

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
21 April 2005 (21.04.2005)

PCT

(10) International Publication Number
WO 2005/035552 A2

(51) International Patent Classification⁷: **C07K**
(21) International Application Number:
PCT/US2004/033278
(22) International Filing Date: 6 October 2004 (06.10.2004)
(25) Filing Language: English
(26) Publication Language: English
(30) Priority Data:
60/509,601 7 October 2003 (07.10.2003) US

(71) Applicant (for all designated States except US): **GENZYME CORPORATION** [US/US]; 500 Kendall Street, Cambridge, MA 02142 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **NICOLETTE, Charles, A.** [US/US]; 4 Mill Street, Framingham, MA 01701 (US).

(74) Agent: **KONSKI, Antoinette, F.**; Bingham McCutchen LLP, Three Embarcadero Center, Suite 1800, San Francisco, CA 94111-4067 (US).

(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, 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, IT, 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:

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

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

(54) Title: TRANSDUCING COMBINATORIAL PEPTIDE LIBRARY

Representative Structure of Members of Peptide Library

Construction of D-peptide Library



Complexity ~ 90 billion

X=A(2)*, G(3)*, E(1), K(1), L(1), P(1), Q(1), R(1), W(1)

*molar ratios

* Fluorescent label

(57) Abstract: This invention provides methods for isolating bio-active molecules from complex, rationally designed peptide libraries with a reduction in assay time and increase in throughput. The molecules have the ability to facilitate transport across biological membranes. The methods also identify bio-active peptides of therapeutic value such as vaccines in treating cancers, viral diseases, and autoimmune diseases, as well as to identify useful clinical diagnostic reagents.

TRANSDUCING COMBINATORIAL PEPTIDE LIBRARY

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority under 35 U.S.C. § 119 (e) to U.S. Provisional Application Serial No. 60/509,601, filed October 7, 2003, the contents of which are
5 hereby incorporated by reference into the present disclosure.

TECHNICAL FIELD

The present invention relates generally to methods for selecting peptides with good pharmacokinetic properties.

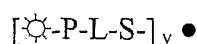
BACKGROUND

10 One of the drawbacks to using peptide libraries as opposed to small molecule libraries for drug discovery is that *in vivo*, peptides tend to have poor pharmacokinetics. Peptides are rapidly degraded upon importation into the interior of the cell by endocytosis or other forms of receptor-mediated uptake. The rapid degradation accounts in part for their poor bioavailability. Thus, peptides, in general,
15 are not well suited to modulating intracellular targets.

A need exists to prepare targeted peptide libraries for use in drug discovery, wherein all members of the libraries have been pre-selected for acceptable pharmacokinetic properties. This invention satisfies this need and provides related advantages as well.

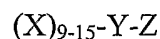
20 DISCLOSURE OF THE INVENTION

This invention provides a random peptide library comprising compounds of the structure:



wherein ⊙ is absent or optionally a fluorescent label,

P is a D-peptide comprised of the structure:



wherein each X is a D-amino acid subunit which varies independently from any other subunit and is an amino acid selected from the group A, G, E, K, L, P, Q, R and W,

wherein Y is a neutral amino acid subunit which varies independently from any other subunit,

wherein Z is a basic amino acid subunit which is the same for each member of the peptide library,

wherein L is a linker,

wherein S is absent or is optionally biotin,

wherein y is a plurality, and

• is a solid support comprising multiple copies of a single species of P.

In one aspect, the random peptide contains a P, wherein P is prepared by a method comprising the step of:

randomly introducing a D-amino acid subunit for each of the solid phase supports wherein the relative molar ratio of each D-amino acid subunit is provided according to the formula:

$$A=2,$$

$$G=3,$$

$$E, K, L, P, Q, R \text{ and } W = 1.$$

In yet another embodiment, the random peptide library contains the maximum number of possible P species.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows a representative structure of members of a peptide library of the invention.

Figure 2 shows a representative structure of members of a peptide library that selects peptides that target intracellular components.

Figure 3 (comprising three panels) are scans of targeted peptides identified using a biotinylated library.

MODES FOR CARRYING OUT THE INVENTION

Throughout this disclosure, various publications, patents and published patent specifications are referenced by an identifying citation. The disclosures of these publications, patents and published patent specifications are hereby incorporated by
5 reference into the present disclosure to more fully describe the state of the art to which this invention pertains.

As used herein, certain terms have the following defined meanings.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of immunology, molecular biology, microbiology, cell
10 biology and recombinant DNA, which are within the skill of the art. See, *e.g.*, Sambrook, Fritsch and Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd edition (1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F. M. Ausubel, et al. eds., (1987)); the series METHODS IN ENZYMOLOGY (Academic Press, Inc.): PCR 2: A PRACTICAL APPROACH (M.J. MacPherson, B.D. Hames and G.R. Taylor eds.
15 (1995)), Harlow and Lane, eds. (1988) ANTIBODIES, A LABORATORY MANUAL, and ANIMAL CELL CULTURE (R.I. Freshney, ed. (1987)).

As used in the specification and claims, the singular form “a”, “an” and “the” include plural references unless the context clearly dictates otherwise. For example, the term “a cell” includes a plurality of cells, including mixtures thereof.

20 As used herein, the term “comprising” is intended to mean that the compositions and methods include the recited elements, but not excluding others. “Consisting essentially of” when used to define compositions and methods, shall mean excluding other elements of any essential significance to the combination. Thus, a composition consisting essentially of the elements as defined herein would
25 not exclude trace contaminants from the isolation and purification method and pharmaceutically acceptable carriers, such as phosphate buffered saline, preservatives, and the like. “Consisting of” shall mean excluding more than trace elements of other ingredients and substantial method steps for administering the compositions of this invention. Embodiments defined by each of these transition terms are within the
30 scope of this invention.

The term “polypeptide” is used synonymously with “peptide” and each is used in its broadest sense to refer to a compound of two or more subunit amino acids,

amino acid analogs, or peptidomimetics. The subunits may be linked by peptide bonds. In another embodiment, the subunit may be linked by other bonds, *e.g.* ester, ether, *etc.* As used herein the term "amino acid" refers to either natural and/or unnatural or synthetic amino acids, including glycine and both the D or L optical isomers, and amino acid analogs and peptidomimetics. A peptide of three or more amino acids is commonly called an oligopeptide if the peptide chain is short. If the peptide chain is long, the peptide is commonly called a polypeptide or a protein.

The term "isolated" means separated from constituents, cellular and otherwise, in which the polynucleotide, peptide, polypeptide, protein, antibody, or fragments thereof, are normally associated with in nature. In one aspect of this invention, an isolated polynucleotide is separated from the 3' and 5' contiguous nucleotides with which it is normally associated with in its native or natural environment, *e.g.*, on the chromosome. As is apparent to those of skill in the art, a non-naturally occurring polynucleotide, peptide, polypeptide, protein, antibody, or fragments thereof, does not require "isolation" to distinguish it from its naturally occurring counterpart. In addition, a "concentrated", "separated" or "diluted" polynucleotide, peptide, polypeptide, protein, antibody, or fragments thereof, is distinguishable from its naturally occurring counterpart in that the concentration or number of molecules per volume is greater than "concentrated" or less than "separated" than that of its naturally occurring counterpart. A polynucleotide, peptide, polypeptide, protein, antibody, or fragments thereof, which differs from the naturally occurring counterpart in its primary sequence or for example, by its glycosylation pattern, need not be present in its isolated form since it is distinguishable from its naturally occurring counterpart by its primary sequence, or alternatively, by another characteristic such as glycosylation pattern. Thus, a non-naturally occurring polynucleotide or peptide is provided as a separate embodiment from the isolated naturally occurring polynucleotide or peptide. A protein produced in a bacterial cell is provided as a separate embodiment from the naturally occurring protein isolated from a eucaryotic cell in which it is produced in nature.

The terms "polynucleotide" and "oligonucleotide" are used interchangeably, and refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof. Polynucleotides may have any three-dimensional

structure, and may perform any function, known or unknown. The following are non-limiting examples of polynucleotides: a gene or gene fragment, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated
5 DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs. If present, modifications to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified
10 after polymerization, such as by conjugation with a labeling component. The term also refers to both double- and single-stranded molecules. Unless otherwise specified or required, any embodiment of this invention that is a polynucleotide encompasses both the double-stranded form and each of two complementary single-stranded forms known or predicted to make up the double-stranded form.

15 As used herein, "solid phase support" or "solid support", used interchangeably, is not limited to a specific type of support. Rather a large number of supports are available and are known to those of skill in the art. Solid phase supports include silica gels, resins, derivatized plastic films, glass beads, cotton, plastic beads, alumina gels. As used herein, "solid support" also includes synthetic antigen-
20 presenting matrices, cells, and liposomes. A suitable solid phase support may be selected on the basis of desired end use and suitability for various protocols. For example, for peptide synthesis, solid phase support may refer to resins such as polystyrene (*e.g.*, PAM-resin obtained from Bachem Inc., Peninsula Laboratories, *etc.*), POLYHIPE® resin (obtained from Aminotech, Canada), polyamide resin
25 (obtained from Peninsula Laboratories), polystyrene resin grafted with polyethylene glycol (TentaGel®, Rapp Polymere, Tubingen, Germany) or polydimethylacrylamide resin (obtained from Milligen/Biosearch, California).

As used herein, "expression" refers to the process by which polynucleotides are transcribed into mRNA and/or the process by which the transcribed mRNA is
30 subsequently being translated into peptides, polypeptides, or proteins. If the polynucleotide is derived from genomic DNA, expression may include splicing of the mRNA in an eukaryotic cell.

“Hybridization” refers to a reaction in which one or more polynucleotides react to form a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues. The hydrogen bonding may occur by Watson-Crick base pairing, Hoogsteen binding, or in any other sequence-specific manner. The complex may comprise two strands forming a duplex structure, three or more strands forming a multi-stranded complex, a single self-hybridizing strand, or any combination of these. A hybridization reaction may constitute a step in a more extensive process, such as the initiation of a PCR reaction, or the enzymatic cleavage of a polynucleotide by a ribozyme.

Hybridization reactions can be performed under conditions of different “stringency”. In general, a low stringency hybridization reaction is carried out at about 40 °C in 10 X SSC or a solution of equivalent ionic strength/temperature. A moderate stringency hybridization is typically performed at about 50 °C in 6 X SSC, and a high stringency hybridization reaction is generally performed at about 60 °C in 1 X SSC.

When hybridization occurs in an antiparallel configuration between two single-stranded polynucleotides, the reaction is called “annealing” and those polynucleotides are described as “complementary”. A double-stranded polynucleotide can be “complementary” or “homologous” to another polynucleotide, if hybridization can occur between one of the strands of the first polynucleotide and the second. “Complementarity” or “homology” (the degree that one polynucleotide is complementary with another) is quantifiable in terms of the proportion of bases in opposing strands that are expected to form hydrogen bonding with each other, according to generally accepted base-pairing rules.

A polynucleotide or polynucleotide region (or a polypeptide or polypeptide region) has a certain percentage (for example, 80%, 85%, 90%, or 95%) of “sequence identity” to another sequence means that, when aligned, that percentage of bases (or amino acids) are the same in comparing the two sequences. This alignment and the percent homology or sequence identity can be determined using software programs known in the art, for example those described in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F.M. Ausubel et al., eds., 1987) Supplement 30, section 7.7.18, Table 7.7.1. Preferably, default parameters are used for alignment. One

alignment program is BLAST, using default parameters. Other programs are BLASTN and BLASTP, using the following default parameters: Genetic code = standard; filter = none; strand = both; cutoff = 60; expect = 10; Matrix = BLOSUM62; Descriptions = 50 sequences; sort by = HIGH SCORE; Databases = non-redundant,
 5 GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + SwissProtein + SPupdate + PIR. Details of these programs can be found at the following Internet address: ncbi.nlm.nih.gov/cgi-bin/BLAST.

All numerical designations, *e.g.*, pH, temperature, time, concentration, and molecular weight, including ranges, are approximations which may be varied (+) or (-) by increments of 0.1. It is to be understood, although not always explicitly stated
 10 that all numerical designations are preceded by the term “about”. It also is to be understood, although not always explicitly stated, that the reagents described herein are merely exemplary and that equivalents of such are known in the art.

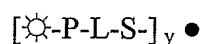
As used herein, the term “transducing domain” identifies a linear sequence of amino acids present in a polypeptide that has been shown to enhance the transport of the molecule across a biological membrane.
 15

A “composition” is intended to mean a combination of active agent and another compound or composition, inert (for example, a detectable agent or label) or active, such as an adjuvant.

20 A “pharmaceutical composition” is intended to include the combination of an active agent with a carrier, inert or active, making the composition suitable for diagnostic or therapeutic use *in vitro*, *in vivo* or *ex vivo*.

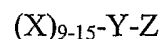
As used herein, the term “pharmaceutically acceptable carrier” encompasses any of the standard pharmaceutical carriers, such as a phosphate buffered saline
 25 solution, water, and emulsions, such as an oil/water or water/oil emulsion, and various types of wetting agents. The compositions also can include stabilizers and preservatives. For examples of carriers, stabilizers and adjuvants, see Martin, REMINGTON’S PHARM. SCI., 15th Ed. (Mack Publ. Co., Easton (1975)).

This invention provides a random peptide library comprising compounds of
 30 the structure:



wherein X is absent or optionally a fluorescent label,

P is a D-peptide comprised of the structure:



wherein each X is a D-amino acid subunit which varies independently from any other subunit and is an amino acid selected from the group A, G, E, K, L, P, Q, R and W,

wherein Y is a neutral amino acid subunit which varies independently from any other subunit,

wherein Z is a basic amino acid subunit which is the same for each member of the peptide library,

wherein L is a linker,

wherein S is absent (*i.e.*, L is then directly bound to the bead) or is optionally biotin, or its equivalent,

wherein y is the number (*i.e.*, a plurality of identical species that are present on the bead, and

• is a solid support comprising multiple copies of a single species of P.

In one aspect, the random peptide contains a P, wherein P is prepared by randomly introducing a D-amino acid subunit for each of the solid phase supports wherein the relative molar ratio of each D-amino acid subunit is provided according to the formula:

A=2,

G=3,

E, K, L, P, Q, R and W = 1.

In one aspect, the present invention provides a combinatorial library of non-hydrolysable peptide polymers coupled to a non-hydrolysable peptide transduction domain. By "combinatorial library" is meant a collection of members. The combinatorial library may or may not include all possible members of the defined set. Libraries of this invention overcome poor pharmacokinetics, such as rapid degradation and inability to access intracellular targets.

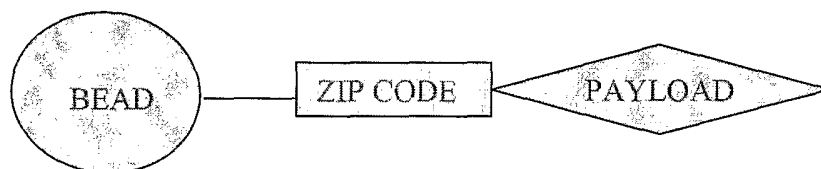
Various solid phase supports (also termed herein "beads") can be used in the practice of the invention. Such solid phase supports must be compatible with the biological assay to be performed, and must be inert to the synthesis of the molecule, *e.g.*, peptide, and if present, a coding molecule. Examples of solid phase supports are

provided above. Additional examples include polystyrene resin, poly(dimethylacryl)amide-grafted styrene-co-divinylbenzene resin, polyamide resin, polystyrene resin grafted with polyethylene glycol, and polydimethylacrylamide resin.

The releasable linker may release upon exposure to an acid, a base, a nucleophile, an electrophile, light, an oxidizing agent, a reducing agent, or an enzyme.

In another aspect, this invention provides peptide libraries that can home or target different intracellular structure. Briefly, a Primary (1°) library identifies the “Transducing Peptides” (also termed “zip codes” herein) that can “home” or “target” to different intracellular structures.

After identification of the Transducing Peptide sequence (“zip code”) of interest, for example, a peptide that targets a particular intracellular structure, then a Secondary (2°) library can be prepared where each member/specie (“candidate sequence”/“payload”) has a fixed structure (sequence) of the zip code (as identified in the Primary library) covalently or non-covalently attached thereto. In constructing the 2° library, a single “Zip Code” sequence is attached to every bead in the library by a linker. The individual members of the library (“candidate payloads”) are then (randomly) synthesized onto the bead-zip code in the same manner as one would normally synthesize a library onto a bead.



In yet one embodiment, the random peptide library contains the maximum number of possible P species.

The amount of peptide on a bead can vary, *e.g.*, from 100 or more picomoles, or alternatively 125 or more picomoles, or alternatively 150 or more picomoles, or alternatively 175 or more picomoles, or alternatively 200 or more picomoles, or alternatively 225 or more picomoles, or alternatively 250 or more picomoles, or alternatively 275 or more picomoles, or alternatively 300 or more picomoles, per bead. One can increase or decrease the amount of peptide on the bead/support, but such modifications would be dependent upon the size of the bead.

The size of the peptide will vary, *e.g.*, 9-15 amino acids in length, or alternatively 12 –17 amino acids in length, or alternatively 13-15 amino acids in length, or alternatively less than 11 amino acids, *e.g.*, 9 amino acids or 10 amino acids, each with molar ratios.

5 In one embodiment, the design of the library or its components have amino acids having positive residues (+). The amino acids lysine “K” and arginine “R” are examples of amino acids having positively charged residues which can assist peptide transport across the cell membrane. In another embodiment, the amino acids having negative residues such as glutamic acid “E” or glutamine “Q”. In another
10 embodiment, the amino acids have limited hydrophobic residues, *e.g.*, leucine “L” or tryptophan “W”.

 In one aspect, the combinatorial library consists of polystyrene beads with each bead containing a minimum 200 pmols of a single (*i.e.*, identical to each other) library specie. A completely degenerate library using 19 amino acids (cysteine
15 excluded for synthetic reasons) far exceeds this synthesis and screening capacity. To help reduce the complexity, the degeneracy can be from about 9 to 15 different amino acids. In one aspect, these amino acids are strategically chosen to bias the library towards cationic residues.

 This invention also provides combinatorial libraries wherein each library
20 specie has the fixed structure of a transduction domain covalently attached to a variable domain (See Figure 2). This library identifies members having pharmacokinetic properties more similar to small molecule drugs (rather than peptides) except that up to 100 million peptides can be manufactured and screened in about 2 weeks.

25 In a further aspect, this invention provides methods to identify transducing peptides that target to different locations within the cell as well as peptides having the selected targeting properties. Using a choice of transducing domains that can target different intracellular structures allows semi-rational library or screen design to meet specific needs. For example, by building in a pre-selected targeting domain, one can
30 produce and screen molecules that interact with the pre-selected target, (*e.g.*, something in the nucleus) to carry the library payload straight into the nucleus.

The transducing peptides can be used to target much larger payloads to an appropriate "zip code" within the cell. For example, a peptide that specifically targets lysosomes could be conjugated to lysosomal storage disease replacement enzymes such as CerazymeTM or to prepare and screen peptides capable of effortlessly crossing the blood-brain barrier, a major problem for many existing drugs.

The peptides generated by the invention can be used in various assays that in turn can identify molecules that: (1) can cross blood/brain barrier, (2) can target any intracellular organelle or compartment, (3) molecules that target the nucleus, (4) can target lysosomes (protein replacement therapy) or (5) can target endosomes.

The libraries can also be modified to identify peptides with specific functional properties, *i.e.*, cancer drugs (*e.g.*, cytotoxic compounds specific to tumor cells, growth inhibitory compounds specific to tumor cells, and oncogene inhibitors (*e.g.*, nuclear p53, H-Ras^{VAL12}, *etc.*), cancer immunomodulatory drugs (*e.g.*, compounds that specifically block selected cytokine production/secretion and TGF- β inhibitors), autoimmune disorder drugs (*e.g.*, compounds that specifically block antibody production/secretion from B cells and compounds that block or stimulate cytokine production/secretion), infectious disease drugs (*e.g.*, antifungal compounds, antibiotic compounds and antiviral compounds).

In one aspect, using a combinatorial library of this invention, compounds can be identified that are capable of crossing the plasma membrane of cells with ultra-fast kinetics, indeed, too fast to be receptor-mediated. These represent cell transducing peptides that are considerably better at crossing the membrane than the known transducing peptide derived from the HIV Tat protein. (see Leifert, et al. (2001) Human Gene Therapy 12:1881-1892). For example, the screened library can have the structure shown in Figure 1.

The screening of all libraries described here can be performed in a routine fashion. Briefly, library beads are arrayed in predetermined pools of up to 10,000/pool. A portion of peptide is cleaved from each bead providing discreet pools of solution-phase peptides that can be tested in any desired biological assay. Iterations of this process directed at producing and reassaying progressively smaller peptide pools identifies individual beads containing bioactive species. Finally,

residual bead-bound peptide can be sequenced by standard Edman degradation or by MALDI-TOF Post-source decay.

Materials and Methods

Peptide Libraries: The procedure provided below is an example to prepare
5 libraries of this invention. Any one of the many combinatorial library technologies described to date can also be employed in the practice of the present invention, including but not limited to synthetic combinatorial peptide or molecule libraries (Needels et al. (1993) Proc. Natl. Acad. Sci. USA 90:10700-4; Ohlmeyer et al. (1993) Proc. Natl. Acad. Sci. USA 90:10922-10926; Lam et al., International Patent
10 Publication No. WO 92/00252; Lebl et al., International Patent Publication No. WO 94/28028, published December 8, 1994, each of which is incorporated herein by reference in its entirety).

In one aspect, the invention employs the solid phase library technique described by Ohlmeyer et al. (1993) Proc. Natl. Acad. Sci. USA 90:10922-26, which
15 is incorporated herein by reference in its entirety. Halogen substituted benzenes linked to tag-linker tert-butyl esters constitute the inert molecular tags that encode the sequence of the unique peptide co-synthesized on any given bead in the library. A brief description of the tag synthesis, peptide synthesis, and encoding/decoding strategy is presented below.

20 Briefly, the molecular tags used as encoding molecules are precipitated from dimethylformamide (DMF) containing 8-bromo-1-octanol and 2,4,6-trichlorophenol by the addition of cesium carbonate. The solution is then heated to 80°C. for 2 hours, washed with 0.5M NaOH, 1M HCl, and finally H₂O, at which point the organic phase is evaporated. The resulting tag alcohol is a colorless oil.

25 The tag alcohol is then added to a 2 M solution of phosgene (in toluene) to produce a crude chloroformate. After evaporation of the solvent, the compound is dissolved in CH₂Cl₂ and pyridine and incubated with tert-Butyl 4-(hydroxymethyl)-3-nitrobenzoate. The resulting tag-linker tert-butyl ester is isolated from the organic phase and purified by chromatography with the product being a clear oil. To generate
30 unique tags, halogen-substituted benzene compounds are reacted with the electrophoric tag. Each derivative has a different gas chromatography retention time. This property confers the ability to encode the unique synthesis history of individual

beads. Electron capture capillary gas chromatography can selectively detect the tags at levels less than 1 pmol.

The peptides can be synthesized on Merrified resin beads (or other suitable resin) such that the peptides are linked by photocleavable crosslinkers by a typical split-synthesis method. As each amino acid is added, a corresponding mixture of acyl carbonate-activated linker tag acids is co-ligated, but with a linker which is not photocleavable. This allows release of the peptide with retention of the coding molecules during the screening procedure. The combination of tag molecules added at each step corresponds to the specific amino acid residue added in that step, thus serving as a record of the synthetic history of any given bead.

When a bead of interest is identified, the sequence of the peptide that was synthesized on it can be deduced by any method known in the art. For the purpose of illustration only, one method requires loading the bead into a Pyrex capillary tube and washed with DMF. It is then suspended in 1 μ l DMF and sealed in the capillary tube and irradiated to release the tag alcohols. The capillary tube is then opened and the tag alcohols are trimethylsilylated with bis(trimethylsilyl) acetamide. The solution above the bead is then injected into an electron capture, capillary gas chromatograph for analysis. The resulting profile of tag elution on the gas chromatogram allows the amino acid sequence of the co-synthesized peptide to be directly determined.

In a specific aspect of the invention, the peptides of a library may comprise a special amino acid at the C-terminus which incorporates either a CO₂H or CONH₂ side chain to simulate a free glycine or a glycine-amide group. Another way to consider this special residue would be as a D or L amino acid analog with a side chain consisting of the linker or bound to the bead. In one embodiment, the pseudo-free C-terminal residue may be of the D or the L optical configuration; in another embodiment, a racemic mixture of D and L-isomers may be used.

In an additional embodiment, pyroglutamate may be included as the N-terminal residue of the peptides of the library. Although pyroglutamate is not amenable to sequence by Edman degradation, identification of the peptide sequence can be accomplished by a coded library strategy, or by limiting substitution to only 50% of the peptides on a given bead with N-terminal pyroglutamate, thus leaving enough non-pyroglutamate peptide on the bead for direct sequencing. One of

ordinary skill would readily recognize that this technique could be used for sequencing of any peptide that incorporates a residue resistant to Edman degradation at the N-terminus. Specific activity of a peptide that comprises a blocked N-terminal group, *e.g.*, pyroglutamate, when the particular N-terminal group is present in 50% of the peptides would readily be demonstrated by comparing activity of a completely
5 (100%) blocked peptide with a non-blocked (0%) peptide.

Non-classical amino acids that induce conformational constraints. The following non-classical amino acids may be incorporated in the peptide library in order to introduce particular conformational motifs: 1,2,3,4-tetrahydroisoquinoline-3-carboxylate (Kazmierski et al., 1991, J. Am. Chem. Soc. 113:2275-2283); (2S,3S)-
10 methyl-phenylalanine, (2S,3R)-methyl-phenylalanine (2R,3S)-methyl-phenylalanine and (2R,3R)-methyl-phenylalanine (Kazmierski and Hruby, 1991, Tetrahedron Lett.); 2-aminotetrahydronaphthalene-2-carboxylic acid (Landis, 1989, Ph.D. Thesis, University of Arizona); hydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylate
15 (Miyake et al., 1989, J. Takeda Res. Labs. 43:53-76); .beta.-carboline (D and L) (Kazmierski, 1988, Ph.D. Thesis, University of Arizona); HIC (histidine isoquinoline carboxylic acid) (Zechel et al., 1991, Int. J. Pep. Protein Res. 43); and HIC (histidine cyclic urea) (Dharanipragada).

The following amino acid analogs and peptidomimetics may be incorporated
20 into a library to induce or favor specific secondary structures: LL-Acp (LL-3-amino-2-propenidone-6-carboxylic acid), a .beta.-turn inducing dipeptide analog (Kemp et al., 1985, J. Org. Chem. 50:5834-5838); .beta.-sheet inducing analogs (Kemp et al., 1988, Tetrahedron Lett. 29:5081-5082); .beta.-turn inducing analogs (Kemp et al., 1988, Tetrahedron Lett. 29:5057-5060); .alpha.-helix inducing analogs (Kemp et al.,
25 1988, Tetrahedron Lett. 29:4935-4938); .gamma.-turn inducing analogs (Kemp et al., 1989, J. Org. Chem. 54:109-115); and analogs provided by the following references: Nagai and Sato, 1985, Tetrahedron Lett. 26:647-650; DiMaio et al., 1989, J. Chem. Soc. Perkin Trans. P. 1687; also a Gly-Ala turn analog (Kahn et al., 1989, Tetrahedron Lett. 30:2317); amide bond isostere (Jones et al., 1988, Tetrahedron Lett.
30 29:3853-3856); tetrazole (Zabrocki et al., 1988, J. Am. Chem. Soc. 110:5875-5880); DTC (Samanen et al., 1990, Int. J. Protein Pep. Res. 35:501-509); and analogs taught in Olson et al., 1990, J. Am. Chem. Sci. 112:323-333 and Garvey et al., 1990, J. Org.

Chem. 56:436. Conformationally restricted mimetics of beta turns and beta bulges, and peptides containing them, are described in U.S. Pat. No. 5,440,013, issued Aug. 8, 1995 to Kahn.

5 Determination of the sequence of peptides that incorporate such non-classical amino acids is readily accomplished by the use of a coded library. Alternatively, a combination of initial Edman degradation followed by amino acid analysis of the residual chain can be used to determine the structure of a peptide with desired activity. Mass spectral analysis can also be employed.

10 *Solid phase supports and linkers.* A solid phase support for use in the present invention will be inert to the reaction conditions for synthesis. A solid phase support for use in the present invention must have reactive groups in order to attach a monomer subunit, or for attaching a linker or handle which can serve as the initial binding point for a monomer subunit. In one embodiment, the solid phase support may be suitable for *in vivo* use, *i.e.*, it may serve as a carrier for or support for direct
15 applications of the library (*e.g.*, TENTAGEL.RTM., Rapp Polymere, Tubingen, Germany).

As used herein, solid phase support is not limited to a specific type of support. Rather a large number of supports are available and are known to one of skill in the art. Solid phase supports include silica gels, resins, derivatized plastic films, glass
20 beads, cotton, plastic beads, alumina gels. A suitable solid phase support may be selected on the basis of desired end use and suitability for various synthetic protocols. For example, for peptide synthesis, solid phase support may refer to resins such as polystyrene (*e.g.*, PAM-resin obtained from Bachem Inc., Peninsula Laboratories, *etc.*), POLYHIPE.RTM. resin (obtained from Aminotech, Canada), polyamide resin
25 (obtained from Peninsula Laboratories), polystyrene resin grafted with polyethylene glycol (TENTAGEL.RTM., Rapp Polymere, Tubingen, Germany) or polydimethylacrylamide resin (obtained from Milligen/Bioscience, Calif.). In one aspect for peptide synthesis, solid phase support refers to polydimethylacrylamide resin.

30 The solid phase supports of the invention also comprise a cleavable linker. As used herein, a cleavable linker refers to any molecule that provides spatial distance between the support and the peptide to be synthesized, and which can be cleaved to

provide for release of the peptide from the support into solution. Linkers can be covalently attached on the solid phase support prior to coupling with a N^{alpha}. -Boc or N^{alpha}. -Fmoc or otherwise appropriately protected amino acids. Various linkers can be used to attach the oligomer to solid phase support. Examples of spacer linkers
5 include aminobutyric acid aminocaproic acid, 7-aminoheptanoic acid, and 8-aminocaprylic acid. Fmoc-aminocaproic acid is commercially available from Bachem Biochem. In a further embodiment, linkers can additionally comprise one or more ()-alanines as spacers.

In addition, the solid-support could be modified to meet specific requirements
10 for the particular purpose of bioassay or detection. Modification of solid phase support may be made by incorporation of a specific linker. For example, modified solid phase support could be made acid-sensitive, base-sensitive, nucleophilic-sensitive, electrophilic sensitive, photosensitive, oxidation sensitive or reduction sensitive.

In addition to the linkers described above, selectively cleavable linkers may be employed. For example, an ultraviolet light sensitive linker, ONb, can be used (see Barany and Albericio (1985) J. Am. Chem. Soc. 107:4936-4942). Other cleavable linkers require hydrogenolysis or photolysis. Examples of photosensitive (photocleavable) linkers are found in Wang (1976) J.Org. Chem. 41:32-58, Hammer
20 et al. (1990) Int. J. Pept. Protein Res. 36:31-45; and Kreib-Cordonier et al. (1990) in PEPTIDES--CHEMISTRY. STRUCTURE AND BIOLOGY, Rivier and Marshall, eds., pp. 895-897. Landen (1977) Methods Enzym. 47:145-149, used aqueous formic acid to cleave Asp-Pro bonds; this approach has been used, to characterize T-cell determinants in conjunction with the Geysen pin synthesis method (Van der Zee et al.
25 (1989) Eur.J.Immunol. 191:4347). Other potential linker groups cleavable under basic conditions include those based on p-(hydroxylmethyl) benzoic acid (Atherton et al. (1981) J. Chem. Soc. Perkin 1:538-546) and hydroxyacetic acid (Baleaux et al. (1986) Int. J. Pept. Protein Res. 28:22-28). Geysen et al. ((1990) J. Immunol. Methods 134:23-33) reported peptide cleavage by a diketopiperazine mechanism. An
30 enzyme may specifically cleave a linker that comprises a sequence that is sensitive or a substrate for enzyme cleavage, e.g., protease cleavage of a peptide

In certain instances, one may derivatize 10-90% of the resin by substitution with the cleavable linker, and the remaining 90-10% substituted with a noncleavable linker to ensure that after cleavage of linker enough peptide will remain for sequencing. Preferably, however, a cleavable linker is used in combination with a coded library strategy. Combinations of cleavable linkers can also be used to allow sequential cleaving from a single bead.

Detection Methods: The following are procedures that can be used to identify peptides or species as described herein.

Biotinylated Detection Protocol

Reagents:

- GenPoint Biotinylated Tyramide Kit (DAKO, Cat.# K0620)
- Strept-Avidin-HRP1 (Boehringer Maneheim)
- Goat serum (Sigma)
- Strept-avidin-FITC (Molecular probes).
- LAMP-1 antibody (SANTA Cruz)
- Anti-Rabbit-Cy3 antibody (Jackson ImmunoResearch)
- TBST (Tris Buffered Saline/Tween-20) (DAKO, Cat.# S3306).
- Non-fat milk Block (Bio-rad, Cat#170-6404)
- PBS (Genzyme)
- TBS (DAKO)
- Universal Block (KKL, Cat.# 71-00-61)6404)
- DAPI/Anti-quench Mounting medium (Genzyme)

Peptide Blocking Buffer: (Made fresh day of staining):

- Peptide Block -50 mM Tris, pH 7.5, 300 mM NaCl, 0.15% Triton-X 100 with 0.5% blocking agent.
- Peptide Wash -50 mM Tris, pH 7.5, 300 mM NaCl, 0.15% Triton-X 100

Antibody(LAMP-1) Blocking Buffer: (Made fresh day of staining):

- Antibody Block -PBS/0.1 % saponin/5% Goat serum
- Antibody Wash -PBS/0.05% Tween-20

Procedure:**Biotinylated Peptide Detection:**

1. Obtain fixed cells in PBS.
2. Wash the cells 2X 5 minute each in PBS at room temperature.
- 5 3. Add universal block or methanol/0.3% Hydrogen Peroxide/distilled water, incubate 60 minutes at room temperature. (Block endogenous peroxidase)
4. Rinse in TBS at room temperature, then add peptide block and incubate at 37°C for 30 minutes.
- 10 5. Remove block by inverting the plate, add 0.20 ml primary Streptavidin-HRP, diluted in peptide block, incubate at 37°C for 60 minutes.
6. Wash 2 times for 10 minutes each in peptide wash at 42°C.
7. Add biotinylated-tyramide(Genpoint, DAKO), incubate 10 minutes at room temperature in the dark.
- 15 8. Wash 3 times 7 minutes each in 1X TBST at 42°C.
9. Add 200 μ l of the Avidin-FITC diluted in peptide blocking buffer, cover and incubate for 15 minutes at 37°C.
10. Wash 3 times 7 minutes each in wash at 42°C.
11. Fix the cells by adding 0.5% Formaldehyde/PBS overnight at 4°C, or 4%Formaldehyde/PBS for 20 minutes at room temperature.
- 20 12. After the incubation with fix, wash the cells 2X in PBS at room temperature 5 minutes each.
13. If additional Biomarkers are to be used go to next section (LAMP-1 detection) step 1, if the peptides are being detected without additional antibody staining, go to step 14 below.
- 25 14. Add mounting medium with DAPI/Antiquenching agent to all of the wells.

LAMP-1 Detection:

1. Invert 96 well plates, shake PBS from the wells, add 200 μ l fresh
30 PBS at room temperature, incubate 5', invert to remove.

2. Add 0.2 ml antibody block (PBS/0.1% Saponin/5% Goat Serum) at room temperature, incubate 10 minutes at room temperature, invert to remove.
3. Dilute LAMP-1 antibody to a final concentration of 1 $\mu\text{g/ml}$ in antibody block, add 0.1 ml to each well and incubate at room temperature for 60 minutes, invert to remove.
4. Wash the wells with antibody wash at room temperature 2X 5 minutes each, rinse 1X with PBS at room temperature. (Carefully draw the wash up down 2 times).
5. Dilute the anti-rabbit-Cy3 labeled secondary antibody in the antibody block, add 0.5 mls to the wells, incubate at room temperature for 30 minutes.
6. Wash the wells with antibody wash at room temperature 2X 5 minutes each, rinse 1X with PBS at room temperature. (Carefully draw the wash up down 2 times).
7. Add 0.5 ml 0.5% formaldehyde/PBS and place at 4°C, overnight, or 4% formaldehyde for 20 minutes at room temperature.
8. After the incubation with fix, wash the cells 2X in PBS at room temperature 5 minutes each.
9. Add mounting medium with DAPI/Antiquenching agent to all of the wells.

Imaging:

The three components of the assay; peptide, antibody/LAMP-1 and nuclei detection, can be achieved by using fluorochromes that fluoresce at different excitation/emission spectrums. This allows each of the components to be detected separately, then combined at the end to determine the level of overlap with image analysis software.

Image acquisition software is currently available that captures up to 5 different fluorescent channels, by utilizing fluorescent filters with specific excitation/emission filters. The following filter excitation/emissions filters can be used: DAPI(nuclei EX:345/EM:425) filter, Cy3(Antibody, EX:575/EM:605), FITC (Peptide, EX:490/EM:525).

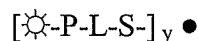
- To identify peptides that reside in the nucleus, one could capture the nuclear component, create a “mask” that would identify the region occupied by the nuclei, then overlay the peptide image of the same cells and determine how much of the peptide signal resides within the nuclear mask. To
5 determine peptides that co-localize to the biomarkers, the antibody or LAMP-1 image could be used to create the “mask” and the peptide image would be overlaid for determination or degree of co-localization. Any biomarker available can be utilized to identify specific cellular regions of interest.

It is to be understood that while the invention has been described in
10 conjunction with the above embodiments, that the foregoing description and examples are intended to illustrate and not limit the scope of the invention. Other aspects, advantages and modifications within the scope of the invention will be apparent to those skilled in the art to which the invention pertains.

CLAIMS

What is claimed is:

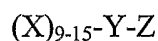
1. A random peptide library comprising compounds of the structure:



5 wherein

⊙ is absent or optionally a fluorescent label,

P is a D-peptide comprised of the structure:



10 wherein, each X is a D-amino acid subunit which varies independently from any other subunit and is selected from the group A, G, E, K, L, P, Q, R and W,

Y is a neutral amino acid subunit which varies independently from any other subunit, Z is a basic amino acid subunit which is the same for each member of the peptide library,

L is a linker,

15 S is absent or is optionally biotin,

y is a plurality, and

• is a solid support comprising a multiple copies of a single species of P.

2. The random peptide library of claim 1, wherein P is prepared by a method comprising the step of:

20 randomly introducing a D-amino acid subunit for each of the solid phase supports wherein the relative molar ratio of each D-amino acid subunit is provided according to the formula:

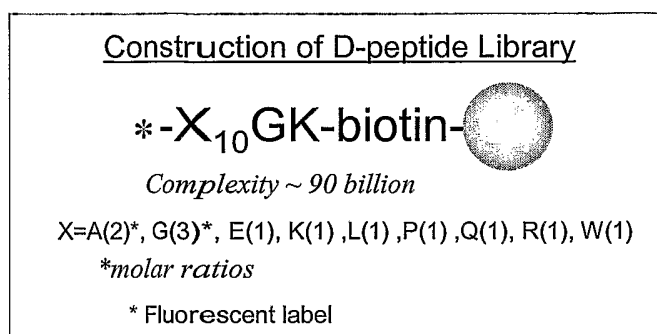
$$A=2,$$

$$G=3,$$

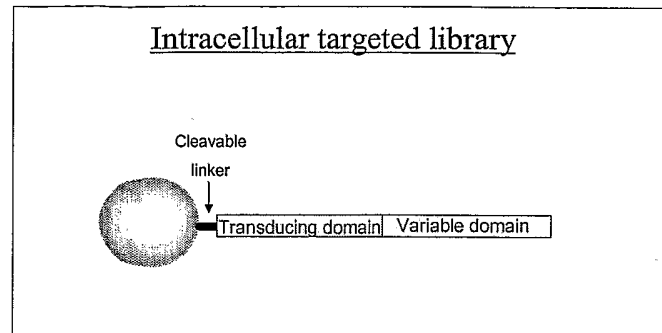
25 E, K, L, P, Q, R and W = 1.

3. The random peptide library of claims 1 or 2, comprising the maximum number of possible P species.

1/3

Representative Structure of Members of Peptide Library**Figure 1**

2/3

**Figure 2**

3/3

**Examples Of Targeted Peptides Identified
Using The '*X₁₀GK-Biotin-Bead*' Library**

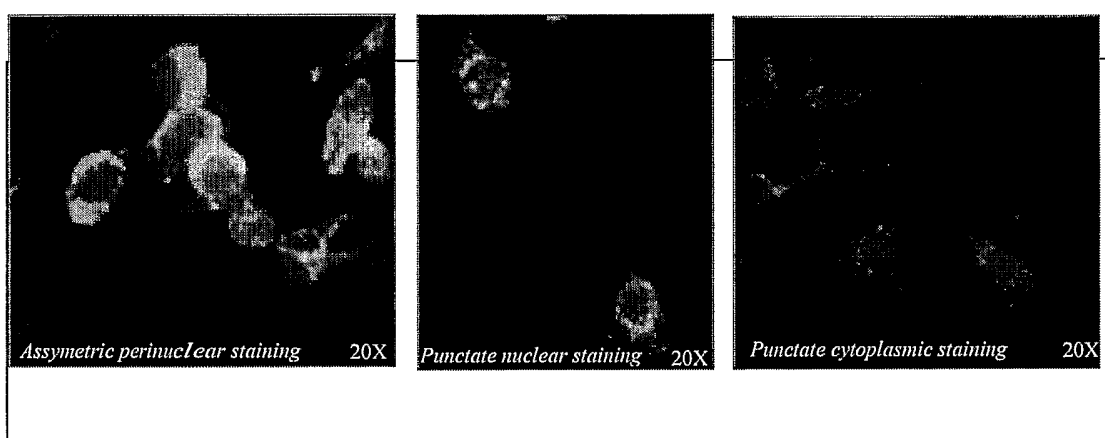


Figure 3