments of this aspect of the invention, the CEA fusion protein comprises a sequence of amino acids selected from the group consisting of: SEQ ID NOs: 8, 10, 13, 15, 45, 46, 51 and 53.

Included in the present invention are DNA sequences that hybridize to SEQ ID NOs: 7, 9, 11, 12, 14, 21, 25, 49, 50, or 52 under stringent conditions. By way of example, and not limitation, a procedure using conditions of high stringency is as follows. Prehybridization of filters containing DNA is carried out for about 2 hours to overnight at about 65°C in buffer composed of 6x SSC, 5x Denhardt's solution, and 100 µg/ml denatured salmon sperm DNA. Filters are hybridized for about 12 to 48 hrs at 65°C in prehybridization mixture containing 100 µg/ml denatured salmon sperm DNA and 5-20 x 10⁶ cpm of ³²P-labeled probe. Washing of filters is done at 37°C for about 1 hour in a solution containing 2x SSC, 0.1% SDS. This is followed by a wash in 0.1x SSC, 0.1% SDS at 50 °C for 45 minutes before autoradiography. Other procedures using conditions of high stringency would include either a hybridization step carried out in 5x SSC, 5x Denhardt's solution, 50% formamide at about 42°C for about 12 to 48 hours or a washing step carried out in 0.2x SSPE, 0.2% SDS at about 65°C for about 30 to 60 minutes. Reagents mentioned in the foregoing procedures for carrying out high stringency hybridization are well known in the art. Details of the composition of these reagents can be found in Sambrook et al., Molecular Cloning: A Laboratory Manual 2nd Edition; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, (1989) or Sambrook and Russell, Molecular Cloning: A Laboratory Manual, 3rd Edition. Cold Spring Harbor Laboratory Press, Plainview, NY (2001). In addition to the foregoing, other conditions of high stringency which may be used are also well known in the art.

An expression vector containing a CEA fusion protein-encoding nucleic acid molecule may be used for high-level expression of CEA fusion protein in a recombinant host cell. Expression vectors may include, but are not limited to, cloning vectors, modified cloning vectors, specifically designed plasmids or viruses. Also, a variety of bacterial expression vectors may be used to express recombinant CEA fusion sequences in bacterial cells if desired. In addition, a variety of fungal cell expression vectors may be used to express recombinant CEA fusion sequences in fungal cells. Further, a variety of insect cell expression vectors may be used to express recombinant protein in insect cells.

The present invention also relates to host cells transformed or transfected with vectors comprising the nucleic acid molecules of the present invention. Recombinant host cells may

be prokaryotic or eukaryotic, including but not limited to, bacteria such as *E. coli*, fungal cells such as yeast, mammalian cells including, but not limited to, cell lines of bovine, porcine, monkey and rodent origin; and insect cells including but not limited to *Drosophila* and silkworm derived cell lines. Such recombinant host cells can be cultured under suitable conditions to produce a CEA fusion protein or a
5 biologically equivalent form. In a preferred embodiment of the present invention, the host cell is human. As defined herein, the term "host cell" is not intended to include a host cell in the body of a transgenic human being, human fetus, or human embryo.

As noted above, an expression vector containing DNA encoding a CEA fusion protein may be used for expression of CEA fusion protein in a recombinant host cell. Therefore,
10 another aspect of this invention is a process for expressing a CEA fusion protein in a recombinant host cell, comprising: (a) introducing a vector comprising a nucleic acid comprising a sequence of nucleotides that encodes a CEA fusion protein into a suitable human host cell, wherein the CEA fusion protein comprises a CEA protein or variant thereof, fused to a substantial portion of an immunoenhancing element or subunit thereof, wherein the immunoenhancing element or subunit
15 thereof is selected from the group consisting of: heat shock protein (HSP) 70, lysosome-associated membrane protein (LAMP), fragment C of tetanus toxoid (FrC), the N-terminal domain of FrC (DOM), the heavy fragment of constant chain of immune globulin G1 (FcIgG), the vesicular stomatitis virus glycoprotein (VSV-G), cholera toxin (CT) from *Vibrio cholerae*, and heat labile enterotoxin of *E. coli* (LT); and wherein the fusion protein is capable of producing an immune
20 response in a mammal; and, (b) culturing the host cell under conditions which allow expression of said CEA fusion protein.

Preferred immunoenhancing elements for use in this aspect of the invention are selected from the group consisting of: LTB, LTA, DOM, and FcIgG.

In a further preferred embodiment of this aspect of the invention, the nucleotide
25 sequence of the CEA portion of the fusion and/or the immunoenhancing element portion of the fusion are codon-optimized for high-level expression in human cells.

This invention also provides a process for expressing a CEA-LT fusion protein in a recombinant host cell, comprising: (a) introducing a vector comprising a nucleic acid comprising a sequence of nucleotides that encodes a CEA-LT fusion protein into a suitable human host cell,
30 wherein the CEA fusion protein comprises a CEA protein or variant thereof, fused to a substantial portion of an LT subunit, and wherein the fusion protein is capable of producing an immune response in a mammal; and, (b) culturing the host cell under conditions which allow expression of said CEA-LT fusion protein.

In preferred embodiments of the process for expressing a CEA-LT fusion protein described above, the LT subunit is a substantial portion of LTB, wherein the LTB sequence is deleted of its signal sequence. In other embodiments, the LT subunit is LTA, or a substantial portion thereof.

This invention also provides a process for expressing a CEA-DOM fusion protein in a recombinant host cell, comprising: (a) introducing a vector comprising a nucleic acid comprising a sequence of nucleotides that encodes a CEA-DOM fusion protein into a suitable human host cell, wherein the CEA fusion protein comprises a CEA protein or variant thereof, fused to a substantial portion of the N-terminal domain of fragment C of tetanus toxin (DOM), and wherein the fusion protein is capable of producing an immune response in a mammal; and, (b) culturing the host cell under conditions which allow expression of said CEA-DOM fusion protein.

In preferred embodiments of the process for expressing a CEA-DOM fusion protein described above, the DOM portion is codon-optimized for high-level expression in human cells. In other preferred embodiments, the CEA portion of the CEA fusion is codon-optimized for high-level expression in human cells. In still further preferred embodiments, both the CEA and the DOM portions are codon-optimized for high-level expression in human cells.

Following expression of a CEA fusion in a host cell, CEA fusion protein may be recovered to provide CEA fusion protein in active form. Several protein purification procedures are available and suitable for use. Recombinant protein may be purified from cell lysates and extracts by various combinations of, or individual application of salt fractionation, ion exchange chromatography, size exclusion chromatography, hydroxylapatite adsorption chromatography and hydrophobic interaction chromatography. In addition, recombinant CEA fusion protein can be separated from other cellular proteins by use of an immunoaffinity column made with monoclonal or polyclonal antibodies specific for a CEA protein, or polypeptide fragments of a CEA protein.

The nucleic acid molecules comprising CEA fusions and the encoded fusion proteins of this invention were designed to enhance the CEA-specific immune response, relative to full-length cDNA encoding CEA, for use in vaccine development. To further enhance the immunogenic properties of the CEA fusion sequences of the present invention, in some embodiments described herein, the polynucleotides encoding CEA fusion proteins comprise optimized codons for further high level expression in a host cell, as described below. In these embodiments, at least a portion of the codons of the CEA fusions are designed so as to use the codons preferred by the projected host cell, which in preferred embodiments is a human cell. The optimized CEA fusions may be used for the development of recombinant adenovirus or plasmid-based DNA vaccines, which provide effective immunoprophylaxis against CEA-associated cancer through neutralizing antibody and cell-mediated immunity. The synthetic molecules may be used as an immunogenic composition. This invention

provides codon-optimized CEA fusion polynucleotides which, when directly introduced into a vertebrate *in vivo*, including mammals such as primates and humans, induce the expression of encoded proteins within the animal.

As stated above, in some embodiments of the present invention, the synthetic
5 molecules comprise a sequence of nucleotides, wherein some of the nucleotides have been altered so as to use the codons preferred by a human cell, thus allowing for high-level fusion protein expression in a human host cell. The synthetic molecules may be used as a source of a CEA fusion protein, for example, CEA-LTB fusion protein, which may be used in a cancer vaccine to provide effective immunoprophylaxis against CEA-associated carcinomas through neutralizing antibody and cell-
10 mediated immunity. The nucleic acid molecules disclosed herein may also serve as the basis for a DNA-based cancer vaccine.

A "triplet" codon of four possible nucleotide bases can exist in over 60 variant forms. Because these codons provide the message for only 20 different amino acids (as well as transcription initiation and termination), some amino acids can be coded for by more than one codon, a
15 phenomenon known as codon redundancy. For reasons not completely understood, alternative codons are not uniformly present in the endogenous DNA of differing types of cells. Indeed, there appears to exist a variable natural hierarchy or "preference" for certain codons in certain types of cells. As one example, the amino acid leucine is specified by any of six DNA codons including CTA, CTC, CTG, CTT, TTA, and TTG. Exhaustive analysis of genome codon frequencies for microorganisms has
20 revealed endogenous DNA of *E. coli* most commonly contains the CTG leucine-specifying codon, while the DNA of yeasts and slime molds most commonly includes a TTA leucine-specifying codon. In view of this hierarchy, it is generally believed that the likelihood of obtaining high levels of expression of a leucine-rich polypeptide by an *E. coli* host will depend to some extent on the frequency of codon use. For example, it is likely that a gene rich in TTA codons will be poorly
25 expressed in *E. coli*, whereas a CTG rich gene will probably be highly expressed in this host. Similarly, a preferred codon for expression of a leucine-rich polypeptide in yeast host cells would be TTA.

The implications of codon preference phenomena on recombinant DNA techniques are manifest, and the phenomenon may serve to explain many prior failures to achieve high
30 expression levels of exogenous genes in successfully transformed host organisms--a less "preferred" codon may be repeatedly present in the inserted gene and the host cell machinery for expression may not operate as efficiently. This phenomenon suggests that synthetic genes which have been designed to include a projected host cell's preferred codons provide an optimal form of foreign genetic material for practice of recombinant DNA techniques. Thus, one aspect of this invention is a CEA fusion gene

that is codon-optimized for expression in a human cell. In a preferred embodiment of this invention, it has been found that the use of alternative codons encoding the same protein sequence may remove the constraints on expression of exogenous CEA fusion protein in human cells.

In accordance with some embodiments of the present invention, the nucleic acid
5 molecules which encode the CEA fusion proteins are converted to a polynucleotide sequence having an identical translated sequence but with alternative codon usage as described by Lathe, "Synthetic Oligonucleotide Probes Deduced from Amino Acid Sequence Data: Theoretical and Practical Considerations" *J. Molec. Biol.* 183:1-12 (1985), which is hereby incorporated by reference. The methodology generally consists of identifying codons in the wild-type sequence that are not
10 commonly associated with highly expressed human genes and replacing them with optimal codons for high expression in human cells. The new gene sequence is then inspected for undesired sequences generated by these codon replacements (e.g., "ATTTA" sequences, inadvertent creation of intron splice recognition sites, unwanted restriction enzyme sites, etc.). Undesirable sequences are eliminated by substitution of the existing codons with different codons coding for the same amino
15 acid. The synthetic gene segments are then tested for improved expression.

The methods described above were used to create synthetic gene sequences which encode CEA fusion proteins, resulting in a gene comprising codons optimized for high level expression. While the above procedure provides a summary of our methodology for designing codon-
20 optimized genes for use in cancer vaccines, it is understood by one skilled in the art that similar vaccine efficacy or increased expression of genes may be achieved by minor variations in the procedure or by minor variations in the sequence.

One of skill in the art will also recognize that additional nucleic acid molecules may be constructed that provide for high levels of CEA fusion expression in human cells, wherein only a portion of the codons of the DNA molecules are codon-optimized. For example, in some
25 embodiments of the present invention, codons comprising the CEA portion of the CEA fusion are optimized for high-level expression in human cells, and codons comprising the adjuvant portion of the CEA fusion are substantially similar to the wild-type adjuvant-encoding nucleotide sequence. In other embodiments of the present invention, codons comprising the adjuvant portion of the CEA fusion are optimized for high-level expression in human cells, and codons comprising the CEA
30 portion of the CEA fusion are substantially similar to a wild-type CEA gene. In still other embodiments of the present invention, both the CEA and the adjuvant portions of the CEA fusion are codon-optimized for high-level expression in human cells. CEA fusions in which only a subset of codons are optimized within the CEA and/or the adjuvant portion of the CEA fusion are also contemplated by this invention.

The nucleic acids of the present invention may be assembled into an expression cassette which comprises sequences designed to provide for efficient expression of the protein in a human cell. The cassette preferably contains CEA fusion protein-encoding gene, with related transcriptional and translations control sequences operatively linked to it, such as a promoter, and termination sequences. In a preferred embodiment, the promoter is the cytomegalovirus promoter without the intron A sequence (CMV), although those skilled in the art will recognize that any of a number of other known promoters such as the strong immunoglobulin, or other eukaryotic gene promoters may be used. A preferred transcriptional terminator is the bovine growth hormone terminator, although other known transcriptional terminators may also be used. The combination of CMV-BGH terminator is particularly preferred.

In accordance with this invention, the CEA fusion expression cassette is inserted into a vector. The vector is preferably an adenoviral or plasmid vector, although linear DNA linked to a promoter, or other vectors, such as adeno-associated virus or a modified vaccinia virus, retroviral or lentiviral vector may also be used.

If the vector chosen is an adenovirus, it is preferred that the vector be a so-called first-generation adenoviral vector. These adenoviral vectors are characterized by having a non-functional E1 gene region, and preferably a deleted adenoviral E1 gene region. In some embodiments, the expression cassette is inserted in the position where the adenoviral E1 gene is normally located. In addition, these vectors optionally have a non-functional or deleted E3 region. It is preferred that the adenovirus genome used be deleted of both the E1 and E3 regions ($\Delta E1\Delta E3$). The adenoviruses can be multiplied in known cell lines which express the viral E1 gene, such as 293 cells, or PERC.6 cells, or in cell lines derived from 293 or PERC.6 cell which are transiently or stably transformed to express an extra protein. For examples, when using constructs that have a controlled gene expression, such as a tetracycline regulatable promoter system, the cell line may express components involved in the regulatory system. One example of such a cell line is T-Rex-293; others are known in the art.

For convenience in manipulating the adenoviral vector, the adenovirus may be in a shuttle plasmid form. This invention is also directed to a shuttle plasmid vector which comprises a plasmid portion and an adenovirus portion, the adenovirus portion comprising an adenoviral genome which has a deleted E1 and optional E3 deletion, and has an inserted expression cassette comprising a CEA fusion protein encoding nucleotide sequence. In preferred embodiments, there is a restriction site flanking the adenoviral portion of the plasmid so that the adenoviral vector can easily be removed. The shuttle plasmid may be replicated in prokaryotic cells or eukaryotic cells.

In a preferred embodiment of the invention, the expression cassette is inserted into the pMRKAd5-HV0 adenovirus plasmid (*See* Emini et al., WO 02/22080, which is hereby incorporated by reference). This plasmid comprises an Ad5 adenoviral genome deleted of the E1 and E3 regions. The design of the pMRKAd5-HV0 plasmid was improved over prior adenovectors by
5 extending the 5' cis-acting packaging region further into the E1 gene to incorporate elements found to be important in optimizing viral packaging, resulting in enhanced virus amplification. Advantageously, this enhanced adenoviral vector is capable of maintaining genetic stability following high passage propagation.

Standard techniques of molecular biology for preparing and purifying DNA
10 constructs enable the preparation of the adenoviruses, shuttle plasmids, and DNA immunogens of this invention.

It has been determined in accordance with the present invention that the CEA-LT fusion protein-encoding molecules described herein (e.g. SEQ ID NO:12), which comprise a substantial portion of the LTA or LTB subunits of *E.coli* heat labile enterotoxin, are expressed with
15 equivalent efficiency compared to the corresponding wild type CEA sequence (*See* EXAMPLE 4). It has also been shown herein that plasmids pV1J/hCEA-LTA and pV1J/hCEA-LTB elicited a greater antibody response than pV1J/hCEA, confirming the adjuvant effect exerted by the LT subunits on the CEA specific immune response (*See* EXAMPLE 11). Thus, the data described herein demonstrate that fusion of the CEA coding sequence to the LTA or LTB cDNA results in an increase the CEA
20 specific immune response. It appears that LTB exerts a greater enhancing effect on the immune response with a prevalent induction of CD8⁺ T cells, whereas LTA elicits a predominant CD4⁺ response.

It has also been shown in accordance with the present invention that tolerance to the CEA self antigen can be broken more efficiently, relative to the full-length wild-type CEA cDNA, due
25 to the increased immunogenic properties of the CEA-LTB fusion. The enhancing effect of LTB on the immunogenic properties of CEA was also observed upon injection of a plasmid carrying a fully codon optimized cDNA of the CEA-LTB fusion. Lastly, the results described herein, using adenovirus vectors carrying CEA-LT fusions, indicate that that enhanced immunogenicity of CEA-LT fusions is not limited to plasmid DNA immunization (*see* EXAMPLE 13).

It has further been shown in accordance with the present invention that plasmids
30 pV1J/hCEA-DOM and pV1J/hCEA-FcIgG elicited a greater CEA-specific cell-mediated and humoral immune response than CEA (*See* EXAMPLE 15). It has also been shown in accordance with the present invention that tolerance to the CEA self antigen can be broken more efficiently with the DOM and FcIgG CEA fusions described herein, relative to the full-length wild-type CEA cDNA, due to the

increased immunogenic properties of the CEA fusions. The enhanced immunogenic properties of these fusion proteins were observed upon immunization with DNA or Ad vectors, indicating that enhanced immunogenicity of CEA-LT fusions is not limited to plasmid DNA immunization (*see* EXAMPLE 16).

5 Therefore, the vectors described above may be used in immunogenic compositions and vaccines for preventing the development of adenocarcinomas associated with aberrant CEA expression and/or for treating existing cancers. The vectors of the present invention allow for vaccine development and commercialization by eliminating difficulties with obtaining high expression levels of exogenous CEA in successfully transformed host organisms and by providing a CEA fusion protein
10 which can elicit an enhanced immune response when administered to a mammal such as a human being.

 To this end, one aspect of the instant invention is a method of preventing or treating CEA-associated cancer comprising administering to a mammal a vaccine vector comprising a
15 polynucleotide comprising a sequence of nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a CEA protein or variant thereof, fused to a substantial portion of an immunoenhancing element selected from the group consisting of: heat shock protein (HSP) 70, lysosome-associated membrane protein (LAMP), fragment C of tetanus toxoid (FrC), the N-terminal domain of FrC (DOM), the heavy fragment of constant chain of immune globulin G1 (FcIgG), the
20 vesicular stomatitis virus glycoprotein (VSV-G), cholera toxin (CT) from *Vibrio cholerae*, and heat labile enterotoxin of *E.coli* (LT); and wherein the fusion protein is capable of producing an immune response in a mammal.

 In preferred embodiments of the methods described herein, the immunoenhancing element is selected from the group consisting of: LTA, LTB, DOM and FcIgG.

 In accordance with the method described above, the vaccine vector may be
25 administered for the treatment or prevention of a cancer in any mammal, including but not limited to: lung cancer, breast cancer, and colorectal cancer. In a preferred embodiment of the invention, the mammal is a human.

 Further, one of skill in the art may choose any type of vector for use in the treatment and prevention method described. Preferably, the vector is an adenovirus vector or a plasmid vector.
30 In a preferred embodiment of the invention, the vector is an adenoviral vector comprising an adenoviral genome with a deletion in the adenovirus E1 region, and an insert in the adenovirus E1 region, wherein the insert comprises an expression cassette comprising: (a) a sequence of nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a CEA protein or variant thereof, fused to an immunoenhancing element or substantial portion thereof; wherein the

immunoenhancing element is selected from the group consisting of: heat shock protein (HSP) 70, lysosome-associated membrane protein (LAMP), fragment C of tetanus toxoid (FrC), the N-terminal domain of FrC (DOM), the heavy fragment of constant chain of immune globulin G1 (FcIgG), the vesicular stomatitis virus glycoprotein (VSV-G), cholera toxin (CT) from *Vibrio cholerae*, and heat labile enterotoxin of *E.coli* (LT); and wherein the fusion protein is capable of producing an immune response in a mammal; and (b) a promoter operably linked to the polynucleotide.

The instant invention further relates to an adenovirus vaccine vector comprising an adenoviral genome with a deletion in the E1 region, and an insert in the E1 region, wherein the insert comprises an expression cassette comprising: (a) a sequence of nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a CEA protein or variant thereof, fused to a substantial portion of an immunoenhancing element; wherein the immunoenhancing element is selected from the group consisting of: HSP70, LAMP, FrC, DOM, FcIgG, VSV-G, CT, LTA and LTB; and wherein the fusion protein is capable of producing an immune response in a mammal; and (b) a promoter operably linked to the polynucleotide.

In a preferred embodiment of this aspect of the invention, the adenovirus vector is an Ad 5 vector.

In another preferred embodiment of the invention, the adenovirus vector is an Ad 6 vector.

In yet another preferred embodiment, the adenovirus vector is an Ad 24 vector.

Also contemplated for use in the present invention is an adenovirus vaccine vector comprising a adenovirus genome that naturally infects a species other than human, including, but not limited to, chimpanzee adenoviral vectors. A preferred embodiment of this aspect of the invention is a chimp Ad 3 vaccine vector.

In another aspect, the invention relates to a vaccine plasmid comprising a plasmid portion and an expression cassette portion, the expression cassette portion comprising: (a) a sequence of nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a CEA protein or variant thereof, fused to an immunoenhancing element or substantial portion thereof, selected from the group consisting of: HSP70, LAMP, FrC, DOM, the FcIgG, VSV-G, CT, LTA and LTB; and wherein the fusion protein is capable of producing an immune response in a mammal; and (b) a promoter operably linked to the polynucleotide.

In some embodiments of this invention, the recombinant adenovirus and plasmid-based polynucleotide vaccines disclosed herein are used in various prime/boost combinations in order to induce an enhanced immune response. In this case, the two vectors are administered in a "prime and boost" regimen. For example the first type of vector is administered one or more times, then after

a predetermined amount of time, for example, 2 weeks, 1 month, 2 months, six months, or other appropriate interval, a second type of vector is administered one or more times. Preferably the vectors carry expression cassettes encoding the same polynucleotide or combination of polynucleotides. In the embodiment where a plasmid DNA is also used, it is preferred that the vector
5 contain one or more promoters recognized by mammalian or insect cells. In a preferred embodiment, the plasmid would contain a strong promoter such as, but not limited to, the CMV promoter. The synthetic CEA fusion gene or other gene to be expressed would be linked to such a promoter. An example of such a plasmid would be the mammalian expression plasmid V1Jns as described (J. Shiver *et. al.* in *DNA Vaccines*, M. Liu *et al.* eds., N.Y. Acad. Sci., N.Y., 772:198-208 (1996), which is
10 herein incorporated by reference).

As stated above, an adenoviral vector vaccine and a plasmid vaccine may be administered to a vertebrate as part of a single therapeutic regime to induce an immune response. To this end, the present invention relates to a method of protecting a mammal from a CEA-associated cancer comprising: (a) introducing into the mammal a first vector comprising: i) a sequence of
15 nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a CEA protein or variant thereof, fused to a substantial portion of an immunoenhancing element selected from the group consisting of: HSP70, LAMP, FrC, DOM, the FcIgG, VSV-G, CT, LTA and LTB; and wherein the fusion protein is capable of producing an immune response in a mammal; and ii) a promoter operably linked to the polynucleotide; (b) allowing a predetermined amount of time to pass;
20 and (c) introducing into the mammal a second vector comprising: i) a sequence of nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a CEA protein or variant thereof, fused to a substantial portion of an immunoenhancing element selected from the group consisting of: HSP70, LAMP, FrC, DOM, the FcIgG, VSV-G, CT, LTA and LTB; and wherein the fusion protein is capable of producing an immune response in a mammal; and ii) a promoter operably
25 linked to the polynucleotide.

In one embodiment of the method of protection described above, the first vector is a plasmid and the second vector is an adenovirus vector. In an alternative embodiment, the first vector is an adenovirus vector and the second vector is a plasmid.

In the method described above, the first type of vector may be administered more than
30 once, with each administration of the vector separated by a predetermined amount of time. Such a series of administration of the first type of vector may be followed by administration of a second type of vector one or more times, after a predetermined amount of time has passed. Similar to treatment with the first type of vector, the second type of vector may also be given one time or more than once, following predetermined intervals of time.

The instant invention further relates to a method of treating a mammal suffering from a CEA-associated adenocarcinoma comprising: (a) introducing into the mammal a first vector comprising: i) a sequence of nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a CEA protein or variant thereof, fused to a substantial portion of an immunoenhancing element selected from the group consisting of: HSP70, LAMP, FrC, DOM, the FcIgG, VSV-G, CT and LT; and wherein the fusion protein is capable of producing an immune response in a mammal; and ii) a promoter operably linked to the polynucleotide; (b) allowing a predetermined amount of time to pass; and (c) introducing into the mammal a second vector comprising: i) a sequence of nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a CEA protein or variant thereof, fused to a substantial portion of an immunoenhancing element selected from the group consisting of: HSP70, LAMP, FrC, DOM, the FcIgG, VSV-G, CT and LT; and wherein the fusion protein is capable of producing an immune response in a mammal; and ii) a promoter operably linked to the polynucleotide.

In one embodiment of the method of treatment described above, the first vector is a plasmid and the second vector is an adenovirus vector. In an alternative embodiment, the first vector is an adenovirus vector and the second vector is a plasmid.

In preferred embodiments of the methods described above, the vectors comprise a sequence of nucleotides that encode a CEA-LT fusion protein, wherein the CEA fusion protein comprises a CEA protein or variant thereof, fused to a substantial portion of a LT subunit. In further preferred embodiments, the vector comprises a sequence of nucleotides that encodes a CEA-LTB fusion protein. In further preferred embodiments of the methods described above, the vectors comprise a sequence of nucleotides that encode a CEA-DOM fusion protein, wherein the CEA fusion protein comprises a CEA protein or variant thereof, fused to a substantial portion of a DOM subunit. In still further preferred embodiments, the vector comprises a sequence of nucleotides that encodes a CEA-FcIgG fusion protein.

The amount of expressible DNA or transcribed RNA to be introduced into a vaccine recipient will depend partially on the strength of the promoters used and on the immunogenicity of the expressed gene product. In general, an immunologically or prophylactically effective dose of about 1 ng to 100 mg, and preferably about 10 µg to 300 µg of a plasmid vaccine vector is administered directly into muscle tissue. An effective dose for recombinant adenovirus is approximately 10^6 – 10^{12} particles and preferably about 10^7 — 10^{11} particles. Subcutaneous injection, intradermal introduction, impression through the skin, and other modes of administration such as intraperitoneal, intravenous, intramuscular or inhalation delivery are also contemplated.

In preferred embodiments of the present invention, the vaccine vectors are introduced to the recipient through intramuscular injection.

The vaccine vectors of this invention may be naked, i.e., unassociated with any proteins, or other agents which impact on the recipient's immune system. In this case, it is desirable for the vaccine vectors to be in a physiologically acceptable solution, such as, but not limited to, sterile saline or sterile buffered saline. Alternatively, it may be advantageous to administer an agent which assists in the cellular uptake of DNA, such as, but not limited to calcium ion. These agents are generally referred to as transfection facilitating reagents and pharmaceutically acceptable carriers. Those of skill in the art will be able to determine the particular reagent or pharmaceutically acceptable carrier as well as the appropriate time and mode of administration.

All publications mentioned herein are incorporated by reference for the purpose of describing and disclosing methodologies and materials that might be used in connection with the present invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

Having described preferred embodiments of the invention with reference to the accompanying drawings, it is to be understood that the invention is not limited to those precise embodiments, and that various changes and modifications may be effected therein by one skilled in the art without departing from the scope or spirit of the invention as defined in the appended claims.

The following examples illustrate, but do not limit the invention.

EXAMPLE 1

Construction of CEA Fusion Proteins

To determine the immunogenicity of CEA fusion proteins, a series of vectors were constructed encoding amino acids (hereinafter aa) 1 to 679 of the human CEA protein fused to a panel of selected polypeptides (see EXAMPLE 2). These sequences were chosen in view of their reported immunoenhancing properties, which have been demonstrated in a variety of experimental systems. The CEA fusions were engineered by joining the cDNA of the CEA protein deleted of the GPI anchoring sequence to the foreign polypeptides (exemplary constructs are depicted in FIGURE 18A). The tumor antigen was linked to the HSP70, FcIgG or LAMP sequences to determine whether enhancement of antigen uptake or retargeting to the endosomal compartment would lead to an increased immune response. Similarly, fusion to fragment C of tetanus toxin (FrC) or to a minimal domain devoid of potentially competitive MHC class I binding epitopes (DOM, see FIGURE 29) (Rice et al. *J. Immunol.* 169: 3908-13 (2002)) was constructed to promote humoral and CD4⁺ T cell

responses. CEA was also linked to the VSV-G coding sequence to determine whether fusion to a viral glycoprotein would influence the immunogenic properties of CEA.

The coding sequences of these CEA-fusions were cloned into vectors pV1Jns under the control of the human CMV/intron A promoter plus the bovine growth hormone (BGH) polyadenylation signal (EXAMPLE 2). Plasmids pV1J/CEA-FRC, pV1J/CEA-DOM, pV1J/CEA-FcIgG, pV1J/CEA-LAMP, pV1J/CEA-VSV-G, and pV1J/CEA-HSP70 carry the wild type cDNA of CEA fused to the coding sequences of the indicated foreign polypeptides. Exemplary nucleotide and amino acid sequences of hCEA-DOM and hCEA-FcIgG fusions are shown in FIGURES 26, 27, and 30.

To assess the effect of LTA and LTB subunits of the *E. coli* heat labile enterotoxin on CEA immunogenicity, a series of additional fusion constructs were generated encoding amino acids 1 to 679 of CEA protein fused to either the LTA (aa 18 to 259) or the LTB (aa 21 to 125) coding sequence. A schematic representation of the structure of exemplary CEA-LTA and CEA-LTB fusions developed for this study are shown in FIGURE 1. Exemplary nucleotide and amino acid sequences of CEA-LT fusions are shown in FIGURES 2-6.

CEA-LT fusions were engineered by joining the cDNA of the CEA protein deleted of the anchoring sequence to the LT subunits to which the signal peptide coding sequence had been removed. The CEA-fusions coding sequences were cloned into vectors pV1Jns under the control of the human cytomegalovirus (CMV)/intron A promoter plus the bovine growth hormone (BGH) polyadenylation signal. Plasmids pV1J/hCEA-LTA and pV1J/hCEA-LTB carry the wild type cDNA of CEA fused to the coding sequences of LTA and LTB, respectively (see EXAMPLE 2).

All constructs carrying the CEA-LTB fusion were generated by fusing the CEA cDNA from nt 1 to 2037, with the LTB cDNA fragment encompassing nt 64 to 375. The LTB coding sequence was obtained by PCR amplification of *E. coli* genomic DNA using sequence specific primers LTB-S1 5'-T A T T C T A G A T G C T C C C C A G A C T A T T A C A G A A -3' (SEQ ID NO:1) and LTB-A1 5'-T A T G C G G C C G C C T A G T T T T C C A T A C T G A T T G C C G C -3' (SEQ ID NO:2). The amplified DNA was introduced at the 3' end of the CEA coding sequence generating plasmids.

EXAMPLE 2

Plasmid Constructs

pV1J/CEA_{opt} and pV1J/CEA: These two constructs carry the codon usage optimized and wild type cDNA of CEA, respectively. The CEA coding sequence is located between the CMV/intronA immediate early promoter of cytomegalovirus and the bovine growth hormone

polyadenylation signal. For generation of pV1J/hCEA_{opt}, plasmid pCR-hCEA_{opt} was digested with *EcoRI* for 1 hr at 37°C. The resulting 2156 bp insert was purified and cloned into the *EcoRI* site of plasmid pV1JnsB ((Montgomery et al. *DNA Cell Biol* 12(9): 777-83 (1993)).

For generation of pV1J/hCEA, plasmid pCI/hCEA (Song et al. Regulation of T-helper-1 versus T-helper-2 activity and enhancement of tumour immunity by combined DNA-based vaccination and nonviral cytokine gene transfer. *Gene Therapy* 7: 481-492 (2000)) was digested with *EcoRI*. The resulting 2109 bp insert was cloned into the *EcoRI* site of plasmid pV1JnsA (Montgomery et al., *supra*).

pV1J/hCEA-LTB and pV1J/hCEA_{opt}-LTB: The codon optimized cDNA of LT B was synthesized by oligonucleotide assembly (Geneart GmbH, Regensburg, Germany) and cloned in pCR-script vector (Stratagene, LA Jolla, CA). To generate pV1J/hCEA_{opt}-LTB_{opt}, LT B_{opt} was amplified by PCR using the following PCR primers: LTB_{opt}-5'XbaI (5' end) 5' – G C T C T A G A G C C C C C A G A G C A T C A C C G A G C T G T G C - 3' (SEQ ID NO:3) and LTB_{opt}-3'BglII (3' end) 5' – G C T C T A G A A C C C C T C A G A A C A T C A C C G A T C T G T G C G C C - 3' (SEQ ID NO:4). The amplified product was then inserted into the *XbaI/BglII* sites of plasmid pV1J/hCEA_{opt}.

pV1J/hCEA-LTA: The LTA coding sequence corresponding to nt 54 to 774 that encode aa 18 to 259, was amplified by PCR from genomic DNA of *E. coli* using sequence specific primers LTA-S1 5' – T A T T C T A G A T A A T G G C G A C A A A T T A T A C C G - 3' (SEQ ID NO:5) and LTA-A1 5' – T A T G C G G C C G C T C A T A A T T C A T C C C G A A T T C T G T T - 3' (SEQ ID NO:6). The amplified DNA was digested with appropriate restriction enzymes and inserted into plasmid pV1J/hCEA.

pV1J/rhCEA_{opt}-LTB: A 3' fragment of the rhesus monkey CEA cDNA (nt 1641 to 2026), which was codon-optimized for high level expression in human cells, was amplified by PCR from pV1J-rhCEA_{opt}. The amplified cDNA lacked the GPI anchor coding sequence and carried the *XbaI/BglII* restriction sites. This fragment was inserted into the *PstI* site of pCR-blunt-rhCEA_{opt}, thus obtaining the intermediate pCR-blunt-rhCEA_{opt} *XbaI/BglIII*. rhCEA_{opt} was extracted as a *BglIII/SalI* fragment and cloned in the same sites in pV1J-nsB, thus obtaining pV1J-rhCEA_{opt} *XbaI/BglII*. LT B_{opt} was amplified by PCR from pCR-script-LT B_{opt} adding *XbaI* and *BglII* sites at 5' and 3' ends, respectively, and was cloned in pV1J-rhCEA_{opt} *XbaI/BglII*, thus obtaining pV1J-rhCEA_{opt}-LT B_{opt}.

pV1J/CEA-FrC, pV1J/CEA-DOM, pV1J/CEA-FcIgG, pV1J/CEA-LAMP, pV1J/CEA-HSP70 and pV1J/CEA-VSV-G: All the constructs encoding the referenced CEA fusion proteins were generated by fusing the CEA cDNA from nt 1 to nt 2037 (SEQ ID NO:22, FIGURE

28A), corresponding to aa 1 to aa 679 (SEQ ID NO:23, FIGURE 28B), with the cDNA fragment corresponding to the following: fragment C of tetanus toxoid (CEA-FrC, SEQ ID NO:24), the N-terminal domain of FrC (CEA-DOM, SEQ ID NOs:21 and 49), the heavy fragment of constant chain of immune globulin G1 (CEA-FcIgG, SEQ ID NO:25), the lysosome-associated membrane protein (CEA-LAMP, SEQ ID NO:26), the heat shock protein 70 (CEA-HSP70, SEQ ID NO:27), or the vesicular stomatitis virus glycoprotein (CEA-VSV-G, SEQ ID NO:28).

FrC and DOM coding sequences were obtained by PCR amplification from pRep-TeT.C plasmid as described in Rice et al. (*J. Immunol.* 169: 3908-13 (2002)). FcIgG was obtained from total RNA of human PBMC. VSV-G and HSP70 were obtained from p-FAST-VSV-G and from plasmid pY3111, respectively. LAMP1 was obtained by gene assembly. Amplifications were carried out using the following primers: FrC sense (5'-T A T T C T A G A T T C A A C A C C A A T T C C A T T T T C T T A T T C -3' (SEQ ID NO:29) FrC antisense (5'-G C G G C C G C T A G A A T C A T T T G T C C A T C C T T C A T C -3' (SEQ ID NO:30), DOM sense (5'-T A T T C T A G A T T C A A C A C C A A T T C C A T T T T C T T A T T C -3' (SEQ ID NO:31) DOM antisense (5'-T T A G C G G C C G C T A G T T C T G T A T C A T A T C G T A A A G G G -3' (SEQ ID NO:32), FcIgG sense (5'-T C T A G A T A A A A C T C A C A C A T G C C C A -3' (SEQ ID NO:33) FcIgG antisense (5'-G C C G A C T C A T T T A C C C G G A G A C A G G G A G -3' (SEQ ID NO:34), LAMP sense (5'-T C T A G A T T T G A T C C C C A T T G C T G T G G G C G G T G C C C T G -3' (SEQ ID NO:35) LAMP antisense (5'-G G C G T G A C T C C T C T T C C T G C C A A T G A G G T A G G C A A T G A G -3' (SEQ ID NO:36), VSV-G sense (5'-A T A T C T A G A T T T C A C C A T A G T T T T T C C A C A C A A C C -3' (SEQ ID NO:37) VSV-G antisense (5'-G C G G C C G C C T T C C T T C C A A G T C G G T T C A T C T C T A T G -3' (SEQ ID NO:38), HSP70 sense (5'-G C T C T A G A T A T G G C T C G T G C G G T C G G G A T C G A C C -3' (SEQ ID NO:39)) and HSP70 antisense (5'-G C C G C G G C C G C T C A C T T G G C C T C C C G G C C G T C G T C G -3' (SEQ ID NO:40). The amplified DNA was introduced at the 3' end of the CEA coding sequence generating plasmids pV1J/CEA-FrC, pV1J/CEA-DOM, pV1J/CEA-FcIgG, pV1J/CEA-LAMP, pV1J/CEA-HSP70 and pV1J/CEA-VSV-G.

pV1J/CEA-DOM_{opt} and pV1J/CEA-FcIgG_{opt}: The codon usage optimized cDNA of DOM and FcIgG were synthesized by oligonucleotide assembly (Geneart GmbH, Regensburg, Germany) and cloned in pCR-script vector (Stratagene, La Jolla, CA). To generate pV1J/CEA-DOM_{opt}, DOM_{opt} was amplified by PCR using the following primers: Dom_{opt} sense (5'-G T T A T C T A G A A G C A C C C C C A T C C C -3' (SEQ ID NO:41)) and Dom_{opt} reverse (5'-T T A A G A T C T C T A A G A T C T G G T G T C G T A T C T C A G G G G -3' (SEQ ID NO:42). The amplified product was then inserted into the *XbaI/BglII* sites of plasmid pV1J/CEA_{opt}. To generate

pV1J/CEA-FcIgGopt, FcIgGopt was amplified by PCR using the following primers: FcIgGopt sense (5'-T T A T C T A G A A A G A C C C A C A C C T G C C C C C C T T G C -3' (SEQ ID NO:43)) and as FcIgGopt reverse (5'- T A T A G A T C T T A G G G T A C C T T A C T T G C C G G G -3' (SEQ ID NO:44)) the amplified product was inserted into *XbaI/BglIII* sites of plasmid pV1J/CEAopt.

EXAMPLE 3

Adenovirus Vectors

Ad5/hCEAopt: Plasmid pCR-hCEAopt was digested with *EcoRI*. The resulting 2156 bp insert was purified and cloned into the *EcoRI* of the polyMRK-Ad5 shuttle plasmid.

Ad5/CEA: The shuttle plasmid pMRK-hCEA for generation of Ad5 vector was obtained by digesting plasmid pDelta1sp1B/hCEA with *SspI* and *EcoRV*. The 9.52kb fragment was then ligated with a 1272 bp *BglIII/BamHI*-restricted, Klenow-treated product from plasmid polyMRK. A *PacI/StuI* fragment from pMRK-hCEA and pMRK-hCEAopt containing the expression cassette for hCEA and E1 flanking Ad5 regions was recombined to *ClaI* linearized plasmid pAd5 in BJ5183 *E. coli* cells. The resulting plasmids were pAd5-hCEA and pAd5-hCEAopt, respectively. Both plasmids were cut with *PacI* to release the Ad ITRs and transfected in PerC-6 cells. Ad5 vectors amplification was carried out by serial passage. MRKAd5/hCEA and MRKAd5/hCEAopt were purified through standard CsCl gradient purification and extensively dialyzed against A105 buffer (5mM Tris-Cl pH 8.0, 1mM MgCl₂, 75 mM NaCl, 5% Sucrose, 0.005 Tween 20).

Ad5/hCEAopt-LTB: Plasmid pMRK-hCEAopt-LTB was constructed by cutting polyMRK-Ad5 shuttle plasmid with *SwaI* and by ligating the linearized vector with the 2300 bp DNA fragment derived from pV1J/hCEAopt-LTB that had been restricted with *EcoRI*, *BglIII* and treated with Klenow. The pMRK-hCEAopt-LTB was linearized and recombined into the Ad genome as indicated above.

Ad5/CEA-DOMopt and Ad5/CEA-FcIgGopt: Plasmid pMRK-CEA-DOMopt and pMRK-CEA-FcIgGopt were constructed by cutting polyMRK-Ad5 shuttle plasmid with *SwaI* and by ligating the linearized vector with the 2.9 kb DNA fragment derived from pV1J/CEA-DOMopt or ligating the linearized vector with the 2700 bp DNA fragment derived from pV1J/CEA-FcIgG1opt that had been restricted with *EcoRI*, *BglIII* and treated with Klenow. pMRK-CEA-FcIgGopt and pMRK-CEA-DOMopt were linearized and recombined into the Ad genome as indicated above.

EXAMPLE 4

Comparative Expression Efficiency of Various CEA-LT Fusion Constructs

The use of codon optimized cDNAs for genetic vaccination against viral diseases has been shown to elicit a greater immune response due, at least in part, to an increased expression of the target protein. To verify whether the LTB coding sequence would also enhance the immunogenic properties of the CEA cDNA designed to incorporate human-preferred (humanized) codons for each amino acid residue, plasmid pV1J/hCEAopt-LTB was also constructed. Finally, a fully codon optimized version of the CEA-LTB fusion was also constructed using a synthetic codon optimized cDNA of LTB to generate plasmid pV1J/hCEA-LTBopt.

To determine whether the LTB effect on CEA immunogenicity was not limited to plasmid DNA immunization, an Adenovirus type 5 vector encoding the CEAopt-LTB fusion flanked by the CMV/intron A promoter and the BGH polyadenylation signal was also constructed. The molecular mass of the CEA fusion proteins expressed by both plasmid and Ad vectors did not differ from that derived from the corresponding vectors encoding the full length form of CEA cDNA (data not shown).

To compare the efficiency of expression of the vectors encoding the CEA-LTA and CEA-LTB fusions and that of the cDNA of full length CEA, HeLa cells were transfected with plasmids pV1J/hCEA-LTA, and pV1J/hCEA-LTB. The CEA expression of these constructs was compared to that of the corresponding plasmid carrying the wt cDNA of CEA, pV1J/hCEA. Similarly, plasmid pV1J/hCEAopt-LTB expression efficiency was compared to that of pV1J/hCEAopt. Expression efficiency of these constructs were determined two days post transfection by monitoring the amount of CEA protein in cell extracts.

Transfection of plasmids pV1J/hCEA-LTA and pV1J/hCEA-LTB yielded approximately two fold higher amounts of CEA protein (183 and 139 $\mu\text{g/l}$, respectively, FIGURE 8A) detected in the culture supernatant as compared to plasmid pV1J/CEA (91 $\mu\text{g/l}$). Similarly, the expression efficiency of constructs pV1J/hCEAopt and pV1J/hCEAopt-LTB was also comparable (113 and 136 $\mu\text{g/l}$, respectively; FIGURE 8B). Finally, the expression efficiency of the Ad/hCEAopt and Ad/hCEAopt-LTB was also compared by infecting HeLa cells at different moi. The CEA expression efficiency of these two vectors was comparable at moi 1000 (1790 and 1400 $\mu\text{g/l}$, respectively, FIGURE 9) whereas at moi 100, vector Ad/hCEAopt-LTB yielded approximately four fold lower amounts of CEA protein detectable in the culture supernatant than Ad/hCEAopt (390 and 1500 $\mu\text{g/l}$, respectively).

Thus, these results indicate that the cDNA encoding the CEA-LTA and CEA-LTB fusion proteins are expressed with equivalent efficiency to that of the corresponding cDNA encoding the full length CEA protein. Additionally, the comparable CEA expression of these cDNAs is not influenced by the type of the gene transfer vehicle utilized for their delivery.

EXAMPLE 5

Detection of CEA Expression.

CEA expression by plasmid and Ad vectors was monitored by Western blot analysis and ELISA. Plasmids were transfected in HeLa cells with Lipofectamine 2000 (Life Technologies). Adenovirus infections of HeLa cells were performed in serum free medium for 30 min at 37°C, and then fresh medium was added. After 48hr incubation, whole cell lysates were harvested. The CEA protein present in the cell lysates was detected by Western blot analysis using a rabbit polyclonal antiserum. The protein was detected as a 180-200 kDa band. The amount of expressed CEA was detected in the cell lysates using the Direct Elisa CEA Kit (DBC-Diagnostics Biochem Canada Inc).

Expression of the fusion proteins in transfected cells was examined by Western blot analysis using antibodies specific for CEA, VSV-G, FcIgG, tetanus toxin, or HSP70. HeLa cells were either transfected with the indicated plasmid or infected with the selected Ad vector. After 48 hr incubation, whole cell lysates and culture supernatant were harvested.

CEA expression in cell lysate or supernatant was also monitored using the Direct Elisa CEA Kit (DBC-Diagnostics Biochem Canada Inc). CEA protein was detected with the antibody specific for the fused polypeptide in transfected cell lysates, whereas no expression of the target antigen was observed in the mock transfected control samples (FIGURE 18B). The molecular mass of the fusion proteins did not differ significantly from that of CEA. This apparent lack of difference in molecular mass between the various CEA polypeptides is probably due to the high degree of glycosylation of the tumor antigen.

To compare the efficiency of expression of the vectors encoding the CEA-fusions to that of pV1J/CEA, HeLa cells were transfected with the different plasmids and CEA expression of these constructs was determined two days post transfection by ELISA. Plasmids pV1J/CEA-FrC, pV1J/CEA-DOM, pV1J/CEA-FcIgG, pV1J/CEA-LAMP, pV1J/CEA-VSV-G and pV1J/CEA-HSP70 expressed CEA with comparable efficiency as pV1J/CEA (FIGURE 19A). Most of the fusion proteins were secreted and detected in the cell supernatant; however, CEA-LAMP was not released from the transfected cells, probably due to its re-routing to the lysosomal compartment. (FIGURE 19B). Thus, these results indicate that the cDNA encoding the CEA-FrC, CEA-DOM, CEA-VSV-G, CEA-FcIgG, CEA-HSP70 and CEA-LAMP fusion proteins are expressed with equivalent efficiency to that of the cDNA encoding the full length CEA protein.

EXAMPLE 6

Peptides

Lyophilized hCEA peptides were purchased from Bio-Synthesis and resuspended in DMSO at 40 mg/ml. Pools of peptides 15 aa long overlapping by 11 residues were assembled as described (Facciabene et al. *J. Virol.* 78: 8663-72 (2004)). Final concentrations were the following: pool A=1.2 mg/ml, pool B 0.89 mg/ml, pool C 0.89 mg/ml, pool D 0.8 mg/ml. Peptides were stored at -80°C. Immune response to DOM was monitored using the tetanus toxoid peptide p30 (F947NNFTVSFWLRVPKVSASHLE967 (SEQ ID NO:54)) (Rice et al. *J. Immunol.* 167: 1558-65 (2001)).

EXAMPLE 7

10 Mice Immunization and tumor challenge.

All animal studies were approved by the IRBM institutional animal care and use committee. Female C57BL/6 mice (H-2^b) were purchased from Charles River (Lecco, Italy). HLA-A2.1 mice (HHD) were kindly provided by F. Lemmonier (Institute Pasteur, Paris, France). C57BL/DR4 mice were purchased from Taconic (Germantown, NY). CEA.tg mice (H-2^b) were provided by J. Primus (Vanderbilt University) and kept in standard conditions (Clarke et al. *Cancer Res.* 58:1469-77 (1998)). Fifty micrograms of plasmid DNA were electroinjected in a 50µl volume in mice quadriceps as previously described (Rizzuto et al. *Proc. Natl. Acad. Sci. U.S.A.* 96(11): 6417-22 (1999)). Ad injections were carried out in mice quadriceps in 50µl volume. Humoral and cell mediated immune response were analyzed at the indicated time.

20 C57BL/6 mice were subjected to two DNA injections in quadriceps muscle followed by electrical stimulation as previously described (Rizzuto et al. *supra*). Injections were carried out at three-week intervals. CEA transgenic mice were subjected to either 5 weekly injections of plasmid DNA (50µg/injection), 2 injections of Ad vectors (1x10⁹ viral particles/injection), or 5 weekly injections followed by a boost with Ad. Two weeks after the last injection, humoral and cell mediated immune response were analyzed. Mice were also challenged with a subcutaneous (s.c.) injection of 5x10⁵ MC38-CEA cells (Clarke et al., *supra*). At weekly intervals, mice were examined for tumor growth.

EXAMPLE 8

30 Antibody Detection and Titration.

Sera for antibody titration were obtained by retro-orbital bleeding. ELISA plates (Nunc maxisorp) were coated with 100ng/well of highly purified CEA protein (Fitzgerald), diluted in coating buffer (50mM NaHCO₃, pH 9.4) and incubated O/N at 4°C as previously described (Facciabene et al., *supra*). Plates were then blocked with PBS containing 5% BSA for 1 hr at 37°C.

Mouse sera were diluted in PBS 5% BSA (dilution 1/50 to evaluate seroconversion rate; dilutions from 1:10 to 1:31,2150 to evaluate titer). Pre-immune sera were used as background. Diluted sera were incubated O/N at 4°C. Washes were carried out with PBS 1% BSA, 0.05% Tween 20.

Secondary antibody (goat anti-mouse, IgG Peroxidase, Sigma) was diluted 1/2000 in PBS, 5% BSA and incubated 2-3 hr at RT on a shaker. After washing, plates were developed with 100µl/well of TMB substrate (Pierce Biotechnology, Inc., Rockford, IL). Reaction was stopped with 25 µl/well of 1M H₂SO₄ solution and plates were read at 450nm/620 nm. Anti-CEA serum titers were calculated as the reciprocal limiting dilution of serum producing an absorbance at least 3-fold greater than the absorbance of autologous pre-immune serum at the same dilution.

EXAMPLE 9

IFN-γ ELISPOT Assay

Assays were carried out using mouse splenocytes and CEA-specific peptides as previously described (Facciabene et al., *supra*). Ninety-six wells MAIP plates (Millipore Corp., Billerica, MA) were coated with 100 µl/ well of purified rat anti-mouse IFN-γ (IgG1, clone R4-6A2, Pharmingen) diluted to 2.5 µg/ml in sterile PBS. After washing with PBS, blocking of plates was carried out with 200 µl/well of R10 medium for 2 hrs at 37°C.

Splenocytes were obtained by removing the spleen from the euthanized mice in a sterile manner and by spleen disruption by grating on a metal grid. Red blood cells were removed by osmotic lysis by adding 1 ml of 0.1X PBS to the cell pellet and vortexing for approximately 15s. One ml of 2x PBS was then added and the volume was brought to 4ml with 1x PBS. Cells were pelleted by centrifugation at 1200 rpm for 10 min at RT, and the pellet was resuspended in 1 ml R10 medium. Viable cells were counted using Türks staining.

Splenocytes were plated at 5×10^5 and 2.5×10^5 cells/well in duplicate and incubated for 20h at 37°C with 1µg/ml suspension of each peptide. Concanavalin A (ConA) was used as positive internal control for each mouse at 5µg/ml. After washing with PBS, 0.05% Tween 20, plates were incubated O/N at 4°C with 50µl/well of biotin-conjugated rat anti-mouse IFNγ (RatIgG1, clone XMG 1.2, PharMingen) diluted to 1:2500 in assay buffer. After extensive washing, plates were developed by adding 50 µl/well NBT/B-CIP (Pierce Biotechnology, Inc., Rockford, IL) until development of spots was clearly visible. The reaction was stopped by washing plates thoroughly with distilled water. Plates were air dried and spots were then counted using an automated ELISPOT reader.

EXAMPLE 10

Intracellular Cytokine Staining.

One to two million mouse splenocytes or PBMC in 1ml RPMI 10% FCS were incubated with pool of peptides (5-6 µg/ml final concentration of each peptide) and brefeldin A (1 µg/ml; BD Pharmingen cat #555028/2300kk) at 37°C and 5% CO₂ for 12-16 hours as previously described (Facciabene et al., *supra*). Cells were then washed with FACS buffer (PBS 1% FBS, 0.01% NaN₃) and incubated with purified anti-mouse CD16/CD32 Fc block (BD Pharmingen cat # 553142) for 15 min at 4°C. Cells were then washed and stained with surface antibodies: CD4-PE conjugated anti-mouse (BD Pharmingen, cat.# 553049), PercP CD8 conjugated anti mouse (BD Pharmingen cat# 553036) and APC- conjugated anti-mouse CD3e (BD Pharmingen cat# 553066) for 30 minutes at room temperature in the dark. After the washing cells were fixed and permeabilized with Cytofix-Cytoperm Solution (BD Pharmingen cat #555028/2300kk) for 20 min at 4°C in the dark. After washing with PermWash Solution (BD Pharmingen cat #555028/2300kk) cells were incubated with the IFNγ-FITC antibodies (BD Pharmingen). Cells were then washed, fixed with formaldehyde 1% in PBS and analyzed on a FACS-Calibur flow cytometer, using CellQuest software (Becton Dickinson, San Jose, CA).

EXAMPLE 11

Immunogenicity of CEA-LT Fusions

To examine the immune responses induced by the plasmids encoding the CEA-LTA and CEA-LTB fusions, groups of 9 C57BL/6 mice were immunized with two injections i.m. of 50 µg each of plasmids pV1J/hCEA, pV1J/hCEA-LTA and pV1J/hCEA-LTB. Additionally, to verify whether coexpression of the CEA-LTA and CEA-LTB fusion proteins could have an additive effect on the immunogenicity of the CEA protein, a group of mice was immunized by coinjecting 25 µg each of plasmids pV1J/hCEA-LTA and pV1J/hCEA-LTB. Immunizations were administered three weeks apart. The plasmid DNA was routinely electroinjected into mouse skeletal muscle in view of the enhanced transduction and immunogenicity connected with this particular procedure (Zucchelli et al. *J. Virol.* 74: 11598-11607 (2000); Widera et al. *J. Immunol.* 164: 4635-4640 (2000)).

The cellular immunity elicited by the different plasmids was measured by ELISPOT assay 2 weeks after the last injection. Antigen-specific IFNγ secretion from stimulated splenocytes was measured using four pools of 15mer peptides overlapping by 11 aa and encompassing the entire CEA glycoprotein. Pool A covers aa 1 to 147, pool B aa 137 to 237, pool C aa 317 to 507, and pool D aa 497 to 703. As a negative control, cytokine production was also measured upon stimulation of the splenocytes with DMSO at the same concentration utilized to solubilize the CEA peptides.

The immune response elicited by DNA vaccination in C57BL/6 mice was primarily biased towards the C-terminal region of the protein since the SFC values detected with the peptide pool A were slightly above background with all constructs (FIGURE 10). The pV1J/hCEA-LTB vaccination regimen was superior to that elicited by pV1J/hCEA as indicated by the higher geometric mean values of the SFC detected with peptide pools B, C and D (pV1J/hCEA-LTB: 482, 1436, and 2054 SFC/10⁶ splenocytes, respectively; pV1J/hCEA: 45, 350, and 264 SFC/10⁶ splenocytes, respectively). Similarly, plasmid pV1J/hCEA-LTA had also an enhancing effect on the CEA specific immune response when compared to pV1J/hCEA. However, the increase in immune response was only observed with peptide pools C and D (925 and 528 SFC/10⁶ splenocytes, respectively), while the immune response measured with peptide pool B was low (15 SFC/10⁶ splenocytes). Additionally, coinjection of plasmids pV1J/hCEA-LTA and pV1J/hCEA-LTB did not have a significant synergic effect on the immune response to CEA when compared to the immune response measured in the pV1J/hCEA-LTB treated group, but rather, it resulted in a decrease of the SFC values detected with peptides pool B and D (210 and 528 SFC/10⁶ splenocytes, respectively).

To define the T-cell specificity elicited upon vaccination with the different CEA constructs, IFN γ intracellular staining was carried out on pooled splenocytes from injected mice using peptide pool D. A CD8⁺-specific response was detected in mice injected with pV1J/hCEA-LTB (4.5%) superior to that detected with pV1J/hCEA-LTA and pV1J/hCEA (0.14% and 0.8%, respectively, FIGURE 10B). In contrast, pV1J/hCEA-LTA elicited a strong CD4⁺-specific response (1.21%) greater than that observed with pV1J/hCEA-LTB and pV1J/hCEA (0.55% and 0.58%, respectively).

The induction of the humoral immune response to CEA was examined by measuring antigen specific antibodies (FIGURE 11). Both plasmids pV1J/hCEA-LTA and pV1J/hCEA-LTB elicited a greater antibody response than pV1J/hCEA, confirming the adjuvant effect exerted by the LT subunits on the CEA specific immune response. Thus, these data demonstrate that fusion of the CEA coding sequence to the LTA or LTB cDNA results in an increase the CEA specific immune response. However, LTB appears to have a greater enhancing effect on the immune response with a prevalent induction of CD8⁺ T cells, whereas LTA elicits a predominant CD4⁺ response.

EXAMPLE 12

Immunogenicity of CEA-LTB Fusions in Different Mouse Strains.

To determine whether the enhancing effect of the LT subunits on the CEA specific immune response was not limited to a single mouse genetic background, DNA based immunizations were carried out in BALB/c, C57/DR4 and HLA-A2.1 (HHD) mice. The BALB/c mice were chosen

in view of their immunocompetence, being a mouse strain extremely reactive to immunization regimens of various sorts. The HHD transgenic mice express the human MHC class I genes. Similarly, C57/DR4 transgenic mice carry the human MHC class II genes. Thus, these two transgenic mouse strains may provide information as to the immunoreactivity of the CEA-LT fusions in the context of human MHC class I and II haplotypes.

The CEA specific immune response in BALB/c mice was first assessed by ELISPOT assay. Enhancement of the antigen specific immune response upon immunization with plasmid pV1J/hCEA-LTB was detected with peptide pools A, B, C, D (pV1J/hCEA-LTB: 166, 1353, 796, 899 SFC/10⁶ splenocytes, respectively; pV1J/hCEA: 57, 312, 327, 318, SFC/10⁶ splenocytes respectively, FIGURE 12). As observed in the C57BL/6 mice, the N-terminal region of the CEA protein appeared to be the least immunogenic as compared to other sections of the tumor antigen. pV1J/hCEA-LTA immunization also yielded an increase in the antigen specific immune response as compared to pV1J/hCEA. The increase in the immune response was detected with peptide pools B, C and D (936, 727, and 650 SFC/10⁶ splenocytes, respectively). Additionally, coinjection of the two plasmids pV1J/hCEA-LTA and pV1J/hCEA-LTB yielded a significant additive effect that was detected mainly with peptide pools C and D (1783 and 2141 SFC/10⁶ splenocytes, respectively).

The CEA specific immune response in C57/DR4 mice was considerably enhanced by the immunization with pV1J/hCEA-LTB, and was detected only peptide pool D (FIGURE 13). IFN γ intracellular staining performed on pooled PBMC from injected mice showed that the CD8⁺ response to CEA was highest in mice immunized with pV1J/hCEA-LTB (15.32%), whereas was very weak in the pV1J/hCEA treated group (0.5%). pV1J/hCEA-LTA immunization increased the antigen specific immune response only moderately (0.43%), and did not further enhance the CEA immunogenicity when coinjected with the construct encoding the CEA-LTB fusion (13.44 %). Interestingly, no significant CD4⁺ T cell response was detected in the immunized mice (data not shown).

The immune response elicited by the different CEA encoding plasmids was assessed in HHD mice by performing IFN γ intracellular staining on pooled PBMC. The immune response was only detected with peptide pools B and D, and as shown in FIGURE 14, immunization with pV1J/hCEA-LTB resulted in more than 10 fold increase in the CD8⁺ response to the target antigens. In contrast, no increase in the immune response was detected using pV1J/hCEA-LTA either alone or upon coinjection with pV1J/hCEA-LTB. No CD4⁺ T cell response was detected in the immunized mice (data not shown).

Taken together, these data confirm that fusion of the LTB coding sequence to CEA results in a considerable increase in the antigen specific immune response. Interestingly, this

response is predominantly CD8⁺-specific and can be observed in different mouse strains, thus indicating that the enhancing effect exerted by the LT subunit is not genotype restricted.

EXAMPLE 13

5 Tolerance to Human CEA in Transgenic Mice.

To determine whether the enhanced immunogenic properties of the hCEA-LTB fusion would break tolerance more efficiently to human CEA, hCEA transgenic mice were immunized with vectors carrying either the fully codon optimized cDNA of hCEA or CEA-LTB. These transgenic mice carry the entire human CEA gene and flanking sequences and express the hCEA
10 protein in the intestine, mainly in the cecum and colon. Thus, this mouse line is a useful model for studying the safety and efficacy of immunotherapy strategies directed against this tumour self antigen (Clarke et al., *Cancer Research* 58: 1469-1477 (1998)).

Immunization with pV1J/hCEA-LTB_{opt} resulted in a significant increase in the CEA specific immune response measured by IFN γ intracellular staining on PBMC of the injected mice
15 (FIGURE 15A). The enhancement of the T cell response was detected with peptide pool D and was predominantly CD8⁺. Additionally, also the CEA specific humoral response was increased in the CEA-LTB treated mice as shown by the 47 fold increase in the geometric mean values of the Ab titer as compared to the pV1J/hCEA_{opt} treated group (FIGURE 15B).

To determine whether the enhancing effect exerted by LTB on the CEA specific
20 immune response could also be observed upon immunization with vectors other than plasmid DNA, groups of 12 CEA tg mice were immunized with Ad5/hCEA_{opt}-LTB and Ad/hCEA_{opt} at a dose of 1x10⁷, 1x10⁸, and 1x10⁹ vp. Mice were subjected to two injections two weeks apart and the immune response was measured by IFN γ intracellular staining on PBMC two weeks after the last injection. The immune response was assessed using the peptide pool D. Ad/hCEA_{opt}-LTB was more
25 immunogenic than Ad/hCEA_{opt} since significant immune responses to CEA could be detected with the 1x10⁸vp dose, whereas 1x10⁹ vp of Ad/hCEA_{opt} were necessary to break tolerance to the target antigen (FIGURE 16). No CD4⁺ response could be detected in any of the immunized mice (data not shown).

These data confirm that tolerance to this self antigen can be broken more efficiently
30 due to the increased immunogenic properties of the CEA-LTB fusion. Furthermore, the enhancing effect of LTB on the immunogenic properties of CEA is also observable upon injection of plasmid carrying the fully codon optimized cDNA of the CEA-LTB fusion. Lastly, these results indicate that that enhanced immunogenicity of CEA-LTB is not limited to plasmid DNA immunization.

EXAMPLE 14

Tumor Growth Kinetics in CEA Transgenic Mice Immunized with CEA-LTB Fusions

It was deemed appropriate to ascertain whether the increased immunogenicity of the CEA-LTB fusion would also lead to an enhanced therapeutic effect capable of interfering with tumor progression. For this purpose, groups of 10 CEA-tg mice were subjected to 5 weekly injections of plasmids pV1J/hCEAopt or pV1J/CEAopt-LTB followed by a final boost with 1×10^{10} vp of the corresponding Ad vector. In view of recent reports that indicate that high levels of cellular immunity can be induced against viral and bacterial antigens by utilizing plasmid DNA prime-Ad boost modality, the same immunization protocol was employed in this study. Two weeks after the last immunization, the CEA tg mice were challenged with a subcutaneous injection of 5×10^5 MC38-CEA tumor cells. This syngenic cell line was derived from a chemically induced colon cancer and expresses CEA. Tumor development in mock treated mice was detected by 22 days post challenge as all the treated mice were no longer tumor free (FIGURE 17A). Additionally, there was a concomitant increase in the average size of the tumor mass that reached significant volume by 34 days post challenge. Mice vaccinated with vectors encoding pV1J/hCEAopt showed a partial resistance to tumor development since 2 out of 10 treated mice remained tumor free at day 34 post challenge. The average size of the tumors of this group was smaller than that observed in the mock treated mice. Immunization with vectors encoding the CEAopt-LTB fusion resulted in a significant protective effect from tumor development. Five out of 10 treated mice remained tumor free at day 34 post challenge, and the average size of the tumor mass in this group was significantly smaller than that observed in the mock or pV1J/hCEAopt treated mice. Thus, these results indicate that the enhanced CEA-specific immune response associated with vectors encoding the CEA-LTB fusion correlates with a significant antitumor effect resulting in partial protection from tumor growth and reduced growth kinetics of the tumor mass.

EXAMPLE 15

CEA-DOM and CEA-FcIgG fusions enhance the immunogenicity of the CEA protein.

To examine the immune responses induced by the plasmids encoding CEA-FRC, CEA-DOM, CEA-VSV-G, CEA-FcIgG, CEA-HSP70 and CEA-LAMP fusions, groups of 9 C57BL/6 mice were immunized with two i.m. injections of 50 or 5 μ g of each plasmid. The immunizations were three weeks apart. In view of the enhanced transduction and immunogenicity reported with electroporation (Zucchelli et al. *J. Virology* 74:11598 (2000), Widera et al., *J. Immunol.* 164: 4635 (2000)), plasmid DNAs were routinely electroporated (DNA-EP) into mouse skeletal muscle.

The immune response elicited by different plasmids was measured by IFN γ ELISPOT assay, 2 weeks after the last injection. Antigen-specific IFN γ secretion from stimulated splenocytes was measured using a pool of 15mer peptides overlapping by 11 aa and encompassing the C-terminal region of CEA (pool D, aa 497-703) (Zucchelli et al., *supra*). The analysis of the immune response to CEA was carried out with peptide pool D since the cellular immune response to CEA in C57BL/6 mice is primarily biased towards the C-terminal region of this protein (Zuccelli et al., *supra*). As a negative control, cytokine production was also measured upon stimulation of the splenocytes with DMSO at the same concentration utilized to solubilize the CEA peptides.

Injection of pV1J/CEA-DOM or pV1J/CEA-Fc elicited a greater immune response to CEA as compared to pV1J/CEA. The greater immunogenicity of these two fusion proteins resulted in higher geometric mean values of spot forming cells (SFC) per 10⁶ splenocytes (FIGURE 20A). Plasmids pV1J/CEA-DOM and pV1J/CEA-FcIgG had similar immunogenic properties and exerted a 3- to 4-fold increase in CEA-specific immune responses upon injection of 5 or 50 μ g of plasmid DNA (pV1J/CEA-DOM: 590 and 1098 SFC/10⁶ splenocytes, pV1J/CEA-FcIgG: 510 and 1160, pV1J/CEA: 146 and 264 SFC/10⁶ splenocytes, respectively). No significant differences were noted between the SFC values elicited by the pV1J/CEA-FrC, pV1J/CEA-LAMP, pV1J/CEA-HSP70 and pV1J/CEA. No CEA specific immune responses were detected in negative control samples.

To determine the effect of the CEA-fusions on the humoral response to CEA, sera from immunized mice were tested in ELISA using purified CEA protein as substrate (FIGURE 20B). An increase in CEA-specific antibody titer was observed upon injection of 50 μ g of plasmids pV1J/CEA-DOM, pV1J/CEA-FcIgG, pV1J/CEA-FrC and pV1J/CEA-HSP70. On the contrary, injection of pV1J/CEA-LAMP and pV1J/CEA-VSV-G resulted in a CEA-specific antibody response similar to that observed upon immunization with pV1J/CEA. Taken together, these data demonstrate that fusion of the CEA coding sequence to the DOM or FcIgG cDNA results in an increase in the CEA-specific cell mediated and humoral immune response.

EXAMPLE 16

CEA-DOM and CEA-FcIgG fusions break tolerance to target antigen in CEA transgenic mice.

Tolerance to the target antigen is one of the hurdles that a cancer vaccine must overcome to elicit an immune response and to exert an efficient antitumor effect. Thus, it was deemed appropriate to determine whether the enhanced immunogenic properties of CEA-DOM and CEA-FcIgG fusions would break tolerance to CEA more efficiently than the CEA protein. To this end, CEA transgenic mice were utilized to perform comparative immunization studies. These transgenic mice carry the entire human CEA gene and flanking sequences and express the CEA

protein in the intestine, mainly in the cecum and colon. This mouse line is a useful model for studying the safety and efficacy of immunotherapy strategies directed against this tumor self antigen (Clarke et al., *supra*).

In view of the enhanced immunogenic properties of vectors carrying the codon usage optimized cDNA (cDNAopt) of CEA, both plasmid and Adenovirus vectors were engineered to carry the cDNAopt of the CEA-DOM (CEA-DOMopt) or CEA-FcIgG (CEA-FcIgGopt) fusions. As observed for CEA, CEA-DOMopt and CEA-FcIgGopt cDNAs were shown to be expressed with a greater efficiency of the corresponding wild type cDNA leading to an enhanced immune response to CEA (data not shown).

The immunogenicity of these two fusion proteins was compared to that of CEA by a series of immunization studies based on the use of plasmid DNA and Ad vectors administered either alone or in combination. Cohorts of CEA transgenic mice were immunized with the following varying regimens: i) 5 injections at weekly intervals of 50 µg of plasmid DNA (DNA/DNA), ii) 2 biweekly injections of Adenovirus in doses ranging from 1×10^7 to 1×10^9 viral particles (vp) of Adenovirus (Ad/Ad), or iii) 5 weekly injections of plasmid DNA followed by a final injection of 1×10^9 vp of Adenovirus (DNA/Ad). Immune responses were analyzed by intracellular IFN γ staining on PBMC or splenocytes of each immunized mouse using pool D peptides. Additionally, the induction of CEA-specific antibodies was monitored by ELISA.

DNA/DNA immunization of the CEA transgenic mice revealed that the CEA-DOMopt and CEA-FcIgGopt vectors exerted a measurable CD8⁺ T cell response to the target antigen (FIGURE 21A). Thus, both constructs were able to break tolerance to CEA in these mice. The antigen specific response elicited by CEA-DOM and CEA-FcIgG fusion proteins was comparable as indicated by the average values of IFN γ intracellular staining (0.22 and 0.34%, respectively). Nonetheless, the immune response elicited by these two constructs was greater than that observed upon vaccination with pV1J/CEAopt (0.07%). Similarly, anti-CEA humoral response was also greater upon vaccination with the fusion proteins. CEA-specific antibody titer was detected in all mice immunized with pV1J/CEA-DOMopt and pV1J/CEA-FcIgGopt and the average of the antibody titer was 56,136 and 24,725, respectively. By contrast, the pV1J/CEAopt immunized group showed an at least a 77 fold lower average titer of CEA-specific antibody (318) (FIGURE 21B).

CEA transgenic mice treated with the Ad/Ad vaccination regimen also showed a better efficiency in breaking tolerance to CEA upon vaccination with CEA-DOMopt and CEA-FcIgGopt Ad vectors than with Ad-CEAopt. A CEA-specific CD8⁺ T cell response could be observed in the vaccinated mice upon injection of a little as 10^7 vp of Ad-CEA-DOM or Ad-CEA-FcIgG, the CEA-specific response was comparable between the two antigens, and increased upon injection of

10⁹ vp (1.55% and 1.15%, respectively). By contrast, 10⁹ vp of Ad-CEAopt were necessary to elicit significant CD8⁺ T-cell precursor frequencies (2.1%) (FIGURE 21C). CEA-specific antibodies were detected in all mice immunized with Ad-CEA-DOMopt and Ad-CEA-FcIgGopt. The averages of the antibody titer were 19,600 and 33,000, respectively. Injection of Ad-CEAopt resulted in a measurable
5 CEA-specific response in only 2 of the treated mice, and the antibody titer was significantly lower (Zucchelli et al., *supra*) (FIGURE 21D). Interestingly, the DNA/Ad immunization showed reduced differences in the CD8⁺ T cell precursor frequencies elicited by CEA, CEA-DOM and CEA-FcIgG vectors (FIGURE 22A). However, averages of CEA-specific antibody titers were greater upon vaccination with vectors expressing CEA-DOM and CEA-FcIgG than CEA (31,200, 26,120 and 412,
10 respectively) (FIGURE 22B).

Interestingly, regardless of the antigen, no obvious CD4⁺ cell Th1 response to CEA was detected in any of the three vaccination regimens (data not shown). However, significant CD4⁺ cell Th1 response against to the helper epitope, p30, present within DOM sequences (Rice et al., *J. Immunol.* 167: 1558-65 (2001)) were detected after DNA/DNA vaccination (0.4%) (FIGURE 23).

15 Thus, these data demonstrate that the CEA-DOM and CEA-FcIgG fusion proteins can break tolerance to CEA in transgenic mice with greater efficacy than the CEA protein. The enhanced immunogenic properties of these fusion proteins can be observed upon immunization with DNA or Ad vectors. However, the greater ability of these two fusion proteins in eliciting CD8⁺ T cells to CEA can be overcome, at least in part, by DNA/Ad vaccination regimen.

20 EXAMPLE 17

T-cell depletion studies.

Immunized animals were depleted of CD4⁺ T cells, CD8⁺ T cells, NK cells, by i.p. injection of anti-CD4 (GK1.5 hybridoma), anti-CD8 (Lyt 2.2 hybridoma), or anti Asialo GM1 (Wako
25 Chemicals, Richmond, VA) as described (Perricone et al., *J. Immunother.* 27(4):273-81 (2004); Yoon et al., *J. Ethnopharmacol.* 93 (2-3):247-53 (2004)). Antibodies (100 μ l diluted ascitic fluid/dose) were injected on day -7 relative to the tumor challenge and then injected every week for 3 weeks after injection of 5x10⁵ MC38-CEA cells. Depletion conditions were validated by flow cytometry analysis of peripheral blood using phycoerythrin-conjugated MAbs anti-CD4, anti-CD8, and anti-NK
30 (PharMingen, San Diego, CA); 99% of the relevant cell subset was depleted, whereas all other subsets remained within normal levels.

EXAMPLE 18

CEA-DOM immunization exerts an antitumor effect in CEA transgenic mice.

We next ascertained whether the increased immunogenicity of the CEA-DOM or CEA-FcIgG fusions would also lead to an enhanced therapeutic effect, capable of interfering with tumor progression. For this purpose, groups of 10 CEA transgenic mice were subjected to the DNA/DNA, Ad/Ad or DNA/Ad immunization regimens with the CEA-DOM, CEA-FcIgG, or CEA vectors. Two weeks after the last immunization, the CEA transgenic mice were challenged with a subcutaneous injection of 5×10^5 MC38-CEA cells, a syngenic tumor cell line that expresses CEA (Clarke et al., *supra*). Immunization with DNA/DNA or Ad/Ad modalities did not result in any significant antitumor effect, regardless of the protein expressed by the injected vectors (FIGURE 24). In contrast, DNA-EP/Ad immunization with vectors encoding the CEA-DOM fusion protein resulted in a significant antitumor effect with 7 out of 10 treated mice remaining tumor free by day 34 post challenge. Thus, these results indicate that the enhanced CEA-specific immune response associated with the CEA-DOMopt cDNA, and the DNA/Ad vaccination regimen correlate with a significant antitumor effect in CEA transgenic mice.

EXAMPLE 19

The CEA-DOM antitumor effect is dependent on $CD4^+$ T cells, $CD8^+$ T cells and NK cells.

The effector cells involved in the antitumor effect observed upon DNA-EP and Ad immunization with vectors encoding CEA-DOM fusion were characterized. After DNA/Ad immunization, but prior to tumor challenge, mice were depleted of $CD4^+$, $CD8^+$ T cells, or NK cells by MAbs. Antibodies were given during the course of tumor challenge to ensure continued depletion of the relevant NK and T cell subsets. The depletion of all three cell types was monitored by flow cytometry analysis using antibodies specific for cell surface markers (data not shown). Depletion of $CD4^+$, $CD8^+$ T cells, or NK cells had a negative effect on survival of the immunized mice resulting in the drastic reduction of tumor-free mice as compared to the vaccinated group (FIGURE 25). Thus, these data indicate that NK, $CD4^+$ and $CD8^+$ T cells play an important role in the antitumor effect exerted by CEA-DOM vaccination.

EXAMPLE 20

Statistical analysis.

Where indicated, results were analyzed by the log rank or two tailed Student t test. A p value < 0.05 was considered significant.

WHAT IS CLAIMED IS :

1. A nucleic acid molecule comprising a sequence of nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a C-terminally truncated human CEA protein which lacks amino acids 679 - 702 of SEQ ID NO:20, fused to an immunoenhancing element which is the B subunit of the heat labile enterotoxin of *E. coli* (LTB) optionally truncated of its signal sequence; and wherein the fusion protein is capable of producing an immune response in a mammal.
2. The nucleic acid molecule of claim 1, wherein the LT subunit B is truncated of its signal sequence.
3. The nucleic acid molecule of claim 2, wherein the sequence of nucleotides comprises a sequence of nucleotides as set forth in SEQ ID NO:9, SEQ ID NO:11, or SEQ ID NO:12.
4. The nucleic acid molecule of claim 3, wherein the sequence of nucleotides is as set forth in SEQ ID NO:12.
5. The nucleic acid molecule of claim 2, wherein the C-terminal end of the CEA protein is fused to the N-terminal end of LT subunit B.
6. A vector comprising the nucleic acid molecule of claim 1.
7. The vector of claim 6, wherein the vector is an adenovirus vector or a plasmid vector.
8. The vector of claim 7, wherein the vector is an Ad 5 vector.
9. The vector of claim 7, wherein the vector is an Ad 6 vector or an Ad 24 vector.
10. The vector of claim 7, wherein the vector is a chimp Ad vector.
11. The vector of claim 7, wherein the vector is pV1JnsB.
12. A host cell comprising the vector of claim 7.
13. A process for expressing a CEA fusion protein in a recombinant host cell, comprising:
 - (a) introducing a vector comprising the nucleic acid molecule of any one of claims 1-5 into a suitable host cell; and,
 - (b) culturing the host cell under conditions which allow expression of said CEA fusion protein.

14. A purified CEA fusion protein encoded by the nucleic acid molecule of any one of claims 1-5.

15. The purified CEA fusion protein of claim 14, wherein the fusion protein comprises a sequence of amino acids selected from the group consisting of: SEQ ID NOS:10 and 13.

16. An adenovirus vaccine vector comprising an adenoviral genome with a deletion in the E1 region, and an insert in the E1 region, wherein the insert comprises an expression cassette comprising:

(a) a polynucleotide comprising a sequence of nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a C-terminally truncated human CEA protein which lacks amino acids 679 - 702 of SEQ ID NO:20, fused to an immunoenhancing element which is the B subunit of the heat labile enterotoxin of *E. coli* (LTB) optionally truncated of its signal sequence; and wherein the fusion protein is capable of producing an immune response in a mammal; and

(b) a promoter operably linked to the polynucleotide.

17. The adenovirus vector of claim 16, which is an Ad 5 vector, an Ad 6 vector or an Ad 24 vector.

18. A vaccine plasmid comprising a plasmid portion and an expression cassette portion, the expression cassette portion comprising:

(a) a polynucleotide comprising a sequence of nucleotides that encodes a CEA fusion protein, wherein the CEA fusion protein comprises a C-terminally truncated human CEA protein which lacks amino acids 679 - 702 of SEQ ID NO:20, fused to an immunoenhancing element which is the B subunit of the heat labile enterotoxin of *E. coli* (LTB) optionally truncated of its signal sequence; and wherein the fusion protein is capable of producing an immune response in a mammal; and,

(b) a promoter operably linked to the polynucleotide.

19. A vaccine vector comprising the nucleic acid molecule of any one of claims 1-5 for use in preventing or treating cancer in a mammal.

20. A vaccine vector according to claim 19, wherein the vector is in accordance with any one of claims 16-18.

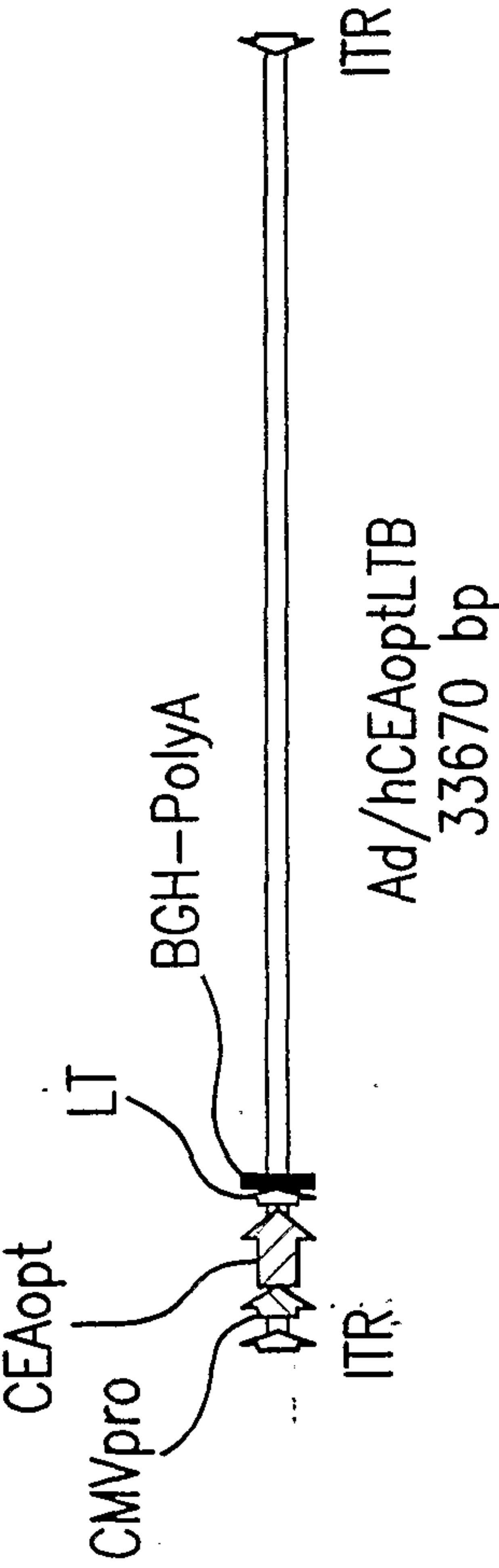
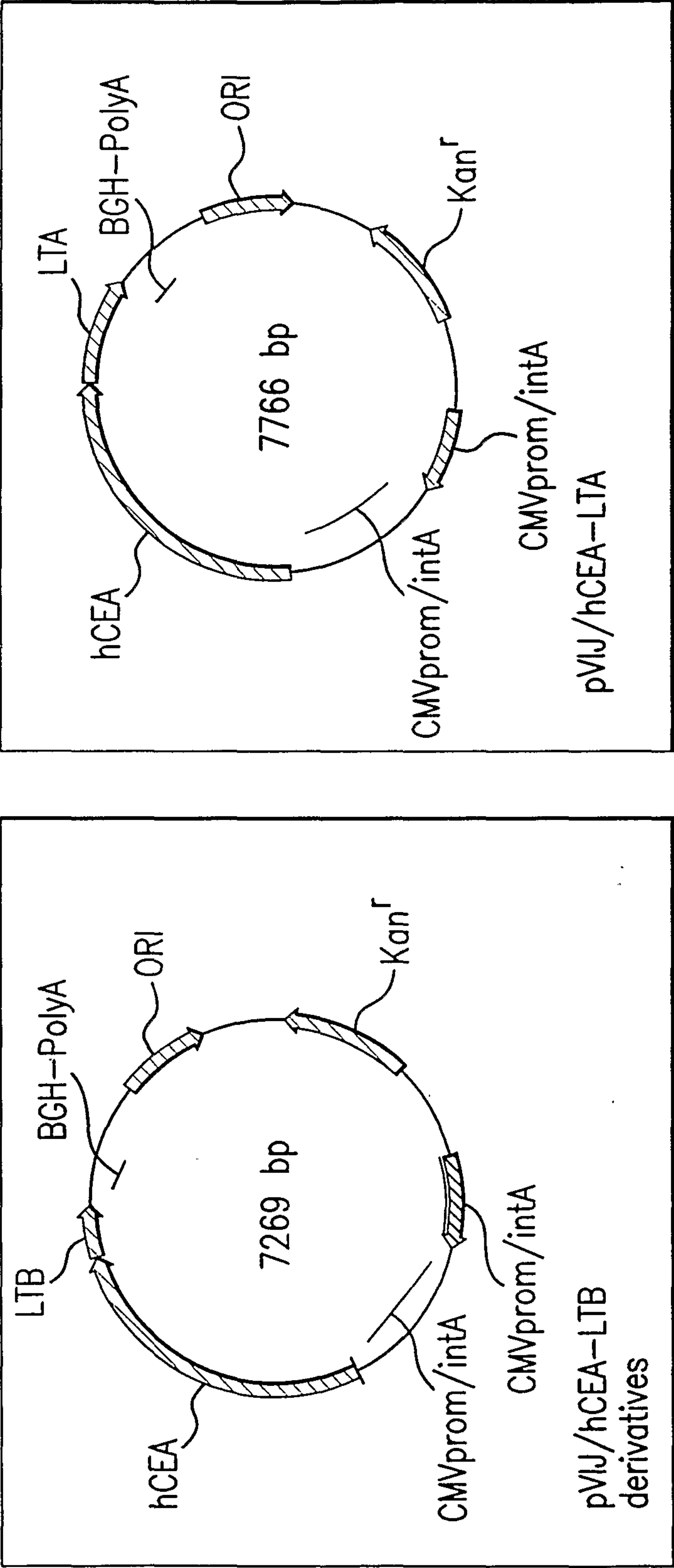


FIG.1

2/62

hCEA-LTA Nucleotide Sequence

```
1   ATGGAGTCTC CCTCGGCCCC TCCCCACAGA TGGTGCATCC CCTGGCAGAG
51  GCTCCTGCTC ACAGCCTCAC TTCTAACCTT CTGGAACCCG CCCACCACTG
101 CCAAGCTCAC TATTGAATCC ACGCCGTTCA ATGTCGCAGA GGGGAAGGAG
151 GTGCTTCTAC TTGTCCACAA TCTGCCCCAG CATCTTTTGT GCTACAGCTG
201 GTACAAAGGT GAAAGAGTGG ATGGCAACCG TCAAATTATA GGATATGTAA
251 TAGGAACTCA ACAAGCTACC CCAGGGCCCG CATACAGTGG TCGAGAGATA
301 ATATACCCCA ATGCATCCCT GCTGATCCAG AACATCATCC AGAATGACAC
351 AGGATTCTAC ACCCTACACG TCATAAAGTC AGATCTTGTG AATGAAGAAG
401 CAACTGGCCA GTTCCGGGTA TACCCGGAGC TGCCCAAGCC CTCCATCTCC
451 AGCAACAACCT CCAAACCCGT GGAGGACAAG GATGCTGTGG CCTTCACCTG
501 TGAACCTGAG ACTCAGGACG CAACCTACCT GTGGTGGGTA AACAATCAGA
551 GCCTCCCGGT CAGTCCCAGG CTGCAGCTGT CCAATGGCAA CAGGACCCTC
601 ACTCTATTCA ATGTCACAAG AAATGACACA GCAAGCTACA AATGTGAAAC
651 CCAGAACCCA GTGAGTGCCA GGCGCAGTGA TTCAGTCATC CTGAATGTCC
701 TCTATGGCCC GGATGCCCCC ACCATTTCCC CTCTAAACAC ATCTTACAGA
751 TCAGGGGAAA ATCTGAACCT CTCCTGCCAC GCAGCCTCTA ACCCACCTGC
801 ACAGTACTCT TGGTTTGTCA ATGGGACTTT CCAGCAATCC ACCCAAGAGC
851 TCTTTATCCC CAACATCACT GTGAATAATA GTGGATCCTA TACGTGCCAA
901 GCCCATAACT CAGACACTGG CCTCAATAGG ACCACAGTCA CGACGATCAC
951 AGTCTATGCA GAGCCACCCA AACCCTTCAT CACCAGCAAC AACTCCAACC
1001 CCGTGGAGGA TGAGGATGCT GTAGCCTTAA CCTGTGAACC TGAGATTTCAG
1051 AACACAACCT ACCTGTGGTG GGTAATAAAT CAGAGCCTCC CGGTCAGTCC
1101 CAGGCTGCAG CTGTCCAATG ACAACAGGAC CCTCACTCTA CTCAGTGTCA
1151 CAAGGAATGA TGTAGGACCC TATGAGTGTG GAATCCAGAA CGAATTAAGT
1201 GTTGACCACA GCGACCCAGT CATCCTGAAT GTCCTCTATG GCCCAGACGA
1251 CCCACCATTT TCCCCCTCAT ACACCTATTA CCGTCCAGGG GTGAACCTCA
1301 GCCTCTCCTG CCATGCAGCC TCTAACCCAC CTGCACAGTA TTCTTGGCTG
1351 ATTGATGGGA ACATCCAGCA ACACACACAA GAGCTCTTTA TCTCCAACAT
1401 CACTGAGAAG AACAGCGGAC TCTATACCTG CCAGGCCAAT AACTCAGCCA
1451 GTGGCCACAG CAGGACTACA GTCAAGACAA TCACAGTCTC TGC GGAGCTG
1501 CCCAAGCCCT CCATCTCCAG CAACAACCTC AAACCCGTGG AGGACAAGGA
1515 TGCTGTGGCC TTCACCTGTG AACCTGAGGC TCAGAACACA AACTACCTGT
1601 GGTGGGTAAA TGGTCAGAGC CTCCCAGTCA GTCCCAGGCT GCAGCTGTCC
1651 AATGGCAACA GGACCCTCAC TCTATTCAAT GTCACAAGAA ATGACGCAAG
1701 AGCCTATGTA TGTGGAATCC AGAACTCAGT GAGTGCAAAC CGCAGTGACC
1751 CAGTCACCCT GGATGTCCTC TATGGGCCGG ACACCCCATC CATTTCCTCC
1801 CCAGACTCGT CTTACCTTTC GGGAGCGAAC CTCAACCTCT CCTGCCACTC
1851 GGCCTCTAAC CCATCCCCGC AGTATTCTTG GCGTATCAAT GGGATACCGC
1901 AGCAACACAC ACAAGTTCTC TTTATCGCCA AAATCACGCC AAATAATAAC
1951 GGGACCTATG CCTGTTTTGT CTCTAACTTG GCTACTGGCC GCAATAATTC
```

FIG.2A-1

3/62

2001 CATAGTCAAG AGCATCACAG TCTCTGCATC TGGAACCTCTA GTTAATGGCG
2051 ACAAATTATA CCGTGCTGAC TCTAGACCCC CAGATGAAAT AAAACGTTCC
2101 GGAGGTCTTA TGCCCAGAGG GCATAATGAG TACTTCGATA GAGGAACTCA
2151 AATGAATATT AATCTTTATG ATCACGCGAG AGGAACACAA ACCGGCTTTG
2201 TCAGATATGA TGACGGATAT GTTTCCACTT CTCTTAGTTT GAGAAGTGCT
2251 CACTTAGCAG GACAGTCTAT ATTATCAGGA TATTCCACTT ACTATATATA
2301 TGTTATAGCG ACAGCACCAA ATATGTTTAA TGTTAATGAT GTATTAGGCG
2351 TATACAGCCC TCACCCATAT GAACAGGAGG TTTCTGCGTT AGGTGGAATA
2401 CCATATTCTC AGATATATGG ATGGTATCGT GTTAATTTTG GTGTAATTGA
2451 TGAACGATTA CATCGTAACA GGAATATAG AGACCGGTAT TACAGAAATC
2501 TGAATATAGC TCCGGCAGAG GATGGTTACA GATTAGCAGG TTTCCCACCG
2551 GATCACCAAG CTTGGAGAGA AGAACCCTGG ATTCATCATG CACCACAAGG
2601 TTGTGGAAAT TCATCAAGAA CAATTACAGA TGATACTTGT AATGAGGAGA
2651 CCCAGAATCT GAGCACAATA TATCTCAGGA AATATCAATC AAAAGTTAAG
2701 AGGCAGATAT TTTCAGACTA TCAGTCAGAG GTTGACATAT ATAACAGAAT
TCGGGATGAA TTATGA (SEQ ID NO:7)

FIG.2A-2

4/62

CEA-LTA Amino Acid Sequence

1	MESPSAPPHR	WCIPWQRLLL	TASLLTFWNP	PTTAKLTIES	TPFNVAEGKE
51	VLLLVHNL PQ	HLFGYSWYKG	ERVDGNRQII	GYVIGTQQAT	PGPAYSGREI
101	IYPNASLLIQ	NIIQNDTG FY	TLHVIKSDLV	NEEATGQFRV	YPELPKPSIS
151	SNNSKPVEDK	DAVAFTCEPE	TQDATYLWWV	NNQSLPVSPR	LQLSNGNRTL
201	TLFNVTRNDT	ASYKCETQNP	VSARRSDSVI	LNVL YGPDAP	TISPLNTSYR
251	SGENLNLSCH	AASNPPAQYS	W FVNGTFQQS	TQELFIPNIT	VNNSGSYTCQ
301	AHNSDTGLNR	TTVTTITVYA	EPPKPFITSN	NSNPVEDEDA	VALTCEPEIQ
351	NTTYLWWVNN	QSLPVSPRLQ	LSNDNRTLTL	LSVTRNDVGP	YECGIQNELS
401	VDHSDPVILN	VLYGPDDPTI	SPSYTYRPG	VNLSLSCHAA	SNPPAQYSWL
451	IDGNIQQHTQ	ELFISNITEK	NSGLYTCQAN	NSASGHSRTT	VKTITVSAEL
501	PKPSISSNNS	KPVEDKDAVA	FTCEPEAQNT	TYLWWVNGQS	LPVSPRLQLS
551	NGNRTLTLFN	VTRNDARAYV	CGIQNSVSAN	RSDPVTLDVL	YGPDTPIISP
601	PDSSYLSGAN	LNLSCHSASN	PSPQYSWRIN	GIPQQHTQVL	FIAKITPNNN
651	GTYACFVSNL	ATGRNNSIVK	SITVSASGTL	VNGDKLYRAD	SRPPDEIKRS
701	GGLMPRGHNE	YFDRGTQMNI	NLYDHARGTQ	TGFVRYDDGY	VSTSLSLRSA
751	HLAQQSILSG	YSTYYIYVIA	TAPNMFNVND	VLGVYSPHPY	EQEVSALGGI
801	PYSQIYGWYR	VNFGVIDERL	HRNREYRDRY	YRNLNIAPAE	DGYRLAGFPP
851	DHQAWREEPW	IHHAPQGCGN	SSRTITDDTC	NEETQNLSTI	YLRKYQSKVK
901	RQIFSDYQSE	VDIYNRIRDE	L (SEQ ID NO:8)		

FIG.2B

5/62

hCEA-LTB Coding Sequence

1	ATGGAGTCTC	CCTCGGCCCC	TCCCCACAGA	TGGTGCATCC	CCTGGCAGAG
51	GCTCCTGCTC	ACAGCCTCAC	TTCTAACCTT	CTGGAACCCG	CCCACCACTG
101	CCAAGCTCAC	TATTGAATCC	ACGCCGTTCA	ATGTCGCAGA	GGGGAAGGAG
151	GTGCTTCTAC	TTGTCCACAA	TCTGCCCCAG	CATCTTTTTG	GCTACAGCTG
201	GTACAAAGGT	GAAAGAGTGG	ATGGCAACCG	TCAAATTATA	GGATATGTAA
251	TAGGAACTCA	ACAAGCTACC	CCAGGGCCCCG	CATACAGTGG	TCGAGAGATA
301	ATATACCCCA	ATGCATCCCT	GCTGATCCAG	AACATCATCC	AGAATGACAC
351	AGGATTCTAC	ACCCTACACG	TCATAAAGTC	AGATCTTGTG	AATGAAGAAG
401	CAACTGGCCA	GTTCCGGGTA	TACCCGGAGC	TGCCCAAGCC	CTCCATCTCC
451	AGCAACAAC	TCAAACCCGT	GGAGGACAAG	GATGCTGTGG	CCTTCACCTG
501	TGAACCTGAG	ACTCAGGACG	CAACCTACCT	GTGGTGGGTA	AACAATCAGA
551	GCCTCCCGGT	CAGTCCCAGG	CTGCAGCTGT	CCAATGGCAA	CAGGACCCTC
601	ACTCTATTCA	ATGTCACAAG	AAATGACACA	GCAAGCTACA	AATGTGAAAC
651	CCAGAACCCA	GTGAGTGCCA	GGCGCAGTGA	TTCAGTCATC	CTGAATGTCC
701	TCTATGGCCC	GGATGCCCCC	ACCATTTCCT	CTCTAAACAC	ATCTTACAGA
751	TCAGGGGAAA	ATCTGAACCT	CTCCTGCCAC	GCAGCCTCTA	ACCCACCTGC
801	ACAGTACTCT	TGGTTTGTCA	ATGGGACTTT	CCAGCAATCC	ACCCAAGAGC
851	TCTTTATCCC	CAACATCACT	GTGAATAATA	GTGGATCCTA	TACGTGCCAA
901	GCCCATAACT	CAGACACTGG	CCTCAATAGG	ACCACAGTCA	CGACGATCAC
951	AGTCTATGCA	GAGCCACCCA	AACCCTTCAT	CACCAGCAAC	AACTCCAACC
1001	CCGTGGAGGA	TGAGGATGCT	GTAGCCTTAA	CCTGTGAACC	TGAGATTGAG
1101	AACACAACCT	ACCTGTGGTG	GGTAAATAAT	CAGAGCCTCC	CGGTGAGTCC
1151	CAGGCTGCAG	CTGTCCAATG	ACAACAGGAC	CCTCACTCTA	CTCAGTGTCA
1201	CAAGGAATGA	TGTAGGACCC	TATGAGTGTG	GAATCCAGAA	CGAATTAAGT
1251	GTTGACCACA	GCGACCCAGT	CATCCTGAAT	GTCCTCTATG	GCCCAGACGA
1301	CCCCACCAT	TCCCCCTCAT	ACACCTATTA	CCGTCCAGGG	GTGAAGCTCA
1351	GCCTCTCCTG	CCATGCAGCC	TCTAACCCAC	CTGCACAGTA	TTCTTGGCTG
1401	ATTGATGGGA	ACATCCAGCA	ACACACACAA	GAGCTCTTTA	TCTCGAACAT
1451	CACTGAGAAG	AACAGCGGAC	TCTATACCTG	CCAGGCCAAT	AACTCAGCCA
1501	GTGGCCACAG	CAGGACTACA	GTCAAGACAA	TCACAGTCTC	TGCGGAGCTG
1551	CCCAAGCCCT	CCATCTCCAG	CAACAACCTC	AAACCCGTGG	AGGACAAGGA
1601	TGCTGTGGCC	TTCACCTGTG	AACCTGAGGC	TCAGAACACA	ACCTACCTGT
1651	GGTGGGTAAA	TGGTCAGAGC	CTCCCAGTCA	GTCCCAGGCT	GCAGCTGTCC
1701	AATGGCAACA	GGACCCTCAC	TCTATTCAAT	GTCACAAGAA	ATGACGCAAG
1751	AGCCTATGTA	TGTGGAATCC	AGAACTCAGT	GAGTGCAAAC	CGCAGTGACC
1801	CAGTCACCCT	GGATGTCCTC	TATGGGCCGG	ACACCCCAT	CATTTCCCCC
1851	CCAGACTCGT	CTTACCTTTC	GGGAGCGAAC	CTCAACCTCT	CCTGCCACTC
	GGCCTCTAAC	CCATCCCCGC	AGTATTCTTG	GCGTATCAAT	GGGATACCGC

FIG.3A-1

6/62

1901 AGCAACACAC ACAAGTTCTC TTTATCGCCA AAATCACGCC AAATAATAAC
1951 GGGACCTATG CCTGTTTTGT CTCTAACTTG GCTACTGGCC GCAATAATTC
2001 CATAGTCAAG AGCATCACAG TCTCTGCATC TGGAACTCTA GATGCTCCCC
2051 AGTCTATTAC AGAACTATGT TCGGAATATC GCAACACACA AATATATACG
2101 ATAAATGACA AGATACTATC ATATACGGAA TCGATGGCAG GTAAAAGAGA
2151 AATGGTTATC ATTACATTTA AGAGCGGCGC AACATTTTCAG GTCGAAGTCC
2201 CGGGCAGTCA ACATATAGAC TCCCAAAAAA AAGCCATTGA AAGGATGAAG
2251 GACACATTAA GAATCACATA TCTGACCGAG ACCAAAATTG ATAAATTATG
2301 TGTATGGAAT AATAAAACCC CCAATTCAAT TGC GGCAATC AGTATGGAAA
ACTAG (SEQ ID NO:9)

FIG.3A-2

7/62

CEA-LTB Amino Acid Sequence

1	MESPSAPPHR	WCIPWQRLLL	TASLLTFWNP	PTTAKLTIES	TPFNVAEGKE
51	VLLL VHNLPQ	HLFGYSWYKG	ERVDGNRQII	GYVIGTQQAT	PGPAYSGREI
101	IYPNASLLIQ	NIIQNDTGfY	TLHVIKSDLV	NEEATGQFRV	YPELPKPSIS
151	SNNSKPVEDK	DAVAFTCEPE	TQDATYLWWV	NNQSLPVSPR	LQLSNGNRTL
201	TLFNVTRNDT	ASYKCETQNP	VSARRSDSVI	LVLYGPDAP	TISPLNTSYR
251	SGENLNLSCH	AASNPPAQYS	WfVNGTFQQS	TQELFIPNIT	VNNSGSYTCQ
301	AHNSDTGLNR	TTVTTITVYA	EPPKPFITSN	NSNPVEDEDA	VALTCEPEIQ
351	NTTYLWWVNN	QSLPVSPRLQ	LSNDNRTLTL	LSVTRNDVGP	YECGIQNELS
401	VDHSDPVILN	VLYGPDDPTI	SPSYTYRPG	VNLSLSCHAA	SNPPAQYSWL
451	IDGNIQQHTQ	ELFISNITEK	NSGLYTCQAN	NSASGHSRTT	VKTITVSAEL
501	PKPSISSNNS	KPVEDKDAVA	FTCEPEAQNT	TYLWWVNGQS	LPVSPRLQLS
551	NGNRTLTLFN	VTRNDARAYV	CGIQNSVSAN	RSDPVTLDVL	YGPDTPIISP
601	PDSSYLSGAN	LNLSCHSASN	PSPQYSWRIN	GIPQQHTQVL	FIAKITPNNN
651	GTYACFVSNL	ATGRNNSIVK	SITVSASGTL	DAPQSITELC	SEYRNTQIYT
701	INDKILSYTE	SMAGKREIVI	ITFKSGATFQ	VEVPGSQHID	SQKKAIERMK
751	DTLRITYLTE	TKIDKLCVWN	NKTPNSIAAI	SMEN (SEQ ID NO:10)	

FIG.3B

8/62

CEAopt-LTB Nucleotide Sequence

1	ATGGAGAGCC	CCAGCGCCCC	CCCCCACC GC	TGGTGCATCC	CCTGGCAGCG
	CCTGCTGCTG	ACCGCCAGCC	TGCTGACCTT	CTGGAACCCC	CCCACCACCG
101	CCAAGCTGAC	CATCGAGAGC	ACCCCCTTCA	ACGTGGCCGA	GGGCAAGGAG
	GTGCTGCTGC	TGGTGCACAA	CCTGCCCCAG	CACCTGTTCG	GCTACAGCTG
201	GTACAAGGGC	GAGCGCGTGG	ACGGCAACCG	CCAGATCATC	GGCTACGTGA
	TCGGCACCCA	GCAGGCCACC	CCCGGCCCCG	CCTACAGCGG	CCGCGAGATC
301	ATCTACCCCA	ACGCCAGCCT	GCTGATCCAG	AACATCATCC	AGAACGACAC
	CGGCTTCTAC	ACCCTGCACG	TGATCAAGAG	CGACCTGGTG	AACGAGGAGG
401	CCACCGGCCA	GTTCCGCGTG	TACCCCGAGC	TGCCCAAGCC	CAGCATCAGC
	AGCAACAACA	GCAAGCCCGT	GGAGGACAAG	GACGCCGTGG	CCTTCACCTG
501	CGAGCCCGAG	ACCCAGGACG	CCACCTACCT	GTGGTGGGTG	AACAACCAGA
	GCCTGCCCCG	GAGCCCCCGC	CTGCAGCTGA	GCAACGGCAA	CCGCACCCTG
601	ACCCTGTTCA	ACGTGACCCG	CAACGACACC	GCCAGCTACA	AGTGCGAGAC
	CCAGAACCCC	GTGAGCGCCC	GCCGCAGCGA	CAGCGTGATC	CTGAACGTGC
701	TGTACGGCCC	CGACGCCCCC	ACCATCAGCC	CCCTGAACAC	CAGCTACCGC
	AGCGGCGAGA	ACCTGAACCT	GAGCTGCCAC	GCCGCCAGCA	ACCCCCCGC
801	CCAGTACAGC	TGGTTCGTGA	ACGGCACCTT	CCAGCAGAGC	ACCCAGGAGC
	TGTTTCATCCC	CAACATCACC	GTGAACAACA	GCGGCAGCTA	CACCTGCCAG
901	GCCCACAACA	GCGACACCGG	CCTGAACCGC	ACCACCGTGA	CCACCATCAC
	CGTGTACGCC	GAGCCCCCCA	AGCCCTTCAT	CACCAGCAAC	AACAGCAACC
1001	CCGTGGAGGA	CGAGGACGCC	GTGGCCCTGA	CCTGCGAGCC	CGAGATCCAG
	AACACCACCT	ACCTGTGGTG	GGTGAACAAC	CAGAGCCTGC	CCGTGAGCCC
1101	CCGCCTGCAG	CTGAGCAACG	ACAACCGCAC	CCTGACCCTG	CTGAGCGTGA
	CCCGCAACGA	CGTGGGCCCC	TACGAGTGCG	GCATCCAGAA	CGAGCTGAGC
1201	GTGGACCACA	GCGACCCCGT	GATCCTGAAC	GTGCTGTACG	GCCCCGACGA
	CCCCACCATC	AGCCCCAGCT	ACACCTACTA	CCGCCCCGGC	GTGAACCTGA
1301	GCCTGAGCTG	CCACGCCGCC	AGCAACCCCC	CCGCCAGTA	CAGCTGGCTG
	ATCGACGGCA	ACATCCAGCA	GCACACCCAG	GAGCTGTTCA	TCAGCAACAT
1401	CACCGAGAAG	AACAGCGGCC	TGTACACCTG	CCAGGCCAAC	AACAGCGCCA
	GCGGCCACAG	CCGCACCACC	GTGAAGACCA	TCACCGTGAG	CGCCGAGCTG
1501	CCCAAGCCCA	GCATCAGCAG	CAACAACAGC	AAGCCCGTGG	AGGACAAGGA
	CGCCGTGGCC	TTCACCTGCG	AGCCCGAGGC	CCAGAACACC	ACCTACCTGT
1601	GGTGGGTGAA	CGGCCAGAGC	CTGCCCGTGA	GCCCCCGCCT	GCAGCTGAGC
	AACGGCAACC	GCACCCTGAC	CCTGTTCAAC	GTGACCCGCA	ACGACGCCCG
1701	CGCCTACGTG	TGCGGCATCC	AGAACAGCGT	GAGCGCCAAC	CGCAGCGACC
	CCGTGACCCT	GGACGTGCTG	TACGGCCCCG	ACACCCCAT	CATCAGCCCC
1801	CCCGACAGCA	GCTACCTGAG	CGGCGCCAAC	CTGAACCTGA	GCTGCCACAG
	CGCCAGCAAC	CCCAGCCCCC	AGTACAGCTG	GCGCATCAAC	GGCATCCCCC

FIG.4A

9/62

1901 AGCAGCACAC CCAGGTGCTG TTCATCGCCA AGATCACCCC CAACAACAAC
GGCACCTACG CCTGCTTCGT GAGCAACCTG GCCACCGGCC GCAACAACAG
2001 CATCGTGAAG AGCATCACCG TGAGCGCCAG CGGCACCTCT AGAGCTCCCC
AGACTATTAC AGAACTATGT TCGGAATATC GCAACACACA AATATATACG
2101 ATAAATGACA AGATACTATC ATATACGGAA TCGATGGCAG GCAAAAGAGA
AATGGTTATC ATTACATTTA AGAGCGGCCA AACATTTTCAG GTCGAAGTCC
2201 CGGGCAGTCA ACATATAGAC TCCCAGAAAA AAGCCATTGA AAGGATGAAG
GACACATTAA GAATCACATA TCTGACCGAG ACCAAAATTG ATAAATTATG
2301 TGTATGGAAT AATAAAACCC CCAATTCAAT TGCGGCAATC AGTATGGAAA
ACTAG (SEQ ID NO:11)

FIG.4B

10/62

hCEA-LTBopt Coding Sequence

1	ATGGAGAGCC	CCAGCGCCCC	CCCCCACC GC	TGGTGCATCC	CCTGGCAGCG
	CCTGCTGCTG	ACCGCCAGCC	TGCTGACCTT	CTGGAACCCC	CCCACCACCG
101	CCAAGCTGAC	CATCGAGAGC	ACCCCCTTCA	ACGTGGCCGA	GGGCAAGGAG
	GTGCTGCTGC	TGGTGCACAA	CCTGCCCCAG	CACCTGTTCG	GCTACAGCTG
201	GTACAAGGGC	GAGCGCGTGG	ACGGCAACCG	CCAGATCATC	GGCTACGTGA
	TCGGCACCCA	GCAGGCCACC	CCCGGCCCCG	CCTACAGCGG	CCGCGAGATC
301	ATCTACCCCA	ACGCCAGCCT	GCTGATCCAG	AACATCATCC	AGAACGACAC
	CGGCTTCTAC	ACCCTGCACG	TGATCAAGAG	CGACCTGGTG	AACGAGGAGG
401	CCACCGGCCA	GTTCCGCGTG	TACCCCGAGC	TGCCCAAGCC	CAGCATCAGC
	AGCAACAACA	GCAAGCCCGT	GGAGGACAAG	GACGCCGTGG	CCTTCACCTG
501	CGAGCCCGAG	ACCCAGGACG	CCACCTACCT	GTGGTGGGTG	AACAACCAGA
	GCCTGCCCGT	GAGCCCCCGC	CTGCAGCTGA	GCAACGGCAA	CCGCACCCTG
601	ACCCTGTTCA	ACGTGACCCG	CAACGACACC	GCCAGCTACA	AGTGCGAGAC
	CCAGAACCCC	GTGAGCGCCC	GCCGCAGCGA	CAGCGTGATC	CTGAACGTGC
701	TGTACGGCCC	CGACGCCCCC	ACCATCAGCC	CCCTGAACAC	CAGCTACCGC
	AGCGGCGAGA	ACCTGAACCT	GAGCTGCCAC	GCCGCCAGCA	ACCCCCCCGC
801	CCAGTACAGC	TGGTTCGTGA	ACGGCACCTT	CCAGCAGAGC	ACCCAGGAGC
	TGTTTCATCCC	CAACATCACC	GTGAACAACA	GCGGCAGCTA	CACCTGCCAG
901	GCCCACAACA	GCGACACCGG	CCTGAACCGC	ACCACCGTGA	CCACCATCAC
	CGTGTACGCC	GAGCCCCCCA	AGCCCTTCAT	CACCAGCAAC	AACAGCAACC
1001	CCGTGGAGGA	CGAGGACGCC	GTGGCCCTGA	CCTGCGAGCC	CGAGATCCAG
	AACACCACCT	ACCTGTGGTG	GGTGAACAAC	CAGAGCCTGC	CCGTGAGCCC
1101	CCGCCTGCAG	CTGAGCAACG	ACAACCGCAC	CCTGACCCTG	CTGAGCGTGA
	CCCGCAACGA	CGTGGGCCCC	TACGAGTGCG	GCATCCAGAA	CGAGCTGAGC
1201	GTGGACCACA	GCGACCCCGT	GATCCTGAAC	GTGCTGTACG	GCCCCGACGA
	CCCCACCATC	AGCCCCAGCT	ACACCTACTA	CCGCCCCGGC	GTGAACCTGA
1301	GCCTGAGCTG	CCACGCCGCC	AGCAACCCCC	CCGCCCAGTA	CAGCTGGCTG
	ATCGACGGCA	ACATCCAGCA	GCACACCCAG	GAGCTGTTCA	TCAGCAACAT
1401	CACCGAGAAG	AACAGCGGCC	TGTAGACCTG	CCAGGCCAAC	AACAGCGCCA
	GCGGCCACAG	CCGCACCACC	GTGAAGACCA	TCACCGTGAG	CGCCGAGCTG
1501	CCCAAGCCCA	GCATCAGCAG	CAACAACAGC	AAGCCCGTGG	AGGACAAGGA
	CGCCGTGGCC	TTCACCTGCG	AGCCCGAGGC	CCAGAACACC	ACCTACCTGT
1601	GGTGGGTGAA	CGGCCAGAGC	CTGCCCGTGA	GCCCCCGCCT	GCAGCTGAGC
	AACGGCAACC	GCACCCTGAC	CCTGTTCAAC	GTGACCCGCA	ACGACGCCCG
1701	CGCCTACGTG	TGCGGCATCC	AGAACAGCGT	GAGCGCCAAC	CGCAGCGACC
	CCGTGACCCT	GGACGTGCTG	TACGGCCCCG	ACACCCCAT	CATCAGCCCC
1801	CCCGACAGCA	GCTACCTGAG	CGGCGCCAAC	CTGAACCTGA	GCTGCCACAG
	CGCCAGCAAC	CCCAGCCCCC	AGTACAGCTG	GCGCATCAAC	GGCATCCCCC

FIG.5A-1

11/62

1901 AGCAGCACAC CCAGGTGCTG TTCATCGCCA AGATCACCCC CAACAACAAC
GGCACCTACG CCTGCTTCGT GAGCAACCTG GCCACCGGCC GCAACAACAG
2001 CATCGTGAAG AGCATCACCG TGAGCGCCAG CGGCACCTCT AGAGCCCCC
AGAGCATCAC CGAGCTGTGC AGCGAGTACC GGAACACCCA GATCTACACC
2101 ATCAACGACA AGATCCTGAG CTACACCGAG AGCATGGCCG GCAAGAGGGA
GATGGTGATC ATCACCTTCA AGAGCGGCGC CACCTTCCAG GTGGAGGTGC
2201 CCGGCAGCCA GCACATCGAC AGCCAGAAGA AGGCCATCGA GCGGATGAAG
GACACCCTGC GGATCACCTA CCTCACCGAG ACCAAGATCG ACAAGCTGTG
2301 CGTGTGGAAC AACAAGACCC CCAACAGCAT CGCCGCCATC AGCATGGAGA
ATTGATAA (SEQ ID NO:12)

FIG.5A-2

12/62

hCEA-LTB Amino Acid Sequence

1	MESPSAPPHR	WCIPWQRLLL	TASLLTFWNP	PTTAKLTIES	TPFNVAEGKE
51	VLLLVHNLPO	HLFGYSWYKG	ERVDGNRQII	GYVIGTQQAT	PGPAYSGREI
101	IYPNASLLIQ	NIIQNDTGFY	TLHVIKSDLV	NEEATGQFRV	YPELPKPSIS
151	SNNSKPVEDK	DAVAFTCEPE	TQDATYLLWW	NNQSLPVSPR	LQLSNGNRTL
201	TLFNVTRNDT	ASYKCETQNP	VSARRSDSVI	LNVLVGPDAF	TISPLNTSYR
251	SGENLNLSCH	AASNPPAQYS	WVFNVTGTFQS	TQELFIPNI	VNNSGSYTCQ
301	AHNSDTGLNR	TTVTITITVYA	EPPKPFITSN	NSNPVEDEDA	VALTCEPEIQ
351	NTTYLWWVNN	QSLPVSPRLQ	LSNDNRTLTL	LSVTRNDVGP	YECGIQNELS
401	VDHSDPVILN	VLYGPDDPTI	SPSYTYRPG	VNLSLSCHAA	SNPPAQYSWL
451	IDGNIQQHTQ	ELFISNITEK	NSGLYTCQAN	NSASGHSRTT	VKTITVSAEL
501	PKPSISSNNS	KPVEDKDAVA	FTCEPEAQNT	TYLWWVNGQS	LPVSPRLQLS
551	NGNRTLTLFN	VTRNDARAYV	CGIQNSVSAN	RSDPVTLDVL	YGPDTPIISP
601	PDSSYLSGAN	LNLSCHSASN	PSPQYSWRIN	GIPQQHTQVL	FIAKITPNNN
651	GTYACFVSNL	ATGRNNSIVK	SITVSASGTS	RAPQSITELC	SEYRNTQIYT
701	INDKILSYTE	SMAGKREIVI	ITFKSGATFQ	VEVPGSQHID	SQKKAIERMK
751	DTLRITYLTE	TKIDKLCVWN	NKTPNSIAAI	SMEN (SEQ ID NO:13)	

FIG.5B

Rhesus CEAopt-LTBopt Coding Sequence

ATGGGCAGCC CCAGCGCCCC CCTGCACCGC TGGTGCAATCC CCTGGCAGAC CCTGCTGCTG ACGGCCAGCC
TGCTGACCTT CTGGAACCCC CCCACCACCG CCCAGCTGAC CATCGAGAGC CGCCCCCTTCA ACGTGGCCGA
GGCAAGGAG GTGCTGCTGC TGGCCACAA CGTGAGCCAG AACCTGTTG GCTACATCTG GTACAAGGGC
GAGCGGTGG ACGCCAGCCG CCGCATCGG AGCTGCGTGA TCCGCACCCA GCAGATCACC CCCGGCCCCG
CCCACAGCGG CCGCGAGACC ATCGACTTCA ACGCCAGCCT GCTGATCCAC AACGTGACCC AGAGCGACAC
CGCAGCTAC ACCATCCAGG TGATCAAGGA GGACCTGGTG AACGAGGAGG CCACCGGCCA GTTCCGCGTG
TACCCCGAGC TGCCCAAGCC CTACATCAGC AGCAACAACA GCAACCCCGT GGAGGACAAG GACGCCGTGG
CCCTGACCTG CGAGCCCGAG ACCCAGGACA CCACCTACCT GTGGTGGGTG AACAACCAGA GCCTGCCCGT
GAGCCCCCGC CTGGAGCTGA GCAGCGACAA CCGCACCCCTG ACCGTGTTCA ACATCCCCCG CAACGACACC
ACCAGCTACA AGTGGGAGC CCAGAACCCC GTGAGCGTGC GCCGCAGCGA CCCCCTGACC CTGAACGTGC
TGTAACGGCC CGACGCCCCC ACCATCAGCC CCTGAACAC CCCCTACCGC GCCGGCGAGA ACCTGAACCT
GACCTGCCAC GCCGCCAGCA ACCCCACCGC CCAGTACTTC TGGTTCGTGA ACGGCACCTT CCAGCAGAGC
ACCCAGGAGC TGTTCTCCC CAACATCACC GTGAACAACA GCGGCAGCTA CATGTGCCAG GCCCACAAACA
GGCCACCCGG CCTGAACCGC ACCACCGTGA CCGCCATCAC CGTGACGCC GAGCTGCCCA AGCCCTACAT
CACCAGCAAC AACAGCAACC CCATCGAGGA CAAGGACGCC GTGACCCTGA CCTGCGAGCC CGAGACCCAG
GACACCACT ACCTGTGGTG GGTGAACAAC CAGAGCCTGA GCGTGAGCAG CCGCCTGGAG CTGAGCAACG
ACAACCGCAC CCGACCGGTG TTCAACATCC CCGCAACGA CACCACCTTC TACGAGTGCG AGACCCAGAA
CCCCGTGAGC GTGGCGCGCA GCGACCCCGT GACCTGAAC GTGCTGTACG GCCCGGACGC CCCACCATC
AGCCCCCTGA ACAACCCCTA CCGCGCCGGC GAGAACCTGA ACCTGAGCTG CCACGCCGCC AGCAACCCCG
CCGCCQAGTA CAGCTGGTTC GTGAACGGCA CCTTCCAGCA GAGCACCCAG GAGCTGTTCA TCCCCAACAT
CACCGTGAAC AACAGCGGCA GCTACATGTG CCAGGCCAC AACAGCGCCA CCGCCTGAA CCGCACCACC
GTGACCGCCA TCAACCGTGA CGTGGAGCTG CCCAAGCCCT ACATCAGCAG CAACAACAGC ACCCCATCG

FIG. 6A-1

AGGACAAGGA CGCCGTGACC CTGACCTGCG AGCCCGTGCG CGAGAACACC ACCTACCTGT GGTGGGTGAA
CAACCAGAGC CTGAGCGTGA GCCCCCGCCT GCAGCTGAGC AACGGCAACC GCATCCTGAC CCTGCTGAGC
GTGACCCGCA ACGACACCGG CCCTTACGAG TCGGCGATCC AGAACAGCGA GAGGCCAAG CGCAGCGACC
CCGTGACCCT GAACGTGACC TACGGCCCCG ACACCCCAT CATCAGCCCC CCCGACCTGA GCTACCGCAG
CGGCGCCAAC CTGAACCTGA GCTGCCACAG CGACAGCAAC CCCAGCCCCC AGTACAGCTG GCTGATCAAC
GGCACCCCTG GCCAGCACAC CCAGGTGCTG TTCATCAGCA AGATCACCCAG CAACAACAGC GGGCCCTACG
CCTGCTTCGT GAGCAACCTG GCCACCGGCC GCAACAACAG CATCGTGAAG AACATCAGCG TGAGCAGCGG
CGACAGCTCT AGAGCCCCC AGAGCATCAC CGAGCTGTGC AGCGAGTACC GGAACACCCA GATCTACACC
ATCAACGACA AGATCCTGAG CTACACCGAG AGCATGGCCG GCAAGAGGGA GATGGTGATC ATCACCTTCA
AGAGCGGCGC CACCTTCCAG GTGGAGGTGC CCGGCAGCCA GCACATCGAC AGCCAGAAGA AGGCCATCGA
GCGGATGAAG GACACCCTGC GGATCACCTA CCTCACCAG ACCAAGATCG ACAAGCTGTG CGTGTGGAAC
AACAAGACCC CCAACAGCAT CGCCGCCATC AGCATGGAGA ATTGATAA (SEQ ID NO:14)

FIG. 6A-2

15/62

RhCEAopt-LTBopt Amino Acid Sequence

```
1   MGSPSAPLHR WCIPWQTLLE TASLLTFWNP PTTAQLTIES RPFNVAEGKE
51  VLLLAHNVSQ NLFGYIWYKG ERVDASRRIG SCVIRTQQIT PGPAHSGRET
101 IDFNASLLIH NVTQSDTGSY TIQVIKEDLV NEEATGQFRV YPELKPYPIS
151 SNNSNPVEDK DAVALTCEPE TQDTTYLWWV NNQSLPVSPR LELSSDNRTL
201 TVFNIPRNDT TSYKCETQNP VSVRRSDPVT LNVLYGPDAP TISPLNTPYR
251 AGENLNLTCH AASNPTAQYF WFNVTGTFQQS TQELFIPNIT VNNSGSYMCO
301 AHNSATGLNR TTVTAITVYA ELPKPYITSN NSNPIEDKDA VTLTCEPETQ
351 DTTYLWWVNN QSLSVSSRLE LSNDNRTLTV FNIPRNDTTF YECETQNPVS
401 VRRSDPVTLN VLYGPDAPTI SPLNTPYRAG ENLNLSCHAA SNPAAQYSWF
451 VNGTFQQSTQ ELFIPNITVN NSGSYMCOAH NSATGLNRTT VTAITVYVEL
501 PKPYISSNNS NPIEDKDAVT LTCEPVAENT TYLWWVNNQS LSVSPRLQLS
551 NGNRILTLLS VTRNDTGPEY CGIQNSSESAK RSDPVTLNVT YGPDTPIIIS
601 PDLSYRSGAN LNLSCHSDSN PSPQYSWLIN GTLRQHTQVL FISKITSNNS
651 GAYACFVSNL ATGRNNSIVK NISVSSGDSS RAPQSITELC SEYRNTQIYT
701 INDKILSYTE SMAGKREMVI ITFKSGATFQ VEVPGSQHID SQKKAIERMK
751 DTLRITYLTE TKIDKLCVWN NKTPNSIAAI SMEN (SEQ ID NO:15)
```

FIG.6B

16/62

Nucleotide Sequence of First Rhesus Monkey CEA

```

1   ATGGGGTCTC CCTCAGCCCC TCTTCACAGA TGGTGCATCC CCTGGCAGAC
51  GCTCCTGCTC ACAGCCTCAC TTCTAACCTT CTGGAACCCG CCCACCACTG
101 CCCAGCTCAC TATTGAATCC AGGCCGTTCA ATGTTGCAGA GGGGAAGGAG
151 GTTCTTCTAC TTGCCCACAA TGTGTCCCAG AATCTTTTTG GCTACATTTG
201 GTACAAGGGA GAAAGAGTGG ATGCCAGCCG TCGAATTGGA TCATGTGTAA
251 TAAGAACTCA ACAAATTACC CCAGGGCCCG CACACAGCGG TCGAGAGACA
301 ATAGACTTCA ATGCATCCCT GCTGATCCAC AATGTCACCC AGAGTGACAC
351 AGGATCCTAC ACCATACAAG TCATAAAGGA AGATCTTGTG AATGAAGAAG
401 CAACTGGCCA GTTCCGGGTA TACCCGGAGC TGCCCAAGCC CTACATCTCC
451 AGCAACAACCT CCAACCCCGT GGAGGACAAG GATGCTGTGG CCTTAACCTG
501 TGAACCTGAG ACTCAGGACA CAACCTACCT GTGGTGGGTA AACAAATCAGA
551 GCCTCCCGGT CAGTCCCAGG CTGGAGCTGT CCAGTGACAA CAGGACCCTC
601 ACTGTATTCA ATATTCCAAG AAATGACACA ACATCCTACA AATGTGAAAC
651 CCAGAACCCA GTGAGTGTCA GACGCAGCGA CCCAGTCACC CTGAACGTCC
701 TCTATGGCCC GGATGCGCCC ACCATTTCCC CTCTAAACAC ACCTTACAGA
751 GCAGGGGAAA ATCTGAACCT CACCTGCCAC GCAGCCTCTA ACCCAACTGC
801 ACAGTACTTT TGGTTTGTCA ATGGGACGTT CCAGCAATCC ACACAAGAGC
851 TCTTTATACC CAACATCACC GTGAATAATA GCGGATCCTA TATGTGCCAA
901 GCCCATAACT CAGCCACTGG CCTCAATAGG ACCACAGTCA CGGCGATCAC
951 AGTCTACGCG GAGCTGCCCA AGCCCTACAT CACCAGCAAC AACTCCAACC
1001 CCATAGAGGA CAAGGATGCT GTGACCTTAA CCTGTGAACC TGAGACTCAG
1051 GACACAACCT ACCTGTGGTG GGTAAACAAT CAGAGCCTCT CGGTGAGTTC
1101 CAGGCTGGAG CTGTCCAATG ACAACAGGAC CCTCACTGTA TTCAATATTC
1151 CAAGAAACGA CACAACGTTC TACGAATGTG AGACCCAGAA CCCAGTGAGT
1201 GTCAGACGCA GCGACCCAGT CACCCTGAAT GTCCTCTATG GCCCGGATGC
1251 GCCCACCATT TCCCCTCTAA ACACACCTTA CAGAGCAGGG GAAAATCTGA
1301 ACCTCTCCTG CCACGCAGCC TCTAACCCAG CTGCACAGTA CTCTTGTTT
1351 GTCAATGGGA CGTTCCAGCA ATCCACACAA GAGCTCTTTA TACCCAAACAT
1401 CACCGTGAAT AATAGCGGAT CCTATATGTG CCAAGCCCAT AACTCAGCCA
1451 CTGGCCTCAA TAGGACCACA GTCACGGCGA TCACAGTCTA TGTGGAGCTG
1501 CCCAAGCCCT ACATCTCCAG CAACAACCTC AACCCCATAG AGGACAAGGA
1551 TGCTGTGACC TTAACCTGTG AACCTGTGGC TGAGAACACA ACCTACCTGT
1601 GGTGGGTAAA CAATCAGAGC CTCTCGGTCA GTCCCAGGCT GCAGCTCTCC
1651 AATGGCAACA GGATCCTCAC TCTACTCAGT GTCACACGGA ATGACACAGG
1701 ACCCTATGAA TGTGGAATCC AGAACTCAGA GAGTGCAAAA CGCAGTGACC
1751 CAGTCACCCT GAATGTCACC TATGGCCCAG ACACCCCAT CATATCCCCC
1801 CCAGACTTGT CTTACCGTTC GGGAGCAAAC CTCAACCTCT CCTGCCACTC

```

FIG.7A-1

17/62

1851 GGACTCTAAC CCATCCCCGC AGTATTCTTG GCTTATCAAT GGGACACTGC
1901 GGCAACACAC ACAAGTTCTC TTTATCTCCA AAATCACATC AAACAATAGC
1951 GGGGCCTATG CCTGTTTTGT CTCTAACTTG GCTACCGGTC GCAATAACTC
2001 CATAGTCAAG AACATCTCAG TCTCCTCTGG CGATTCAGCA CCTGGAAGTT
2051 CTGGTCTCTC AGCTAGGGCT ACTGTCGGCA TCATAATTGG AATGCTGGTT
2101 GGGGTTGCTC TGATGTAG (SEQ ID NO:16)

FIG.7A-2

18/62

Nucleotide Sequence of Second Rhesus Monkey CEA

```
1   ATGGGGTCTC CCTCAGCCCC TCTTCACAGA TGGTGCATCC CCTGGCAGAC
51  GCTCCTGCTC ACAGCCTCAC TTCTAACCTT CTGGAACCCG CCCACEACTG
101 CCCAGCTCAC TATTGAATCC AGGCCGTTCA ATGTTGCAGA GGGGAAGGAG
151 GTTCTTCTAC TTGCCCACAA TGTGTCCCAG AATCTTTTTG GCTACATTTG
201 GTACAAGGGA GAAAGAGTGG ATGCCAGCCG TCGAATTGGA TCATGTGTAA
251 TAAGAACTCA ACAAATTACC CCAGGGCCCCG CACACAGCGG TCGAGAGACA
301 ATAGACTTCA ATGCATCCCT GCTGATCCAC AATGTCACCC AGAGTGACAC
351 AGGATCCTAC ACCATACAAG TCATAAAGGA AGATCTTGTG AATGAAGAAG
401 CAACTGGCCA GTTCCGGGTA TACCCGGAGC TGCCCAAGCC CTACATCTCC
451 AGCAACAAC TCCAACCCCGT GGAGGACAAG GATGCTGTGG CCTTAACCTG
501 TGAACCTGAG ACTCAGGACA CAACCTACCT GTGGTGGGTA AACAATCAGA
551 GCCTCCCGGT CAGTCCCAGG CTGGAGCTGT CCAGTGACAA CAGGACCCTC
601 ACTGTATTCA ATATTCCAAG AAATGACACA ACATCCTACA AATGTGAAAC
651 CCAGAACCCA GTGAGTGTCA GACGCAGCGA CCCAGTCACC CTGAACGTCC
701 TCTATGGCCC GGATGCGCCC ACCATTTCCC CTCTAAACAC ACCTTACAGA
751 GCAGGGGAAA ATCTGAACCT CACCTGCCAC GCAGCCTCTA ACCCAACTGC
801 ACAGTACTTT TGGTTTGTCA ATGGGACGTT CCAGCAATCC ACACAAGAGC
851 TCTTTATACC CAACATCACC GTGAATAATA GCGGATCCTA TATGTGCCAA
901 GCCCATAACT CAGCCACTGG CCTCAATAGG ACCACAGTCA CGGCGATCAC
951 AGTCTACGCG GAGCTGCCCC AGCCCTACAT CACCAGCAAC AACTCCAACC
1001 CCATAGAGGA CAAGGATGCT GTGACCTTAA CCTGTGAACC TGAGACTCAG
1051 GACACAACCT ACCTGTGGTG GGTAAACAAT CAGAGCCTCT CGGTCAGTTC
1101 CAGGCTGGAG CTGTCCAATG ACAACAGGAC CCTCACTGTA TTCAATATTC
1151 CAAGAAACGA CACAACGTTC TACGAATGTG AGACCCAGAA CCCAGTGAGT
1201 GTCAGACGCA GCGACCCAGT CACCCTGAAT GTCCTCTATG GCCCGGATGC
1251 GCCCACCATT TCCCCTCTAA ACACACCTTA CAGAGCAGGG GAAATCTGA
1301 ACCTCTCCTG CCACGCAGCC TCTAACCCAG CTGCACAGTA CTTTGTGTTT
1351 GTCAATGGGA CGTTCCAGCA ATCCACACAA GAGCTCTTTA TACCAACAT
1401 CACCGTGAAT AATAGCGGAT CCTATATGTG CCAAGCCCAT AACTCAGCCA
1451 CTGGCCTCAA TAGGACCACA GTCACGGCGA TCACAGTCTA TGTGGAGCTG
1501 CCCAAGCCCT ACATCTCCAG CAACAACCTC AACCCCATAG AGGACAAGGA
1551 TGCTGTGACC TTAACCTGTG AACCTGTGGC TGAGAACACA ACCTACCTGT
1601 GGTGGGTAAC CAATCAGAGC CTCTCGGTCA GTCCCAGGCT GCAGCTCTCC
1651 AATGGCAACA GGATCCTCAC TCTACTCAGT GTCACACGGA ATGACACAGG
1701 ACCCTATGAA TGTGGAATCC AGAACTCAGA GAGTGCAAAA CGCAGTGACC
1751 CAGTCACCCT GAATGTCACC TATGGCCCAG ACACCCCAT CATATCCCCC
1801 CCAGACTTGT CTTACCGTTC GGGAGCAAAC CTCAACCTCT CCTGCCACTC
```

FIG.7B-1

19/62

1851 GGACTCTAAC CCATCCCCGC AGTATTCTTG GCTTATCAAT GGGACACTGC
1901 GGCAACACAC ACAAGTTCTC TTTATCTCCA AAATCACATC AAACAATAAC
1951 GGGGCCTATG CCTGTTTTGT CTCTAACTTG GCTACCGGTC GCAATAACTC
2001 CATAGTCAAG AACATCTCAG TCTCCTCTGG CGATTCAGCA CCTGGAAGTT
2051 CTGGTCTCTC AGCTAGGGCT ACTGTCGGCA TCATAATTGG AATGCTGGTT
2101 GGGGTTGCTC TGATGTAG (SEQ ID NO:17)

FIG.7B-2

20/62

Amino Acid Sequence of First Rhesus Monkey CEA Protein

```

1   MGSPSAPLHR WCIPWQTLLL TASLLTFWNP PTTAQLTIES RPFNVAEGKE
51  VLLLAHNVSQ NLFGYIWYKG ERVDASRRIG SCVIRTQQIT PGPAHSGRET
101 IDFNASLLIH NVTQSDTGSY TIQVIKEDLV NEEATGQFRV YPELPPKPYIS
151 SNNSNPVEDK DAVALTCEPE TQDTTYLWWV NNQSLPVSPR LELSSDNRTL
201 TVFNIPRNDT TSYKCETQNP VSVRRSDPVT LNVLYGPDAP TISPLNTPYR
251 AGENLNLTCH AASNPTAQYF WFFVNGTFQQS TQELFIPNIT VNNSGSYMCO
301 AHNSATGLNR TTVTAITVYA ELPKPYITSN NSNPIEDKDA VTLTCEPETQ
351 DTTYLWWVNN QSLSVSSRLE LSNDNRTLTV FNIPRNDTTF YECETQNPVS
401 VRRSDPVTLN VLYGPDAPTI SPLNTPYRAG ENLNLSCHAA SNPAAQYSWF
451 VNGTFQQSTQ ELFIPNITVN NSGSYMCAH NSATGLNRTT VTAITVYVEL
501 PKPYISSNNS NPIEDKDAVT LTCEPVAENT TYLWWVNNQS LSVSPRLQLS
551 NGNRILTLLS VTRNDTGPYE CGIQNSAK RSDPVTNLVT YGPDTPIISP
601 PDLSYRSGAN LNLSCHSDSN PSPQYSWLIN GTLRQHTQVL FISKITSNNS
651 GAYACFVSNL ATGRNNSIVK NISVSSGDSA PGSSGLSARA TVGIIIGMLV
701 GVALM (SEQ ID NO:18)

```

FIG.7C

Amino Acid Sequence of Second Rhesus Monkey CEA Protein

```

1   MGSPSAPLHR WCIPWQTLLL TASLLTFWNP PTTAQLTIES RPFNVAEGKE
51  VLLLAHNVSQ NLFGYIWYKG ERVDASRRIG SCVIRTQQIT PGPAHSGRET
101 IDFNASLLIH NVTQSDTGSY TIQVIKEDLV NEEATGQFRV YPELPPKPYIS
151 SNNSNPVEDK DAVALTCEPE TQDTTYLWWV NNQSLPVSPR LELSSDNRTL
201 TVFNIPRNDT TSYKCETQNP VSVRRSDPVT LNVLYGPDAP TISPLNTPYR
251 AGENLNLTCH AASNPTAQYF WFFVNGTFQQS TQELFIPNIT VNNSGSYMCO
301 AHNSATGLNR TTVTAITVYA ELPKPYITSN NSNPIEDKDA VTLTCEPETQ
351 DTTYLWWVNN QSLSVSSRLE LSNDNRTLTV FNIPRNDTTF YECETQNPVS
401 VRRSDPVTLN VLYGPDAPTI SPLNTPYRAG ENLNLSCHAA SNPAAQYFWF
451 VNGTFQQSTQ ELFIPNITVN NSGSYMCAH NSATGLNRTT VTAITVYVEL
501 PKPYISSNNS NPIEDKDAVT LTCEPVAENT TYLWWVNNQS LSVSPRLQLS
551 NGNRILTLLS VTRNDTGPYE CGIQNSAK RSDPVTNLVT YGPDTPIISP
601 PDLSYRSGAN LNLSCHSDSN PSPQYSWLIN GTLRQHTQVL FISKITSNNN
651 GAYACFVSNL ATGRNNSIVK NISVSSGDSA PGSSGLSARA TVGIIIGMLV
701 GVALM (SEQ ID NO:19)

```

FIG.7D

21/62

Amino Acid Sequence of Human CEA Protein

1	MESPSAPPHR	WCIPWQRLLL	TASLLTFWNP	PTTAKLTIES	TPFNVAEGKE
51	VLLL VHNLPQ	HLFGYSWYKG	ERVDGNRQII	GYVIGTQQAT	PGPAYSGREI
101	IYPNASLLIQ	NIIQNDTGFY	TLHVIKSDLV	NEEATGQFRV	YPELPKPSIS
151	SNNSKPVEDK	DAVAFTCEPE	TQDATYLWWV	NNQSLPVSPR	LQLSNGNRTL
201	TLFNVTRNDT	ASYKCETQNP	VSARRSDSVI	LNVL YGPDAP	TISPLNTSYR
251	SGENLNLSCH	AASNPPAQYS	WFVNGTFQQS	TQELFIPNIT	VNNSGSYTCQ
301	AHNSDTGLNR	TTVTTITVYA	EPPKPFITSN	NSNPVEDEDA	VALTCEPEIQ
351	NTTYLWWVNN	QSLPVSPRLQ	LSNDNRTLTL	LSVTRNDVGP	YECGIQNELS
401	VDHSDPVILN	VLYGPDDPTI	SPSYTYRPG	VNLSLSCHAA	SNPPAQYSWL
451	IDGNIQQHTQ	ELFISNITEK	NSGLYTCQAN	NSASGHSRTT	VKTITVSAEL
501	PKPSISSNNS	KPVEDKDAVA	FTCEPEAQNT	TYLWWVNGQS	LPVSPRLQLS
551	NGNRTLTLFN	VTRNDARAYV	CGIQNSVSAN	RSDPVTLDVL	YGPDTPIISP
601	PDSSYLSGAN	LNLSCHSASN	PSPQYSWRIN	GIPQQHTQVL	FIAKITPNNN
651	GTYACFVSNL	ATGRNNSIVK	SITVSASGTS	PGLSAGATVG	IMIGULVGVA
701	LI (SEQ ID NO:20)				

FIG.7E

22/62

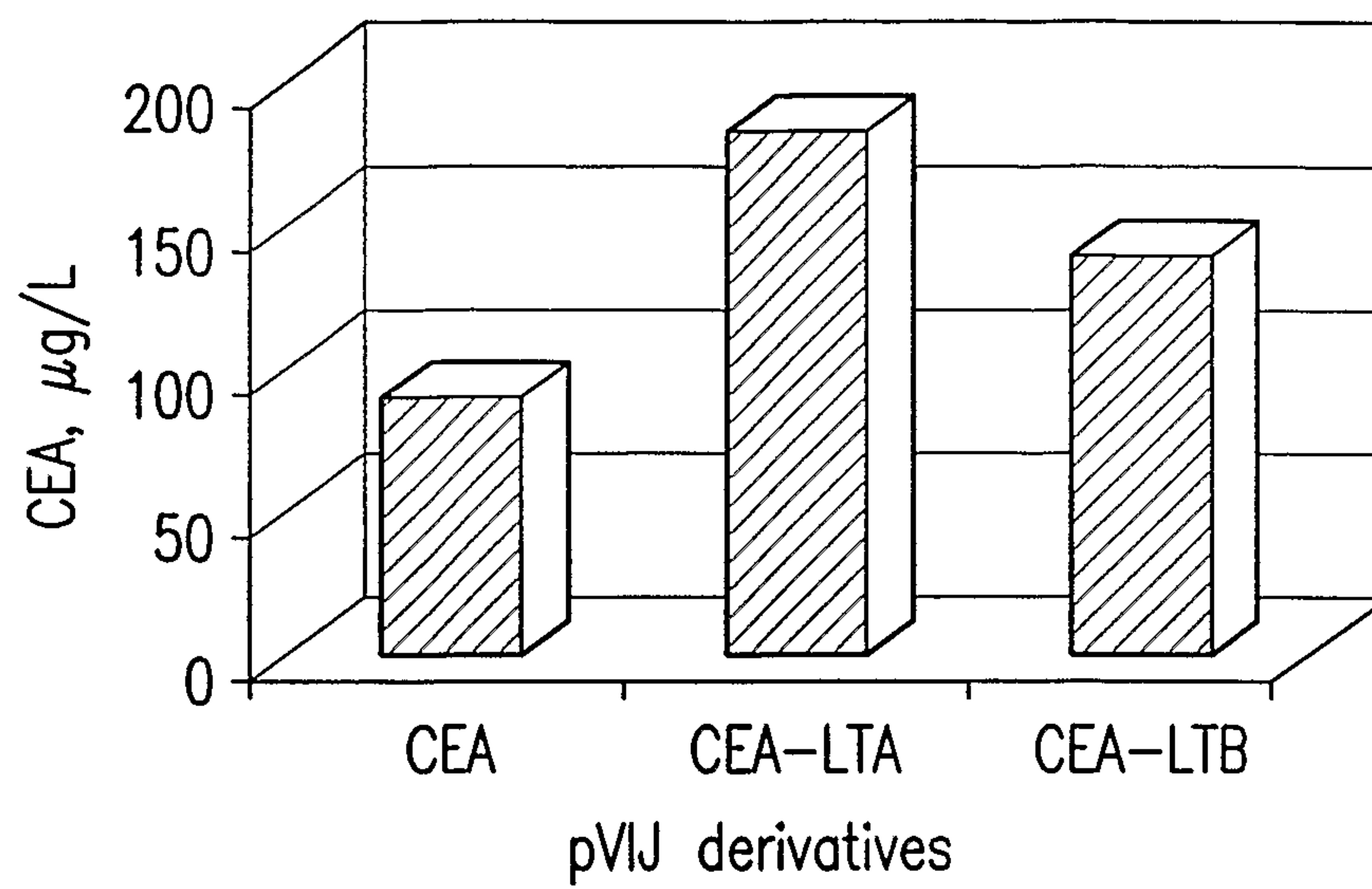


FIG.8A

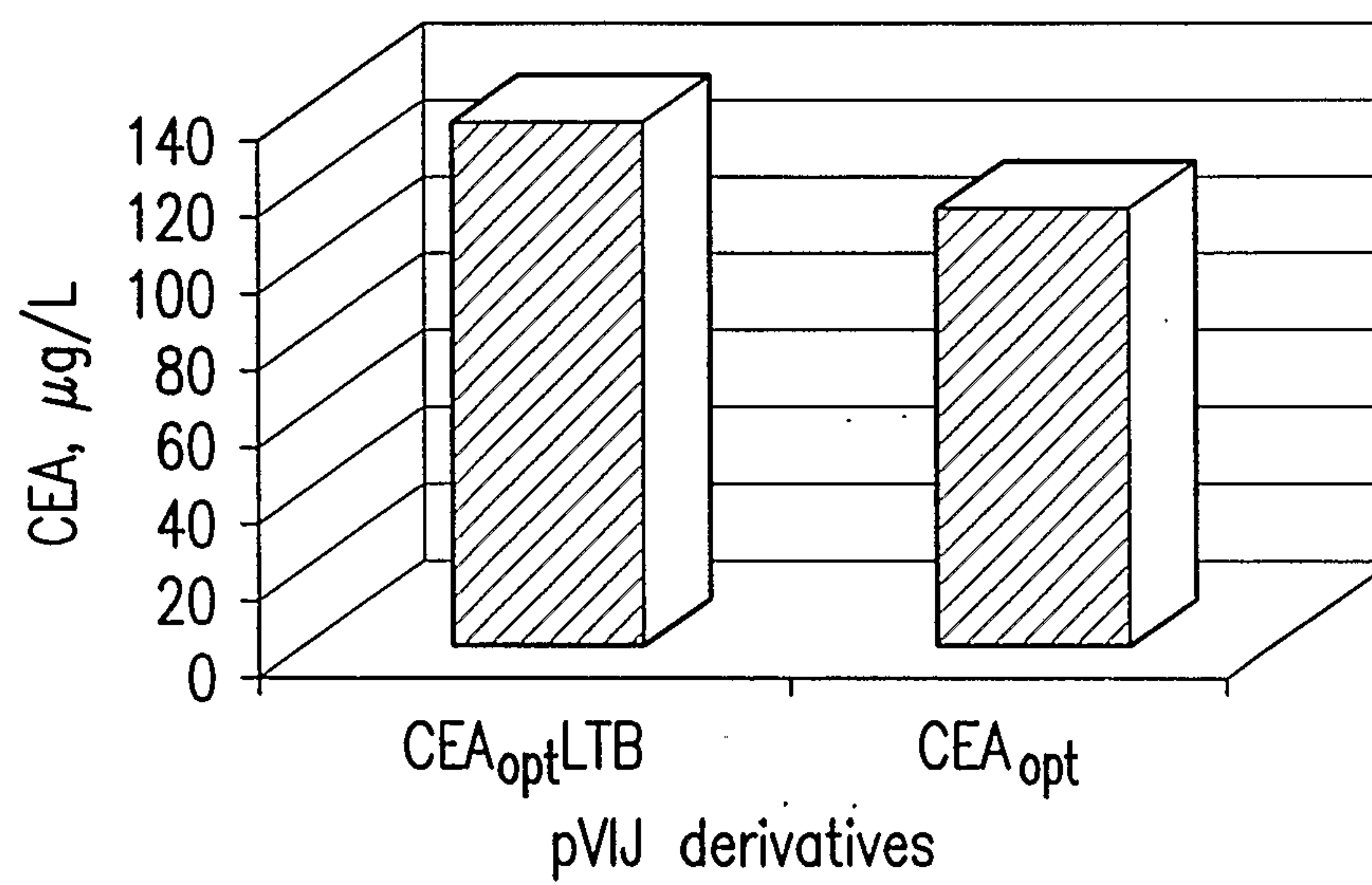
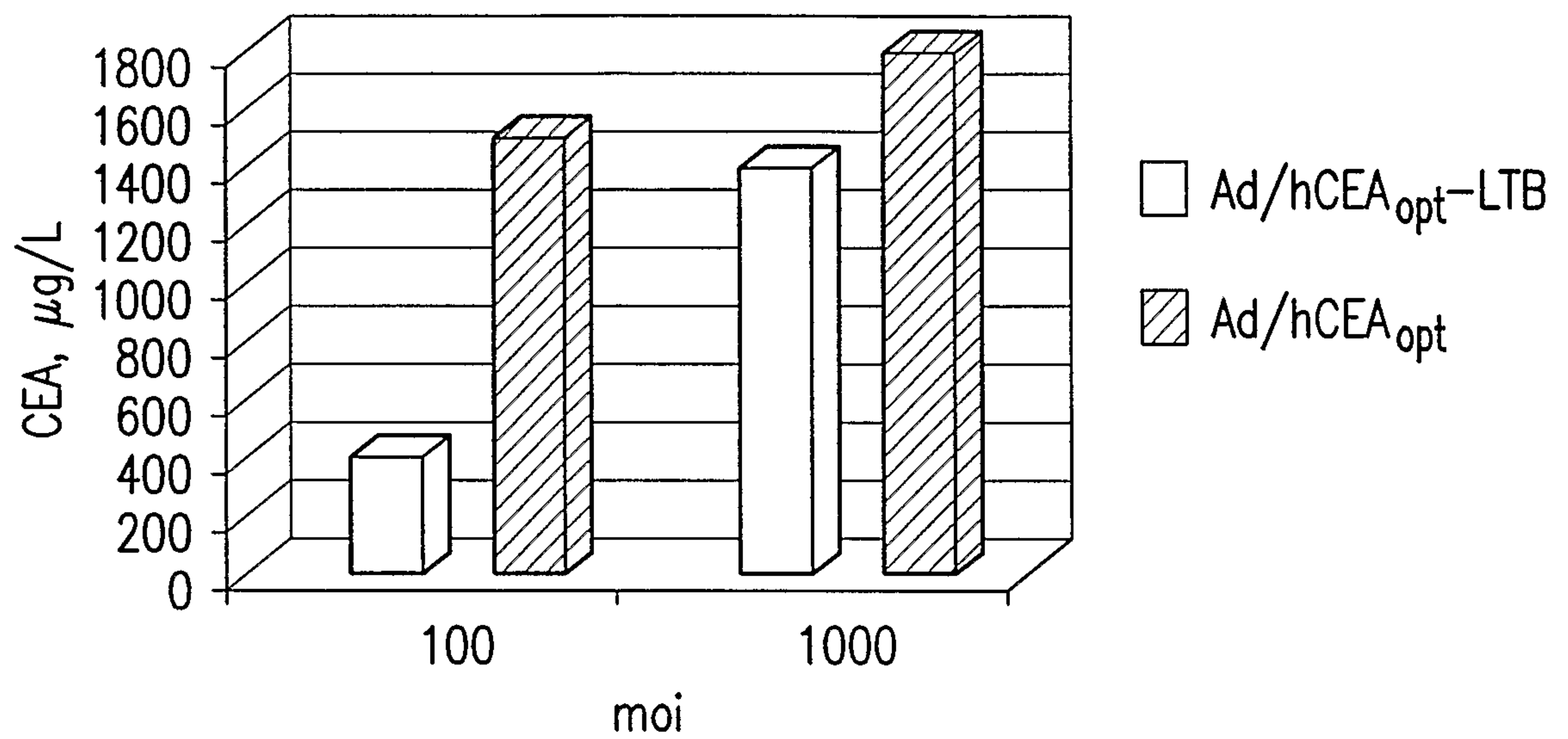


FIG.8B

23/62



Ad vectors
HeLa cells

FIG.9

24/62

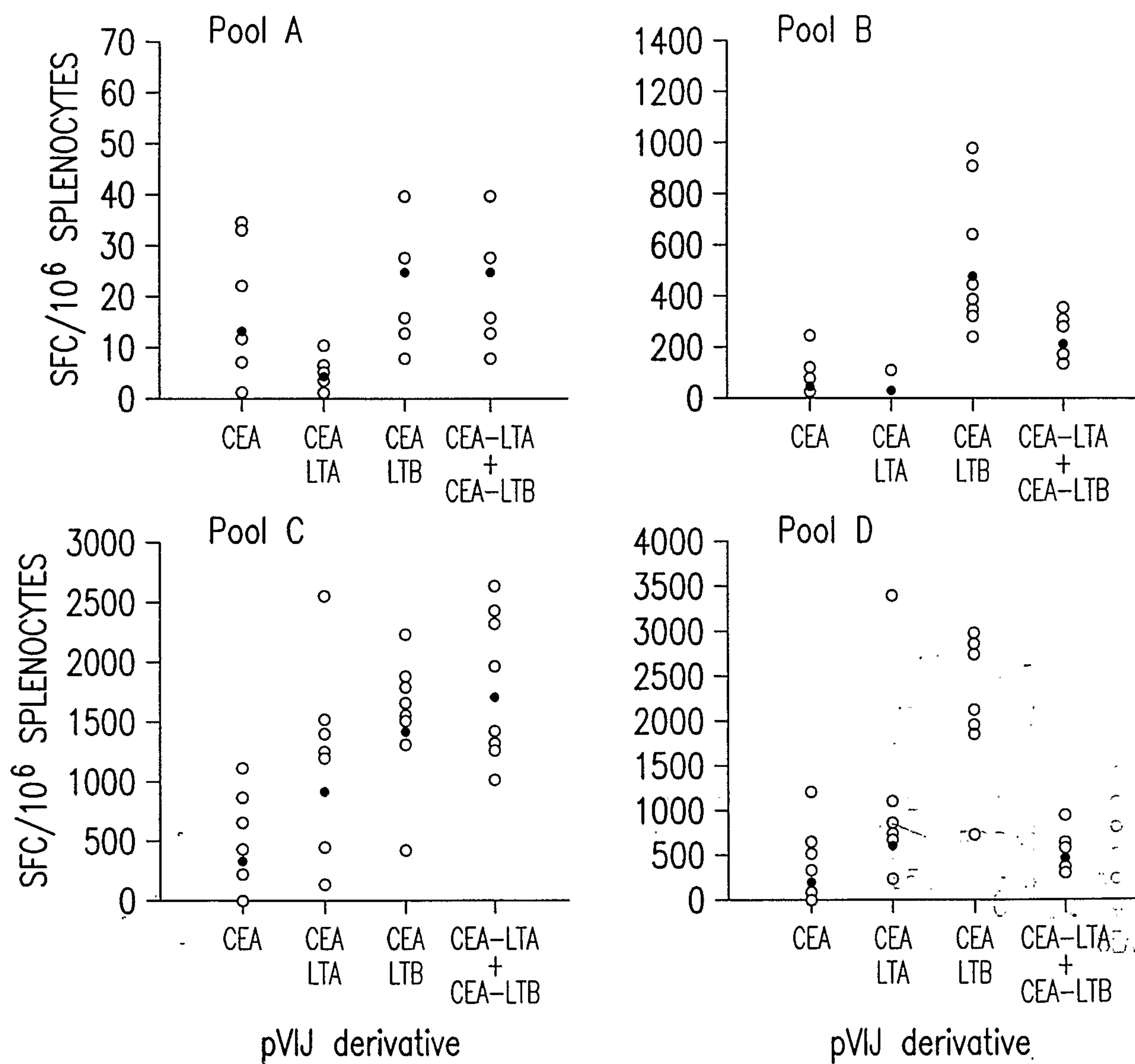


FIG.10A

25/62

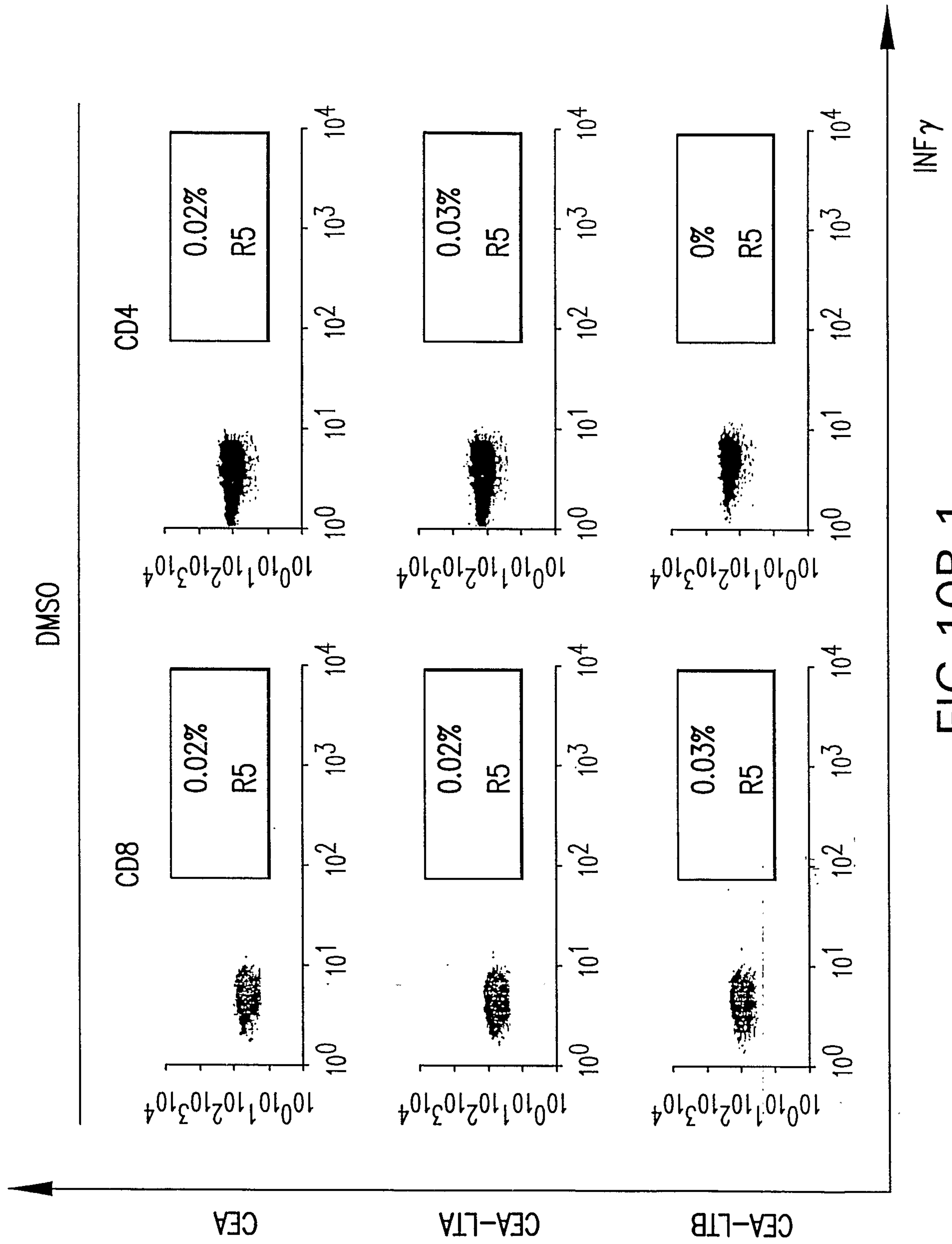


FIG. 10B-1

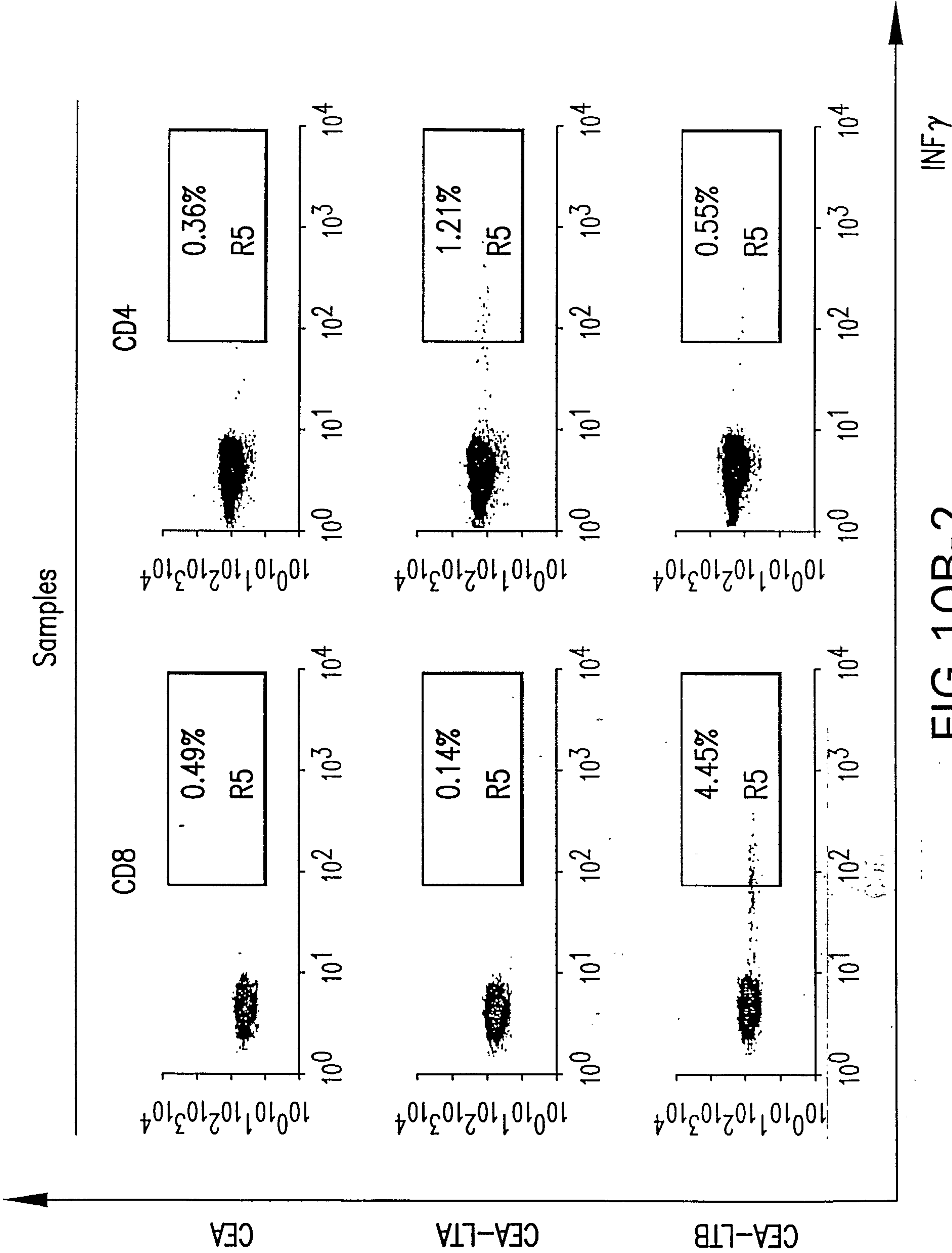


FIG.10B-2

27/62

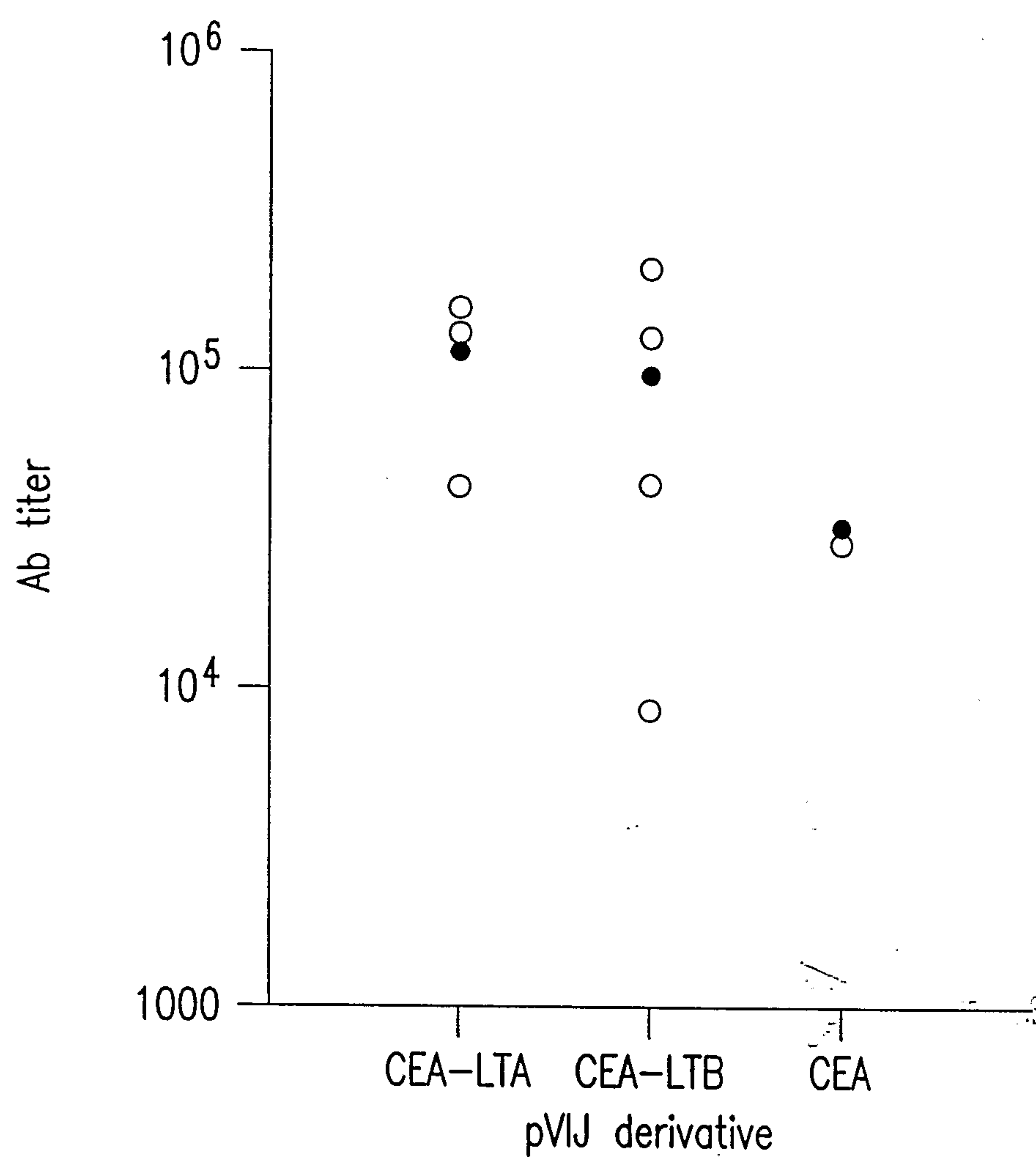


FIG. 11

28/62

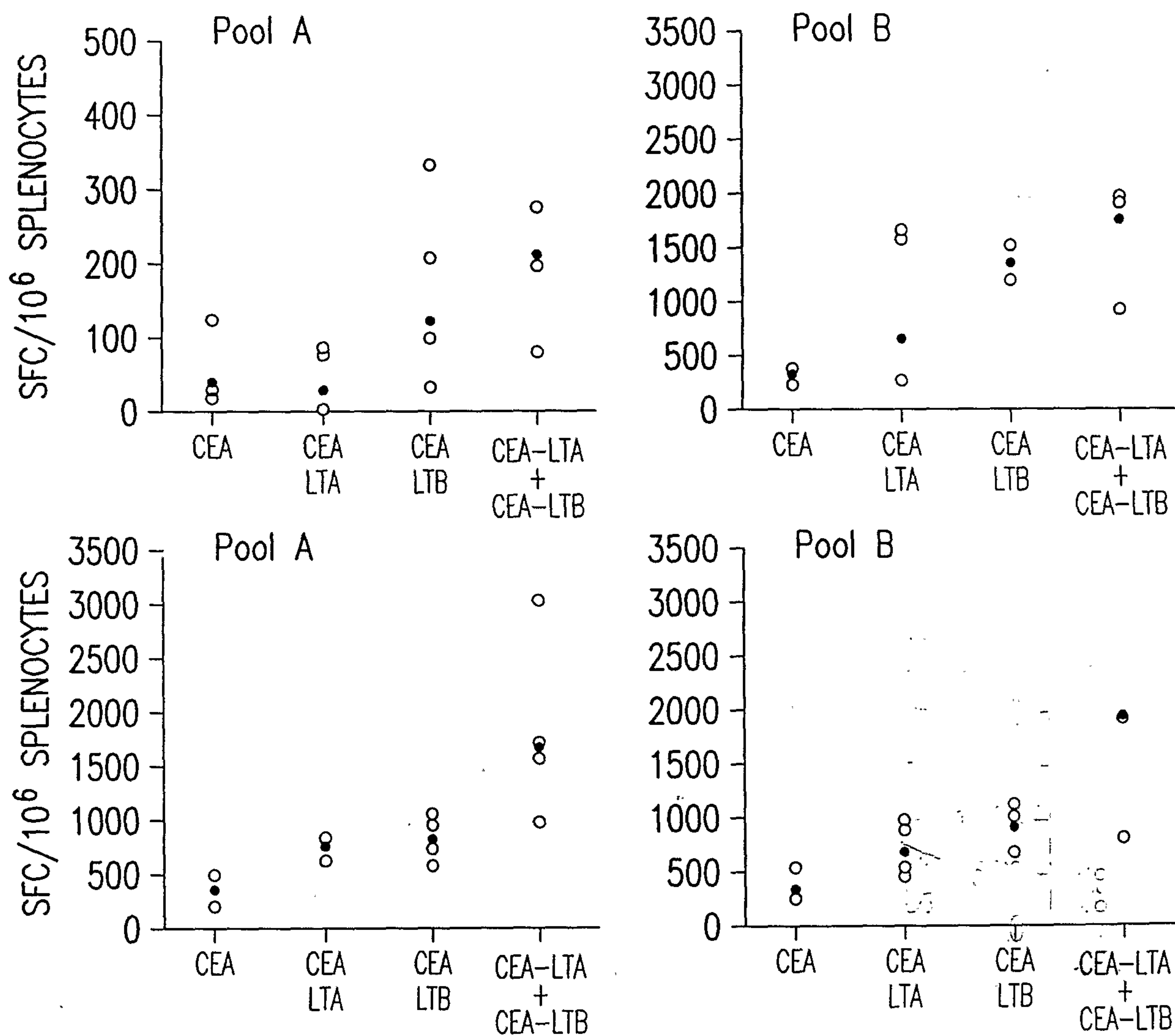


FIG.12

29/62

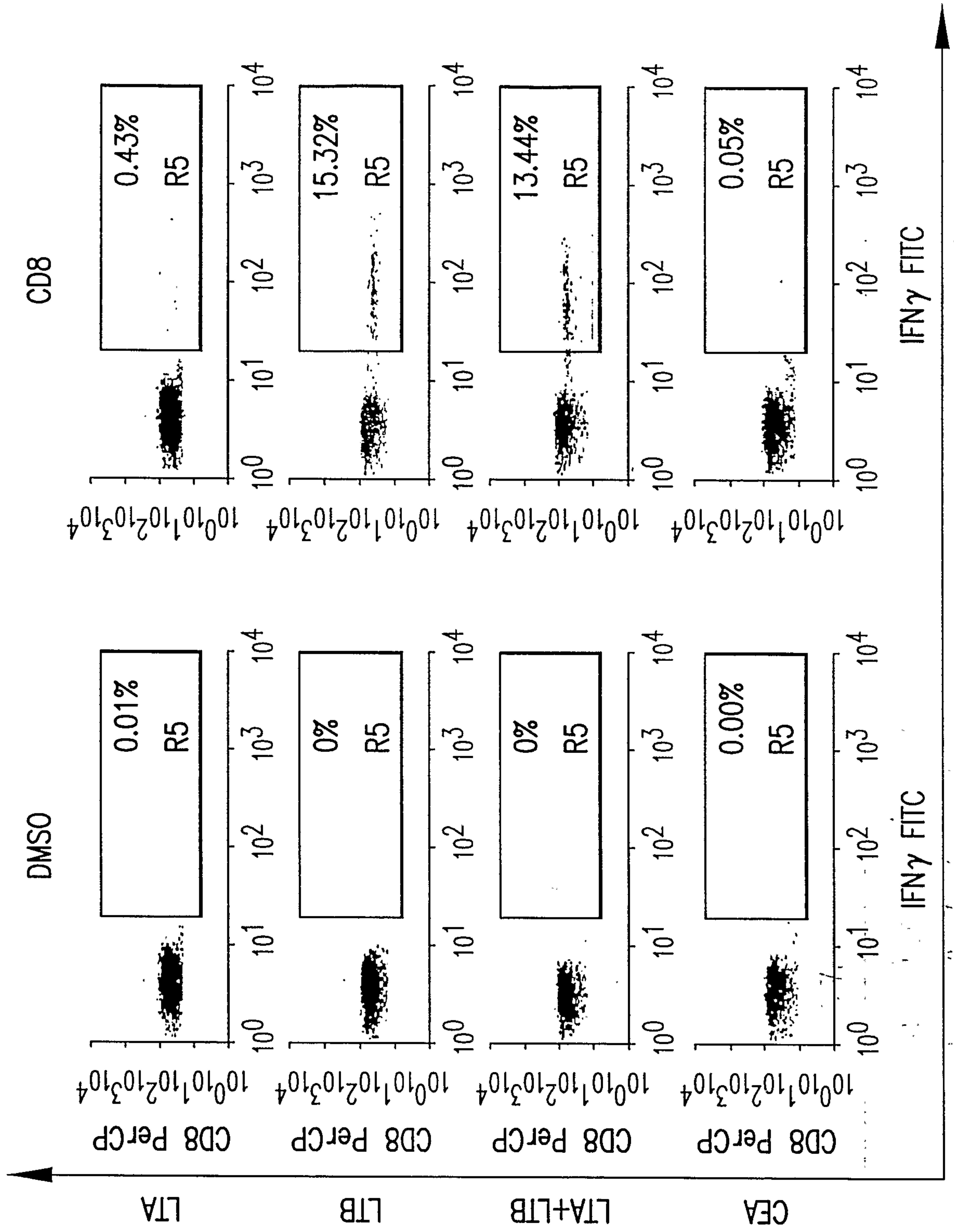


FIG. 13

30/62

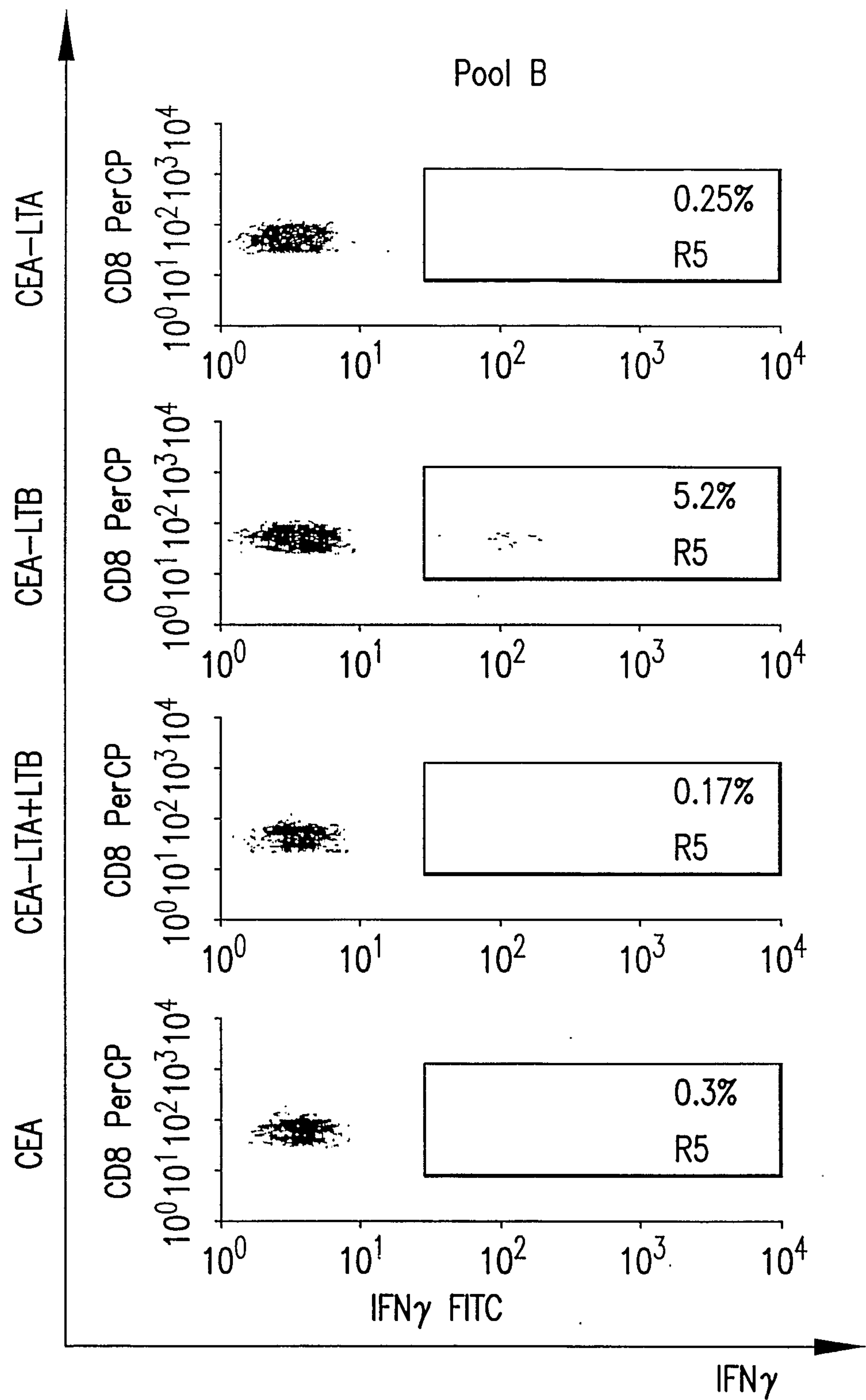


FIG. 14A

31/62

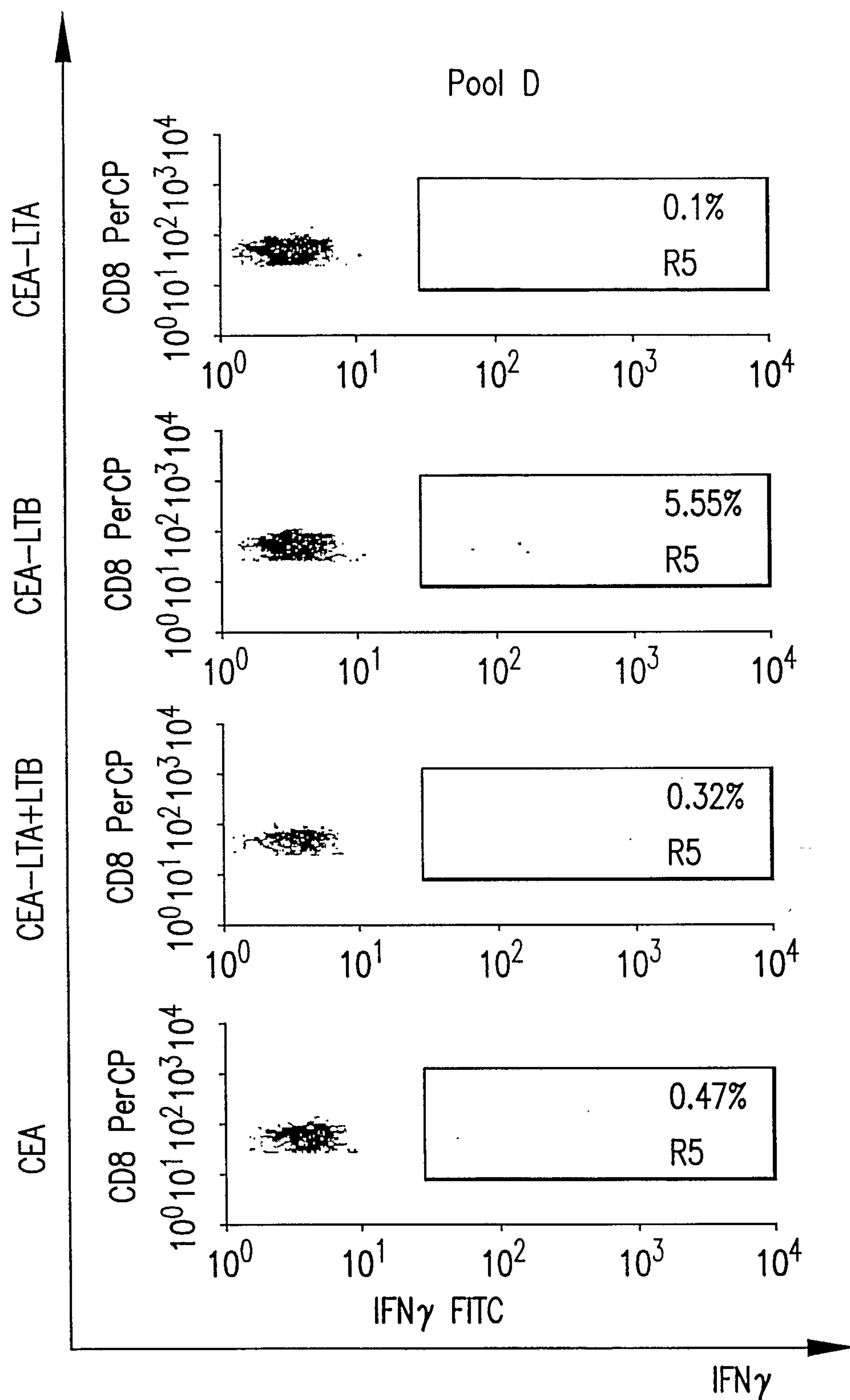


FIG. 14B

32/62

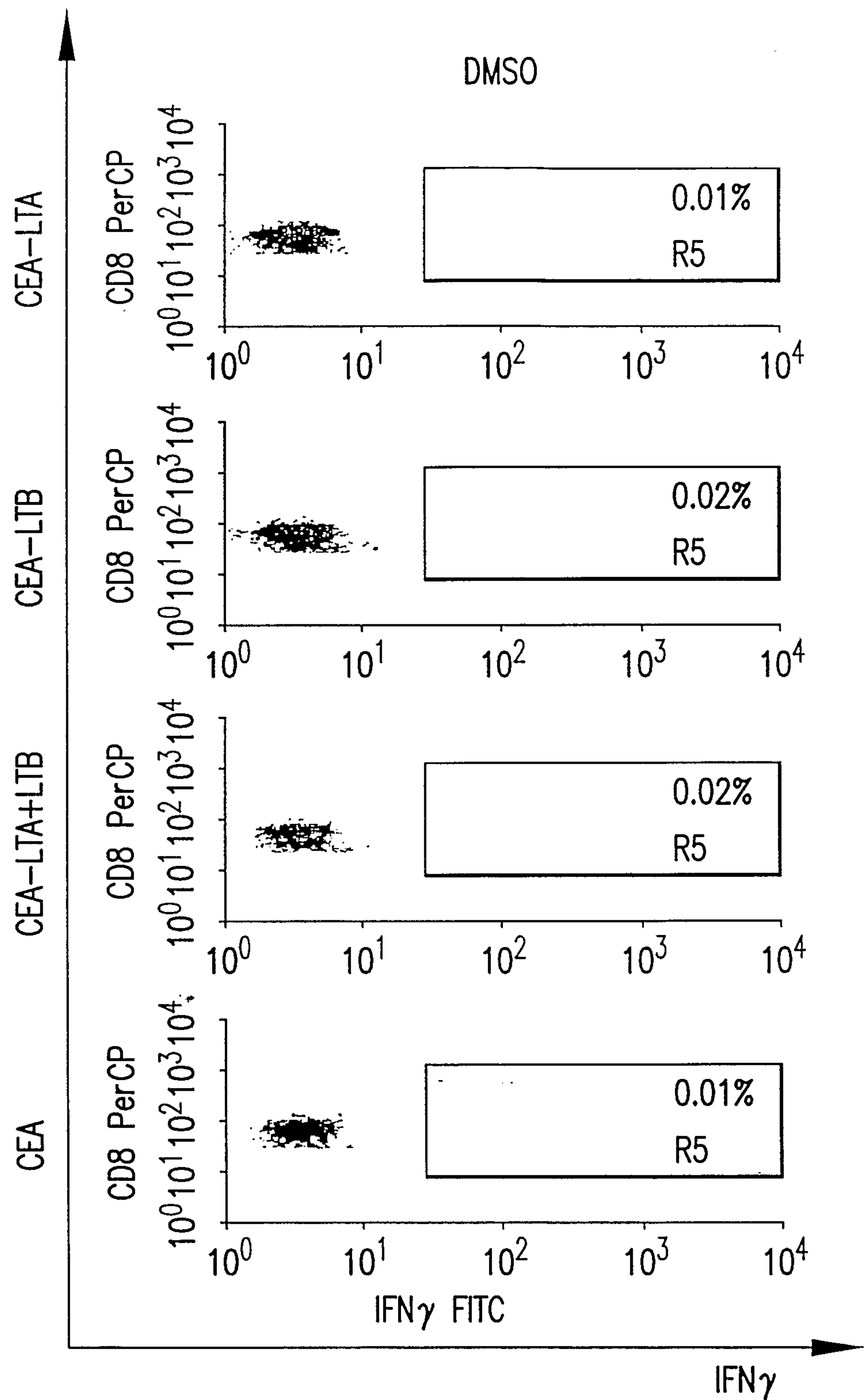


FIG. 14C

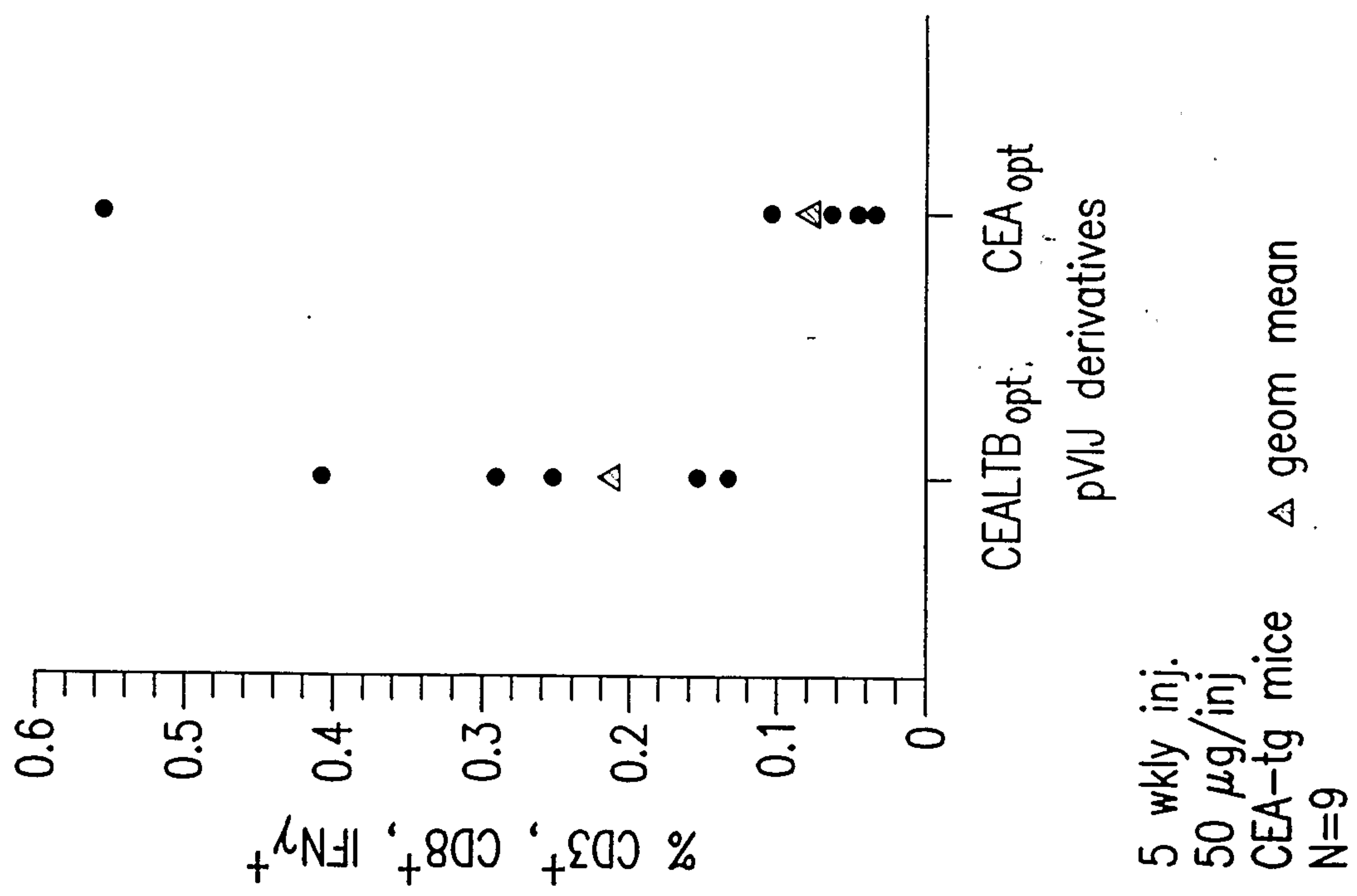


FIG. 15A

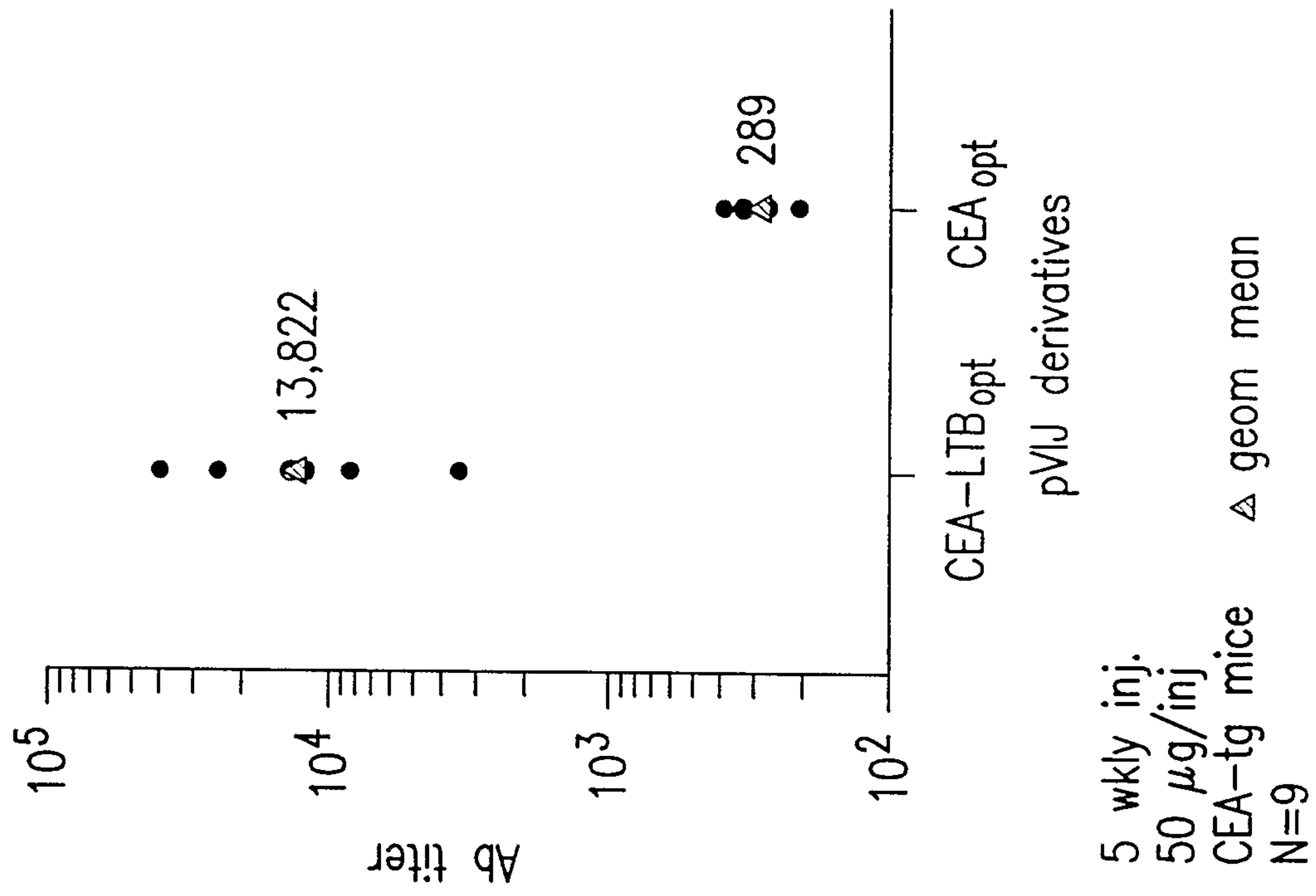


FIG. 15B

34/62

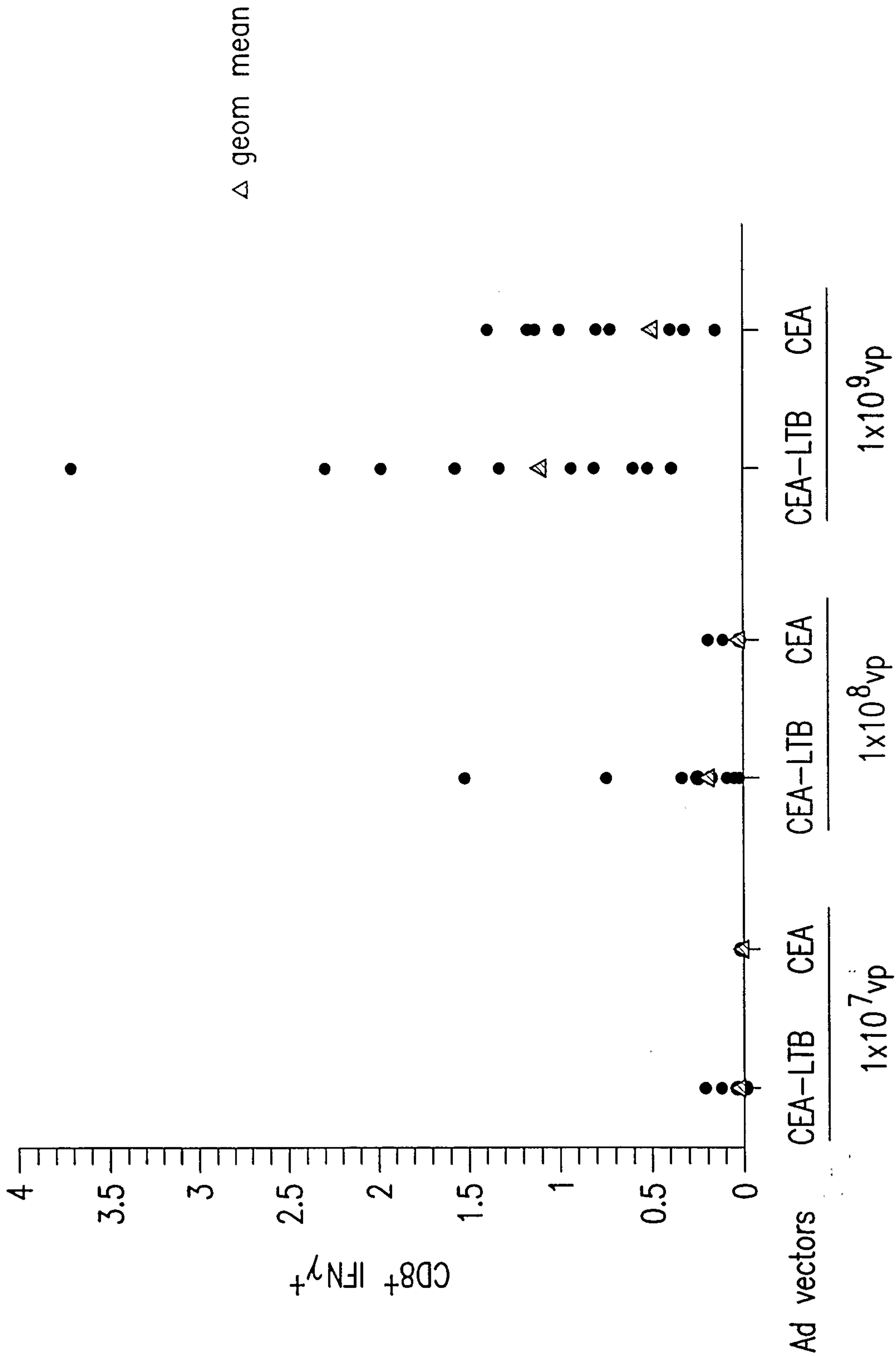


FIG. 16

35/62

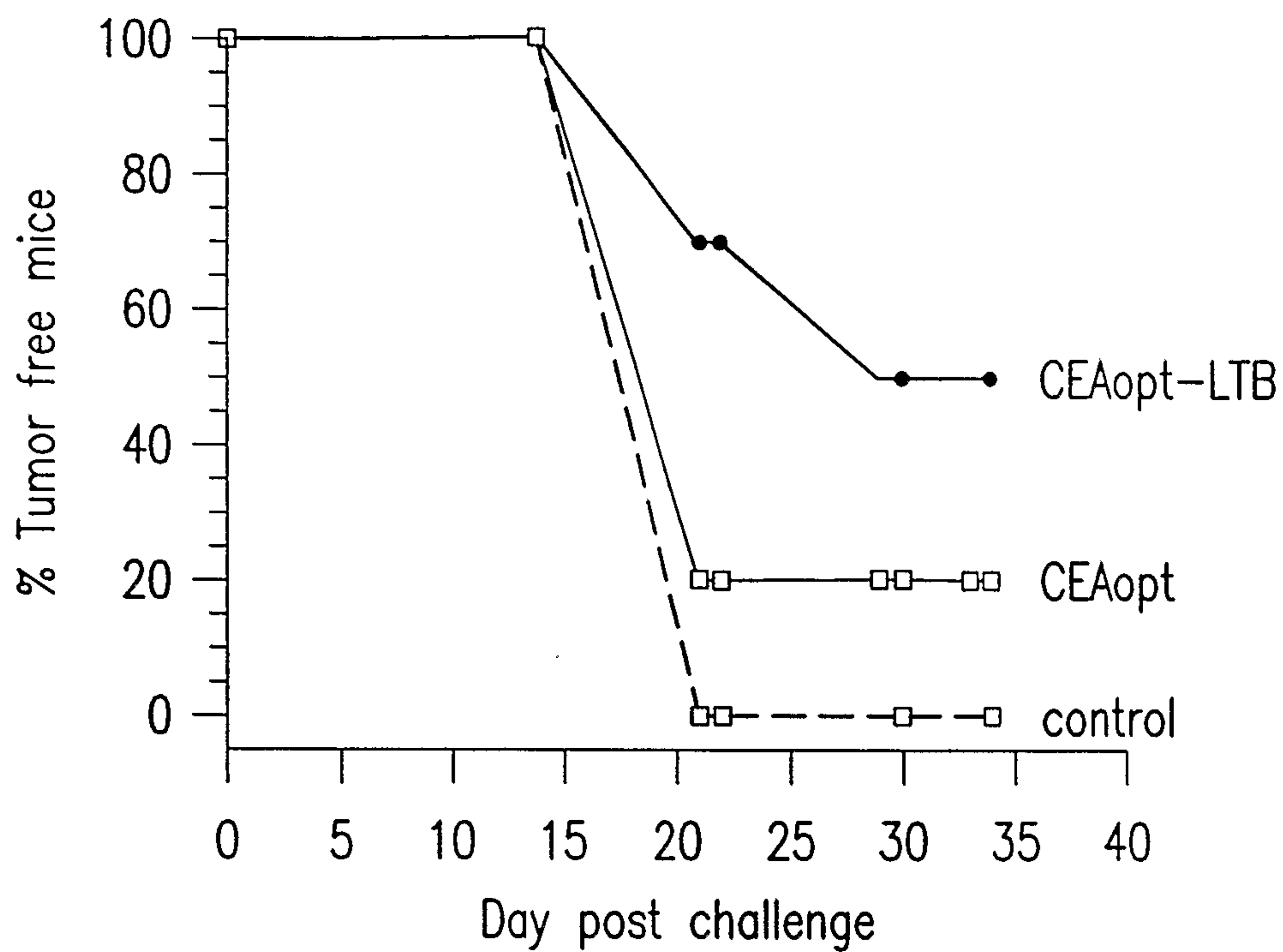


FIG.17A

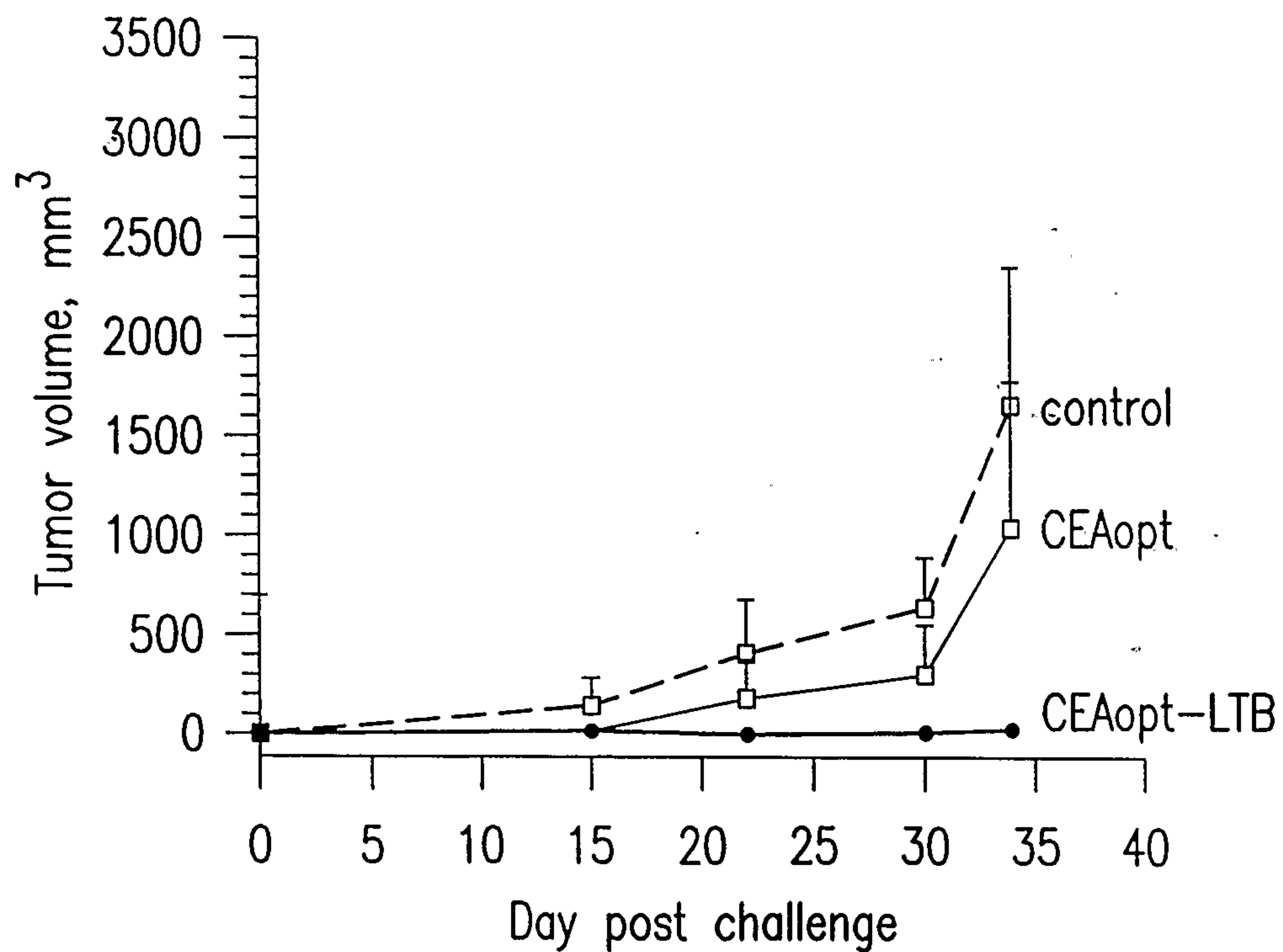


FIG.17B

FIG. 18A

FIG. 18B

37/62

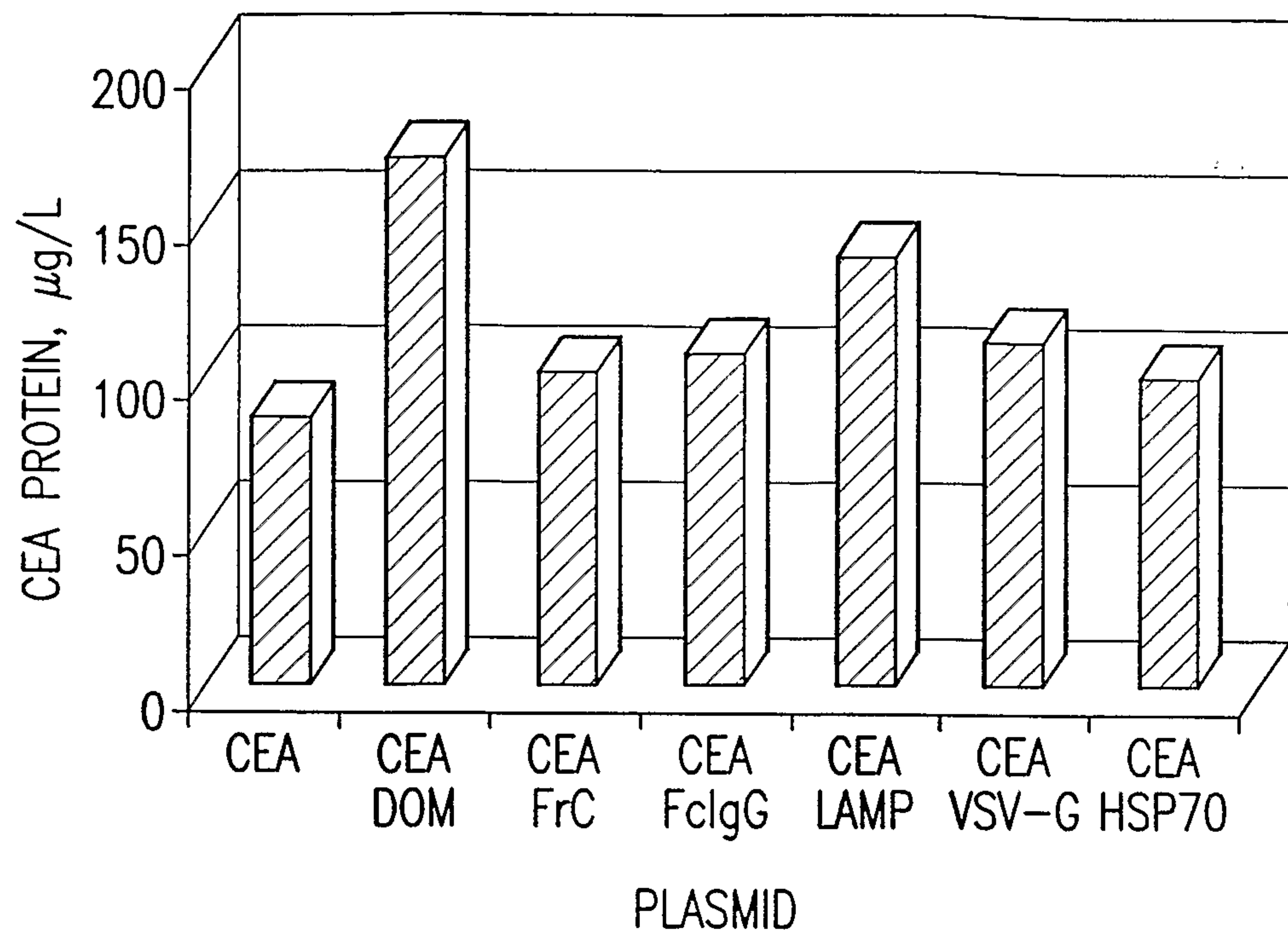


FIG.19A

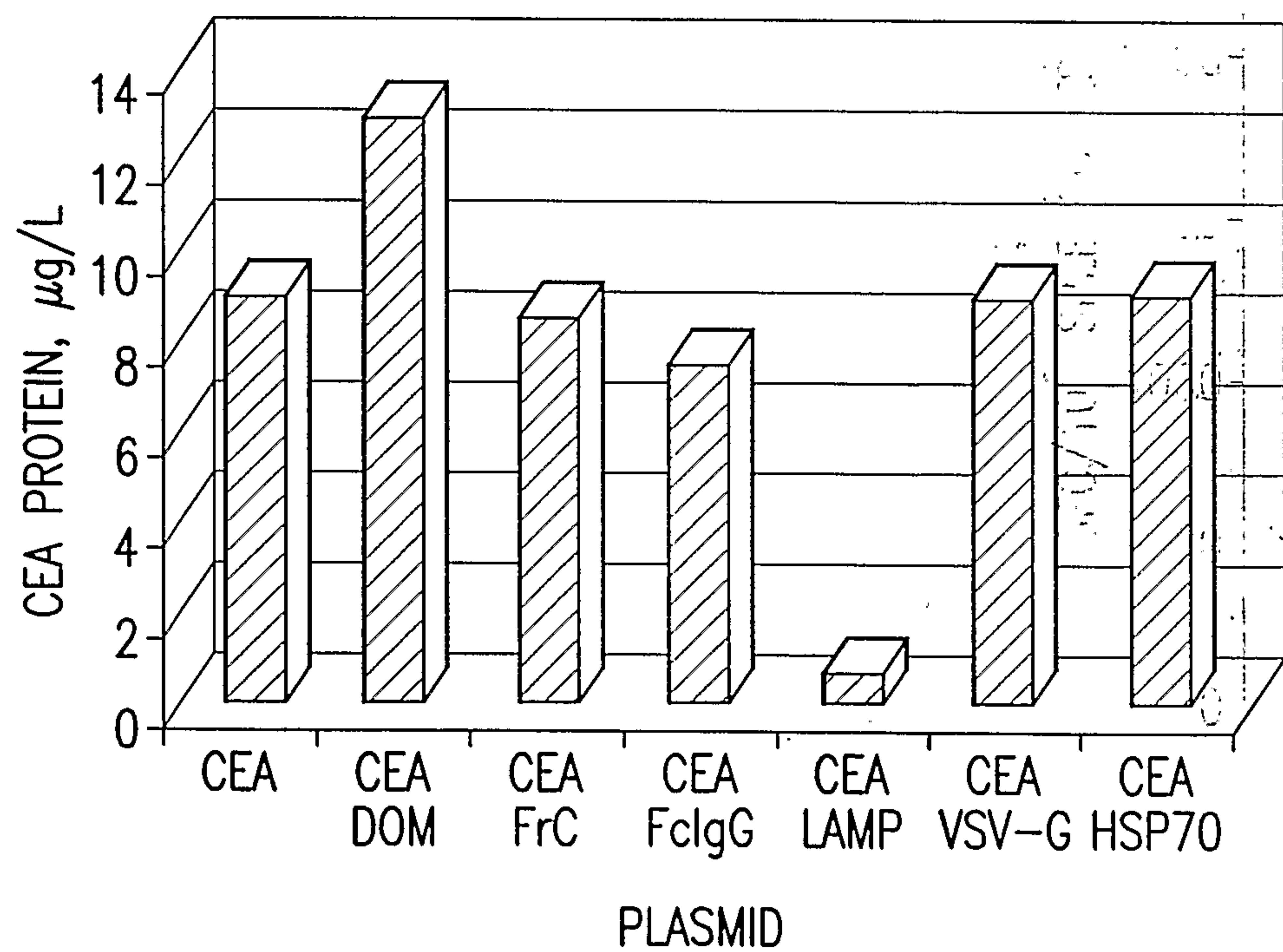


FIG.19B

38/62

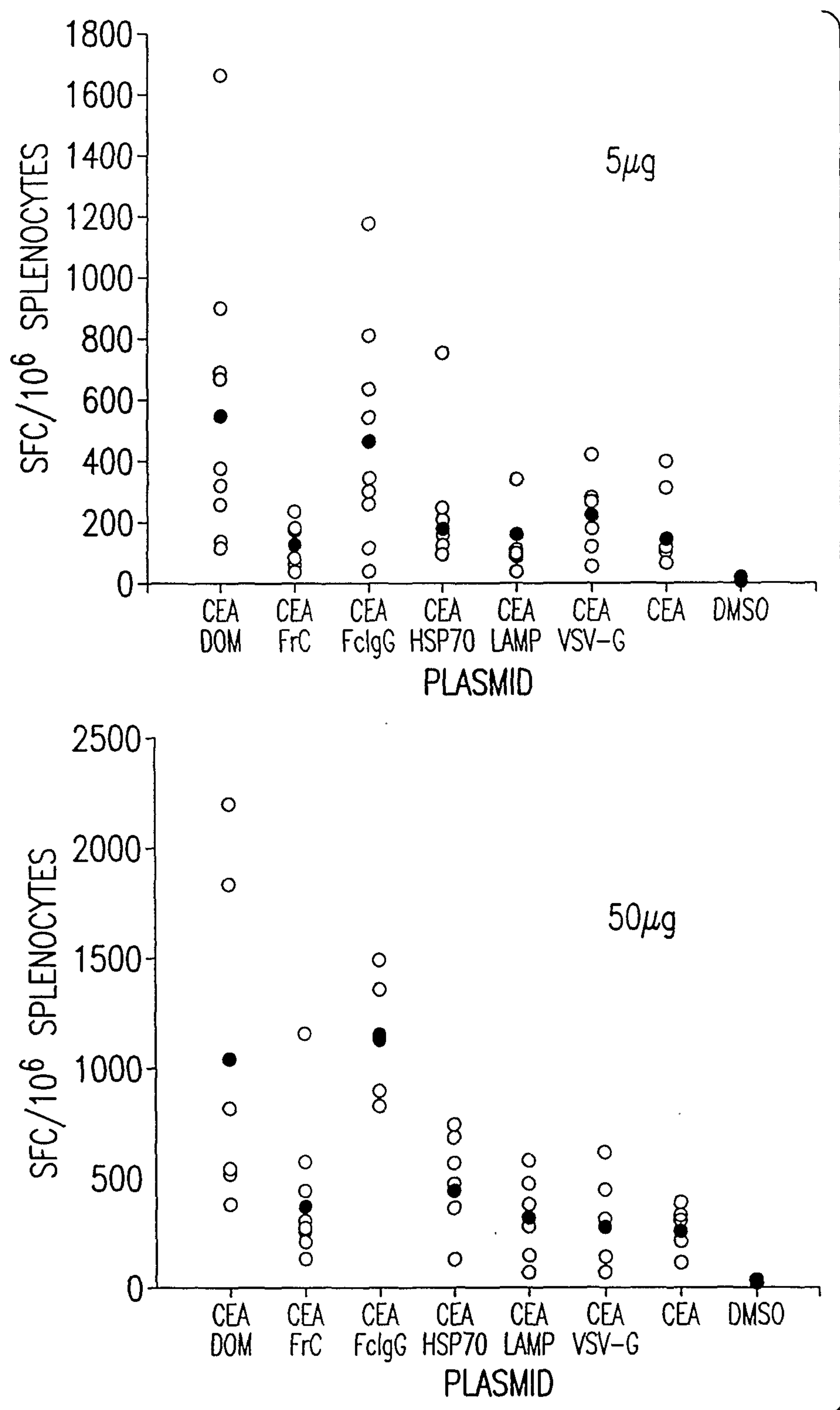


FIG. 20A

39/62

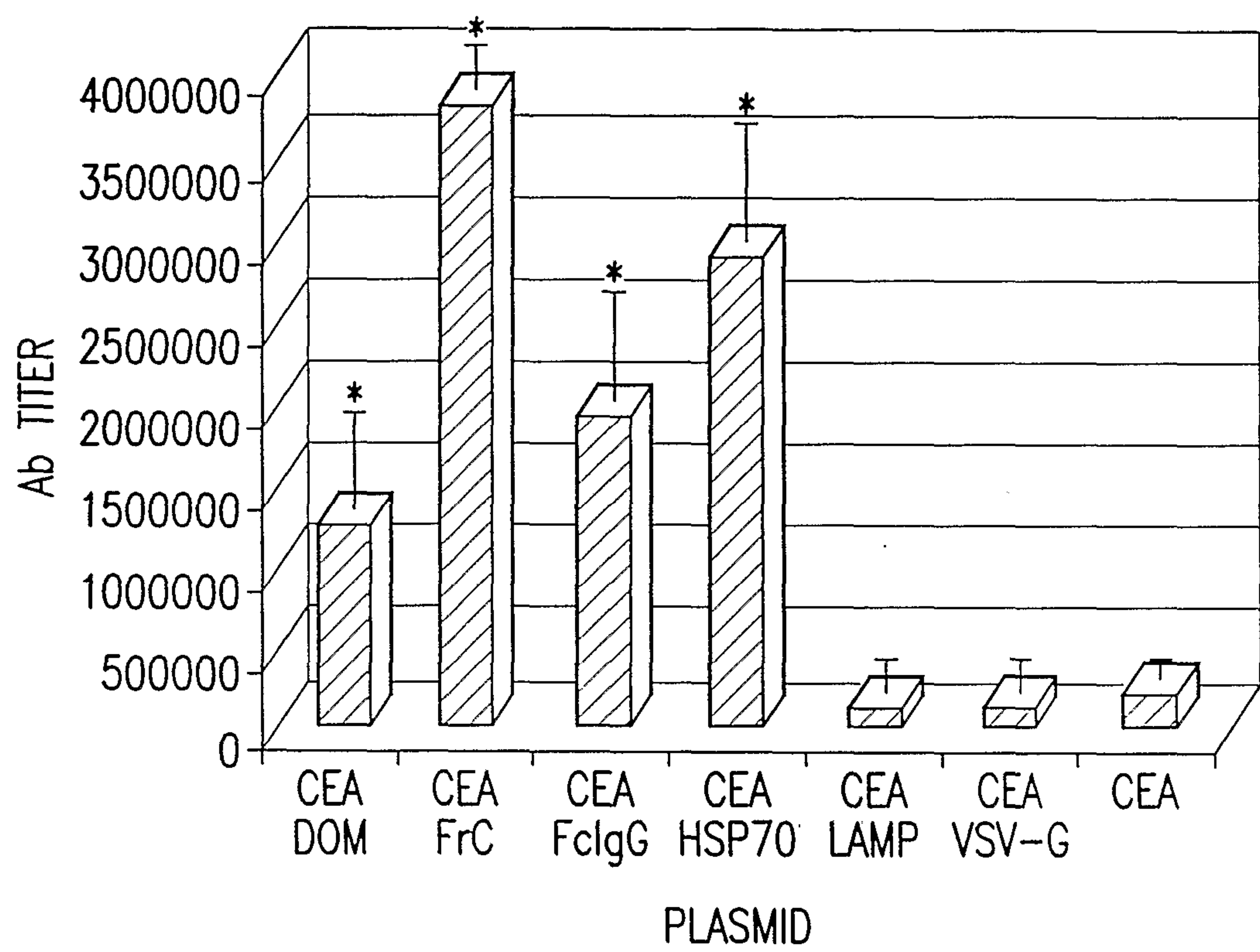


FIG.20B

40/62

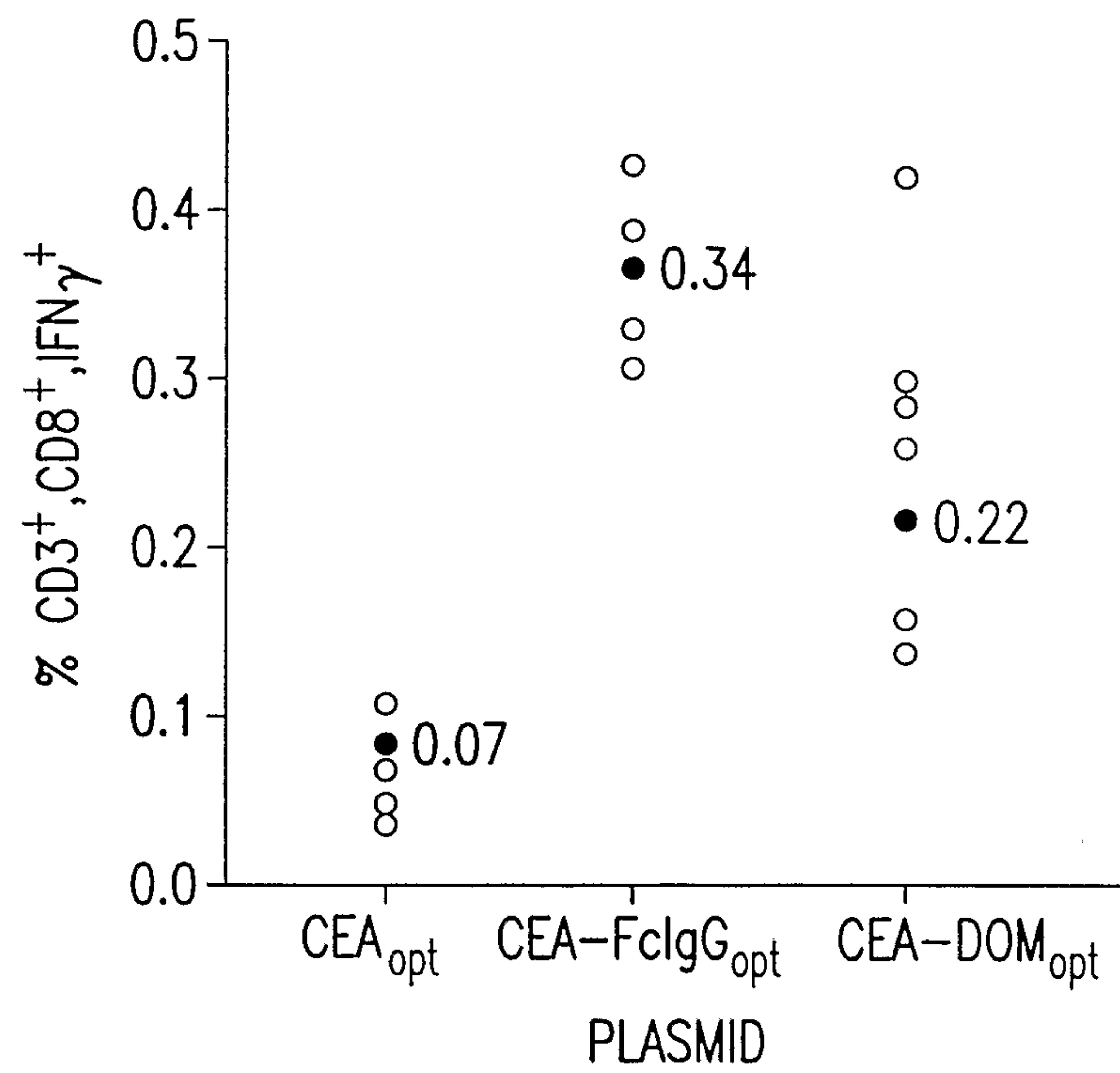


FIG.21A

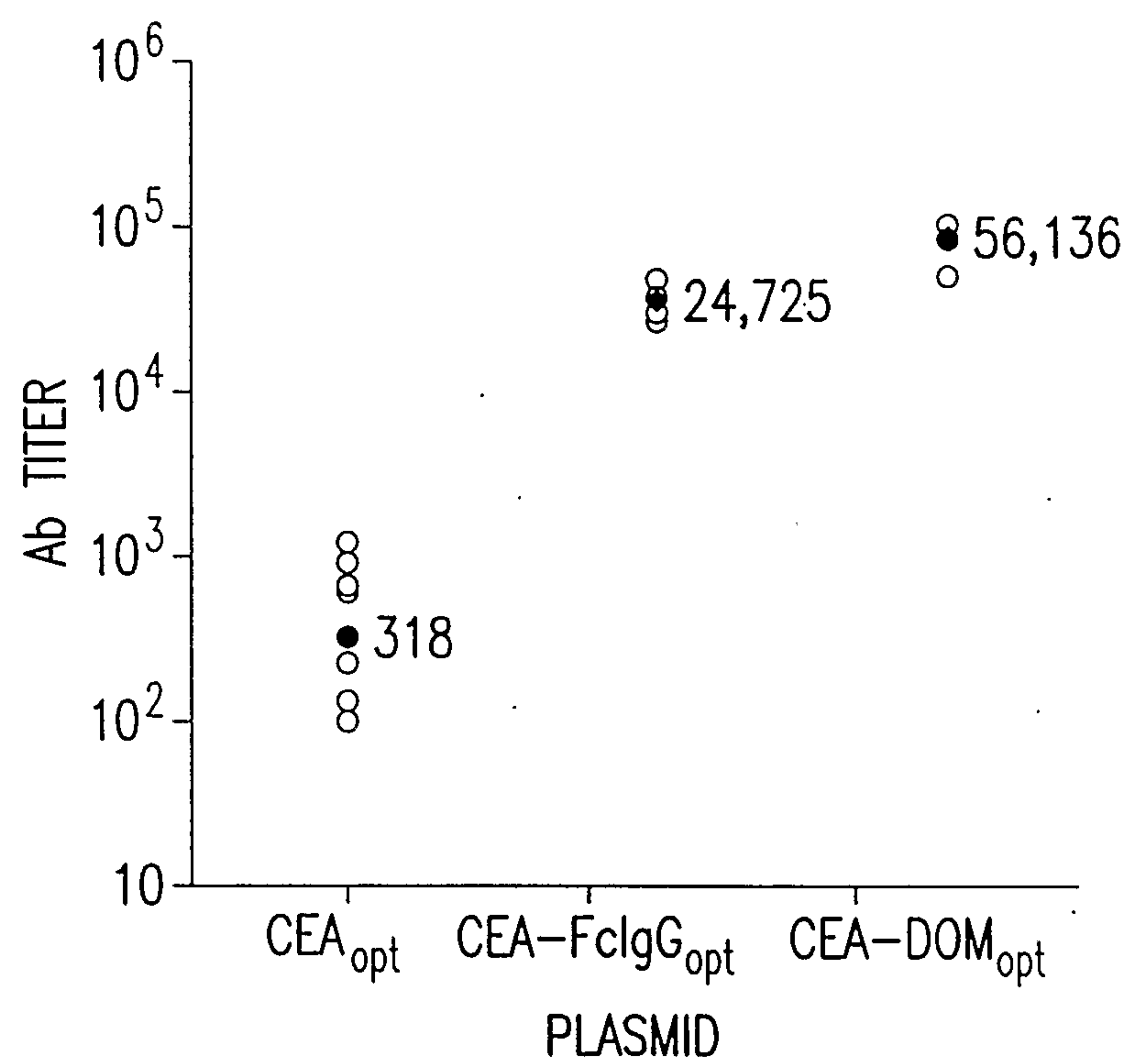


FIG.21B

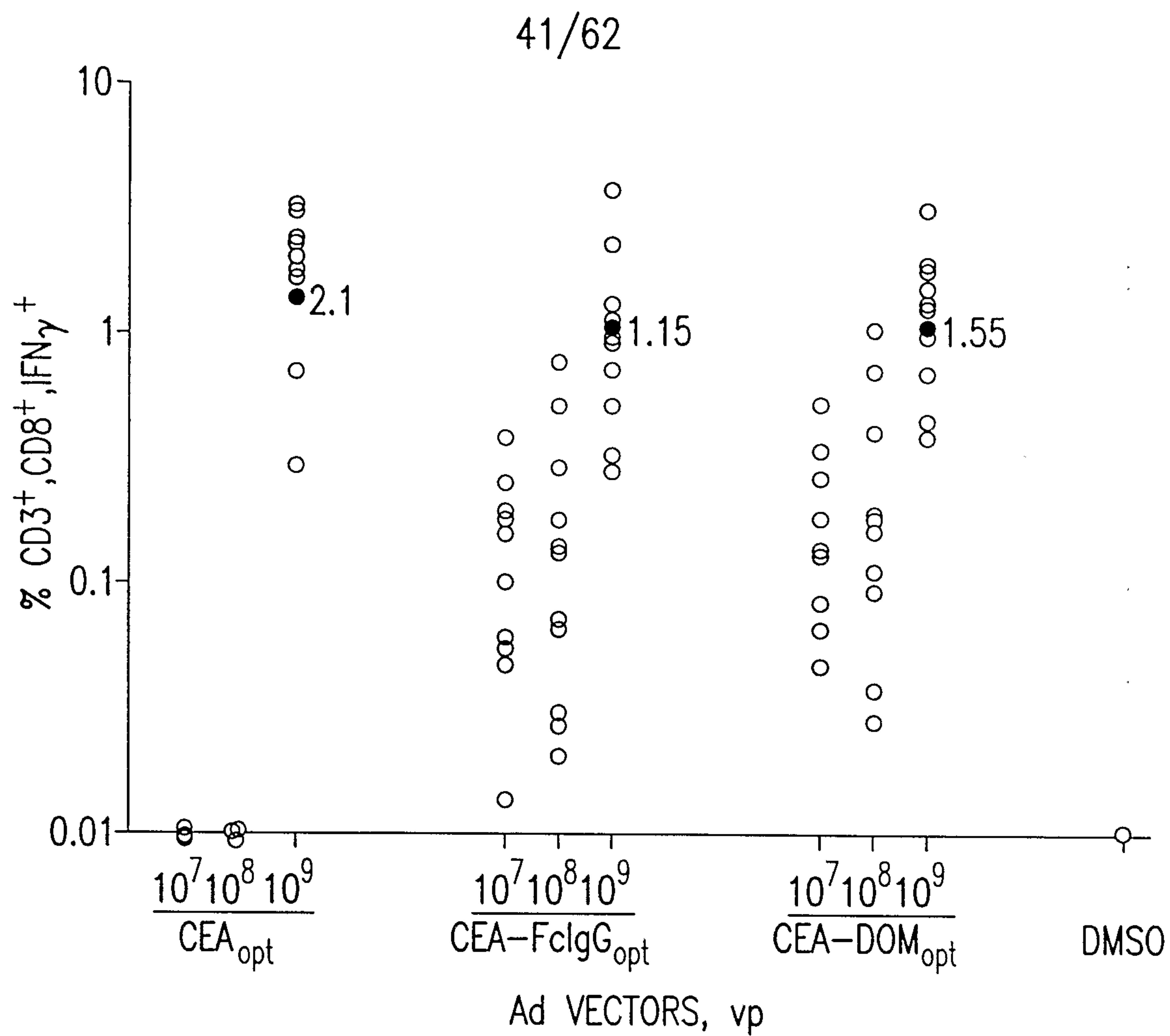


FIG.21C

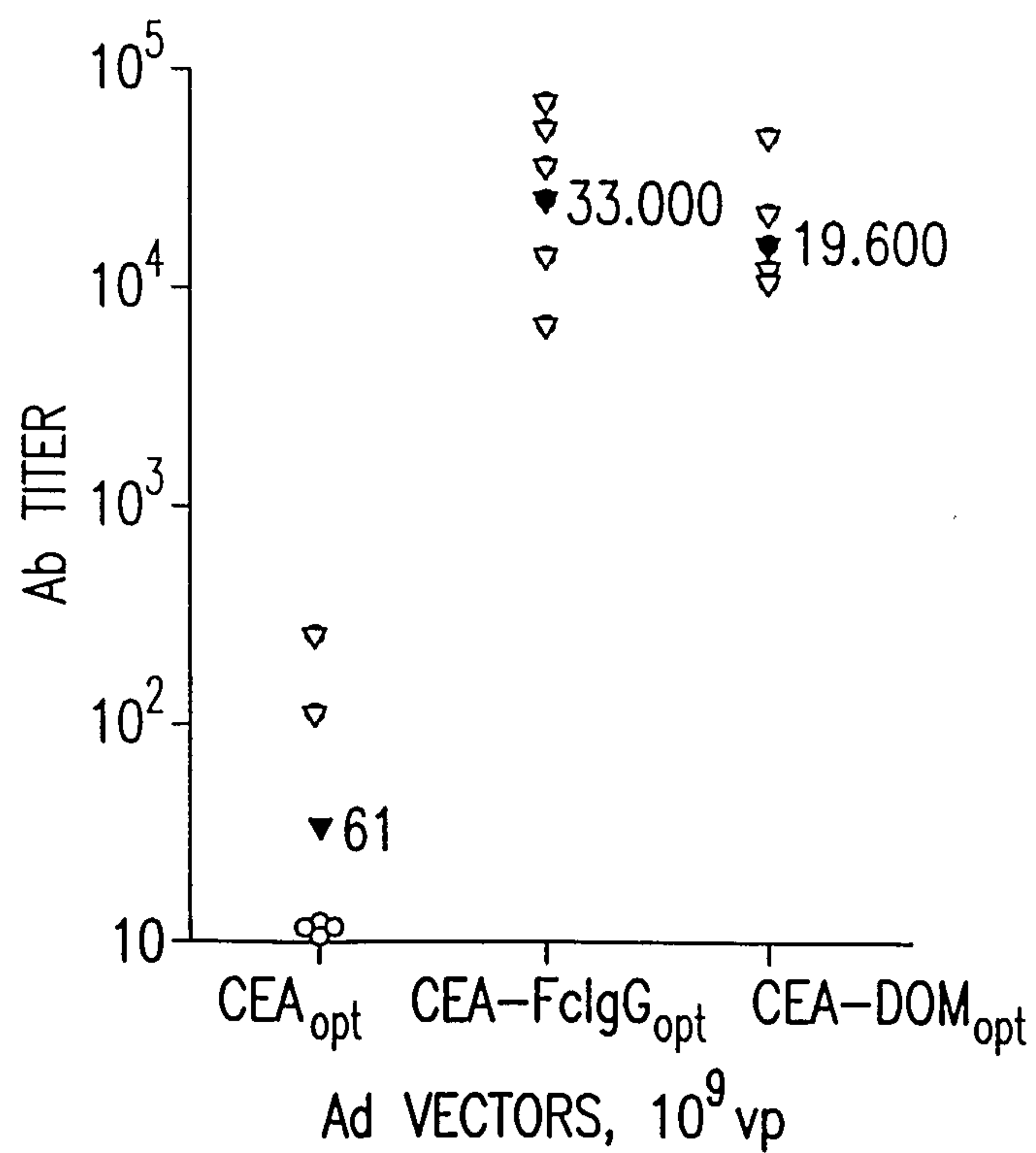


FIG.21D

42/62

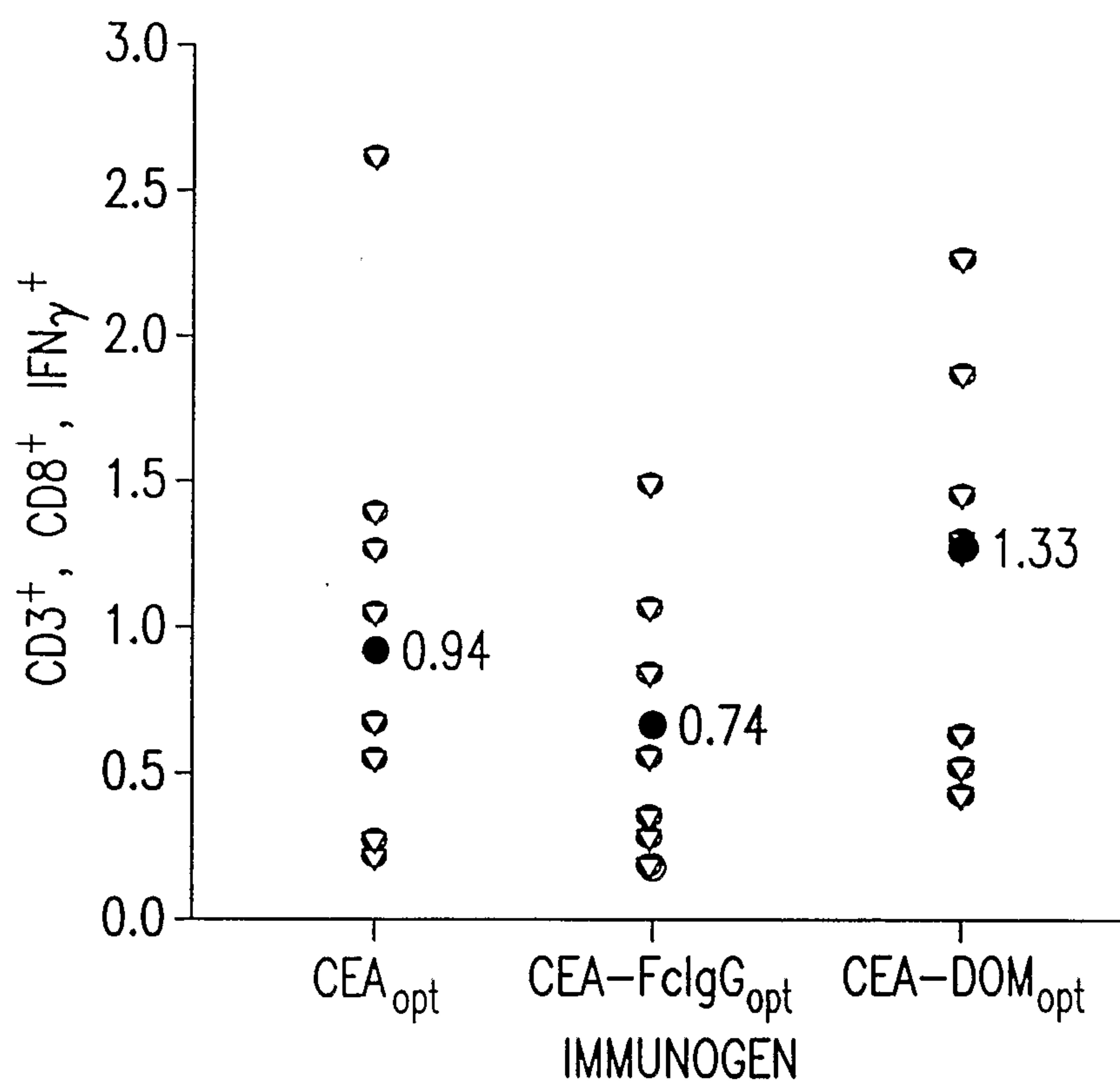


FIG.22A

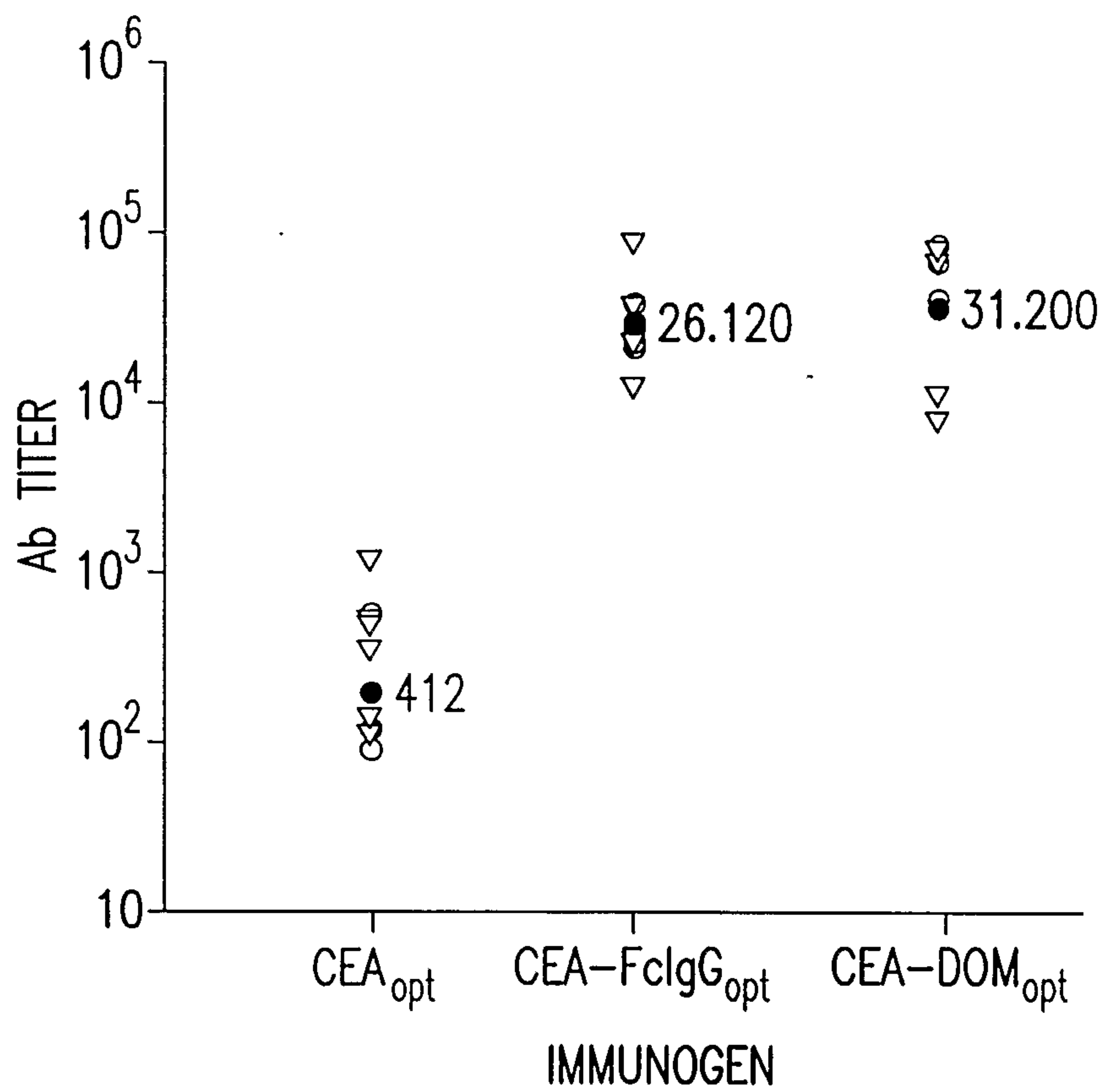


FIG.22B

43/62

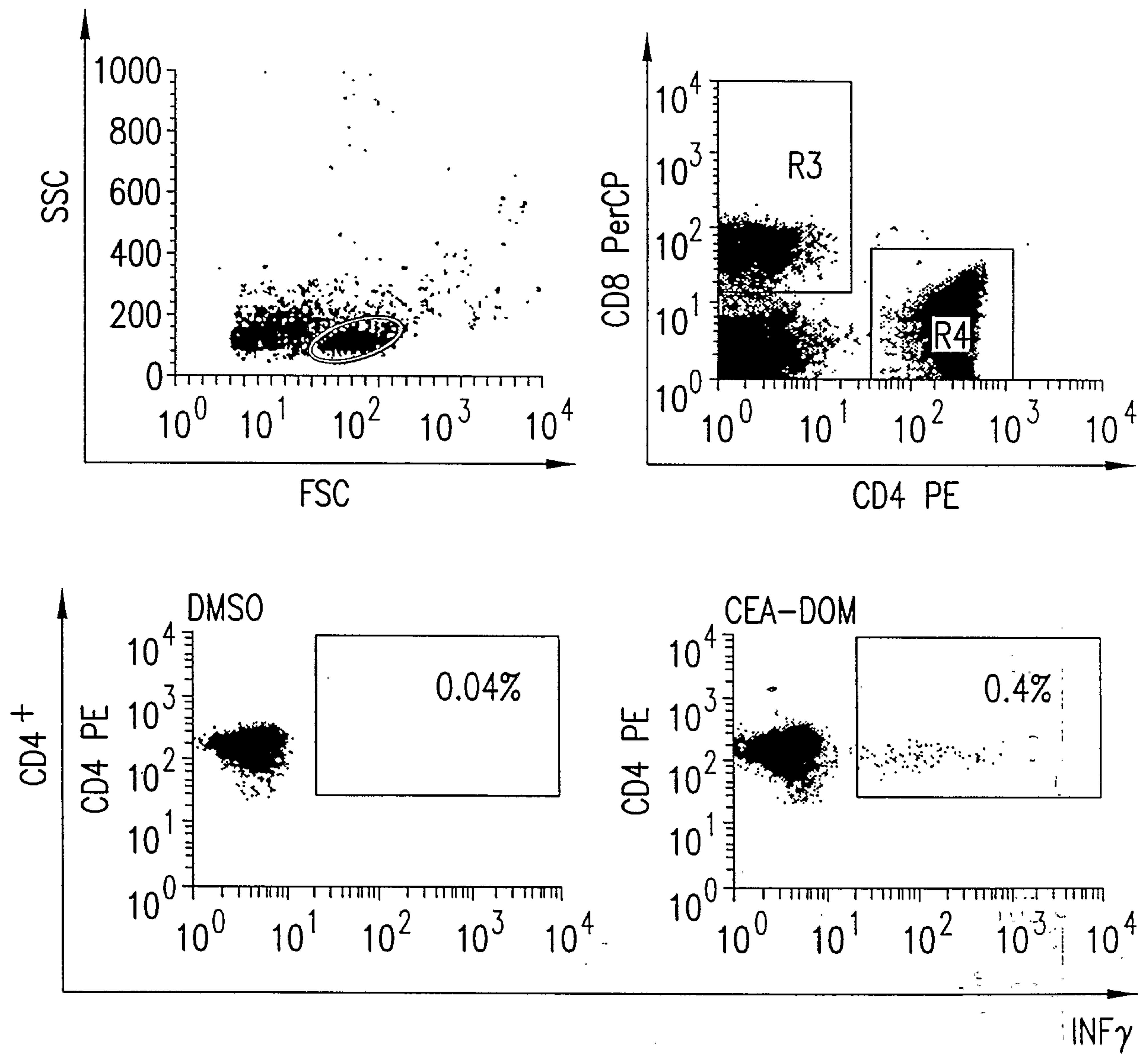


FIG.23

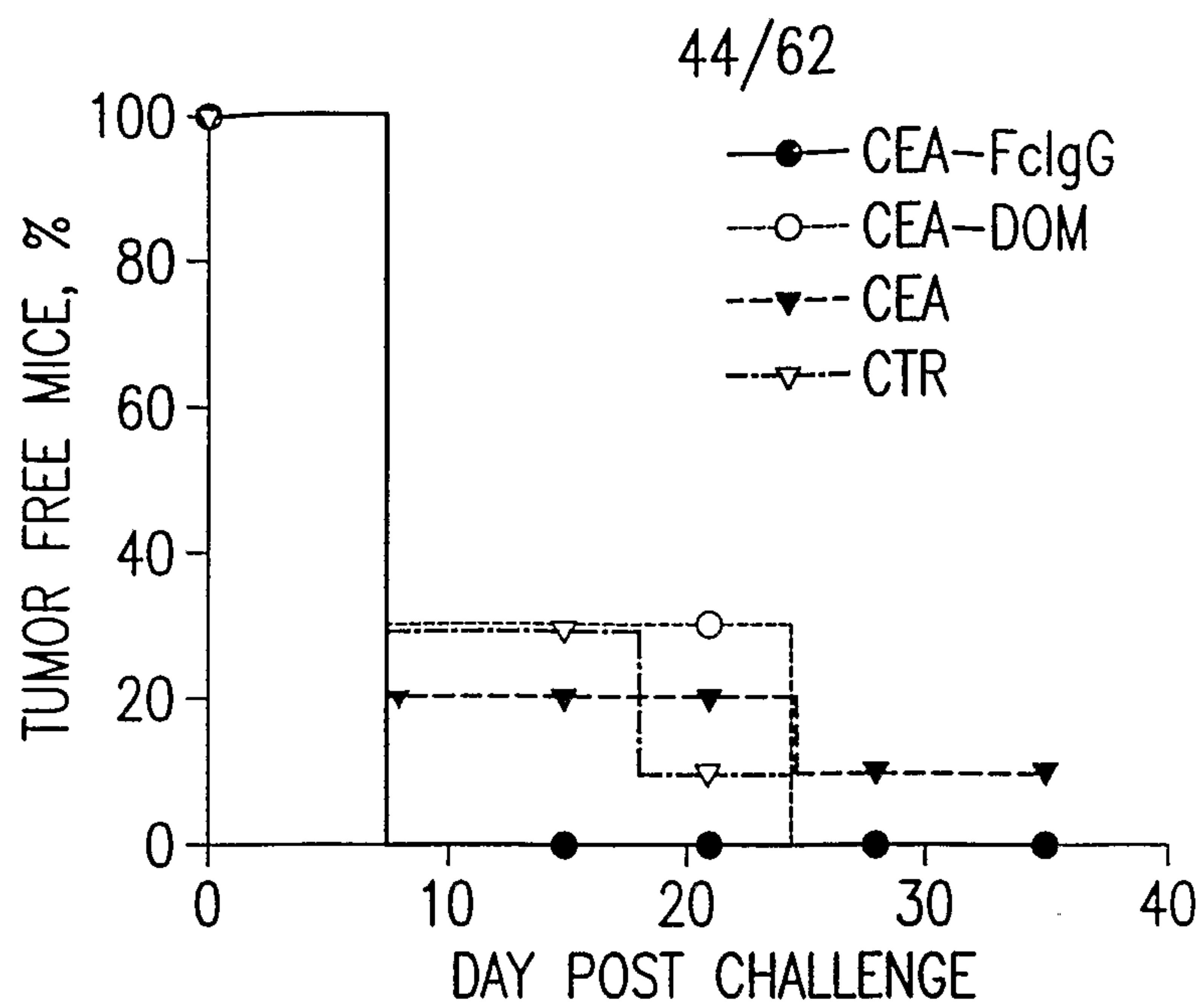


FIG. 24A

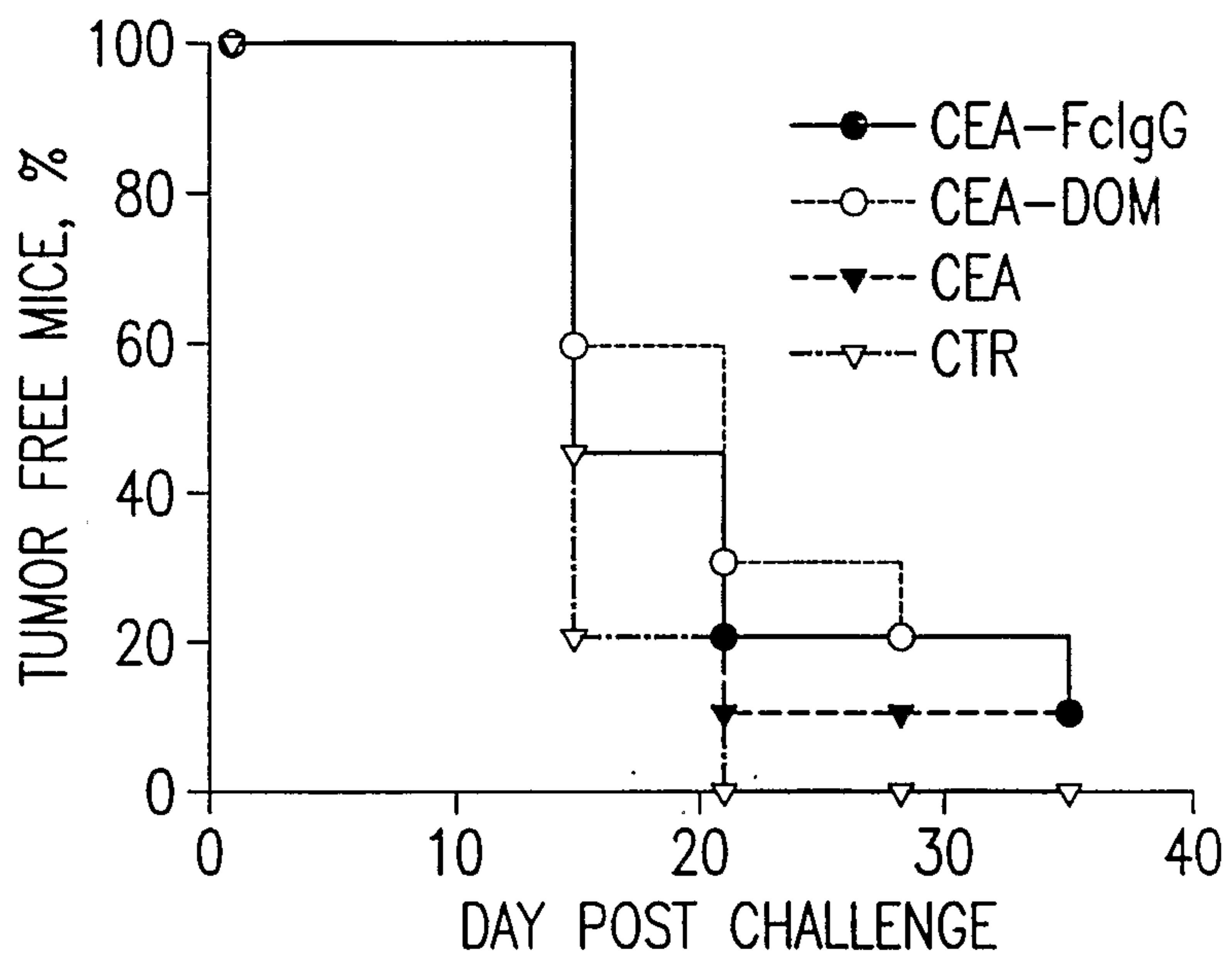


FIG. 24B

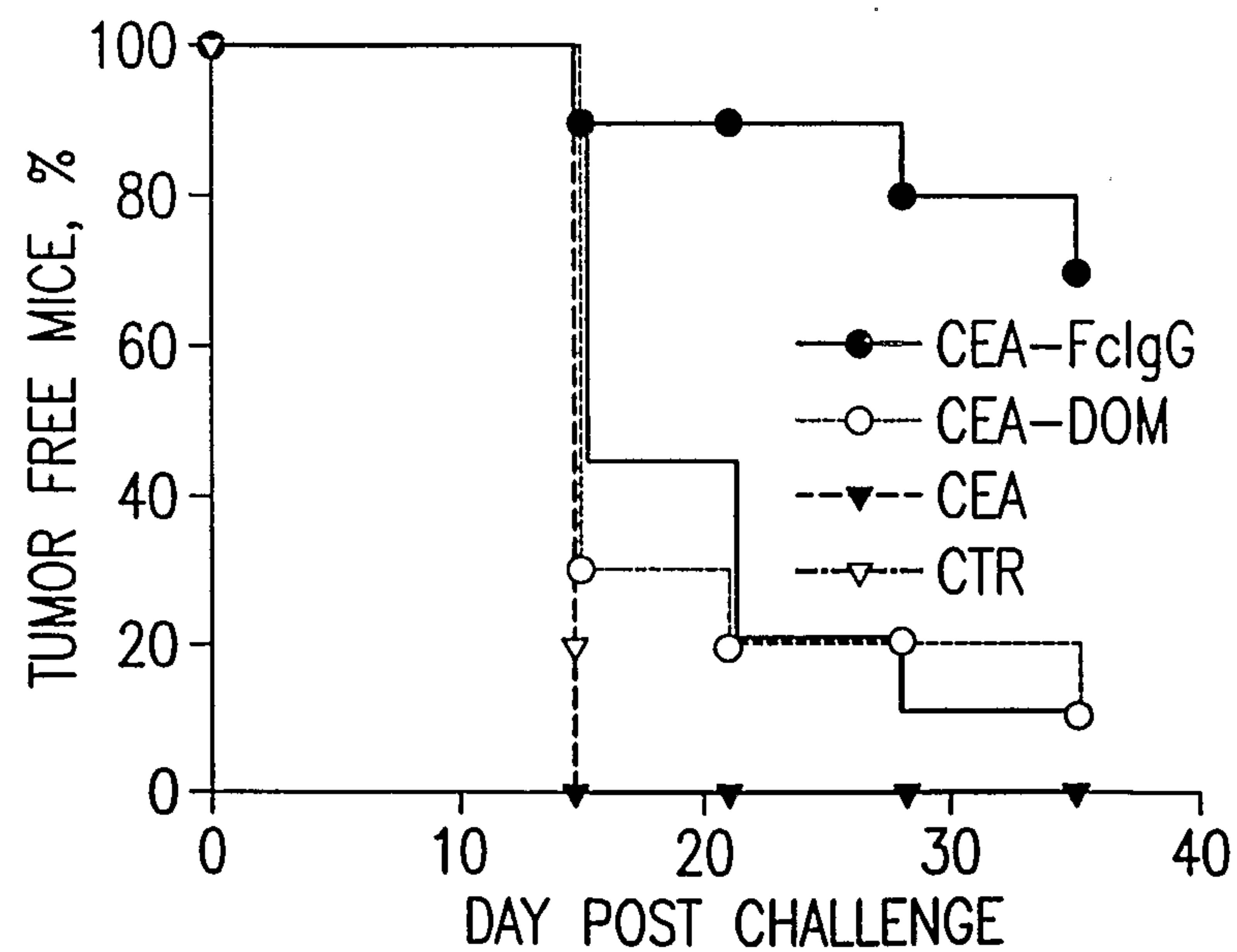


FIG. 24C

45/62

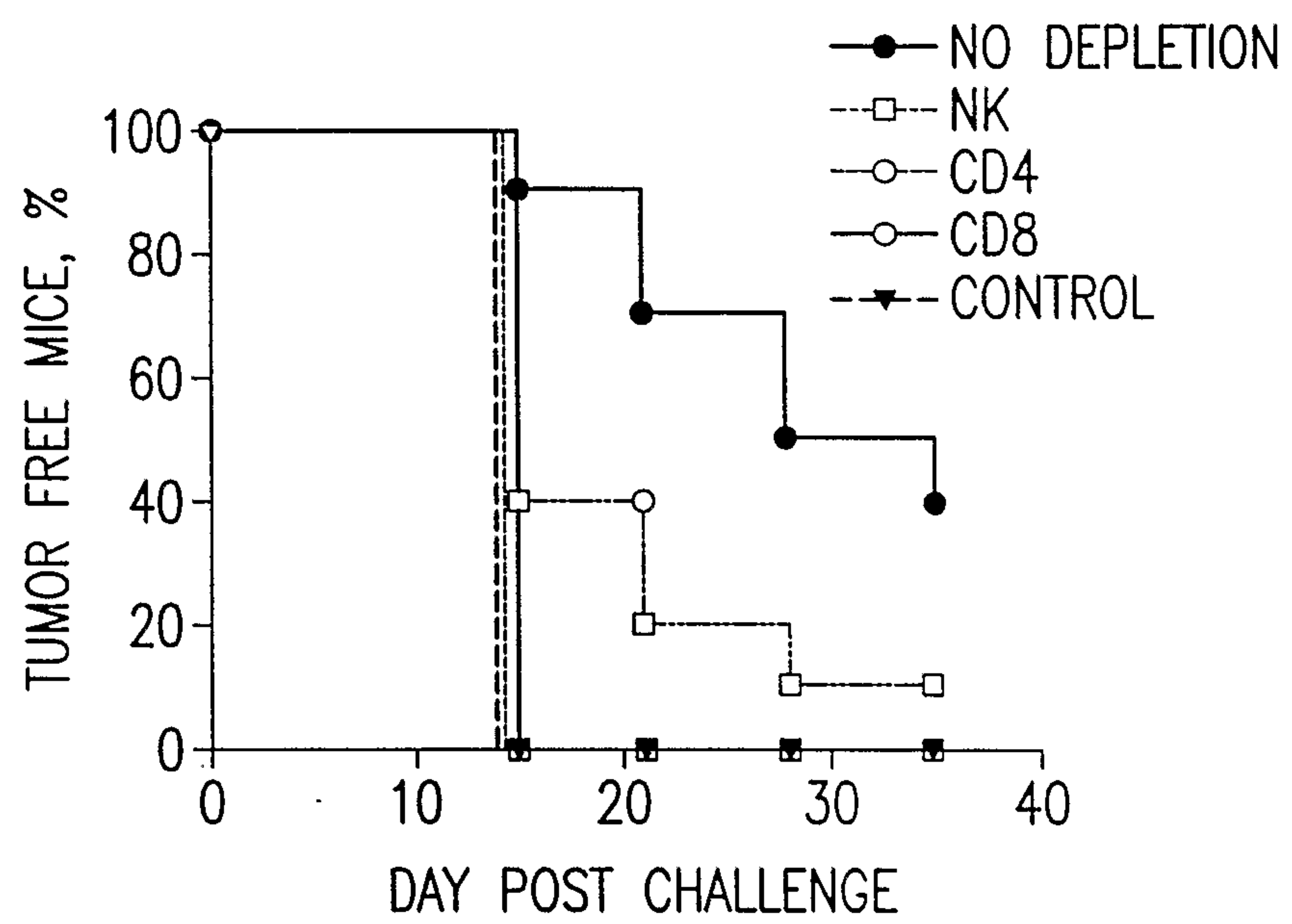


FIG.25

46/62

M E S P S A P P H R W C I P W Q R .
 1 ATGGAGAGCC CCAGCGCCCC CCCCCACCGC TGGTGCATCC CCTGGCAGCG
 · L L L T A S L L T F W N P P T T A
 51 CCTGCTGCTG ACCGCCAGCC TGCTGACCTT CTGGAACCCC CCCACCACCG
 · K L T I E S T P F N V A E G K E
 101 CCAAGCTGAC CATCGAGAGC ACCCCCTTCA ACGTGGCCGA GGGCAAGGAG
 V L L L V H N L P Q H L F G Y S W .
 151 GTGCTGCTGC TGGTGCACAA CCTGCCCCAG CACCTGTTCG GCTACAGCTG
 · Y K G E R V D G N R Q I I G Y V I
 201 GTACAAGGGC GAGCGCGTGG ACGGCAACCG CCAGATCATC GGCTACGTGA
 · G T Q Q A T P G P A Y S G R E I
 251 TCGGCACCCA GCAGGCCACC CCCGGCCCCG CCTACAGCGG CCGCGAGATC
 I Y P N A S L L I Q N I I Q N D T .
 301 ATCTACCCCA ACGCCAGCCT GCTGATCCAG AACATCATCC AGAACGACAC
 · G F Y T L H V I K S D L V N E E A
 351 CGGCTTCTAC ACCCTGCACG TGATCAAGAG CGACCTGGTG AACGAGGAGG
 · T G Q F R V Y P E L P K P S I S
 401 CCACCGGCCA GTTCCGCGTG TACCCCGAGC TGCCCAAGCC CAGCATCAGC
 S N N S K P V E D K D A V A F T C .
 451 AGCAACAACA GCAAGCCCGT GGAGGACAAG GACGCCGTGG CCTTCACCTG
 · E P E T Q D A T Y L W W V N N Q S
 501 CGAGCCCGAG ACCCAGGACG CCACCTACCT GTGGTGGGTG AACAACCAGA
 · L P V S P R L Q L S N G N R T L
 551 GCCTGCCCGT GAGCCCCCGC CTGCAGCTGA GCAACGGCAA CCGCACCTG
 T L F N V T R N D T A S Y K C E T .
 601 ACCCTGTTCA ACGTGACCCG CAACGACACC GCCAGCTACA AGTGCGAGAC
 · Q N P V S A R R S D S V I L N V L
 651 CCAGAACCCC GTGAGCGCCC GCCGCAGCGA CAGCGTGATC CTGAACGTGC
 · Y G P D A P T I S P L N T S Y R
 701 TGTACGGCCC CGACGCCCCC ACCATCAGCC CCCTGAACAC CAGCTACCGC
 S G E N L N L S C H A A S N P P A .
 751 AGCGGCGAGA ACCTGAACCT GAGCTGCCAC GCCGCCAGCA ACCCCCCCGC
 · Q Y S W F V N G T F Q Q S T Q E L
 801 CCAGTACAGC TGGTTCGTGA ACGGCACCTT CCAGCAGAGC ACCCAGGAGC
 · F I P N I T V N N S G S Y T C Q
 851 TGTTTCATCCC CAACATCACC GTGAACAACA GCGGCAGCTA CACCTGCCAG
 A H N S D T G L N R T T V T T I T .
 901 GCCCACAACA GCGACACCGG CCTGAACCGC ACCACCGTGA CCACCATCAC
 · V Y A E P P K P F I T S N N S N P
 951 CGTGTACGCC GAGCCCCCA AGCCCTTCAT CACCAGCAAC AACAGCAACC
 · V E D E D A V A L T C E P E I Q

FIG.26A-1

47/62

1001 CCGTGGAGGA CGAGGACGCC GTGGCCCTGA CCTGCGAGCC CGAGATCCAG
 N T T Y L W W V N N Q S L P V S P .
 1051 AACACCACCT ACCTGTGGTG GGTGAACAAC CAGAGCCTGC CCGTGAGCCC
 . R L Q L S N D N R T L T L L S V T .
 1101 CCGCCTGCAG CTGAGCAACG ACAACCGCAC CCTGACCCTG CTGAGCGTGA
 . R N D V G P Y E C G I Q N E L S
 1151 CCCGCAACGA CGTGGGCCCC TACGAGTGCG GCATCCAGAA CGAGCTGAGC
 V D H S D P V I L N V L Y G P D D .
 1201 GTGGACCACA GCGACCCCGT GATCCTGAAC GTGCTGTACG GCCCCGACGA
 . P T I S P S Y T Y Y R P G V N L S .
 1251 CCCCACCATC AGCCCCAGCT ACACCTACTA CCGCCCCGGC GTGAACCTGA
 . L S C H A A S N P P A Q Y S W L
 1301 GCCTGAGCTG CCACGCCGCC AGCAACCCCC CCGCCCAGTA CAGCTGGCTG
 I D G N I Q Q H T Q E L F I S N I .
 1351 ATCGACGGCA ACATCCAGCA GCACACCCAG GAGCTGTTCA TCAGCAACAT
 . T E K N S G L Y T C Q A N N S A S .
 1401 CACCGAGAAG AACAGCGGCC TGTACACCTG CCAGGCCAAC AACAGCGCCA
 . G H S R T T V K T I T V S A E L
 1451 GCGGCCACAG CCGCACCACC GTGAAGACCA TCACCGTGAG CGCCGAGCTG
 P K P S I S S N N S K P V E D K D .
 1501 CCCAAGCCCA GCATCAGCAG CAACAACAGC AAGCCCGTGG AGGACAAGGA
 . A V A F T C E P E A Q N T T Y L W .
 1551 CGCCGTGGCC TTCACCTGCG AGCCCGAGGC CCAGAACACC ACCTACCTGT
 . W V N G Q S L P V S P R L Q L S
 1601 GGTGGGTGAA CGGCCAGAGC CTGCCCCTGA GCCCCCGCCT GCAGCTGAGC
 N G N R T L T L F N V T R N D A R .
 1651 AACGGCAACC GCACCCTGAC CCTGTTCAAC GTGACCCGCA ACGACGCCCG
 . A Y V C G I Q N S V S A N R S D P .
 1701 CGCCTACGTG TCGGGCATCC AGAACAGCGT GAGCGCCAAC GGGAGGGAGC
 . V T L D V L Y G P D T P I I S P .
 1751 CCGTGACCCT GGACGTGCTG TACGGCCCCG ACACCCCATC GATCAGCCCC
 P D S S Y L S G A N L N L S C H S .
 1801 CCCGACAGCA GCTACCTGAG CGGCGCCAAC CTGAACCTGA GCTGCCACAG
 . A S N P S P Q Y S W R I N G I P Q .
 1851 CGCCAGCAAC CCCAGCCCCC AGTACAGCTG GCGCATCAAC GGCATCCCCC
 . Q H T Q V L F I A K I T P N N N
 1901 AGCAGCACAC CCAGGTGCTG TTCATCGCCA AGATCACCCC CAACAACAAC
 G T Y A C F V S N L A T G R N N S .
 1951 GGCACCTACG CCTGCTTCGT GAGCAACCTG GCCACCGGCC GCAACAACAG
 . I V K S I T V S A S G T S R S T P .
 2001 CATCGTGAAG AGCATCACCG TGAGCGCCAG CGGCACCTCT AGAAGCACCC
 . I P F S Y S K N L D C W V D N E

FIG.26A-2

48/62

2051 CCATCCCATT CAGCTACAGC AAGAACCTGG ACTGCTGGGT GGACAACGAG
E D I D V I L K K S T I L N L D I .
2101 GAGGACATCG ACGTGATCCT GAAGAAGAGC ACCATCCTGA ACCTGGACAT
· N N D I I S D I S G F N S S V I T .
2151 CAACAACGAC ATCATCAGCG ACATCAGCGG CTTCAACAGC AGCGTGATCA
· Y P D A Q L V P G I N G K A I H
2201 CCTACCCCGA CGCCCAGCTG GTGCCC GGCA TCAACGGCAA GGCCATCCAC
L V N N E S S E V I V H K A M D I .
2251 CTGGTGAACA ACGAGAGCAG CGAGGTGATC GTGCACAAGG CCATGGACAT
· E Y N D M F N N F T V S F W L R V .
2301 CGAGTACAAC GACATGTTCA ACAACTTCAC CGTGAGCTTC TGGCTGAGAG
· P K V S A S H L E Q Y G T N E Y
2351 TGCCTAAGGT GAGCGCCAGC CACCTGGAGC AGTACGGCAC CAACGAGTAC
S I I S S M K K H S L S I G S G W .
2401 AGCATCATCA GCAGCATGAA GAAGCACAGC CTGAGCATCG GCAGCGGCTG
· S V S L K G N N L I W T L K D S A .
2451 GAGCGTGAGC CTGAAGGGCA ACAACCTCAT CTGGACCCTG AAGGATAGCG
· G E V R Q I T F R D L P D K F N
2501 CCGGAGAGGT GAGACAGATC ACCTTCAGAG ACCTGCCCGA CAAGTTCAAT
A Y L A N K W V F I T I T N D R L .
2551 GCCTACCTGG CCAACAAGTG GGTGTTTCATC ACCATCACCA ACGACAGACT
· S S A N L Y I N G V L M G S A E I .
2601 GAGCAGCGCC AACCTGTACA TCAACGGCGT GCTCATGGGC AGCGCCGAGA
· T G L G A I R E D N N I T L K L
2651 TCACCGGCCT GGGCGCCATC AGAGAGGACA ACAACATCAC CCTGAAGCTG
D R C N N N N Q Y V S I D K F R I .
2701 GACAGATGCA ACAACAACAA CCAGTACGTG AGCATCGACA AGTTCCGGAT
· F C K A L N P K E I E K L Y T S Y .
2751 CTTCTGCAAG GCCCTGAACC CCAAGGAGAT CGAGAAGCTG TACACCAGCT
· L S I T F L R D F W G N P L R Y
2801 ACCTGAGCAT CACCTTCCTG AGAGACTTCT GGGGCAACCC CCTGAGATAC
D T * (SEQ ID NO:45)
2851 GACACCTAG (SEQ ID NO:21)

FIG.26A-3

49/62

M E S P S A P P H R W C I P W Q R L L L
 1 ATGGAGAGCC CCAGCGCCCC CCCCCACCGC TGGTGCATCC CCTGGCAGCG CCTGCTGCTG
 T A S L L T F W N P P T T A K L T I E S
 61 ACCGCCAGCC TGCTGACCTT CTGGAACCCC CCCACCACCG CCAAGCTGAC CATCGAGAGC
 T P F N V A E G K E V L L L V H N L P Q
 121 ACCCCCTTCA ACGTGGCCGA GGGCAAGGAG GTGCTGCTGC TGGTGCACAA CCTGCCCCAG
 H L F G Y S W Y K G E R V D G N R Q I I
 181 CACCTGTTCG GCTACAGCTG GTACAAGGGC GAGCGCGTGG ACGGCAACCG CCAGATCATC
 G Y V I G T Q Q A T P G P A Y S G R E I
 241 GGCTACGTGA TCGGCACCCA GCAGGCCACC CCCGGCCCCG CCTACAGCGG CCGCGAGATC
 I Y P N A S L L I Q N I I Q N D T G F Y
 301 ATCTACCCCA ACGCCAGCCT GCTGATCCAG AACATCATCC AGAACGACAC CGGCTTCTAC
 T L H V I K S D L V N E E A T G Q F R V
 361 ACCCTGCACG TGATCAAGAG CGACCTGGTG AACGAGGAGG CCACCGGCCA GTTCCGCGTG
 Y P E L P K P S I S S N N S K P V E D K
 421 TACCCCGAGC TGCCCAAGCC CAGCATCAGC AGCAACAACA GCAAGCCCGT GGAGGACAAG
 D A V A F T C E P E T Q D A T Y L W W V
 481 GACGCCGTGG CCTTCACCTG CGAGCCCGAG ACCCAGGACG CCACCTACCT GTGGTGGGTG
 N N Q S L P V S P R L Q L S N G N R T L
 541 AACAAACCAGA GCCTGCCCCG GAGCCCCCGC CTGCAGCTGA GCAACGGCAA CCGCACCCCTG
 T L F N V T R N D T A S Y K C E T Q N P
 601 ACCCTGTTCA ACGTGACCCG CAACGACACC GCCAGCTACA AGTGCGAGAC CCAGAACCCC
 V S A R R S D S V I L N V L Y G P D A P
 661 GTGAGCGCCC GCCGCAGCGA CAGCGTGATC CTGAACGTGC TGTACGGCCC CGACGCCCCC
 T I S P L N T S Y R S G E N L N L S C H
 721 ACCATCAGCC CCCTGAACAC CAGCTACCGC AGCGGCGAGA ACCTGAACCT GAGCTGCCAC
 A A S N P P A Q Y S W F V N G T F Q Q S
 781 GCCGCCAGCA ACCCCCCCGC CCAGTACAGC TGGTTCGTGA ACGGCACCTT CCAGCAGAGC
 T Q E L F I P N I T V N N S G S Y T T C Q
 841 ACCCAGGAGC TGTTTCATCCC CAACATCACC GTGAACAACA GCGGCAGCTA CACCTGCCAG
 A H N S D T G L N R T T V T T I T V Y A
 901 GCCACAACA GCGACACCGG CCTGAACCGC ACCACCGTGA CCACCATCAC CGTGTACGCC
 E P P K P F I T S N N S N P V E D E D A
 961 GAGCCCCCA AGCCCTTCAT CACCAGCAAC AACAGCAACC CCGTGGAGGA CGAGGACGCC
 V A L T C E P E I Q N T T Y L W W V N N
 1021 GTGGCCCTGA CCTGCGAGCC CGAGATCCAG AACACCACCT ACCTGTGGTG GGTGAACAAC
 Q S L P V S P R L Q L S N D N R T L T L
 1081 CAGAGCCTGC CCGTGAGCCC CCGCCTGCAG CTGAGCAACG ACAACCGCAC CCTGACCCTG
 L S V T R N D V G P Y E C G I Q N E L S
 1141 CTGAGCGTGA CCCGCAACGA CGTGGGCCCC TACGAGTGCG GCATCCAGAA CGAGCTGAGC
 V D H S D P V I L N V L Y G P D D P T I

FIG.27A-1

50/62

1201 GTGGACCACA GCGACCCCGT GATCCTGAAC GTGCTGTACG GCCCCGACGA CCCACCATC
 S P S Y T Y Y R P G V N L S L S C H A A
 1261 AGCCCCAGCT ACACCTACTA CCGCCCCGGC GTGAACCTGA GCCTGAGCTG CCACGCCGCC
 S N P P A Q Y S W L I D G N I Q Q H T Q
 1321 AGCAACCCCC CCGCCCAGTA CAGCTGGCTG ATCGACGGCA ACATCCAGCA GCACACCCAG
 E L F I S N I T E K N S G L Y T C Q A N
 1381 GAGCTGTTCA TCAGCAACAT CACCGAGAAG AACAGCGGCC TGTACACCTG CCAGGCCAAC
 N S A S G H S R T T V K T I T V S A E L
 1441 AACAGCGCCA GCGGCCACAG CCGCACCACC GTGAAGACCA TCACCGTGAG CGCCGAGCTG
 P K P S I S S N N S K P V E D K D A V A
 1501 CCCAAGCCCA GCATCAGCAG CAACAACAGC AAGCCCGTGG AGGACAAGGA CGCCGTGGCC
 F T C E P E A Q N T T Y L W W V N G Q S
 1561 TTCACCTGCG AGCCCGAGGC CCAGAACACC ACCTACCTGT GGTGGGTGAA CGGCCAGAGC
 L P V S P R L Q L S N G N R T L T L F N
 1621 CTGCCCCGTGA GCCCCCGCCT GCAGCTGAGC AACGGCAACC GCACCCTGAC CCTGTTCAAC
 V T R N D A R A Y V C G I Q N S V S A N
 1681 GTGACCCGCA ACGACGCCCG CGCCTACGTG TGCGGCATCC AGAACAGCGT GAGCGCCAAC
 R S D P V T L D V L Y G P D T P I I S P
 1741 CGCAGCGACC CCGTGACCCT GGACGTGCTG TACGGCCCCG ACACCCCAT CATCAGCCCC
 P D S S Y L S G A N L N L S C H S A S N
 1801 CCCGACAGCA GCTACCTGAG CGGCGCCAAC CTGAACCTGA GCTGCCACAG CGCCAGCAAC
 P S P Q Y S W R I N G I P Q Q H T Q V L
 1861 CCCAGCCCCC AGTACAGCTG GCGCATCAAC GGCATCCCCC AGCAGCACAC CCAGGTGCTG
 F I A K I T P N N N G T Y A C F V S N L
 1921 TTCATCGCCA AGATCACCCC CAACAACAAC GGCACCTACG CCTGCTTCGT GAGCAACCTG
 A T G R N N S I V K S I T V S A S G T S
 1981 GCCACCGGCC GCAACAACAG CATCGTGAAG AGCATCACCG TGAGCGCCAG CGGCACCTCT
 R K T H T C P P C P A P E L L G G P S V
 2041 AGAAAGACCC ACACCTGCCC CCCTTGCCCT GCCCCTGAGC TGCTGGGCGG ACCCAGCGTG
 F L F P P K P K D T L M I S R T P E V T
 2101 TTCCTGTTCC CCCCCAAGCC TAAGGACACC CTCATGATCA GCAGAACCCC CGAGGTGACC
 C V V V D V S H E D P E V K F N W Y V D
 2161 TGCGTGGTGG TGGACGTGAG CCACGAGGAT CCCGAGGTGA AGTTCAACTG GTACGTGGAC
 G V E V H N A K T K P R E E Q Y N S T Y
 2221 GCGGTGGAGG TGCACAATGC CAAGACCAAG CCCAGAGAGG AGCAGTACAA CAGCACCTAC
 R V V S V L T V L H Q D W L N G K E Y K
 2281 AGAGTGGTGA GCGTGCTCAC CGTGCTGCAC CAGGATTGGC TGAACGGCAA GGAGTACAAG
 C K V S N K A L P A P I E K T I S K A K
 2341 TGCAAGGTGA GCAACAAGGC CCTGCCTGCC CCCATCGAGA AAACCATCAG CAAGGCCAAG
 G Q P R E P Q V Y T L P P S R D E L T K

FIG.27A-2

51/62

2401 GGCCAGCCCA GAGAGCCCA GGTGTACACC CTGCCCCCTA GCAGAGATGA GTTGACCAAG
N Q V S L T C L V K G F Y P S D I A V E
2461 AACCAGGTGA GCCTCACATG CCTGGTGAAG GGCTTCTACC CCAGCGACAT CGCCGTGGAG
W E S N G Q P E N N Y K T T P P V L D S
2521 TGGGAGAGCA ACGGCCAGCC CGAGAACAAC TACAAGACCA CCCCCCTGT GCTGGACAGC
D G S F F L Y S K L T V D K S R W Q Q G
2581 GATGGCAGCT TCTTCCTGTA CAGCAAGCTC ACCGTGGACA AGAGCAGATG GCAGCAGGGC
N V F S C S V M H E A L H N H Y T Q K S
2641 AACGTGTTCA GCTGCAGCGT GATGCACGAG GCCCTGCACA ATCACTACAC CCAGAAGAGC
L S L S P G K * (SEQ ID NO:46)
2701 CTGAGCCTGA GCCCCGGCAA GTAA (SEQ ID NO:25)

FIG.27A-3

52/62

```

atggagtctc cctcggcccc tccccacaga tgggtgcatcc cctggcagag gctcctgctc 60
acagcctcac ttctaacctt ctggaacccg cccaccactg ccaagctcac tattgaatcc 120
acgccgttca atgtcgcaga ggggaaggag gtgcttctac ttgtccacaa tctgccccag 180
catctttttg gctacagctg gtacaaagggt gaaagagtgg atggcaaccg tcaaattata 240
ggatatgtaa taggaactca acaagctacc ccaggggccc catacagtgg tcgagagata 300
atatacccca atgcatccct gctgatccag aacatcatcc agaatgacac aggattctac 360
accctacacg tcataaagtc agatcttgtg aatgaagaag caactggcca gttccgggta 420
taccgggagc tgcccaagcc ctccatctcc agcaacaact ccaaaccctg ggaggacaag 480
gatgctgtgg ccttcacctg tgaacctgag actcaggacg caacctacct gtggtgggta 540
aacaatcaga gcctcccggg cagtcccagg ctgcagctgt ccaatggcaa caggaccctc 600
actctattca atgtcacaag aaatgacaca gcaagctaca aatgtgaaac ccagaaccca 660
gtgagtgcc a ggcgcagtga ttcagtcac ctgaatgtcc tctatggccc ggatgcccc 720
accatttccc ctctaaacac atcttacaga tcaggggaaa atctgaacct ctctgcccac 780
gcagcctcta acccacctgc acagtactct tggtttgtca atgggacttt ccagcaatcc 840
accaagagc tctttatccc caacatcact gtgaataata gtggatccta tacgtgccaa 900
gcccataact cagacactgg cctcaatagg accacagtca cgacgatcac agtctatgca 960
gagccacca aacccttcat caccagcaac aactccaacc ccgtggagga tgaggatgct 1020
gtagccttaa cctgtgaacc tgagattcag aacacaacct acctgtggtg ggtaaataat 1080
cagagcctcc cggtcagtcc caggctgcag ctgtccaatg acaacaggac cctcactcta 1140
ctcagtgtca caaggaatga ttaggaccc tatgagtgtg gaatccagaa cgaattaagt 1200
gttgaccaca gcgaccagc catcctgaat gtcctctatg gccagacga cccaccatt 1260
tccccctcat acacctatta ccgtccaggg gtgaacctca gcctctcctg ccatgcagcc 1320
tctaaccac ctgcacagta ttcttggtg attgatggga acatccagca acacacaaa 1380
gagctcttta tctccaacat cactgagaag aacagcggac tctataacct ccaggccaat 1440
aactcagcca gtggccacag caggactaca gtcaagacaa tcacagtctc tgcggagctg 1500
cccaagccct ccatctccag caacaactcc aaaccctgg aggacaagga tgctgtggcc 1560
ttcacctgtg aacctgaggc tcagaacaca acctacctgt ggtgggtaaa tggctcagagc 1620
ctcccagtc a gtcccaggct gcagctgtcc aatggcaaca ggaccctcac tctattcaat 1680
gtcacaagaa atgacgcaag agcctatgta tgtggaatcc agaactcagt gagtgcaaac 1740
cgcagtgacc cagtcaccct ggatgtcctc tatgggccgg acaccccat catttcccc 1800
ccagactcgt cttacctttc gggagcgaac ctcaacctct cctgccactc ggcctctaac 1860
ccatccccgc agtattcttg gcgtatcaat gggataccgc agcaacacac acaagttctc 1920
tttatcgcca aaatcacgcc aaataataac gggacctatg cctgttttgt ctctaacttg 1980
gctactggcc gcaataattc catagtcaag agcatcacag tctctgcac tggaact(SEQ ID NO: 22)

```

FIG. 28A

53/62

1 MESPSAPP^hHR WCIPWQRLLL TASLLTFWNP PTTAKLTIES TPFNVAEGKE
51 VLLL^hVHNL^hPQ HLF^hGYSWYKG ERVDG^hNRQII GYVIGTQQAT PGPAYSGREI
101 IYPNASLLIQ NIIQNDTGFY TLHVIKSDLV NEEATGQFRV YPELPKPSIS
151 SNNSKPVEDK DAVAFTCEPE TQDATYLWWV NNQSLPVSPR LQLSNGNRTL
201 TLFNVTRNDT ASYKCETQNP VSARRSDSVI LNVLYGPDAP TISPLNTSYR
251 SGENLNLSCH AASNPPAQYS W^hFVNGTFQQS TQELFIPNIT VNNSGSYTCQ
301 AHNSDTGLNR TTVTTITVYA EPPKPFITSN NSNPVEDEDA VALTCEPEIQ
351 NTTYLWWVNN QSLPVSPRLQ LSNDNRTLTL LSVTRNDVGP YECGIQNELS
401 VDHSDPVILN VLYGPDDPTI SPSYTYRPG VNLSLSCHAA SNPPAQYSWL
451 IDGNIQQHTQ ELFISNITEK NSGLYTCQAN NSASGHSRTT VKTITVSAEL
501 PKPSISSNNS KPVEDKDAVA FTCEPEAQNT TYLWWVNGQS LPVSPRLQLS
551 NGNRTLTLFN VTRNDARAYV CGIQNSVSAN RSDPVTLDVL YGPDTPIIISP
601 PDSSYLSGAN LNLSCHSASN PSPQYSWRIN GIPQQHTQVL FIAKITPNNN
651 GTYACFVSNL ATGRNNSIVK SITVSASGT (SEQ ID NO:23)

FIG.28B

54/62

D S T P I P F S Y S K N L D C W V D N E
1 GATTCAACAC CAATTCCATT TTCTTATTCT AAAAATCTGG ATTGTTGGGT TGATAATGAA
E D I D V I L K K S T I L N L D I N N D
61 GAAGATATAG ATGTTATATT AAAAAAGAGT ACAATTTTAA ATTTAGATAT TAATAATGAT
I I S D I S G F N S S V I T Y P D A Q L
121 ATTATATCAG ATATATCTGG GTTTAATTCA TCTGTAATAA CATATCCAGA TGCTCAATTG
V P G I N G K A I H L V N N E S S E V I
181 GTGCCCGGAA TAAATGGCAA AGCAATACAT TTAGTAAACA ATGAATCTTC TGAAGTTATA
V H K A M D I E Y N D M F N N F T V S F
241 GTGCATAAAG CTATGGATAT TGAATATAAT GATATGTTTA ATAATTTTAC CGTTAGCTTT
W L R V P K V S A S H L E Q Y G T N E Y
301 TGGTTGAGGG TTCCTAAAGT ATCTGCTAGT CATTTAGAAC AATATGGCAC AAATGAGTAT
S I I S S M K K H S L S I G S G W S V S
361 TCAATAATTA GCTCTATGAA AAAACATAGT CTATCAATAG GATCTGGTTG GAGTGTATCA
L K G N N L I W T L K D S A G E V R Q I
421 CTTAAAGGTA ATAACTTAAT ATGGACTTTA AAAGATTCCG CGGGAGAAGT TAGACAAATA
T F R D L P D K F N A Y L A N K W V F I
481 ACTTTTAGGG ATTTACCTGA TAAATTTAAT GCTTATTTAG CAAATAAATG GGTTTTTATA
T I T N D R L S S A N L Y I N G V L M G
541 ACTATTACTA ATGATAGATT ATCTTCTGCT AATTTGTATA TAAATGGAGT ACTTATGGGA
S A E I T G L G A I R E D N N I T L K L
601 AGTGCAGAAA TTA CTGGTTT AGGAGCTATT AGAGAGGATA ATAATATAAC ATTAAACTA
D R C N N N N Q Y V S I D K F R I F C K
661 GATAGATGTA ATAATAATAA TCAATACGTT TCTATTGATA AATTTAGGAT ATTTTGCAAA
A L N P K E I E K L Y T S Y L S I T F L
721 GCATTAAATC CAAAAGAGAT TGAAAAATTA TACACAAGTT ATTTATCTAT AACCTTTTTA
R D F W G N P L R Y D T D R *(SEQ ID NO:48)-
781 AGAGACTTCT GGGGAAACCC TTTACGATAT GATACAGATA GGTAG (SEQ ID NO:47)-

FIG.29

55/62

1 ATGGAGTCTC CCTCGGCCCC TCCCCACAGA TGGTGCATCC CCTGGCAGAG GCTCCTGCTC
61 ACAGCCTCAC TTCTAACCTT CTGGAACCCG CCCACCACTG CCAAGCTCAC TATTGAATCC
121 ACGCCGTTCA ATGTCGCAGA GGGGAAGGAG GTGCTTCTAC TTGTCCACAA TCTGCCCCAG
181 CATCTTTTGT GCTACAGCTG GTACAAAGGT GAAAGAGTGG ATGGCAACCG TCAAATTATA
241 GGATATGTAA TAGGAACTCA ACAAGCTACC CCAGGGCCCCG CATAAGTGG TCGAGAGATA
301 ATATACCCCA ATGCATCCCT GCTGATCCAG AACATCATCC AGAATGACAC AGGATTCTAC
361 ACCCTACACG TCATAAAGTC AGATCTTGTG AATGAAGAAG CAACTGGCCA GTTCCGGGTA
421 TACCCGGAGC TGCCCAAGCC CTCCATCTCC AGCAACAACCT CCAAACCCGT GGAGGACAAG
481 GATGCTGTGG CCTTCACCTG TGAACCTGAG ACTCAGGACG CAACCTACCT GTGGTGGGTA
541 AACAATCAGA GCCTCCCGGT CAGTCCCAGG CTGCAGCTGT CCAATGGCAA CAGGACCCTC
601 ACTCTATTCA ATGTCACAAG AAATGACACA GCAAGCTACA AATGTGAAAC CCAGAACCCA
661 GTGAGTGCCA GCGGCAGTGA TTCAGTCATC CTGAATGTCC TCTATGGCCC GGATGCCCCC
721 ACCATTTCCC CTCTAAACAC ATCTTACAGA TCAGGGGAAA ATCTGAACCT CTCCTGCCAC
781 GCAGCCTCTA ACCCACCTGC ACAGTACTCT TGGTTTGTCA ATGGGACTTT CCAGCAATCC
841 ACCCAAGAGC TCTTTATCCC CAACATCACT GTGAATAATA GTGGATCCTA TACGTGCCAA
901 GCCCATAACT CAGACACTGG CCTCAATAGG ACCACAGTCA CGACGATCAC AGTCTATGCA
961 GAGCCACCCA AACCTTCAT CACCAGCAAC AACTCCAACC CCGTGGAGGA TGAGGATGCT
1021 GTAGCCTTAA CCTGTGAACC TGAGATTCAG AACACAACCT ACCTGTGGTG GGTAATAAT
1081 CAGAGCCTCC CGGTCAGTCC CAGGCTGCAG CTGTCCAATG ACAACAGGAC CCTCACTCTA
1141 CTCAGTGTCA CAAGGAATGA TGTAGGACCC TATGAGTGTG GAATCCAGAA CGAATTAAGT
1201 GTTGACCACA GCGACCCAGT CATCCTGAAT GTCCTCTATG GCCCAGACGA CCCCACCATT
1261 TCCCCCTCAT ACACCTATTA CCGTCCAGGG GTGAACCTCA GCCTCTCCTG CCATGCAGCC
1321 TCTAACCAC CTGCACAGTA TTCTTGGCTG ATTGATGGGA ACATCCAGCA ACACACACAA
1381 GAGCTCTTTA TCTCCAACAT CACTGAGAAG AACAGCGGAC TCTATACCTG CCAGGCCAAT
1441 AACTCAGCCA GTGGCCACAG CAGGACTACA GTCAAGACAA TCACAGTCTC TGC GGAGCTG
1501 CCCAAGCCCT CCATCTCCAG CAACAACCTC AAACCCGTGG AGGACAAGGA TGCTGTGGCC
1561 TTCACCTGTG AACCTGAGGC TCAGAACACA ACCTACCTGT GGTGGGTAAA TGGTCAGAGC
1621 CTCCCAGTCA GTCCCAGGCT GCAGCTGTCC AATGGCAACA GGACCCTCAC TCTATTCAAT
1681 GTCACAAGAA ATGACGCAAG AGCCTATGTA TGTGGAATCC AGAACTCAGT GAGTGCAAAC
1741 CGCAGTGACC CAGTCACCCT GGATGTCCTC TATGGGCCGG ACACCCCAT CATTTCCCCC
1801 CCAGACTCGT CTTACCTTTC GGGAGCGAAC CTCAACCTCT CCTGCCACTC GGCCTCTAAC
1861 CCATCCCCGC AGTATTGTTG GCGTATCAAT GGGATACCGC AGCAACACAC ACAAGTTCTC
1921 TTTATCGCCA AAATCACGCC AAATAATAAC GGGACCTATG CCTGTTTTGT CTCTAACTTG
1981 GCTACTGGCC GCAATAATTC CATAGTCAAG AGCATCACAG TCTCTGCATC TGGAACCTA
2041 GATTCAACAC CAATTCCATT TTCTTATTCT AAAAATCTGG ATTGTTGGGT TGATAATGAA
2101 GAAGATATAG ATGTTATATT AAAAAAGAGT ACAATTTTAA ATTTAGATAT TAATAATGAT
2161 ATTATATCAG ATATATCTGG GTTTAATTCA TCTGTAATAA CATATCCAGA TGCTCAATTG
2221 GTGCCCCGAA TAAATGGCAA AGCAATACAT TTAGTAAACA ATGAATCTTC TGAAGTTATA
2281 GTGCATAAAG CTATGGATAT TGAATATAAT GATATGTTTA ATAATTTTAC CGTTAGCTTT
2341 TGGTTGAGGG TTCCTAAAGT ATCTGCTAGT CATTTAGAAC AATATGGCAC AAATGAGTAT
2401 TCAATAATTA GCTCTATGAA AAAACATAGT CTATCAATAG GATCTGGTTG GAGTGTATCA

FIG.30A-1

56/62

2461 CTTAAAGGTA ATAACCTTAAT ATGGACTTTA AAAGATTCCG CGGGAGAAGT TAGACAAATA
2521 ACTTTTAGGG ATTTACCTGA TAAATTTAAT GCTTATTTAG CAAATAAATG GGTTTTTATA
2581 ACTATTACTA ATGATAGATT ATCTTCTGCT AATTTGTATA TAAATGGAGT ACTTATGGGA
2641 AGTGCAGAAA TTACTGGTTT AGGAGCTATT AGAGAGGATA ATAATATAAC ATTAAAACTA
2701 GATAGATGTA ATAATAATAA TCAATACGTT TCTATTGATA AATTTAGGAT ATTTTGCAAA
2761 GCATTAAATC CAAAAGAGAT TGAAAAATTA TACACAAGTT ATTTATCTAT AACCTTTTTA
2821 AGAGACTTCT GGGGAAACCC TTTACGATAT GATATAG (SEQ ID NO:49)

FIG.30A-2

57/62

1	ATGGGCAGCC	CCAGCGCCCC	CCTGCACCGC	TGGTGCATCC	CCTGGCAGAC
	CCTGCTGCTG	ACCGCCAGCC	TGCTGACCTT	CTGGAACCCC	CCCACCACCG
101	CCCAGCTGAC	CATCGAGAGC	CGCCCCTTCA	ACGTGGCCGA	GGGCAAGGAG
	GTGCTGCTGC	TGGCCCACAA	CGTGAGCCAG	AACCTGTTCG	GCTACATCTG
201	GTACAAGGGC	GAGCGCGTGG	ACGCCAGCCG	CCGCATCGGC	AGCTGCGTGA
	TCCGCACCCA	GCAGATCACC	CCCGGCCCCG	CCCACAGCGG	CCGCGAGACC
301	ATCGACTTCA	ACGCCAGCCT	GCTGATCCAC	AACGTGACCC	AGAGCGACAC
	CGGCAGCTAC	ACCATCCAGG	TGATCAAGGA	GGACCTGGTG	AACGAGGAGG
401	CCACCGGCCA	GTTCCGCGTG	TACCCCGAGC	TGCCCAAGCC	CTACATCAGC
	AGCAACAACA	GCAACCCCGT	GGAGGACAAG	GACGCCGTGG	CCCTGACCTG
501	CGAGCCCGAG	ACCCAGGACA	CCACCTACCT	GTGGTGGGTG	AACAACCAGA
	GCCTGCCCGT	GAGCCCCCGC	CTGGAGCTGA	GCAGCGACAA	CCGCACCCTG
601	ACCGTGTTCA	ACATCCCCCG	CAACGACACC	ACCAGCTACA	AGTGCGAGAC
	CCAGAACCCC	GTGAGCGTGC	GCCGCAGCGA	CCCCGTGACC	CTGAACGTGC
701	TGTACGGCCC	CGACGCCCCC	ACCATCAGCC	CCCTGAACAC	CCCCTACCGC
	GCCGGCGAGA	ACCTGAACCT	GACCTGCCAC	GCCGCCAGCA	ACCCACCGC
801	CCAGTACTTC	TGGTTCGTGA	ACGGCACCTT	CCAGCAGAGC	ACCCAGGAGC
	TGTTCATCCC	CAACATCACC	GTGAACAACA	GCGGCAGCTA	CATGTGCCAG
901	GCCCACAACA	GCGCCACCGG	CCTGAACCGC	ACCACCGTGA	CCGCCATCAC
	CGTGTAAGCC	GAGCTGCCCA	AGCCCTACAT	CACCAGCAAC	AACAGCAACC
1001	CCATCGAGGA	CAAGGACGCC	GTGACCCTGA	CCTGCGAGCC	CGAGACCCAG
	GACACCACCT	ACCTGTGGTG	GGTGAACAAC	CAGAGCCTGA	GCGTGAGCAG
1101	CCGCCTGGAG	CTGAGCAACG	ACAACCGCAC	CCTGACCGTG	TTCAACATCC
	CCCGCAACGA	CACCACCTTC	TACGAGTGCG	AGACCCAGAA	CCCCGTGAGC
1201	GTGCGCCGCA	GCGACCCCGT	GACCCTGAAC	GTGCTGTACG	GCCCCGACGC
	CCCCACCATC	AGCCCCCTGA	ACACCCCTA	CCGCGCCGGC	GAGAACCTGA
1301	ACCTGAGCTG	CCACGCCGCC	AGCAACCCCG	CCGCCCAGTA	CAGCTGGTTC
	GTGAACGGCA	CCTTCCAGCA	GAGCACCCAG	GAGCTGTTCA	TCCCCAACAT
1401	CACCGTGAAC	AACAGCGGCA	GCTACATGTG	CCAGGCCCAC	AACAGCGCCA
	CCGGCCTGAA	CCGCACCACC	GTGACCGCCA	TCACCGTGTA	CGTGAGCTG
1501	CCCAAGCCCT	ACATCAGCAG	CAACAACAGC	AACCCCATCG	AGGACAAGGA
	CGCCGTGACC	CTGACCTGCG	AGCCCGTGGC	CGAGAACACC	ACCTACCTGT
1601	GGTGGGTGAA	CAACCAGAGC	CTGAGCGTGA	GCCCCCGCCT	GCAGCTGAGC
	AACGGCAACC	GCATCCTGAC	CCTGCTGAGC	GTGACCCGCA	ACGACACCGG
1701	CCCCTACGAG	TGCGGCATCC	AGAACAGCGA	GAGCGCCAAG	CGCAGCGACC
	CCGTGACCCT	GAACGTGACC	TACGGCCCCG	ACACCCCAT	CATCAGCCCC
1801	CCCGACCTGA	GCTACCGCAG	CGGCGCCAAC	CTGAACCTGA	GCTGCCACAG
	CGACAGCAAC	CCCAGCCCCC	AGTACAGCTG	GCTGATCAAC	GGCACCCCTGC
1901	GCCAGCACAC	CCAGGTGCTG	TTCATCAGCA	AGATCACCAG	CAACAACAGC
	GGCGCCTACG	CCTGCTTCGT	GAGCAACCTG	GCCACCGGCC	GCAACAACAG
2001	CATCGTGAAG	AACATCAGCG	TGAGCAGCGG	CGACAGCTCT	AGAAGCACCC
	CCATCCCATT	CAGCTACAGC	AAGAACCTGG	ACTGCTGGGT	GGACAACGAG

FIG.31A-1

58/62

2101 GAGGACATCG ACGTGATCCT GAAGAAGAGC ACCATCCTGA ACCTGGACAT
CAACAACGAC ATCATCAGCG ACATCAGCGG CTTCAACAGC AGCGTGATCA
2201 CCTACCCCGA CGCCCAGCTG GTGCCCGGCA TCAACGGCAA GGCCATCCAC
CTGGTGAACA ACGAGAGCAG CGAGGTGATC GTGCACAAGG CCATGGACAT
2301 CGAGTACAAC GACATGTTCA ACAACTTCAC CGTGAGCTTC TGGCTGAGAG
TGCCTAAGGT GAGCGCCAGC CACCTGGAGC AGTACGGCAC CAACGAGTAC
2401 AGCATCATCA GCAGCATGAA GAAGCACAGC CTGAGCATCG GCAGCGGCTG
GAGCGTGAGC CTGAAGGGCA ACAACCTCAT CTGGACCCTG AAGGATAGCG
2501 CCGGAGAGGT GAGACAGATC ACCTTCAGAG ACCTGCCCCGA CAAGTTCAAT
GCCTACCTGG CCAACAAGTG GGTGTTTCATC ACCATCACCA ACGACAGACT
2601 GAGCAGCGCC AACCTGTACA TCAACGGCGT GCTCATGGGC AGCGCCGAGA
TCACCGGCCT GGGCGCCATC AGAGAGGACA ACAACATCAC CCTGAAGCTG
2701 GACAGATGCA ACAACAACAA CCAGTACGTG AGCATCGACA AGTTCCGGAT
CTTCTGCAAG GCCCTGAACC CCAAGGAGAT CGAGAAGCTG TACACCAGCT
2801 ACCTGAGCAT CACCTTCCTG AGAGACTTCT GGGGCAACCC CCTGAGATAC
GACACCTAG (SEQ ID NO:50)

FIG.31A-2

59/62

1 MGSPSAPLHR WCIPWQTLLE TASLLTFWNP PTTAQLTIES RPFNVAEGKE
51 VLLLAHNVSQ NLFQYIWKY ERVDASRRIG SCVIRTQQIT PGPAHSGRET
101 IDFNASLLIH NVTQSDTGSY TIQVIKEDLV NEEATGQFRV YPELKPYPIS
151 SNNSNPVEDK DAVALTCEPE TQDTTYLWWV NNQSLPVSPR LELSSDNRTL
201 TVFNIPRNDT TSYKCETQNP VSVRRSDPVT LNVLYGPDAP TISPLNTPYR
251 AGENLNLTCH AASNPTAQYF WFNVTGTFQQS TQELFIPNIT VNNSGSGYMCQ
301 AHNSATGLNR TTVTAITVYA ELPKPYITSN NSNPIEDKDA VTLTCEPETQ
351 DTTYLWWVNN QSLSVSSRLE LSNDNRTLTV FNIPRNDTTF YECETQNPVS
401 VRRSDPVTLN VLYGPDAPTI SPLNTPYRAG ENLNLSCHAA SNPAAQYSWF
451 VNGTFQQSTQ ELFIPNITVN NSGSGYMCQAH NSATGLNRTT VTAITVYVEL
501 PKPYISSNNS NPIEDKDAVT LTCEPVAENT TYLWWVNNQS LSVSPRLQLS
551 NGNRILTLLS VTRNDTGPEY CGIQNSESAK RSDPVTLNVT YGPDTPIIISP
601 PDLSYRSGAN LNLSCHSDSN PSPQYSWLIN GTLRQHTQVL FISKITSNNS
651 GAYACFVSNL ATGRNNSIVK NISVSSGDSS RSTPIPFSSYS KNLDCWVDNE
701 EDIDVILKKS TILNLDINND IISDISGFNS SVITYPDAQL VPGINGKAIH
751 LVNNESESEVI VHKAMDIEYN DMFNFTVSF WLRVPKVSAS HLEQYGTNEY
801 SIISSMKKHS LSIGSGWSVS LKGNNLIWTL KDSAGEVRQI TFRDLPDKFN
851 AYLANKWVFI TITNDRLLSA NLYINGVLMG SAEITGLGAI REDNNITLKL
901 DRCNNNNQYV SIDKFRIFCK ALNPKEIEKL YTSYLSITFL RDFWGNPLRY
951 DT* (SEQ ID NO:51)

FIG.31B

60/62

1	ATGGGCAGCC	CCAGCGCCCC	CCTGCACCGC	TGGTGCATCC	CCTGGCAGAC
	CCTGCTGCTG	ACCGCCAGCC	TGCTGACCTT	CTGGAACCCC	CCCACCACCG
101	CCCAGCTGAC	CATCGAGAGC	CGCCCCTTCA	ACGTGGCCGA	GGGCAAGGAG
	GTGCTGCTGC	TGGCCCACAA	CGTGAGCCAG	AACCTGTTCG	GCTACATCTG
201	GTACAAGGGC	GAGCGCGTGG	ACGCCAGCCG	CCGCATCGGC	AGCTGCGTGA
	TCCGCACCCA	GCAGATCACC	CCCGGCCCCG	CCCACAGCGG	CCGCGAGACC
301	ATCGACTTCA	ACGCCAGCCT	GCTGATCCAC	AACGTGACCC	AGAGCGACAC
	CGGCAGCTAC	ACCATCCAGG	TGATCAAGGA	GGACCTGGTG	AACGAGGAGG
401	CCACCGGCCA	GTTCCGCGTG	TACCCCGAGC	TGCCCAAGCC	CTACATCAGC
	AGCAACAACA	GCAACCCCGT	GGAGGACAAG	GACGCCGTGG	CCCTGACCTG
501	CGAGCCCGAG	ACCCAGGACA	CCACCTACCT	GTGGTGGGTG	AACAACCAGA
	GCCTGCCCGT	GAGCCCCCGC	CTGGAGCTGA	GCAGCGACAA	CCGCACCCTG
601	ACCGTGTTCA	ACATCCCCCG	CAACGACACC	ACCAGCTACA	AGTGCGAGAC
	CCAGAACCCC	GTGAGCGTGC	GCCGCAGCGA	CCCCGTGACC	CTGAACGTGC
701	TGTACGGCCC	CGACGCCCCC	ACCATCAGCC	CCCTGAACAC	CCCCTACCGC
	GCCGGCGAGA	ACCTGAACCT	GACCTGCCAC	GCCGCCAGCA	ACCCACCGC
801	CCAGTACTTC	TGGTTCGTGA	ACGGCACCTT	CCAGCAGAGC	ACCCAGGAGC
	TGTTTCATCC	CAACATCACC	GTGAACAACA	GCGGCAGCTA	CATGTGCCAG
901	GCCCACAACA	GCGCCACCGG	CCTGAACCGC	ACCACCGTGA	CCGCCATCAC
	CGTGTACGCC	GAGCTGCCCA	AGCCCTACAT	CACCAGCAAC	AACAGCAACC
1001	CCATCGAGGA	CAAGGACGCC	GTGACCCTGA	CCTGCGAGCC	CGAGACCCAG
	GACACCACCT	ACCTGTGGTG	GGTGAACAAC	CAGAGCCTGA	GCGTGAGCAG
1101	CCGCCTGGAG	CTGAGCAACG	ACAACCGCAC	CCTGACCGTG	TTCAACATCC
	CCCGCAACGA	CACCACCTTC	TACGAGTGCG	AGACCCAGAA	CCCCGTGAGC
1201	GTGCGCCGCA	GCGACCCCGT	GACCCTGAAC	GTGCTGTACG	GCCCCGACGC
	CCCCACCATC	AGCCCCCTGA	ACACCCCTA	CCGCGCCGGC	GAGAACCTGA
1301	ACCTGAGCTG	CCACGCCGCC	AGCAACCCCG	CCGCCCAGTA	CAGCTGGTTC
	GTGAACGGCA	CCTTCCAGCA	GAGCACCCAG	GAGCTGTTCA	TCCCCAACAT
1401	CACCGTGAAC	AACAGCGGCA	GCTACATGTG	CCAGGCCCAC	AACAGCGCCA
	CCGGCCTGAA	CCGCACCACC	GTGACCGCCA	TCACCGTGTA	GGTGGAGCTG
1501	CCCAAGCCCT	ACATCAGCAG	CAACAACAGC	AACCCCATCG	AGGAAEAGGA
	CGCCGTGACC	CTGACCTGCG	AGCCCGTGGC	CGAGAACACC	ACCTACCTGT
1601	GGTGGGTGAA	CAACCAGAGC	CTGAGCGTGA	GCCCCCGCCT	GCAGCTGAGC
	AACGGCAACC	GCATCCTGAC	CCTGCTGAGC	GTGACCCGCA	ACGACACCGG
1701	CCCCTACGAG	TGCGGCATCC	AGAACAGCGA	GAGCGCCAAG	CGCAGCGACC
	CCGTGACCCT	GAACGTGACC	TACGGCCCCG	ACACCCCAT	CATCAGCCCC
1801	CCCGACCTGA	GCTACCGCAG	CGGCGCCAAC	CTGAACCTGA	GCTGECACAG
	CGACAGCAAC	CCCAGCCCCC	AGTACAGCTG	GCTGATCAAC	GGCACCCTGC
1901	GCCAGCACAC	CCAGGTGCTG	TTCATCAGCA	AGATCACCAG	CAACAACAGC
	GGCGCCTACG	CCTGCTTCGT	GAGCAACCTG	GCCACCGGCC	GCAACAACAG
2001	CATCGTGAAG	AACATCAGCG	TGAGCAGCGG	CGACAGCTCT	AGAACCCTC
	AGAACATCAC	CGATCTGTGC	GCCGAGTACC	ACAACACCCA	GATCTACACC

FIG.32A-1

61/62

2101 CTGAACGACA AGATCTTCAG CTACACCGAG AGCCTGGCCG GCAAGAGAGA
GATGGCCATC ATCACCTTCA AGAACGGCGC CATCTTCCAG GTGGAGGTGC
2201 CCGGCAGCCA GCACATCGAC AGCCAGAAGA AGGCCATCGA GCGGATGAAG
GACACCCTGC GGATCGCCTA CCTCACCGAG GCCAAGGTGG AGAAGCTGTG
2301 CGTGTGGAAC AACAAGACCC CTCACGCCAT CGCCGCCATC AGCATGGCCA
ATTGATAAG (SEQ ID NO:52)

FIG.32A-2

62/62

1	MGSPSAPLHR	WCIPWQTL	LL	TASLLTFWNP	PTTAQLTIES	RPFNVAEGKE
51	VLLLAHNVSQ	NLFGYIWYKG	ERVDASRRIG	SCVIRTQQIT	PGPAHSGRET	
101	IDFNASLLIH	NVTQSDTG	SY	TIQVIKEDLV	NEEATGQFRV	YPELPKPYIS
151	SNNSNPVEDK	DAVALTCEPE	TQDTTYLWWV	NNQSLPVSPR	LELSSDNRTL	
201	TVFNIPRNDT	TSYKCETQNP	VSVRRSDPVT	LNVL	YGPDAP	TISPLNTPYR
251	AGENLNLTCH	AASNPTAQYF	WVFN	GTFFQQS	TQELFIPNIT	VNNSG
301	AHNSATGLNR	TTVTAITVYA	ELPKPYITSN	NSNPIEDKDA	VTLTCEPETQ	
351	DTTYLWWVNN	QSLSVSSRLE	LSNDNRTLTV	FNIPRNDTTF	YECETQNPVS	
401	VRRSDPVTLN	VLYGPDAPTI	SPLNTPYRAG	ENLNLSCHAA	SNPAAQYSWF	
451	VNGTFQQSTQ	ELFIPNITVN	NSG	SYMCAH	NSATGLNRTT	VT
501	PKPYISSNNS	NPIEDKDAVT	LTCEPVAENT	TYLWWVNNQS	LSVSPRLQLS	
551	NGNRILTLLS	VTRNDTG	PYE	CGIQNSESAK	RSDPVTLNVT	YGPDTPIISP
601	PDLSYRSGAN	LNLSCHSDSN	PSPQYSWLIN	GT	LRQHTQVL	FISKITSNNS
651	GAYACFVSNL	ATGRNNSIVK	NISVSSGDSS	RTPQNITDLC	AEYHNTQIYT	
701	LNDKIFSYTE	SLAGKREMAI	ITFKNGAIFQ	VEVPGSQHID	SQKKAIERMK	
751	DTLRIAYLTE	AKVEKLCVWN	NKTPHAIAAI	SMAN**	(SEQ ID NO:53)	

FIG.32B