

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
2 December 2004 (02.12.2004)

PCT

(10) International Publication Number
WO 2004/104194 A1

(51) International Patent Classification⁷: **C12N 15/00**,
15/09, C07H 21/04

Research Center, 2-41-6 Aomi, Koto-Ku, Tokyo 135-0064 (JP).

(21) International Application Number:
PCT/US2003/018219

(74) Agents: **ESMOND, Robert W.** et al.; Sterne, Kessler, Golstein & Fox P.L.L.C., 1100 New York Avenue, NW, Washington, DC 20005 (US).

(22) International Filing Date: 12 June 2003 (12.06.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/470,448 15 May 2003 (15.05.2003) US

(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, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

(71) Applicants: **INVITROGEN CORPORATION** [US/US]; 1600 Faraday Avenue, Carlsbad, CA 92008 (US). **NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY (AIST)**, **JAPAN BIOLOGICAL INFORMATION RESEARCH CENTER (JBIRC)** [JP/JP]; 41-6 Aomi, 2-Chome, Kohto-Ku., Tokyo, 135-0064 (JP).

(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, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors: **MIURA, Aya**; Japan Biological Information Research Center, 2-41-6 Aomi, Koto-Ku, Tokyo 135-0064 (JP). **SATO, Ryohei**; Japan Biological Information Research Center, 2-41-6 Aomi, Koto-Ku, Tokyo 135-0064 (JP). **NOMURA, Nobuo**; Japan Biological Information Research Center, 2-41-6 Aomi, Koto-Ku, Tokyo 135-0064 (JP). **GOSHIMA, Naoki**; Japan Biological Information

Published:

— with international search 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: COMPOSITIONS AND METHODS FOR SITE-SPECIFIC RECOMBINATION AND EXPRESSION

(57) Abstract: The present invention provides materials and methods for the *in vitro* transcription and translation of nucleic acid sequence of interest. Methods may include a recombination reaction between a nucleic acid comprising a sequence of interest and an additional nucleic acid molecule comprising one or more recombination sites. The invention also includes kits for performing the methods of the invention. The invention also provides methods of conducting a business for providing nucleic acid molecules and polypeptides produced according to the methods of the invention.



WO 2004/104194 A1

COMPOSITIONS AND METHODS FOR SITE-SPECIFIC RECOMBINATION AND EXPRESSION

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] The present invention is in the fields of molecular biology and protein chemistry. In general, the invention relates to *in vitro* production of proteins. Methods of the invention may include one or more recombination reactions to create one or more nucleic acid molecules, which may then be used as templates for *in vitro* transcription and translation.

Related Art

[0002] The techniques of modern molecular biology have made possible manipulation of DNA as well as the other cellular components expressed from the genes contained in the DNA of an organism. In living cells, DNA is transcribed to make mRNA which is then used as a template for translation whereby a protein whose sequence is determined by the DNA is made. In developing the modern tools of molecular biology, research has been directed toward ways of performing various steps of the transcription and/or translation processes *in vitro* under controlled conditions and with defined inputs. These procedures mimic, in essence, similar processes that occur in a much more heterogeneous mixture in living cells.

[0003] *In vitro* protein synthesis has among its advantages specifically producing desired proteins without producing undesired proteins that may be required for maintaining cells used for protein production in *in vivo* or cellular systems. Cell-free systems are very popular because there are standard protocols available for their preparation and because they are commercially available from a number of sources.

[0004] Because it is essentially free from cellular regulation of gene expression, *in vitro* protein synthesis has advantages in the production of cytotoxic, unstable and/or insoluble proteins. The over-production of protein beyond a predetermined concentration can be difficult to obtain *in vivo*, because the expression levels are often regulated by the intracellular concentration of product. The concentration of protein accumulated in the cell generally affects the viability of the cell, so that over-production of the desired protein is often difficult to obtain. In an isolation and purification process, many proteins are insoluble or unstable and are either degraded by intracellular proteases or aggregate in inclusion bodies, so that the loss rate is high. *In vitro* synthesis circumvents many of these problems. Also, through simultaneous and rapid expression of various proteins in a multiplexed configuration, this technology can provide a valuable tool for development of combinatorial arrays for research, and for screening of proteins. In addition, various kinds of unnatural amino acids can be efficiently incorporated into proteins for specific purposes (Noren *et al.*, *Science* 244: 182-188 (1989)).

[0005] Protein synthesis is guided by an RNA polynucleotide template that encodes the desired protein. A DNA template comprising a sequence of interest encoding the desired protein is transcribed to produce the RNA polynucleotide that encodes the desired protein. The DNA template therefore minimally includes the DNA sequence of interest to be transcribed as well as a binding site for RNA polymerase (*e.g.*, a promoter) that catalyzes transcription of the mRNA template that is subsequently translated to produce protein. When the transcription and translation are coupled in one system, the system is called an *in vitro* Transcription Translation (IVTT) system.

[0006] IVTT or protein synthesis using cell-free extracts is becoming an important tool for analysis of proteins. The availability of complete genome sequences for a variety of organisms provides a wealth of information on the molecular structure and organization of a myriad of genes and open reading frames whose functions are not known or are only poorly understood. Thus, the utility of IVTT and more generally, protein synthesis *in vitro*, is expected to be even more important in the future for rapid and efficient protein synthesis and functional analysis.

- [0007] Even before the advent of modern genetic engineering, cell extracts were developed which allowed the synthesis of protein *in vitro* from purified mRNA transcripts. To date, several systems have become available for the study of protein synthesis and RNA structure and function. To use these systems to synthesize a protein under investigation, a translation extract must be "programmed" with an mRNA corresponding to the gene and protein under investigation. The mRNA can be produced from DNA, or the mRNA can be added exogenously in purified form. Historically, such mRNA templates were purified from natural sources or, using more recently developed technologies, prepared synthetically from cloned DNA using bacteriophage RNA polymerases in an *in vitro* reaction.
- [0008] More recently, techniques using coupled or complementary transcription and translation systems in which synthesis of both RNA and protein are carried out in the same reaction have been developed (IVTT). The cell extracts used for the modern techniques must contain all the components necessary both for transcription (to produce RNA) and for translation (to synthesize protein) in a single system. In such a system, the input template is DNA, which is normally much easier to obtain than RNA and much more readily manipulable.
- [0009] An early coupled system was based on a bacterial extract (Lederman and Zubay, *Biochim. Biophys. Acta*, 149: 253 (1967)). Since prokaryotes normally carry out a coupled reaction within their cytoplasm, this bacterial based system closely reflected the *in vivo* process. This general system has seen widespread use for the study of prokaryotic genes. However, this general bacterial system is typically not useful for eukaryotic genes, due to its inefficiency and relatively high nuclease content.
- [0010] In addition to prokaryotic system extracts, systems employing extracts from eukaryotic cells have been developed. These eukaryotic systems may be supplemented with an RNA polymerase, which may be prokaryotic (*e.g.*, *E. coli* RNA polymerase) or eukaryotic (*e.g.*, wheat germ RNA polymerase) in origin, to transcribe exogenous DNA. These systems have been used with limited success for the general study of eukaryotic genes. The use of these systems has been limited due to their low efficiency, and to the fact that they were developed and used prior to the widespread success of cDNA cloning

techniques. Other coupled systems have been developed for the study of viral protein synthesis, but are not generally useful for non-viral templates.

[0011] In the mid-1980s, the development of more efficient *in vitro* transcription systems, particularly ones using phage polymerases such as T7, SP6 and T3, resulted in more defined protein synthesis systems that more efficiently translated cloned mRNA sequences *in vitro* using translation extracts from wheat germ and rabbit reticulocytes. For example, Perara and Lingappa (*J. Cell Biol.*, 101: 2292-2301 (1985)) showed that SP6 RNA polymerase transcription reactions could be added directly to reticulocyte lysate for the production of protein.

[0012] Later, other workers showed that transcription and translation could be coupled in reticulocyte lysate by including a phage polymerase and appropriate transcriptional co-factors in the reaction (Spirin *et al.*, *Science*, 242: 1162-1164 (1988); Craig *et al.*, *Nucleic Acids Res.*, 20: 4987-4995 (1992)). More recently, U.S. Pat. No. 5,324,637 to Thompson *et al.* disclosed a coupled transcription and translation system in eukaryotes. Thompson used reticulocyte lysate and a phage polymerase where the coupling of the two reactions was facilitated by specific conditions, notably the concentration of magnesium ions, which permitted both transcription and translation to occur in the same reaction mix.

[0013] A number of subsequent improvements to *in vitro* expression systems have been reported (see *e.g.*, Kim *et al.*, *Eur. J. Biochem.*, 239: 881-886, (1996); Patnaik and Swartz, *Biotechniques*, 24: 862-868, (1998); and Kim and Swartz, *Biotech. and Bioeng.*, 66: 180-188, (1999)).

[0014] *E. coli* extract for cell-free protein synthesis has been made using a RecD mutant of *E. coli* (Lesley *et al.*, *J. Biol. Chem.*, 266: 2632-2639 (1991)). An RNase I deficient mutant has been used as a potential source of material for IVTT, but with unsatisfactory results with respect to efficient protein production (Kudlicki *et al.*, *Anal. Biochem.*, 206: 389-393 (1992); and Ellman *et al.*, *Methods in Enzymology*, 202: 301-336 (1991)). WO 00/55353 discloses two methods for replenishing ATP necessary for translation in an *in vitro* expression system. According to these methods, PEP (phosphoenolpyruvate) or pyruvate is used to regenerate the ATP energy source. The first method disclosed in WO 00/55353 was previously known in the art and involves PEP

used in conjunction with pyruvate kinase to regenerate ATP from ADP. In the second method of WO 00/55353, pyruvate is used in conjunction with pyruvate oxidase to regenerate ATP.

[0015] It has been shown in a wheat germ cell-free system that phosphatase-immunodepletion improved protein synthesis by reducing ATP and GTP hydrolysis (Kawarasaki *et al.*, *J. Biotech.* 61: 199-208, (1998)). It has also been suggested that PEP, the substrate generally used for regeneration of ATP, is also degraded by phosphatase(s) limiting protein synthesis (Kim and Swartz, *Biotech. and Bioeng.*, 66: 180-188, (1999); Swartz and Kim, U.S. Patent No. 6,168,931 (2001); Kim and Choi, *J. Biotech.* 84: 27-32, (2000)).

[0016] WO 02/072890 and published United States patent application 2002-0168706 A1 describe an improved IVTT system that enhances the production of proteins. The system disclosed provides compositions and methods for stabilizing or maintaining template DNA, stabilizing or maintaining mRNA including mRNA derived from the template(s), conserving energy to be used for synthesis and/or providing sufficient energy regenerating substrates to provide the energy necessary for efficient protein synthesis from the templates.

[0017] Regardless of the transcription/translation system used, art-known methods still require the construction of suitable input DNA template from which transcription may occur. This is typically accomplished using conventional cloning techniques. A typical method may require restriction digestion of a DNA source and vector, isolation of desired fragments, ligation of the desired fragments with a vector, transformation of a competent host, screening of the transformed hosts to identify a desired clone, growth of a stock of host transformed with the desired clone, and isolation of cloned plasmid DNA from the transformed host. This process may be shortened somewhat using more recently developed cloning techniques such as topoisomerase cloning (see, for example, United States patent no. 5,766,891; PCT/US95/16099; and PCT/US98/12372) and/or recombinational cloning (see, for example, United States patent nos. 5,888,732; 6,143,557; 6,171,861; 6,270,969; 6,277,608; and published United States application no. 2002 0007051-A1); however, these methodologies still require the time consuming processes of transformation of a competent host, screening of the transformed

hosts to identify a desired clone, growth of a stock of host transformed with the desired clone, and isolation of cloned plasmid DNA from the transformed host.

[0018] Given the large number of open reading frames (ORFs) and potential open reading frames identified by sequencing the human genome, there exists a need in the art to produce and analyze proteins encoded by these ORFs. Ideally, methods for the production of proteins encoded by these ORFs should be amenable to high throughput applications in order to facilitate the analysis. Thus, there also remains a need in the art for more rapid methods of generating proteins encoded by ORFs. Art-known methods involving IVTT are limited in utility by the methods used to construct an input DNA template molecule from which transcription can occur. Thus, there also remains a need in the art for more rapid methods of producing an input DNA template from which transcription may occur for use in *in vitro* transcription and/or translation systems.

SUMMARY OF THE INVENTION

[0019] The present invention relates to materials and methods for the construction of nucleic acid molecules (*e.g.*, DNA molecules) and, optionally, the use of nucleic acid molecules produced using materials and methods of the invention in the production of nucleic acid molecules (*e.g.*, DNA and/or RNA molecules) and/or proteins. Optionally, one or more methods of the invention may be practiced *in vitro*. In one aspect, methods of the invention may be practiced *in vitro*. The present invention also relates to nucleic acid molecules and/or proteins generated using methods of the invention and to compositions comprising such nucleic acid molecules and/or proteins. The present invention also includes kits comprising components useful in practicing the methods of the invention as well as reaction mixtures used to perform such methods.

[0020] In one aspect, the present invention provides methods of producing nucleic acid molecules (*e.g.*, DNA and/or RNA molecules) and/or polypeptides. Such methods may include contacting a first nucleic acid molecule, comprising a sequence of interest, which may encode one or more

(*e.g.*, two, three, four, five, six, eight, ten, etc.) polypeptides or portion thereof, and comprising one or more (*e.g.*, two, three, four, five, six, eight, ten, etc.) recombination sites, with a second nucleic acid molecule comprising at least one (*e.g.*, two, three, four, five, six, eight, ten, etc.) recombination sites under conditions causing recombination of the first and second nucleic acid molecules to produce a third nucleic acid molecule. The third nucleic acid molecule may be used as a template in subsequent reactions. For example, a third nucleic acid molecule generated using methods of the invention may be suitable for logarithmic amplification (*e.g.*, by PCR) and/or may be suitable for use as a template in an *in vitro* transcription reaction. Optionally, a third nucleic acid molecule may be amplified (*e.g.*, by PCR) and the amplification product may be used as a template for an *in vitro* transcription reaction.

[0021] In one aspect, a second nucleic acid molecule may comprise one or more (*e.g.*, two, three, four, five, six, eight, ten, etc.) transcriptional regulatory sequences. Thus, a third nucleic acid molecule produced using methods of the invention may comprise a transcriptional regulatory sequence and a sequence of interest, which may be operably linked. In one aspect, methods of the invention may be used to generate a third nucleic acid molecule comprising a transcriptional regulatory sequence operably linked to a sequence of interest. Optionally, such a third nucleic acid molecule may be transcribed into RNA. RNA corresponding to a third nucleic acid molecule may be used, for example, as RNA (*e.g.*, antisense RNA, siRNA, etc.) or may be translated into one or more polypeptides. Thus, in some embodiments, all or a portion of the sequence of interest may be transcribed into RNA. The RNA thus produced may be translated into one or more polypeptides. In one aspect, transcription and/or translation may be done *in vitro*.

[0022] Optionally, a third nucleic acid molecule generated according to methods of the invention is not introduced into a cell. In another aspect, a third nucleic acid molecule generated by methods of the invention may be amplified and the amplification product used in an *in vitro* transcription reaction to produce RNA and the RNA translated *in vitro* into one or more polypeptides all without introduction of any of the molecules involved (*e.g.*, first nucleic acid molecule, second nucleic acid molecule, third nucleic acid

molecule, amplification product, corresponding RNA) being introduced into a cell.

[0023] In some instances, the second nucleic acid molecule may be an expression vector. Such a second nucleic acid molecule may be a closed circle, which may be supercoiled, or may be linear. One skilled in the art will appreciate that a closed circular nucleic acid molecule may be converted to a linear molecule, for example, by digestion with one or more restriction enzymes. Linear second nucleic acid molecules may be made by other methods known to those skilled in the art, for example, by PCR. In some embodiments, the second nucleic acid molecule may be an expression vector and may comprise at least two recombination sites. The second nucleic acid molecule may be digested with one or more restriction enzyme prior to recombination with a first nucleic acid molecule. Any restriction enzyme recognizing a site in the second nucleic acid molecule may be used in these embodiments. For example, restriction enzymes suitable for use in methods of the invention include, but are not limited to, *Bgl*II and *Pst*I. In some embodiments, the second nucleic acid molecule may be a vector having two recombination sites and at least one site recognized by a restriction enzyme is located between the two recombination sites. Digestion of such a second nucleic acid molecule results in a linear nucleic acid molecule in which the recombination sites are "tethered" (*i.e.*, are in the same nucleic acid molecule). When such a second nucleic acid molecule is reacted in a recombination reaction with a first nucleic acid molecule comprising two recombination sites, the tethered sites are re-circularized to produce a third nucleic acid molecule that can be amplified using primers that hybridize to the second nucleic acid molecule such that no amplification product can be produced in the absence of a recombination reaction.

[0024] Optionally, transcriptional regulatory sequences in second nucleic acid molecules for use in the present invention may comprise a promoter. Any promoter known to those skilled in the art may be used, for example, the promoter may be a phage promoter. Promoters suitable for use in methods of the invention include, but are not limited to, phage promoters, viral promoters, and other promoters known to those skilled in the art. Examples include, but are not limited to, the T7 promoter and the SP6 promoter. Other

transcriptional regulatory sequences suitable for use in the present invention include, but are not limited to, enhancers, repressors, operators, suppressors, and the like.

[0025] Optionally, methods of the invention may comprise amplifying all or a portion of the third nucleic acid molecule.

[0026] In some embodiments, the present invention provides methods of making one or more desired nucleic acid molecules. Such methods may include contacting a first nucleic acid molecule comprising a sequence of interest, which may encode one or more polypeptides or may not encode a polypeptide, and comprising one or more recombination sites, with a second nucleic acid molecule comprising one or more recombination sites and, optionally, one or more transcriptional regulatory sequence, under conditions causing recombination of the first and second nucleic acid molecules to produce a desired nucleic acid molecule. In the desired nucleic acid molecule, the sequence of interest and transcriptional regulatory sequence (when present) may be operably linked. A method of this type may entail amplifying all or a portion of the desired nucleic acid molecule. In one such aspect, the desired nucleic acid molecule is not introduced into a cell before being amplified.

[0027] In certain such methods of making a desired nucleic acid molecule, the second nucleic acid molecule may be an expression vector and may be circular or linear. In some embodiments, the second nucleic acid molecule is an expression vector and is digested with one or more restriction enzymes prior to recombination. Any restriction enzyme recognizing a site in the second nucleic acid molecule may be used in these embodiments. For example, restriction enzymes suitable for use in methods of the invention include, but are not limited to, *Bgl*II and *Pst*I. In some embodiments, a second nucleic acid molecule may be a vector having two or more recombination sites and at least one site recognized by a restriction enzyme is located between at least two recombination sites.

[0028] Optionally, in methods of making a desired nucleic acid molecule, transcriptional regulatory sequences in second nucleic acid molecules for use in the present invention may comprise a promoter. Promoters suitable for use in methods of the invention include, but are not limited to, phage promoters,

viral promoters, and other promoters known to those skilled in the art. Examples include, but are not limited to, the T7 promoter and the SP6 promoter. Other transcriptional regulatory sequences suitable for use in the present invention include, but are not limited to, enhancers, repressors, operators, suppressors, and the like.

[0029] In another aspect, the present invention provides methods of amplifying one or more nucleic acid molecules. Such methods may include recombining one or more first nucleic acid molecules and one or more second nucleic acid molecules to form one or more third nucleic acid molecules, which may be amplified, for example, by PCR. Methods of this type may entail mixing one or more third nucleic acid molecules with one or more primers and one or more polypeptides having polymerase activity to form a mixture and incubating the mixture under conditions sufficient to amplify the one or more third nucleic acid molecules. In particular, one or more third nucleic acid molecules may be double stranded nucleic acid molecules and such amplification methods may include contacting a first strand of the third nucleic acid molecule with a first primer molecule which is complementary to a portion of the first strand; and contacting a second strand of the third nucleic acid molecule with a second primer molecule which is complementary to a portion of said second strand in the presence of one or more polypeptides having polymerase activity to form a mixture; and incubating the mixture under conditions sufficient to form a third strand complementary to all or a portion of the first strand and a fourth strand complementary to all or a portion of the second strand. Such methods may also include denaturing the first and third and the second and fourth strands and repeating the contacting and incubating described above. In some embodiments, such conditions according to the invention may include incubating the mixture in the presence of one or more nucleotides, one or more buffers or buffering salts, one or more primers, one or more cofactors, and/or one or more polypeptides having a nucleotide polymerase activity.

[0030] The present invention also relates to compositions for carrying out methods of the invention and to compositions made while carrying out methods of the invention. Such compositions may comprise any one or a combination of the elements used in methods of the invention (*e.g.* one or

more recombination proteins, one or more restriction enzymes, one or more promoter sequences, etc.) and/or they may also comprise one or more substrates used during transcription and/or translation reactions (*e.g.* one or more nucleotides, one or more amino acids, one or more polymerases, one or more enzymes, one or more cofactors, one or more buffers and one or more buffering salts). In addition, such compositions may comprise one or more products of the transcription and/or translation reactions such as RNA, peptides, proteins, etc. Preferably, the compositions of the invention comprise at least one component selected from the group consisting of one or more recombination proteins, one or more restriction enzymes, one or more polymerase enzymes, and one or more cell extracts.

[0031] The present invention also encompasses kits for practicing the methods of the invention (*e.g.*, methods of making polypeptides, methods of making nucleic acid molecules, etc). A kit for producing a polypeptide according to the methods of the invention may comprise one or more containers. Such containers may contain a variety of components, for example, one or more recombination proteins, one or more restriction enzymes, and/or one or more cell extracts for *in vitro* translation. In certain such embodiments, kits may comprise one or more containers containing one or more recombination proteins, one or more containers containing one or more restriction enzymes, and one or more containers containing one or more cell extracts. Kits of the invention may comprises one or more additional components such as one or more containers containing one or more components selected from the group consisting of one or more polymerases, one or more buffers, one or more primers, one or more vectors, one or more nucleic acid molecules comprising one or more promoter sequences, and one or more nucleotides.

[0032] Kits for *in vitro* synthesis of nucleic acid molecules and/or polypeptides are also provided by the present invention. Such kits may contain any number or combination of reagents or components for carrying out the invention. Kits of the invention preferably comprise one or more elements selected from the group consisting of one or more nucleotides or derivatives thereof; one or more amino acids or derivatives thereof; one or more transfer RNA molecules, which may be charged with an amino acid; one or more polymerases; one or more cofactors; one or more buffers or buffer

salts; one or more energy sources; one or more cell extracts; one or more nucleic acid templates; one or more reagents to determine the efficiency of the kit or assay for production of the products such as nucleic acid and protein products; and directions or protocols for carrying out the methods of the invention or to use of the kits of the invention and/or its components. Such kits of the invention may comprise one or more of the above components in any number of separate containers, tubes, vials and the like or such components may be combined in various combinations in such containers.

[0033] In another aspect, the present invention relates to methods (*e.g.*, business methods) for providing a product suitable for the practice of one or more of the previously-described methods of the invention (*e.g.*, methods of making polypeptides, methods of making nucleic acid molecules, etc.) to a customer. Such methods for providing a product may comprise receiving an order for a product from a customer; and providing the customer with the product. In some aspects a product may be a kit for practicing one or more of the methods of the invention. For example, a kit may comprise one or more containers comprising one or more recombination proteins and one or more containers comprising one or more cell extracts. Optionally, kits may comprise written instructions. An order received in the practice of the present method may be received by any mode of communication, for example, by telephone, via the internet, etc.

[0034] In another aspect of the invention, the present invention encompasses methods for providing information to a customer as to availability of a product suitable for the practice of a method of the invention (*e.g.*, method of making a polypeptide, method of making a nucleic acid molecule, etc.). Such a method of providing information may comprise transmitting product description data to an output source; transmitting to the output source instructions to publish the product description in media accessible to potential interested parties; and detecting publication of the data in the media, thereby providing information as to availability of the product to parties having potential interest in the availability of the product. In some embodiments, a product may be a kit. Such a kit may comprise one or more containers comprising one or more recombination proteins and one or more containers comprising one or more cell extracts. Such a kit may further comprise written instructions. In some

embodiments, product description data may be published in a print medium and/or may be published on the internet.

[0035] Other preferred embodiments of the present invention will be apparent to one of ordinary skill in light of the following drawings and description of the invention, and of the claims.

BRIEF DESCRIPTION OF THE DRAWINGS/FIGURES

[0036] Fig. 1 is a schematic representation of a recombinational cloning reaction.

[0037] Fig. 2 is a schematic representation of the recombination region of a T7 promoter containing plasmid that may be used in the practice of the invention.

[0038] Fig. 3 is a schematic representation of the arrangement of sequences that may be included in a T7 promoter-containing plasmid of the invention.

[0039] Fig. 4 is a schematic representation of a system for providing a product to a party.

[0040] Fig. 5 provides a schematic representation of a system for advising a party as to the availability of a product.

[0041] Figs. 6A and 6B provide a comparison of the present invention to previously known methods. Fig. 6A is a timeline comparing methods of the present invention to previously available *in vivo* methods. Fig. 6B shows a schematic representation of the reactions performed in one embodiment of the present invention.

[0042] Fig. 7 is a schematic representation of a recombination reaction that may be performed in the practice of the present invention.

[0043] Fig. 8 is a schematic representation of a high throughput *in vitro* synthesis reaction in accordance with the present invention.

DETAILED DESCRIPTION OF THE INVENTION

Definitions

[0044] In the description that follows, a number of terms used in recombinant nucleic acid technology are utilized extensively. In order to provide a clear and more consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided.

[0045] Host: As used herein, the term "host" refers to any prokaryotic or eukaryotic (*e.g.*, mammalian, insect, yeast, plant, avian, animal, etc.) organism that is a recipient of a replicable expression vector, cloning vector or any nucleic acid molecule. The nucleic acid molecule may contain, but is not limited to, a sequence of interest, a transcriptional regulatory sequence (such as a promoter, enhancer, repressor, and the like) and/or an origin of replication. As used herein, the terms "host," "host cell," "recombinant host" and "recombinant host cell" may be used interchangeably. For examples of such hosts, see Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

[0046] Transcriptional Regulatory Sequence: As used herein, the phrase "transcriptional regulatory sequence" refers to any functional stretch of nucleotides contained on a nucleic acid molecule, in any configuration or geometry, that act to regulate the synthesis of an RNA molecule from a DNA template molecule (*i.e.*, transcription). Examples of transcriptional regulatory sequences include, but are not limited to, promoters, enhancers, repressors, operators (*e.g.*, the tet operator), suppressors and the like.

[0047] Promoter: As used herein, a promoter is an example of a transcriptional regulatory sequence, and more specifically is a nucleic acid generally described as the 5'-region of a gene located proximal to the start codon or nucleic acid that encodes untranslated RNA. The transcription of an adjacent nucleic acid segment is initiated at or near the promoter. A repressible promoter's rate of transcription decreases in response to a repressing agent. An inducible promoter's rate of transcription increases in response to an inducing agent. A constitutive promoter's rate of transcription

is not specifically regulated, although it can vary under the influence of general metabolic conditions.

[0048] Recognition Sequence: As used herein, the phrase "recognition sequence" refers to a particular nucleotide sequence that is recognized by and bound by a protein, chemical compound, DNA, or RNA molecule (*e.g.*, restriction endonuclease, a modification methylase, topoisomerases, a recombination protein or a recombinase). In the present invention, a recognition sequence may refer to a recombination site. For example, the recognition sequence for Cre recombinase is loxP, which is a 34 base pair sequence comprising two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence (see Fig. 1 of Sauer, B., *Current Opinion in Biotechnology* 5:521-527 (1994)). Other examples of recognition sequences are *attB*, *attP*, *attL*, and *attR* sequences, which are recognized by the recombinase enzyme λ Integrase. *attB* is an approximately 25 base pair sequence containing two 9 base pair core-type Int binding sites and a 7 base pair overlap region. *attP* is an approximately 240 base pair sequence containing core-type Int binding sites and arm-type Int binding sites as well as sites for auxiliary proteins such as integration host factor (IHF), Fis and excisionase (Xis) (see Landy, *Current Opinion in Biotechnology* 3:699-707 (1993)). Such sites may also be engineered according to the present invention to enhance production of products in the methods of the invention. For example, when such engineered sites lack the P1 or H1 domains to make the recombination reactions irreversible (*e.g.*, *attR* or *attP*), such sites may be designated *attR'* or *attP'* to show that the domains of these sites have been modified in some way.

[0049] Recombination Proteins: As used herein, the phrase "recombination proteins" includes excisive or integrative proteins, enzymes, co-factors or associated proteins that are involved in recombination reactions involving one or more recombination sites (*e.g.*, two, three, four, five, seven, ten, twelve, fifteen, twenty, thirty, fifty, etc.), which may be wild-type proteins (see Landy, *Current Opinion in Biotechnology* 3:699-707 (1993)), or mutants, derivatives (*e.g.*, fusion proteins containing the recombination protein sequences or fragments thereof), fragments, and variants thereof. Examples of

recombination proteins include Cre, Int, IHF, Xis, Flp, Fis, Hin, Gin, Φ C31, Cin, Tn3 resolvase, TndX, XerC, XerD, TnpX, Hjc, SpCCE1, and ParA.

[0050] Recombinases: As used herein, the term "recombinases" is used to refer to the protein that catalyzes strand cleavage and re-ligation in a recombination reaction.

[0051] Site-Specific Recombinase: As used herein, the phrase "site-specific recombinase" refers to a type of recombinase that typically has at least the following four activities (or combinations thereof): (1) recognition of specific nucleic acid sequences; (2) cleavage of said sequence or sequences; (3) topoisomerase activity involved in strand exchange; and (4) ligase activity to reseal the cleaved strands of nucleic acid (see Sauer, B., *Current Opinions in Biotechnology* 5:521-527 (1994)). Conservative site-specific recombination is distinguished from homologous recombination and transposition by a high degree of sequence specificity for both partners. The strand exchange mechanism involves the cleavage and rejoining of specific nucleic acid sequences in the absence of DNA synthesis (Landy, A. (1989) *Ann. Rev. Biochem.* 58:913-949).

[0052] Site-specific recombinases are proteins that are present in many organisms (*e.g.*, viruses and bacteria). These recombinases (along with associated proteins in some cases) recognize specific sequences of bases in a nucleic acid molecule and exchange the nucleic acid segments flanking those sequences. The recombinases and associated proteins are collectively referred to as "recombination proteins" (see, *e.g.*, Landy, A., *Current Opinion in Biotechnology* 3:699-707 (1993)).

[0053] Numerous recombination systems from various organisms have been described. See, *e.g.*, Hoess, *et al.*, *Nucleic Acids Research* 14(6):2287 (1986); Abremski, *et al.*, *J. Biol. Chem.* 261(1):391 (1986); Campbell, *J. Bacteriol.* 174(23):7495 (1992); Qian, *et al.*, *J. Biol. Chem.* 267(11):7794 (1992); Araki, *et al.*, *J. Mol. Biol.* 225(1):25 (1992); Maeser and Kahnmann, *Mol. Gen. Genet.* 230:170-176 (1991); Esposito, *et al.*, *Nucl. Acids Res.* 25(18):3605 (1997). Many of these belong to the integrase family of recombinases (Argos, *et al.*, *EMBO J.* 5:433-440 (1986); Voziyanov, *et al.*, *Nucl. Acids Res.* 27:930 (1999)). Perhaps the best studied of these are the Integrase/*att* system from bacteriophage λ (Landy, A. *Current Opinions in Genetics and Devel.* 3:699-

707 (1993)), the Cre/loxP system from bacteriophage P1 (Hoess and Abremski (1990) In *Nucleic Acids and Molecular Biology*, vol. 4. Eds.: Eckstein and Lilley, Berlin-Heidelberg: Springer-Verlag; pp. 90-109), and the FLP/FRT system from the *Saccharomyces cerevisiae* 2 μ circle plasmid (Broach, *et al.*, *Cell* 29:227-234 (1982)).

[0054] Recombination Site: As used herein, the phrase "recombination site" refers to a recognition sequence on a nucleic acid molecule that participates in an integration/recombination reaction by recombination proteins. Recombination sites are discrete sections or segments of nucleic acid on the participating nucleic acid molecules that are recognized and bound by a site-specific recombination protein during the initial stages of integration or recombination. In addition to the loxP, *attB*, *attP*, *attL*, and *attR* sites described above, other examples of recombination sites include the *attB*, *attP*, *attL*, and *attR* sequences described in co-pending United States patent applications 09/517,466, filed March 2, 2000, and 09/732,914, filed December 11, 2000 (published as US 2002/0007051 A1)—all of which are specifically incorporated herein by reference—and mutants, fragments, variants and derivatives thereof, which are recognized by the recombination protein λ Int and/or by the auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis) (see Landy, *Curr. Opin. Biotech.* 3:699-707 (1993)).

[0055] Recombination sites may be added to molecules by any number of known methods. For example, recombination sites can be added to nucleic acid molecules by blunt end ligation, PCR performed with fully or partially random primers, or inserting the nucleic acid molecules into a vector using a restriction site flanked by recombination sites.

[0056] Recombinational Cloning: As used herein, the phrase "recombinational cloning" refers to a method, such as that described in U.S. Patent Nos. 5,888,732; 6,143,557; 6,171,861; 6,270,969; and 6,277,608 (the contents of which are fully incorporated herein by reference), whereby segments of nucleic acid molecules or populations of such molecules are exchanged, inserted, replaced, substituted or modified, *in vitro* or *in vivo*. Preferably, such cloning method is an *in vitro* method.

[0057] Figure 1 provides a schematic representation of a recombinational cloning reaction between two closed circular nucleic acid molecules.

Recombination sites are represented as squares or circles. Recombination sites capable of recombining with each other are shown as the same shape (*i.e.*, squares recombine with squares and circles recombine with circles). In the presence of recombination proteins, a recombination reaction occurs between sites capable of recombining with each other producing a nucleic acid molecule containing both starting molecules (a cointegrate molecule). Reaction of the second pair of recombination sites resolves the cointegrate into two new nucleic acid molecules. With reference to Figure 1, a first nucleic acid molecule comprising segments A and B is reacted with a second nucleic acid molecule comprising segments C and D to produce two new nucleic acid molecules, a third nucleic acid molecule comprising segments A and D, and a fourth nucleic acid molecule comprising segments B and C.

[0058] Cloning systems that utilize recombination at defined recombination sites have been previously described in U.S. patent no. 5,888,732, U.S. patent no. 6,143,557, U.S. patent no. 6,171,861, U.S. patent no. 6,270,969, and U.S. patent no. 6,277,608, and in pending United States application no. 09/517,466 filed March 2, 2000, and in published United States application no. 2002 0007051-A1, all assigned to Invitrogen Corporation, Carlsbad, CA, the disclosures of which are specifically incorporated herein in their entirety. In brief, the GATEWAY™ Cloning System described in these patents and applications utilizes vectors that contain at least one recombination site to clone desired nucleic acid molecules *in vivo* or *in vitro*. In some embodiments, the system utilizes vectors that contain at least two different site-specific recombination sites that may be based on the bacteriophage lambda system (*e.g.*, *att1* and *att2*) that are mutated from the wild-type (*att0*) sites. Each mutated site has a unique specificity for its cognate partner *att* site (*i.e.*, its binding partner recombination site) of the same type (for example *attB1* with *attP1*, or *attL1* with *attR1*) and will not cross-react with recombination sites of the other mutant type or with the wild-type *att0* site. Different site specificities allow directional cloning or linkage of desired molecules thus providing desired orientation of the cloned molecules. Nucleic acid fragments flanked by recombination sites are cloned and subcloned using the GATEWAY™ system by replacing a selectable marker (for example, an antibiotic resistance gene) flanked by *att* sites on the recipient plasmid

molecule, sometimes termed the Destination Vector. Desired clones are then selected by transformation of a host strain and positive selection for a selectable marker on the recipient molecule. Similar strategies for negative selection (*e.g.*, use of toxic genes such as *ccdB*) can be used in other organisms such as thymidine kinase (TK) in mammals and insects.

[0059] Mutating specific residues in the core region of the *att* site can generate a large number of different *att* sites. As with the *att1* and *att2* sites utilized in GATEWAY™, each additional mutation potentially creates a novel *att* site with unique specificity that will recombine only with its cognate partner *att* site bearing the same mutation and will not cross-react with any other mutant or wild-type *att* site. Novel mutated *att* sites (*e.g.*, *attB* 1-10, *attP* 1-10, *attR* 1-10 and *attL* 1-10) are described in previous patent application serial number 09/517,466, filed March 2, 2000, which is specifically incorporated herein by reference. Other recombination sites having unique specificity (*i.e.*, a first site will recombine with its corresponding site and will not recombine or not substantially recombine with a second site having a different specificity) may be used to practice the present invention. Examples of suitable recombination sites include, but are not limited to, *loxP* sites; *loxP* site mutants, variants or derivatives such as *loxP511* (see U.S. Patent No. 5,851,808); *frt* sites; *frt* site mutants, variants or derivatives; *dif* sites; *dif* site mutants, variants or derivatives; *psi* sites; *psi* site mutants, variants or derivatives; *cer* sites; and *cer* site mutants, variants or derivatives; as well as others described elsewhere herein.

[0060] Repression Cassette: As used herein, the phrase "repression cassette" refers to a nucleic acid segment that contains a repressor or a selectable marker present in the subcloning vector.

[0061] Vector: As used herein, the term "vector" refers to a nucleic acid molecule (preferably DNA) that provides a useful biological or biochemical property to an insert. Examples include plasmids, phages, autonomously replicating sequences (ARS), centromeres, and other sequences that are able to replicate or be replicated *in vitro* or in a host cell, or to convey a desired nucleic acid segment to a desired location within a host cell. A vector can have one or more recognition sites (*e.g.*, two, three, four, five, seven, ten, etc. recombination sites, restriction sites, and/or topoisomerases sites) at which the

sequences can be manipulated in a determinable fashion without loss of an essential biological function of the vector, and into which a nucleic acid fragment can be spliced in order to bring about its replication and cloning. Vectors can further provide primer sites (*e.g.*, for PCR), transcriptional and/or translational initiation and/or regulation sites, recombinational signals, replicons, selectable markers, etc. Clearly, methods of inserting a desired nucleic acid fragment that do not require the use of recombination, transpositions or restriction enzymes (such as, but not limited to, uracil N-glycosylase (UDG) cloning of PCR fragments (U.S. Patent No. 5,334,575 and 5,888,795, both of which are entirely incorporated herein by reference), T:A cloning, and the like) can also be applied to clone a fragment into a cloning vector to be used according to the present invention. The cloning vector can further contain one or more selectable markers (*e.g.*, two, three, four, five, seven, ten, etc.) suitable for use in the identification of cells transformed with the cloning vector.

[0062] Subcloning Vector: As used herein, the phrase "subcloning vector" refers to a cloning vector comprising a circular or linear nucleic acid molecule that includes, preferably, an appropriate replicon. In the present invention, the subcloning vector can also contain functional and/or regulatory elements that are desired to be incorporated into the final product to act upon or with the cloned nucleic acid insert. The subcloning vector can also contain a selectable marker (preferably DNA).

[0063] Primer: As used herein, the term "primer" refers to a single stranded or double stranded oligonucleotide that is extended by covalent bonding of nucleotide monomers during amplification or polymerization of a nucleic acid molecule (*e.g.*, a DNA molecule). In one aspect, the primer may be a sequencing primer (for example, a universal sequencing primer). In another aspect, the primer may comprise a recombination site or portion thereof.

[0064] Template: As used herein, the term "template" refers to a double stranded or single stranded nucleic acid molecule that serves to direct synthesis of another nucleic acid molecule. Examples of synthesis of another nucleic acid molecule from a template include, but are not limited to, amplification, transcription, and sequencing. In the case of a double-stranded DNA molecule, denaturation of its strands to form a first and a second strand

is preferably performed before these molecules may be amplified, synthesized or sequenced, or the double stranded molecule may be used directly as a template. For single stranded templates, a primer complementary to at least a portion of the template hybridizes under appropriate conditions and one or more polypeptides having polymerase activity (*e.g.*, two, three, four, five, or seven DNA polymerases and/or reverse transcriptases) may then synthesize a molecule complementary to all or a portion of the template. Alternatively, for double stranded templates, one or more transcriptional regulatory sequences (*e.g.*, two, three, four, five, seven or more promoters) may be used in combination with one or more polymerases to make nucleic acid molecules complementary to all or a portion of the template. The newly synthesized molecule, according to the invention, may be of equal or shorter length compared to the original template. Mismatch incorporation or strand slippage during the synthesis or extension of the newly synthesized molecule may result in one or a number of mismatched base pairs. Thus, the synthesized molecule need not be exactly complementary to the template. Additionally, a population of nucleic acid templates may be used during synthesis or amplification to produce a population of nucleic acid molecules typically representative of the original template population.

[0065] Incorporating: As used herein, the term "incorporating" means becoming a part of a nucleic acid (*e.g.*, DNA) molecule or primer.

[0066] Library: As used herein, the term "library" refers to a collection of nucleic acid molecules (circular or linear). In one embodiment, a library may comprise a plurality of nucleic acid molecules (*e.g.*, two, three, four, five, seven, ten, twelve, fifteen, twenty, thirty, fifty, one hundred, two hundred, five hundred one thousand, five thousand, or more), that may or may not be from a common source organism, organ, tissue, or cell. In another embodiment, a library is representative of all or a portion or a significant portion of the nucleic acid content of an organism (a "genomic" library), or a set of nucleic acid molecules representative of all or a portion or a significant portion of the expressed nucleic acid molecules (a cDNA library or segments derived therefrom) in a cell, tissue, organ or organism. A library may also comprise nucleic acid molecules having random sequences made by *de novo* synthesis, mutagenesis of one or more nucleic acid molecules, and the like. Such

libraries may or may not be contained in one or more vectors (*e.g.*, two, three, four, five, seven, ten, twelve, fifteen, twenty, thirty, fifty, etc.).

[0067] Amplification: As used herein, the term "amplification" refers to any *in vitro* method for increasing the number of copies of a nucleic acid molecule with the use of one or more polypeptides having polymerase activity (*e.g.*, one, two, three, four or more nucleic acid polymerases or reverse transcriptases). Nucleic acid amplification results in the incorporation of nucleotides into a DNA and/or RNA molecule or primer thereby forming a new nucleic acid molecule complementary to a template. The formed nucleic acid molecule and its template can be used as templates to synthesize additional nucleic acid molecules. As used herein, one amplification reaction may consist of many rounds of nucleic acid replication. DNA amplification reactions include, for example, polymerase chain reaction (PCR). One PCR reaction may consist of 5 to 100 cycles of denaturation and synthesis of a DNA molecule.

[0068] Nucleotide: As used herein, the term "nucleotide" refers to a base-sugar-phosphate combination. Nucleotides are monomeric units of a nucleic acid molecule (DNA and RNA). The term nucleotide includes ribonucleoside triphosphates ATP, UTP, CTG, GTP and deoxyribonucleoside triphosphates such as dATP, dCTP, dITP, dUTP, dGTP, dTTP, or derivatives thereof. Such derivatives include, for example, [α -S]dATP, 7-deaza-dGTP and 7-deaza-dATP. The term nucleotide as used herein also refers to dideoxyribonucleoside triphosphates (ddNTPs) and their derivatives. Illustrated examples of dideoxyribonucleoside triphosphates include, but are not limited to, ddATP, ddCTP, ddGTP, ddITP, and ddTTP. According to the present invention, a "nucleotide" may be unlabeled or detectably labeled by well known techniques. Detectable labels include, for example, radioactive isotopes, fluorescent labels, chemiluminescent labels, bioluminescent labels and enzyme labels.

[0069] Nucleic Acid Molecule: As used herein, the phrase "nucleic acid molecule" refers to a sequence of contiguous nucleotides (riboNTPs, dNTPs, ddNTPs, or combinations thereof) of any length. A nucleic acid molecule may encode a full-length polypeptide or a fragment of any length thereof, or may be non-coding. As used herein, the terms "nucleic acid molecule" and

"polynucleotide" may be used interchangeably and include both RNA and DNA.

[0070] Oligonucleotide: As used herein, the term "oligonucleotide" refers to a synthetic or natural molecule comprising a covalently linked sequence of nucleotides that are joined by a phosphodiester bond between the 3' position of the pentose of one nucleotide and the 5' position of the pentose of the adjacent nucleotide.

[0071] Protein: As used herein, the term "protein" refers to a sequence of contiguous amino acids of any length. The terms "peptide," "oligopeptide," or "polypeptide" may be used interchangeably herein with the term "protein."

[0072] Hybridization: As used herein, the terms "hybridization" and "hybridizing" refer to base pairing of two complementary single-stranded nucleic acid molecules (RNA and/or DNA) to give a double stranded molecule. As used herein, two nucleic acid molecules may hybridize, although the base pairing is not completely complementary. Accordingly, mismatched bases do not prevent hybridization of two nucleic acid molecules provided that appropriate conditions, well known in the art, are used. In some aspects, hybridization is said to be under "stringent conditions." By "stringent conditions," as the phrase is used herein, is meant overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (750 mM NaCl, 75mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1 x SSC at about 65°C.

[0073] Extract: As used herein, the term "extract" refers to a cell lysate or exudate. The cell can be any cell that can be grown for preparing an extract. Both prokaryotic cells and eukaryotic cells can be used to prepare extracts for protein and/or nucleic acid synthesis according to the invention.

[0074] An extract may be prepared by lysing cells using techniques known in the art. Preferably, an extract is processed to remove cellular debris. Centrifugation is a common method for removing such solid material. Filtration, chromatography, or any other separation or purification procedures may be used to produce a desired extract. The extract preferably includes all necessary components for practicing the methods of the invention. An extract can be concentrated using one or more of the many tools of the art. One or

more components (*e.g.*, enzymes, energy sources, and the like) present in the extract can originate in the extracted cell or can be added during the production of the extract. Extracts may also include a mixture of components crafted to imitate a cell lysate or exudate with respect to the components necessary or desired for protein and/or nucleic acid synthesis. An extract thus can be a mixture of components to imitate or improve upon a cell lysate or exudate in protein synthesis reactions and/or to provide components used for synthesis from a nucleic acid template. As will be recognized by one of ordinary skill in the art, such mixtures, can be produced by obtaining a partial extract or fraction thereof and/or by mixing any number of individual components.

[0075] Polymerase: As used herein, the term "polymerase" refers to an enzyme that catalyses synthesis of nucleic acids using a preexisting nucleic acid template. A polymerase may synthesize an RNA and/or a DNA molecule. Enzymes having nucleic acid polymerase activity suitable for use in the present methods include polymerases that are active in the chosen system with the chosen template to synthesize protein. The extract of a cell will usually contain a suitable polymerase, such as RNA polymerase II, SP6 RNA polymerase, T3 RNA polymerase, T7 RNA polymerase, RNA polymerase III and/or phage-derived RNA polymerases. These and other polymerases are available commercially and/or are known in the art and can be readily accessed by the skilled artisan by searching one or more of the public or private databases. Suitable polymerase can also be supplemented in the system. When RNA is to be synthesized from a DNA template, a polymerase active on the DNA molecule of interest should be used. RNA polymerases and transcription factors useful in the invention are well known in the art and will be readily recognized by those skilled in the art.

[0076] *In vivo*: As used herein, the phrase "*in vivo*" means within a living cell or organism.

[0077] *In vitro*: As used herein, the phrase "*in vitro*" means outside a living cell or organism (*e.g.*, in a cell extract, in a test tube or other container, etc.).

[0078] Other terms used in the fields of recombinant nucleic acid technology and molecular and cell biology as used herein will be generally understood by one of ordinary skill in the applicable arts.

Overview

[0079] The present invention relates to materials and methods for the *in vitro* construction of template nucleic acid molecules and their use in *in vitro* synthesis of proteins. The invention includes synthesis systems, methods and kits embodying one or more of the features of the present invention. The synthesis system of the present invention includes the necessary components to synthesize proteins from template nucleic acid molecules. The *in vitro* synthesis system of the invention provides efficient synthesis outside the confines of a cell. The kits of the present invention allow or facilitate the practice of the methods of the invention (*e.g.*, methods of making template nucleic acid molecules, methods of *in vitro* transcription translation, methods of making polypeptides, etc).

[0080] In general, methods of the invention may comprise making a nucleic acid molecule (*e.g.*, a DNA molecule) in a process that may involve a site-specific recombination reaction. Nucleic acid molecules made by methods of the invention may be used in further reactions. For example, nucleic acid molecules of the invention may be used in *in vitro* transcription reactions, *in vitro* translation reactions, coupled *in vitro* transcription/translation reactions, amplification reactions (*e.g.*, PCR), and/or combinations of these reactions. For example, a nucleic acid molecule produced by the methods of the invention may be amplified and used in a coupled transcription/translation reaction.

Recombination Sites

[0081] Recombination sites for use in the invention may be any nucleic acid that can serve as a substrate in a recombination reaction. Such recombination sites may be wild-type or naturally occurring recombination sites, or modified, variant, derivative, or mutant recombination sites. Examples of recombination sites for use in the invention include, but are not limited to, phage-lambda recombination sites (such as *attP*, *attB*, *attL*, and *attR* and mutants or

derivatives thereof) and recombination sites from other bacteriophages such as phi80, P22, P2, 186, P4 and P1 (including lox sites such as loxP and loxP511).

[0082] Preferred recombination proteins and mutant, modified, variant, or derivative recombination sites for use in the invention include those described in U.S. Patent Nos. 5,888,732, 6,143,557, 6,171,861, 6,270,969, and 6,277,608 and in U.S. application no. 09/438,358, filed November 12, 1999, which are specifically incorporated herein by reference. Mutated *att* sites (*e.g.*, *attB* 1-10, *attP* 1-10, *attR* 1-10 and *attL* 1-10) are described in United States application numbers 09/517,466, filed March 2, 2000, and 09/732,914, filed December 11, 2000 (published as US 2002/0007051-A1) the disclosures of which are specifically incorporated herein by reference in their entirety. Other suitable recombination sites and proteins are those associated with the GATEWAY™ Cloning Technology systems available from Invitrogen Corporation, Carlsbad, CA, and are described in the associated product literature, the entire disclosures of all of which are specifically incorporated herein by reference in their entireties.

[0083] Recombination sites that may be used in the present invention include *att* sites. The 15 bp core region of the wildtype *att* site (GCTTTTTTAT ACTAA (SEQ ID NO:)), which is identical in all wildtype *att* sites, may be mutated in one or more positions. Engineered *att* sites that specifically recombine with other engineered *att* sites can be constructed by altering nucleotides in and near the 7 base pair overlap region, bases 6-12, of the core region. Thus, recombination sites suitable for use in the methods, molecules, compositions, and vectors of the invention include, but are not limited to, those with insertions, deletions or substitutions of one, two, three, four, or more nucleotide bases within the 15 base pair core region (see U.S. Patent Nos. 5,888,732 and 6,277,608, which describe the core region in further detail, and the disclosures of which are incorporated herein by reference in their entireties). Recombination sites suitable for use in the methods, compositions, and vectors of the invention also include those with insertions, deletions or substitutions of one, two, three, four, or more nucleotide bases within the 15 base pair core region that are at least 50% identical, at least 55% identical, at least 60% identical, at least 65% identical, at least 70% identical, at least 75%

identical, at least 80% identical, at least 85% identical, at least 90% identical, or at least 95% identical to this 15 base pair core region.

[0084] As a practical matter, whether any particular nucleic acid molecule is at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, a given recombination site nucleotide sequence or portion thereof can be determined conventionally using known computer programs such as DNAsis software (Hitachi Software, San Bruno, California) for initial sequence alignment followed by ESEE version 3.0 DNA/protein sequence software (cabot@trog.mbb.sfu.ca) for multiple sequence alignments. Alternatively, such determinations may be accomplished using the BESTFIT program (Wisconsin Sequence Analysis Package, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711), which employs a local homology algorithm (Smith and Waterman, *Advances in Applied Mathematics* 2: 482-489 (1981)) to find the best segment of homology between two sequences. When using DNAsis, ESEE, BESTFIT or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number of nucleotides in the reference sequence are allowed. Computer programs such as those discussed above may also be used to determine percent identity and homology between two proteins at the amino acid level.

[0085] Analogously, the core regions in *attB1*, *attP1*, *attL1* and *attR1* are identical to one another, as are the core regions in *attB2*, *attP2*, *attL2* and *attR2*. Nucleic acid molecules suitable for use with the invention also include those comprising insertions, deletions or substitutions of one, two, three, four, or more nucleotides within the seven base pair overlap region (TTTATAC, bases 6-12 in the core region). The overlap region is defined by the cut sites for the integrase protein and is the region where strand exchange takes place. Examples of such mutants, fragments, variants and derivatives include, but are not limited to, nucleic acid molecules in which (1) the thymine at position 1 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; (2) the thymine at position 2 of the seven bp overlap

region has been deleted or substituted with a guanine, cytosine, or adenine; (3) the thymine at position 3 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; (4) the adenine at position 4 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or thymine; (5) the thymine at position 5 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; (6) the adenine at position 6 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or thymine; and (7) the cytosine at position 7 of the seven bp overlap region has been deleted or substituted with a guanine, thymine, or adenine; or any combination of one or more (*e.g.*, two, three, four, five, etc.) such deletions and/or substitutions within this seven bp overlap region. The nucleotide sequences of representative seven base pair core regions are set out below.

[0086] Altered *att* sites have been constructed that demonstrate that (1) substitutions made within the first three positions of the seven base pair overlap (TTTATAC) strongly affect the specificity of recombination, (2) substitutions made in the last four positions (TTTATAC) only partially alter recombination specificity, and (3) nucleotide substitutions outside of the seven bp overlap, but elsewhere within the 15 base pair core region, do not affect specificity of recombination but do influence the efficiency of recombination. Thus, nucleic acid molecules and methods of the invention include those comprising or employing one, two, three, four, five, six, eight, ten, or more recombination sites which affect recombination specificity, particularly one or more (*e.g.*, one, two, three, four, five, six, eight, ten, twenty, thirty, forty, fifty, etc.) different recombination sites that may correspond substantially to the seven base pair overlap within the 15 base pair core region, having one or more mutations that affect recombination specificity. Particularly preferred such molecules may comprise a consensus sequence such as NNNATAC wherein "N" refers to any nucleotide (*i.e.*, may be A, G, T/U or C, or an analogue or derivative thereof). Preferably, if one of the first three nucleotides in the consensus sequence is a T/U, then at least one of the other two of the first three nucleotides is not a T/U.

[0087] The core sequence of each *att* site (*attB*, *attP*, *attL* and *attR*) can be divided into functional units consisting of integrase binding sites, integrase

cleavage sites and sequences that determine specificity. Specificity determinants are defined by the first three positions following the integrase top strand cleavage site. These three positions are shown with underlining in the following reference sequence: CAACTTTTTTATAC AAAGTTG (SEQ ID NO:). Modification of these three positions (64 possible combinations) can be used to generate *att* sites that recombine with high specificity with other *att* sites having the same sequence for the first three nucleotides of the seven base pair overlap region. The possible combinations of first three nucleotides of the overlap region are shown in Table 1.

Table 1. Modifications of the First Three Nucleotides of the <i>att</i> Site Seven Base Pair Overlap Region that Alter Recombination Specificity.			
AAA	CAA	GAA	TAA
AAC	CAC	GAC	TAC
AAG	CAG	GAG	TAG
AAT	CAT	GAT	TAT
ACA	CCA	GCA	TCA
ACC	CCC	GCC	TCC
ACG	CCG	GCG	TCG
ACT	CCT	GCT	TCT
AGA	CGA	GGA	TGA
AGC	CGC	GGC	TGC
AGG	CGG	GGG	TGG
AGT	CGT	GGT	TGT
ATA	CTA	GTA	TTA
ATC	CTC	GTC	TTC
ATG	CTG	GTG	TTG
ATT	CTT	GTT	TTT

[0088] Representative examples of seven base pair *att* site overlap regions suitable for use in methods, compositions and vectors of the invention are shown in Table 2. The invention further includes nucleic acid molecules comprising one or more (*e.g.*, one, two, three, four, five, six, eight, ten, twenty, thirty, forty, fifty, etc.) nucleotides sequences set out in Table 2. Thus, for example, in one aspect, the invention provides nucleic acid molecules comprising the nucleotide sequence GAAATAC, GATATAC, ACAATAC, or TGCATAC.

Table 2. Representative Examples of Seven Base Pair <i>att</i> Site Overlap Regions Suitable for use in the recombination sites of the Invention.			
AAAATAC	CAAATAC	GAAATAC	TAAATAC
AACATAC	CACATAC	GACATAC	TACATAC
AAGATAC	CAGATAC	GAGATAC	TAGATAC
AATATAC	CATATAC	GATATAC	TATATAC
ACAATAC	CCAATAC	GCAATAC	TCAATAC
ACCATAC	CCCATAC	GCCATAC	TCCATAC
ACGATAC	CCGATAC	GCGATAC	TCGATAC
ACTATAC	CCTATAC	GCTATAC	TCTATAC
AGAATAC	CGAATAC	GGAATAC	TGAATAC
AGCATAC	CGCATAC	GGCATAC	TGCATAC
AGGATAC	CGGATAC	GGGATAC	TGGATAC
AGTATAC	CGTATAC	GGTATAC	TGTATAC
ATAATAC	CTAATAC	GTAATAC	TTAATAC
ATCATAC	CTCATAC	GTCATAC	TTCATAC
ATGATAC	CTGATAC	GTGATAC	TTGATAC
ATTATAC	CTTATAC	GTTATAC	TTTATAC

[0089] As noted above, alterations of nucleotides located 3' to the three base pair region discussed above can also affect recombination specificity. For example, alterations within the last four positions of the seven base pair overlap can also affect recombination specificity.

[0090] For example, mutated *att* sites that may be used in the practice of the present invention include *attB1* (AGCCTGCTTT TTTGTACAAA CTTGT (SEQ ID NO:)), *attP1* (TACAGGTCAC TAATACCATC TAAGTAGTTG ATTCATAGTG ACTGGATATG TTGTGTTTTC CAGTATTATG TAGTCTGTTT TTTATGCAAA ATCTAATTTC ATATATTGAT ATTTATATCA TTTTACGTTT CTCGTTCAGC TTTTTTGTAC AAAGTTGGCA TTATAAAAA GCATTGCTCA TCAATTTGTT GCAACGAACA GGTCACATC AGTCAAAATA AAATCATTAT TTG (SEQ ID NO:)), *attL1* (CAAATAATGA TTTTATTTTG ACTGATAGTG ACCTGTTCGT TGCAACAAAT TGATAAGCAA TGCTTTTTTTC TAATGCCAAC TTTGTACAAA AAAGCAGGCT (SEQ ID NO:)), and *attR1* (ACAAGTTTGT ACAAAAAAGC TGAACGAGAA ACGTAAAATG ATATAAATAT CAATATATTA AATTAGATTT TGCATAAAAA ACAGACTACA TAATACTGTA AAACACAACA TATCCAGTCA CTATG (SEQ ID NO:)). Table 3 provides the sequences of the regions surrounding the core region for the wild type *att* sites (*attB0*, *P0*, *R0*, and *L0*) as well as a variety of other suitable recombination sites. Those

skilled in the art will appreciate that the remainder of the site may be the same as the corresponding site (B, P, L, or R) listed above.

Table 3. Nucleotide sequences of <i>att</i> sites.		
<i>att</i> B0	AGCCTGCTTT TTTATACTAA CTTGAGC	(SEQ ID NO:)
<i>att</i> P0	G TTCAGCTTT TTTATACTAA GTTGGCA	(SEQ ID NO:)
<i>att</i> L0	AGCCTGCTTT TTTATACTAA GTTGGCA	(SEQ ID NO:)
<i>att</i> R0	G TTCAGCTTT TTTATACTAA CTTGAGC	(SEQ ID NO:)
<i>att</i> B1	AGCCTGCTTT TTTGTACAAA CTTGT	(SEQ ID NO:)
<i>att</i> P1	G TTCAGCTTT TTTGTACAAA GTTGGCA	(SEQ ID NO:)
<i>att</i> L1	AGCCTGCTTT TTTGTACAAA GTTGGCA	(SEQ ID NO:)
<i>att</i> R1	G TTCAGCTTT TTTGTACAAA CTTGT	(SEQ ID NO:)
<i>att</i> B2	ACCCAGCTTT CTTGTACAAA GTGGT	(SEQ ID NO:)
<i>att</i> P2	G TTCAGCTTT CTTGTACAAA GTTGGCA	(SEQ ID NO:)
<i>att</i> L2	ACCCAGCTTT CTTGTACAAA GTTGGCA	(SEQ ID NO:)
<i>att</i> R2	G TTCAGCTTT CTTGTACAAA GTGGT	(SEQ ID NO:)
<i>att</i> B5	CAACTTTATT ATACAAAGTT GT	(SEQ ID NO:)
<i>att</i> P5	G TTCAACTTT ATTATACAAA GTTGGCA	(SEQ ID NO:)
<i>att</i> L5	CAACTTTATT ATACAAAGTT GGCA	(SEQ ID NO:)
<i>att</i> R5	G TTCAACTTT ATTATACAAA GTTGT	(SEQ ID NO:)
<i>att</i> B11	CAACTTTTCT ATACAAAGTT GT	(SEQ ID NO:)
<i>att</i> P11	G TTCAACTTT TCTATACAAA GTTGGCA	(SEQ ID NO:)
<i>att</i> L11	CAACTTTTCT ATACAAAGTT GGCA	(SEQ ID NO:)
<i>att</i> R11	G TTCAACTTT TCTATACAAA GTTGT	(SEQ ID NO:)
<i>att</i> B17	CAACTTTTGT ATACAAAGTT GT	(SEQ ID NO:)
<i>att</i> P17	G TTCAACTTT TGTATACAAA GTTGGCA	(SEQ ID NO:)
<i>att</i> L17	CAACTTTTGT ATACAAAGTT GGCA	(SEQ ID NO:)
<i>att</i> R17	G TTCAACTTT TGTATACAAA GTTGT	(SEQ ID NO:)

Table 3. Nucleotide sequences of <i>att</i> sites.		
<i>att</i> B19	CAACTTTTTC GTACAAAGTT GT	(SEQ ID NO:)
<i>att</i> P19	GTTCAACTTT TTCGTACAAA GTTGGCA	(SEQ ID NO:)
<i>att</i> L19	CAACTTTTTC GTACAAAGTT GGCA	(SEQ ID NO:)
<i>att</i> R19	GTTCAACTTT TTCGTACAAA GTTGT	(SEQ ID NO:)
<i>att</i> B20	CAACTTTTTC GTACAAAGTT GT	(SEQ ID NO:)
<i>att</i> P20	GTTCAACTTT TTGGTACAAA GTTGGCA	(SEQ ID NO:)
<i>att</i> L20	CAACTTTTTC GTACAAAGTT GGCA	(SEQ ID NO:)
<i>att</i> R20	GTTCAACTTT TTGGTACAAA GTTGT	(SEQ ID NO:)
<i>att</i> B21	CAACTTTTTC ATACAAAGTT GT	(SEQ ID NO:)
<i>att</i> P21	GTTCAACTTT TTAATACAAA GTTGGCA	(SEQ ID NO:)
<i>att</i> L21	CAACTTTTTC ATACAAAGTT GGCA	(SEQ ID NO:)
<i>att</i> R21	GTTCAACTTT TTAATACAAA GTTGT	(SEQ ID NO:)

[0091] Other recombination sites having unique specificity (*i.e.*, a first site will recombine with its corresponding site and will not substantially recombine with a second site having a different specificity) are known to those skilled in the art and may be used to practice the present invention. Corresponding recombination proteins for these systems may be used in accordance with the invention with the indicated recombination sites. Other systems providing recombination sites and recombination proteins for use in the invention include the FLP/FRT system from *Saccharomyces cerevisiae*, the resolvase family (*e.g.*, $\gamma\delta$, TndX, TnpX, Tn3 resolvase, Hin, Hjc, Gin, SpCCE1, ParA, and Cin), and IS231 and other *Bacillus thuringiensis* transposable elements. Other suitable recombination systems for use in the present invention include the XerC and XerD recombinases and the psi, dif and cer recombination sites in *E. coli*. Other suitable recombination sites may be found in United States patent no. 5,851,808 issued to Elledge and Liu which is specifically incorporated herein by reference.

Construction of Nucleic Acid Molecules

[0092] As discussed below in more detail, in one aspect, the invention provides methods for constructing nucleic acid molecules having particular functions or activities (*e.g.*, containing a transcriptional regulatory sequence operably linked to a sequence encoding a protein of interest). In one general embodiment of the invention, nucleic acid molecules of the invention are prepared as follows.

[0093] A first nucleic acid molecule comprising a sequence of interest may be prepared by methods known to those skilled in the art, such as by conventional cloning, recombinational cloning, topoisomerase-mediated cloning, PCR or other amplification technique, solid phase synthesis, etc. Preferably, a first nucleic acid molecule may comprise one or more recombination sites in addition to a sequence of interest. One or more recombination sites may be located at any position in the first nucleic acid molecule and may be located upstream (*i.e.*, in the 5'-direction of the sense strand) of the sequence of interest. A sequence of interest may be encode one or more polypeptides or may code for an RNA molecule (*e.g.*, a an antisense RNA, an siRNA, etc.). In certain such embodiments, at least one recombination site may be located adjacent to and/or upstream of the first base of the sequence of interest.

[0094] A second nucleic acid molecule comprising the sequence of one or more transcriptional regulatory sequences may be prepared using methods known to those skilled in the art. Preferably, a second nucleic acid molecule may also comprise one or more recombination sites and/or one or more transcriptional regulatory sequences. Recombination sites may be located at any position in the second nucleic acid molecule. For example, a second nucleic acid molecule may comprise a transcriptional regulatory sequence adjacent to a recombination site. The recombination site may be located downstream (*i.e.*, in the 3'-direction) of the transcriptional regulatory sequence.

[0095] Methods of construction of nucleic acid molecules of the invention may include a recombination reaction. For example, a first nucleic acid molecule having at least one recombination site may be recombined with a second nucleic acid molecule having at least one recombination site. In some

embodiments of this type, the first nucleic acid molecule may have a sequence encoding a nucleic acid molecule and/or protein of interest adjacent to and downstream of a recombination site. The first nucleic acid molecule may be contacted with a second nucleic acid molecule optionally having a recombination site located adjacent to and downstream of a transcriptional regulatory sequence. The first and the second nucleic acid molecules may be contacted under conditions sufficient to cause recombination between the recombination sites on the first and second nucleic acid molecules resulting in the production of a third nucleic acid molecule. In the third nucleic acid molecule, the coding sequence from the first nucleic acid molecule may be operably linked to the transcriptional regulatory sequence (when present) from the second nucleic acid molecule.

[0096] As discussed above, the first and the second nucleic acid molecules may be prepared by any technique known to those skilled in the art. In some embodiments, the first nucleic acid molecule may be a vector and the sequence encoding the protein of interest may be flanked by recombination sites. Examples of suitable nucleic acid molecules that may serve as first nucleic acid molecules are sequences of interest in Entry Vectors available from Invitrogen Corporation, Carlsbad, CA, such as the clones in the ULTIMATE™ ORF Collection, also available from Invitrogen Corporation, Carlsbad, CA.

[0097] A second nucleic acid molecule of the present invention may be a vector, for example, a Destination Vector compatible with the GATEWAY™ recombinational cloning technology of Invitrogen Corporation, Carlsbad, CA. In some embodiments, a second nucleic acid molecule may be a linear molecule. For example, a closed circular molecule (*e.g.*, a plasmid) may be digested with one or more restriction enzymes to produce one or more linear fragments, one of which may be used as a second nucleic acid molecule. Alternatively, a linear nucleic acid molecule to be used as a second nucleic acid molecule according to the present invention may be prepared by other techniques known in the art (*e.g.*, PCR, chemical synthesis, etc.). Examples of suitable second nucleic acid molecules include, but are not limited to, the T7-based pEXP1-DEST and pEXP2-DEST vectors supplied with the Expressway™ *In Vitro* Protein Synthesis System, Invitrogen Corporation,

Carlsbad, CA, catalog nos. K9600-01, K9600-02, K9605-01, and K9605-02. A schematic representation of the recombination region from these vectors after insertion of a sequence encoding a protein of interest (shown as GENE in the figure) is provided in Figure 2.

[0098] In some embodiments, a third nucleic acid molecule produced by the methods of the invention may comprise a promoter recognized by T7 polymerase and additional sequences that facilitate *in vitro* transcription of a nucleic acid molecule corresponding to the third nucleic acid molecule and/or translation of the corresponding nucleic acid molecule. The additional sequences may have been present on either the first or the second nucleic acid molecule or may have been partially contained on both. For example, a third nucleic acid molecule may comprise a T7 promoter, an initiation codon, and a prokaryotic Shine-Dalgarno ribosome binding site (RBS) upstream of the sequence encoding a protein of interest. The sequence encoding the protein of interest may preferably comprise a stop codon. Other suitable additional sequences include, but are not limited to, purification tag sequences (*e.g.*, six histidines), epitopes (*e.g.*, V5 epitope) and the like. Figure 3 provides a schematic representation of a suitable arrangement of T7 promoter and additional sequences for use in the present invention.

[0099] Third nucleic acid molecules may also comprise a sequence upstream of the T7 promoter containing a minimum of 6-10 nucleotides (nt) for efficient promoter binding. This sequence need not be specific.

[0100] Third nucleic acid molecules may also comprise a sequence following the T7 promoter containing a minimum of 15-20 nt that forms a potential stem-and-loop structure as described by Studier *et al.*, 1990, *Meth. Enzymol.* 185, 60-89.

[0101] Third nucleic acid molecules may also comprise a sequence of 9-11 nt between the RBS and the ATG initiation codon for optimal translation efficiency of the protein of interest. This sequence need not be specific.

[0102] Third nucleic acid molecules may also comprise a T7 terminator located 4-100 nt downstream of the sequence encoding the protein of interest for efficient transcription termination and message stability.

[0103] Optionally, a third nucleic acid molecule produced as described above may be amplified using techniques well known in the art, for example,

polymerase chain reaction (PCR). PCR, a well known DNA amplification technique, is a process by which DNA polymerase and deoxyribonucleoside triphosphates are used to amplify a target DNA template. In PCR reactions, two primers, one complementary to the 3' termini (or near the 3'-terminus) of the first strand of the DNA molecule to be amplified, and a second primer complementary to the 3' termini (or near the 3'-terminus) of the second strand of the DNA molecule to be amplified, are hybridized to their respective DNA strands. After hybridization, DNA polymerase, in the presence of deoxyribonucleoside triphosphates, allows synthesis of a third DNA strand complementary to the first strand and a fourth DNA strand complementary to the second strand of the DNA molecule to be amplified. This synthesis results in two double stranded DNA molecules. Such double stranded DNA molecules may then be used to provide DNA templates for synthesis of additional DNA molecules by providing a DNA polymerase, primers, and deoxyribonucleoside triphosphates. As is well known, the additional synthesis is carried out by "cycling" the original reaction (with excess primers and deoxyribonucleoside triphosphates) allowing multiple denaturing and synthesis steps. Typically, denaturing of double stranded DNA molecules to form single stranded DNA templates is accomplished by high temperatures. DNA polymerases (including thermostable DNA polymerases) that may be used in the present invention include, but are not limited to, *Taq* DNA polymerase, *Tne* DNA polymerase, *Tma* DNA polymerase, *Pfu* DNA polymerase, *Tfl* DNA polymerase, *Tth* DNA polymerase, *Tbr* DNA polymerase, *Pwo* DNA polymerase, *Bst* DNA polymerase, *Bca* DNA polymerase, VENT DNA polymerase, T7 DNA polymerase, T5 DNA polymerase, DNA polymerase III, Klenow fragment DNA polymerase, Stoffel fragment DNA polymerase, and mutants, fragments or derivatives thereof.

[0104] Other suitable amplification techniques are known to those skilled in the art and include, but are not limited to, Strand Displacement Amplification (SDA; U.S. Patent No. 5,455,166; EP 0 684 315), and Nucleic Acid Sequence-Based Amplification (NASBA; U.S. Patent No. 5,409,818; EP 0 329 822).

[0105] In another aspect of the invention, the present invention provides methods of amplifying one or more nucleic acid molecules. Such methods may comprise:

(a) recombining one or more first nucleic acid molecules and one or more second nucleic acid molecules to form one or more third nucleic acid molecules;

(b) mixing one or more third nucleic acid molecules with one or more primers and one or more polypeptides having polymerase activity to form a mixture; and

(c) incubating the mixture under conditions sufficient to amplify all or a portion of the one or more third nucleic acid molecules. Upon recombination of the first and second nucleic acid molecules, tethered recombination sites on the second nucleic acid molecule may be reconnected. This permits amplification of all or a portion of the third nucleic acid molecule using primers that hybridize to sequences present on the second nucleic acid molecule. Typically, the sequences on the second nucleic acid molecule to which the primers hybridize are arranged such that no amplification of the third nucleic acid molecule is possible in the absence of recombination between the first and second nucleic acid molecule. For example, a linear second nucleic acid molecule may be used and primer sites selected such that extension of each of the primers is away from the binding site of the other primer. For example, sites may be selected near the ends of the second nucleic acid molecule and the direction of extension may be toward the near end. Thus, in the absence of recombination, each primer is extended off the end of the second nucleic acid molecule and no amplification product is produced.

[0106] In particular, one or more third nucleic acid molecules may be double stranded nucleic acid molecules and such amplification methods may comprise:

(a) contacting a first strand of the third nucleic acid molecule with a first primer molecule which is complementary to a portion of the first strand and a second strand of the third nucleic acid molecule with a second primer molecule which is complementary to a portion of said second strand in the presence of one or more polypeptides having polymerase activity to form a mixture;

(b) incubating the mixture under conditions sufficient to form a third strand complementary to all or a portion of the first strand and a fourth strand complementary to all or a portion of the second strand;

(c) denaturing the first and third and the second and fourth strands; and

(d) repeating (a) through (c) one or more times. In some embodiments, such conditions according to the invention may include one or more nucleotides, one or more buffers or buffering salts, one or more primers, one or more cofactors, and/or one or more additional polypeptides having a nucleotide polymerase activity.

Recombination Reactions

[0107] Those skilled in the art can readily optimize the conditions for conducting the recombination reactions described above without the use of undue experimentation, based on the guidance provided herein and available in the art (see, *e.g.*, U.S. Patent Nos. 5,888,732 and 6,277,608, which are specifically incorporated herein by reference in their entireties). In a typical reaction from, about 50 ng to about 1000 ng of a second nucleic acid molecule may be contacted with a first nucleic acid molecule under suitable reaction conditions. Each nucleic acid molecule may be present in a molar ratio of from about 25:1 to about 1:25 first nucleic acid molecule:second nucleic acid molecule. In some embodiments, a first nucleic acid molecule may be present at a molar ratio of from about 10:1 to 1:10 first nucleic acid molecule:second nucleic acid molecule. In a preferred embodiment, each nucleic acid molecule may be present at a molar ratio of about 1:1 first nucleic acid molecule:second nucleic acid molecule.

[0108] Typically, the nucleic acid molecules may be dissolved in an aqueous buffer and added to the reaction mixture. One suitable set of conditions is 4 μ l CLONASETM enzyme mixture (*e.g.*, Invitrogen Corporation, Cat. Nos. 11791-019 and 11789-013), 4 μ l 5X reaction buffer and nucleic acid and water to a final volume of 20 μ l. This will typically result in the inclusion of about 200 ng of Int and about 80 ng of IHF in a 20 μ l BP reaction and about 150 ng Int, about 25 ng IHF and about 30 ng Xis in a 20 μ l LR reaction.

[0109] Proteins for conducting an LR reaction may be stored in a suitable buffer, for example, LR Storage Buffer, which may comprise about 50 mM Tris at about pH 7.5, about 50 mM NaCl, about 0.25 mM EDTA, about 2.5 mM Spermidine, and about 0.2 mg/ml BSA. When stored, proteins for an LR

reaction may be stored at a concentration of about 37.5 ng/μl INT, 10 ng/μl IHF and 15 ng/μl XIS. Proteins for conducting a BP reaction may be stored in a suitable buffer, for example, BP Storage Buffer, which may comprise about 25 mM Tris at about pH 7.5, about 22 mM NaCl, about 5 mM EDTA, about 5 mM Spermidine, about 1 mg/ml BSA, and about 0.0025% Triton X-100. When stored, proteins for an BP reaction may be stored at a concentration of about 37.5 ng/μl INT and 20 ng/μl IHF. One skilled in the art will recognize that enzymatic activity may vary in different preparations of enzymes. The amounts suggested above may be modified to adjust for the amount of activity in any specific preparation of enzymes.

[0110] A suitable 5X reaction buffer for conducting recombination reactions may comprise 100 mM Tris pH 7.5, 88 mM NaCl, 20 mM EDTA, 20 mM Spermidine, and 4 mg/ml BSA. Thus, in a recombination reaction, the final buffer concentrations may be 20 mM Tris pH 7.5, 17.6 mM NaCl, 4 mM EDTA, 4 mM Spermidine, and 0.8 mg/ml BSA. Those skilled in the art will appreciate that the final reaction mixture may incorporate additional components added with the reagents used to prepare the mixture, for example, a BP reaction may include 0.005% Triton X-100 incorporated from the BP Clonase™.

[0111] In some preferred embodiments, particularly those in which *attL* sites are to be recombined with *attR* sites, the final reaction mixture may include about 50 mM Tris HCl, pH 7.5, about 1 mM EDTA, about 1 mg/ml BSA, about 75 mM NaCl and about 7.5 mM spermidine in addition to recombination enzymes and the nucleic acids to be combined. In other preferred embodiments, particularly those in which an *attB* site is to be recombined with an *attP* site, the final reaction mixture may include about 25 mM Tris HCl, pH 7.5, about 5 mM EDTA, about 1 mg/ml bovine serum albumin (BSA), about 22 mM NaCl, and about 5 mM spermidine.

[0112] In some preferred embodiments, particularly those in which *attL* sites are to be recombined with *attR* sites, the final reaction mixture may include about 40 mM Tris HCl, pH 7.5, about 1 mM EDTA, about 1 mg/ml BSA, about 64 mM NaCl and about 8 mM spermidine in addition to recombination enzymes and the nucleic acids to be combined. One of skill in the art will

appreciate that the reaction conditions may be varied somewhat without departing from the invention. For example, the pH of the reaction may be varied from about 7.0 to about 8.0; the concentration of buffer may be varied from about 25 mM to about 100 mM; the concentration of EDTA may be varied from about 0.5 mM to about 2 mM; the concentration of NaCl may be varied from about 25 mM to about 150 mM; and the concentration of BSA may be varied from 0.5 mg/ml to about 5 mg/ml. In other preferred embodiments, particularly those in which an *attB* site is to be recombined with an *attP* site, the final reaction mixture may include about 25 mM Tris HCl, pH 7.5, about 5 mM EDTA, about 1 mg/ml bovine serum albumin (BSA), about 22 mM NaCl, about 5 mM spermidine and about 0.005% detergent (*e.g.*, Triton X-100).

Expression of Proteins

[0113] In general, expression of a protein using the methods of the invention may comprise contacting a nucleic acid molecule (*e.g.*, a template such as a third nucleic acid molecule as described above) with a cell extract under conditions sufficient to cause the expression of the protein. The general system includes a nucleic acid molecule (*e.g.*, a third nucleic acid molecule as described above) that functions as a template and encodes one or more desired proteins. The desired protein(s) can be any polymer of amino acids encodable by a nucleic acid template to produce a protein molecule. The protein can be further processed (*e.g.*, modified, purified, etc) coincident with or after synthesis. When desired, the system can be altered as known in the art such that an amino acid not normally incorporated into a polypeptide is incorporated into a polypeptide in response to specified codons in an mRNA molecule. Thus, for example, codons in an mRNA that normally encode a first amino acid may result in incorporation into a polypeptide of a second amino acid. In some embodiments, a second amino acid may be a normally occurring amino acid (*e.g.*, Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, or Val). In some embodiments, a second amino acid may be an unconventional or unnatural amino acids (*e.g.*, detectably labeled amino acids). This may be accomplished, for example, by

including tRNA molecules charged with molecules not normally incorporated into polypeptides by cells (see, for example, United States patent nos. 5,536,652 and 5,370,995, which are specifically incorporated herein by reference).

[0114] Typically, a template nucleic acid molecule produced as described above may comprise one or more sequences that function as promoters when contacted with a DNA-dependent RNA polymerase enzyme. To produce proteins according to the invention, a template nucleic acid molecule comprising a promoter may be contacted with a polypeptide having DNA-dependent RNA polymerase activity under conditions causing the synthesis of one or more RNA molecules corresponding to the sequence of the template nucleic acid molecule encoding one or more protein of interest. Suitable conditions may include one or more nucleotides (*e.g.*, ribonucleotides), one or more buffers, one or more metal cofactors (*e.g.*, Mg^{2+}), one or more nuclease inhibitors (*e.g.*, RNase inhibitors) and the like. The RNA molecules produced may be purified using standard techniques. Optionally, the reaction mixture containing the RNA molecules may be added directly to a protein synthesis reaction, which may comprise cell extract, as described in more detail below. Alternatively, a template nucleic acid molecule may be contacted with the polypeptide having DNA-dependent RNA polymerase activity in the presence of all the necessary components to synthesize a protein and the RNA produced may be directly translated into protein without any further manipulation, *i.e.*, a coupled *in vitro* transcription/translation reaction (IVTT) may be performed.

[0115] Extracts to be used for the synthesis of proteins can be made from any suitable cells. Suitable cells are those that have components for protein and/or nucleic acid synthesis, optionally produced under selected growth conditions or with modification(s) and/or mutation(s) that inactivate or reduce or substantially reduce unwanted properties detrimental to *in vitro* synthesis (*e.g.*, nuclease activities, phosphatase activities, etc.). Cells from which an extract may be prepared and used according to the invention include, but are not limited to, bacterial cells, fungal and yeast cells, plant cells and animal cells. Preferred bacterial cells include, for example, gram positive and gram negative bacteria, and especially *Escherichia* spp. cells particularly *E. coli* cells and

most particularly *E. coli* strains DH10B, Stbl2, DH5 α , DB3, DB3.1 (preferably *E. coli* LIBRARY EFFICIENCY[®] DB3.1[™] Competent Cells; Invitrogen Corporation, Carlsbad, CA), DB4 and DB5 (*see* U.S. Application No. 09/518,188, filed on March 2, 2000, the disclosure of which are incorporated by reference herein in their entireties), BL21, HB101, INV110, INV α F', MC1061/P3, PIR1, PIR2, TOP10, TOP10F', TOP10/P3, etc. Other preferred bacterial host cells include, but are not limited to, *Bacillus* spp. cells (particularly *B. subtilis* and *B. megaterium* cells), *Streptomyces* spp. cells, *Erwinia* spp. cells, *Klebsiella* spp. cells, *Serratia* spp. cells (particularly *S. marcessans* cells), *Pseudomonas* spp. cells (particularly *P. aeruginosa* cells), and *Salmonella* spp. cells (particularly *S. typhimurium* and *S. typhi* cells). Preferred animal host cells include insect cells (especially, *Drosophila melanogaster* cells and more especially S2 cells, *Spodoptera frugiperda* Sf9 and Sf21 cells and *Trichoplusia* High-Five cells), nematode cells (particularly *C. elegans* cells), avian cells, amphibian cells (particularly *Xenopus laevis* cells), reptilian cells, and mammalian cells (most particularly NIH3T3, CHO, COS, C127, VERO, BHK, HeLa, 293, Per-C6, Bowes melanoma, and human, rabbit, mouse, rat, hamster, pig, bovine and gerbil cells generally). Preferred fungal and yeast host cells include *Aspergillus* spp. cells, *Saccharomyces cerevisiae* cells and *Pichia pastoris* cells. These and other suitable host cells are available commercially, for example, from Invitrogen Corporation (Carlsbad, CA), American Type Culture Collection (Manassas, Virginia), and Agricultural Research Culture Collection (NRRL; Peoria, Illinois), the catalogues of each being incorporated in their entireties by reference. Plant cells are exemplified by protoplasts, tobacco, potato and other tuberous plants, grasses including maize, cotton and other fibrous plants, annuals and perennials, monocots and dicots, and especially plant cells that can be transformed and/or grown in culture.

[0116] The cell extract can be supplemented to provide components not present or not present in sufficient quantities after extraction.

[0117] The extract can be prepared by any method used in the art that maintains the integrity of the transcription/translation system or, if the process damages one or more component necessary for any stage of

transcription/translation, the damaged component can be replaced or substituted for after the extract preparation.

[0118] Extracts can be prepared according to the method of Zubay (1973), *Annu. Rev. Genet.* 7, 267-287. The ordinarily skilled artisan will recognize that many modifications to the extraction process are possible within the scope of the present invention. In some embodiments, extracts may be supplemented with one or more polymerase enzymes, for example, T7 polymerase, SP6 polymerase and the like. Other methods of preparing and/or supplementing extracts suitable for IVTT reactions are known in the art, for example, as disclosed in WO 02/072890 and published United States patent application 2002-0168706 A1.

[0119] Other embodiments may utilize extracts prepared from eukaryotic organisms, for example, wheat germ. Suitable eukaryotic systems are well known in the art (see, for example, United States Patent nos. 5,665,563 and 5,492,817). Suitable extracts of wheat germ can be prepared by a methods known in the art (see, for example, Roberts, B. E. and Paterson, B. M., *Proc. Natl. Acad. Sci. USA* Vol. 70, No. 8, pp. 2330-2334, (1973); and C. W., *et al. Meth. Enzymol.* Vol. 101, p. 635, (1983)). In brief, wheat germ extract may be prepared by grinding wheat germ in an extraction buffer, followed by centrifugation to remove cell debris. The supernatant is then separated by chromatography from endogenous amino acids and plant pigments that are inhibitory to translation. The extract may also be treated with one or more nuclease enzymes (*e.g.*, micrococcal nuclease) to destroy endogenous mRNA, thereby reducing the production of undesired proteins. The extract may contain and/or may be supplemented with the cellular components necessary for protein synthesis, such as tRNA, rRNA and initiation, elongation, and termination factors. The extract may be supplemented with one or more components, (*e.g.*, energy generating systems, metal ions, buffers, etc). Suitable eukaryotic cell extracts are commercially available, for example, the TNT® Coupled Wheat Germ Extract System and TNT® Coupled Reticulocyte Lysate System, from Promega Corporation, Madison, WI; and the Retic Lysate IVT™ and Wheat Germ IVT™, from Ambion, Inc. Austin, TX. Other suitable systems and protocols are described in *Methods in*

Molecular Biology, Vol. 37: *In Vitro Transcription and Translation Protocols*, Edited by: M.J. Tymms, Copyright 1995, Human Press Inc., Totowa, NJ.

- [0120] A non-limiting example of a suitable set of reaction conditions for conducting a transcription/translation reaction is 57 mM Hepes/KOH pH 8.2, 230 mM K-glutamate, 1.2 mM ATP, 0.85 mM each of GTP, UTP, CTP, 30 mM PEP, 1.7 mM DTT, 12 mM Mg(OAc)₂, 0.17 mg/ml *E. coli* total RNA mixture, 34 µg/ml folinic acid, 66 µg/ml T7 RNA polymerase, 1.25 mM each of amino acids, 14 mg/ml extract, 3 µl of ³⁵S-Met (15 µCi/µl) if labeled proteins are desired and various amount of DNA template.
- [0121] Other reaction conditions for conducting *in vitro* transcription/translation reactions are known to those skilled in the art. For example, a reaction buffer may comprise all the required amino acids and an ATP regenerating system (see, for example, Kim *et al.*, (1996) *Eur. J. Biochem.* 239, 881-886; Lesley, *et al.* (1991) *J. Biol. Chem.* 266, 2632-2638; and Pratt (1984) *Transcription and Translation* (Oxford: S.J. IRL Press).
- [0122] The skilled artisan will recognize that concentrations of the various components of the incubation medium can be adjusted as is known in the art while still maintaining the synthetic function. For example, the pH can range from about 5.6 to about 8.8 or more preferably about 6.1 to 8.5 or even about 7.2 to 8.4 depending on the product to be synthesized. Likewise the concentration of K-glutamate can be easily varied between about 80 mM and 320 mM or more preferably about 120 to 280 mM or even about 180 to 250 mM; the concentration of ATP can be varied from about 0.1 to 3.0 mM or more preferably about 0.3 to 2.0 mM or even about 0.8 to 1.5 mM; the concentrations of GTP, UTP CTP and TTP can vary from about 0 or 0.1 to 2 mM or more preferably about 0.4 to 1.2 mM or even about 0.7 to 1.0 mM; PEP can vary from about 10 to 60 mM or more preferably about 20 to 50 mM or even about 25 to 40 mM; DTT can vary from about 0.5 to 3.0 mM or more preferably about 1.0 to 2.5 mM or even about 1.5 to 2.0 mM; magnesium can vary from about 7.5 to 20 mM or more preferably about 10 to 15 mM or even about 11.5 to 14 mM; total RNA mixture can vary from about 0.1 to 0.5 mg/ml or more preferably about 0.12 to 0.3 mg/ml or even 0.15 to 0.2 mg/ml; folinic acid can vary from about 15 to 50 µg/ml or more preferably about 25 to

40 µg/ml or even about 30 to 35 µg/ml; polymerase can vary from about 0 or 1.0 to 300 µg/ml or more preferably about 20 to 200 µg/ml or even about 30 to 150 µg/ml, about 40 to 120 µg/ml, about 50 to 100 µg/ml, or about 60 to 75 µg/ml; amino acids can vary independently from about 0.4 to 5 mM or more preferably about 0.5 to 2.0 mM or even 0.8 to 1.5 mM; extract can vary from about 2 to 40 mg protein/ml or more preferably about 5 to 25 mg protein/ml or even about 10 to 18 mg protein/ml; and a synthesis marker, if desired, (³⁵S-Met in the example described above) can be selected as desired and used in any amount detectable that does not unduly compromise synthesis. The skilled artisan will readily recognize that many of the components have known substitutes or equivalents that might be used either in the same concentration or in a concentration that produces a qualitatively and/or quantitatively similar effect.

[0123] To produce proteins according to the methods of the invention, a third nucleic acid molecule as described above may be purified and resuspended in a suitable buffer (*e.g.*, 10 mM Tris, 1 mM EDTA, pH 7.0 to 8.0) or water such that the final concentration is greater than 500 ng/µl. It is desirable to ensure that the purified third nucleic acid molecule is not contaminated with ethanol, salt, or RNases. A third nucleic acid molecule may be purified using techniques well known in the art, for example, commercial DNA purification kits (*e.g.*, the S.N.A.P.[™] MiniPrep Kit (Catalog no. K1900-01 available from Invitrogen Corporation, Carlsbad, CA). For protocols to purify DNA, refer to published reference sources (Ausubel *et al.*, (1994) *Current Protocols in Molecular Biology*; Sambrook *et al.*, (1989) *Molecular Cloning: A Laboratory Manual*, Second Edition (Plainview, New York: Cold Spring Harbor Laboratory Press). A third nucleic acid molecule should not be purified by gel electrophoresis as purified DNA solution obtained from agarose gels significantly inhibits the protein synthesis reaction. Also, ammonium acetate is not recommended for use in DNA precipitation as any residual contamination may inhibit translation. Sodium acetate may be used in DNA precipitation. Care should be taken to avoid contamination of the solution containing the third nucleic acid molecule with nucleases, *e.g.*, RNases. Standard techniques, *e.g.*, wearing gloves and using RNase-free reagents can

be employed. Purified third nucleic acid molecule should be free of excess ethanol or salt as both can inhibit translation. When ethanol precipitation is used to purify DNA, DNA should be carefully washed with 70% ethanol to remove excess salt and dried.

[0124] To synthesize protein in an IVTT, a third nucleic acid molecule may be mixed with an appropriate amount of cell extract, optionally supplemented with one or more ingredients (*e.g.*, buffers, polymerases, amino acids, etc.) and incubated, for example, at 37 °C for from about 10 minutes to about 10 hours, from about 30 minutes to about 8 hours, from about 1 hour to about 8 hours, from about 1 hour to about 7 hours, from about 1 hour to about 6 hours, from about 1 hour to about 5 hours, from about 1 hour to about 4 hours, from about 1 hour to about 3 hours, or from about 1 hour to about 2 hours. Optionally, the reaction may be shaken, for example, using an Eppendorf Thermomixer (Fisher, Catalog no. 05-400-200) or other equivalent means.

[0125] After protein is produced, the protein may be further processed (*e.g.*, purified, analyzed etc.) using techniques known in the art. Suitable techniques include, but are not limited to, Coomassie[®]-stained protein gel, Western blot, enzymatic assays, affinity purification (if affinity tag is present), and the like.

Kits

[0126] In another aspect, the invention provides kits that may be used in conjunction with methods the invention. Kits according to this aspect of the invention may comprise one or more containers, which may contain one or more components selected from the group consisting of one or more nucleic acid molecules, one or more primers, one or more polymerases, one or more recombination proteins (or other enzymes for carrying out the methods of the invention), one or more buffers, one or more detergents, one or more restriction endonucleases, one or more nucleotides, pyrophosphatase, and the like.

[0127] A wide variety of nucleic acid molecules can be used with the invention. Examples of nucleic acid molecules that can be supplied in kits of the invention include those that contain or encode promoters, signal peptides, enhancers, repressors, suppressors, selection markers, transcription signals,

translation signals, primer hybridization sites (*e.g.*, for sequencing or PCR), recombination sites, restriction sites and polylinkers, sequences that encode domains and/or regions (*e.g.*, 6 His tag) for the preparation of fusion proteins, and the like. Similarly, libraries can be supplied in kits of the invention. These libraries may be in the form of replicable nucleic acid molecules or they may comprise nucleic acid molecules that are not associated with an origin of replication. As one skilled in the art would recognize, the nucleic acid molecules of libraries, as well as other nucleic acid molecules that are not associated with an origin of replication, either could be inserted into other nucleic acid molecules that have an origin of replication or could be provided without being associated with an origin or replication in the kits of the invention.

[0128] Further, in some embodiments, libraries supplied in kits of the invention may be used as first nucleic acid molecules as described above and may comprise two components: (1) the nucleic acid molecules of these libraries and (2) 5' and/or 3' recombination sites. In some embodiments, when the nucleic acid molecules of a library are supplied with 5' and/or 3' recombination sites, it will be possible to recombine these molecules with a second nucleic acid molecule as described above to produce a third nucleic acid molecule, using recombination reactions. In other embodiments, recombination sites can be attached to the nucleic acid molecules of the libraries before use (*e.g.*, by the use of a ligase, which may also be supplied with the kit). In such cases, nucleic acid molecules that contain recombination sites or primers that can be used to generate recombination sites may be supplied with the kits.

[0129] Kits of the invention may further comprise containers containing one or more recombination proteins. Suitable recombination proteins have been disclosed above and include, but are not limited to, Cre, Int, IHF, Xis, Flp, Fis, Hin, Gin, Cin, Tn3 resolvase, Φ C31, TndX, XerC, and XerD.

[0130] Kits of the invention can also comprise one or more primers. These primers will generally be designed to anneal to molecules having specific nucleotide sequences. For example, these primers can be designed for use in PCR to amplify a particular nucleic acid molecule. Further, primers supplied with kits of the invention can be sequencing primers designed to hybridize to

vector sequences. Thus, such primers will generally be supplied as part of a kit for sequencing nucleic acid molecules that have been inserted into a vector.

[0131] One or more buffers (*e.g.*, one, two, three, four, five, eight, ten, fifteen) may be supplied in kits of the invention. These buffers may be supplied at a working concentrations or may be supplied in concentrated form and then diluted to the working concentrations. These buffers will often contain salt, metal ions, co-factors, metal ion chelating agents, etc. for the enhancement of activities of the stabilization of either the buffer itself or molecules in the buffer. Further, these buffers may be supplied in dried or aqueous forms. When buffers are supplied in a dried form, they will generally be dissolved in water prior to use.

[0132] Kits of the invention may contain virtually any combination of the components set out above or described elsewhere herein, or other components known to those of ordinary skill to be useful in carrying out methods of the invention. As one skilled in the art would recognize, the components supplied with kits of the invention will vary with the intended use for the kits. Thus, kits may be designed to perform various functions set out in this application and the components of such kits will vary accordingly.

[0133] Kits of the invention may further comprise one or more pages of written instructions for carrying out the methods of the invention. For example, instructions may comprise methods steps necessary to carryout recombinational cloning of an ORF provided with recombination sites and a vector also comprising recombination sites and optionally further comprising one or more functional sequences (*e.g.*, transcriptional regulatory sequences). Instructions may comprise the steps necessary to carry out an *in vitro* transcription/translation reaction and/or an amplification reaction.

Business Methods

[0134] The present invention also provides a system and method of providing company products to a party outside of the company, for example, a system and method for providing a customer or a product distributor a product of the company such as a kit containing materials for practicing the methods of the invention and/or instructions for practicing such methods. Figure 4 provides a

schematic diagram of a product management system. In practice, the blocks in Figure 4 can represent an intra-company organization, which may include departments in a single building or in different buildings, a computer program or suite of programs maintained by one or more computers, a group of employees, a computer I/O device such as a printer or fax machine, a third party entity or company that is otherwise unaffiliated with the company, or the like.

[0135] The product management system as shown in Figure 4 is exemplified by company 100, which receives input in the form of an order from a party outside of the company, *e.g.*, distributor 150 or customer 140, to order department 126, or in the form of materials and parts 130 from a party outside of the company; and provides output in the form of a product delivered from shipping department 119 to distributor 150 or customer 140. Company 100 system is organized to optimize receipt of orders and delivery of a products to a party outside of the company in a cost efficient manner, particularly instructions or a kit of the present invention, and to obtain payment for such product from the party in a timely manner.

[0136] With respect to the methods of the present invention, the term "materials and parts" refers to items that are used to make a device, other component, or product, which generally is a device, other component, or product that company sells to a party outside of the company. As such, materials and parts include, for example, recombinases, starting nucleic acid molecules with one or more recombination sites, nucleic acid segments, nucleotides, host cells, polymerases, amino acids, buffers, paper, ink, reaction vessels, etc.

[0137] In comparison, the terms "devices", "other components", and "products" refer to items sold by the company. Other components are exemplified by instructions, including instructions for the practice of the methods of the invention. Other components also can be items that may be included in a kit, *e.g.*, a kit product, for example, reagents for manipulating nucleic acid molecules (*e.g.*, performing recombinational cloning). Kits may include multiple marker molecules and instructions for use. Reagents may include, for example, buffers, salts, cofactors and the like. Other components may include host cell extracts. As such, it will be recognized that an item

useful as materials and parts as defined herein further can be considered another component, which can be sold by the company.

[0138] The term "commercial products" refers to devices, other components, or combinations thereof, including combinations with additional materials and parts, that are sold or desired to be sold or otherwise provided by a company to one or more parties outside of the company. Commercial products are exemplified herein by kits, which can contain instructions according to the present invention, and one or more nucleic acid molecules, one or more cell extracts, one or more enzymes (*e.g.*, polymerases), one or more primers, one or more reagents, one or more buffers or buffer salts, and/or combinations thereof. Nucleic acid molecules may comprise one or more recombination sites.

[0139] Referring to Figure 4, company 100 includes manufacturing 110 and administration 120. Devices 112 and other components 114 are produced in manufacturing 110, and can be stored separately therein such as in device storage 113 and other component storage 115, respectively, or can be further assembled and stored in product storage 117. Materials and parts 130 can be provided to company 100 from an outside source and/or materials and parts 114 can be prepared in company, and used to produce devices 112 and other components 116, which, in turn, can be assembled and sold as a product. Manufacturing 110 also includes shipping department 119, which, upon receiving input as to an order, can obtain products to be shipped from product storage 117 and forward the product to a party outside the company.

[0140] For purposes of the present invention, product storage 117 can be used to store instructions for practicing methods of the invention; or can store kits, which can contain at least one nucleic acid molecule, one or more polymerases, one or more primers and/or one or more cell extracts, and, optionally, instructions as disclosed herein; or can store a combination of such instructions and/or kits. Upon receiving input from order department 126, for example, that customer 140 has ordered such a kit and instructions, shipping department 119 can obtain from product storage 117 such kit for shipping, and can further obtain such instructions in a written form to include with the kit, and ship the kit and instructions to customer 140 (and providing input to billing department 124 that the product was shipped; or shipping department

119 can obtain from product storage 117 the kit for shipping, and can further provide the instructions to customer 140 in an electronic form, by accessing a database in company 100 that contains the instructions, and transmitting the instructions to customer 140 via the internet (not shown).

[0141] As further exemplified in Figure 4, administration 120 includes order department 126, which receives input in the form of an order for a product from customer 140 or distributor 150. Order department 126 then provides output in the form of instructions to shipping department 119 to fill the order, *i.e.*, to forward products as requested to customer 140 or distributor 150. Shipping department 119, in addition to filling the order, further provides input to billing department 124 in the form of confirmation of the products that have been shipped. Billing department 124 then can provide output in the form of a bill to customer 140 or distributor 150 as appropriate, and can further receive input that the bill has been paid, or, if no such input is received, can further provide output to customer 140 or distributor 150 that such payment may be delinquent.

[0142] Additional optional components of company 100 include customer service department 122, which can receive input from customer 140 and can provide output in the form of feedback or information to customer 140. Furthermore, although not shown in Figure 4, customer service 122 can receive input or provide output to any other component of company. For example, customer service department 122 can receive input from customer 140 indicating that an ordered product was not received, wherein customer service department 122 can provide output to shipping department 119 and/or order department 126 and/or billing department 124 regarding the missing product, thus providing a means to assure customer 140 satisfaction. Customer service department 122 also can receive input from customer 140 in the form of requested technical information, for example, for confirming that instructions of the invention can be applied to the particular need of customer 140, and can provide output to customer 140 in the form of a response to the requested technical information.

[0143] As such, the components of company 100 are suitably configured to communicate with each other to facilitate the transfer of materials and parts, devices, other components, products, and information within company 100,

and company 100 is further suitably configured to receive input from or provide output to an outside party. For example, a physical path can be utilized to transfer products from product storage 117 to shipping department 119 upon receiving suitable input from order department 126. Order department 126, in comparison, can be linked electronically with other components within company 100, for example, by a communication network such as an intranet, and can be further configured to receive input, for example, from customer 140 by a telephone network, by mail or other carrier service, or via the internet. For electronic input and/or output, a direct electronic link such as a T1 line or a direct wireless connection also can be established, particularly within company 100 and, if desired, with distributor 150 or materials or parts 130 provider, or the like.

[0144] Although not illustrated, company 100 may comprise one or more data collection systems, including, for example, a customer data collection system, which can be realized as a personal computer, a computer network, a personal digital assistant (PDA), an audio recording medium, a document in which written entries are made, any suitable device capable of receiving data, or any combination of the foregoing. Data collection systems can be used to gather data associated with a customer 140 or distributor 150, including, for example, a customer's shipping address and billing address, as well as more specific information such as the customer's ordering history and payment history, such data being useful, for example, to determine that a customer has made sufficient purchases to qualify for a discount on one or more future purchases.

[0145] Company 100 can utilize a number of software applications to provide components of company 100 with information or to provide a party outside of company access to one or more components of company 100, for example, access to order department 126 or customer service department 122. Such software applications can comprise a communication network such as the Internet, a local area network, or an intranet. For example, in an internet-based application, customer 140 can access a suitable web site and/or a web server that cooperates with order department 126 such that customer 140 can provide input in the form of an order to order department 126. In response, order department 126 can communicate with customer 140 to confirm that the order has been received, and can further communicate with shipping

department 119, providing input that products such as a kit of the invention, which contains, for example, one or more nucleic acid molecules having one or more recombination sites, one or more polymerases, one or more primers and/or one or more cell extracts and instructions for use, should be shipped to customer 140. In this manner, the business of company 100 can proceed in an efficient manner.

[0146] In a networked arrangement, billing department 124 and shipping department 119, for example, can communicate with one another by way of respective computer systems. As used herein, the term "computer system" refers to general purpose computer systems such as network servers, laptop systems, desktop systems, handheld systems, personal digital assistants, computing kiosks, and the like. Similarly, in accordance with known techniques, distributor 150 can access a web site maintained by company 100 after establishing an online connection to the network, particularly to order department 126, and can provide input in the form of an order. If desired, a hard copy of an order placed with order department 126 can be printed from the web browser application resident at distributor 150.

[0147] The various software modules associated with the implementation of the present invention can be suitably loaded into the computer systems resident at company 100 and any party outside of company 100 as desired, or the software code can be stored on a computer-readable medium such as a floppy disk, magnetic tape, or an optical disk. In an online implementation, a server and web site maintained by company 100 can be configured to provide software downloads to remote users such as distributor 150, materials and parts 130, and the like. When implemented in software, the techniques of the present invention are carried out by code segments and instructions associated with the various process tasks described herein.

[0148] Accordingly, the present invention further includes methods for providing various aspects of a product (e.g., a kit and/or instructions of the invention), as well as information regarding various aspects of the invention, to parties such as the parties shown as customer 140 and distributor 150 in Figure 4. Thus, methods for selling devices, products and methods of the invention to such parties are provided, as are methods related to those sales, including customer support, billing, product inventory management within the

company, etc. Examples of such methods are shown in Figure 4, including, for example, wherein materials and parts 130 can be acquired from a source outside of company 100 (*e.g.*, a supplier) and used to prepare devices used in preparing a composition or practicing a method of the invention, for example, kits, which can be maintained as an inventory in product storage 117. It should be recognized that devices 112 can be sold directly to a customer and/or distributor (not shown), or can be combined with one or more other components 116, and sold to a customer and/or distributor as the combined product. The other components 116 can be obtained from a source outside of company 100 (materials and parts 130) or can be prepared within company 100 (materials and parts 114). As such, the term "commercial product" is used generally herein to refer an item sent to a party outside of the company (a customer, a distributor, etc.) and includes items such as devices 112, which can be sent to a party alone or as a component of a kit or the like.

[0149] At the appropriate time, the product is removed from product storage 117, for example, by shipping department 119, and sent to a requesting party such as customer 140 or distributor 150. Typically, such shipping occurs in response to the party placing an order, which is then forwarded the within the organization as exemplified in Figure 4, and results in the ordered product being sent to the party. Data regarding shipment of the product to the party is transmitted further within the organization, for example, from shipping department 119 to billing department 124, which, in turn, can transmit a bill to the party, either with the product, or at a time after the product has been sent. Further, a bill can be sent in instances where the party has not paid for the product shipped within a certain period of time (*e.g.*, within 30 days, within 45 days, within 60 days, within 90 days, within 120 days, within from 30 days to 120 days, within from 45 days to 120 days, within from 60 days to 120 days, within from 90 days to 120 days, within from 30 days to 90 days, within from 30 days to 60 days, within from 30 days to 45 days, within from 60 days to 90 days, etc.). Typically, billing department 124 also is responsible for processing payment(s) made by the party. It will be recognized that variations from the exemplified method can be utilized; for example, customer service department 122 can receive an order from the party, and transmit the order to

shipping department 119 (not shown), thus serving the functions exemplified in Figure 4 by order department 126 and the customer service department 122.

[0150] The methods of the invention also include providing technical service to parties using a product, particularly a kit of the invention. While such a function can be performed by individuals involved in product research and development, inquiries related to technical service generally are handled, routed, and/or directed by an administrative department of the organization (*e.g.*, customer service department 122). Often communications related to technical service (*e.g.*, solving problems related to use of the product or individual components of the product) require a two way exchange of information, as exemplified by arrows indicating pathways of communication between customer 150 and customer service department 122.

[0151] As mentioned above, any number of variations of the process exemplified in Figure 4 are possible and within the scope of the invention. Accordingly, the invention includes methods (*e.g.*, business methods) that involve (1) the production of products (*e.g.*, nucleic acid and/or protein molecules, kits that contain instructions for performing methods of the invention, etc.); (2) receiving orders for these products; (3) sending the products to parties placing such orders; (4) sending bills to parties obliged to pay for products sent to such; and/or (5) receiving payment for products sent to parties. For example, methods are provided that comprise two or more of the following: (a) obtaining parts, materials, and/or components from a supplier; (b) preparing one or more first products (*e.g.*, one or more nucleic acid and/or protein molecules); (c) storing the one or more first products of (b); (d) combining the one or more first products of (b) with one or more other components to form one or more second products (*e.g.*, a kit); (e) storing the one or more first products of (b) or one or more second products of (d); (f) obtaining an order a first product of (b) or a second product of (d); (g) shipping either the first product of (b) or the second product of (d) to the party that placed the order of (f); (h) tracking data regarding to the amount of money owed by the party to which the product is shipped in (g); (i) sending a bill to the party to which the product is shipped in (g); (j) obtaining payment for the product shipped in (g) (generally, but not necessarily, the payment is made by the party to which the product was shipped in (g); and (k) exchanging

technical information between the organization and a party in possession of a product shipped in (d) (typically, the party to which the product was shipped in (g)).

[0152] The present invention also provides a system and method for providing information as to availability of a product (*e.g.*, a device product, a kit product, and the like) to parties having potential interest in the availability of the kit product. Such a method of the invention, which encompasses a method of advertising to the general or a specified public, the availability of the product, particularly a product comprising instructions and/or a kit of the present invention, can be performed, for example, by transmitting product description data to an output source, for example, an advertiser; further transmitting to the output source instructions to publish the product information data in media accessible to the potential interested parties; and detecting publication of the data in the media, thereby providing information as to availability of the product to parties having potential interest in the availability of the product.

[0153] Accordingly, the present invention provides methods for advertising and/or marketing devices, products, and/or methods of the invention, such methods providing the advantage of inducing and/or increasing the sales of such devices, products, and/or methods. For example, advertising and/or marketing methods of the invention include those in which technical specifications and/or descriptions of devices and/or products; methods of using the devices and/or products; and/or instructions for practicing the methods and/or using the devices and/or products are presented to potential interested parties, particularly potential purchasers of the product such as customers, distributors, and the like. In particular embodiments, the advertising and/or marketing methods involve presenting such information in a tangible form or in an intangible to the potential interested parties. As disclosed herein and well known in the art, the term "intangible form" means a form that cannot be physically handled and includes, for example, electronic media (*e.g.*, e-mail, internet web pages, etc.), broadcasts (*e.g.*, television, radio, etc.), and direct contacts (*e.g.*, telephone calls between individuals, between automated machines and individuals, between machines, etc.); whereas the term "tangible form" means a form that can be physically handled.

[0154] Figure 5 provides a schematic diagram of an information providing management system as encompassed within the present invention. In practice, the blocks in Figure 5 can represent an intra-company organization, which can include departments in a single building or in different buildings, a computer program or suite of programs maintained by one or more computers, a group of employees, a computer I/O device such as a printer or fax machine, a third party entity or company that is otherwise unaffiliated with the company, or the like.

[0155] The information providing management system as shown in Figure 5 is exemplified by company 200, which makes, purchases, or otherwise makes available devices and methods 210 that alone, or in combination, provide products 220, for example, instructions, devices and/or kits of the present invention, that company 200 wishes to sell to interested parties. To this end, product descriptions 230 are made, providing information that would lead potential users to believe that products 220 can be useful to user. In order to effect transfer of product descriptions 230 to the potential users, product descriptions 230 is provided to advertising agency 240, which can be an entity separate from company 200, or to advertising department 260, which can be an entity related to company 200, for example, a subsidiary. Based on the product descriptions 230, advertisement 250 is generated and is provided to media accessible to potential purchasers of products 260, whom may then contact company 200 to purchase products 220.

[0156] By way of example, product descriptions 230 can be in a tangible form such as written descriptions, which can be delivered (*e.g.*, mailed, sent by courier, etc) to advertising agency 240 and/or advertising department 250, or can be in an intangible form such as entered into and stored in a database (*e.g.*, on a computer, in an electronic media, etc.) and transmitted to advertising agency 240 and/or advertising department 250 over a telephone line, T1 line, wireless network, or the like. Similarly, advertisement 250 can be a tangible or intangible form such that it conveniently and effectively can be provided to potential parties of interest (*e.g.*, potential purchasers of product 260). For example, advertisement 250 can be provided in printed form as flyers (*e.g.*, at a meeting or other congregation of potential interested parties) or as printed pages (or portions thereof) in magazines known to be read by the potential

interested parties (*e.g.*, trade magazines, journals, newspapers, etc.). In addition, or alternatively, advertisement 250 can be provided in the form of directed mailing of computer media containing the advertisement (*e.g.*, CDs, DVDs, floppy discs, etc.) or of e-mail (*i.e.*, mail or e-mail that is sent only to selected parties, for example, parties known to members of an organization that includes or is likely to include potential users of products 220); of web pages (*e.g.*, on a website provided by company 200, or having links to the company 200 website); or of pop-up or pop-under ads on web pages known to be visited by potential purchaser of products 260, and the like. Potential purchasers of products 260, upon being apprised of the availability of the products 220, for example, the kits of the present invention, then can contact company 200 and, if so desired, can order said products 220 for company 200 (see Figure 4)

[0157] It will be understood by one of ordinary skill in the relevant arts that other suitable modifications and adaptations to the methods and applications described herein are readily apparent from the description of the invention contained herein in view of information known to the ordinarily skilled artisan, and may be made without departing from the scope of the invention or any embodiment thereof. Having now described the present invention in detail, the same will be more clearly understood by reference to the following examples, which are included herewith for purposes of illustration only and are not intended to be limiting of the invention.

[0158] The entire disclosures of U.S. Appl. No. 08/486,139, (now abandoned), filed June 7, 1995, U.S. Appl. No. 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), U.S. Appl. No. 09/233,492, filed January 20, 1999, (now U.S. Patent No. 6,270,969), U.S. Appl. No. 09/233,493, filed January 20, 1999, (now U.S. Patent No. 6,143,557), U.S. Appl. No. 09/005,476, filed January 12, 1998, (now U.S. Patent No. 6,171,861), U.S. Appl. No. 09/432,085 filed November 2, 1999, U.S. Appl. No. 09/498,074 filed February 4, 2000, U.S. Appl. No. 60/065,930, filed October 24, 1997, U.S. Appl. No. 09/177,387, filed October 23, 1998, U.S. Appl. No. 09/296,280, filed April 22, 1999, (now U.S. Patent No. 6,277,608), U.S. Appl. No. 09/296,281, filed April 22, 1999, (now abandoned), U.S. Appl. No. 09/648,790, filed August 28, 2000, U.S. Appl. No. 09/855,797, filed May 16, 2001, U.S. Appl. No.

09/907,719, filed July 19, 2001, U.S. Appl. No. 09/907,900, filed July 19, 2001, U.S. Appl. No. 09/985,448, filed November 2, 2001, U.S. Appl. No. 60/108,324, filed November 13, 1998, U.S. Appl. No. 09/438,358, filed November 12, 1999, U.S. Appl. No. 60/161,403, filed October 25, 1999, U.S. Appl. No. 09/695,065, filed October 25, 2000, U.S. Appl. No. 09/984,239, filed October 29, 2001, U.S. Appl. No. 60/122,389, filed March 2, 1999, U.S. Appl. No. 60/126,049, filed March 23, 1999, U.S. Appl. No. 60/136,744, filed May 28, 1999, U.S. Appl. No. 09/517,466, filed March 2, 2000, U.S. Appl. No. 60/122,392, filed March 2, 1999, U.S. Appl. No. 09/518,188, filed March 2, 2000, U.S. Appl. No. 60/169,983, filed December 10, 1999, U.S. Appl. No. 60/188,000, filed March 9, 2000, U.S. Appl. No. 09/732,914, filed December 11, 2001, U.S. Appl. No. 60/284,528, filed April 19, 2001, U.S. Appl. No. 60/291,973, filed May 21, 2001, U.S. Appl. No. 60/318,902, filed September 14, 2001, U.S. Appl. No. 60/333,124, filed November 27, 2001, and U.S. Appl. No. 10/005,876, filed December 7, 2001, are herein incorporated by reference.

[0159] It will be readily apparent to those skilled in the relevant arts that other suitable modifications and adaptations to the methods and applications described herein are obvious and may be made without departing from the scope of the invention or any embodiment thereof. Having now described the present invention in detail, the same will be more clearly understood by reference to the following examples, which are included herewith for purposes of illustration only and are not intended to be limiting of the invention.

EXAMPLES

[0160] Facing the post-genomic sequence era, genome research is progressing from sequence analysis focusing on genome sequences or cDNA sequences to post-sequence analysis such as exhaustive and systematic protein function and structural analysis.

[0161] As an underlying protein expression system that is fast and efficient, a cell-free protein synthesis system, for example, one that uses a wheat germ extract, may be used. Cell-free protein synthesis systems that use wheat germ

extracts duplicate eukaryotic protein synthesis systems *in vitro*, and have high protein synthesis functionality. Moreover, the expression of proteins that affect cell physiology is also possible with such a system, and high-throughput synthesis is easy due to the *in vitro* test system.

[0162] With the advent of recombinational cloning techniques (*e.g.*, GATEWAY™), large numbers of clones of open reading frames adjacent to recombination sites have been prepared. In order to express with high throughput the proteins necessary for functional analysis from these clones, proteins may be synthesized using the following procedures: (1) construction of expression clones through an LR reaction, (2) amplification of template DNA through PCR, (3) *in vitro* mRNA synthesis, and (4) protein synthesis using a wheat germ cell-free protein synthesis system.

[0163] The methods of the invention disclosed herein enable *in vitro* performance of all processes from the construction of expression clones to protein synthesis. The methods described herein permit the completion of the entire process, expression clone through protein synthesis, more rapidly than was previously possible. Moreover, due to the fact that it is possible to process all clones with the same operations, operations may be conducted in a high-throughput method (*e.g.*, using a 96-well plate format).

[0164] The present invention overcomes many of the problems associated with prior art methodologies, such as those that require introduction of a template containing nucleic acid molecule into a host cell. In some countries, for example, Japan, experiments requiring the introduction of genetic material into a cell (*e.g.*, into *E. coli*) are classified as recombinant DNA experiments and are subject to governmental restrictions. Previously known methods are, therefore, subject to these restrictions making it difficult for scientists and students. Methods of the present invention do not require introduction of genetic material into a host cell and, therefore, are not subject to the same restrictions.

[0165] With reference to Figures 6A and 6B, a comparison of previously known methods (left side of both figures) and the present methods (right side of both figures) is presented. Fig. 6A provides a timeline of the production of a polypeptide. Fig. 6B shows a schematic representation of the reactions performed in one embodiment of the present invention. In Fig. 6B, it can be

seen that in known methods, an entry clone is reacted with a destination vector in an LR reaction (shown as (2) in Fig. 6B), the reaction product is transformed into *E. coli*, plasmid is purified and the purified plasmid is used as a template in an amplification reaction. In one embodiment of the present invention, a destination vector is first digested with a restriction enzyme (shown as (1) in Fig. 6B), an LR reaction is conducted (2), a PCR reaction is conducted on the LR reaction product (3), the PCR product (4) may be used in an *in vitro* transcription reaction (5), and the RNA transcripts thus produced may be used in an *in vitro* translation reaction (6). In another embodiment, the PCR product may be used in a coupled transcription/translation reaction.

[0166] With the conventional method, after expression clones for the wheat germ cell-free protein synthesis system are constructed through a GATEWAY™ LR reaction, they are transformed into a host cell (*e.g.*, *E. coli*) and a plasmid is prepared using standard techniques. Once prepared, the plasmid may be used as an amplification (*e.g.*, PCR) template. Use of a host cell introduces several potential problems. For example, proteins having toxicity or proteins with high physiological activity adversely affect the growth of *E. coli*, which can result in variation in plasmid yield. This variable yield makes it difficult to handle such clones in a high throughput system for protein expression. In addition, plasmid DNA prepared from a host cell may contain RNase contamination resulting in the degradation of mRNA produced from the template. The present invention obviates these difficulties.

EXAMPLE 1

[0167] With reference to Figure 7, first nucleic acid molecules comprising a sequence of interest encoding a polypeptide and comprising one or more recombination sites are prepared. In Figure 7, a first nucleic acid molecule is shown as an Entry Clone for use in a GATEWAY™ LR reaction. The sequence of interest (indicated as Gene(ORF) in Figure 7) is flanked with recombination sites (shown as *attL1* and *attL2* in Figure 7). The first nucleic acid molecule is shown as a plasmid with an origin of replication (*ori*) and a selectable marker (*e.g.*, the kanamycin resistance gene, Km). A second nucleic acid molecule comprising at least one transcriptional regulatory sequence and one or more

recombination sites is also prepared. In Figure 7, a second nucleic acid molecule is shown as a plasmid comprising a transcriptional regulatory sequence (*e.g.*, the SP6 promoter) and an additional sequence that affects transcription and/or translation (*e.g.*, the Ω translation enhancer sequence). The sequence of vector pEU3'His is provided in Table 4. This vector contains the Ω sequence, a translational enhancer from 5'-untranslated sequence of tobacco mosaic virus. The Ω sequence is a 71-mer having the following sequence:

5'-GTATTTTACAAACAATTACCAACAACAACAACAAACAACA
ACATTACATTTTACATTCTACAACACTACA-3'

Table 4. Nucleotide sequence of vector pEU3'His

atttaggtgacactatagaaTACACGGAATTCGCGTTTTTATTTTAAATTTTCTTTCAAATAC
 TTCTAGCTAGAGTATTTTTACAACAATTACCAACAACAACAACAACAACAACATTACAT
 TTTACATTCTACAACCTACAGATATCAACAAGTTTGTACAAAAAAGCTGAACGAGAAACGTAAA
 ATGATATAAATATCAATATATTAAATTAGATTTTGCATAAAAAACAGACTACATAACTGTGA
 AAACACAACATATCCAGTCACTATGGCGGCCGCATTAGGCACCCAGGCTTTACACTTTATGC
 TTCCGGCTCGTATAATGTGTGGATTTTGTAGTTAGGATCCGGCGAGATTTTCAGGAGCTAAGGA
 AGCTAAAATTGGAGAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAA
 AGAACATTTTGGAGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTTCAGCTGGA
 TATTACGGCCTTTTTTAAAGACCGTAAAGAAAAATAAGCACAGTTTTTATCCGGCCTTTATTCA
 CATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCCGTATGGCAATGAAAGACGGTGAGCT
 GGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTTC
 ATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGAAGATGT
 GGGCTGTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGT
 CTCAGCCAATCCCTGGGTGAGTTTACCAGTTTTGATTTAAACGTGGCCAATATGGACAACCTT
 CTTCCGCCCCGTTTTTACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCT
 GGCGATTTCAGGTTTCATCATGCCGTCTGTGATGGCTTCCATGTCCGCAGAAATGCTTAATGAATT
 ACAACAGTACTGCGATGAGTGGCAGGGCGGGCGTAAAgatcTGGATCCGGCTTACTAAAAGC
 CAGATAACAGTATGCGTATTTGCGCGCTGATTTTTGCGGTATAAGAATATATACTGATATGTA
 TACCCGAAGTATGTCAAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACAGTTGACAGC
 GACAGCAATCAGTTGCTCAAGGCATATATGATGTCAATATCTCCGGTCTGGTAAGCACAAACA
 TGCAGAATGAAGCCCGTCTGCTGCGTGCCGAACGCTGGAAGCGGAAAAATCAGGAAGGGATGG
 CTGAGGTGCGCCGTTTTATTGAAATGAACGGCTCTTTTGCTGACGAGAACAGGGACTGGTGAA
 ATGCAGTTTAAAGGTTTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAG
 AGTGATATTATTGACACGCCCGGGCGACGGATGGTGATCCCCCTGGCCAGTGACAGCTCTGCTG
 TCAGATAAAGTCTCCCGTGAACCTTTACCCGGTGGTGCATATCGGGGATGAAAGCTGGCGCATG
 ATGACCACCGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGGAAGAAAGTGGCTGATCTCAGC
 CACCGCGAAAATGACATCAAAAACGCCATTAACCTGATGTTCTGGGGAATATAAATGTCAGGC
 TCCCTTATACACAGCCAGTCTGCAGGTCGACCATAGTGACTGGATATGTTGTGTTTTACAGTA
 TTATGTAGTCTGTTTTTTATGCAAAATCTAATTTAATATATTGATATTTATATCATTTTACGT
 TTCTCGTTTCAGCTTTCTTGTACAAAGTGGTTCCACATCACCATCACCATTAAGATATCACTAG
 TTCTCGAGCTCGGTACCTGTCCGCGGTGCGGACGTACGCGGGCGGCCCATAAATTGGATCC
 ATATATAGGGCCCGGTTATAATTACCTCAGGTCGACGTCCCATGGTTTTGTATAGAATTTAC
 GGCTAGCGCCGATGCGACGCCGGTCCGCTTTATCCGGCCTTCTATATCAGGCTGTGTTTA
 AGACGCCGCGCTTCGCCCCAAATCCTTATGCCGGTTTCGACGGCTGGACAAAAATCTGTTTATC
 TTCCCAGCGCAGGCAGGTTAATGTACCACCCAGCAGCAGCCGGTATCCAGCGCGTATATACC
 TTCCGGCGTACCTTTGCCCTCCAGCGATGCCAGTGACCAAAGGCGATGCTGTATTCTTCAGC
 GACAGGGCCAGGAATCGCAAACCACGGTTTCAGTGGGGCAGGGGCCCTTCCGGCGATTCTTA
 CTAGCTAGTATGCATAGGTGCTGAAATATAAAGTTTGTGTTTCTAAAACACACGTGGTACGTA
 CGATAACGTACAGTGTTTTTCCCTCCACTTAAATCGAAGGGTAGTGTCTTGAGCGCGCGGAG
 TAAACATATATGGTTTCATATATGTCCGTAGGCACGTAAAAAAGCGAGGGATTTCGAATCCCC
 CGGAACCCCCGGTTGGGGCCCCACGCCTCGATCGAGCAAAAAAAAAAAAAAAAAAGAAAAAAAAA
 AAAAAAGCTTTCCCGCGGCCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCGTGTGAAA
 TTGTTATCCGCTCACAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGG
 TGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGG
 AAACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCCTAT
 TGGGCGCTCTTCCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTTCGGCTGCGGCGAGC
 GGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAA
 GAACATGTGAGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT
 TTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCG
 AAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC
 TGTTCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCT
 TTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTGTTTCGCTCCAAGCTGGGCTG
 TGTGCACGAACCCCCCGTTACGCCCCACCGCTGCGCCTTATCCGGTAACCTATCGTCTTGAGTC
 CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGC
 GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG
 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTC
 TTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTAC

Table 4. continued.

GCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTTGACGCTCAGTGG
AACGAAAACACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATC
CTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGAC
AGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCTGTTCCATCCATA
GTTGCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGT
GCTGCAATGATACCGCGAGACCCACGCTCACC GGCTCCAGATTTATCAGCAATAAACAGCCA
GCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAAT
TGTTGCCGGAAGCTAGAGTAAGTAGTTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATT
GCTACAGGCATCGTGGTGTACGCTCGTCTTGGTATGGCTTCATTTCAGCTCCGTTCCCAA
CGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCCT
CCGATCGTTGTGCAAGTAAGTTGGCCGAGTGTTATCACTCATGGTTATGGCAGCACTGCAT
AATTCTCTTACTGTTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG
TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGCGTCAATACGGGATAAT
ACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTGGAACCGTTCTTCGGGGCGAAAA
CTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGA
TCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCC
GCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTCAATAT
TATGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAA
ATAAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTCTAAGAAACCA
TTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCTGCTCTCGCGCGTT
TCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGCTGT
AAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTCTGGG
GCTGGCTTAATACTATGCGGCATCAGAGCAGATTGTAAGTGCAGCATATCGACGCTCTCC
CTTATGCGACTCCTGCATTAGGAAGCAGCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGC
CGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCTGCCAC
CATACCACGCGGAAACAAGCGCTCATGAGCCCGAAGTGCGGAGCCCGATCTTCCCCATCGGT
GATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCG
TCCGGCGTAGAGGATCTGGCTAGCGATGACCCTGCTGATTGGTTGCTGACCATTTCCGGGGT
GCGGAACGGCGTTACCAGAACTCAGAAGGTTTCGTCCAACCAAACCGACTCTGACGGCAGTTT
ACGAGAGAGATGATAGGGTCTGCTTCAGTAAGCCAGATGCTACACAATTAGGCTTGTACATAT
TGTCGTTAGAACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACG

[0168] The first and second nucleic acid molecule may be contacted under conditions causing recombination of the first and second nucleic acid molecules to produce a third nucleic acid molecule in which the sequence of interest and transcriptional regulatory sequence are operably linked. In Figure 7, the third nucleic acid molecule is indicated as an Expression Clone. Those skilled in the art will appreciate that the third nucleic acid molecule shown in Figure 7 is a plasmid resulting from the recombination of the *attL1* and *attL2* sites in the first nucleic acid molecule and the *attR1* and *attR2* sites in the second nucleic acid molecule. It is not necessary that recombination occur in both sets of sites. In the configuration shown in Figure 7, as long as the *attL1* site and the *attR1* site recombine, a third nucleic acid molecule having a transcriptional regulatory sequence operably linked to a sequence of interest will be produced. Thus other second nucleic acid molecules having only a single recombination site may be used in the practice of the invention. Although both the first and second nucleic acid molecules are shown as plasmids in Figure 7, either or both may not be a plasmid (*e.g.*, may be a PCR fragment, restriction fragment, synthetic molecule, etc.).

[0169] In some embodiments, it may be desirable to use a linear nucleic acid molecule as a second nucleic acid molecule. For example, when the recombination reaction mixture is to be directly amplified, linearization of the second nucleic acid molecule may reduce background in the amplification reaction. With reference to Figure 7, the destination vector pEU3'His (for construction of an expression vector for use in a wheat germ cell-free protein synthesis system) was digested at the restriction enzyme sites indicated (*Bgl*III and *Pst*I) in the insertion region resulting in the generation of a linear nucleic acid molecule. After a recombination reaction with a first nucleic acid molecule comprising a sequence of interest, a third nucleic acid molecule (*e.g.*, the expression clone of Figure 7) is created. The third nucleic acid molecule may be amplified, for example, using primers that flank the recombination region of the second nucleic acid molecule. As a result of the linearization of the second nucleic acid molecule, only expression clones into which the ORF of the target genes has been inserted through the GATEWAY™ LR reaction function as a PCR template.

[0170] A GATEWAY™ LR reaction with GATEWAY™-adapted entry clones and a destination vector adapted for IVTT (*e.g.*, for the wheat germ cell-free protein synthesis system) was conducted, resulting in the construction of expression clones into which the targeted gene ORF was inserted through site-specific recombination reactions between *attL1* and *attR1*, and *attL2* and *attR2*. The GATEWAY™ LR reaction was performed in a 5 µl reaction volume using the restriction digested destination vector pEU3'His reaction and various entry clones for 3 hours. Entry clones are pDONR201-based plasmids (Invitrogen Corporation, Carlsbad, CA) harboring ORFs of EGFP and human full-length cDNA. Methods of the invention were tested on about 4000 clones. The sequences of some of the clones tested are provided in Tables 5-8.

Table 5. Sequence of EGFP clone.

ATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTTCGAGCTGGACGGC
GACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAG
CTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACC
ACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTTC
TTCAAGTCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGC
AACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTG
AAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACATAAC
AGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTTCAAGATC
CGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATC
GGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAA
GACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCGGGATCACT
CTCGGCATGGACGAGCTGTACAAGTAT

Table 6. Sequence of clone FLJ20768.

ATGTTCCGGTGTGGAGGCCTGGCGGCGGGTGCTTTGAAGCAGAAGCTGGTGCCCTTGGTGCGG
ACCGTGTGCGTCCGAAGCCGAGGCAGAGGAACCGGCTCCAGGCAACTTGTTCCAGCGATGG
CATGTTCTCTAGAACTCCAGATGACAAGACAAATGGCTAGCTCTGGTGCATCAGGGGGCAAA
ATCGATAATTCTGTGTTAGTCCTTATTGTGGGCTTATCAACAGTAGGAGCTGGTGCCATATGCC
TACAAGACTATGAAAGAGGACGAAAAAGATACAATGAAAGAATTTCAAGGTTAGGGCTGACA
CCAGAACAGAAACAGAAAAAGGCCGCTTATCTGCTTCAGAAGGAGAGGAAGTTCCTCAAGAC
AAGGCGCCAAGTCATGTTCTTTCTGCTAATTGGTGGAGGCACAGCTGCTTTTGCTGCAGCC
AGATCCATCCGGGCTCGGGATCCTGGGGCCAGGGTACTGATTGTATCTGAAGATCCTGAGCTG
CCGTACATCGCAGCTCCTCTTTCAAAGAAGTGTGGTTTTCAGATGACCCAAATGTCACAAAG
ACATCGCGATTCAAACAGTGGAATGGAAAAGAGAGAAGCATATATTTCCAGCCACCTTCTTTT
TATGTCTCTGCTCAGGACCTGCCTCATATTGAGAATGGTGGTGTGGCTGTCTCACTGGGAAG
AAGGTAGTACAGCTGGATGTGAGAGACAACATGGTGAACTTAATGATGGCTCTCAAATAACC
TATGAAAAGTGCTTGATTGCAACAGGAGGTACTCCAAGAAGTCTGTCTGCCATTGATAGGGCT
GGAGCAGAGGTGAAGAGTAGAACAACGCTTTTCAGAAAGATTGGAGACTTTAGAAGCTTGGAG
AAGATTTACGGGAAGTCAAATCAATTACGATTATCGGTGGGGGCTTCCTTGGTAGCGAACTG
GCCTGTGCTCTTGGCAGAAAGGCTCGAGCCTTGGGCACAGAAGTGATTCAACTCTTCCCCGAG
AAAGGAAATATGGGAAAGATCCTCCCCGAATACCTCAGCAACTGGACCATGGAAAAAGTCAGA
CGAGAGGGGGTTAAGGTGATGCCCAATGCTATTGTGCAATCCGTTGGAGTCAGCAGTGGCAAG
TTACTTATCAAGCTGAAAGACGGCAGGAAGGTAGAACTGACCACATAGTGGCAGCTGTGGGC
CTGGAGCCCAATGTTGAGTTGGCCAAGACTGGTGGCCTGGAAATAGACTCAGATTTTGGTGGC
TTCCGGGTAAATGCAGAGCTACAAGCACGCTCTAACATCTGGGTGGCAGGAGATGCTGCATGC
TTCTACGATATAAAGTTGGGAAGGAGGCGGGTAGAGCACCATGATCACGCTGTTGTGAGTGGA
AGATTGGCTGGAGAAAAATGACTGGAGCTGCTAAGCCGTACTGGCATCAGTCAATGTTCTGG
AGTGATTTGGGCCCCGATGTTGGCTATGAAGCTATTGGTCTTGTGGACAGTAGTTTGGCCACA
GTTGGTGTTTTTGCAAAAGCAACTGCACAAGACAACCCCAAATCTGCCACAGAGCAGTCAGGA
ACTGGTATCCGATCAGAGAGTGAGACAGAGTCCGAGGCCCTCAGAAATTACTATTCCTCCCAGC
ACCCCGGCAGTTCCACAGGCTCCCGTCCAGGGGGAGGACTACGGCAAAGGTGTCATCTTCTAC
CTCAGGGACAAAGTGGTCGTGGGGATTGTGCTATGGAACATCTTTAACCGAATGCCAATAGCA
AGGAAGATCATTAAGGACGGTGAGCAGCATGAAGATCTCAATGAAGTAGCCAACTATTCAAC
ATTCATGAAGACTGT

Table 7. Sequence of clone FLJ20819.

ATGACCTCTAGGAAGAAAGTGTGCTGAAGGTTATCATCCTGGGAGATTCTGGAGTCGGGAAG
ACATCACTCATGAACCAGTATGTGAATAAGAAATTCAGCAATCAGTACAAAGCCACAATAGGA
GCTGACTTTCTGACCAAGGAGGTGATGGTGGATGACAGGCTAGTCACAATGCAGATATGGGAC
ACAGCAGGACAGGAACGGTTCAGTCTCTCGGTGTGGCCTTCTACAGAGGTGCAGACTGCTGC
GTTCTGGTATTTGATGTGACTGCCCCAACACATTCAAAACCTAGATAGCTGGAGAGATGAG
TTTCTCATCCAGGCCAGTCCCCGAGATCCTGAAAACCTCCCATTTGTTGTGTGGGAAACAAG
ATTGACCTCGAAAACAGACAAGTGGCCACAAAGCGGGCACAGGCCTGGTGCTACAGCAAAAAC
AACATTCCTTACTTTGAGACCAGTGCCAAGGAGGCCATCAACGTGGAGCAGGCGTTCCAGACG
ATTGCACGGAATGCACTTAAGCAGGAAACGGAGGTGGAGCTGTACAACGAATTTCTGAACTT
ATCAAACTGGACAAGAATGACCGGGCCAAGGCCTCGGCAGAAAGCTGCAGTTGCTGT

Table 8. Sequence of clone FLJ20855.

ATGGTTCATCAGGTGCTCTACCGGGCGCTGGTCTCCACCAAGTGGCTGGCGGAGTCCATCAGG
ACTGGCAAGCTGGGGCCCGGCCTGCGGGTGCTGGACGCGTCCTGGTACTCACCAGGCACCCGA
GAGGCCCGCAAGGAGTACCTCGAGCGCCACGTACCCGGCGCCTCTTTCTTTGACATAGAAGAG
TGCCGGGACACGGCGTCGCCCTACGAGATGATGCTGCCCAGCGAGGCTGGCTTCGCCGAGTAT
GTGGGCCCGCCTGGGCATCAGCAACCACACGCACGTGGTGGTGTATGATGGTGAACACCTGGGC
AGCTTCTATGCTCCCCGGGTCTGGTGGATGTTCCGTGTGTTTGGCCACCGCACCGTATCAGTG
CTCAATGGTGGCTTCCGGAACCTGGCTGAAGGAGGGCCACCCGGTGACATCCGAGCCCTCACGC
CCAGAACCGGCCGTCTTCAAAGCCACACTGGACCGTCCCTGCTCAAGACCTACGAGCAGGTG
CTGGAGAACCTTGAATCTAAGAGGTTCCAGCTGGTGGATTCAAGGTCTCAAGGGCGGTTCCCTG
GGCACCGAGCCGGAGCCGGATGCAGTAGGACTGGACTCGGGCCATATCCGTGGTGCCGTCAAC
ATGCCTTTTCATGGACTTCTGACTGAGGATGGCTTCGAGAAGGGCCCAGAAGAGCTCCGTGCT
CTGTTCCAGACCAAGAAGGTGGATCTCTCGCAGCCTCTCATTGCCACGTGCCGCAAGGGAGTC
ACCGCCTGCCACGAGGCCTTGGCTGCCTACCTCTGCGGCAAGCCTGATGTGGCCGTGTACGAT
GGCTCCTGGTCCGAGTGGTTTTGCGCGGGCCCCCCCAGAGAGCCGTGTGTCCAGGGAAAGTCT
GAGAAGGCCTGT

[0171] As discussed above, after a third nucleic acid molecule is constructed it may be used as a template in further reactions (*e.g.*, amplification reactions, transcription reactions, etc.). For example, an amplification reaction may be conducted using the reaction product of the LR reaction described above as a template. The amplification may be conducted in a 50 μ l reaction mixture using KOD Dash as a polymerase and 1 μ l of the LR reaction as a template and the following primers

Forward primer : 5'-ACACATACGATTAGGTGACACT-3'

Reverse primer: 5'-GTCCACCCCGTAGAAAAGA-3'

Primers were selected to amplify the target gene transfer regions containing the SP6 promoter region and the amplification reactions were conducted in a high throughput format (*e.g.*, in a 96 well format).

[0172] An amplified third nucleic acid molecule may be used as a template for *in vitro* transcription. The transcription reaction may be performed in a 50 μ l reaction mixture using an appropriate polymerase. The transcription reaction was conducted using SP6 polymerase at 37 °C for 4 hours. A typical mRNA synthesis reaction may contain:

5 x TB Buffer	10.0 μ l
25 mM NTPs MIX	5.0 μ l
RNase inhibitor (40 U/ μ l, TOYOBO)	1.0 μ l
SP6 RNA polymerase (50 U/ μ l, TAKARA, catalog no. 2520A)	3.0 μ l
PCR product	5.0 μ l
water	26.0 μ l
Total	50.0 μ l

[0173] The transcription reaction may be conducted in a high throughput format, *e.g.*, in a 96-well plate format.

[0174] An RNA molecule corresponding to the third nucleic acid molecule, may be used to direct *in vitro* translation to produce a polypeptide encoded by a sequence of interest present on the first nucleic acid molecule. Any suitable *in vitro* translation protocol may be employed to translate the RNA into

protein. One suitable method is a bilayer method (Figure 8). For example, an upper layer (*e.g.*, of about 125 μ l volume) containing a suitable substrate mixture, which may include, for example, creatin phosphate, GTP, ATP, and/or amino acids) may be added to a lower layer (*e.g.*, of about 25 μ l volume) comprising an extract suitable for *in vitro* translation and RNA corresponding to the sequence of interest encoding a polypeptide. One suitable extract for *in vitro* translation is PROTEIOSTM (TOYOBO) wheat germ extract. The RNA may be purified from the *in vitro* transcription reaction, for example, by ethanol precipitation, prior to being added to the extract. Typically, approximately 3 μ l out of a 20 μ l volume *in vitro* transcription reaction is added. The amount of RNA added can be varied as is known to those skilled in the art. The layers may be incubated at a suitable temperature, for example, at 26 °C, for a suitable period of time, for example, 24 hours.

[0175] Optionally the transcription and translation reactions may be coupled, *i.e.*, a PCR product may be added to a mixture containing a suitable RNA polymerase and the requisite materials for translation (*e.g.*, a cell extract and suitable substrate mixture).

[0176] In another embodiment, the reactions of methods of the invention may be performed sequentially, for example, in the same reaction vessel (*e.g.*, microfuge tube, well of a microtiter plate, etc.). For example, a recombination reaction may be conducted in a small volume (*e.g.*, 5 μ l). The reaction may be stopped, for example, by heating to inactivate the recombination proteins, and then the material for an amplification reaction may be added directly to the recombination reaction. For example, buffers, primers, nucleotides, and a DNA polymerase may be added up to a second volume (*e.g.*, 25 μ l). The amplification reaction may be conducted and then materials for an *in vitro* transcription reaction may be added (*e.g.*, RNA polymerase and ribonucleotides) up to a third volume (*e.g.*, 50 μ l). After the transcription reaction is conducted, a materials for an *in vitro* translation reaction (*e.g.*, a cell extract and a substrate mixture) may be added up to a fourth volume.

[0177] After incubation, the products of the translation reaction may be analyzed and/or recovered using any suitable technology, for example, by

SDS-PAGE, column chromatography, etc. For example, the translation reaction mixture may be clarified by centrifugation. The mixture may be centrifuged at 14,500 rpm (19,000xg) for 20 min at 4 °C and the supernatant may be recovered.

[0178] To analyze the translation reaction, aliquots of the total reaction and the supernatant were analyzed by SDS-PAGE. Each aliquot was mixed with a 4×LDS Sample Buffer (e.g., NuPAGE LDS sample buffer Invitrogen Corporation, Carlsbad, CA, Cat. No. NP0007), 2-mercaptoethanol, and heat-treated for 10 minutes at 70 °C. A 10.9 µL aliquot was mixed with 4.1 µL of loading buffer to make a 15 µL sample and the sample was loaded on the gel.

[0179] The samples were applied to an SDS-polyacrylamide gel and separated using standard techniques. The proteins were transferred to a membrane and analyzed using antibodies to a 6-histidine tag encoded by the second nucleic acid molecule (*i.e.*, the destination vector pEU3'His in Figure 7). The primary antibody was an anti-6-His antibody (His-probe (G-18), Santa Cruz Biotechnology, Inc., Santa Cruz, CA), used at 1/1000 dilution in PBST-3% BSA (phosphate buffered saline containing Tween-20 at 0.1% (v/v)) The secondary antibody was an anti-rabbit Ig, HRP-Linked Whole Ab, used at 1/2000 dilution in PBST-10% milk (Amersham Biosciences Anti-rabbit Ig, HRP-linked whole antibody (Donkey)cat. no. NA934). Detection was performed the detection reaction with ECL-Plus (Amersham Biosciences ECL PLUS Western blotting detection reagent, cat. no. RPN2132), and detected with BIO-RAD Fluor-S MAX according to the manufacturers instructions.

[0180] Using the above-described protocol, a recombination reaction was used to create a third nucleic acid molecule and the third nucleic acid molecule was directly used as a PCR template. The amplified third nucleic acid molecule was then used for an *in vitro* transcription reaction and the RNA produced was used in an *in vitro* translation reaction. Thus, all procedures necessary to produce a polypeptide from a sequence of interest were performed *in vitro*. In addition, by not introducing the nucleic acid molecules into a host cell at any point, it is possible to express polypeptides that adversely impact the growth of the host cell. Further, by not introducing the nucleic acid molecule into a host cell, it is not necessary to purify the desired nucleic acid molecules from

host cells prior to subsequent manipulations. Thus, the problem of RNase contamination from the host cell was solved. An additional benefit provided by the present invention is the reduction in the time required to produce a polypeptide from a sequence of interest. Using the above-described methods, a polypeptide may be produced much more rapidly than was previously possible. As shown in Figure 6, the present methods can substantially reduce the time required to produce proteins. The above-described methods are particularly well suited to the parallel processing of a large number of sequences of interest, *i.e.*, high throughput processing.

[0181] Having now fully described the present invention in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious to one of ordinary skill in the art that the same can be performed by modifying or changing the invention within a wide and equivalent range of conditions, formulations and other parameters without affecting the scope of the invention or any specific embodiment thereof, and that such modifications or changes are intended to be encompassed within the scope of the appended claims.

[0182] All publications, patents and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains, and are herein incorporated by reference to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference.

WHAT IS CLAIMED IS:

1. A method of producing a polypeptide, comprising:
contacting a first nucleic acid molecule comprising a sequence of interest encoding the polypeptide and comprising one or more recombination sites with a second nucleic acid molecule comprising at least one transcriptional regulatory sequence and one or more recombination sites under conditions causing recombination of the first and second nucleic acid molecules to produce a third nucleic acid molecule in which the sequence of interest and transcriptional regulatory sequence are operably linked;
transcribing the sequence of interest on the third nucleic acid molecule into RNA; and
translating the RNA into a polypeptide, wherein transcribing and translating are done *in vitro*.
2. The method according to claim 1, wherein the third nucleic acid molecule is not introduced into a cell prior to transcription.
3. The method according to claim 1, wherein the second nucleic acid molecule is an expression vector.
4. The method according to claim 3, wherein the expression vector is linear.
5. The method according to claim 3, wherein the expression vector is cleaved with a restriction enzyme.
6. The method according to claim 5, wherein the restriction enzyme is selected from a group consisting of *Bgl*II and *Pst*I.
7. The method according to claim 1, wherein the transcriptional regulatory sequence is a promoter.
8. The method according to claim 7, wherein the promoter is a phage promoter.

9. The method according to claim 7, wherein the promoter is selected from a group consisting of a T7 promoter and an SP6 promoter.

10. The method according to claim 1, further comprising amplifying all or a portion of the third nucleic acid molecule.

11. A method of making a desired nucleic acid molecule, comprising:

contacting a first nucleic acid molecule comprising a sequence of interest encoding a polypeptide and comprising one or more recombination sites with a second nucleic acid molecule comprising at least one transcriptional regulatory sequence and one or more recombination sites under conditions causing recombination of the first and second nucleic acid molecules to produce a desired nucleic acid molecule in which the sequence of interest and transcriptional regulatory sequence are operably linked; and

amplifying the desired nucleic acid molecule, wherein the desired nucleic acid molecule is not introduced into a host cell before amplifying.

12. The method according to claim 11, wherein the second nucleic acid molecule is an expression vector.

13. The method according to claim 12, wherein the expression vector is linear.

14. The method according to claim 12, wherein the expression vector is cleaved with a restriction enzyme.

15. The method according to claim 14, wherein the restriction enzyme is selected from a group consisting of *Bgl*II and *Pst*I.

16. The method according to claim 11, wherein the transcriptional regulatory sequence is a promoter.

17. The method according to claim 16, wherein the promoter is a phage promoter.

18. The method according to claim 16, wherein the promoter is selected from a group consisting of a T7 promoter and an SP6 promoter.

19. A kit for producing a polypeptide, comprising one or more containers containing one or more recombination proteins, one or more containers contain one or more restriction enzymes, and one or more containers comprising an extract for *in vitro* translation.

20. The kit according to claim 19, further comprising one or more containers containing one or more components selected from the group consisting of polymerases, buffers, primers, vectors, nucleic acid molecules comprising promoter sequences, and nucleotides.

21. A method for providing a product suitable for the practice of a method according to any one of claims 1-18 to a customer, comprising:
receiving an order for a product from a customer; and
providing the customer with the product.

22. The method of claim 21, wherein the product is a kit.

23. The method of claim 22, wherein the kit comprises one or more containers comprising one or more recombination proteins and one or more containers comprising one or more cell extracts.

24. The method according to claim 23, wherein the kit further comprises written instructions.

25. The method according to claim 21, wherein the order is received by telephone.

26. The method according to claim 21, wherein the order is received via the internet.

27. A method for providing information as to availability of a product suitable for the practice of a method according to any one of claims 1-18 to a customer, comprising:

transmitting product description data to an output source;

transmitting to the output source instructions to publish the product description in media accessible to potential interested parties; and

detecting publication of the data in the media, thereby providing information as to availability of the product to parties having potential interest in the availability of the product.

28. The method of claim 27, wherein the product is a kit.

29. The method of claim 28, wherein the kit comprises one or more containers comprising one or more recombination proteins and one or more containers comprising one or more cell extracts.

30. The method according to claim 29, wherein the kit further comprises written instructions.

31. The method according to claim 27, wherein the product description data is published in a print medium.

32. The method according to claim 21, wherein the product description data is published on the internet.

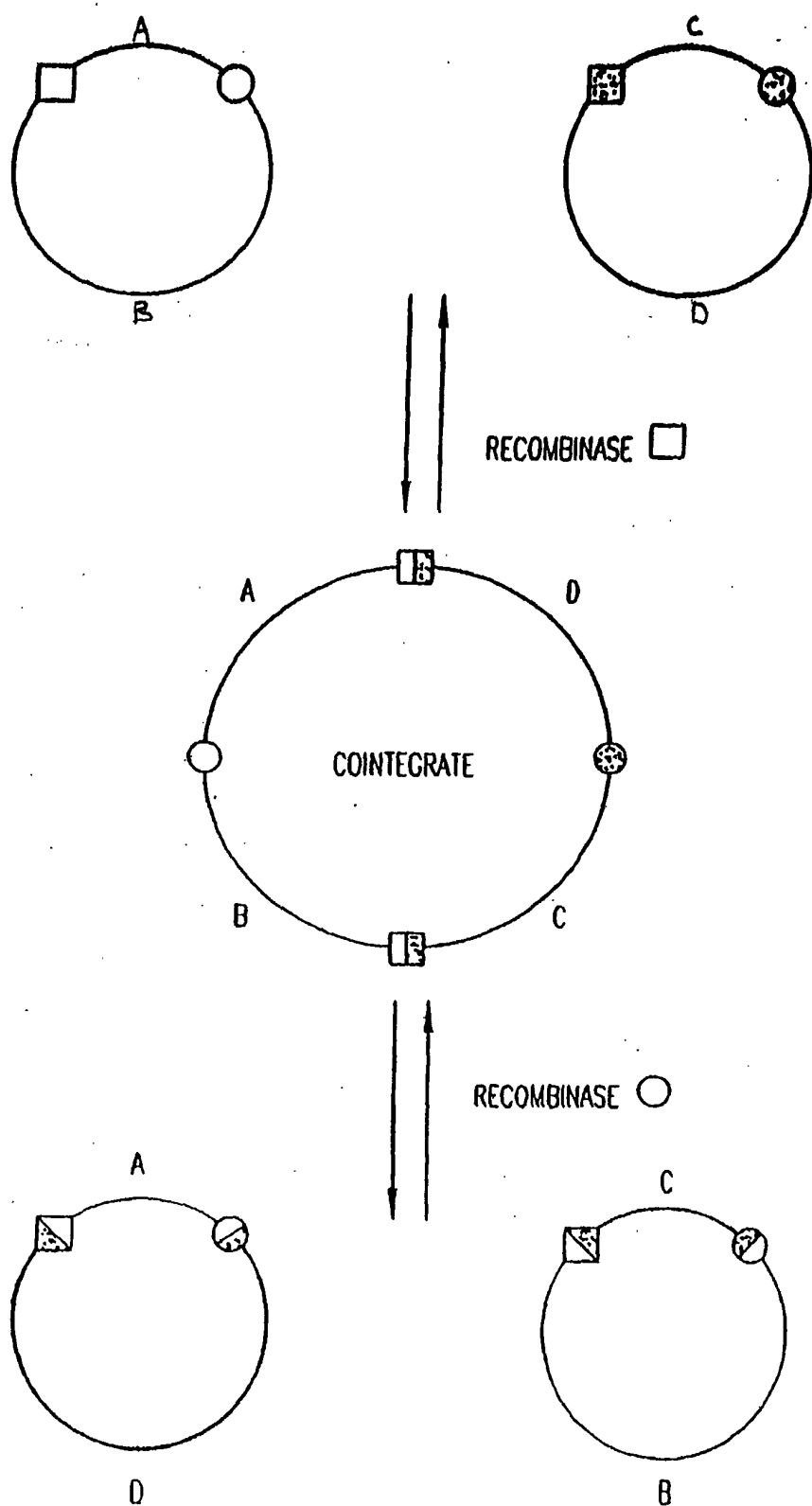


FIG. 1

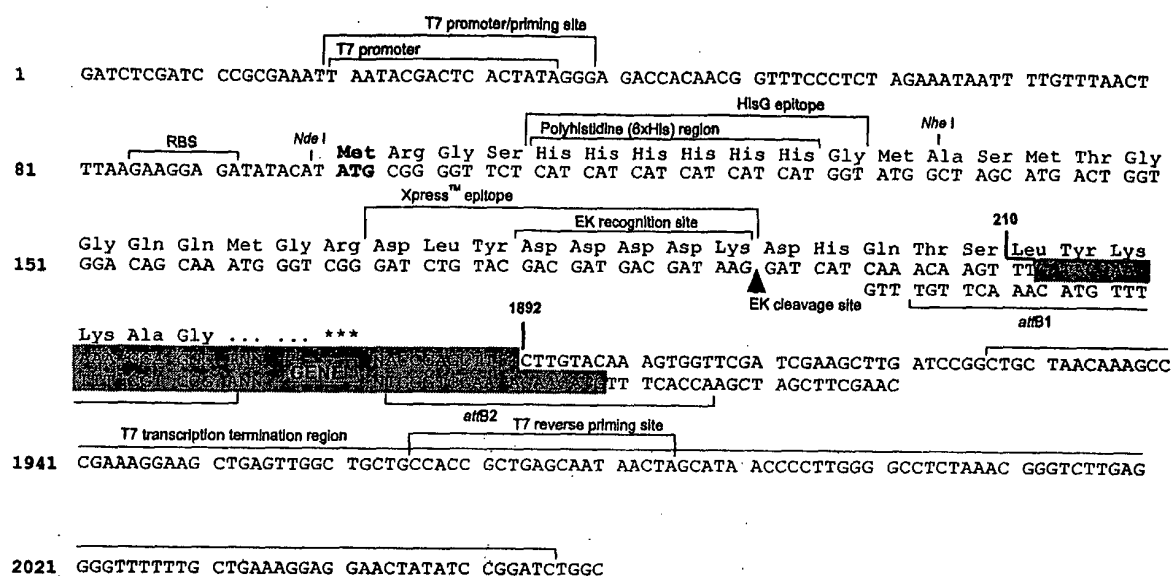
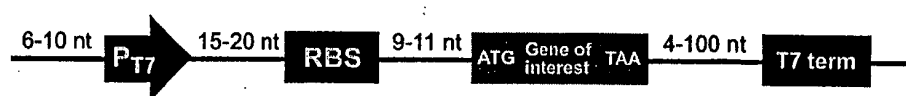


FIG. 2

**FIG. 3**

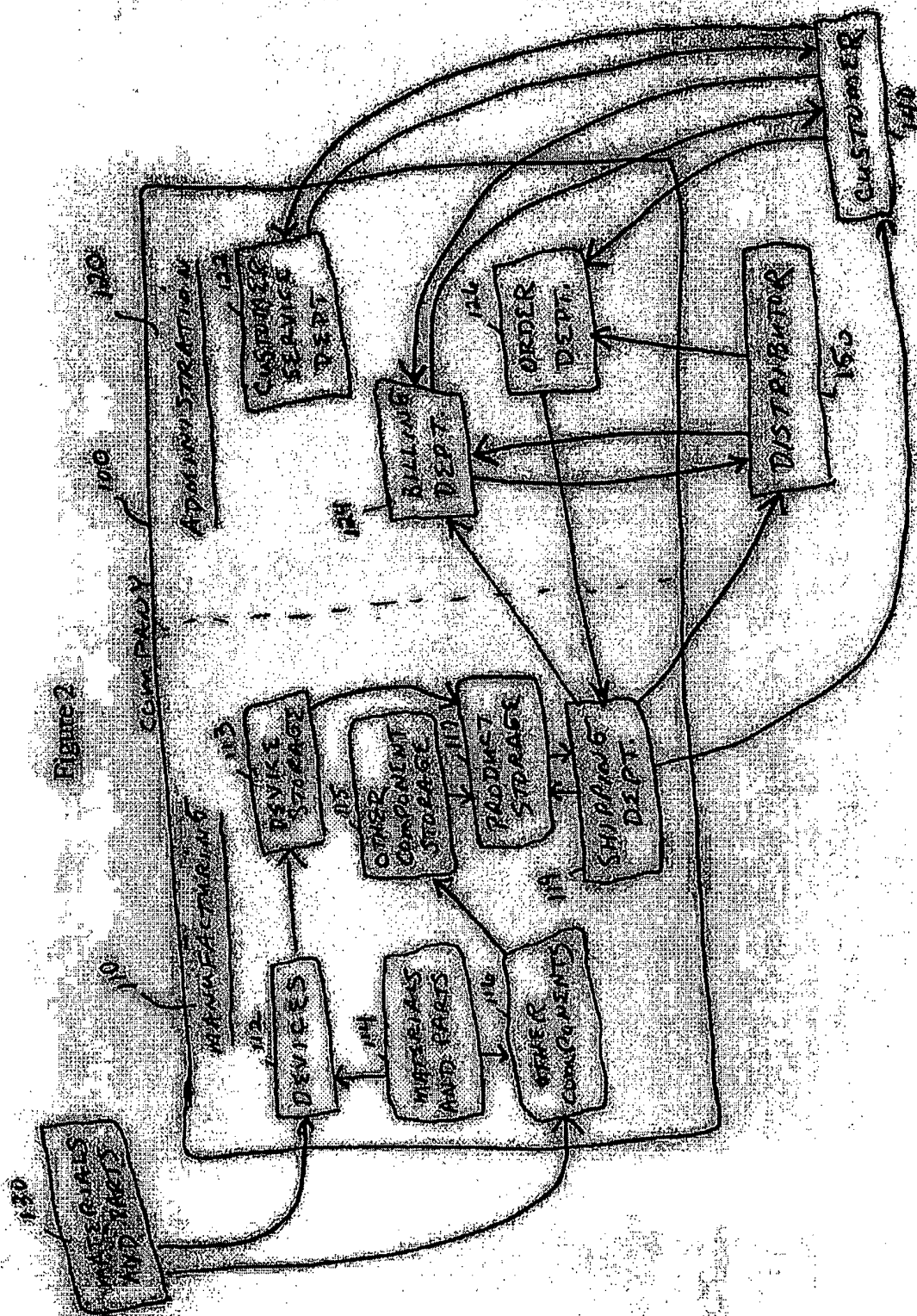


FIG. 4

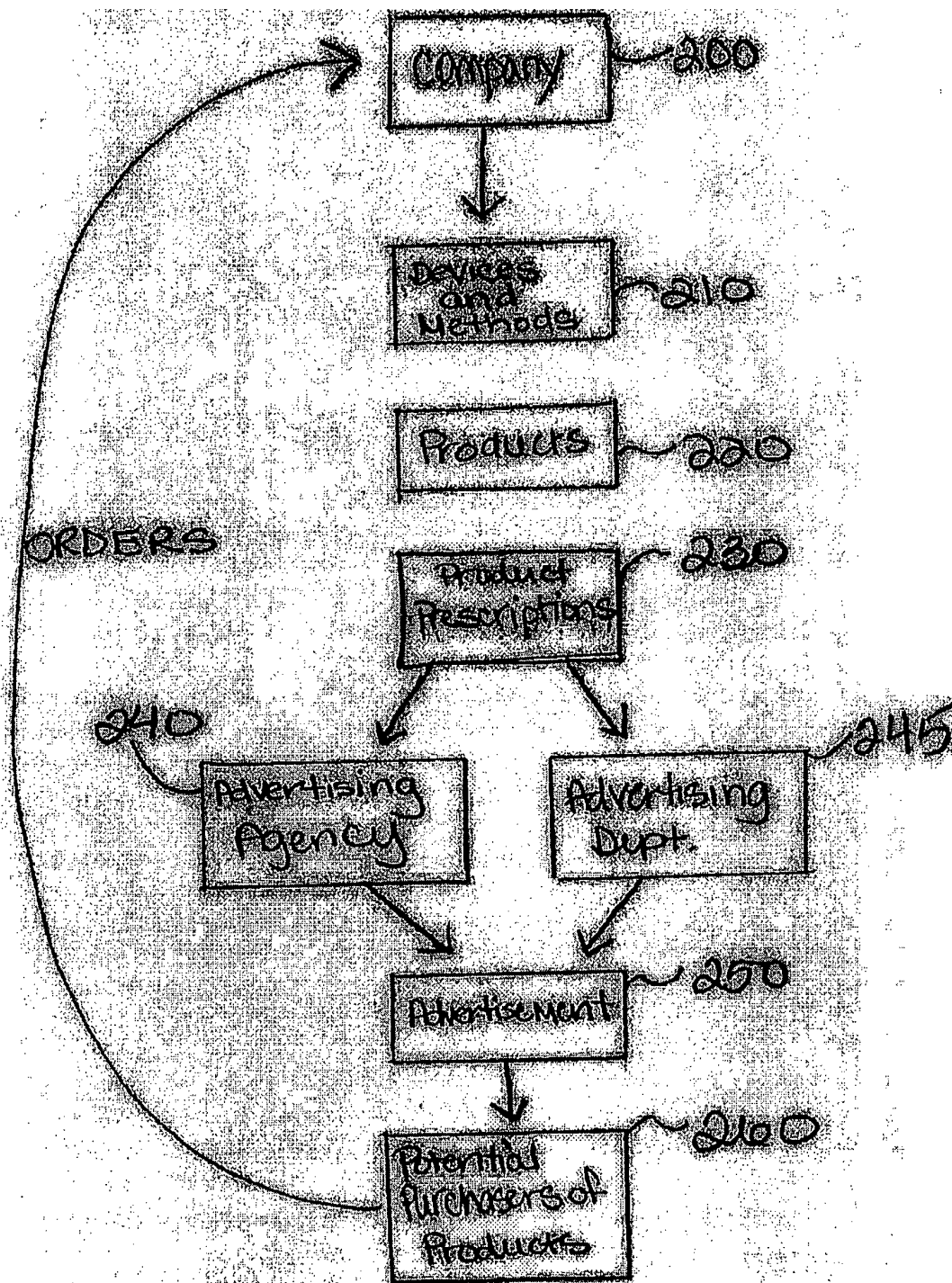
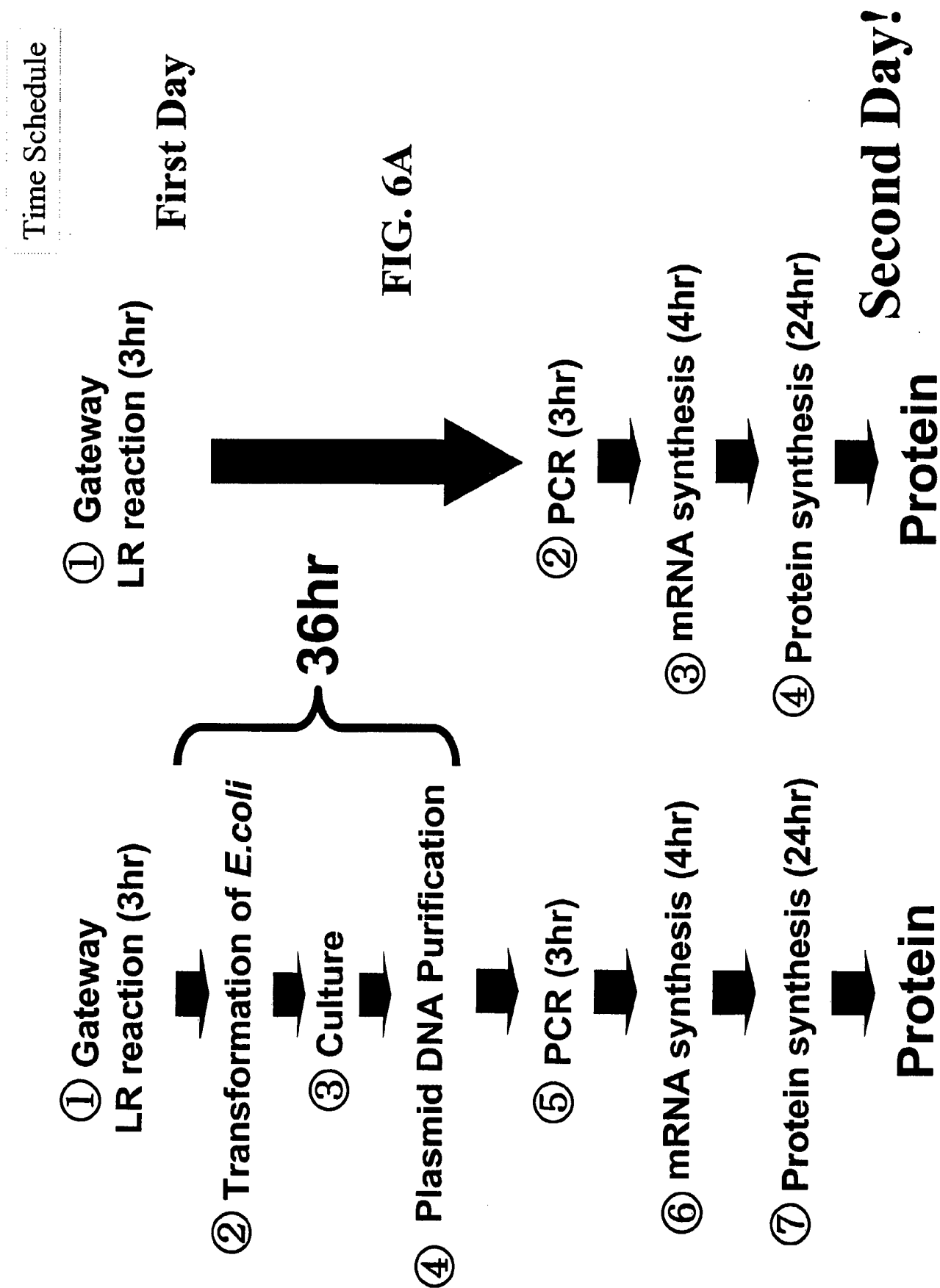


FIG. 5



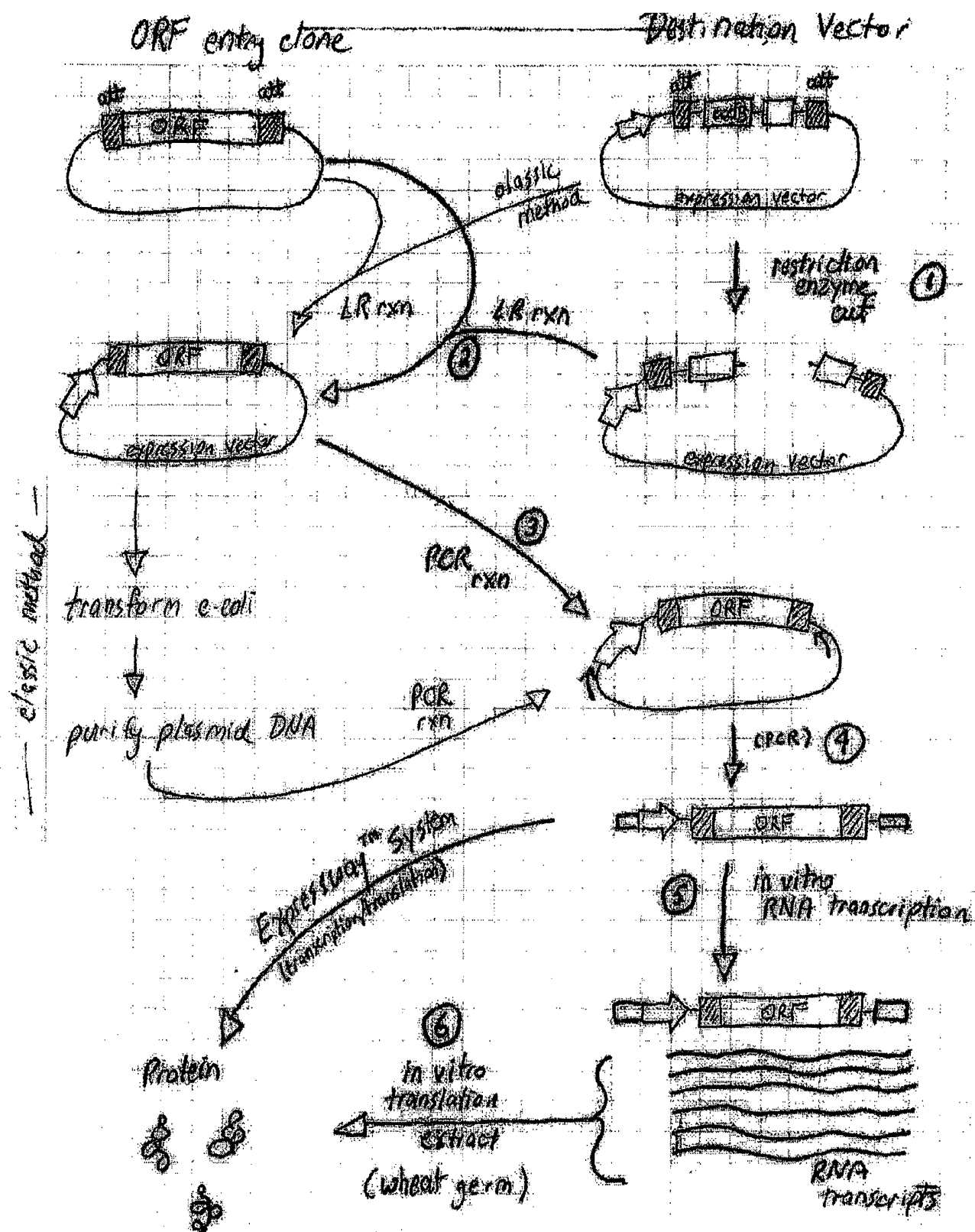
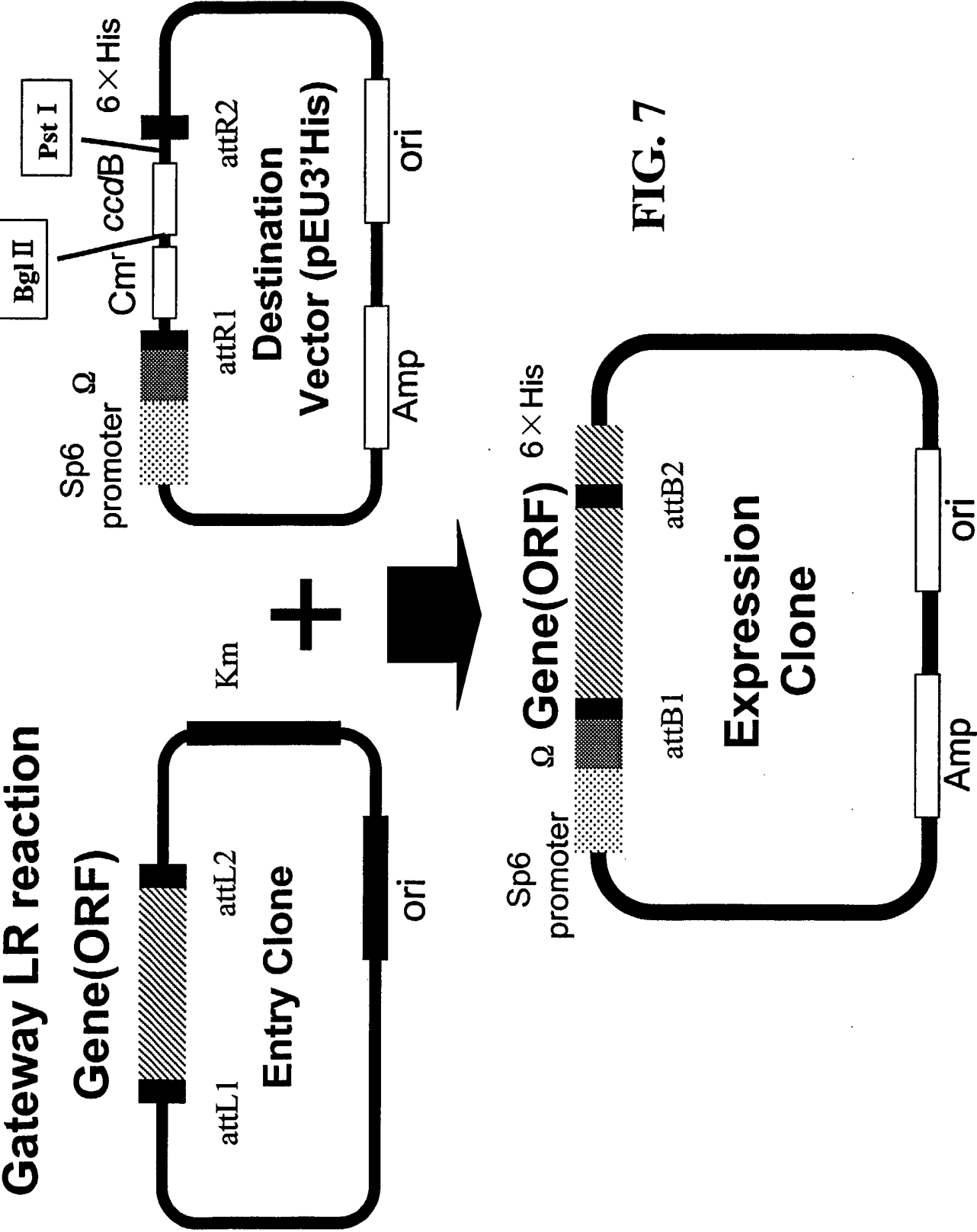
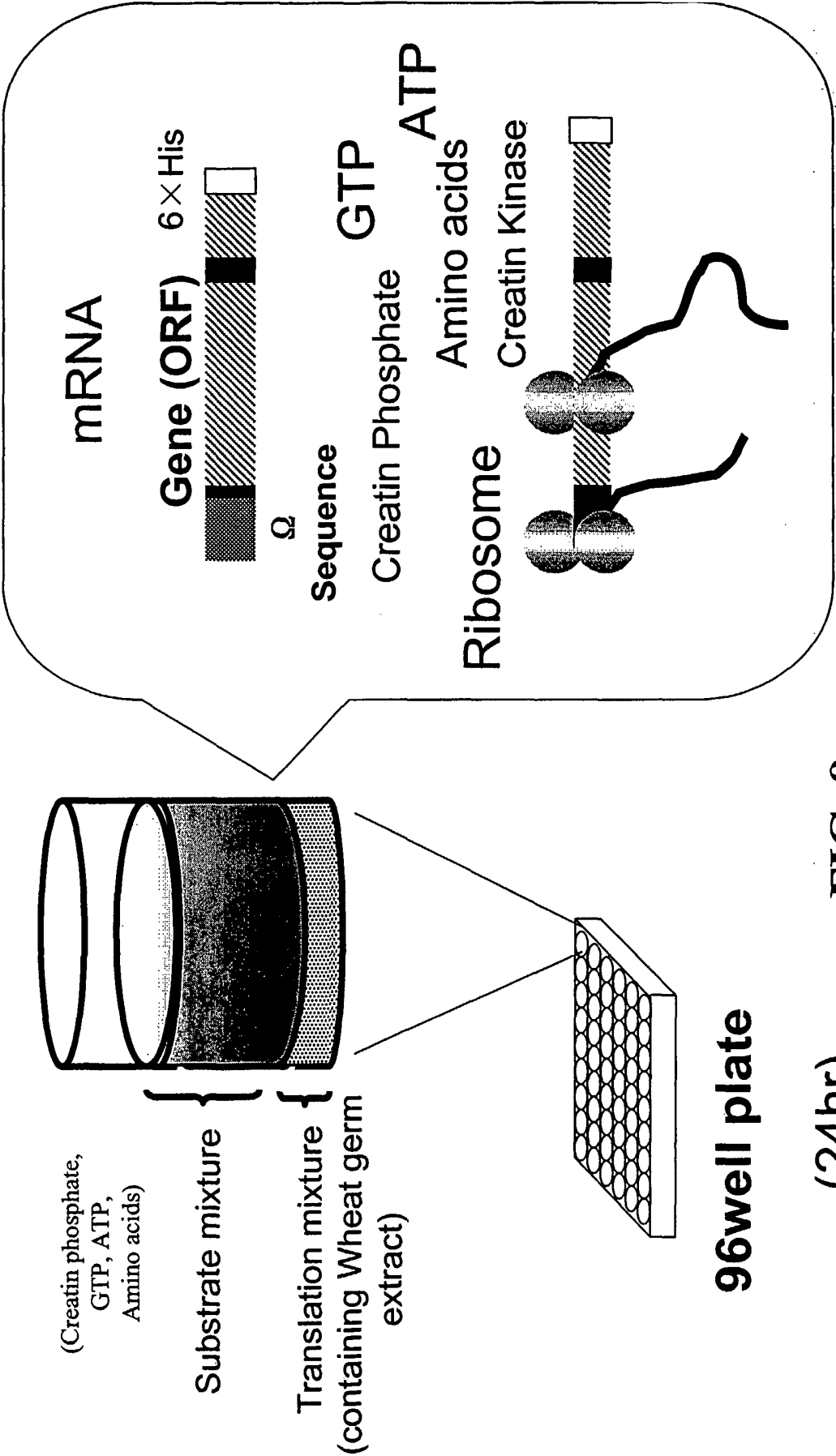


FIG. 6B



Wheat germ cell free
protein synthesis



96well plate
(24hr) **FIG. 8**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US03/18219

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12N 15/00, 15/09; C07H 21/04

US CL : 435/6, 29, 91.4, 320.1, 69.1; 536/23.1

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, 29, 91.4, 320.1, 69.1; 536/23.1

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)
Please See Continuation Sheet**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 6,277,608 B1 (HARTLEY et al) 21 August 2001 (21.08.01), see the entire document, especially Figure 1, columns 4-10.	1-20 ----- 21-32
X --- Y	US 5,888,732 A (HARTLEY et al) 30 March 1999 (30.03.1999), see the entire patent, especially Figure 1, columns 4-5.	1-20 ----- 21-32

☐ 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	"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
"E" earlier application or patent published on or after the international filing date	"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
"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)	"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
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

28 September 2003 (28.09.2003)

Date of mailing of the international search report

17 NOV 2003

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Facsimile No. (703)305-3230

Authorized officer

Gerald G Leffers Jr., PhD
Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

PCT/US03/18219

Continuation of B. FIELDS SEARCHED Item 3:

EAST: JPO, EPO, USPAT, US PGPUBS, DERWENT; STN: MEDLINE, EMBASE, BIOSIS, CAPLUS
search terms: inventors' names, recombinational cloning, site-specific recombinase, recombination, cloning