

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
13 February 2003 (13.02.2003)

PCT

(10) International Publication Number
WO 03/012146 A1

- (51) International Patent Classification⁷: **C12Q 1/68**, (74) Agent: **ELBING, Karen, L.**; Clark & Elbing LLP, 101 C07H 21/04, 21/02, G01N 15/06 Federal Street, Boston, MA 02110 (US).
- (21) International Application Number: PCT/US02/24180 (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, 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, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (22) International Filing Date: 31 July 2002 (31.07.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/309,231 31 July 2001 (31.07.2001) US (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application: US 60/309,231 (CIP) Filed on 31 July 2001 (31.07.2001)
- (71) Applicant (*for all designated States except US*): **PHYLOS, INC.** [US/US]; 128 Spring Street, Lexington, MA 02421 (US). **Published:**
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **KURZ, Markus** [DE/US]; 44 Pine Street, Newton, MA 02465 (US). **LOHSE, Peter** [CH/US]; 50 Golden Ball Road, Weston, MA 02493 (US).
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: MODULAR ASSEMBLY OF NUCLEIC ACID-PROTEIN FUSION MULTIMERS

(57) Abstract: Described herein are methods and reagents for generating nucleic acid-protein fusion multimers, and methods for using such fusion molecule multimers to select an interaction between a protein or a peptide and a compound.



WO 03/012146 A1

MODULAR ASSEMBLY OF NUCLEIC ACID-PROTEIN FUSION MULTIMERS

5 Background of the Invention

In general, the present invention relates to methods of generating and using nucleic acid-protein fusion multimers.

Certain macromolecules, such as proteins, are known to interact specifically with other molecules based on their three-dimensional shapes and electronic distributions. For example, proteins interact selectively with other proteins, nucleic acids, and small-molecules. The identification of proteins that interact with a target molecule lays the groundwork for the development of compounds to treat diseases and their associated symptoms. However, the discovery of a single drug candidate can require the screening of thousands of compounds, for example, proteins. It is therefore important to be able to screen large numbers of candidates rapidly and efficiently.

Summary of the Invention

The present invention features methods for making nucleic acid-protein fusion multimers. Such multimeric fusion complexes can be used for the *in vitro* selection of multidomain peptides or proteins with desired properties (for example, a desired binding property). A fusion multimer contains a dimerization (or multimerization) domain, which can reside in either the nucleic acid or protein portion of the fusion molecule. In addition, the fusion multimer includes a potential target (or compound) recognition domain in its protein portion, which may be randomized for selection purposes. Once dimerization or multimerization occurs, the recognition domains cooperatively interact with the compound of choice. In certain cases, this may strengthen the dimerization or multimerization of the fusions because of

additional binding forces. The compound recognition domains may be, for example, rather simple and unstructured peptide sequences, or may be antibody-like CDR loops or DNA binding motifs, such as zinc finger domains.

- 5 Accordingly, in general, in a first aspect, the present invention features novel nucleic acid-protein fusion multimers. One such multimer includes two or more fusion molecules of nucleic acid covalently bound to protein, the nucleic acid of at least one of the fusion molecules encoding the covalently bound protein, wherein the fusion molecules are hybridized to each other
10 through complementary nucleic acid sequences.

Another nucleic acid-protein fusion multimer includes two or more fusion molecules of nucleic acid covalently bound to protein, the fusion molecules not including streptavidin, wherein the fusion molecules are hybridized to each other through complementary nucleic acid sequences.

- 15 Another nucleic acid-protein fusion multimer includes two or more fusion molecules of nucleic acid covalently bound at the 3' end to protein, wherein the fusion molecules are hybridized to each other through complementary nucleic acid sequences.

- Another nucleic acid-protein fusion multimer includes two or more
20 fusion molecules of nucleic acid covalently bound to protein through a peptide acceptor, wherein the fusion molecules are hybridized to each other through complementary nucleic acid sequences.

- Yet another nucleic acid-protein fusion multimer includes two or more fusion molecules of nucleic acid covalently bound to protein, wherein the
25 covalent bond is not a thiol-maleimide bond, and wherein the fusion molecules are hybridized to each other through complementary nucleic acid sequences.

Another nucleic acid-protein fusion multimer includes (a) two or more fusion molecules of nucleic acid covalently bound to protein, the nucleic acid

of at least one of the fusion molecules encoding the covalently bound protein; and (b) an oligonucleotide, wherein a sequence of the nucleic acid of each of the fusion molecules is hybridized to a complementary sequence of the oligonucleotide.

5 Another nucleic acid-protein fusion multimer includes (a) two or more fusion molecules of nucleic acid covalently bound to protein, the fusion molecules not including streptavidin; and (b) an oligonucleotide, wherein a sequence of the nucleic acid of each of the fusion molecules is hybridized to a complementary sequence of the oligonucleotide.

10 Another nucleic acid-protein fusion multimer includes (a) two or more fusion molecules of nucleic acid covalently bound at the 3' end to protein; and (b) an oligonucleotide, wherein a sequence of the nucleic acid of each of the fusion molecules is hybridized to a complementary sequence of the oligonucleotide.

15 Yet another nucleic acid-protein fusion multimer includes (a) two or more fusion molecules of nucleic acid covalently bound to protein through a peptide acceptor; and (b) an oligonucleotide, wherein a sequence of the nucleic acid of each of the fusion molecules is hybridized to a complementary sequence of the oligonucleotide.

20 Another nucleic acid-protein fusion multimer includes (a) two or more fusion molecules of nucleic acid covalently bound to protein, wherein the covalent bond is not a thiol-maleimide bond; and (b) an oligonucleotide, wherein a sequence of the nucleic acid of each of the fusion molecules is hybridized to a complementary sequence of the oligonucleotide.

25 For any of the above multimers that includes an oligonucleotide, that oligonucleotide may have a linear, bi-directional, or branched structure.

The invention further features additional multimers. One such nucleic acid-protein fusion multimer includes (a) two or more fusion molecules of nucleic acid covalently bound to protein; and (b) an oligonucleotide having a

bi-directional or branched structure, wherein a sequence of the nucleic acid of each of the fusion molecules is hybridized to a complementary sequence of the oligonucleotide.

Another nucleic acid-protein fusion multimer includes (a) two or more
5 fusion molecules of nucleic acid covalently bound to protein, wherein the nucleic acid of each of the fusion molecules includes a polypurine tract; and (b) an oligonucleotide including at least two polypyrimidine tracts, wherein the polypurine tracts of the fusion molecules are hybridized to the polypyrimidine tracts of the oligonucleotide, and wherein binding of the
10 fusion molecules to the oligonucleotide occurs in opposite directions to form a triple helical structure. In this embodiment, the oligonucleotide may be circular, it may form a clamp-like structure, and/or it may include one or more polyamide nucleic acids.

In preferred embodiments of any of the above multimers of the
15 invention, the fusion molecules may be cross-linked through cross-linking moieties positioned within the nucleic acid of the fusion molecules or the oligonucleotide; and the cross-linking moiety may be psoralen.

The invention further features nucleic acid-protein fusion multimers that include two or more fusion molecules of nucleic acid covalently bound to
20 protein, wherein the protein portions of the fusion molecules each include a multimerization domain, the multimerization domains interacting through non-covalent bond formation.

In preferred embodiments of this aspect of the invention, the multimerization domains interact to form homodimers, heterodimers, trimers,
25 or tetramers; at least two of the multimerization domains include leucine zipper binding regions; the leucine zipper binding regions are derived from a Fos, Jun, or GCN4 leucine zipper binding region; and/or at least one of the multimerization domains includes a tetrazipper binding region.

In further aspects, the invention features a nucleic acid-protein fusion multimer that includes two or more fusion molecules of nucleic acid covalently bound to protein, wherein the protein of each of the fusion molecules includes a multimerization domain that includes a functional
5 group, the functional group of one fusion molecule being linked to the functional group of another fusion molecule through a covalent bond.

In preferred embodiments of this aspect of the invention, the covalent linkage involves an external cross-linking agent; at least one functional group includes an amine or a thiol; at least two functional groups include a thiol and
10 the covalent bond is a disulfide bond; and/or the multimerization domain includes an antibody constant region.

In preferred embodiments of any of the multimers according to the invention, the protein of at least one of the fusion molecules further includes a compound recognition domain. This compound recognition domain may
15 include an antibody variable region and/or a randomized domain. The compound recognition domain may interact with DNA (for example, it may be a zinc finger binding domain). Moreover, any multimer of the invention may include at least one nucleic acid-protein fusion molecule that is an RNA-protein fusion molecule or a DNA-protein fusion molecule.

20 In yet other embodiments, the invention features an RNA-protein fusion multimer, the multimer including two or more fusion molecules of RNA covalently bound to protein, wherein the fusion molecules are hybridized to each other through complementary nucleic acid sequences.

In preferred embodiments of this aspect of the invention, the RNA of at
25 least one of the fusion molecules encodes the covalently bound protein; the

fusion molecules are cross-linked through cross-linking moieties positioned within the RNA of the fusion molecules; and the cross-linking moiety is psoralen.

In a second general aspect, the invention further features methods for
5 preparing the multimers described herein. One such method involves the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound to protein, the nucleic acid of at least one of the fusion molecules encoding the covalently bound protein; and (b) hybridizing the fusion molecules to each other through complementary nucleic acid
10 sequences, thereby forming a nucleic acid-protein fusion multimer.

Another method involves the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound to protein, the fusion molecules not including streptavidin; and (b) hybridizing the fusion molecules to each other through complementary nucleic acid sequences, thereby forming a
15 nucleic acid-protein fusion multimer.

Another method for multimer preparation involves the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound at the 3' end to protein; and (b) hybridizing the fusion molecules to each other through complementary nucleic acid sequences, thereby forming a nucleic
20 acid-protein fusion multimer.

Another method for multimer preparation involves the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound to protein through a peptide acceptor; and (b) hybridizing the fusion molecules to each other through complementary nucleic acid sequences, thereby forming
25 a nucleic acid-protein fusion multimer.

Yet another method for multimer preparation involves the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound to protein, wherein the covalent bond is not a thiol-maleimide bond; and (b)

hybridizing the fusion molecules to each other through complementary nucleic acid sequences, thereby forming a nucleic acid-protein fusion multimer.

Another method for multimer preparation involves the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound to protein, the nucleic acid of at least one of the fusion molecules encoding the covalently bound protein; (b) providing an oligonucleotide, wherein the oligonucleotide includes a plurality of sequences that are complementary to a partner sequence within each of the fusion molecules; and (c) hybridizing the oligonucleotide to each of the fusion molecules, thereby forming a nucleic acid-protein fusion multimer.

Another method for multimer preparation involves the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound to protein, the fusion molecules not including streptavidin; (b) providing an oligonucleotide, wherein the oligonucleotide includes a plurality of sequences that are complementary to a partner sequence within each of the fusion molecules; and (c) hybridizing the oligonucleotide to each of the fusion molecules, thereby forming a nucleic acid-protein fusion multimer.

Yet another method for multimer preparation involves the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound at the 3' end to protein; (b) providing an oligonucleotide, wherein the oligonucleotide includes a plurality of sequences that are complementary to a partner sequence within each of the fusion molecules; and (c) hybridizing the oligonucleotide to each of the fusion molecules, thereby forming a nucleic acid-protein fusion multimer.

Yet another method for multimer preparation involves the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound to protein through a peptide acceptor, wherein the fusion molecules are hybridized to each other through complementary nucleic acid sequences; (b)

providing an oligonucleotide, wherein the oligonucleotide includes a plurality of sequences that are complementary to a partner sequence within each of the fusion molecules; and (c) hybridizing the oligonucleotide to each of the fusion molecules, thereby forming a nucleic acid-protein fusion multimer.

- 5 Another method for multimer preparation involves the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound to protein, wherein the covalent bond is not a thiol-maleimide bond; (b) providing an oligonucleotide, wherein the oligonucleotide includes a plurality of sequences that are complementary to a partner sequence within each of the fusion molecules; and (c) hybridizing the oligonucleotide to each of the fusion molecules, thereby forming a nucleic acid-protein fusion multimer.

For any preparative methods involving an oligonucleotide, that oligonucleotide may have a linear, bi-directional, or branched structure.

- Yet another method for multimer preparation according to the invention involves the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound to protein; (b) providing an oligonucleotide having a bi-directional or branched structure, wherein the oligonucleotide includes a plurality of sequences that are complementary to a partner sequence within each of the fusion molecules; and (c) hybridizing the oligonucleotide to each of the fusion molecules, thereby forming a nucleic acid-protein fusion multimer.

- Another method for multimer preparation involves the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound to protein, wherein the nucleic acid of each of the fusion molecules includes a polypurine tract; (b) providing an oligonucleotide including at least two polypyrimidine tracts; and (c) hybridizing the polypurine tracts to the

polypyrimidine tracts, wherein binding of the fusion molecules to the oligonucleotide occurs in opposite directions to form a triple helical structure, thereby forming a nucleic acid-protein multimer.

In this embodiment of the invention, the oligonucleotide may be
5 circular, may form a clamp-like structure, and/or may include one or more polyamide nucleic acids.

In any preparative technique according to the invention, the method may further involve cross-linking the fusion molecules to each other or to the oligonucleotide. This cross-linking may be carried out by functionalizing
10 the fusion molecules or the oligonucleotide with psoralen and irradiating the nucleic acid-protein fusion multimer.

Yet another inventive method for multimer preparation involves the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound to protein, wherein the protein portions of the fusion
15 molecules each include a multimerization domain; and (b) combining the fusion molecules under conditions that allow non-covalent interactions between the multimerization domains, thereby forming a nucleic acid-protein fusion multimer.

In preferred embodiments of this aspect of the invention, the
20 multimerization domains interact to form homodimers, heterodimers, trimers, or tetramers; at least two of the multimerization domains include leucine zipper binding regions; the leucine zipper binding regions are derived from a Fos, Jun, or GCN4 leucine zipper binding region; and/or at least one of the multimerization domains includes a tetrazipper binding region.

25 In yet another method according to the invention, a nucleic acid-protein fusion multimer is prepared by the steps of: (a) providing two or more fusion molecules of nucleic acid covalently bound to protein, wherein the protein of each of the fusion molecules includes a multimerization domain that includes a functional group; and (b) combining the fusion molecules under conditions

that allow the functional group of one fusion molecule to be linked to the functional group of another fusion molecule through a covalent bond, thereby forming a nucleic acid-protein fusion multimer.

In preferred embodiments of this method, the covalent linkage involves an external cross-linking agent; at least one functional group includes an amine or a thiol; at least two functional groups include a thiol and the covalent bond is a disulfide bond; and/or the multimerization domain includes an antibody constant region.

In preferred embodiments of any of the preparative methods of the invention, the protein of at least one of the fusion molecules further includes a compound recognition domain; the compound recognition domain includes an antibody variable region; and/or the compound recognition domain includes a randomized domain. Such a compound recognition domain may interact with DNA (for example, it may be a zinc finger binding domain. In addition, for any preparative method described herein, at least one of the nucleic acid-protein fusion molecules may be an RNA-protein fusion molecule or a DNA-protein fusion molecule.

In yet another method according to the invention, an RNA-protein fusion multimer is prepared by the steps including: (a) providing two or more fusion molecules of RNA covalently bound to protein; and (b) hybridizing the fusion molecules to each other through complementary nucleic acid sequences, thereby forming an RNA-protein fusion multimer.

In a third general aspect, the invention further features a method of selecting a protein that interacts with a compound, the method involving the steps of: (a) providing a population of candidate nucleic acid-protein fusion multimers, the multimers including two or more hybridized or covalently bound fusion molecules of nucleic acid covalently bound to protein; (b) providing a compound; (c) contacting the compound with the population of candidate nucleic acid-protein fusion multimers under conditions that allow

an interaction between the compound and the candidate nucleic acid-protein fusion multimers; and (d) selecting a nucleic acid-protein fusion multimer that interacts with the compound, thereby selecting a protein that interacts with the compound.

- 5 In one preferred embodiment, the compound is immobilized on a column.

 In another preferred embodiment, the method further involves the steps of: (e) dissociating the nucleic acid-protein fusion multimers that do not interact with the compound; (f) recombining the dissociated nucleic acid-
10 protein fusion multimers; (g) contacting the compound with the recombined nucleic acid-protein fusion multimers; and (h) selecting a recombined nucleic acid-protein fusion multimer that interacts with the compound, thereby selecting a protein that interacts with the compound. In this method, the dissociation and recombination of the nucleic acid-protein fusion multimers
15 that do not interact with the compound may occur through heating and cooling of the fusion molecules or through denaturation by a reduction reaction, followed by recombination under oxidative conditions. Steps (e) through (h) may be repeated any number of times.

 In another preferred approach of the selection methods, the population,
20 in step (a), is maintained under equilibrium conditions, whereby the individual fusion molecules of the nucleic acid-protein fusion multimers rapidly dissociate and associate with other individual fusion molecules, thereby forming new nucleic acid-protein fusion multimers.

 In yet another preferred embodiment, the method having steps (a) - (d)
25 described above further involves the steps of: (e) amplifying the nucleic acids of the nucleic acid-protein fusion multimers selected in step (d); (f) generating, from the amplified nucleic acids, fusion molecules of nucleic acid

covalently bound to protein; (g) generating from those fusion molecules a second population of nucleic acid-protein fusion multimers; and (h) repeating steps (b) through (d).

In other preferred embodiments of the selection methods, the compound
5 interacts with the nucleic acid-protein fusion multimer in solution and is subsequently immobilized on a solid phase; the compound is detectably labeled; the compound is detectably labeled with biotin and the solid support is a streptavidin resin; and/or at least one of the fusion molecules is an RNA-protein fusion molecule or a DNA-protein fusion molecule.

10 As used herein, by “nucleic acid-protein fusion molecule” is meant a molecule comprising a nucleic acid covalently bound directly or indirectly to a protein. This molecule may also include additional components, for example, a non-nucleosidic spacer or psoralen. The nucleic acid molecule may be an RNA or DNA molecule, or may include RNA or DNA analogs at
15 one or more positions in the sequence. Alternatively, the nucleic acid portion of the fusion may be partially or wholly composed of a PNA sequence. The “protein” portion of the fusion is composed of two or more naturally occurring or modified amino acids joined by one or more peptide bonds. “Protein” and “peptide” are used interchangeably herein. Exemplary RNA-
20 protein fusion molecules are described, for example, by Roberts and Szostak (Proc. Natl. Acad. Sci. USA 94:12297-302, 1997); Szostak et al. (WO 98/31700; WO 00/47775); and Gold et al. (U.S. Patent No. 5,843,701). Exemplary DNA-protein fusion molecules are described, for example, in U.S. Patent No. 6,416,950; WO 00/32823, all of which are hereby incorporated by
25 reference.

By “nucleic acid-protein fusion multimer” is meant two or more identical or different nucleic acid-protein fusions bound together covalently or non-covalently.

By “functionalize” is meant to chemically modify a molecule in a manner that results in the attachment of a functional group or moiety. For example, a nucleic acid-protein fusion molecule can be functionalized with a cross-linking moiety such as psoralen, an azido compound, or a sulfur-
5 containing nucleoside.

By “hybridize” is meant to associate complementary nucleic acid sequences with one another. Standard hybridization techniques are known in the art (see, for example, Ausubel et al, *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY, 1998). Preferably, the nucleic
10 acid sequences hybridize under low temperature (for example, less than or equal to room temperature), medium to high ionic strength buffer (for example, in a buffer containing 100 mM or greater salt concentration), and neutral pH (for example, pH 7-8).

By a “multimerization domain” of a nucleic acid-protein fusion
15 molecule is meant a portion of a nucleic acid-protein fusion molecule that binds that fusion molecule to one or more additional nucleic acid-protein fusion molecules to form a nucleic acid-protein fusion multimer. Binding may be covalent or non-covalent. The multimerization domain may be located in the protein portion of the fusion molecule, where multimerization
20 of the fusion molecules occurs through interaction of the protein multimerization domains. Examples of protein multimerization domains include, but are not limited to, domains that are functionalized such that they interact with one another, domains capable of forming homodimers, heterodimers, trimers, or tetramers (for example leucine zipper binding
25 regions or tetrazipper binding regions), and antibody constant regions. Alternatively, the multimerization domain may be located in the nucleic acid portion of the fusion molecule, where multimerization of the fusion molecules occurs, for example, through hybridization. “Multimerization domains” also

may be located on an external oligonucleotide, in which case the domains hybridize with the nucleic acid portions of fusion molecules to form a multimer.

By a “compound recognition domain” is meant a portion of a nucleic acid-protein fusion molecule that facilitates interaction of that molecule (or its associated complex) with a compound (e.g., by binding or by catalyzing a binding event). The compound recognition domain may be located in the protein portion of the fusion molecule and may include, for example, a randomized amino acid sequence, a naturally occurring or optimized domain capable of interacting with DNA (for example, a zinc finger binding domain), or an antibody variable region. Compound recognition domains may facilitate compound interactions alone or preferably in association with the compound recognition domains of other associated fusion molecules in a complex.

By a “peptide acceptor” is meant any molecule capable of being added to the C-terminus of a growing protein chain by the catalytic activity of the ribosomal peptidyl transferase function. Typically, such molecules contain (i) a nucleotide or nucleotide-like moiety (for example, adenosine or an adenosine analog (di-methylation at the N-6 amino position is acceptable)), (ii) an amino acid or amino acid-like moiety (for example, any of the 20 D- or L-amino acids or any amino acid analog thereof (for example, O-methyl tyrosine or any of the analogs described by Ellman et al., Meth. Enzymol. **202**:301, 1991), and (iii) a linkage between the two (for example, an ester, amide, or ketone linkage at the 3' position or, less preferably, the 2' position); preferably, this linkage does not significantly perturb the pucker of the ring from the natural ribonucleotide conformation. Peptide acceptors may also possess a nucleophile, which may be, without limitation, an amino group, a hydroxyl group, or a sulfhydryl group. In addition, peptide acceptors may be composed of nucleotide mimetics, amino acid mimetics, or mimetics of the

combined nucleotide-amino acid structure. Exemplary peptide acceptors are described, for example, in Szostak et al., WO 98/31700.

By “selecting,” as it applies to proteins, is meant identifying, detecting, or substantially isolating a protein in a nucleic acid-protein fusion multimer
5 that interacts with a compound. A protein may be selected, for example, by immobilizing a compound on a column, running a solution containing candidate nucleic acid-protein fusion multimers through the column, binding the fusion multimers of interest to the column by affinity interactions with the compound, removing non-specific fusions from the column, and eluting the
10 multimers that bound to the compound. The nucleic acid-protein fusion multimers contained in the eluate are those that contain proteins that are selected. Thus, by “selecting” a nucleic acid-protein multimer that interacts with a compound, one can “select” a protein that interacts with the compound.

By a “compound” or “ligand” is meant a chemical molecule, be it
15 naturally-occurring or artificially-derived, and includes, for example, peptides, proteins, synthetic organic molecules, naturally-occurring organic molecules, nucleic acid molecules, and components thereof.

The advantages of *in vitro* selection of proteins or peptides using nucleic acid-protein fusion multimers are many. For example, multimerization allows
20 the preparation of candidate pools with increased complexity by allowing the combination of identical copies of the same nucleic acid-protein fusion molecule with different partners. In addition, equilibrium association and dissociation of the pool members can allow dynamic recombination, thereby further increasing the pool complexity by repeatedly producing new
25 combinations of nucleic acid-protein fusion multimers.

Furthermore, in contrast to certain gene recombination methods, the present invention provides a method for domain shuffling on the phenotype level. Dynamic recombination of nucleic acid-protein fusion molecules can take place at any time during the course of an *in vitro* selection experiment.

This potential for inherent steady recombination of nucleic acid-protein fusion molecules, as described below, can offer advantages for the *in vitro* selection of proteins with novel properties.

Other features and advantages of the invention will be apparent from the following detailed description, and from the claims.

Detailed Description

The drawings will first briefly be described.

Brief Description of the Drawings

FIGURE 1A is a schematic representation of the heterodimerization of nucleic acid-protein fusion molecules through duplex formation of the nucleic acid portion of the fusion molecule or the linker DNA.

FIGURE 1B is a schematic representation of the homodimerization of nucleic acid-protein fusion molecules through duplex formation of the nucleic acid portion of the fusion molecule or the linker DNA. Homodimers are formed when the nucleic acid portions of the fusion molecules or the DNA linkers contain palindromic sequences.

FIGURE 2A is a schematic representation of the multimerization of nucleic acid-protein fusion molecules through ternary complex formation with a connector oligonucleotide. The connector oligonucleotide may be a linear sequence that simultaneously binds to more than one fusion molecule by Watson-Crick base pairing.

FIGURE 2B is a schematic representation of the multimerization of nucleic acid-protein fusion molecules through ternary complex formation with a connector oligonucleotide. The connector oligonucleotide may be a bi-directional oligonucleotide.

FIGURE 2C is a schematic representation of the multimerization of nucleic acid-protein fusion molecules through ternary complex formation with a connector oligonucleotide. The connector oligonucleotide may be a branched oligonucleotide.

- 5 FIGURE 3A is a schematic representation of the multimerization of nucleic acid-protein fusion molecules through ternary complex formation with a connector oligonucleotide. The connector oligonucleotide may be a circular triplex-forming oligonucleotide. Target sequences on the nucleic acid-protein fusion molecules are designed to contain polypurine tracts that
- 10 bind to two interconnected polypyrimidine tracts contained in the oligonucleotide. Each polypyrimidine tract binds one nucleic acid-protein fusion molecule through antiparallel Watson-Crick base pairing, while binding the other nucleic acid-protein fusion molecule in a parallel Hoogsteen mode.
- 15 FIGURE 3B is a schematic representation of the multimerization of nucleic acid-protein fusion molecules through ternary complex formation with a connector oligonucleotide. The connector oligonucleotide may be a triplex-forming oligonucleotide having a clamp-like structure. Target sequences on the nucleic acid-protein fusion molecules are designed to
- 20 contain polypurine tracts that bind to two interconnected polypyrimidine tracts contained in the oligonucleotide. Each polypyrimidine tract binds one nucleic acid-protein fusion molecule through antiparallel Watson-Crick base pairing, while binding the other nucleic acid-protein fusion molecule in a parallel Hoogsteen mode.
- 25 FIGURE 4A is a schematic representation of the protein-based multimerization of nucleic acid-protein fusion molecules. In their protein portions, the fusion molecules contain multimerization domains that bind one another.

FIGURE 4B is a schematic representation of the protein-based multimerization of nucleic acid-protein fusion molecules occurring when the protein portion of the fusion molecules includes an antibody Fab fragment. The protein portions of the fusion molecules contain constant regions C_H and C_L (i.e., multimerization domains) which mediate association followed by disulfide bond formation, allowing the randomized variable regions V_H and V_L (i.e., compound recognition domains) to be correctly positioned for recognition and binding of potential antigens.

FIGURE 4C is a schematic representation of the protein-based multimerization of nucleic acid-protein fusion molecules occurring when the protein portion of the nucleic acid-protein fusion molecule contains a DNA binding domain (i.e., compound recognition domain), and a double stranded DNA target molecule is provided. The dimer is formed by the interaction of the DNA binding domain, for example, a zinc finger domain, with the DNA target molecule. In this format, the multimerization and compound recognition domains may overlap. Alternatively, the multimerization domain and compound recognition domain may differ. For example, candidate DNA binding complexes may exploit a leucine zipper domain for multimerization purposes and a zinc finger domain (or randomized or mutagenized domain) for DNA recognition purposes.

FIGURE 5 is a schematic representation illustrating how the complexity of nucleic acid-protein fusion multimer libraries can be increased through multiple nucleic acid-protein fusion libraries. Two independent sub-libraries (e.g., sub-libraries A and B) are generated, each library containing multiple copies of each nucleic acid-protein fusion molecule. The sub-libraries are then combined, and dimerization of the fusion molecules of the two different sub-libraries occurs. During the dimerization step, each particular nucleic

acid-protein fusion molecule from sub-library A will be combined with a different nucleic acid-protein fusion molecule from sub-library B, resulting in a library of unique members.

FIGURE 6 is a schematic representation illustrating how library complexity is increased through repeated recombination during the selection step. A library of nucleic acid-protein fusion molecules is generated and multimerized. The library is then passed over a column containing the desired immobilized affinity ligand. Those nucleic acid-protein fusion multimers that bind the ligand are retained on the column, while the remaining nucleic acid-protein fusion multimers are recovered in the eluate. In the next step, the multimers are dissociated, and then recombined to form new multimers. These newly formed multimerized nucleic acid-protein fusion complexes are then passed over the column.

FIGURE 7 is a schematic representation illustrating how the complexity of a nucleic acid-protein fusion multimer library can be increased through dynamic recombination. The nucleic acid-protein fusion complexes of this library dimerize through weak, non-covalent interactions. In an equilibrium state, those molecules rapidly associate and dissociate, thereby constantly creating new multimeric species. Suitable complexes then jointly bind to a ligand, which increases the overall complex stability and removes this complex from equilibrium to separate the selected nucleic acid-protein fusion complexes from unbound species.

FIGURE 8 is a schematic representation illustrating how the diversity of a nucleic acid-protein fusion multimer library is repeatedly generated when proceeding from one molecular generation to the next. In generation n, specific nucleic acid-protein fusion molecule multimers are selected for by binding a target. As part of the preparation of the generation n+1 for the next round of selection, those fusion molecules that were selected for in the previous step as binding a target are generated again and added to the n+1

generation of nucleic acid-protein fusion multimers to provide additional diversity and to combine advantageous binding features to shape tighter binders.

Described herein are methods for making nucleic acid-protein fusion
5 multimers and using such fusion complexes to select desired proteins and peptides in the form of nucleic acid-protein fusion molecules. Techniques for carrying out each method of the invention are now described in detail, using particular examples. These examples are provided for the purpose of illustrating the invention, and should not be construed as limiting.

10 Example 1: Formation of nucleic acid-protein fusion molecules

As described above, nucleic acid-protein fusion molecules useful in the invention may include any nucleic acid or nucleic acid analog covalently bonded to any naturally occurring or modified peptide sequences. To generate RNA-protein fusions in which the RNA encodes the associated
15 protein, individual RNA sequences (or a plurality of sequences) may be translated *in vitro*, and fusions formed, for example, according to the methods of Roberts and Szostak (supra) and Szostak et al. (supra). The RNA for the *in vitro* translation reaction may be generated by any standard approach, including normal cellular synthesis, recombinant techniques, chemical
20 synthesis, and enzymatic synthesis (e.g., *in vitro* transcription using, for example, T7 RNA polymerase), and useful RNA libraries according to the invention include, without limitation, cellular RNA, mRNA libraries, and random synthetic RNA libraries. A peptide acceptor in the method of Szostak et al. (supra) (for example, puromycin) is bonded to the RNA through a
25 nucleic acid or nucleic acid analog linker. Almost any spacer unit may be used to bind the peptide acceptor to the RNA. For example, spacer units provided by Glen Research (Sterling, VA) may be utilized, particularly, triethylene glycol phosphate (Spacer 9), hexaethylene glycol phosphate

(Spacer 18), propylene phosphate (Spacer C3), and dodecamethylene phosphate (Spacer C12) spacers. Additional exemplary nucleic acid analogs include, for example, a polyamide nucleic acid (PNA; Nielsen et al., Science 254:1497, 1991), a P-RNA (Krishnamurthy, Angew. Chem. 35:1537, 1996),
5 or a 3'N phosphoramidate (Gryaznov and Letsinger, Nucleic Acids Res. 20:3403, 1992). Such peptide acceptor molecules may be generated by any standard technique, for example, the techniques described in Roberts and Szostak (supra) and Szostak et al. (supra).

An RNA-protein fusion molecule preferably consists of an mRNA
10 molecule that includes a translation initiation sequence and a start codon operably linked to a candidate protein coding sequence and a peptide acceptor at the 3' end of the candidate protein coding sequence. A DNA or RNase resistant nucleic acid analog linker sequence is included between the end of the message and the peptide acceptor. If desired, groups or collections of
15 RNA sequences, for example, from a particular source or of a given type, may be translated together in a single reaction mixture according to the same general procedure.

DNA-protein fusion molecules may also be used to carry out the present invention. Such DNA-protein fusion molecules are similar to the above-
20 described RNA-protein fusion molecules, except that the DNA, for example, a cDNA is covalently attached to the protein portion of the fusion molecule. Preferably the DNA contains the genetic information for the protein to which it is bonded. DNA-protein fusion molecules may be generated, for example, according to the methods provided by U.S. Patent No. 6,416,950; WO
25 00/32823.

For selection purposes, the fusion molecules so produced preferably include a protein domain having a candidate compound recognition domain that facilitates interaction with a target compound, either alone or in association with the compound recognition domains from other associated

fusions in a fusion multimer. In addition, either the nucleic acid or protein portion of the fusion molecule includes a multimerization domain that facilitates the formation of multimeric complexes with other fusion molecules in a population.

5 Example 2: Formation of a nucleic acid-protein fusion multimer through direct hybridization of the nucleic acid portions of the nucleic acid-protein fusion molecules

Nucleic acid-protein fusion molecules can be multimerized, for example, through duplex formation of nucleic acid portions located in either
10 the RNA or linker DNA portions of RNA-protein fusion molecules or located in DNA-protein fusion molecules. As shown in Figure 1A, the base-pairing regions can be specifically designed to allow the formation of heterodimers (for example, A-B heterodimers) or pools of fusion molecules (for example, A_{a-z}; B_{a-z}). Alternatively, the use of palindromic sequences in the DNA
15 portion of the DNA-protein fusion molecule or in the DNA linker allows the formation of homodimers (A-A; see Figure 1B). This strategy is not restricted to dimerization events, as multiple nucleic acid-protein fusion molecules can be connected by direct hybridization. Preferably, multimerization domains in the nucleic acid sequences hybridize under low
20 temperature (for example, less than or equal to room temperature), medium to high ionic strength buffer (for example, in a buffer containing 100 mM or greater salt concentration), and neutral pH (for example, pH 7-8). In addition, the hybridizing sequence length for each nucleic acid portion of the fusion molecule is preferably between 15 and 25 nucleotides.

25 The direct hybridization of nucleic acid-protein fusion molecules to one another requires careful purification of the nucleic acid-protein fusion molecules from unfused nucleic acids or DNA linkers. Contaminating unfused nucleic acids or DNA linkers would also be subject to hybridization

with nucleic acid-protein fusion molecules and therefore interfere with dimerization. In addition, the careful choice of length and sequence of the hybridization domains permits the tailoring of the thermodynamic stability of the multimeric fusion molecule complex. This is especially important in applications where equilibrium association and dissociation of the fusion molecules are crucial, as described below. The nucleic acid-protein fusion molecules can be purified according to the methods of Roberts and Szostak (supra).

Example 3: Formation of a nucleic acid-protein fusion multimer through ternary complex formation using a connector oligonucleotide

The multimerization of individual nucleic acid-protein fusion molecules can also be mediated by an external oligonucleotide. For example, multimerization domains in a simple linear oligonucleotide sequence can be utilized which simultaneously bind to the multiple domains in more than one nucleic acid-protein fusion molecule through Watson-Crick base pairing (Figure 2A). Heteromultimeric or homomultimeric nucleic acid-protein fusion molecules may be generated by designing an oligonucleotide that contains sequences that hybridize to a sequence contained in each of the desired members of the multimeric fusion molecule.

In a similar approach, nucleic acid-protein fusion multimers can be formed by hybridizing the nucleic acid multimerization domain of each fusion molecule to multimerization domains in a bi-directional template oligonucleotide (Figure 2B). If desired, the protein portions of the nucleic acid-protein fusion may be positioned near each other, by designing the template oligonucleotide to hybridize to the 3' end of the nucleic acid portion of the fusion molecule, or, more preferably, to nucleic acid sequences adjacent to the puromycin at the 3'-end of the linker portions of the fusion molecules.

The required reversal of sequence polarity of the bi-directional oligonucleotides to be used in this method can be easily introduced by standard DNA synthesis of one half of the template molecule, followed by synthesis of the second half using 5'-phosphoramidites (Glen Research).

- 5 Preferably the length of the hybridizing sequence of the nucleic acid-protein fusion molecule and the oligonucleotide is between 15 and 25 nucleotides for each fusion molecule, resulting in a total hybridization sequence length of 30 to 50 nucleotides if the multimeric fusion molecule contains two fusion molecules. Optimal hybridization conditions include low
10 temperature (for example, less than or equal to room temperature), medium to high ionic strength buffer (for example, in a buffer containing 100 mM or greater salt concentration), and neutral pH (for example, pH 7-8).

Example 4: Formation of a nucleic acid-protein fusion multimer through ternary complex formation using a branched connector oligonucleotide

- 15 Nucleic acid-protein fusion multimers can also be generated by hybridizing individual fusion molecules to a branched connector oligonucleotide (Figure 2C). For example, a branching amidite (reagents available from Clontech, Palo Alto, CA) may be used to synthesize a branched oligonucleotide, wherein each branch of the oligonucleotide
20 contains a multimerization domain that hybridizes to a multimerization domain in the nucleic acid portion or the linker of one or more individual fusion molecules. Branched connector oligonucleotides consisting of multiple branching points may also be designed and used to form multimeric nucleic acid-protein fusion molecules. The optimal length of the hybridizing
25 sequence is 15 to 25 nucleotides for each fusion molecule of the multimeric nucleic acid-protein fusion molecule. The optimal hybridization conditions are as in Example 3.

Example 5: Formation of a nucleic acid-protein fusion multimer using a circular triplex-forming oligonucleotide

Nucleic acid-protein fusion multimers can also be formed by using circular triplex-forming oligonucleotides (TFO) as templates to which individual fusion molecules bind (Selvasekaran and Turnbull, Nucleic Acids Res. 27:624, 1999). Target sequences (multimerization domains) on the nucleic acid-protein fusion molecules, preferably in the linker DNA, can be designed to consist of polypurine tracts. These tracts bind to polypyrimidine tracts contained within the TFO; each polypyrimidine tract in the TFO targets one nucleic acid-protein fusion molecule through antiparallel Watson-Crick base-pairing while binding the second nucleic acid-protein fusion molecule in a parallel Hoogsteen mode. Thus, the two nucleic acid-protein fusion molecules can be bound in opposite directions, allowing the peptide moieties to be positioned next to each other. Preferably the length of each polypurine and polypyrimidine tract is between 15 and 25 nucleotides. Hybridization of the polypurine and polypyrimidine tracts occurs at low temperatures using a medium to high ionic strength buffer containing multivalent cations (e.g., Mg^{2+} , spermine, or spermidine) at a slightly acid pH (e.g., pH approximately 5 to 6). In addition, more than two nucleic acid-protein fusion molecules may be bound together with one triplex-forming oligonucleotide template.

For this approach, polyamide nucleic acids (PNAs) would be a suitable alternative to conventional DNA, because of their favorable behavior in triplex formation (Uhlman et al., Angew. Chem. Int. Ed. 37:2796-2823, 1998).

Example 6: Formation of a nucleic acid-protein fusion multimer using a clamp-like triplex-forming oligonucleotide

Nucleic acid-protein fusion multimers can also be formed using a method similar to that described above. In this method, multimerization

domains in individual nucleic acid-protein fusion molecules are hybridized to multimerization domains in a TFO oligonucleotide that has a clamp-like structure. As in the above Example, polyamide nucleic acids (PNAs) may be used in the synthesis of such TFOs. Preferably the polypurine and

5 polypyrimidine tracts are between 10 and 15 nucleotides in length, with hybridization occurring at low temperature (for example, less than or equal to room temperature), using a low to medium ionic strength buffer (for example, in a buffer containing less than or equal to 50 mM) at a slightly acidic pH (e.g., pH approximately 5 to 6).

10 It is important to point out that the purity of the nucleic acid-protein fusion molecules should be high for each of the above-described methods of generating multimeric fusion complexes, as contaminating unfused nucleic acids or DNA linkers may hybridize with the nucleic acid-protein fusion molecules, thereby interfering with the multimerization process.

15 Example 7: Formation of a nucleic acid-protein fusion multimer through covalent bond formation between the nucleic acid portions of fusion molecules

If desired, an enhancement of the methods of any of the above Examples may be implemented by adding stability to the hybridized fusion

20 molecules and oligonucleotides through the formation of covalent bonds between those nucleobases involved in the hybridization. For example, oligonucleotide connectors functionalized to carry psoralen moieties allow the introduction of cross-links to complementary nucleic acids in the fusion molecule upon irradiation with long wave UV-light. Such psoralen moieties

25 can be positioned at strategic positions within (Pieles and Englisch, Nucleic Acids Res. 17:285, 1989) or adjacent to (Pieles et al., Nucleic Acids Res. 17:8967, 1989) an oligonucleotide to be base paired with the fusion molecule(s).

Example 8: Covalent bond formation between the peptide portions of nucleic acid-protein fusion molecules

Nucleic acid-protein fusion multimers can also be formed through covalent interactions between the protein portions of individual fusion molecules. For example, suitable functional groups (e.g., -NH₂ and -SH) that are present in the multimerization domains of the peptide portions of the fusion molecules can be used to covalently cross-link individual fusion molecules, using commercial standard cross-linking reagents (-NH₂ to -NH₂: Disuccinimidyl suberate (DSS) and related reagents, as described by Mattson et al., Molecular Biology Reports 17:167-183, 1993; -NH₂ to -SH: (*N*-[ε-Maleimidocaproyloxy]succinimide ester (EMCS), *N*-succinimidyl 3-(2-pyridyldithiol)propionate (SPDP), Succinimidyl 4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate (SMCC), *m*-Maleimidobenzoyl-*N*-hydroxysuccinimide ester (MBS) and related reagents, according to Peeters et al., J. Immunol. Methods 120:133-143, 1989; and -SH to -SH: 1,4-bis-Maleimidyl-2,3-dihydroxybutane (BMDB) and related reagents; Pierce, Rockford, IL). Alternatively, two -SH groups can be directly linked through disulfide bond formation.

Preferably in these cross-linking methods, amino acid(s) are utilized that carry the desired functional groups and that are positioned at useful locations in the peptide. Alternatively, such amino acid(s) may be introduced at specified positions within the peptide. In addition, it is preferable that such functional groups be present only once per nucleic acid-protein fusion molecule, or be positioned in a way such that a higher relative reactivity in cross-link formation, compared to potential competitors, is established. This could be achieved, for example, by employing the above-described nucleic acid hybridization methods to position the desired reactive groups such that a specific cross-linking reaction is promoted.

Dynamic equilibrium state multimer formation (as described in Example 14) is advantageous to correct improper orientation or positioning.

Example 9: Formation of nucleic acid-protein fusion multimers through the association of the protein portions of the nucleic acid-protein fusion

5 molecules

The formation of nucleic acid-protein fusion multimers can also be accomplished through the association of the multimerization domains of the protein portions of the fusion molecules. In this approach, each protein portion of a fusion molecule contains two regions, a defined multimerization domain and a compound recognition domain (for example, a randomized domain). The nucleic acid-protein fusion molecules may be generated, for example, by synthesizing the nucleic acid molecule that contains nucleic acids encoding the defined multimerization domain and the randomized compound recognition domain. Alternatively, a nucleic acid encoding the
10 desired randomized compound recognition domain may be ligated to a selected nucleic acid sequence encoding the defined multimerization domain. Dimeric or multimeric nucleic acid-protein fusion molecules are then formed through the interaction of the defined multimerization domains.

Multimerization domains may be chosen to form, for example, homodimers
20 (e.g., GCN4 leucine zippers, as described by O'Shea et al., Science 243:538-42, 1989; and O'Shea et al., Science 254:539-44, 1991), heterodimers (e.g., Jun-Fos according to O'Shea et al., Science 245:646-8, 1989), trimers, or tetramers (Harbury et al., Science 262:1401-7, 1993; and Graddis et al., Biochemistry 32:12664-71, 1993). Alternatively,
25 multimerization domains may also be antibody constant regions (e.g., C_H-C_L, as described by Muller et al., FEBS Lett. 422:259-64, 1998). If desired,

interactions between the multimerization domains of the fusion molecules can be further strengthened by the formation of covalent disulfide bridges between the multimerization domains.

Preferably the relative orientation of the multimerization domains of each nucleic acid-protein fusion molecule are chosen such that the randomized compound recognition domains are in close proximity to each other without interfering with the nucleic acid portions of the fusion molecules. This can be achieved, for example, by introducing parallel multimerization domains (e.g., zinc finger domains, leucine zipper domains, for example, Jun-Fos leucine zipper regions, antibody C_H-C_L regions, tetrazipper regions, or other such binding domains known in the art) at the carboxy-terminus of the protein portion, with the randomized compound recognition domains at the amino-terminus (see Figures 4A and 4B).

Example 10: Formation of multimeric nucleic acid-antibody Fab-fragment complexes

The formation of nucleic acid-protein fusion multimers through interaction between multimerization domains in the protein portions of the fusion molecules can be made versatile by designing those protein portions to consist of antibody Fab fragments. Dimerization is mediated through association between the constant regions, C_H and C_L (i.e., multimerization domains), followed by disulfide bond formation, allowing the randomized variable regions V_H and V_L (i.e. compound recognition domains) to be correctly positioned for recognition and binding of potential antigens. This method can be used for the construction of antibody Fab libraries (Gao et al., Proc. Natl. Acad. Sci. USA 96:6025, 1999).

Example 11: Formation of nucleic acid-protein fusion multimers for the selection of protein multimers that bind to a desired DNA double helical element

DNA binding proteins (Pomerantz et al., Biochemistry 37:965, 1998) can also be discovered by *in vitro* selection using nucleic acid-protein fusion multimers. As discussed above in Example 9, the C-termini of the protein portions of the fusion molecules can be fused to an appropriate dimerization domain (e.g., a leucine zipper domain or antibody constant region as shown in Fig. 4C) or nucleic acid sequence in the context of nucleic acid-protein fusion molecules. The molecule further contains a compound recognition domain that recognizes and binds DNA (e.g., a zinc finger domain). When the fusion protein dimer is contacted with a double-stranded target DNA element, the fusion proteins will recognize and bind two adjacent sequence domains on the target DNA (Figure 4C). The use of zinc finger domains containing randomized sequences can be used to select for protein dimers that bind to a desired DNA double helical element.

Example 12: Increasing library complexity through assembly of multiple nucleic acid-protein fusion molecules

A random library of nucleic acid-protein fusion multimers can be formed by combining different libraries of nucleic acid-protein fusion molecules. For example, for dimeric constructs, the construction of a random library begins with the preparation of two independent pools of randomized template DNA (e.g., A_{1-n} ; B_{1-n}). Amplification of the DNA molecules, followed by formation of nucleic acid-protein fusion molecules, for example, according to the methods of Roberts and Szostak (*supra*) and Szostak et al. (*supra*) leads to multiple copies of each member of the two pools. During the dimerization step, each particular molecule (e.g., A_i) is randomly combined with a different molecule from the other pool (e.g., B_n), resulting in a library

of potentially unique members (A_1 - B_1 ; A_1 - B_2 , A_1 - B_3 ,...) (Figure 5). Thus, the library complexity is maximized, and virtually equals the number of possible dimeric nucleic acid-protein fusion molecule combinations that are present in the pools.

5 Example 13: Increasing library complexity through repeated recombination during the selection step

Depending on the nature and strength of the forces that govern the multimerization of the nucleic acid-protein fusion molecules, dissociation and recombination steps can be introduced during the selection process, thereby
10 continuously producing new multimers approaching the maximal molecular diversity.

An exemplary scheme for the *in vitro* selection of nucleic acid-protein fusion molecules with high binding affinities for a given ligand now follows (Figure 6). Nucleic acid-protein fusion molecules A and B are generated
15 separately, and are subsequently multimerized using any of the methods described above. This library is then passed over a column with the desired immobilized affinity ligand, for example, a candidate compound. Those nucleic acid-protein fusion multimers that bind the ligand are retained on the column and the unbound multimeric fusions are recovered in the eluate.

20 In the next step, the unbound multimeric fusion complexes contained in the eluate are dissociated and recombined. This is achieved, for example, through a brief heating-cooling process for multimers that combined by nucleic acid hybridization or non-covalent protein-protein interactions. In another example, nucleic acid-protein fusion multimers held together by
25 covalent disulfide bonds can be dissociated through reduction, and reassociated under oxidative conditions. Depending on the secondary and tertiary structure requirements, such conditions may need to allow proper refolding of the peptide domains. Once the re-multimerization takes place,

the newly formed fusion complexes are applied to the affinity column described above. This process can be repeated, as desired, and where applicable, simplified by automation. When the desired number of selection rounds have been completed, the flowthrough is discarded and the selected
5 nucleic acid-protein fusion complexes are eluted from the affinity column. The selected fusion complexes can be affinity eluted by incubating the affinity column with free ligand for an extended period of time. Alternatively, and more preferably, elution may occur by denaturation of the nucleic acid-protein fusion complexes through acid or base treatment, for
10 example, dilute acid, as described by Roberts and Szostak (supra). The column format may be replaced by other suitable methods known in the art.

Example 14: Increasing library diversity through dynamic recombination

Another approach to increasing the diversity of nucleic acid-protein fusion multimers is through a modified form of the dynamic recombination
15 described by Eliseev & Lehn, Curr Top Microbiol Immunol 243:159-72, 1999. This modified method employs individual nucleic acid-protein fusion complexes that are held together by rather weak, non-covalent interactions. In an equilibrium state, these molecules rapidly associate and dissociate, thereby constantly creating new multimeric species. Upon formation of a
20 fusion multimer with the appropriate affinity for a ligand, the multimer binds to the ligand, resulting in an increase in the overall complex stability. By binding to the ligand, the multimer is removed from the equilibrium, and separated from those nucleic acid-protein fusion complexes that are still dissociating and recombining (Figure 7).

25 A specific example of a nucleic acid-protein fusion multimer that may be used in this Example is one that is held together through nucleic acid hybridization, as described above. Depending on the length and sequence of the hybridization domain, the T_m (and hence the association/dissociation

equilibrium state) can be adjusted to the temperature range where the selection process is performed. This results in the dissociation and recombination described above, increasing library complexity.

Example 15: Repeated generation of new diversity through recombination

5 when proceeding from one molecular generation to another

The *in vitro* selection steps described above yield a subset of the original pool of nucleic acid-protein fusion multimers that bind to a desired ligand. As representatives of any multimer, this is herein described for dimers (e.g., A_1-B_1 ; A_2-B_2). During the preparation of the next molecular generation for an additional round of selection, every selected nucleic acid-protein fusion molecule (representing a single domain A or B) is again generated in multiple copies through PCR and transcription, and added to the pool of fusion molecules to be used in the next round of selection. Thus, new diversity is generated for subsequent recombination events (e.g., A_1-B_1 ; A_1-B_2 ; A_2-B_1 ; A_2-B_2 ; Figure 8). Repeated rounds of selection eventually lead to the enrichment of the combination of domains that optimally bind to the desired ligand.

Example 16: Deconvolution of optimal ligand binding domains and final product design

20 To continue the representative example with dimers, after several rounds of *in vitro* selection and selective enrichment, each of the individual pools (A_x , B_x) is reduced to a limited number of individual nucleic acid-protein fusion molecules (e.g., $A_{m,n,o,\dots}$, $B_{p,q,r,\dots}$) that interact with a desired ligand. Yet, the final assignment of the correct dimerization partners (e.g., A_n-B_q) might not be obvious in all the cases. If most of the individual molecules fall into a few classes, for example, they have the same binding domain in the protein portion of the fusion, nucleic acid-protein fusion

multimers can be prepared and tested selectively. If, however, the number of individual nucleic acid-protein molecules is large, any one sequence (e.g., A_n) can be held constant and *in vitro* selection with the whole population (e.g., $B_{p,q,r,\dots}$) can be continued for a few more rounds, to identify the appropriate partner(s) for the particular fusion molecule held constant.

Once the final assignment of dimerization partners is accomplished, it may be desirable to engineer the individual segments into one common molecule. This can be achieved, for example, by mounting the selected peptide domains onto a suitable scaffold (e.g., small organic template molecules, or peptide linkers). In addition, selected light and heavy chains of antibody Fabs can be grafted onto F_c fragments to generate complete antibody structures (IgG).

Example 17: Rational design of nucleic acid-protein fusion multimers

The multimerization techniques described above can also be used for the design of multi-domain peptides or proteins with defined spatial arrangements of their subunits. This can easily be achieved through hybridization of the nucleic acid portions of the fusion molecules with other nucleic acid-protein fusion molecules or with suitable nucleic acid templates. This approach can be used to induce proximity of the multidomain peptides or proteins prior to inter-domain chemical reactions, for example, in a modification of the template-assembled synthetic proteins (TASP) methods for the creation of artificial multifunctional peptides and multi-domain receptors (Tuchscherer et al., Methods Mol. Biol. 36:261-85, 1994). This method can also be used to create antibody constructs analogous to the multivalent miniantibodies described by Plückthun and Pack (Immunotechnology 3:83, 1997).

In addition, the hybridization of nucleic acid-protein fusion molecules with nucleic acid templates does not necessarily require the RNA portion, but may also be completed with only the DNA-linker and protein portions. In

such cases, the RNA can be safely removed with ribonucleases devoid of DNase activity, (e.g., RNase I, Ambion (Austin, TX)) prior to complex assembly.

From the foregoing description, it will be apparent that variations and
5 modifications may be made to the invention described herein to adopt it to various usages and conditions. Such embodiments are also within the scope of the following claims.

All publications and patents mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication
10 or patent was specifically and individually indicated to be incorporated by reference.

Other embodiments are within the claims.

What is claimed is:

Claims

1. A nucleic acid-protein fusion multimer, said multimer comprising two or more fusion molecules of nucleic acid covalently bound to protein, the nucleic acid of at least one of said fusion molecules encoding the covalently
5 bound protein, wherein said fusion molecules are hybridized to each other through complementary nucleic acid sequences.

2. A nucleic acid-protein fusion multimer, said multimer comprising two or more fusion molecules of nucleic acid covalently bound at the 3' end to protein, wherein said fusion molecules are hybridized to each other through
10 complementary nucleic acid sequences.

3. A nucleic acid-protein fusion multimer, said multimer comprising:
(a) two or more fusion molecules of nucleic acid covalently bound to protein, the nucleic acid of at least one of said fusion molecules encoding the covalently bound protein; and
15 (b) an oligonucleotide, wherein a sequence of the nucleic acid of each of said fusion molecules is hybridized to a complementary sequence of said oligonucleotide.

4. A nucleic acid-protein fusion multimer, said multimer comprising:
(a) two or more fusion molecules of nucleic acid covalently bound at the
20 3' end to protein; and
(b) an oligonucleotide, wherein a sequence of the nucleic acid of each of said fusion molecules is hybridized to a complementary sequence of said oligonucleotide.

5. A nucleic acid-protein fusion multimer, said multimer comprising:
(a) two or more fusion molecules of nucleic acid covalently bound to protein; and

(b) an oligonucleotide having a bi-directional or branched structure,
5 wherein a sequence of the nucleic acid of each of said fusion molecules is hybridized to a complementary sequence of said oligonucleotide.

6. A nucleic acid-protein fusion multimer, said multimer comprising:
(a) two or more fusion molecules of nucleic acid covalently bound to protein, wherein the nucleic acid of each of said fusion molecules comprises a
10 polypurine tract; and

(b) an oligonucleotide comprising at least two polypyrimidine tracts,
wherein said polypurine tracts of said fusion molecules are hybridized to said polypyrimidine tracts of said oligonucleotide, and wherein binding of said fusion molecules to said oligonucleotide occurs in opposite directions to form
15 a triple helical structure.

7. The nucleic acid-protein fusion multimer of claim 6, wherein said oligonucleotide is circular.

8. The nucleic acid-protein fusion multimer of claim 6, wherein said oligonucleotide forms a clamp-like structure.

20 9. The nucleic acid-protein fusion multimer of claim 6, wherein said oligonucleotide comprises polyamide nucleic acids.

10. A nucleic acid-protein fusion multimer, said multimer comprising two or more fusion molecules of nucleic acid covalently bound to protein, wherein the protein portions of said fusion molecules each comprise a multimerization domain, the multimerization domains interacting through
5 non-covalent bond formation.

11. The nucleic acid-protein fusion multimer of claim 10, wherein said multimerization domains interact to form homodimers, heterodimers, trimers, or tetramers.

12. A nucleic acid-protein fusion multimer, said multimer comprising
10 two or more fusion molecules of nucleic acid covalently bound to protein, wherein the protein of each of said fusion molecules comprises a multimerization domain that includes a functional group, the functional group of one fusion molecule being linked to the functional group of another fusion molecule through a covalent bond.

15 13. The nucleic acid-protein fusion multimer of claim 12, wherein said multimerization domain comprises an antibody constant region.

14. The nucleic acid-protein fusion multimer of claim 1 or 3, wherein the protein of at least one of said fusion molecules further comprises a compound recognition domain.

20 15. The nucleic acid-protein fusion multimer of claim 14, wherein said compound recognition domain comprises an antibody variable region.

16. The nucleic acid-protein fusion multimer of claim 14, wherein said compound recognition domain comprises a randomized domain.

17. The nucleic acid-protein fusion multimer of claim 14, wherein said compound recognition domain interacts with DNA.

18. The nucleic acid-protein fusion multimer of claim 17, wherein said compound recognition domain comprises a zinc finger binding domain.

5 19. An RNA-protein fusion multimer, said multimer comprising two or more fusion molecules of RNA covalently bound to protein, wherein said fusion molecules are hybridized to each other through complementary nucleic acid sequences.

20. The RNA-protein fusion multimer of claim 19, wherein the RNA of
10 at least one of said fusion molecules encodes the covalently bound protein.

21. The RNA-protein fusion multimer of claim 19, wherein said fusion molecules are cross-linked through cross-linking moieties positioned within the RNA of said fusion molecules.

22. The RNA-protein fusion multimer of claim 21, wherein said cross-
15 linking moiety is psoralen.

23. A method for preparing a nucleic acid-protein multimer, said method comprising the steps of:

(a) providing two or more fusion molecules of nucleic acid covalently bound to protein, the nucleic acid of at least one of said fusion molecules
20 encoding the covalently bound protein; and

(b) hybridizing said fusion molecules to each other through complementary nucleic acid sequences, thereby forming a nucleic acid-protein fusion multimer.

24. A method for preparing a nucleic acid-protein fusion multimer, said method comprising the steps of:

- (a) providing two or more fusion molecules of nucleic acid covalently bound at the 3' end to protein; and
- 5 (b) hybridizing said fusion molecules to each other through complementary nucleic acid sequences, thereby forming a nucleic acid-protein fusion multimer.

25. A method for preparing a nucleic acid-protein fusion multimer, said method comprising the steps of:

- 10 (a) providing two or more fusion molecules of nucleic acid covalently bound to protein, the nucleic acid of at least one of said fusion molecules encoding the covalently bound protein;
- (b) providing an oligonucleotide, wherein said oligonucleotide comprises a plurality of sequences that are complementary to a partner
- 15 sequence within each of said fusion molecules; and
- (c) hybridizing said oligonucleotide to each of said fusion molecules, thereby forming a nucleic acid-protein fusion multimer.

26. A method for preparing a nucleic acid-protein fusion multimer, said method comprising the steps of:

- 20 (a) providing two or more fusion molecules of nucleic acid covalently bound at the 3' end to protein;
- (b) providing an oligonucleotide, wherein said oligonucleotide comprises a plurality of sequences that are complementary to a partner sequence within each of said fusion molecules; and
- 25 (c) hybridizing said oligonucleotide to each of said fusion molecules, thereby forming a nucleic acid-protein fusion multimer.

27. A method for preparing a nucleic acid-protein fusion multimer, said method comprising the steps of:

(a) providing two or more fusion molecules of nucleic acid covalently bound to protein through a peptide acceptor, wherein said fusion molecules
5 are hybridized to each other through complementary nucleic acid sequences;

(b) providing an oligonucleotide, wherein said oligonucleotide comprises a plurality of sequences that are complementary to a partner sequence within each of said fusion molecules; and

(c) hybridizing said oligonucleotide to each of said fusion molecules,
10 thereby forming a nucleic acid-protein fusion multimer.

28. A method for preparing a nucleic acid-protein fusion multimer, said method comprising the steps of:

(a) providing two or more fusion molecules of nucleic acid covalently bound to protein;

15 (b) providing an oligonucleotide having a bi-directional or branched structure, wherein said oligonucleotide comprises a plurality of sequences that are complementary to a partner sequence within each of said fusion molecules; and

(c) hybridizing said oligonucleotide to each of said fusion molecules,
20 thereby forming a nucleic acid-protein fusion multimer.

29. A method for preparing a nucleic acid-protein fusion multimer, said method comprising the steps of:

(a) providing two or more fusion molecules of nucleic acid covalently bound to protein, wherein the nucleic acid of each of said fusion molecules
25 comprises a polypurine tract;

(b) providing an oligonucleotide comprising at least two polypyrimidine tracts; and

(c) hybridizing said polypurine tracts to said polypyrimidine tracts, wherein binding of said fusion molecules to said oligonucleotide occurs in opposite directions to form a triple helical structure, thereby forming a nucleic acid-protein multimer.

5 30. A method for preparing a nucleic acid-protein fusion multimer, said method comprising the steps of:

(a) providing two or more fusion molecules of nucleic acid covalently bound to protein, wherein the protein portions of said fusion molecules each comprise a multimerization domain; and

10 (b) combining said fusion molecules under conditions that allow non-covalent interactions between said multimerization domains, thereby forming a nucleic acid-protein fusion multimer.

31. A method for preparing a nucleic acid-protein fusion multimer, said method comprising the steps of:

15 (a) providing two or more fusion molecules of nucleic acid covalently bound to protein, wherein the protein of each of said fusion molecules comprises a multimerization domain that includes a functional group; and

(b) combining said fusion molecules under conditions that allow the functional group of one fusion molecule to be linked to the functional group
20 of another fusion molecule through a covalent bond, thereby forming a nucleic acid-protein fusion multimer.

32. The method of claim 31, wherein the covalent linkage comprises an external cross-linking agent.

33. A method for preparing an RNA-protein fusion multimer, said method comprising the steps of:

- (a) providing two or more fusion molecules of RNA covalently bound to protein; and
- 5 (b) hybridizing said fusion molecules to each other through complementary nucleic acid sequences, thereby forming an RNA-protein fusion multimer.

34. A method of selecting a protein that interacts with a compound, said method comprising the steps of:

- 10 (a) providing a population of candidate nucleic acid-protein fusion multimers, said multimers comprising two or more hybridized or covalently bound fusion molecules of nucleic acid covalently bound to protein;
- (b) providing a compound;
- (c) contacting said compound with said population of candidate nucleic
- 15 acid-protein fusion multimers under conditions that allow an interaction between said compound and said candidate nucleic acid-protein fusion multimers; and
- (d) selecting a nucleic acid-protein fusion multimer that interacts with said compound, thereby selecting a protein that interacts with said compound.

- 20 35. The method of claim 34, wherein said compound is immobilized on a column.

36. The method of claim 35, further comprising the steps of:
- (e) dissociating the nucleic acid-protein fusion multimers that do not interact with said compound;
 - (f) recombining said dissociated nucleic acid-protein fusion multimers;
 - 5 (g) contacting said compound with said recombined nucleic acid-protein fusion multimers; and
 - (h) selecting a recombined nucleic acid-protein fusion multimer that interacts with said compound, thereby selecting a protein that interacts with said compound.
- 10 37. The method of claim 34, wherein said population, in step (a), is maintained under equilibrium conditions, whereby the individual fusion molecules of said nucleic acid-protein fusion multimers rapidly dissociate and associate with other individual fusion molecules, thereby forming new nucleic acid-protein fusion multimers.
- 15 38. The method of claim 34, further comprising the steps of:
- (e) amplifying the nucleic acids of said nucleic acid-protein fusion multimers selected in step (d);
 - (f) generating, from said amplified nucleic acids, fusion molecules of nucleic acid covalently bound to protein;
 - 20 (g) generating from those fusion molecules a second population of nucleic acid-protein fusion multimers; and
 - (h) repeating steps (b) through (d).
39. The method of claim 34, wherein said compound interacts with said nucleic acid-protein fusion multimer in solution and is subsequently
- 25 immobilized on a solid phase.

Figure 1A

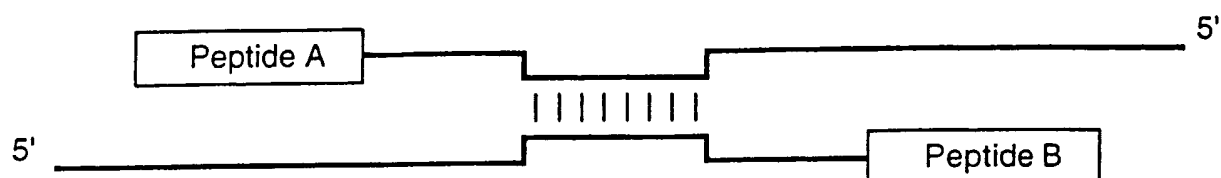


Figure 1B

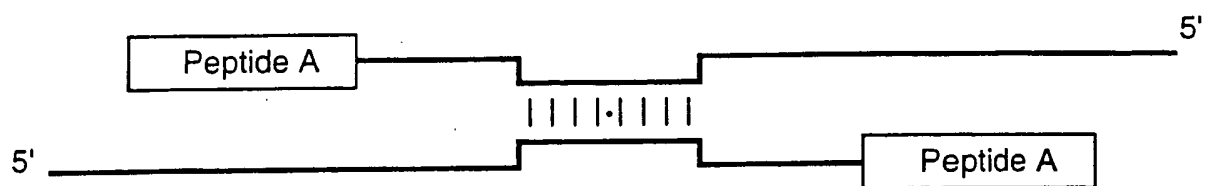


Figure 2A

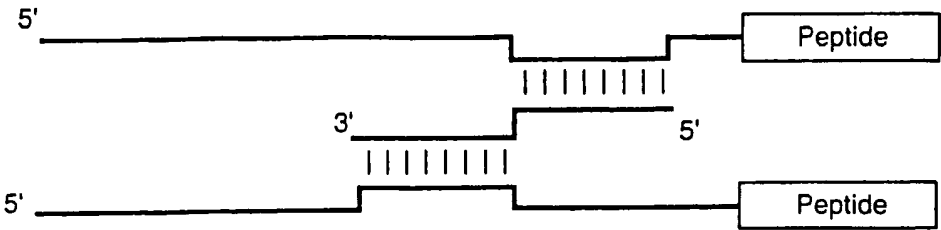


Figure 2B

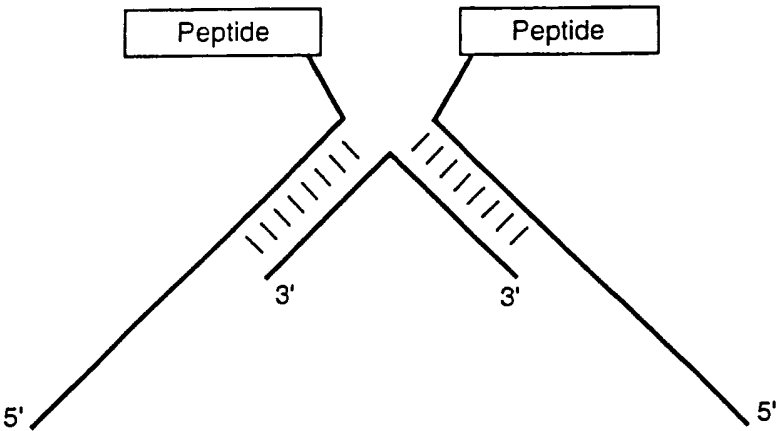


Figure 2C

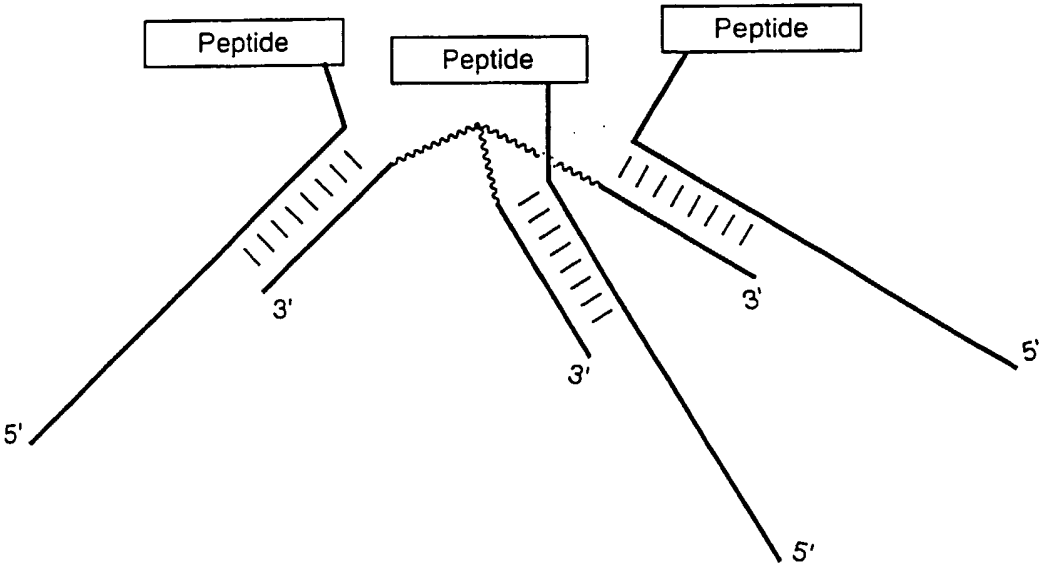


Figure 3A

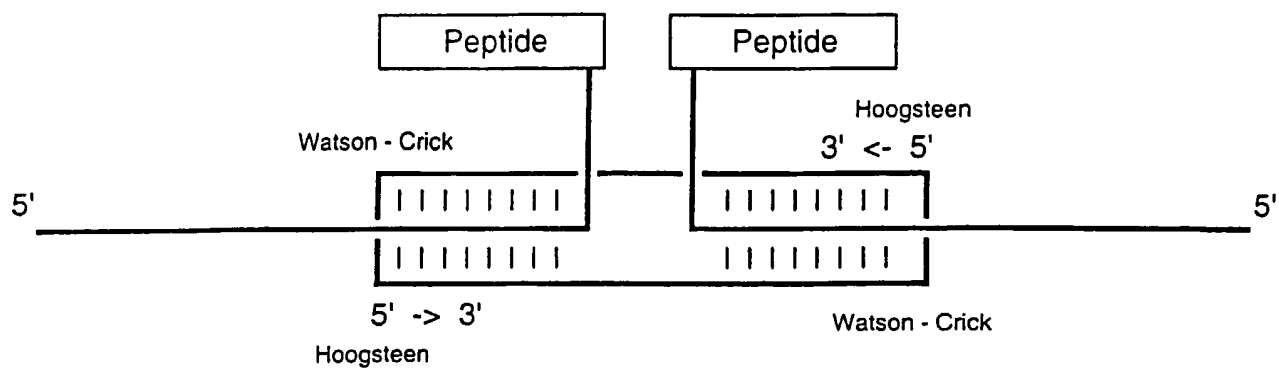


Figure 3B

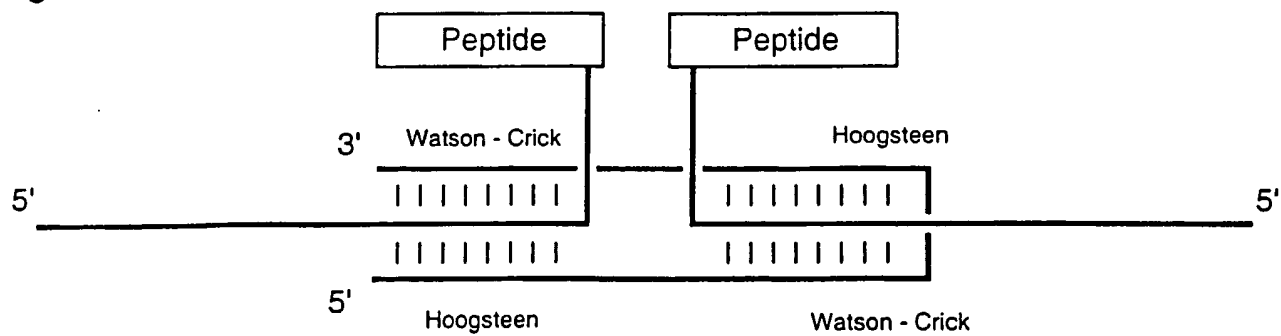


Figure 4A

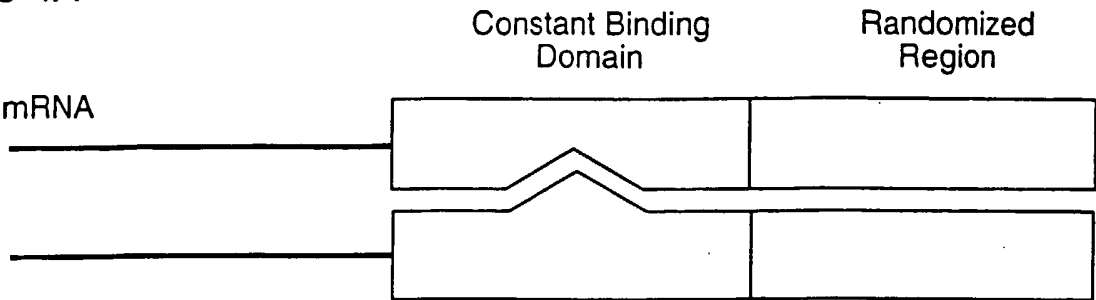


Figure 4B Antibody FAB

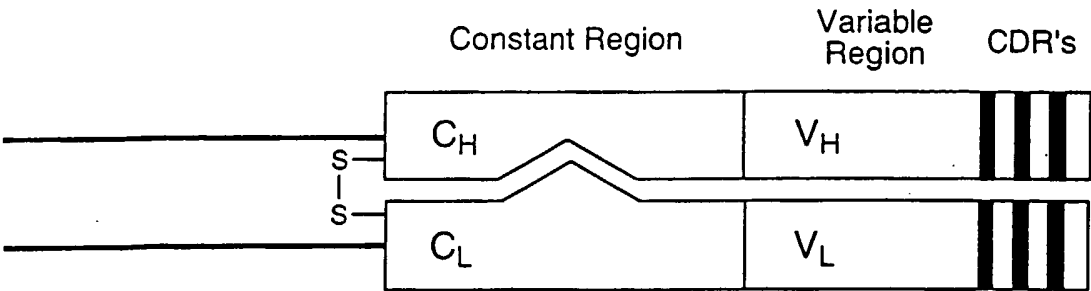
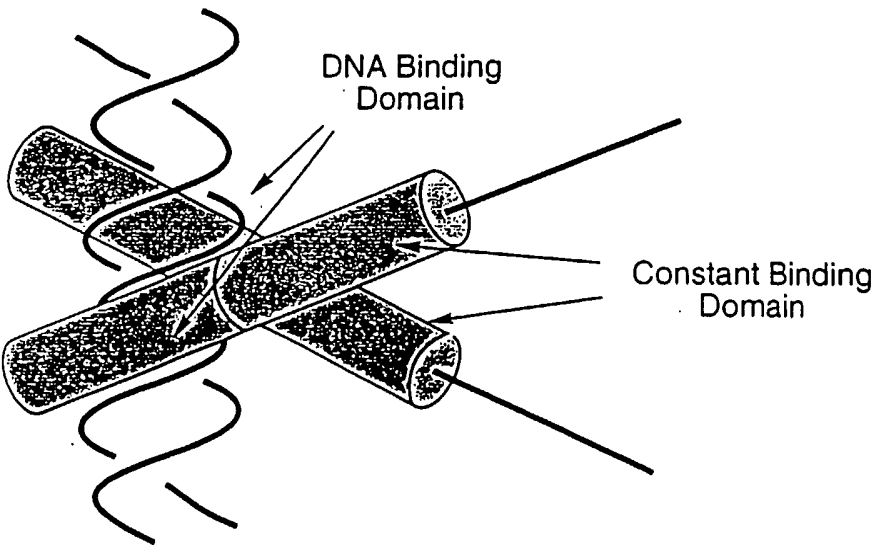


Figure 4C



5/8

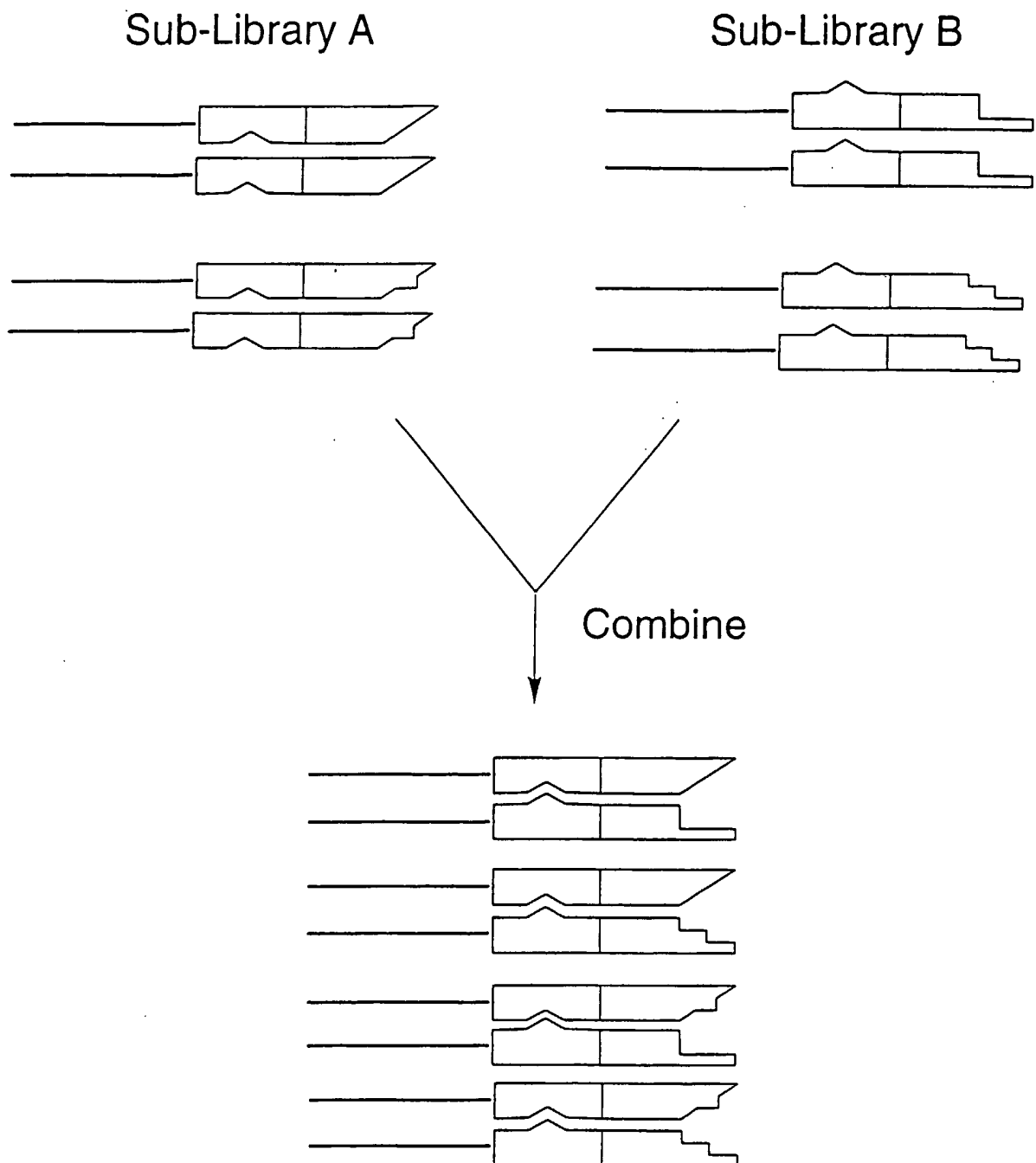


Figure 5

6/8

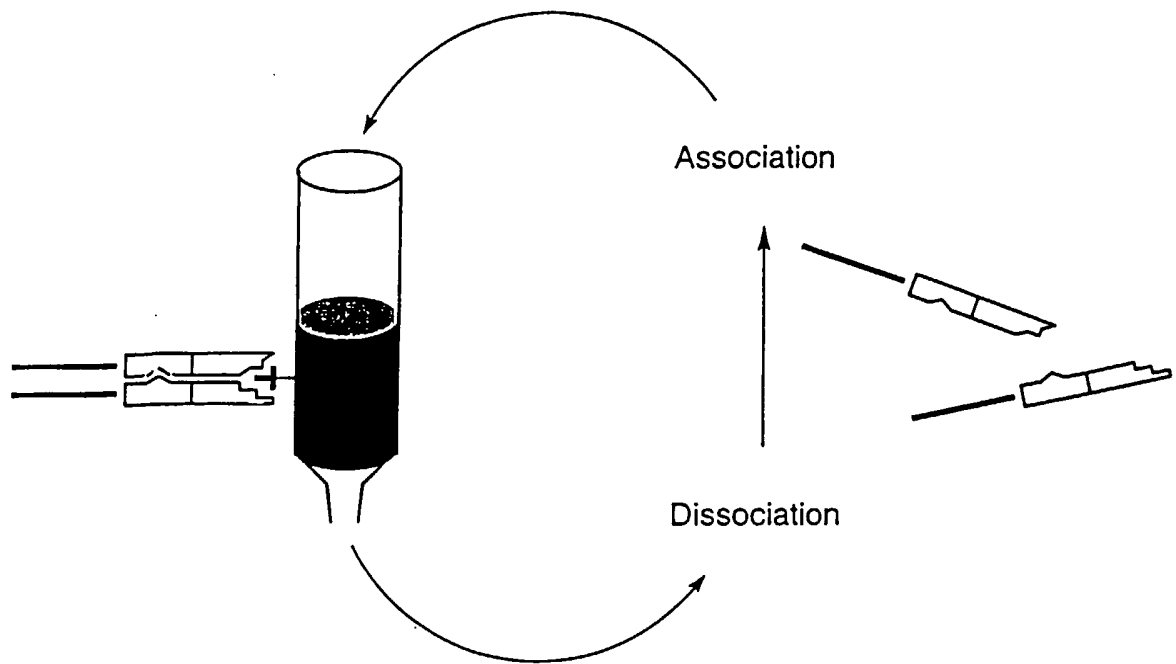


Figure 6

7/8

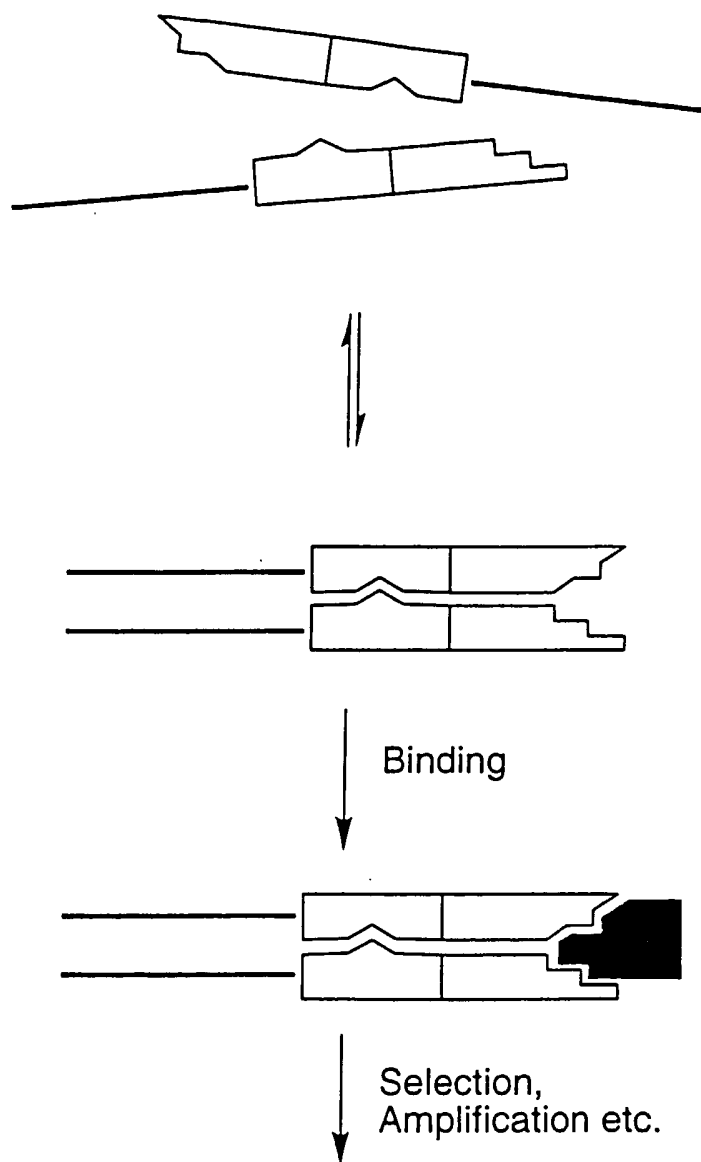


Figure 7

8/8

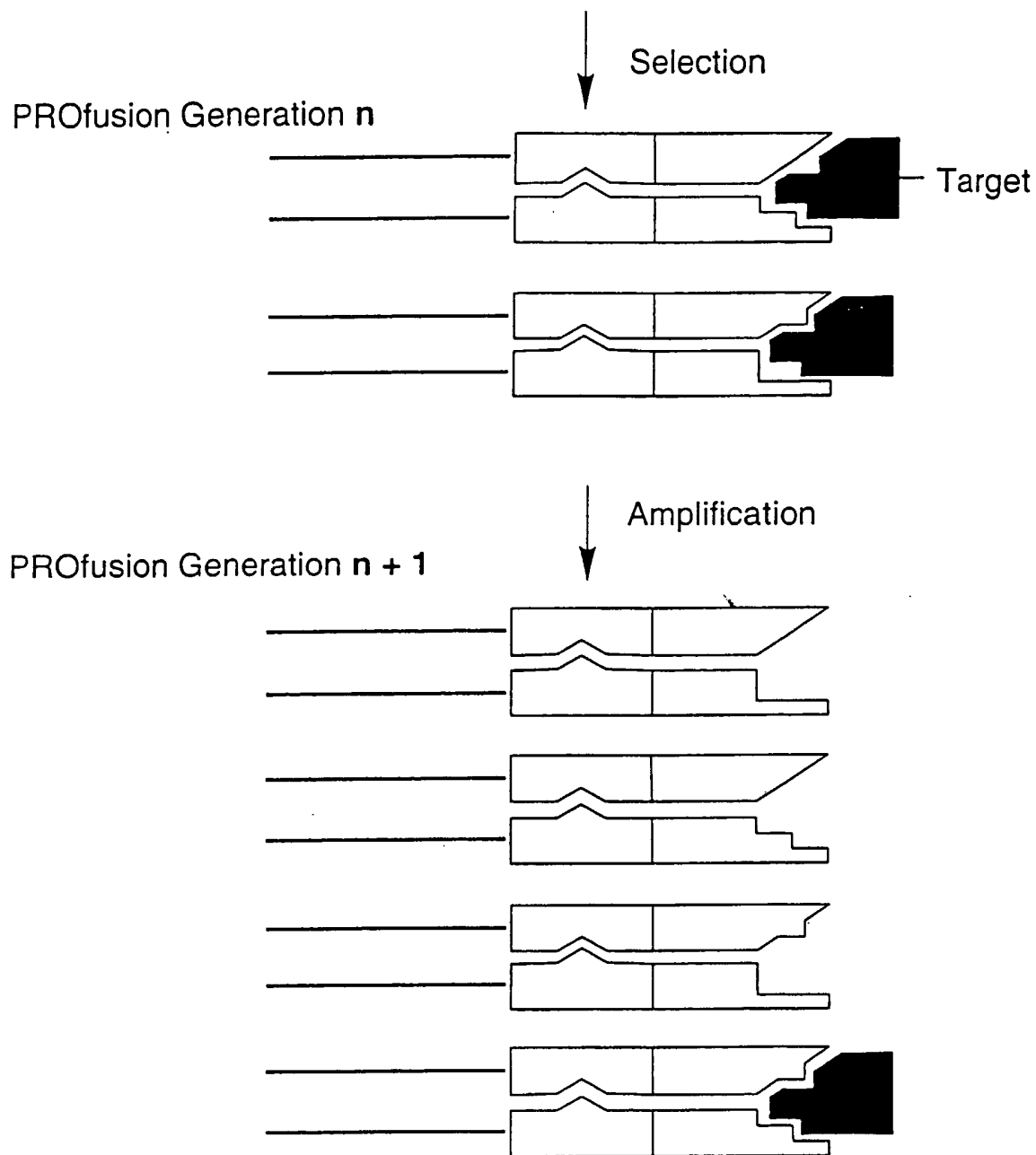


Figure 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/24180

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12Q 1/68; C07H 21/04, 21/02; G01N 15/06

US CL : 435/6; 536/23.1, 530/300

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. : 435/6; 536/23.1, 530/300

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)

WEST 2.0. CAPLUS, MEDLINE, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,843,701 A1 (GOLD et al) 01 Dec. 1998 (01.12.1998) see entire patent, especiall columns 1-14.	16
A,P	US 6,416,950 B1, 09 July 2002 (09.07.2002) See entire patent.	5, 6, 16, 18, 20, 21,22
A,E	US 6,436,665 B1 (KUIMELIS) 20 Aug. 2002 (20.08. 2002) see entire patent.	10-13, 20
A, E	US 6,429,300 B1 (KURZ et al.) 06 Aug. 2002 (02.08.2002) see entire patent.	8-10, 22
Y	SELVASEKARAN et al. Chemical ligation of oligonucleotides on circular DNA templates. Nucleic acids Research, Vol. 27, No. 2, pages 624-627, 1999.	7
Y	BRENNER AND LERNER, Encoded combinatorial chemistry, Proc. Natl. Acad. Sci. USA. June 1992, Vol. 89, pages 5381-5383, see entire reference.	5, 6
X	US 6,261,804 B1 (SZOSTAK et al) 17 July 2001 (17.07.2001) see entire reference especially columns 1-4.	1-4, 14, 17, 19, 20
---		-----
Y		5-13, 15, 16, 18, 21, 22



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

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

"E" earlier application or patent published on or after the international filing date

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

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

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

"T"

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

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

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

"&"

document member of the same patent family

Date of the actual completion of the international search

03 November 2002 (03.11.2002)

Date of mailing of the international search report

18 DEC 2002

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703)305-3230

Authorized officer

Valerie Bell-Harris for
Gary Benzion

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

PCT/US02/24180

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 6,214,553 B1 (SZOSTAK et al) 10 April 2001 (10.04.2001) entire patent, especially columns 1-4.	1-4, 14, 17, 19, 20 ----- 5-13, 15, 16, 18, 21, 22

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/24180

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

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

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-22

Remark on Protest

☐
☐

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

PCT/US02/24180

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group 1, claim(s) 1-22, drawn to a nucleic acid-protein fusion multimer.

Group 2, claim(s) 23-33, drawn to method of making nucleic acid-protein multimer.

Group 3, claim(s) 34-39, drawn to a method for selecting protein that interacts with compound.

The inventions listed as Groups 1-3 do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: The groups 1-3 do not relate to a single general inventive concept because the "special technical" is not new and does not constitute a contribution over the art. PCT Rule 13.2 requires that unity of invention only exist when a shared technical feature is a contribution of the art (and therefore constitutes a "special technical" feature). As nucleic acid-protein fusion multimer were known in the art (see Szostak et al. [US 6,214,553 B1, April 10, 2001, column 4]), the technical feature does not constitute a special technical feature and unity of the invention is lacking.