AU9339788

(12) PATENT ABRIDGMENT (11) Document No. AU-B-39788/93
(19) AUSTRALIAN PATENT OFFICE (10) Acceptance No. 686999

(54) Title
PROCESS FOR PREPARING MODIFIED PROTEINS

International Patent Classification(s)
(51)[5] **C07K 003/08** **C07K 011/00** **C12N 009/50**

(21) Application No. : **39788/93** (22) Application Date : **26.03.93**

(87) PCT Publication Number : **WO93/20098**

(30) Priority Data

(31) Number	(32) Date	(33) Country
865368	**07.04.92**	**US UNITED STATES OF AMERICA**

(43) Publication Date : **08.11.93**

(44) Publication Date of Accepted Application : **19.02.98**

(71) Applicant(s)
THE SCRIPPS RESEARCH INSTITUTE

(72) Inventor(s)
STEPHEN B.H. KENT; MARTINA SCHNOLZER-RACKWITZ

(74) Attorney or Agent
GRIFFITH HACK , GPO Box 4164, SYDNEY NSW 2001

(56) Prior Art Documents
US 5066716
AU 21798/92 C07K 3/00;1/00

(57) Claim

1. A process for preparing modified proteins which comprises the steps of:

a. sequentially coupling amino acids or amino acid analogs to a terminal amino acid or amino acid analog bound to a first resin support to form a first peptide segment-resin, the first peptide segment peptide comprising from about two to about one hundred amino acid residues;

b. covalently attaching a haloacyl moiety to the N-terminus of the peptide segment-resin to form a haloacylpeptide segment bound to the first resin support;

c. cleaving the haloacylpeptide segment from the first resin support;

d. sequentially coupling amino acids or amino acid analogs to a terminal amino acid or amino acid analog bound to a second resin-support through a sulfur, nitrogen or a seleniumcontaining bond to form a second peptide segment-resin, the second peptide segment comprising from about two to about one hundred amino acid residues, with a

.../2

proviso that the nitrogen containing bond is employed for binding the terminal amino acid or amino acid analog to the second resin support if and only if the terminal amino acid or amino acid analog is cysteine or an S-H containing analog of cysteine;

e. cleaving the second peptide segment-resin to form a second peptide segment having a thiol-, selenol-, or nitrogen containing group at the C-terminus thereof, with a proviso that the C-terminus of the second peptide-segment has a nitrogen-containing group if and only if the C-terminal amino acid or amino acid analog is cysteine or an S-H containing analog of cysteine; and

f. coupling the haloacylpeptide segment and the second peptide segment.

5. The process according to claim 1, wherein, in step (d), the sulfur-containing bond linking the second peptide segment and the second resin support is a thiol ester linkage and in step (e), the thiol-containing group of the second peptide segment is carbonylthiol.

20. A modified protein represented by the formula:

R-L-R'

wherein R and R' are the same or different and are each peptides or pseudopeptides having at least two amino acid residues; and L represents a thiol ester or selenol ester linkage.

OPI DATE 08/11/93 APPLN. ID 39788/93
AOJP DATE 13/01/94 PCT NUMBER PCT/US93/02846

AU9339788

(51) International Patent Classification 5 :		(11) International Publication Number:
C07K 3/04	A1	WO 93/20098
		(43) International Publication Date: 14 October 1993 (14.10.93)

(21) International Application Number: PCT/US93/02846

(22) International Filing Date: 26 March 1993 (26.03.93)

(30) Priority data:
07/865,368 7 April 1992 (07.04.92) US

(71) Applicant: THE SCRIPPS RESEARCH INSTITUTE [US/US]; 10666 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventors: KENT, Stephen, B., H. ; 2766 Costebelle Drive, La Jolla, Ca 92037 (US). SCHNOLZER-RACKWITZ, Martina ; 8533-C Villa La Jolla Drive, La Jolla, CA 92037 (US).

(74) Agents: FITTING, Thomas et al.; The Scripps Research Institute, 10666 North Torrey Pines Road, TPC-8, La Jolla, CA 92037 (US).

(81) Designated States: AU, CA, JP, KR, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

686999

(54) Title: PROCESS FOR PREPARING MODIFIED PROTEINS

(57) Abstract

Process for the preparation of modified proteins comprising the coupling of a first peptide segment having a haloacyl group at the N-terminus thereof with a second peptide segment having a carbonylthiol group at the C-terminus thereof are disclosed. Novel modified proteins produced by the process are also disclosed.

PROCESS FOR PREPARING MODIFIED PROTEINS

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

This invention is in the area of modified biomolecules and methods of making such modified biomolecules. More particularly, this invention relates to protein engineering by chemical means to produce modified proteins where one or more peptide bonds are substituted by non-peptide linkage and one or more encoded amino acids may be replaced by unnatural amino acids or amino acid analogs or any other non-coded structure.

RELATED ART

Numerous attempts have been made to develop a successful methodology for synthesizing modified biomolecules such as proteins, glycoproteins, nucleotides, polysaccharides, and other biopolymers. Such modified biomolecules are invaluable for study of structure-activity relationships of native biomolecules and there is a growing number of commercial applications of these molecules for diagnostic or therapeutic purposes.

Structural modification of proteins and peptides, normally referred to as "protein engineering" involves the rationally designed alteration of structure with the aim of understanding protein structure and function and of creating a protein with new desirable properties. In the past, this has been principally carried out by site-directed mutagenesis or other

techniques involving genetic manipulation. The major drawbacks of these prior art approaches are that amino acids replacing native amino acids are those that must be coded genetically. As a result, other structural variants such as unnatural amino acids or amino acid analogs cannot be introduced in the protein backbone. However, recent findings (Ellman, *et al.*, *Science*, 255:197, 1992; Noren, *et al.*, *Science*, 24:182, 1989) would allow unnatural amino acids or amino acid analogs to be incorporated into proteins in a site-specific manner. In this approach, a codon encoding an amino acid to be replaced is substituted by the nonsense codon TAG by means of oligonucleotide-directed mutagenesis. A suppressor tRNA directed against this codon is then chemically aminoacylated with the desired unnatural amino acid. Addition of the amino acylated tRNA to an *in vitro* protein synthesizing system programmed with the mutagenized DNA directs the insertion of the prescribed amino acid into the protein at the target site. Taking the enzyme T4 lysozyme, the above authors, incorporated a wide variety of amino acid analogs into the enzyme at alanine 82 position with a few exceptions, for example, ~~of~~ D-alanine not being incorporated.

While Schultz's approach partially solves problems posed by biosynthetic protein engineering, it does not allow the alteration of the protein backbone at more than one target site to incorporate two or more different non-coded structural units. Also, by the very nature of the system, i.e., the fact that it relies upon a living system to produce the engineered protein, many substitutions or alterations, such as those which would result in a lethal mutation, cannot

be done. The chemical synthesis would overcome the shortcomings left by the Schultz techniques. (reviewed by R. E. Offord, *Protein Eng.*, 1:151, 1987). However, chemical synthesis is fraught with many difficulties such as the need of protection of unwanted reactive groups.

Overall, there is a definite need for a simple and efficient method for making a modified protein which ~~posses~~ possesses desired properties. The present invention addresses such need and provides novel modified proteins.

SUMMARY OF THE INVENTION

This invention provides new and useful modified biomolecules. It also provides a new process for producing such modified biomolecules. In general, the modified biomolecules of this invention comprise two molecular segments, each selected from peptides, pseudopeptides, or non-peptide linear molecules linked through a non-amido linkage to form a peptide or pseudopeptide backbone, wherein one of the segment contains at least one non-coded structural unit and the non-coded structural unit does not form a part of the non-amido linkage. The chemical bonding of the two segments is by means of terminal reactive groups on one segment which react with reactive groups of the other segment molecule.

The process of this invention provides a directed ligation of the two molecular segments to create a desired bond at the ligation point(s) and comprises the steps of:

a. providing a first segment having at least one non-coded structural unit and attaching a first chemoselective synthon to the first segment at the terminal position thereof;

b. providing a second segment optionally containing at least one non-coded structural unit, and second chemoselective synthon at the terminal position thereof, the second chemoselective synthon being complementary to the first chemoselective synthon of the first segment; and

c. ligating the first segment and the second segment, whereby the first synthon of the first segment and the second synthon of the second segment form a non-peptide linkage, wherein the first segment and the second segment are each selected from peptides, pseudopeptides, or non-peptide linear molecules, provided that both segments are not non-peptide linear molecules at the same time.

The above sequence a-c can be repeated by using a first modified biomolecule as the first segment to which a second segment or a second biomolecule is ligated. The present process also may include the step of ligating additional segments with the first and second segments which have been provided with additional terminal synthons that are compatible with the first and second synthons and chemoselective to synthons of the additional segments.

The present invention is therefore applicable in the chemical synthesis of various protein conjugates, such as proteins with reporter molecules, radionuclides, cytotoxic agents, nucleotides, antibodies, and non-protein micromolecules.

Preferably, the process of this invention involves a series of steps comprising:

a. sequentially coupling selected amino acids or amino acid analogs to a terminal amino acid or amino acid analog bound to a first resin support to form a first peptide segment-resin, the first peptide segment having about two to about one hundred amino acid residues;

b. covalently attaching a haloacyl moiety to the N-terminus of the first peptide segment-resin to form a haloacylpeptide segment bound to the first resin support;

c. cleaving the haloacylpeptide peptide segment from the first resin support;

d. sequentially coupling amino acids or amino acid analogs to a terminal amino acid or amino acid analog bound to a second resin support through a sulfur, nitrogen or a selenium-containing bond to form a second peptide segment-resin, the second peptide segment comprising from about two to about one hundred amino acid residues, with a proviso that the nitrogen containing bond is employed for bind ng the terminal amino acid or amino acid analog to the second resin support if and only if the terminal amino acid or amino acid analog is cysteine or an S-H containing analog of cysteine;

e. cleaving the second peptide segment-resin to form a second peptide segment having a thiol-, selenol-, or nitrogen containing group at the C-terminus thereof, with a proviso that the C-terminus of the second peptide segment has a nitrogen containing group if and only if the C-terminal amino acid or amino acid analog is cysteine or an S-H containing analog of cysteine; and

f. coupling the haloacylpeptide peptide segment and the second peptide segment to form a modified polypeptide.

The order of the sequence of steps a-b-c-d-e- is not critical to this invention. The sequence of steps a-b-c and the sequence of steps d-e may be conducted successively or separately. The entire sequence can be repeated in a chain-reaction manner.

Optionally, any reactive groups such as thiol that may be present in the peptide segments can be protected prior to step (f) and deprotected after step (f) is complete.

This invention, in its broadest sense, encompasses a biologically active protein comprising two molecular segments, each selected from peptides, pseudopeptides, or non-peptide linear molecules linked through a non-amido linkage to form a peptide or pseudopeptide backbone, wherein one of the segments contains at least one non-coded structural unit and the non-coded structural unit does not form a part of the non-amido linkage, provided that both segments are not a non-peptide linear molecule at the same time.

This invention further provides a modified protein represented by the formula:

R-L-R'

wherein R and R' are the same or different and are each peptides or pseudopeptides having at least two amino acid residues; and L represents a thiol ester or selenol ester linkage.

The above objects and features of the invention will become more fully apparent from the description of the preferred embodiments in conjunction with the accompanying figures.

BRIEF DESCRIPTION OF THE FIGURES

FIGURE 1 shows a synthetic strategy for the total chemical synthesis of HIV-1 PR analogs in accordance with this invention where "A" schematically represents the coupling of the N-terminal segment HIV-1 PR (1-50, Gly^{51} SH) and the C-terminal segment bromoacetyl (5399) HIV-1 PR, and "B" schematically represents the coupling of the N-terminal segment HIV-1 PR (1-50, Cys^{51}amide) and the same C-terminal segment.

FIGURE 2 shows an elution profile of aliquots taken at t = 0, 45 min., 3 h, and 48 h from the ligation mixtures containing HIV-1 PR (1-50, $GlY^{51}SH$) and bromoacetyl (53-99) HIV-1 PR by the practice this invention using reverse phase HPLC (absorbance 214nm).

FIGURE 3A shows the ion spray mass spectrum of the purified ($[NHCH_2COSCH_2CO]^{51-52}Aba^{67,95}$) HIV-1 PR where the labeled peaks represent a single molecular species differing in the number of excess protons.

FIGURE 3B shows the deconvoluted mass spectrum of the purified ([$NHCH_2COSCH_2CO$]$^{51-52}$ Aba ^{62,95}) HIV-1 PR where the peak of the molecular weight of the enzyme is located at 10, 769 dalton.

FIGURE 4 shows an elution profile of aliquots taken at t = 0, 30 min, and 120 min from the ligation mixture containing HIV-1 PR(1-50, Cys amide) and bromoacetyl (53-99)] HIV-1 PR.

FIGURE 5 shows an elution profile of the hexapeptide Ac-Thr-Ile-Nle-Nle-Gln-Arg; NH_2 before and after treatment with the ligation mixture taken at t = 3h as shown in FIGURE 2 using reverse phase HPLC with absorbance monitored at 214 nm, where the upper panel shows peaks before the treatment and the lower panel, peaks after the treatment, respectively.

FIGURE 6 shows a fluorogenic assay of aliquots of the ligation mixture taken at the times indicated as shown in FIGURE 2 where data points illustrating fluorescence units were read from continuous chart recorder tracings.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

This invention is based on a conceptually novel approach to synthesizing large biomolecules such as proteins by a convergent type synthesis. Thus, when applied to the synthesis of modified proteins, this invention involves the coupling of at least two peptide segments to create a linkage which can be a non-amido bond. The peptide segments may be synthesized using solid phase peptide synthesis,

solution phase synthesis, or by other techniques known in the art including combinations of the foregoing methods.

The present process involves the steps of:

a. providing a first segment having at least one non-coded structural unit, and a first chemoselective synthon at the terminal position thereof;

b. providing a second segment optionally containing at least one non-coded structural unit, and a second chemoselective synthon at the terminal position thereof, the second chemoselective synthon being complementary to the first chemoselective synthon of the first segment; and

c. ligating the first segment and the second segment, whereby the first synthon of the first segment and the second synthon of the second segment form a non-peptide linkage, wherein the first segment and the second segment are each selected from peptides, pseudopeptides, or non-peptide linear molecules, provided that both segments are not non-peptide linear molecules at the same time.

Preferably, both of the first and second segments are peptides or pseudopeptides having from about two to about one hundred amino acid residues. More

preferably, the segments constitute from about forty to sixty amino acid residues.

Preferably, the non-peptide linkage formed is represented by one of the following linking moieties:

$-CH_2-S-$, $-CH_2-Se-$, $-CO-S-$, $-CO-Se-$, $-CH_2-NH-$, or $-CH(OH)-CH_2-$.

The present process allows at least two different non-coded structural units to be incorporated to the segments as well as allowing the same non-coded structural units to be incorporated at two sites.

In view of the broad and varied class of modified proteins which may be bound by this invention, all chemically modified proteins having such characteristics as defined above are deemed to be within the scope of this invention. However, for purposes of illustrative clarity and ease of comprehension, this invention will be described herein in more detail utilizing a modified HIV-1 protease as an embodiment. It should be understood that the use of this particular enzyme for descriptive purposes shall not restrict nor limit the use of other proteins or peptides.

As employed herein, the term "modified protein" is intended to include oligopeptides, oligopseudopeptides, polypeptides, pseudopolypeptides, and modified native proteins-synthetic or otherwise derived. The term "pseudopeptide" means a peptide where one or more peptide bonds are replaced by non-amido bonds such as ester or one or more amino acids

are replaced by amino acid analogs. The term "peptides" refers not only to those comprised of all natural amino acids, but also to those which contain unnatural amino acids or other non-coded structural units. The terms "peptides", when used alone, include pseudopeptides. The "modified proteins" have utility in many biomedical applications because of increased stability toward *in vivo* degradation, superior pharmacokinetics, and enhanced or diminished immunogenecity compared to their native counterparts.

HIV-1 protease (HIV-1 PR) is a virally-encoded enzyme which cuts polypeptide chains with high specificity and which is essential for the replication of active virions (N.E. Kohl, *et al., Proc. Natl. Acad. Sci., U.S.A.*, 85:4686, 1988). The 21,500 dalton HIV-PR molecule is made up of two identical 99 amino acid polypeptide chains.

Comparison of the crystal structures of the empty (A. Wlodawer, *et al., Science*, 245:616, 1989) and inhibitor-bound (for example, M. Miller, *et al., Science*, 246:1149, 1989) enzyme revealed that on binding a substrate-derived inhibitor the HIV-1 molecule undergoes significant conformational changes which are particularly pronounced in two exterior, functionally-important "flap" regions. From these crystallography studies it appears that peptide bonds in the flap regions of the HIV-1 PR polypeptide backbone are involved in the formation of ß-sheet/ß-turn structure, in the interaction which occurs between the two subunits of the active dimer at the tip of each flap in the enzyme-inhibitor (substrate) complex, and in hydrogen bonding interactions with

bound peptide inhibitors (and, presumably, substrates). Mutagenesis studies carried out with recombinant HIV-1 PR showed that the flap region is highly sensitive to changes in the amino acid sequence (D.D. Loeb, *et al.*, *Nature*, 340:397, 1989).

These observations make the flap region especially interesting as a target for protein backbone modifications to investigate the role of peptide bond interactions in HIV-1 protease activity.

In the case of HIV-1 PR where modifications of the flap region are desired, this invention allows pseudo-peptide bonds to be introduced into the region. The Gly^{51}-Gly^{52} bond of HIV-1 PR is particularly preferred for bond manipulation for two reasons. First, glycine is the only achiral amino acid and therefore there is no concern about loss of optical purity; and second, the Gly^{51}-Gly^{52} bond is located near the middle of the 99 amino acid HIV-1 PR monomer polypeptide chain. This means that the two peptide segments which are to be coupled are each about 50 residues in length.

The solid phase peptide synthesis method is generally described in the following references: Merrifield, *J. Am. Chem. Soc.*, 888:2149, 1963; Barany and Merrifield, *In the Peptides*, E. Gross and J. Meinenhofer, *Eds.*, *Academic Press*, New York, 3:285 (1980); S.B.H. Kent, *Annu. Rev. Biochem.*, 57:957 (1988). By the solid phase peptide synthesis method, a peptide of a desired length and sequence can be produced through the stepwise addition of amino acids to a growing peptide chain which is covalently bound to a solid resin

particle. Automated synthesis may be employed in this method.

Accordingly, these embodiments can be accomplished by the steps of:

1. sequentially coupling amino acids to a terminal amino acid bound to a resin support to form a peptide segment-resin; and

2. cleaving the peptide segment from the resin support.

In the preferred application of this method, the C-terminal end of the growing peptide chain is covalently bound to a $-OCH_2$ PAM resin and amino acids having protected α-amino groups are added in the stepwise manner indicated above. A preferred α-amino protecting group is the tert-butyloxycarbonyl (BOC) group, which is stable to the condensation conditions and yet is readily removable without destruction of the peptide bonds or racemization of chiral centers in the peptide chain. At the end of the procedure the product peptide is cleaved from the resin, and any remaining protecting groups are removed by treatment under acidic conditions such as, for example, with a mixture of hydrobromic acid and trifluoroacetic acid, with trifluoromethane sulfonic acid or with liquified hydrogen fluoride.

In the case of HIV-1 protease, the C-terminal segment comprises the following amino acid sequence: F^{53} IKVRQYD^{60} QIPVEICGHK^{70} AIGTVLVGPT^{80}PVNIIGRNLL^{90}TQIGCTLNF^{99}. The N-terminal

segment comprises the following amino acid sequence: $P^{1}QITLWQRPL^{10}$ VTIRIGGQLK^{20} EALLDTGADD^{30} TVLEEMNLPG^{42} KWKPKMIGGI$^{50}G^{51}$. The synthesis of these segments normally require about 15 hours at a standard cycle speed.

If desired, any amino acids in the above sequences can be replaced by amino acid analogs or amino acid mimetic compounds known in the art. Suitable amino acid substitutes include β-alanine, L-α-aminoisobutyric acid, L-α-amino-n-butyric acid (Aba), 3.4-dehydroproline, homoarginine, homocysteine, homoproline, homoserine, 3-mercaptopropionic acid, norleucine (Nle), penicillamine, pyroglutamic acid and sarcosine.

L-amino acids and D-amino acids may be used in this invention, particularly, D-amino acids are useful for the formation of a "reversible peptide sequence".

The "reversed" or "retro" peptide sequence as discussed above refers to that part of an overall sequence of covalently-bonded amino acid residues (or analogs or mimetics thereof) wherein the normal carboxyl-to-amino direction of peptide bond formation in the amino acid backbone has been reversed such that, reading in the conventional left-to-right direction, the amino portion of the peptide bond precedes (rather than follows) the carbonyl portion (see, generally, Goodman, M. and M. Chorev, *Accounts of Chem. Res.*, 12:423, 1979).

The "reversed" peptides are within the meaning of the "peptides" used throughout the specification. D-amino

acids, amino acid analogs and amino acid mimetic compounds are collectively referred to herein as "non-coded amino acids." Although the solid phase peptide synthesis theoretically enables one skilled in the art to prepare a peptide backbone of any length, the efficiency of coupling amino acids (addition of an amino acid in successive cycles) would necessarily limit the use of this technique when a peptide to be synthesized has greater than 150 residues. Preferably, for practical and economic reasons, the process of this invention employs from about two to about one hundred cycles.

Accordingly, the modified proteins obtained by the process of this invention have the general structure set forth in formula (1) and R and R' are as previously defined. If a larger peptide segment (containing greater than one hundred amino acid residues) is desired, such peptide segment may be available from naturally occurring proteins (native proteins) by enzymic or chemical degradation. Alternatively, the present process can be repeated with different ligation modes for building such large peptide segments.

An important feature of this invention is the bond formation linking two peptide segments (namely R and R'). A variety of bond forming reactions can be used to covalently link the two peptide segments. Representative combinations of such groups are carbonylthiol with halo to form a thiol ester bond between the two segments, carbony selenol with melo to form a seleno ester, or thiol with thiol to form a disulfide bond, thiol with halo to form a thioether

bond, selenol with halo to form a selenol ether, amino with isothiocyanate to form a thiourea bond, amino with aldehyde to form an imine bond which can be reduced to a carbon-nitrogen bond, thiol with maleimide to form a thioether bond, and gem-diol with boron to form oxygen-boron bonds, and hydroxyl with carboxyl to form an ester bond. Among these, preferred linkages are those already enumerated by way of their structure. A preferred linkage is a sulfur-containing or selenium-containing linkage which does not readily hydrolyze *in vivo* or which is more ~~liable~~ labile than an amido bond *in vivo*.

The most preferred linkage L is a thiol ester linkage. This linkage can be accomplished by first attaching a facile leaving group to a first peptide segment and by attaching carbonylthiol functionality to a second peptide segment followed by nucleophilic substitution where the sulfur nucleophile attacks the leaving group. Preferably, a haloacyl (e.g., haloacetyl) such as iodo, chloro, or bromoacetyl is attached to the N-terminus of the first peptide segment. The suitable haloacyl group may be straight or branched (substituted by alkyl). This step is conveniently carried out while the first peptide segment is still bound to a resin support.

To introduce a bromoacetyl group to the first peptide segment, suitable activated forms of bromoacetic acid may be employed in this invention. The particularly preferred agent for this purpose is bromoacetic anhydride.

When the aforementioned C-terminal segment of HIV-1 PR is derivati ed with the N-bromoacetylating agent, the unprotected N-terminus of the protected peptide segment on resin is condensed with bromoacetic anhydride to produce bromoacetyl (53-99) HIV-1 PR. Deprotection and release of the product peptide segment from the resin support can be accomplished by standard conditions (e.g., treatment with anhydrous HF containing 10% p-cresol at 0°C for several hours). The product peptide segment is precipitated, dried by lyophilization, and purified, if desired, by reverse phase HPLC, according to standard techniques known in the art.

To introduce carbonylthiol functionality to the second peptide segment, its terminal carboxylic acid moiety may be converted to a carbonylthiol group.

When the aforementioned N-terminal segment of HIV-1 PR is to be derivatized accordingly, 4-[α-(Boc-Gly-S) benzyl] phenoxyacetamidomethyl-resin is used as the resin support. If any amino acid or amino acid analog other than glycine is desired, that amino acid can be loaded on an aminomethyl-resin in the form of 4-[α-(Boc-X-S) benzyl] phenoxyacetic acid wherein X represents the amino acid.

The N-terminal derivatized peptide segment can thus readily be prepared by the stepwise solid phase synthesis. The second product peptide segment is cleaved from the resin support, deprotected, and isolated in the manner described above.

The thiol ester linkage can be generated by coupling the two segments under normal ligation conditions. The formation of the thiol ester linkage is highly chemoselective and thus compatible with most reactive groups that may be present in the molecules.

A typical ligation reaction is carried out by mixing the unprotected N- and C- terminal segments in 6M guanidine hydrochloride buffer at about pH 3-6. In this buffer, the solubility of the unprotected peptide segments is very high, thus eliminating the major drawback of the prior art techniques having to use protected peptide fragments despite their limited solubility in ligation media. Other denaturants such as urea, detergents, and sodium dodecyl sulfate can be used as the ligation buffer.

The coupling of HIV-1 PR (1-50, Gly-SH) with bromoacetyl (53-99) HIV-1 PR is complete in several hours. The product peptide [(NH $CH_2COSCH_2CO)^{51-52}$] HIV-1 PR can be isolated, purified and characterized by standard techniques.

When one of the peptide segments has any reactive groups which may interfere with the thiol ester formation, those groups may be protected by suitable protecting groups known in the art or the amino acids bearing such groups can be substituted by other amino acids or amino acid analogs which are reaction inert and yet do not adversely affect the biological activity of a product protein.

For example, in the Examples, the C-terminal segment of HIV-1 PR has two cysteine residues at positions 67

and 95. These cysteine positions have been shown to be replaceable by L-α-amino-n-butyric acid (Aba) without causing the loss of the enzymic activity of the native protease. In one embodiment of this invention, bromoacetyl [(53-99)$Aba^{67,95}$] HIV-1 PR is used. Alternatively, to block a thiol group of the cysteine residues before the coupling reaction, the use of protecting groups that are compatible with ligation conditions are preferred. However, these precautions may not be necessary since, in present experience, the thiol ester group of the N-terminal segment attacks a bromoacetyl group even at low pH conditions where a thiol side chain of a cysteine residue is unreactive.

In another embodiment of the invention, the linkage L can be a thioether bond. This linkage can be accomplished in substantially the same manner as that described for the formation of the thiol ester linkage, except that thiol functionality is attached to a second peptide segment. This is conveniently accomplished by utilizing a thiol side chain of a cysteine residue. Thus, cysteine can be employed as the C-terminal amino acid of the second peptide segment (N-terminal segment). HIV-1 PR(1-50, Cys^{51} amide) described herein represents one example of such derivatized second peptide segment where the cysteine's carboxylic acid is blocked as amide. When this peptide segment is subjected to the ligation conditions described earlier in the presence of bromoacetyl (53-99) HIV-1 PR, the coupling of the two peptide segments takes place rapidly to afford {[$NHCH(CONH_2)CH_2SCH_2CO$]$^{51-52}$}HIV-1 PR. The ligated product protein can be isolated, purified, and

characterized by standard techniques. In a further embodiment of the invention, the linkage can be a selenium-containing bond such as selenol ester and seleno ether. In like manner, a selenocysteine can be used in place of cysteine.

The process of this invention is applicable to the synthesis of any biomolecules and is not limited to proteins insofar as two constituent fragments are available by chemical synthesis or other biosynthetic methods. In particular, the present process: (1) allows rapid synthesis of modified proteins;(2) avoids the use of protecting groups in at least the critical bond formation stage;and (3) incorporates into the protein backbone structural units such as D-amino acids and amino acid analogs.

This invention will be described in further detail below by way of the aforeindicated embodiments, but these embodiments should not be taken as limiting the scope of the invention.

EXAMPLE 1

SYNTHESIS OF [$NHCH_2COSCH_2CO)^{51-51}Aba^{67,95}$] HIV-1 PR

Two peptide segments, HIV-1 PR (1-50, $Gly^{51}SH$) (PREPARATION 3) and bromoacetyl (53-99) HIV-1 PR (PREPARATION 2) were coupled by ligating the segments in 6M guanidine hydrochloride 0.1 M sodium phosphate at pH4.3. The segments were separately dissolved in ligation buffer at a concentration of 20 mg/ml.

The process of the ligation was followed by reverse phase HPLC on a Vydac C_{18} column using a linear gradient of 30-60% buffer B (90% acetonitrile/0.09% trifluoroacetic acid) in buffer A (0.1% trifluoroacetic acid) in 30min. The flow rate was 1 ml/min and absorbance was monitored at 214 nm. Results are shown in FIGURE 2.

The resulting ligated peptide was purified by reverse phase HPLC on a semipreparative Vydac C_{18} column using various linear gradients of 90% acetonitrile/0.09% TFA in 0.1% aq TFA. The purity of the product peptide was checked by analytical HPLC as well as by ion spray mass spectrometry. The product peptide had the expected molecular weight of 10.768.6 ± 1.1 Da (Calcd. for monoisotopic; 10.763.9 Da; (average) 10, 770.8 Da). The mass spectra of the title peptide are shown in FIGURE 3A and 3B.

Example 2

SYNTHESIS OF [(NHCH(CO NH_2)CH_2SCH_2CO)$^{51-52}$Aba^{67,95}] HIV-1 PR

Two peptide segments HIV-1 PR (1-50, Cys^{51} amide) (PREPARATION 1) and bromoacetyl (53-99) HIV-1 PR (PREPARATION 2) were coupled substantially according to the procedure of EXAMPLE 1, except that the pH was over 7.0. The ligation reaction was monitored by HCLP, with the results shown in FIGURE 3. After work-up and purification, the product peptide had the expected molecular weight of 10,800.31 ± 0.75 Da (Calcd. for monoisotopic 10.792.9 Da; (average) 10,799.8 Da).

EXAMPLE 3

ENZYMATIC ACTIVITY OF [($NHCH_2COSCH_2CO$)$^{51-52}$Aba^{67,95}] HIV-1 PR

An aliquot of the ligation reaction mixture after 3 hours ligation (EXAMPLE 1) was treated with a peptide substrate (1 mg/ml) having the sequence of Ac-Thr-Ile-Nle-Nle-Gln-Arg. NH_2 (Nle:L-norleucine) at pH 6.5. Reaction was monitored by reverse phase HPLC using a Vydac C18 column with the results shown in FIGURE 5. In FIGURE 5 the upper panel shows the peptide substrate peaks before treatment. The lower panel shows the peaks of cleavage products after 15 min. treatment.

The cleavage products were separated by reverse phase HPLC (Vydac C_{18} column; linear gradient of 0 to 40%

buffer B (90% acetonitrile/0.09% trifluoroacetic acid) in buffer A (0.1% trifluoroacetic acid) over 20 min; flow rate, 1 ml/min; absorbance monitored at 214 nm). The peptide products were identified by ion spray mass spectrometry as (H)-Nle-Gln-Arg. NH_2 (early eluting) and Ac-Thr-Ile-Nle-(OH) (late eluting).

Thus, the title peptide shows substantially the same specificity as the native enzyme HIV-1 PR against the substrate peptide employed in this assay.

Further, an aliquot of the ligation reaction mixture was treated with a solution containing a fluorogenic substrate in 100 mM MES buffer at pH 6.5. The fluorogenic substrate used was 2-aminobenzoyl-Thr-Ile-Nle-Phe(p-NO_2)-Gln-Arg.-NH_2. Fluorescence was recorded as time of incubation. Date points shown in FIGURE 6B were read from continuous chart recorder tracings. Quantitative comparison of ([$NHCH_2COSCH_2CO$]$^{51-52}$Aba^{67,95}) HIV-1 PR with an equal amount of [Aba$^{67, 95, 167, 195}$] HIV-1 PR showed identical activities. Neither segment alone namely, HIV-1 PR(1-50, Gly^{51}SH), nor bromoacetyl (53-99; Aba^{67,95}) HIV-1 PR showed any activity (detection limit, $<$1 part in 1,000)

PREPARATION 1

HIV-1 PR (1-50, Cys.^{51} amide)

The synthesis of the title peptide was performed on an Applied Biosystem 430A peptide synthesizer. The instrument was modified by removal of delivery-line filters to the top and bottom Reaction Vessel valve

blocks, and the filter in delivery-line to the Cartridge needle. Larger (1.2mm internal diameter) tubing was used from the DMF reservoir to the bottom reaction vessel valve block. A direct transfer line was used from the Activator (cartridge) valve block to the reaction vessel block (i.e., the Concentrator valve block was bypassed to allow direct transfer of the activated amino acid solution from the cartridge to the reaction vessel). Reagent positions 2 and 3 were connected to act as a common reservoir for TFA, giving capacity of 900 ml (40+ cycles). The A-regulator pressure for TFA delivery was doubled to 5 psi for a delivery rate of 18ml/min. Polypropylene bottle seals for the TFA bottles were replaced with Viton gaskets to prevent leaks. The metering loop for reagent position 7 (DIEA for delivery to the Cartridge) was cut to length to deliver 0.72g (5.60 mmol) DIEA. The metering loop for reagent position 8 (0.48 M HBTU in DMF) was cut to length to deliver 2.22 ml. Reagent position 1 was DIEA and was used only to deliver DIEA to fraction collector via the resin sampler. Reagent position 9 was DCM and was used to rinse residual DMF out of the cartridge prior to drying for reuse. Reagent position 10 was DMF. Positions 4, 5, and 6 were not used. A custom-designed 20ml reaction vessel was used for all syntheses.

Synthetic cycles for use on the 430A were written from scratch. The Begin cycles only set toggle functions and rinsed the starting resin with DMF. The End cycle rinsed the product peptide-resin twice with DMF, then with DCM and dried the peptide-resin under nitrogen. special functions were written for the Concentrator and Activator programs, to allow the Activator program (i.e., events in the cartridge) to be coordinated with

the Reaction Vessel program. These coordinating functions were similar in principle to the special function described by Reid and Simpson and the functions used in earlier rapid double couple cycles (Kent, et al., in "Peptides: Chemistry and Biology", *Procc. 10th Amer. Peptide Syn.*). A special function in the Activator program was also written to deliver Reagent 8 (0.48 MHBTU) to the cartridge. The synthetic protocols used are the following:

Time/Mode		**Synthesis Cycle**
Deprotect	100% TFA	2 x 30 sec flow wash 1 x 1 min, vortex
Wash	DMF	1 x 1 min flow wash drain, minimize hold-up
Couple	activated Boc-AA*	10 min, vortex
Wash	DMF	1 x 30 sec flow wash tal resin sample for ninhydrin test, drain

* Activation of Boc-AA (scale: 0.2-0.5 mmol resin)

dissolve 2.25 mmol Boc-AA
in 4.44 ml 0.48 M HBTU (2.13 mmol) in DMF for 5 min
add 5.60 mmol DIEA
activate for 2.5 min
transfer from cartridge to reaction vessel

The first TFA treatment consisted of 2x30 sec each upward vortexing flow washes followed by brief pulses of gas. After the batch treatment for 1 min, TFA was removed from the Reaction Vessel by a 1 min upward vortexing flow wash with DMF. The line from the resin

sampler switching valve was rinsed into the Reaction Vessel with DMF in the middle of this DMF flow wash. The peptide-resin was drained by filtration under nitrogen pressure prior to the coupling step, in order to minimize the dilution of the activated Boc-amino acid. The activated Boc-amino acid solution was 0.25 M at transfer. Hold-up of DMF in the swollen peptide-resin reduced the concentration to about 0.15-0.2M in the reaction vessel. Brief DMF line rinses (0.3 ml each) were used to clean delivery lines after delivery of HBTU solution to the cartridge needle, and after the addition of the DIEA to the cartridge. The transfer line from the Activator valve block to the Reaction vessel was also rinsed with DMF before and after transfer of the activated Boc-amino acid solution. Peptide-resin samples (5-7mg) were taken in DMF and stored in DIEA/DMF until analysis. Samples were washed thoroughly with DCM-MeOH (1:1 v/v) before drying and quantitative ninhydrin analysis. The overall cycle time for the addition of a single amino acid residue was 16 min 30 sec plus transfer times (Cartridge to Reaction Vessel, including rinses: ~2 min), plus the artefactual deadtime of 45 sec at the beginning of each cycle imposed by the operating software of the 430A. Total cycle time was 19 min 15 sec (about 75 residues per day).

The following chain protecting groups were used: D and E, cyclohexyl; N, xanthyl; Q and M, unprotected; K, 2-chlorobenzyloxycarbonyl (2CIZ); R, tosyl; Y, bromobenzyloxycarbonyl (BrZ); H, 2,4-dinitrophenyl (DNP); T, benzyl. The formyl group was used to protect W(Trp). The peptide was removed from the resin by treatment with HF in the presence of 10% p-cresol for 1

hr at 0°C. Nα-Boc group had been removed prior to the HF treatment. 4-MeBHA-resin (4-methylbenzylhydrylamine) was used as the resin support in the above procedure.

The resulting peptide was precipitated with ether and finally dissolved in 50% acetic acid, diluted with water and lyophilized. Thus, starting with 216 mg of the resin, 1.36 g of the peptide resin resulted. The title peptide (181 mg) was obtained from 385 mg of the peptide resin.

The peptide was purified by reverse phase HPLC on a semipreparative Vydac C_{18} column using various linear gradients of 90% acetonitrile/0.09% TFA in 0.1% aq TFA. This purity was checked by analytical HPLC as well as by ion spray mass spectrometry.

PREPARATION 2

BROMOACETYL [(53-99) $Aba^{67,95}$] HIV-1 PR

By employing the solid phase protein synthesis procedure as described in PREPARATION 1, the title peptide was prepared. Boc Phe (OCH_2)-phenylacetamidomethyl-resin (PAM resin) was used as the resin support. Thus, starting with 214 mg of the resin and through stepwise elaboration (46 cycles) of the peptide-resin was obtained. DNP and N^{α}-Boc groups were removed prior to bromoacetylation. A protected peptide (Phe^{53}-Phe^{99}) was bromoacetylated using the conditions according to F.A. Robey, *et al., Anal. Biochem.*, 177:373, (1989).

To 525 mg protected peptide-resin 0.4 mmol bromoacetic anhydridl in 2 ml DCM/DMF (1:1 v/v) was added. The

reaction was allowed to proceed as 25°C for 30 min with shaking, after which time the reaction appeared to be complete as evidenced by the disappearance of free amine. The resin was filtered, washed with DMF, then DCM/T60W (50:50 v/v) and dried in vacuum. The peptide was removed from the resin by treatment with HF in the presence of 10% p-cresol for 1 hr at 0°C. The resulting peptide was precipitated with ether and finally dissolved in 50% acetic acid, diluted with water and lyophilized to produce 259 mg of the title peptide.

The peptide was purified by reverse phase HPLC on a semipreparative Vydac C_{18} column using various linear gradients of 90% acetonitrile/0.09% TFA in 0.1% aq TFA. The purity was checked by analytical HPLC as well as by ion spray mass spectrometry. The purified sample had the molecular mass of 5235.8 ± 0.2 daltons [Calc:(monoisotoic)5231.8 daltons; (average) 5236.0 daltons].

PREPARATION 3

HIV-1 PR (1-50, $Gly^{51}SH$)

By employing the solid phase protein synthesis procedure as described in PREPARATION 1, the title peptide was prepared. 4-[α(Boc-Gly-S) benzyl]-phenoxyacetamido methyl-resin was prepared as described by D. Yamashiro, *et al., Int. J. Peptide Protein Res.*, 31:322, (1988) and used as the resin support. Starting with 359 mg 4-[α(Boc-Gly-S) benzyl]-phenoxy acetamido methyl-resin and through stepwise elaborating (50 cycles) of the protected peptide chain, 1.1 g of the peptide-resin was obtained. In this synthesis, Gln, Trp, and Met were

unprotected. The resulting peptide was precipitated with ether and finally dissolved in 50% acetic acid, diluted with water and lyophilized. The peptide was purified by reverse phase HPLC on a semipreparative Vydac C_{18} column using various linear gradients of 90% acetonitrile/0.09% TFA in 0.1% aq TFA. This purity was checked by analytical HPLC as well as by ion spray mass spectrometry.

The invention now being fully described, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made without departing from the spirit or scope of the invention.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A process for preparing modified proteins which comprises the steps of:

a. sequentially coupling amino acids or amino acid analogs to a terminal amino acid or amino acid analog bound to a first resin support to form a first peptide segment-resin, the first peptide segment peptide comprising from about two to about one hundred amino acid residues;

b. covalently attaching a haloacyl moiety to the N-terminus of the peptide segment-resin to form a haloacylpeptide segment bound to the first resin support;

c. cleaving the haloacylpeptide segment from the first resin support;

d. sequentially coupling amino acids or amino acid analogs to a terminal amino acid or amino acid analog bound to a second resin-support through a sulfur, nitrogen or a seleniumcontaining bond to form a second peptide segment-resin, the second peptide segment comprising from about two to about one hundred amino acid residues, with a proviso that the nitrogen containing bond is employed for binding the terminal amino acid or amino acid analog to the second resin support if and only if the terminal amino acid or amino acid analog is cysteine or an S-H containing analog of cysteine;

e. cleaving the second peptide segment-resin to form a second peptide segment having a thiol-, selenol-, or nitrogen containing group at the C-terminus thereof, with a proviso that the C-terminus of the second peptide-segment has a nitrogen-containing group if and only if the C-terminal amino acid or amino acid analog is cysteine or an S-H containing analog of cysteine; and

f. coupling the haloacylpeptide segment and the second peptide segment.

2. The process according to claim 1, wherein the haloacyl is bromoacetyl.

3. The process according to claim 2, wherein the bromoacetylpeptide segment is bromoacetyl(53-99)HIV-1 PR.

4. The process according to claim 2, wherein the bromoacetylpeptide segment is bromoacetyl[(53-99), $Aba^{67,95}$]HIV-1 PR.

5. The process according to claim 1, wherein, in step (d), the sulfur-containing bond linking the second peptide segment and the second resin support is a thiol ester linkage and in step (e), the thiol-containing group of the second peptide segment is carbonylthiol.

6. The process according to claim 5, wherein the second peptide segment is HIV-1 PR(1-50,Gly^{51}SH).

7. The process according to claim 1, wherein the nitrogen-containing bond linking the second segment peptide and the second resin support is an amide linkage.

8. The process according to claim 7, wherein the second peptide segment is HIV-1 PR(1-50, Cys^{51}amide).

9. The process of claim 7, wherein the second peptide segment is HIV-1 PR(1-50, Cys^{51}).

10. The process according to claim 1, wherein the coupling between the haloacyl peptide segment and the second peptide segment is a selenol ester or a seleno ether bond.

11. The process according to claim 1, wherein the first resin support is a $COCH_2$-PAM resin.

12. The process according to claim 1, wherein the second resin support is 4-[α-(Boc-amino acid-S) benzyl] phenoxyacetomidomethyl resin.

13. The process according to claim 1, wherein the haloacylpeptide segment is unprotected.

14. The process according to claim 1, wherein the second peptide segment is unprotected.

15. The process according to claim 5, wherein the modified protein has the linkage of -CO-S-CH_2-CO-.

16. The process according to claim 15, wherein the modified protein is $[(NHCH_2COSCH_2CO)^{51-52}\ Aba^{67,95}]$ HIV-1 PR.

17. The process according to claim 15, wherein the modified protein is $[(NHCH_2COSCH_2CO)^{51-52}]$ HIV-1 PR.

18. The process according to claim 7, wherein the modified protein has the linkage of -NH-CH(CO-NH_2)CH_2-S-CH_2-CO-.

19. The process according to claim 18 wherein the modified protein is $\{[NHCH(CONH_2)CH_2SCH_2CO]^{51-52}Aba^{67,95}]\}$ HIV-1 PR.

20. A modified protein represented by the formula:

R-L-R'

wherein R and R' are the same or different and are each peptides or pseudopeptides having at least two amino acid residues; and L represents a thiol ester or selenol ester linkage.

21. The modified protein according to claim 20, wherein the protein is a protease.

22. The modified protein according to claim 21, wherein the thiol ester linkage is -CO-S-CH_2-CO-.

23. The modified protein according to claim 22, wherein both peptide and pseudopeptide comprise about two to about one hundred amino acid residues.

24. The modified protein according to claim 22, wherein the peptide is a native peptide and the pseudopeptide is an analog of a native peptide.

25. The modified protein according to claim 20, wherein R is HIV-1 PR(1-51).

26. The modified protein according to claim 20, wherein R' is HIV-1 PR(53-99).

27. The modified protein according to claim 20, wherein R' is HIV-1 PR[(53-99)$Aba^{67,95}$].

28. A modified protein characterized by [$NHCH_2COSCH_2CO$)$^{51-52}$ $Aba^{67,95}$] HIV-1 PR.

29. A modified protein characterized by {[NHCH($CONH_2$) CH_2SCH_2CO]$^{51-52}Aba^{67,95}$} HIV-1 PR.

30. A process for preparing modified proteins which comprises the steps of:

a. sequentially coupling amino acids or amino acid analogs to a terminal amino acid or amino acid analog bound to a first resin support to form a first peptide segment resin, the first peptide segment comprising from about two to about one hundred amino acid residues;

b. covalently attaching a haloacyl moiety to the N-terminus of the peptide segment-resin to form a haloacylpeptide segment bound to the first resin support;

c. cleaving the haloacylpeptide segment from the first resin support;

d. sequentially coupling amino acids or amino acid analogs to a terminal amino acid or amino acid analog bound to a second resin support through a sulfur, nitrogen or a selenium containing bond to form a second peptide segment-resin, the second peptide segment comprising from about two to about one hundred amino acid residues;

e. cleaving the second peptide segment-resin to form a second peptide segment having a nitrogen-, thiol- or selenol-containing group at the C-terminus thereof; and

f. coupling the haloacylpeptide segment and the second peptide segment, and

substantially as described herein with reference to Example 1 or Example 2.

31. A modified protein represented by the formula:

R-L-R'

wherein R and R' are the same or different and are each a residue of a peptide or pseudopeptide; and L represents a thiol ester or selenol ester linkage, the modified protein being substantially as described herein with reference to Example 1 or Example 2.

32. A biologically active polypeptide comprising two molecular segments, each selected from peptides, pseudopeptides, or non-peptide linear molecules linked through a non-amido linkage to form a peptide or pseudopeptide backbone wherein one of the segments contains at least one non-coded structural unit and the non-coded structural unit does not form

a part of the non-amido linkage, provided that both segments are not a non-peptide linear molecule at the same time, the biologically active polypeptide being substantially as herein described with reference to Example 1 or Example 2.

33. A process for preparing a modified biomolecule which comprises the steps of:

a. providing a first segment having at least one non-coded structural unit, and a first chemoselective synthon at the terminal position thereof;

b. providing a second segment optionally containing at least one non-coded structural unit, and a second chemoselective synthon at the terminal position thereof, the second chemoselective synthon being complementary to the first chemoselective synthon of the first segment; and

c. ligating the first segment and the second segment, whereby the first synthon of the first segment and the second synthon of the second segment forms a non-peptide linkage, wherein the first segment and the second segment are each selected from peptides, pseudopeptides, or non-peptide linear molecules, provided that both segments are not non-peptide linear molecules at the same time,

the process being substantially as described herein with reference to Example 1 or Example 2.

DATED this 10th day of November 1997

THE SCRIPPS RESEARCH INSTITUTE

By their Patent Attorneys

GRIFFITH HACK

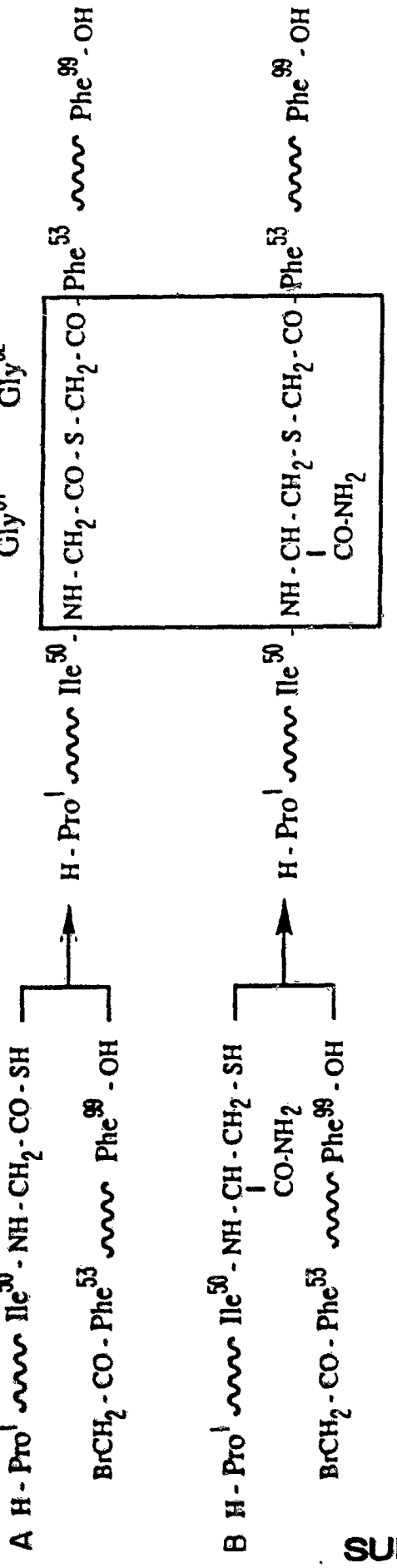
A H - Pro¹ ~~~ Ile⁵⁰ - NH - CH₂ - CO - SH
BrCH₂ - CO - Phe⁵³ ~~~ Phe⁹⁹ - OH
B H - Pro¹ ~~~ Ile⁵⁰ - NH - CH - CH₂ - SH
CO-NH₂
BrCH₂ - CO - Phe⁵³ ~~~ Phe⁹⁹ - OH
H - Pro¹ ~~~ Ile⁵⁰ -
Gly⁵¹ Gly⁵²
NH - CH₂ - CO - S - CH₂ - CO - Phe⁵³ ~~~ Phe⁹⁹ - OH
NH - CH - CH₂ - S - CH₂ - CO - Phe⁵³ ~~~ Phe⁹⁹ - OH
CO-NH₂

FIG. 1

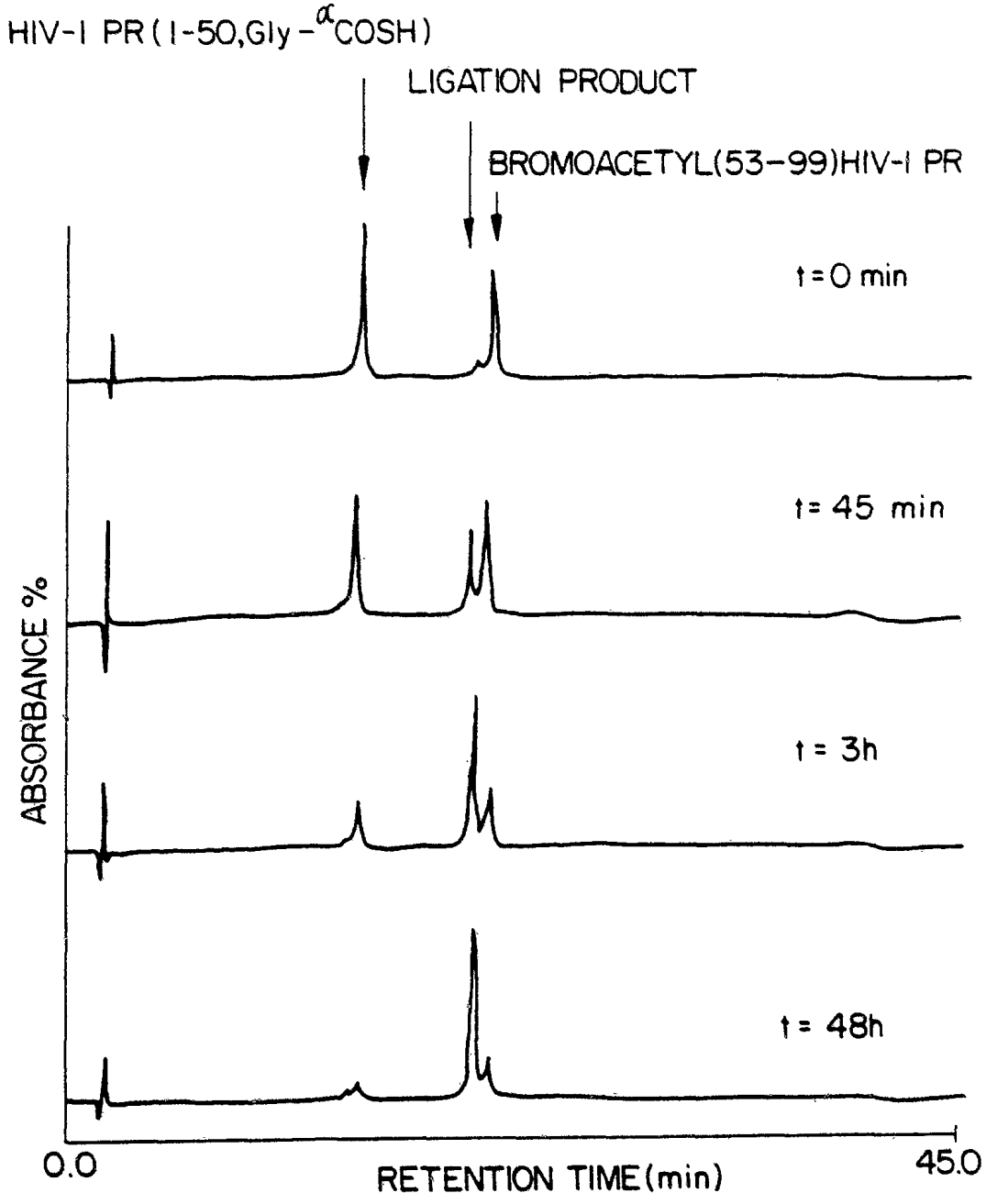
HIV-1 PR(1-50,Gly-αCOSH)
LIGATION PRODUCT
BROMOACETYL(53-99)HIV-1 PR
t=0 min
t=45 min
t=3h
t=48h
ABSORBANCE %
0.0
RETENTION TIME(min)
45.0

FIG. 2

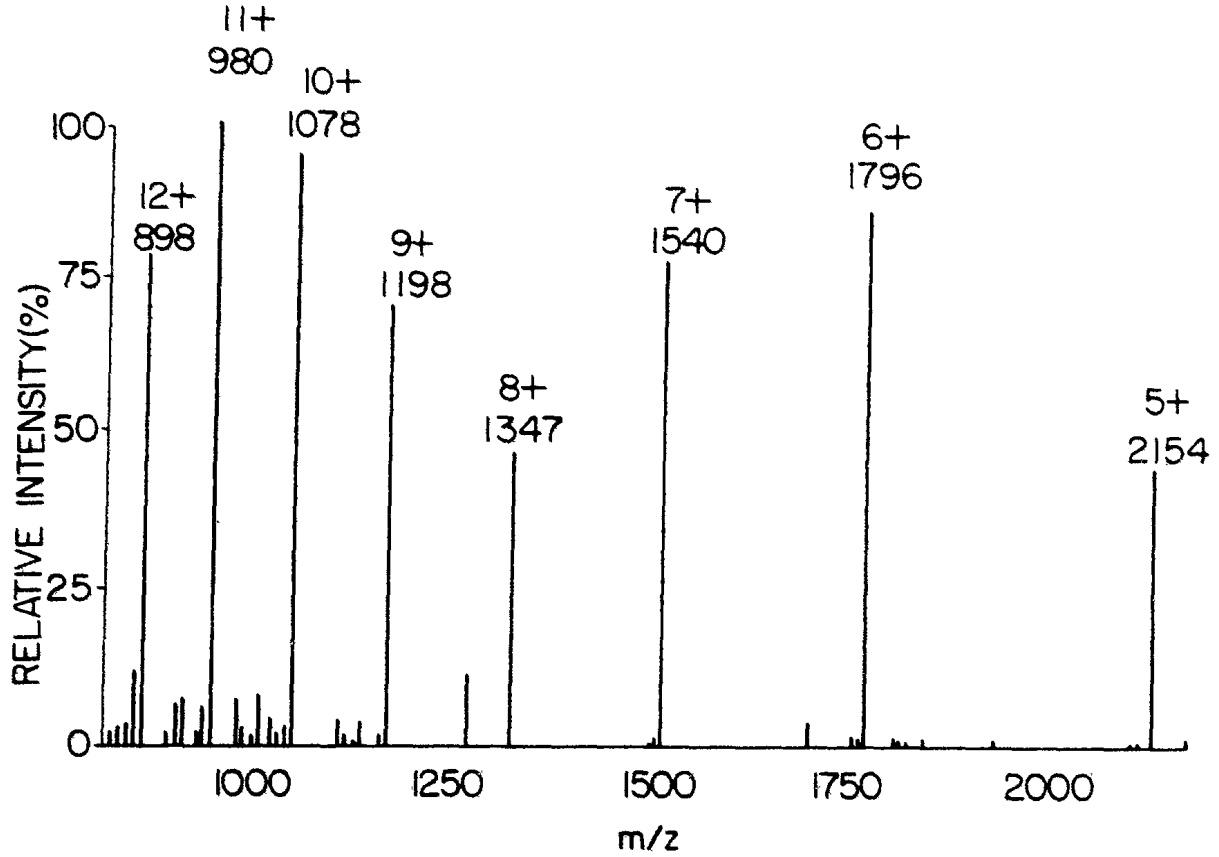
11+
980
10+
1078
12+
898
9+
1198
8+
1347
7+
1540
6+
1796
5+
2154
100
75
50
25
0
RELATIVE INTENSITY(%)
1000
1250
1500
1750
2000
m/z

FIG. 3A

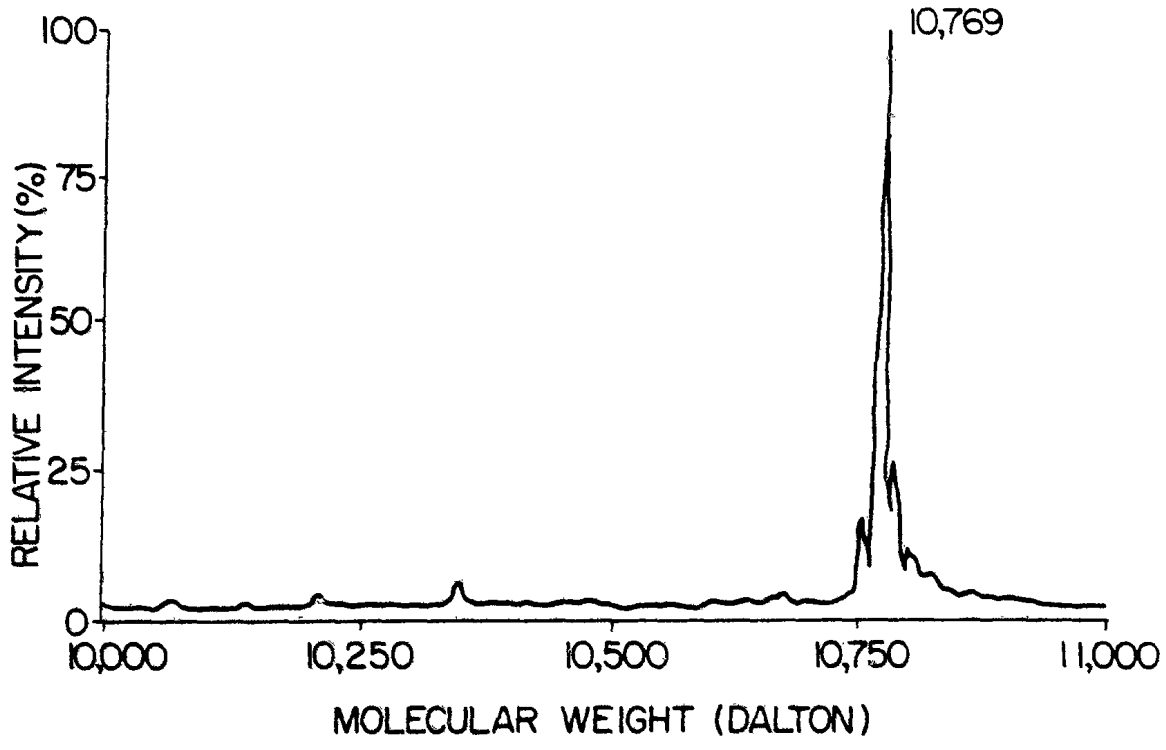
10,769
100
75
50
25
0
RELATIVE INTENSITY (%)
10,000
10,250
10,500
10,750
11,000
MOLECULAR WEIGHT (DALTON)

FIG. 3B

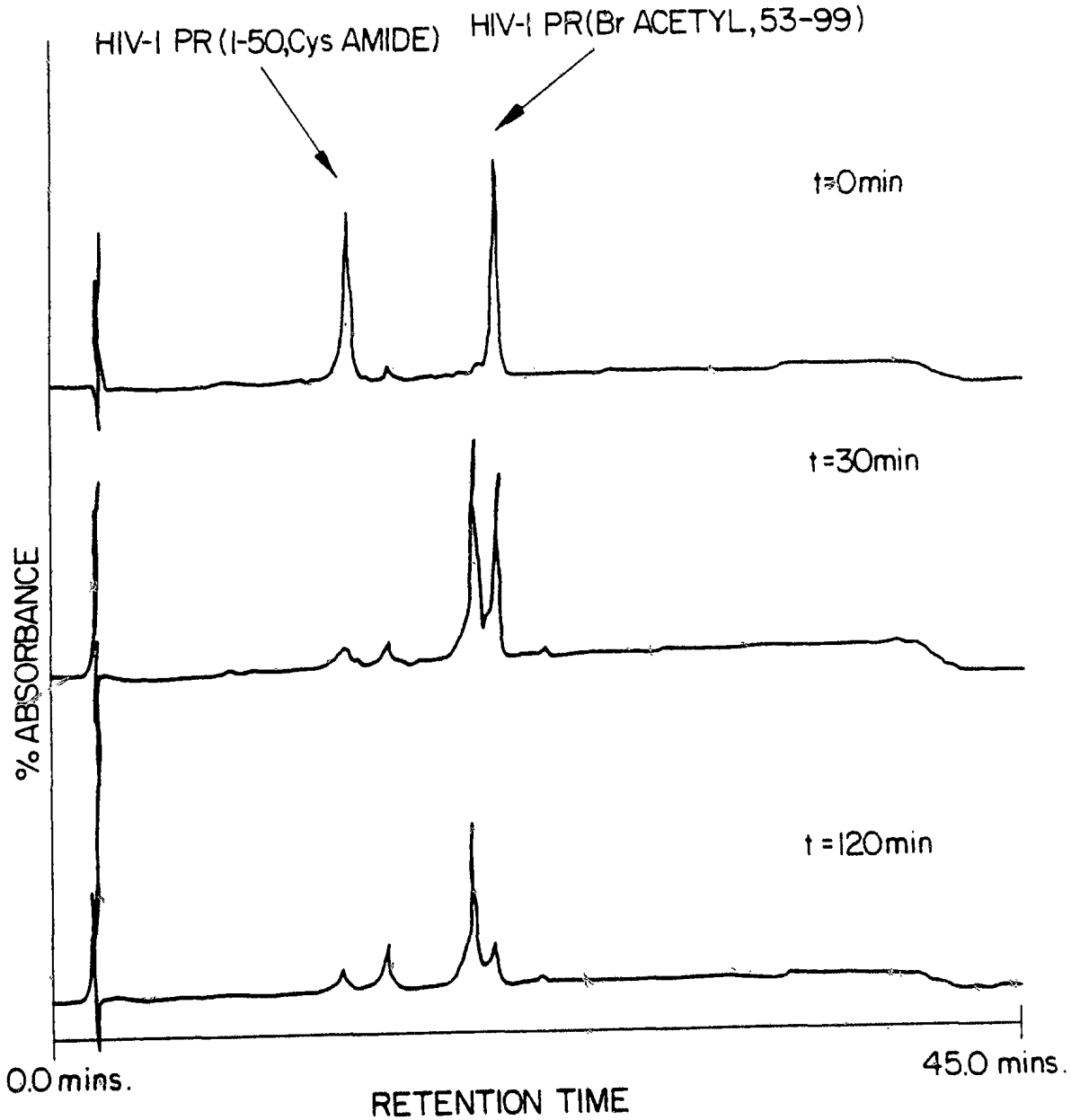
HIV-1 PR (1-50,Cys AMIDE)
HIV-1 PR(Br ACETYL,53-99)
t=0min
t=30min
t=120min
% ABSORBANCE
0.0 mins.
45.0 mins.
RETENTION TIME

FIG. 4

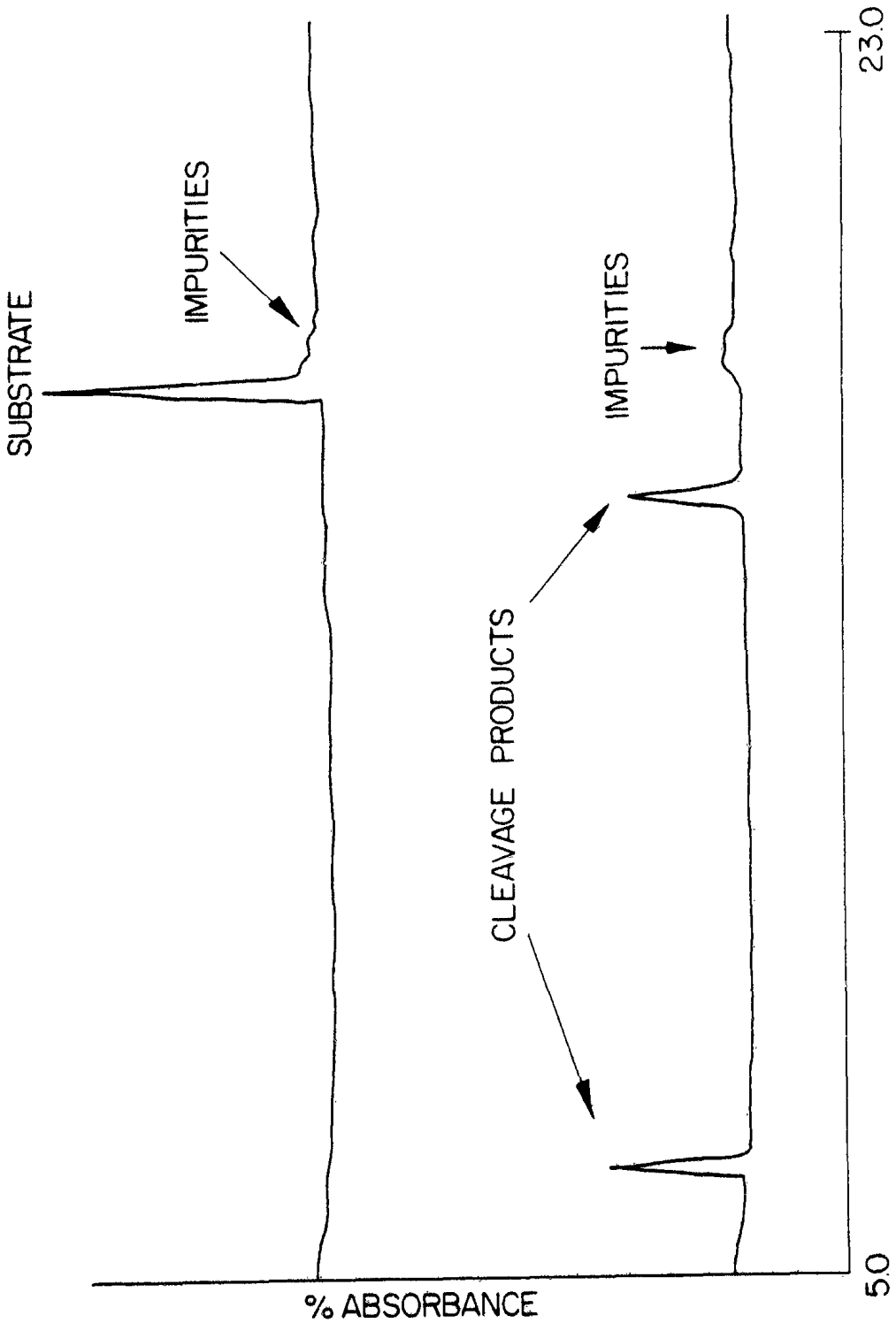
SUBSTRATE
IMPURITIES
CLEAVAGE PRODUCTS
IMPURITIES
23.0
5.0
% ABSORBANCE
RETENTION TIME (min)

FIG. 5

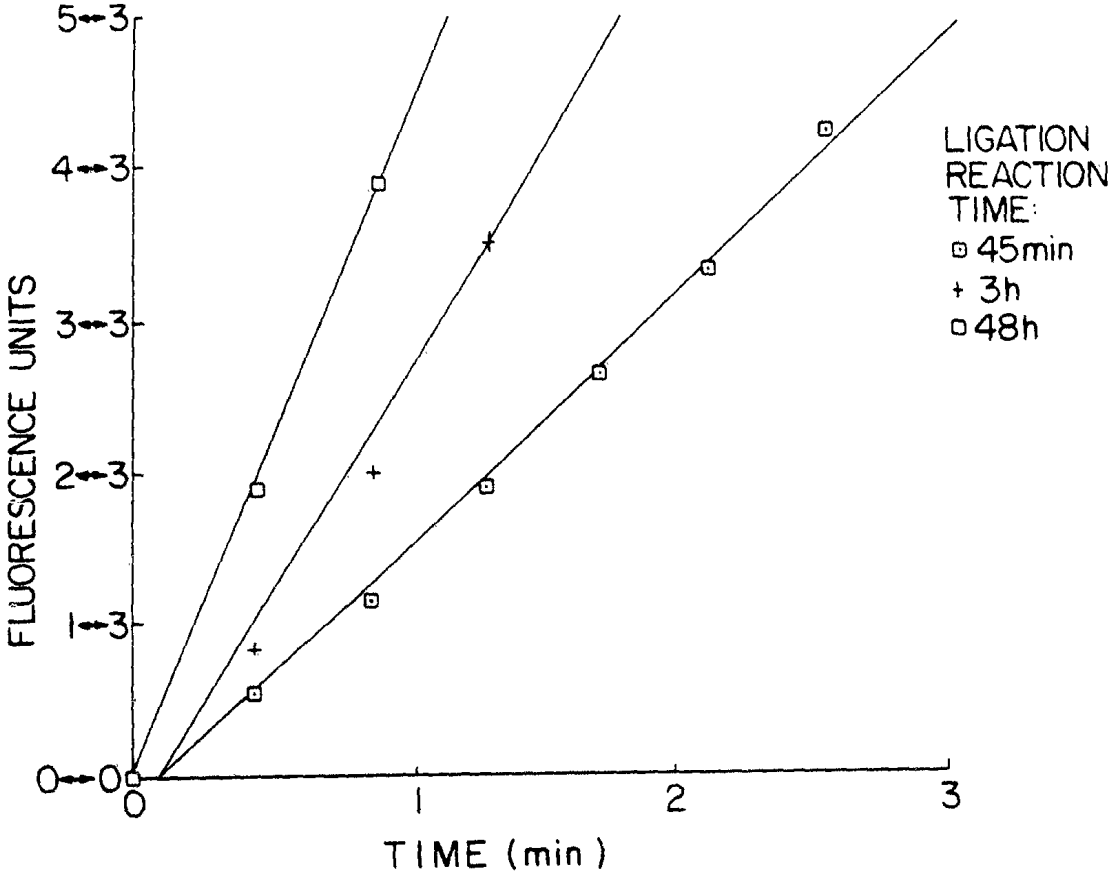
5→3
4→3
3→3
2→3
1→3
0→0
FLUORESCENCE UNITS
0
1
2
3
TIME (min)
LIGATION REACTION TIME:
45min
3h
48h

FIG. 6

INTERNATIONAL SEARCH REPORT

International application No
PCT/US93/02846

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :C 07 K 3/04
US CL :530/333

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/333

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CAS ONLINE, MEDLINE, APS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SCIENCE, volume 255, issued 10 January 1992, J.A. Ellman et al., "Site-Specific Incorporation of Novel Backbone Structures into Proteins", pages 197-200, see entire document.	1-38
A	SCIENCE, volume 244, issued 14 April 1989, C.N. Noren et al., "A General Method for Site-Specific Incorporation of Unnatural Amino Acids into Proteins", pages 182-189, see entire document.	1-38

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 29 JUNE 1993	Date of mailing of the international search report 08 JUL 1993
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Authorized officer CAROL A. SALATA, PH.D. Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Proceedings of the National Academy of Sciences, volume 85, issued July 1988, N.E. Kohl et al., "Active human immunodeficiency virus protease is required for viral infectivity", pages 4686-4690, see entire document.	1-38
A	Tetrahedron Letters, volume 31, Number 32, issued 1990, P. Mascagni et al., "Protein Engineering of HIV Viral Proteins by Total Chemical Synthesis: the C-Terminal 104 Residue Peptide GAG p24", pages 4637-4640, see entire document.	1-38
A,P	SCIENCE, volume 256, issued 10 April 1992, M. Schnolzer, "Constructing Proteins by Dovetailing Unprotected Synthetic Peptides: Backbone-Engineered HIV Protease", pages 221-225, see entire document.	1-38
A	NATURE, volume 352, issued 08 August 1991, K.L. Beattie et al. "Solid Phase Gene Assembly", pages 548-549, see entire document.	1-38