



(86) Date de dépôt PCT/PCT Filing Date: 2002/06/26

(87) Date publication PCT/PCT Publication Date: 2003/01/09

(85) Entrée phase nationale/National Entry: 2003/12/01

(86) N° demande PCT/PCT Application No.: EP 2002/007060

(87) N° publication PCT/PCT Publication No.: 2003/002736

(30) Priorité/Priority: 2001/06/27 (01115424.2) EP

(51) Cl.Int.⁷/Int.Cl.⁷ C12N 9/16, C12N 15/10, C12N 9/06

(71) Demandeur/Applicant:

F. HOFFMANN-LA ROCHE AG, CH

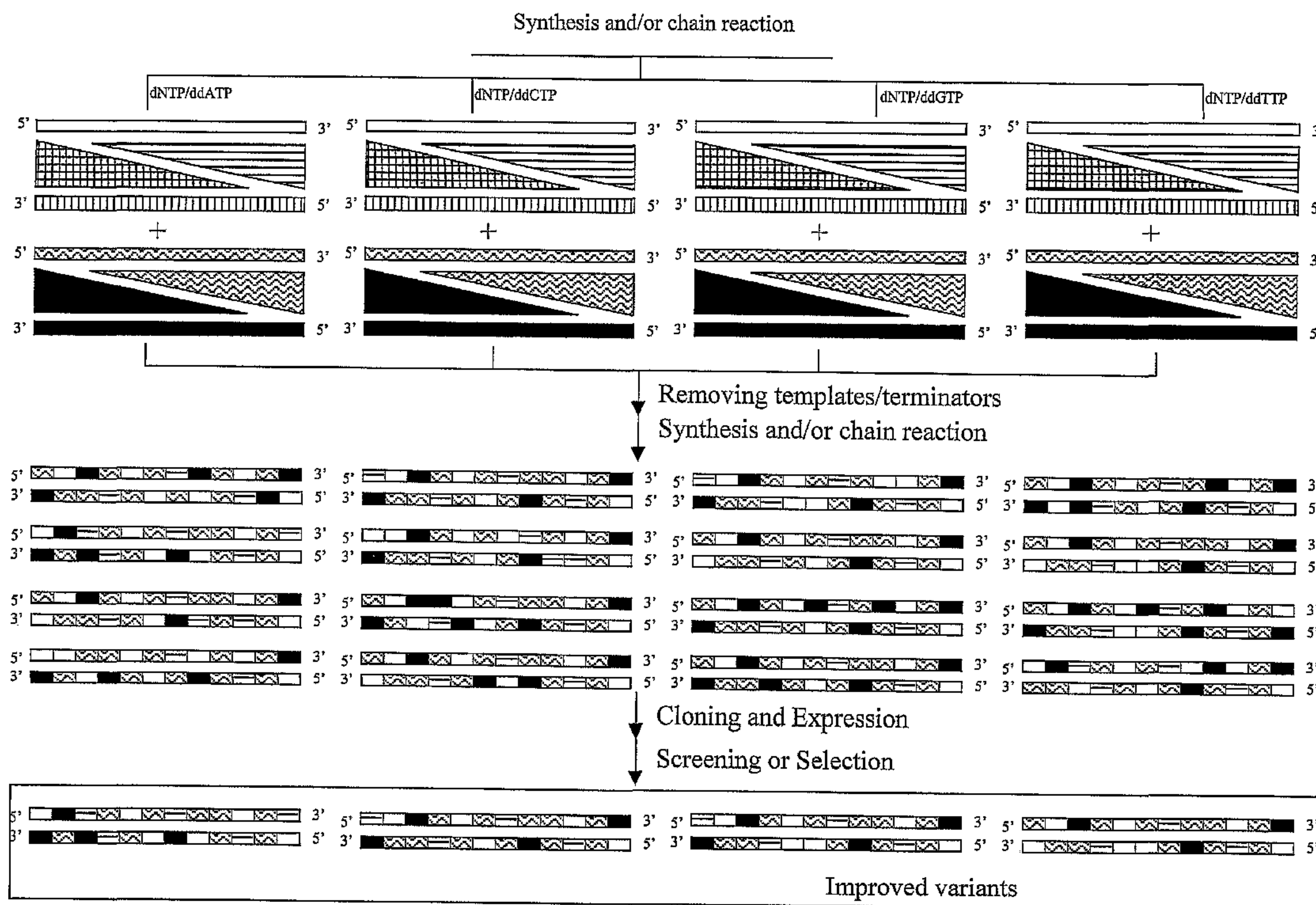
(72) Inventeurs/Inventors:

SHAO, ZHIXIN, DE;
KRATZSCH, PETER, DE;
SCHMUCK, RAINER, DE;
VON DER ELTZ, HERBERT, DE;
KENKLIES, JANET, DE

(74) Agent: BORDEN LADNER GERVAIS LLP

(54) Titre : TECHNIQUE PAR REVISION STRUCTUREE POUR LA RECOMBINAISON IN VITRO DE SEQUENCES
POLYNUCLEOTIDIQUES

(54) Title: A WALK-THROUGH TECHNIQUE FOR IN VITRO RECOMBINATION OF POLYNUCLEOTIDE SEQUENCES



(57) **Abrégé/Abstract:**

A simple and efficient method for random in vitro mutagenesis and recombination of polynucleotide sequences based on walk-through with chain-elongating molecules and chain terminating molecules e.g. dNTPs/ddNTPs followed by reassembly and amplification is described. The utility of this method for improving protein structure and/or function is demonstrated by creating novel sarcosine oxidase variants and by recombining human placental alkaline phosphatases and calf intestinal alkaline phosphatase for better mutants with improved properties.

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
9 January 2003 (09.01.2003)

PCT

(10) International Publication Number
WO 03/002736 A3(51) International Patent Classification⁷: C12N 9/16,
9/06, 15/10

(21) International Application Number: PCT/EP02/07060

(22) International Filing Date: 26 June 2002 (26.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
01115424.2 27 June 2001 (27.06.2001) EP(71) Applicant (for DE only): ROCHE DIAGNOSTICS
GMBH [DE/DE]; Sandhofer Strasse 116, 68305
Mannheim (DE).(71) Applicant (for all designated States except DE, US): F.
HOFFMANN-LA ROCHE AG [CH/CH]; Grenzacher-
strasse 124, CH-4070 Basel (CH).

(72) Inventors; and

(75) Inventors/Applicants (for US only): SHAO, Zhixin
[CN/DE]; Westend 43, 82377 Penzberg (DE).KRATZSCH, Peter [DE/DE]; Dorfstrasse 5, 82387
Antdorf (DE). SCHMUCK, Rainer [DE/DE]; Am
Klosterweiher 7, 83671 Benediktbeuern (DE). VON DER
ELTZ, Herbert [DE/DE]; In der Au 21, 82362 Weilheim
(DE). KENKLIES, Janet [DE/DE]; Am alten Bahnhof 1,
82377 Penzberg (DE).

(81) Designated States (national): CA, JP, US.

(84) Designated States (regional): European patent (AT, BE,
CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,
NL, PT, SE, TR).

Declarations under Rule 4.17:

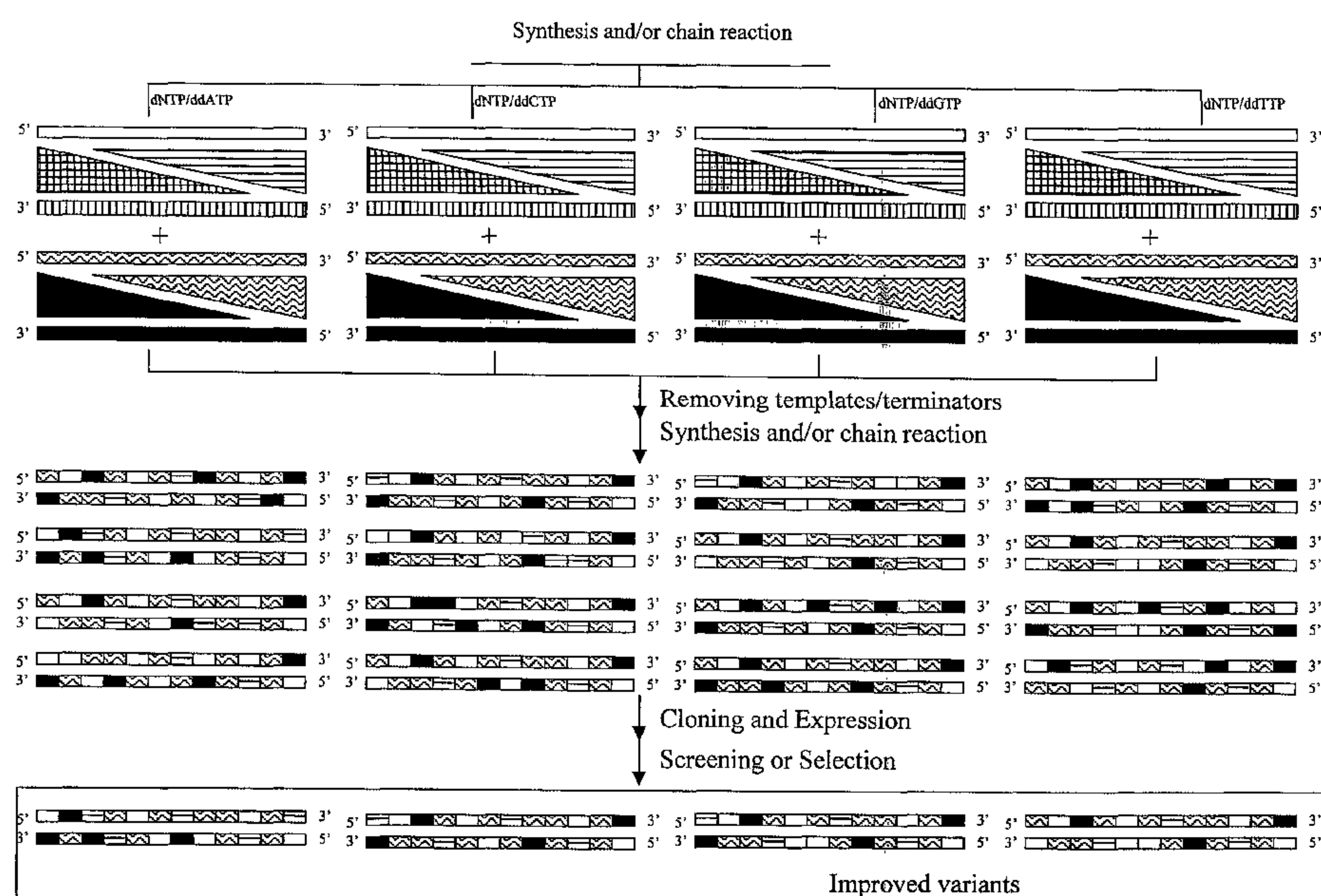
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations CA, JP, European patent (AT, BE, CH, CY, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR)
- of inventorship (Rule 4.17(iv)) for US only

Published:

- with international search report

(88) Date of publication of the international search report:
31 July 2003

[Continued on next page]

(54) Title: A WALK-THROUGH TECHNIQUE FOR *IN VITRO* RECOMBINATION OF POLYNUCLEOTIDE SEQUENCES

(57) **Abstract:** A simple and efficient method for random *in vitro* mutagenesis and recombination of polynucleotide sequences based on walk-through with chain-elongating molecules and chain terminating molecules e.g. dNTPs/ddNTPs followed by reassembly and amplification is described. The utility of this method for improving protein structure and/or function is demonstrated by creating novel sarcosine oxidase variants and by recombining human placental alkaline phosphatase and calf intestinal alkaline phosphatase for better mutants with improved properties.

WO 03/002736 A3



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.

A walk-through technique for *in vitro* recombination of polynucleotide sequences

A simple and efficient method for random *in vitro* mutagenesis and recombination of polynucleotide sequences based on walk-through with chain-elongating molecules and chain terminating molecules e.g. dNTPs/ddNTPs followed by reassembly and amplification is described. The utility of this method for improving protein structure and/or function is demonstrated by creating novel sarcosine oxidase variants and by recombining human placental alkaline phosphatase and calf intestinal alkaline phosphatase.

Various optimization procedures such as genetic algorithms (Holland, J. H. 1975. Adaptation in natural and artificial systems. The University Press, Ann Arbor, Goldberg, D. E. 1989. Genetic algorithms in search, optimization and machine learning. Addison-Wesley. Reading) and evolutionary strategies (Eigen, M. 1971. *Naturwissenschaften* 58: 465-523, Rechenberg, I. 1973. *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Frommann-Holzboog, Stuttgart) have been inspired by natural evolution. These procedures employ mutation, which makes small random changes in each single member of the population, as well as crossover, which combines properties of different individuals, to achieve a specific optimization goal. There also exist strong interplay's between mutation and crossover, as shown by computer simulations of different optimization problems (Brady, R. M. 1985. *Nature* 317: 804-806, Muhlenbein, H. 1991. *Parallel Computing* 17:619-632, Pál, K. F. 1993. *Bio. Cybern.* 69: 539-546, Pál, K. F. 1995. *Bio.Cybern.* 73, 335-341). Developing efficient and practical experimental techniques to mimic these key processes is a scientific challenge. The application of such techniques should allow us, for example, to explore and optimize the functions of biological molecules such as proteins and nucleic acids, *in vivo* or even completely free from the constraints of a living system (Joyce, G. F. 1992. *Scientific American*, 267:90-97, Shao, Z. and Arnold F. H. 1996. *Curr. Opin. Struct. Biol.* 6: 513-518).

Proteins are engineered with the goal of better understanding the molecular basis for their functions as well as much improving their performance for practical applications. A variety of approaches, involving both 'rational' and 'irrational' design, have been successfully used to optimize protein func-

tions (Shao, Z. and Arnold F. H. 1996. *Curr. Opin. Struct. Biol.* 6: 513-518). The choice of approach for a given optimization problem will depend upon the degree of understanding of the relationships between sequence, structure and function. The rational redesign of an enzyme's catalytic site, for example, often requires extensive knowledge of the enzyme structure, the structures of its complexes with various ligands and analogs of reaction intermediate and details of the catalytic mechanism. Such information is available only for a very few well-studied systems; little is known about the vast majority of potentially interesting enzymes.

For optimizing protein functions, several 'irrational' optimization approaches are being used (recently reviewed by Shao, Z. and Arnold F. H. 1996. *Curr. Opin. Struct. Biol.* 6: 513-518). These optimization procedures do not require much specific knowledge about the enzyme itself, only values of its function to be optimized. These procedures are even fairly tolerant to inaccuracies and noise in the function evaluation (Moore, J. C. and Arnold, F. H. 1996. *Chem. Eng. Sci.* 51:5091-5102). A particularly promising approach is to use both mutation and crossover synchronously to probe populations of solutions simultaneously for a specific optimization problem. Through multiple generations of screening or selection, bad local optima are evaded and better and better solutions are "bred" (Stemmer, W. P. C. 1994a. *Nature*, 370: 389-391, 1994; Stemmer, W. P. C. 1994b. *Proc. Natl. Acad. Sci., USA*, 91:10747-10751). During the past few years, more and more people have recognized the merits of evolutionary search strategies. Only a few practical techniques, however, such as

- sequential mutagenic PCR (Moore, J. C. and Arnold, F. H. 1996. *Chem. Eng. Sci.* 51:5091-5102),
- combinatorial cassette mutagenesis (Reidhaar-Olson, J. F. and Sauer, R. T. 1988. *Science* 241: 53-57),
- 'DNA shuffling' (Stemmer, W. P. C. 1994a. *Nature*, 370: 389-391; Stemmer, W. P. C. 1994b. *Proc. Natl. Acad. Sci., USA*, 91:10747-10751),
- random-priming recombination (Shao, Zhixin, H. Zhao, Lori Giver and Frances H. Arnold. 1998. *Nucleic Acids Res.* 26(2): 681-683),
- staggered-extension (StEP). H. Zhao et al. 1998, *Nature Biotechnol.* 16(3):258-2661),
- artificial recombination (Shao, Zhixin. 1999. *Protein Science* 8(Suppl.2): 55; Volkov, Alexander, Zhixin Shao, and Frances H. Arnold. 1999 *Nucleic Acids Res.* 27(18)e18:i-vi)

have been successfully applied to these problems.

One of the most commonly used technique of evolutionary optimization in procedure is "DNA-shuffling" (US 5,834,252; US 5,605,793; US 5,830,721; US 5,837,458; US 5,811,238) DNA-shuffling comprises the following steps:

- 1) creation of random double-stranded fragments from different double-stranded template polynucleotides wherein said different template polynucleotides contain preferably areas of identity and areas of heterology
- 2) reassembling of the random fragments
- 3) cloning
- 4) expression and
- 5) screening.

The different double-stranded template polynucleotides may belong to the same family of nucleic acids or proteins (i.e. are related) but which differ in their sequence (i.e. are not identical) and hence in their biological activity. The use of DNase I is described for random fragmentation in the prior art.

The DNA-shuffling using DNase fragmentation described by Stemmer et al. has a number of disadvantages. First of all, DNase I cleaves the DNA non-randomly. During an early phase of reaction cleavage occurs preferably in the middle of a DNA sequence and in the later phase of fragmentation cleavage occurs preferably between purine – and pyrimidine analogues. This leads to a fragmentation procedure which can be hardly controlled. Furthermore the minimal length of the desired gene is limited because of the non-controllable digestion of the template polynucleotide. Moreover, DNase has to be removed completely after the digestion step because it introduces disturbance into the further reassembly reaction.

The aim of the present invention was, therefore, developing a novel technique for the recombination of DNA sequences.

The subject of the present invention is a novel technique for *in vitro* walk-through recombination of DNA sequences. The technique involves walking through a template gene with a mixture of chain-elongating molecules and chain-terminating molecules, e.g. dNTPs/ddNTPs, to generate a pool of 3'

end-randomly distributed DNA fragments with a low level of point mutations. Thus, a fragment ladder is generated similar to fragment ladders during sequencing reactions. After removing the terminating molecules, e.g. ddNMP ends, these short DNA fragments can prime one another under appropriate reaction conditions based on homology and thus can be reassembled to form full-length genes by repeated thermocycling in the presence of thermostable DNA polymerase. These genes can be further amplified by a conventional PCR and cloned into a proper vector for expression of the encoded proteins. Screening or selection of the expressed mutants leads to new variants with improved or even novel functions. These variants can be immediately used as partial solutions to a practical problem, or they can serve as new starting points for further cycles of walk-through mutagenesis and recombination. This technique has been proven by creating *Bacillus subtilis* BMTU3420 sarcosine oxidase variants and by recombining human placental alkaline phosphatase and calf intestinal alkaline phosphatase. It was found that the technique is both – simple and efficient.

Therefore, subject of the present invention is a method for forming a polynucleotide sequence comprising the steps of

- a) generating a nucleic acid fragment ladder by nucleic acid synthesis which is carried out in presence of a reaction mixture comprising at least one primer, an enzyme having nucleic acid-synthesizing activity and a mixture of nucleic acid chain-terminating and nucleic acid chain-elongating molecules and a template polynucleotide,
- b) removing the chain-terminating molecules or changing it into non-terminating molecules,
- c) reassembly of the polynucleotide by hybridizing these fragments to one another or to template polynucleotides in presence of a thermostable enzyme having nucleic acid-synthesizing activity and conducting nucleic acid synthesis either
 - i) in presence of a mixture of nucleic acid chain-terminating molecules and nucleic acid chain elongating molecules or
 - ii) in presence of a reaction mixture of nucleic acid chain elongating molecules which does not contain chain-terminating moleculeswhereas in case of i) steps b) and c) will be repeated subsequently.

The step of removing chain-terminators and the step of fragment reassembly step can also be done at the same time. The subdivision of the process in steps a, b, c does not imply that three separate steps have to be performed. The three steps may also take place at the same time. It is also possible that the whole WalkThrough recombination may be performed in more preferred forms, such as in a single container, vessel or tube.

Even though one primer is enough for the fragment synthesis (or chain elongation), two or more primers are necessary for different fragment synthesis reactions in order to get correctly orientated fragments for the afterwards reassembly. That is, two or more primer are needed for the whole walk through process.

Only one template is enough for the WalkThrough process, as diversity can be introduced during the chain synthesis and fragment reassembly (resulting from, for example, synthesis errors of DNA polymerase). That means in case the template polynucleotide is not a mixture of different polynucleotides the DNA fragments may contain a low level of local mutations due to some synthesis errors and thus new sequences will be created.

Using more than one template may significantly increase the diversity of the variant library, and, consequently, finding potentially positive variants may be easier. Preferably, the template polynucleotide in step a) and/or c) is a mixture consisting of different template polynucleotides which may belong to the same family of nucleic acids or proteins (i.e. related) but which differ in their sequence (i.e. not identical) and hence differ in their biological activity. In this case gene shuffling may occur.

Conditions for the WalkThrough technique may vary from case to case and can be optimized at three main different levels:

- (a) fragment synthesis in the presence of chain-terminators,
- (b) Removing the chain terminators incorporated, and
- (c) Reassembling the fragments from the step (b).

a) Fragment synthesis in the presence of chain-terminators

This step is somehow similar to normal polynucleotide sequencing reaction, and, therefore, various conditions suitable for sequencing different targets may be applied to this step.

To carry out the WalkThrough procedure, the template can be single- or double-stranded polynucleotides in linear or closed circular form. The template may also be in the form of genomic DNA, that is, in native, intact, unpurified, or uncloned forms. Since, in most cases, the template genes are cloned in vectors into which no additional mutations should be introduced, they are usually first cleaved with restriction endonuclease(s) and purified from the vectors through agarose gel electrophoresis. For example linear DNA molecules are denatured, annealed to oligodeoxynucleotides for walkthrough reactions in the presence of an appropriate amount of chain-elongating/chain-terminating molecules as e.g. dNTPs/ddNTPs. Thus the oligonucleotides prime the DNA of interest at different positions along the entire target region extend to generate DNA fragments complementary to each strand of the template DNA. Due to some synthesis errors, these DNA fragments also contain a low level of local mutations. After removing the chain-terminating ends of ddNMP's, these DNA fragments can prime one another under appropriate reaction conditions based on homology and be reassembled into full-length genes by repeated thermocycling in the presence of thermostable DNA polymerase. The resulting full-length genes will have diverse sequences, most of which, however, still resemble that of the original template DNA.

The fragment synthesis can be carried out with either chain elongation under mild conditions using mesophilic DNA polymerases or thermocycle sequencing using thermostable polymerase. To adapt PCR to the WalkThrough synthesis provides a more convenient way for more nascent DNA fragments and makes our technique more robust.

If thermocycling is used for the fragment synthesis, the number of thermocycling can be from about 15 to about 55 cycles, depending on the amount of template and its purity. The concrete cycle number can be easily determined according to the quality and quantity of the newly synthesized fragments.

The amount of initial DNA template(s) may vary according to the chain elongation conditions. Normally, 0.1-10 pmol of ds-DNA is sufficient for this kind of reaction.

The length of the primers used may vary from 14-mer to 28-mer. The lengths of the selected primers should be long enough to prevent annealing to unspecific DNA binding positions and guarantee a good signal-to-noise ratio. We usually use primers of 18- to 21-mer.

The distances between the WalkThrough primers depend on the chain reaction modes and may vary from 100 to 800 nucleotides. The average distance between two primers in our cases is normally around 500 nucleotides.

The concentration of each primer can vary from 0.05 pmol to 2 pmol, depending on what kind of chain reaction is used. Normally, 0.1 pmol primer is used for the chain elongation of *B. subtilis* sarcosine oxidase gene, while 1.0 pmol each primer is used for the chain elongation of human placental alkaline phosphatase and calf intestinal alkaline phosphatase.

There are dozen of polymerases currently available, the synthesis of the nascent DNA fragments with randomly distributed 3' end can be achieved in more different fashions. For example, bacteriophage T4 DNA polymerase (Nossal 1974) or T7 sequenase version 2.0 DNA polymerase (Tabor and Richardson 1987, 1989), and /or thermostable DNA polymerases can be used for the WalkThrough synthesis. More preferably, other modified or mixture of DNA polymerases can be used which incorporate ddNTPs more efficiently than the wild type polymerases. For single-stranded RNA template, reverse transcriptase is preferred for the WalkThrough synthesis. Each of these enzymes is usually used under its optimized conditions to promote the chain elongation. Enzymes having nucleic acid-synthesizing activity can be DNA or RNA dependent polymerases.

It has been known for long time that different DNA polymerases can use 2', 3'-dideoxynucleotide triphosphates (ddNTPs) as substrates (Sanger, F., Nicklen, S., and Coulson, A.R. 1977. *Proc. Natl. Acad. Sci. U.S.A.* 74:5463-5467, Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.M., and Roe, B.A. 1980. *J. Mol. Biol.* 143:161-178) for DNA synthesis on single-stranded templates (Klenow, H. and I. Henningsen. 1970. *Proc. Natl. Acad. Sci.* 65:168).

When a ddNMP is incorporated at the 3' end of the growing chain, chain elongation is terminated at A, C, G, or T because the chain now lacks a 3' hydroxyl group. To generate the four chain ladders, only one of the four possible ddNTPs is included in each of the four reactions. The ddNTP/dNTP

ratio in each reaction is adjusted such that a portion of the elongating chains terminate at each occurrence of the base in the template corresponding to the included complementary ddNTP. In this way, each of the four elongation reactions contains a population of extended chains, all of which have a fixed 5' end determined by the annealed primers and a variable 3' end terminated at a specific dideoxynucleotide. The chain elongation and termination can also be done in one reaction in the presence of proper amount of ddNTPs/dNTPs.

Chain terminating molecules means that these molecules terminate nucleic acid synthesis and, therefore, a fragment ladder is generated. After generation of the fragment ladder suitable chain-terminating molecules can be removed. Suitable chain-terminating nucleotides are e.g. dideoxynucleotides as ddGTP, ddATP, ddCTP, ddTTP or derivatives thereof. Derivatives of dideoxynucleotides are defined as those dideoxynucleotides that are able to be incorporated by a thermostable DNA polymerase into growing DNA molecules that are synthesized in a thermocycling reaction. These dideoxynucleotides and derivatives are preferably used at a concentration of about 20 μ M to 1.0 mM. Such derivatives can include thionucleotides, 7-deaza-2'-dGTP, 7-deaza-2'-dATP as well as deoxyinosine triphosphate that can also be used as a substitute deoxynucleotide for dATP, dGTP, dTTP or dCTP, but are not limited to these. These deoxynucleotides and derivatives are preferably used at a concentration of about 4 μ M to 400 μ M.

The concentrations and the ratios of the ddNTP/dNTP may vary from case to case in order to get shorter or longer fragments. Though a ratio of dNTP/ddNTP between 1/50 and 5/1 works well for the chain elongation, it has to be optimized for different individual template in order to get fragment pools with randomly distributed 3'-ends. This is very important for allowing every nucleotide of the template should be copied at a similar frequency into products, and therefore, providing possibility to recombine or dissect two or more mutations although they may be very close to each other.

Preferred WalkThrough buffer systems normally include Tris-HCl at a concentration of about 50 to 500 mM, preferably of about 100 to 250 mM. The pH values of these buffer systems range from pH6.5 to pH 10.0, depending on polymerases used. MgCl_2 is generally included in these buffer systems at a concentration ranging from 1.0 to 5.0 mM. KCl may also be included at a concentration of 2-80 mM. Certain amount of mercaptoethanol (0.5-1.5%), Tween 20 (0.2-0.4%) and DMSO (1-5%) may also exist in the buffers. Other agents, such as glycerol, betaine, etc., which can lower the

melting point of the templates may also be included in the buffers to facilitate the WalkThrough reaction.

The Walk-Through DNA synthesis is based on the chain elongation guided by template, and the nascent strand is synthesized from the 3'-OH termini at the primers using polymerase and the four deoxynucleoside triphosphates and stop after the ddNTP incorporation. Thus the reaction is independent of the length of the DNA template.

(b) Removing the chain terminators incorporated

This step is the second key step for successful WalkThrough recombination. The end nucleotide or incorporated terminator at each newly generated fragments must be removed so that each such fragment gains a free 3'-OH group necessary for further fragment reassembly.

Methods of removing chain-terminating molecules are known to a person skilled in the art. Chain-terminating molecules can be removed, for example, by DNA polymerases, many DNA polymerases have a 3'-5' exonuclease activity. This activity removes a single nucleotide at a time, releasing a nucleotide 5' monophosphate. In the absence of dNTPs, this activity will catalyze stepwise degradation from 3' end of both single- and double-stranded DNA. The polymerase which may serve this purpose include Klenow fragment of *E. coli* DNA polymerase I (Jacobsen, H., Klenow, H., and Overgaard-Hansen, K. 1974. The N-terminal amino-acid sequences of DNA polymerase I from *Escherichia coli* and of the large and the small fragments obtained by a limited proteolysis. *Eur. J. Biochem.* 45:623-627. Joyce, C.M. and Gridley, N.D.F. 1983. Construction of a plasmid that overproduces the large proteolytic fragment (Klenow fragment) of DNA polymerase I of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 80:1830-1834), bacteriophage T4 DNA polymerase (Nossal, N.G. 1984, Prokaryotic DNA replication systems. *Annu. Rev. Biochem.* 53:581-615. Lin, T.C., Rush, J., Spicer, E. K., and Konigsberg, W.H. 1987. Cloning and expression of T4 DNA polymerase. *Proc. Natl. Acad. Sci. USA* 84:7000-7004), bacteriophage T7 DNA polymerase (Tabor, S. and Richardson, C.C. 1987. DNA sequencing analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA* 84:4767-4771). The 3'-end terminators may be removed from the fragments with DNA polymerase with 3'→5' proof reading activity in the presence of very limited or no dNTPs under mild conditions. Alternatively, chain-terminating molecules can also be removed by exonucleases, such as Exonuclease III. This enzyme is a multifunctional enzyme that catalyzes hydrolysis of several types of

phosphodiester bonds in double stranded DNA. The main application of Exo III is as a 3'-5' double-stranded specific exonuclease that catalyzes release of 3' nucleotides from the 3'- end of double stranded DNA (Roger, S.G. and Weiss, B. 1980. Exonuclease III of Escherichia coli K-12, an AP exonuclease. Meth. Enzymol. 65:201-211).

Preferred reaction buffers used for terminator-removing reactions normally include Tris-HCl at a concentration of about 40 to 200 mM, preferably of about 50 to 100 mM. The pH values of these buffer systems range from pH 6.5 to pH9.0, depending on enzymes with 3'-5' exonuclease activity used. MgCl₂ is generally included in these buffer systems at a concentration ranging from 1.0 to 3.0 mM. KCl may also be included at a concentration of 5-50 mM. Certain amount of mercaptoethanol (0.5-1.5%) may also exist in the buffers.

The reaction normally is carried out at 37°C for 0.5-2.0 hours.

Using thermostable DNA polymerases and thermostable enzymes with 3'->5' exonuclease activity in the proper buffer and under the optimized conditions, the steps of terminator removal and re-assembly can also be combined, so that no separate (B) and (C) steps are necessary. This method of the present invention is particularly preferred, as demonstrated in the example of recombining human placental alkaline phosphatase and calf intestinal alkaline phosphatase.

(c) Reassembling the fragments

The reassembling of the fragments is a PCR like reaction, comprising cycles of DNA-denaturation, annealing and DNA synthesis. During the annealing ssDNA-fragments hybridize in homologues areas. The overlapping ends of the ssDNA are extended by a polymerase. In case ssDNA random fragments from different polynucleotides hybridize, gene-shuffling occurs. Gene-shuffling means recombination between homologues but non-identical sequences. The term "identical" means that two nucleic acid sequences have the same sequence or a complementary sequence. Thus, "areas of identity" means that regions or areas of a nucleic acid fragment or polynucleotide are identical or complementary to another polynucleotide or nucleic acid fragment. The term "homologues" means that one single-stranded (ss) nucleic acid sequence may hybridize to a complementary ss nucleic acid sequence. The degree of hybridization may depend on a number of factors including the amount of

identity between sequences and the hybridization conditions such as temperature and salt concentration as discussed later.

The reassembly step is the most critical one in the whole WalkThrough process, as all synthesized fragments during Step (A) prime one another for elongation without any additional primers added.

Since the concentration of the DNA is the most important variable, it is useful to set up three separate reactions with different concentrations (high, middle, low). According to our experience, the amount of DNA fragments between 0.1 μ g to 2.0 μ g usually gives satisfied reassembly results in a reaction volume of 20-50 μ l.

The deoxynucleotides during reassembly are preferably used at a concentration of about 100 μ M to 400 μ M, and the chosen concentration depends on the cycle number of the reaction.

Preferred reassembly buffers normally include Tris-HCl at a concentration of about 5 to 50 mM, preferably of about 10 mM. The pH values of these buffer systems range from pH 7.5 to pH 10.0, depending on polymerases used. $MgCl_2$ is generally included in these buffer systems at a concentration ranging from 1.0 to 5.0 mM. KCl may also be included at a concentration of 20-80 mM. Certain amount of mercaptoethanol (0.5-1.5%), Tween 20 (0.2-0.4%) or Triton-X 100 (0.1-0.5%) may also exist in the buffers.

A typical reassembly cycle consists of three steps: the first step is heat denaturation of the double-stranded target nucleic acid. The exact conditions required for denaturation of the sample nucleic acid depend on the length and composition of the sample nucleic acid. Typically, an incubation at 90°C-100°C for about 10 seconds up to 5 minutes is efficient to denature the sample nucleic acid. The annealing temperature used in reassembly reaction is about 40°C to 70°C, usually ranging from about 55°C to 65°C and lasting for a period of 15 second to 60 seconds. The elongation is usually done under conditions sufficient to provide for polymerization of nucleotides to the fragment ends. To achieve polymerization conditions, the temperature of the reaction mixture will typically be maintained at a temperature ranging from about 65°C to 75°C, more preferably at 68°C to 72°C for

about 15 second to 2 minutes, preferably for 30 seconds to 1 minute depending on the length of the finally reassembled DNA.

The number of thermocycling can be from about 15 to about 55 cycles, depending on the amount of template and the length of the finally reassembled DNA. The concrete cycle number can be easily determined according to the quality and quantity of the newly synthesized fragments.

Several thermostable polymerases with different synthesis fidelity can be used for the reassembly, depending on what kind of the error rate the final reassembly product should have. Each of these enzymes is usually used under its optimized conditions to promote the reassembly. The goal is that the resulting full-length genes will have diverse sequences, most of which, however, still resemble that of the original template DNA.

These sequences obtained after the reassembling step can be further amplified by a conventional PCR and cloned into a vector for expression. Suitable vectors are known to a person skilled in the art and vectors cited in the following references are herein incorporated by reference Kingsman SM, Kingsman AJ. *Philos Trans. R. Soc. Lond B. Biol. Sci.* 1989, 324(1224):477-485; Bailey JE. *Adv. Biochem. Engl. Biotechnol.* 1993, 48:29-52. Suitable expression systems are known in the art as well and expression systems cited in the following references are herein incorporated by reference Shatzman AR, Rosenberg M. *Methods Enzymol.* 1987; 152:661-73. Screening or selection of the expressed mutants should lead to variants with improved or even new specific functions. Suitable screening and selection systems are known in the art and screening and selection systems cited in the following references are herein incorporated by Kuchner O, Arnold FH. *Trends Biotechnol.* 1997 Dec.; 15(12):523-30; Patel PH, Loeb LA. *Procc. Natl. Acad. Sci. USA*, 2000 May 9; 97(10):5095-100. These variants can be immediately used as partial solutions to a practical problem, or they can serve as new starting points for further cycles of directed evolution.

Compared to other techniques used for protein optimization, such as combinatorial cassette and oligonucleotide-directed mutagenesis (Reidhaar-Olson, J. F. and Sauer, R. T. 1988. Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences. *Science* 241: 53-57; Oliphant, A. R., Nussbaum, A. L., and Struhl, K. 1986. Cloning of random-sequence oligodeoxynucleotides. *Gene* 44: 177-183, Hermes, J. D., Blacklow, S. C., and Knowles, J. R. 1990. Searching

sequence space by definably random mutagenesis - improving the catalytic potency of an enzyme. Proc. Natl. Acad. Sci. USA 87: 696-700), error-prone PCR (Leung, D. W., Chen, E., and Goeddel, D. V. 1989. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. BioTechnique 1: 11-15; Chen K. and Arnold, F. H. 1993 . Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. Proc. Natl. Acad. Sci. USA 90:5681-5622), or DNA 'shuffling' (Stemmer, W. P. C. 1994a. Rapid evolution of a protein *in vitro* by DNA shuffling. Nature, 370: 389-391; Stemmer, W. P. C. 1994b. DNA shuffling by random fragmentation and reassembly -*in vitro* recombination for molecular evolution. Proc. Natl. Acad. Sci., USA, 91:10747-10751), this walk-through based technique shows several advantages for *in vitro* protein optimization:

- (1) Since the template used for walkthrough synthesis may be either single- or double-stranded polynucleotides, one limitation of DNA 'shuffling' (Stemmer, W. P. C. 1994a. Rapid evolution of a protein *in vitro* by DNA shuffling. Nature, 370: 389-391;), which has to employ double-stranded polynucleotide as a template, does not exist any more. With the technique described here, any potential mutations and/or crossovers can be introduced at DNA level by using different DNA-dependent DNA polymerases, or even directly from mRNA level by using different RNA-dependent DNA polymerases. This provides more opportunities and realities to achieve the goal of optimizing protein specific functions.
- (2) Different from the DNA 'shuffling' procedure which requires fragmenting the double-stranded DNA template with DNase I to get random fragments (Stemmer, W. P. C. 1994a. Nature, 370: 389-391; Stemmer, W. P. C. 1994b. Proc. Natl. Acad. Sci., USA, 91:10747-10751), the technique described here employs WalkThrough synthesis to obtain size-controllable DNA fragments with randomly distributed 3' end as "breeding blocks" for further reassembly (Fig. 1).
- (3) Since the WalkThrough chains are a population of fragments that each stops in every position, they are uniform in their positional preference and lack a sequence bias. The sequence heterogeneity allows both, mutations and crossover may happen more randomly than, for example, with error-prone PCR or DNA 'shuffling'.

- (4) Normal error-prone PCR and DNA shuffling can not efficiently recombine or dissect two or more mutations if they are very close to each other. (Stemmer, W. P. C. 1994a. Rapid evolution of a protein in vitro by DNA shuffling. *Nature*, 370: 389-391). In contrast, Walk Through approach allows recombination occurring at every position of templates and therefore, provides possibility to recombine or dissect two or more mutations although they may be very close to each other.
- (5) The Walk-Through DNA synthesis is based on the chain elongation guided by template, and the nascent strands are synthesized from the 3'-OH termini at the primers using polymerase and the four deoxynucleoside triphosphates and stop after the ddNTP incorporation. Thus the reaction is independent of the length of the DNA template. This is particularly useful for engineering small peptides or large enzymes or even enzyme pathways.
- (6) Since DNase I is an endonuclease that hydrolyzes double-stranded DNA preferentially at sites adjacent to pyrimidine nucleotides, its use in DNA shuffling (Stemmer, W. P. C. 1994a.. *Nature*, 370: 389-391; Stemmer, W. P. C. 1994b. *Proc. Natl. Acad. Sci.*, USA, 91:10747-10751) may result in bias (particularly for genes with high G+C or high A+T content) at the step of template gene digestion. Effects of this potential bias on the overall mutation rate and recombination frequency have not yet been investigated, but they may be avoided by using the Walk-Through approach.
- (7) Since there are dozen of polymerases currently available, the synthesis of the nascent DNA fragments with randomly distributed 3' end can be achieved in more different fashions. For example, bacteriophage T4 DNA polymerase (Nossal, N.G. 1974. *J. Biol. Chem.* 249: 5668-5676) or T7 sequenase version 2.0 DNA polymerase (Tabor, S. and Richardson, C. C. 1987. *Proc. Natl. Acad. Sci.*, USA, 84:4767-4771, Tabor, S. and Richardson, C. C. 1989. *J. Biol. Chem.* 264:6447-6458), and /or thermostable DNA polymerases can be used for the WalkThrough synthesis.

For single-stranded polynucleotide template (particularly for RNA template), reverse transcriptase is preferred for the WalkThrough synthesis. Since this enzyme lacks 3'→5' exonuclease activity, it is therefore prone to error. In the presence of high concentrations of dNTPs and

Mn²⁺, about 1 base in every 500 is misincorporated (Sambrook, J., Fritsch, E. F. and Maniatis, T. 1989. Molecular cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

- (8) One of the key steps in our technique is to control the 3' end of the nascent, single-strand DNA synthesized during the WalkThrough process. Under certain conditions, this step may be used for efficient terminal and/or internal insertion/deletion, resulting in molecules with different sizes. Efficient changing target molecule sizes is normally unachievable through error-prone PCR or DNA shuffling.
- (9) By modifying reaction conditions, PCR can be adjusted for the WalkThrough synthesis using thermostable polymerase for the short, nascent DNA fragments. To adapt PCR to the Walk Through synthesis provides a more convenient way for more nascent DNA fragments and makes our technique more robust.

The inventive method of the present invention may, however, be combined with other known methods of the art if this seems to be advantageously.

WalkThrough DNA synthesis was used to generate DNA fragments with randomly distributed 3' ends from denatured, linear, double-stranded DNA (e.g., restriction fragments purified by gel electrophoresis) to *Bacillus subtilis* BMTU3420 sarcosine oxidase gene. The purified DNA, mixed with a molar excess of primers, was denatured, and synthesis was then carried out using the Stoffel fragment. This enzyme lacks 5'→3' exonuclease activity, so that the WalkThrough product was synthesized exclusively by extension and was not degraded by exonuclease (see Example 1).

The WalkThrough recombination technique is further demonstrated with another example of recombining human placental alkaline phosphatase and calf intestinal alkaline phosphatase. (see Example 2). In example 2, the step of removing chain-terminators and the step of fragment re-assembly step can be done at the same time.

The technique described here may be used to explore the vast space of potentially useful catalysts for their optimal performance in a wide range of applications as well as to develop or evolve new enzymes for basic structure-function studies.

Most of the experimental conditions described here may potentially be used with little modification for optimizing native enzymes, part (or all) of the enzyme pathways, as well as other biomolecules. For example, in terms of reducing proteolytic activity, probing activator genes, enhancing secretion mechanisms, or solving other fundamental enzyme engineering problems, the technique described here may find immediate applications.

While this protocol describes using DNA-dependent DNA polymerase and single-stranded DNA as template, alternative protocols are also feasible for using single-stranded RNA as a template. By using specific enzyme mRNA as the template and RNA-dependent DNA polymerase (reverse transcriptase) as the catalyst, our approach may be modified to introduce mutations and crossovers into cDNA clones and to create molecular diversity directly from mRNA level to achieve the goal of optimizing enzyme functions. Reverse transcriptases are derived from retrovirus, such as avian myeloblastosis virus (AMV) or Moloney murine leukemia virus (MMLV), which use them to make DNA copies of their RNA genomes. AMV and MMLV reverse transcriptases (Verma IM. The reverse transcriptase. *Biochim. Biophys. Acta.* 1977 Mar. 21; 473(1):1-38. Roth MJ. Tanese N. Goff SP. Gene product of Moloney murine leukemia virus required for proviral integration is a DNA-binding protein. *J. Mol. Biol.* 1988; 203(1):131-9) are mainly used as RNA-directed DNA polymerase. Specifically, deoxyoligonucleotides are used as primers for extension on RNA (usually messenger RNA) templates to generate complementary DNA.

Figure 1:

Walking through the single-stranded polynucleotide template with dNTPs/ddNTPs to generate fragments with randomly-distributed 3' end, removing ddNMP from the 3' end, reassembly of the full length DNA by thermocycling in the presence of DANN polymerase and nucleotides and amplification of reassembled products by conventional PCR for further cloning and screening.

Figure 2:

Primer arrangement along human placental alkaline phosphatase (hpap) and calf intestinal alkaline phosphatase (ciap) genes

Figure 3:

The N-terminal sequence alignment of parental ciap and hpap genes and their recombination variants (AP01, AP03, AP05, AP06, AP11 and AP15)

Figure 4:

The C-terminal sequence alignment of parental parental ciap and hpap genes and their recombination variants (AP2, AP9 and AP13)

Figure 5:

Sequences of hpap, ciap, AP01, AP03, AP05, AP06, AP11, AP15, AP2, AP9 and AP13.

Example 1

1. The DNA of interest with appropriate restriction endonuclease(s) and purify the DNA fragment of interest was purified by gel electrophoresis using Roche High Pure PCR Prep Kit (Roche Diagnostics GmbH, Germany)). As an example, the *Bacillus subtilis* BMTU3420 sarcosine oxidase gene was cleaved as a 1.2 kb-long *Pst*I-*Nsi*I fragment from the recombinant plasmid pBMTU5823.
2. About 0.5 pmol of the double-stranded DNA dissolved in H₂O was mixed with 0.1 pmol of each SODF1, SODF2, SODR1, and SODR2 primers. After immersion in boiling water for 3 minutes, the mixture was placed immediately in an ice/ethanol bath.
3. Ten µl of 10 x reaction buffer [10X buffer: 900 mM HEPES, pH 6.6; 0.1 M magnesium chloride, 10 mM dithiothreitol, 5 mM each dATP, dCTP, dGTP and dTTP, and 2mM each ddATP, ddCTP, ddGTP and ddTTP) was added to the denatured sample, and the total volume of the re-

action mixture was brought up to 98 μ l with H₂O.

4. Ten units (about 2 μ l) of the Stoffel fragment was added. All the components were mixed by gently tapping the outside of the tube and were centrifuged at 12,000 g for 1-2 seconds in a microfuge to move all the liquid to the bottom. The reaction was carried out under standard PCR cycling sequencing conditions using Taq DNA Sequencing Kit for Standard and Cycle Sequencing Kit (Roche Diagnostics GmbH, Germany).
5. The reaction products were subjected to Wizard DNA Clean up System (Promega, WI, USA) to remove the enzymes, primers, reaction buffer components and dNTPs/ddNTPs. The purified products were incubated in 1 x Klenow buffer and 10 U of Klenow at 37°C for 40 minutes. Two μ l of DpnI (10 U/ μ l) was added to the mixture, and the incubation was continued at 37°C for further 40 minutes.
6. The digested WalkThrough products were purified with High Pure PCR Purification Kit (Roche Diagnostics GmbH, Germany) and was used for whole gene reassembly.

Reassembly of the whole gene

1. For SarcOD gene reassembly by PCR, 5 μ l of the purified WalkThrough DNA fragments, 20 μ l of 2X PCR pre-mix (5-fold diluted cloned *Pfu* buffer, 0.5 mM each dNTP, 0.1U/ μ l cloned *Pfu* polymerase (Stratagene, La Jolla, CA)) and 15 μ l of H₂O were mixed on ice.
2. After incubation at 96°C for 5 min, 40 thermocycles were performed, each with 1.0 min at 95°C, 1.0 min at 55°C and 0 + 5 sec/cycle at 72°C, with the extension step of the last cycle proceeding at 72°C for 10 min, in an Eppendorf Mastercycler Gradient (Eppendorf, Germany).
3. 3 μ l aliquots at cycles 20, 30 and 40 were removed from the reaction mixture and analyzed by agarose gel electrophoresis. The reassembled PCR product at 40 cycles contained the correct size product in a smear of larger and smaller sizes.

Amplification

The correctly reassembled product of this first PCR was further amplified in a second PCR reaction which contained the PCR primers complementary to the ends of the template DNA.

1. 2.0 µl of the PCR reassembly aliquots were used as template in 100-µl standard PCR reactions, which contained 0.2 µM each primers of *pstF* (5' GGTAGAGCGAG-TCTCGAGGGGGGAGATG C 3', SEQ ID NO: 1) and *NsiR* (5' AGCCGGCGTGACGTGGGTCAGC 3', SEQ ID NO: 2), 1.5 mM MgCl₂, 10 mM Tris-HCl [pH 9.0], 50 mM KCl, 200 µM each of the four dNTPs, 2.5 U of *Taq* polymerase (Roche, Germany) and 2.5 U of *Pfu* polymerase (Stratagene, La Jolla, CA).
2. After incubation at 96°C for 5 min, 30 thermocycles were performed, each with 1.0 min at 95°C and 1.0 min at 72°C, with the extension step of the last cycle proceeding at 72°C for 10 min, in an Eppendorf Mastercycler Gradient (Eppendorf GmbH, Germany).
3. The amplification resulted in a large amount of PCR product with the correct size of the sarcosine oxidase whole gene.

Cloning

1. The PCR product of Bs SarcOD gene was digested with *Pst* I and *Nsi* I restriction enzymes, and cloned into pBMTU5823 (*Pst/Nsi*).
2. *E. coli* XL1 F' cells were transformed with the above ligation mixture to form a mutant library.

Results:

After only one round of applying this technique on the wild-type SarcOD gene, two Hind III and Pvu II resistant mutants have been found to be enzymatic active.

Example 2

Recombination of human placental alkaline phosphatase and calf intestinal alkaline phosphatase genes

1. Description of the method

Walk through recombination' is a method used to recombine two or more DNA-sequences based on their homology. Furthermore point mutations can also be introduced during the recombination steps. This method consists mainly of five steps:

Fragment synthesis

Template removal

Terminator removal

Reassembly

Amplification

1.1 Fragment synthesis

Fragment synthesis results from a DNA synthesis reaction where extension is terminated by the incorporation of dideoxy nucleotides. Target DNA sequences selected for recombination serve as templates in this reaction. The number of primers should be chosen depending on the length of the target DNA sequences and the length of the resulting fragments. The fragment length may also be controlled by the reaction conditions. Furthermore, a 'Cycle Sequencing reaction' is favorable because of its higher product yield and easy use.

1.2 Template removal

To avoid interference with the reassembly step, template DNA has to be completely removed after the fragment synthesis. Several methods can be used for this purpose. we have chosen to use U-DNA and Uracil-DNA Glycosylase. Uracil-DNA Glycosylase can be used to remove uracil base at any site where a deoxyuridylate has been incorporated (U-DNA). The resulting abasic site can subsequently be hydrolyzed by alkali-treatment, high temperatures or specific endonucleases. In our approach, a simple temperature treatment is sufficient. U-DNA can be prepared by a PCR reaction using dUTP's instead of dTTP's. After the U-DNA has served as the template for the fragment synthesis, the whole

reaction is subjected to the uracil-DNA glycosylase treatment and temperature treatment. Another control PCR can be used to ensure the complete U-DNA removal.

1.3 Terminator removal

The fragments need to have 3'-OH end for the reassembly reaction. Since our fragments are terminated with dideoxynucleotides, they do not carry 3'-OH ends. These terminators can be removed by 3'-5'-exonuclease activity of several nucleases or DNA polymerases resulting in 3'-OH end. We combine terminator removal with the reassembly by employing a thermostable exonuclease III and *Taq* polymerase for this step. Exonuclease III cuts the terminator by its 3'-5'-exonuclease activity (enzyme activity only at ds DNA, and therefore only after DNA annealing) and *Taq* polymerase extends the fragment by its 5'-3'-polymerase activity.

1.4. Reassembly

Reassembly is a PCR-like reaction without primers, there the fragments anneal to each other based on their homologies and extend. Through all cycles of denaturation, annealing and extension fragments growth or reassemble up to the length of the original DNA sequences.

1.5 Amplification

Amplifying the reassembled DNA takes place in a PCR with sequence flanking primers in order to provide enough material for subsequent cloning and analysing steps.

2. Experimental Protocol

2.1 Preparation of U-DNA

dUTP containing DNA template is generated by PCR using 2.5 mM MgCl₂ according to the manual (,Uracyl-DNA Glycosylase' Roche # 1 444 646), 600 µM dUTP and three other nucleotides dATP, dCTP and dGTP at a concentration of 200 µM each. A typical reaction mixture and cycling conditions used for recombining the alkaline phosphatase are shown below:

Template DNA (plasmid)	40 ng
Primer APhpaF	20 pmol

22

Primer APxbaR	20 pmol	
10 x <i>Taq</i> Puffer	10 μ l	
<i>Taq</i> DNA polymerase	0,5 U	
MgCl ₂	2,5 mM	
dATP, dCTP, dGTP (Roche # 1 969 064)	0,2 mM	each
dUTP (Roche # 1 420 470)	0,6 mM	
H ₂ O	ad 100 μ l	

Cycling conditions:

95 °C	5 min	
95 °C	1 min	
60 °C	1 min	30 x
72 °C	2 min	
72 °C	10 min	
4 °C	∞	

The synthesized U-DNA is separated from the template DNA by preparative agarose gel electrophoresis (1% agarose/TAE) and gel extraction (,QIAquick Gel Extraction Kit' Qiagen # 28706).

2.2 Fragment synthesis

Fragments are synthesized using Cycle sequencing reactions (,DIG Taq DNA Sequencing Kit for Standard and Cycle Sequencing' Roche # 1449443) with some modifications. Twelve reaction mixtures are set up for a total of six primers and two U-DNA templates. A typical reaction mixture with cycling conditions used for recombining the two alkaline phosphatase genes are shown as following:

U-DNA	175 ng
Primer (1 μ M)	1 pmol
Puffer	3 μ l
<i>Taq</i> DNA polymerase (5U/ μ l)	0,6 μ l
Termination mixture ddATP/dGTP	2,5 μ l

Termination mixture ddCTP/dGTP	2,5 µl
Termination mixture ddGTP/dGTP	2,5 µl
Termination mixture ddTTP/dGTP	2,5 µl
H ₂ O	ad 30 µl

Cycling conditions:

95 °C	5 min	
95 °C	30 sec	
60 °C	30 sec	30 x
72 °C	1 min	
4 °C	∞	

The calf intestinal alkaline phosphatase gene (*ciap*) and the human placental alkaline phosphatase gene (*hpap*) show a sequence similarity of 81 % and have lengths of about 1530 bps. Six primers are used for fragment synthesis. Two external primers (APhpaF and APxbaR) share the same sequence. Figure 2 shows the primer arrangement along human placental alkaline phosphatase (*hpap*) and calf intestinal alkaline phosphatase (*ciap*) genes.

Primer sequencesFlanking primers

APhpaF	CTT CGG CGT TCA GTA ACA CGC (SEQ ID NO.: 3)
ApxbaR	GCT TTC GAG GTG AAT TTC GAC C (SEQ ID NO.: 4)

Primers for ciap gene

CIAPint1F	GGT CAC GTC TGT GAT CAA CCG (SEQ ID NO.: 5)
CIAPint1R	CGG TTG ATC ACA GAC GTG AC C (SEQ ID NO.: 6)
CIAPint2F	CGC AAA GCT TAT ATG GCA CTG AC (SEQ ID NO.: 7)

CIAPint2R GTC AGT GCC ATA TAA GCT TTG CC
(SEQ ID NO.: 8)

Primers for hpap gene

HPAPint1F CCA AGA AAG CAG GGA AGT CAG TG
(SEQ ID NO.: 9)

HPAPint1R CAC TGA CTT CCC TGC TTT CTT GG
(SEQ ID NO.: 10)

HPAPint2F CAT GTT CGA CGA CGC CAT TGA G
(SEQ ID NO.: 11)

HPAPint2R CTC AAT GGC GTC GTC GAA CAT G
(SEQ ID NO.: 12)

2.3 Uracil-DNA Glycosylase treatment and temperature treatment

This step is accomplished by addition of 1 µl Uracil-DNA Glycosylase (1 U/µl; Roche , #1775375) to each fragment synthesis mixture and by incubation at 37 °C for 4 h to cleavage the U-bases followed by incubation at 95 °C for 2 min to hydrolyse the abasic sites and to inactivate the enzyme. Another PCR may be used to control the complete removal of the template U-DNA.

2.4 Reassembly

It is important to use proper amount of purified DNA in the critical reassembly step. DNA fragments are purified from the cleavage products by using Microcon (Microcon 50; Millipore # 42415) and are retained in H₂O. A typical mixture and cycling conditions used for recombining the two alkaline phosphatase genes are shown below:

Purified fragments	x µl (proper amount)
dNTP's (Roche # 1 969 064)	0,4 mM each
Expand High Fidelity 10x buffer (Roche # 1 759 175)	5 µl
Taq/ Exo III mix	1 U
H ₂ O	ad 50 µl

After 30 cycles 1 µl of Taq/ExoIII is added to the reaction mixture for further reassembly.

Cycling conditions:

95 °C	5 min	
95 °C	1 min	
45 °C	1 min	50 cycles
72 °C	30 sec	
72 °C	10 min	
4 °C	∞	

2.5 Amplification PCR (also used as control PCR)

The amplification PCR is a standard PCR with the two sequence flanking primers (AphpaF and ApxbaR). The following is a typical mixture with cycling conditions used for recombining the two alkaline phosphatase genes:

Reassembly mixture	0,5 µl
Primer APhpaF	10 pmol
Primer APxbaR	10 pmol
10 x <i>Taq</i> Puffer (+ Mg; Roche # 1 271 318)	5 µl
<i>Taq</i> DNA polymerase (Roche # 1 418 432)	2,5 U
dATP, dCTP, dGTP, dTTP (Roche # 1 969 064)	0,2 mM each
H ₂ O	ad 50 µl

Cycling conditions:

95 °C	5 min	
95 °C	1 min	
60 °C	1 min	30 x
72 °C	2 min	
72 °C	10 min	
4 °C	∞	

2.6 Cloning and expression

The amplification product is purified by preparative gel electrophoresis (1% agarose/TAE) and gel extraction (QIAquick Gel Extraction Kit' Quiagen # 28706), cloned into vector pCR®-XL-TOPO® according to the suppliers manual (Invitrogen # K 475020) and expressed in E. coli TOP 10 cells delivered with the vector.

2.7 Sequencing results

Nine recombination clones have been analyzed by sequencing using one sequencing primer. The results are summarized in Table 1. Eight of these clones show one ore more recombination events within the sequenced region (445 – 645 bp). One variant contains 4 crossovers within the mentioned region. The overall mutation rate is about 0,25 %. Since the reassembly products can be ligated into the vector in two different orientations, sequencing these clones with only one primer could give both sequences of the N-terminal and coding for the C-terminal regions.

Clone	Direction of sequencing	Length of sequenced region	Crossovers	Mutations
AP 1	N-terminus	509 bp	1	1
AP 3	N-terminus	445 bp	1	0
AP 5	N-terminus	624 bp	4	3
AP 6	N-terminus	553 bp	1	1
AP 11	N-terminus	599 bp	2	1
AP 15	N-terminus	528 bp	1	2
AP 2	C-terminus	645 bp	1	3
AP 9	C-terminus	639 bp	0	1
AP13	C-terminus	638 bp	2	1

Table 1: Sequencing results after recombination of two alkaline phosphatase genes

2.8 Sequences

Figure 3 shows the N-terminal sequence alignment of parental ciap and hpap genes and their recombination variants (AP01, AP03, AP05, AP06, AP11 and AP15)

Figure 4 shows the C-terminal sequence alignment of parental parental ciap and hpap genes and their recombination variants (AP2, AP9 and AP13)

Figure 5 shows sequences of hpap, ciap, AP01, AP03, AP05, AP06, AP11, AP15, AP2, AP9 and AP13.

1/22

SEQUENCE LISTING

<110> Roche Diagnostics GmbH

<120> A walk-through technique for in vitro recombination of polynucleotide sequences

<130> 19061-Koe

<160> 36

<170> PatentIn version 3.1

<210> 1

<211> 28

<212> DNA

<213> Artificial primer

<400> 1

ggtagagcga gtctcgaggg ggagatgc

28

<210> 2

<211> 22

<212> DNA

<213> Artificial primer

<400> 2

agccggcgtg acgtgggtca gc

22

<210> 3

2/22

<211> 21

<212> DNA

<213> Artificial primer

<400> 3

cttcggcggtt cagtaacacg c

21

<210> 4

<211> 22

<212> DNA

<213> Artificial primer

<400> 4

gcttttcgagg tgaatttcga cc

22

<210> 5

<211> 21

<212> DNA

<213> Artificial primer

<400> 5

ggtcacgtct gtgatcaacc g

21

<210> 6

<211> 21

<212> DNA

<213> Artificial primer

<400> 6

cggttgatca cagacgtgac c

21

3/22

<210> 7

<211> 23

<212> DNA

<213> Artificial primer

<400> 7

cgcaaagctt atatggcact gac

23

<210> 8

<211> 23

<212> DNA

<213> Artificial primer

<400> 8

gtcagtgcca tataagcttt gcc

23

<210> 9

<211> 23

<212> DNA

<213> Artificial primer

<400> 9

ccaagaaagc agggaagtca gtg

23

<210> 10

<211> 23

<212> DNA

<213> Artificial primer

4/22

<400> 10
cactgacttc cctgctttct tgg 23

<210> 11

<211> 22

<212> DNA

<213> Artificial primer

<400> 11
catgttcgac gacgccattg ag 22

<210> 12

<211> 22

<212> DNA

<213> Artificial primer

<400> 12
ctcaatggcg tcgtcgaaca tg 22

<210> 13

<211> 700

<212> DNA

<213> Human placental phosphatase

<400> 13
gttaacgtgc tgaacagccg ggcatttttt tacgctatac cctacataat aaaaccggag 60
ctaccatgaa taagaaggta ctgacccttt ctgccgtgat ggcaagtctg ttattcggcg 120
cgcacgcgca cgcggcgatc ccagttgagg aggagaacct ggacttctgg aaccgcgagg 180
cagccgaggg cctgggtgcc gccaaagaagc tgcagcctgc acagacagcc gccaaagaacc 240

5/22

aagggcagaa gaaggacaaa ctggggcctg agatacccct ggccatggac cgcttcccat 360
atgtggctct gtccaagaca tacaatgtag acaaacatgt gccagacagt ggagccacag 420
ccacggccta cctgtgcggg gtcaagggca acttccagac cattggcttg agtgcagccg 480
cccgctttaa ccagtgcac acgacacgcg gcaacgaggt catctccgtg atgaatcggg 540
ccaagaaagc aggggaagtca gtgggagtggt taaccaccac acgagtgcag cacgcctcgc 600
cagccggcac ctacgcccac acggtgaacc gcaactggta ctcggaacgcc gacgtgcctg 660
cctcggccccg ccaggagggg tgccaggaca tcgctacgca 700

<210> 14

<211> 700

<212> DNA

<213> Calf intestinal alkaline phosphatase

<400> 14

gttaacgtgc tgaacagccg ggcatttttt tacgctatac cctacataat aaaaccggag 60
ctaccatgaa taagaaggta ctgacccttt ctgccgtgat ggcaagtctg ttattcgggg 120
cccacgcgca cgcggcgatc ccagctgagg aggaaaaccc cgccttcttg aaccgccagg 180
cagcccaggc ccttgatgta gccaagaagt tgcagccgat ccagacagct gccaagaatg 240
tcattcctctt cttgggggat gggatggggg tgcctacggt gacagccact cggatcctaa 300
aggggcagat gaatggcaaa ctgggacctg agacaccctt ggccatggac cagttcccat 360
acgtggctct gtccaagaca tacaacgtgg acagacaggt gccagacagc gcaggcactg 420
ccactgccta cctgtgtggg gtcaagggca actacagaac catcgggtgta agtgcagccg 480
cccgctacaa tcagtgcac acgacacgtg ggaatgaggt cacgtctgtg atcaaccggg 540
ccaagaaagc aggggaaggc gtgggagtggt tgaccaccac cagggtgcag catgcctccc 600
cagccggggc ctacgcgcac acggtgaacc gaaactggta ctcagacgcc gacctgcctg 660
ctgatgcaca gaagaatggc tgccaggaca tcgccgcaca 700

6/22

<210> 15

<211> 509

<212> DNA

<213> Human placental phosphatase and calf intestinal alkaline phosphatase

<400> 15

```

gttaacgtgc tgaacagccg ggcattcttt tacgctatac cctacataat aaaaccggag      60
ctaccatgaa taagaaggta ctgacccttt ctgccgtgat ggcaagtctg ttattcgggg      120
cccacgcgca cgcggcgatc ccagttgagg aggagaacct ggacttctgg aaccgcgagg      180
cagccaaggc cctgggtgcc gccagaagc tgcagcctgc acagacagcc gccagaacc      240
tcatcatctt cctgggcgat gggatggggg tgtctacggt gacagctgcc aggatcctaa      300
aagggcagaa gaaggacaaa ctggggcctg agataccctt ggccatggac cgcttcccat      360
atgtggctct gtccaagaca tacaatgtag acaaacatgt gccagacagt ggagccacag      420
ccacggccta cctgtgcggg gtcaagggca acttccagac cattggcttg agtgcagccg      480
cccgctttaa ccagtgcaac acgacacgc                                         509

```

<210> 16

<211> 445

<212> DNA

<213> Human placental phosphatase and calf intestinal alkaline phosphatase

<400> 16

```

gttaacgtgc tgaacagccg ggcatttttt tacgctatac cctacataat aaaaccggag      60
ctaccatgaa taagaaggta ctgacccttt ctgccgtgat ggcaagtctg ttattcggcg      120
cgcacgcgca cgcggcgatc ccagttgagg aggagaacct ggacttctgg aaccgcgagg      180
cagccaaggc cctgggtgcc gccagaagc tgcagcctgc acagacagcc gccagaacc      240

```

7/22

tcatcatctt cctgggcgat gggatggggg tgtctacggt gacagctgcc aggatcctaa 300
aagggcagaa gaaggacaaa ctggggcctg agatacccct ggccatggac cgcttcccat 360
atgtggctct gtccaagaca tacaatgtag acaaacatgt gccagacagt ggagccacag 420
ccacggccta cctgtgtggg gtcaa 445

<210> 17

<211> 624

<212> DNA

<213> Human placental phosphatase and calf intestinal alkaline phosphatase

<400> 17

gttaacgtgc tgaacagccg ggcatttttt tacgctatac cctacataat aaaaccggag 60
ctaccatgaa taagaaggta ctgacccttt ctgccgtgat gttaacgtgc tgaacagccg 120
ggcatttttt tacgctatac cctacataat aaaaccggag ctaccatgaa taagaaggta 180
ctgacccttt ctgccgtgat gccaagaagc tgcagcctgc acagacagcc gccaagaacc 240
tcgtcatctt cctgggcgat gggatggggg tgcctacggt gacagccact cggaccctaa 300
aagggcagaa gaaggacaaa ctggagcctg agatacccct ggccatggac cgcttcccat 360
atgtggctct gtccaagaca tacaacgtgg acagacaggt gccagacagt ggagccacag 420
ccacggccta cctgtgcggg gtcaagggca acttccagac cattggcttg agtgcagccg 480
cccgctttta ccagtgcaac acgacacgcg gcaacgaggt catctccgtg atgaatcggg 540
ccaagaaagc aggggaagtca gtgggagtggt taaccaccac acgagtgcag cacgcctcgc 600
cagccggcac ctacgcccac acgg 624

<210> 18

<211> 553

<212> DNA

8/22

<213> Human placental phosphatase and calf intestinal alkaline phosphatase

<400> 18

```

gttaacgtgc tgaacagccg ggcatttttt tacgctatac cctacataat aaaaccggag      60
ctaccatgaa taagaaggta ctgacccttt ctgccgtgat ggcaagtctg ttattcggcg      120
cgcacgcgca cgcggcgatc ccagttgagg aggagaacct ggacttctgg aaccgcgagg      180
cagccaaggc cctgggtgcc gccagaagc tgcagcctgc acagacagcc gccagaacc      240
tcatcatctt cctgggcat gggatggggg tgtctacggt gacagctgcc aggatcctaa      300
aagggcagaa gaaggacaaa ctggggcctg agataccctt ggccatggac cgcttcccat      360
atgtggctct gtccaagaca tacaatgtag acaaacatgt gccagacagt ggagccacag      420
ccacggccta cctgtgcggg gtcaaggga acttccagac cattggcttg agtgcagccg      480
cccgctttaa ccagtgaac acgacacgcg gcagcgaggt catctccgtg atgaaccggg      540
ccaagaaagc agg                                         553

```

<210> 19

<211> 599

<212> DNA

<213> Human placental phosphatase and calf intestinal alkaline phosphatase

<400> 19

```

gttaacgtgc tgaacagccg ggcatttttt tacgctatac cctacataat aaaaccggag      60
ctaccatgaa taagaaggta ttgacccttt ctgccgtgat ggcaagtctg ttattcggcg      120
cgcacgcgca cgcggcgatc ccagttgagg aggagaacct ggacttctgg aaccgcgagg      180
cagccaaggc cctgggtgcc gccagaagc tgcagcctgc acagacagcc gccagaacc      240
tcatcatctt cctgggcat gggatggggg tgtctacggt gacagctgcc aggatcctaa      300
aagggcagat gaatggcaaa ctgggacctg agacaccctt ggccatggac cagttcccat      360

```

9/22

acgtggctct gtccaagaca tacaatgtag acaaacatgt gccagacagt ggagccacag 420
 ccacggccta cctgtgcggg gtcaagggca acttccagac cattggcttg agtgcagccg 480
 cccgctttaa ccagtgcaac acgacacgcg gcaacgaggt catctccgtg atgaatcggg 540
 ccaagaaagc agggaagtca gtgggagtgg taaccaccac acgagtgcag cacgcctcg 599

<210> 20

<211> 528

<212> DNA

<213> Human placental phosphatase and calf intestinal alkaline

<400> 20

gttaccgtgc tgaacagccg ggcatttttt tacgctatac cctacataat aaaaccggag 60
 ctaccatgaa taagaaggta ctgacccttt ctgccgtgat ggcaagtctg ttattcgggg 120
 cccacgcgca cgcggcgatc ccagctgagg aggaaaaccc cgccttcttg aaccgccagg 180
 cagcccaggc ccttgatgta gccaagaagt tgcagccgat ccagacagct gccaagaatg 240
 tcacccctct cttgggggat gggatggggg tgtctacggt gacagctgcc aggatcctaa 300
 aagggcagaa gaaggacaaa ctggggcctg agatacccct ggccatggac cgcttcccat 360
 atgtggctct gtccaagaca tacaatgtag acaaacatat gccagacagt ggagccacag 420
 ccacggccta cctgtgcggg gtcaagggca acttccagac cattggcttg agtgcagccg 480
 cccgctttaa ccagtgcaac acgacacgcg gcaacgaggt catctccg 528

<210> 21

<211> 606

<212> DNA

<213> Human placental phosphatase

<400> 21

cttccctgga cccgtctgtg acccatctca tgggtctctt tgagcctgga gacatgaaat 60

10/22

```

acgagatcca ccgagactcc acactggacc cctccctgat ggagatgaca gaggctgccc 120
tgcgcctgct gagcaggaac ccccgcggtt tcttcctctt cgtggagggg ggtcgcatcg 180
accatgggtca tcatgaaagc agggcttacc gggcactgac tgagacgatc atgttcgacg 240
acgccattga gagggcgggc cagctcacca gcgaggagga cacgctgagc ctcgtcactg 300
ggtcctccta tacggaaacg gtccaggcta tgtgctcaag gacggcgccc ggccggatgt 360
taccgagagc gagagcggga gccccgagta tcggcagcag tcagcagtgc ccctggacga 420
agagaccac gcaggcgagg acgtggcggt gttcgcgcg gcgccgcagg cgcacctggt 480
tcacggcgtg caggagcaga ccttcatagc gcacgtcatg gccttcgccg cctgcctgga 540
gccctacacc gcctgcgacc tggcgcccc cgcggcacc accgactgat aaaagcttgg 600
tctaga 606

```

<210> 22

<211> 707

<212> DNA

<213> Calf intestinal alkaline phosphatase

<400> 22

```

cggccgatga ctccagtgtg acacacctca tgggcctctt tgagccggca gacatgaagt 60
ataatgttca gcaagaccac accaaggacc cgacctggc ggagatgacg gaggcggccc 120
tgcaagtgtc gagcaggaac ccccgggggt tctacctctt cgtggaggga ggccgcattg 180
accacgggtca ccatgacggc aaagcttata tggcactgac tgaggcgatc atgtttgaca 240
atgccatcgc caaggctaac gagctcacta gcgaactgga cacgctgatc cttgtcactg 300
cagaccactc ccatgtcttc tcctttggtg gctacacact gcgtgggacc tccattttcg 360
gtctggcccc cggcaaggcc ttagacagca agtcctacac ctccatcctc tatggcaatg 420
gcccaggcta tgcgcttggc gggggctcga ggcccgatgt taatggcagc acaagcgagg 480
aaccctcgta ccggcagcag gcggccgtgc ccctggctag cgagaccac gggggcgaag 540

```

11/22

acgtggcggt gttcgcgcga ggcccgcagg cgcacctggt gcacggcgtg caggaggaga 600
ccttcgtggc gcacatcatg gcctttgcgg gctgcgtgga gccctacacc gactgcaatc 660
tgccagcccc ctccaccgcc accagcatcc ccgactgata atctaga 707

<210> 23

<211> 763

<212> DNA

<213> Human placental phosphatase and calf intestinal alkaline phosphatase

<400> 23

gagatccacc gagactccac actggacccc tccttgatgg agatgacaga ggctgccctg 60
cgcctgctga gcaggaaccc ccgcggcttc ttctcttcg tggaggggtg tcgcatcgac 120
catggtcatc atgaaagcag ggcttaccgg gcactgactg ggacgatcat gttcgacgac 180
gacattgaga gggcggggcca gctcaccagc gaggaggaca cgctgagcct cgtcactgcc 240
gaccactccc acgtcttctc cttcggaggc taccacctgc gaggagctc catcttcggg 300
ctggcccctg gcaaggcccc ggacaggaag gcctacacgg tcctcctata cggaacgggt 360
ccaggctatg tgctcaagga cggcgcccgg ccggatgtta ccgagggcga gagcgggagc 420
cccgagtatc ggcagcagtc agcagtgcc ctggacgaag agaccacgc aggcgaggac 480
gtggcggtgt tcgcgcgcgg ccgcaggcg cacctggtgc acggcgtgca ggaggagacc 540
ttcgtggcgc acatcatggc ctttgcgggc tgcgtggagc cctacaccga ctgcaatctg 600
ccagccccct ccaccgccac cagcatcccc gactgataat ctagaggctc aaattcacct 660
cgaaagcaag ggcgaattct gcagatatcc atcacactgg cggccgctcg agcatgcatc 720
tagagggccc aattcgccct atagtgagtc gtattacaat tca 763

<210> 24

<211> 757

12/22

<212> DNA

<213> Human placental phosphatase and calf intestinal alkaline phosphatase

<400> 24

caccgagact ccacactgga cccctccctg atggagatga cagaggctgc cctgcgcctg	60
ctgaacagga acccccgcgg cttcttcctc tccgtggagg gtggtcgcat cgaccatggt	120
catcatgaaa gcagggctta ccgggcactg actgagacga tcatgttcga cgacgccatt	180
gagagggcgg gccagctcac cagcgaggag gacacgctga gcctcgtcac tgccgaccac	240
tcccacgtct tctccttcgg aggctacccc ctgcgaggga gctccatctt cgggctggcc	300
cctggcaagg cccgggacag gaaggcctac acggtcctcc tatacggaaa cgggccaggc	360
tatgtgctca aggacggcgc ccggccggat gttaccgaga gcgagagcgg gagccccgag	420
tatcggcagc agtcagcagt gcccctggac gaagagaccc acgcaggcga ggacgtggcg	480
gtgttcgcgc gcggcccgca ggcgcacctg gttcacggcg tgcaggagca gaccttcata	540
gcgcacgtca tggccttcgc cgcctgcctg gagccctaca ccgcctgcga cctggcgccc	600
cccgccggca ccaccgactg ataaaagttt ggctctagag gtcgaaattc acctcgaaag	660
caagggcgaa ttctgcagat atccatcaca ctggcggccg ctcgagcatg catctagagg	720
gcccatttcg ccctatagtg agtcgtatta caattca	757

<210> 25

<211> 756

<212> DNA

<213> Human placental phosphatase and calf intestinal alkaline phosphatase

<400> 25

accgagactc cacactggac ccctccctga tggagatgac agaggctgcc ctgcgcctgc	60
tgagcaggaa cccccgcggc ttcttcctct tcgtggaggg tggtcgcata gaccatggtc	120

13/22

atcatgaaag cagggcttac cgggcactga ctgagacgat catgttcgac gacgccatcg	180
agagggcggg ccagctcacc agcgaggagg acacgctgag cctcgtcact gccgaccact	240
cccacgtctt ctcccttcgga ggctaccccc tgcgagggag ctccatcttc gggctggccc	300
ctggcaaggc ccgggacagg aaggcctaca cggtcctcct atacggaaac ggtccaggct	360
atgtgctcaa ggacggcgcc cggccggatg ttaccgagag cgagagcggg agccccgagt	420
atcggcagca gccagcagtg cccctggacg aagagacca cgcaggcgag gacgtggcgg	480
tgttcgcgcg cggcccgcag gcgcacctgg ttacaggcgt gcaggagcag accttcatag	540
cgcacgtcat ggccttcgcc gcctgcctgg agccctacac cgcttgcgac ctggcgcccc	600
ccgccggcac caccgactga taaaagcttg gctctagagg tcgaaattca cctcgaaagc	660
aagggcgaat tctgcagata tccatcacac tggcggccgc tcgagcatgc atctagaggg	720
ccaattcgc cctatagtga gtcgtattac aattca	756

<210> 26

<211> 1650

<212> DNA

<213> Calf intestinal alkaline phosphatase

<400> 26

cttcggcggt cagtaacacg cgттаacgtg ctgaacagcc gggcattttt ttacgctata	60
ccctacataa taaaaccgga gctaccatga ataagaaggт actgaccctt tctgccgtga	120
tggcaagtct gttattcggg gcccacgcgc acgcggcgat cccagctgag gaggaaaacc	180
ccgccttctg gaaccgccag gcagcccagg cccttgatgt agccaagaag ttgcagccga	240
tccagacagc tgccaagaat gtcatcctct tcttggggga tgggatgggg gtgcctacgg	300
tgacagccac tcggatccta aaggggcaga tgaatggcaa actgggacct gagacacccc	360
tggccatgga ccagttccca tacgtggctc tgtccaagac atacaacgtg gacagacagg	420
tgccagacag cgcaggcact gccactgcct acctgtgtgg ggtcaagggc aactacagaa	480

14/22

ccatcggtgt aagtgcagcc gcccgcctaca atcagtgcaa cacgacacgt gggaatgagg 540
tcacgtctgt gatcaaccgg gccaaagaaag caggggaaggc cgtgggagtg gtgaccacca 600
ccaggggtgca gcatgcctcc ccagccggggg cctacgcgca cacggtgaac cgaaactggt 660
actcagacgc cgacctgcct gctgatgcac agaagaatgg ctgccaggac atcgccgcac 720
agctgggtcta caacatggat attgacgtga tcctgggttg aggccgaatg tacatgtttc 780
ctgagggggac cccagaccct gaatacccag atgatgccag tgtgaatgga gtccggaagg 840
acaagcagaa cctgggtgcag gaatggcagg ccaagcacca gggagcccag tatgtgtgga 900
accgcactgc gtccttcag gcggccgatg actccagtgt aacacacctc atgggcctct 960
ttgagccggc agacatgaag tataatgttc agcaagacca caccaaggac ccgaccctgg 1020
cggagatgac ggaggcggcc ctgcaagtgc tgagcaggaa ccccgggggc ttctacctct 1080
tcgtggaggg aggccgcatt gaccacggtc accatgacgg caaagcttat atggcactga 1140
ctgaggcgat catgtttgac aatgccatcg ccaaggctaa cgagctcact agcgaactgg 1200
acacgctgat ccttgtcact gcagaccact cccatgtctt ctcccttggt ggctacacac 1260
tgcgtgggac ctccattttc ggtctggccc ccggcaaggc cttagacagc aagtcctaca 1320
cctccatcct ctatggcaat ggcccaggct atgcgcttgg cgggggctcg aggcccgatg 1380
ttaatggcag cacaagcgag gaaccctcgt accggcagca ggcggccgtg cccctggcta 1440
gcgagacca cggggggcgaa gacgtggcgg tgttcgcgcg aggcccgcag gcgcacctgg 1500
tgcacggcgt gcaggaggag accttcgtgg cgcacatcat ggcctttgcg ggctgcgtgg 1560
agccctacac cgactgcaat ctgccagccc cctccaccgc caccagcatc cccgactgat 1620
aatctagagg tcgaaattca cctcgaaagc 1650

<210> 27

<211> 1649

<212> DNA

<213> Human placental phosphatase

15/22

<400> 27

cttcggcggtt cagtaacacg cgттаacgtg ctgaacagcc gggcattttt ttacgctata	60
ccctacataa taaaaccgga gctaccatga ataagaaggt actgaccctt tctgccgtga	120
tggcaagtct gttattcggc gcgcacgcgc acgcggcgat cccagttgag gaggagaacc	180
cggacttctg gaaccgcgag gcagccgagg ccctgggtgc cgccaagaag ctgcagcctg	240
cacagacagc cgccaagaac ctcatcatct tcctggggcga tgggatgggg gtgtctacgg	300
tgacagctgc caggatccta aaagggcaga agaaggacaa actggggcct gagatacccc	360
tggccatgga ccgcttccca tatgtggctc tgtccaagac atacaatgta gacaaacatg	420
tgccagacag tggagccaca gccacggcct acctgtgcgg ggtcaagggc aacttccaga	480
ccattggctt gagtgcagcc gcccgcttta accagtgcaa cacgacacgc ggcaacgagg	540
tcattctccgt gatgaatcgg gccaaagaaag caggggaagtc agtgggagtg gtaaccacca	600
cacgagtgca gcacgcctcg ccagccggca cctacgcca cacggtgaac cgcaactggt	660
actcggacgc cgacgtgcct gcctcggccc gccaggaggg gtgccaggac atcgctacgc	720
agctcatctc caacatggac attgacgtga tcctaggtgg aggccgaaag tacatgtttc	780
ccatgggaac cccagaccct gagtaccag atgactacag ccaaggtggg accaggctgg	840
acgggaagaa tctggtgcag gaatggctgg cgaagcgcca gggtgcccgg tatgtgtgga	900
accgcactga gctcatgcag gcttccctgg acccgtctgt gacctatctc atgggtctct	960
ttgagcctgg agacatgaaa tacgagatcc accgagactc cacactggac ccctccctga	1020
tggagatgac agaggctgcc ctgcgctgc tgagcaggaa ccccgcggc ttcttcctct	1080
tcgtggaggg tggtcgcac gaccatggtc atcatgaaag cagggttac cgggcactga	1140
ctgagacgat catgttcgac gacgccattg agagggcggg ccagctcacc agcgaggagg	1200
acacgctgag cctcgtcact gccgaccact cccacgtctt ctcttcgga ggctacccc	1260
tgcgagggag ctccatcttc gggctggccc ctggcaaggc ccgggacagg aaggcctaca	1320
cggtcctcct atacggaaac ggtccaggct atgtgctcaa ggacggcgcc cggccggatg	1380
ttaccgagag cgagagcggg agccccgagt atcggcagca gtcagcagtg cccttgga cg	1440

16/22

aagagaccca cgcaggcgag gacgtggcgg tgttcgcgcg cggcccgcag gcgcacctgg 1500
ttcacggcgt gcaggagcag accttcatag cgcacgtcat ggccttcgcc gcctgcctgg 1560
agccctacac cgctgcgac ctggcgcccc ccgccggcac caccgactga taaaagcttg 1620
gtctagaggt cgaaattcac ctcgaaagc 1649

<210> 28

<211> 530

<212> DNA

<213> Calf intestinal alkaline phosphatase and human placental phosphatase

<400> 28

cttcggcgtt cagtaacacg cgttaacgtg ctgaacagcc gggcattctt ttacgctata 60
ccctacataa taaaaccgga gctaccatga ataagaaggt actgaccctt tctgccgtga 120
tggcaagtct gttattcggg gccacgcgc acgcggcgat ccagttgag gaggagaacc 180
cggacttctg gaaccgcgag gcagccaagg ccctgggtgc cgccaagaag ctgcagcctg 240
cacagacagc cgccaagaac ctcatcatct tcctgggcga tgggatgggg gtgtctacgg 300
tgacagctgc caggatccta aaagggcaga agaaggacaa actggggcct gagatacccc 360
tggccatgga ccgcttccca tatgtggctc tgtccaagac atacaatgta gacaaacatg 420
tgccagacag tggagccaca gccacggcct acctgtgcgg ggtcaagggc aacttccaga 480
ccattggctt gagtgcagcc gcccgcttta accagtgcaa cacgacacgc 530

<210> 29

<211> 466

<212> DNA

<213> Calf intestinal alkaline phosphatase and human placental phosphatase

17/22

<400> 29
 cttcggcggtt cagtaacacg cgttaacgtg ctgaacagcc gggcattttt ttacgctata 60
 ccctacataa taaaaccgga gctaccatga ataagaaggt actgaccctt tctgccgtga 120
 tggcaagtct gttattcggc gcgcacgcgc acgcggcgat ccagttgag gaggagaacc 180
 cggacttctg gaaccgcgag gcagccaagg ccctgggtgc cgccaagaag ctgcagcctg 240
 cacagacagc cgccaagaac ctcatcatct tcctggggcga tgggatgggg gtgtctacgg 300
 tgacagctgc caggatccta aaagggcaga agaaggacaa actggggcct gagatacccc 360
 tggccatgga ccgcttccca tatgtggctc tgtccaagac atacaatgta gacaaacatg 420
 tgccagacag tggagccaca gccacggcct acctgtgtgg ggtcaa 466

<210> 30

<211> 645

<212> DNA

<213> Calf intestinal alkaline phosphatase and human placental phosphatase

<400> 30
 cttcggcggtt cagtaacacg cgttaacgtg ctgaacagcc gggcattttt ttacgctata 60
 ccctacataa taaaaccgga gctaccatga ataagaaggt actgaccctt tctgccgtga 120
 tggcaagtct gttattcggc gcgcacgcgc acgcggcgat ccagttgag gaggagaacc 180
 cggacttctg gaaccgcgag gcagccaagg ccctgggtgc cgccaagaag ctgcagcctg 240
 cacagacagc cgccaagaac ctcgatcatct tcctggggcga tgggatgggg gtgcctacgg 300
 tgacagccac tcggacccta aaagggcaga agaaggacaa actggagcct gagatacccc 360
 tggccatgga ccgcttccca tatgtggctc tgtccaagac atacaacgtg gacagacagg 420
 tgccagacag tggagccaca gccacggcct acctgtgcgg ggtcaagggc aacttccaga 480
 ccattggctt gaggtcagcc gcccgcttta accagtgcaa cacgacacgc ggcaacgagg 540
 tcattctcgt gatgaatcgg gccagaagaag cagggaagtc agtgggagtg gtaaccacca 600

18/22

cacgagtgcg gcacgcctcg ccagccggca cctacgcca cacgg 645

<210> 31

<211> 574

<212> DNA

<213> Calf intestinal alkaline phosphatase and human placental phosphatase

<400> 31

cttcggcggt cagtaacacg cgттаacgtg ctgaacagcc gggcattttt ttacgctata 60
ccctacataa taaaaccgga gctaccatga ataagaaggt actgaccctt tctgccgtga 120
tggcaagtct gttattcggc gcgcacgcgc acgcggcgat ccagttgag gaggagaacc 180
cggacttctg gaaccgcgag gcagccaagg ccctgggtgc cgccaagaag ctgcagcctg 240
cacagacagc cgccaagaac ctcatcatct tcctgggcga tgggatgggg gtgtctacgg 300
tgacagctgc caggatccta aaagggcaga agaaggacaa actggggcct gagatacccc 360
tggccatgga ccgcttccca tatgtggctc tgtccaagac atacaatgta gacaaacatg 420
tgccagacag tggagccaca gccacggcct acctgtgcgg ggtcaagggc aacttccaga 480
ccattggctt gagtgcagcc gcccgcttta accagtgcaa cacgacacgc ggcagcgagg 540
tcattctcgt gatgaaccgg gcccaagaaag cagg 574

<210> 32

<211> 620

<212> DNA

<213> Calf intestinal alkaline phosphatase and human placental phosphatase

<400> 32

cttcggcggt cagtaacacg cgттаacgtg ctgaacagcc gggcattttt ttacgctata 60
ccctacataa taaaaccgga gctaccatga ataagaaggt attgaccctt tctgccgtga 120

19/22

tggcaagtct gttattcggc gcgcacgcgc acgcggcgat cccagttgag gaggagaacc	180
cggacttctg gaaccgcgag gcagccaagg ccctgggtgc cgccaagaag ctgcagcctg	240
cacagacagc cgccaagaac ctcatcatct tcctggggcga tgggatgggg gtgtctacgg	300
tgacagctgc caggatccta aaagggcaga tgaatggcaa actgggacct gagacacccc	360
tggccatgga ccagttccca tacgtggctc tgtccaagac atacaatgta gacaaacatg	420
tgccagacag tggagccaca gccacggcct acctgtgcgg ggtcaagggc aacttccaga	480
ccattggctt gagtgcagcc gcccgcttta accagtgcaa cacgacacgc ggcaacgagg	540
tcattctccg gatgaatcgg gcccaagaaag cagggaagtc agtgggagtg gtaaccacca	600
cacgagtgca gcacgcctcg	620

<210> 33

<211> 549

<212> DNA

<213> Calf intestinal alkaline phosphatase and human placental phosphatase

<400> 33

cttcggcggt cagtaacacg cgttaccgtg ctgaacagcc gggcattttt ttacgctata	60
ccctacataa taaaaccgga gctaccatga ataagaaggt actgaccctt tctgccgtga	120
tggcaagtct gttattcggg gccacgcgc acgcggcgat cccagctgag gaggaaaacc	180
ccgccttctg gaaccgccag gcagcccagg cccttgatgt agccaagaag ttgcagccga	240
tccagacagc tgccaagaat gtcattcctt tcttggggga tgggatgggg gtgtctacgg	300
tgacagctgc caggatccta aaagggcaga agaaggacaa actggggcct gagatacccc	360
tggccatgga ccgcttccca tatgtggctc tgtccaagac atacaatgta gacaaacata	420
tgccagacag tggagccaca gccacggcct acctgtgcgg ggtcaagggc aacttccaga	480
ccattggctt gagtgcagcc gcccgcttta accagtgcaa cacgacacgc ggcaacgagg	540
tcattctccg	549

20/22

<210> 34

<211> 667

<212> DNA

<213> Calf intestinal alkaline phosphatase and human placental phosphatase

<400> 34

gagatccacc gagactccac actggacccc tccctgatgg agatgacaga ggctgccctg	60
cgcctgctga gcaggaaccc ccgcggcttc ttcctcttcg tggaggggtgg tcgcatcgac	120
catggtcatc atgaaagcag ggcttaccgg gcactgactg ggacgatcat gttcgacgac	180
gacattgaga gggcggggcca gctcaccagc gaggaggaca cgctgagcct cgtcactgcc	240
gaccactccc acgtcttctc cttcggaggc taccctctgc gagggagctc catcttcggg	300
ctggcccctg gcaaggcccg ggacaggaag gcctacacgg tcctcctata cggaaacggt	360
ccaggctatg tgc tcaagga cggcgcccgg ccggatgtta ccgagggcga gagcgggagc	420
cccgagtatc ggcagcagtc agcagtgcc ctggacgaag agaccacgc aggcgaggac	480
gtggcggtgt tcgcgcgcgg cccgcaggcg cacctgggtgc acggcgtgca ggaggagacc	540
ttcgtggcgc acatcatggc ctttgcgggc tgcgtggagc cctacaccga ctgcaatctg	600
ccagccccct ccaccgccac cagcatcccc gactgataat ctagaggtcg aaattcacct	660
cgaaagc	667

<210> 35

<211> 661

<212> DNA

<213> Calf intestinal alkaline phosphatase and human placental phosphatase

21/22

<400> 35

caccgagact ccacactgga ccctccctg atggagatga cagaggctgc cctgcgcctg	60
ctgaacagga acccccgcggtttcttctc tccgtggagg gtggtcgcat cgaccatggt	120
catcatgaaa gcagggctta ccgggcactg actgagacga tcatgttcga cgacgccatt	180
gagagggcggt gccagctcac cagcgaggag gacacgctga gcctcgtcac tgccgaccac	240
tcccacgtct tctccttcggtaggctacccc ctgcgaggga gctccatctt cgggctggcc	300
cctggcaagg ccggggacag gaaggcctac acggtcctcc tatacggaac cggtccaggc	360
tatgtgctca aggacggcgc ccggccggat gttaccgaga gcgagagcgg gagccccgag	420
tatcggcagc agtcagcagt gcccctggac gaagagaccc acgcaggcga ggacgtggcg	480
gtgttcgcgc gcggcccgca ggcgacactg gttcacggcg tgcaggagca gaccttcata	540
gcgcacgtca tggccttcgc cgctgcctg gagccctaca ccgcctgcga cctggcgccc	600
cccgccggca ccaccgactg ataaaagttt ggctctagag gtcgaaattc acctcgaaag	660
c	661

<210> 36

<211> 660

<212> DNA

<213> Calf intestinal alkaline phosphatase and human placental phosphatase

<400> 36

accgagactc cacactggac ccctccctga tggagatgac agaggctgcc ctgcgcctgc	60
tgagcaggaa ccccgcggttcttctcctc tccgtggagg tggtcgcata gaccatgggtc	120
atcatgaaag cagggcttac ccgggcactga ctgagacgat catgttcgac gacgccatcg	180
agagggcggtt ccagctcacc agcgaggagg acacgctgag cctcgtcact gccgaccact	240
cccacgtctt ctcttcgga ggctaccccc tgcgaggagg ctccatcttc gggctggccc	300
ctggcaaggc ccggggacagg aaggcctaca cggtcctcct atacggaaac ggtccaggct	360

22/22

atgtgctcaa ggacggcgcc cggccggatg ttaccgagag cgagagcggg agccccgagt	420
atcggcagca gccagcagtg cccctggacg aagagacca cgcaggcgag gacgtggcgg	480
tgttcgcgcg cggcccgcag gcgcacctgg ttcacggcgt gcaggagcag accttcatag	540
cgcacgtcat ggccttcgcc gcctgcctgg agccctacac cgctgcgac ctggcgcccc	600
ccgccggcac caccgactga taaaagcttg gctctagagg tcgaaattca cctcgaaagc	660

Claims

- 1) Method for forming a polynucleotide sequence comprising the steps of
 - a) generating a nucleic acid fragment ladder by nucleic acid synthesis which is carried out in presence of a reaction mixture comprising at least one primer, an enzyme having nucleic acid-synthesizing activity and a mixture of nucleic acid chain-terminating and nucleic acid chain-elongating molecules and a template polynucleotide,
 - b) removing the chain-terminating molecules or changing it into non-terminating molecules,
 - c) reassembly of the polynucleotide by hybridizing these fragments to one another in presence of a thermostable enzyme having nucleic acid-synthesizing activity and conducting nucleic acid synthesis either
 - ii) in presence of a mixture of nucleic acid chain-terminating molecules and nucleic acid chain elongating molecules or
 - ii) in presence of a reaction mixture of nucleic acid chain elongating molecules which does not contain chain-terminating moleculeswhereas in case of i) steps b) and c) will be repeated subsequently.
- 2) Method for forming a polynucleotide sequence according to claim 1 whereas the template polynucleotide in step a) and/or c) is a mixture consisting of different template polynucleotides.
- 3) Method for forming a polynucleotide sequence according to claim 1 or 2 wherein the nucleic acid chain-terminating molecule is a ddNTP and the nucleic acid elongating molecule is a dNTP.
- 4) Method for forming a polynucleotide sequence according to one of the claims 1-3 wherein the enzyme having nucleic-acid-synthesizing activity is a DNA polymerase and the template polynucleotide is DNA.

- 5) Method for forming a polynucleotide sequence according to one of the claims 1-4 wherein the enzyme having nucleic acid-synthesizing activity is a RNA-dependent DNA polymerase and the template polynucleotide is RNA.
- 6) A polynucleotide sequence obtained by a method according to any one of the claims 1-5.
- 7) Method for providing mutant polypeptides or proteins comprising the steps according to any one of the claims 1-5 and further comprising the steps of
 - cloning procedures - as cloning the amplified product into a vector and expression of the encoded protein.
 - screening or selection procedures.
- 8) Method for providing mutant polypeptides or proteins comprising the steps according to claim 7 additionally comprising an amplification step of reassembled product
- 9) Mutant polypeptide or protein obtained by a method according to claim 7 or 8.
- 10) Use of a method according to any of the claims 1 to 5 and 7 to 8 for optimizing biomolecules.
- 11) Use of a method according to any of the claims 1 to 5 and 7 to 8 for optimizing enzymes.

Synthesis and/or chain reaction

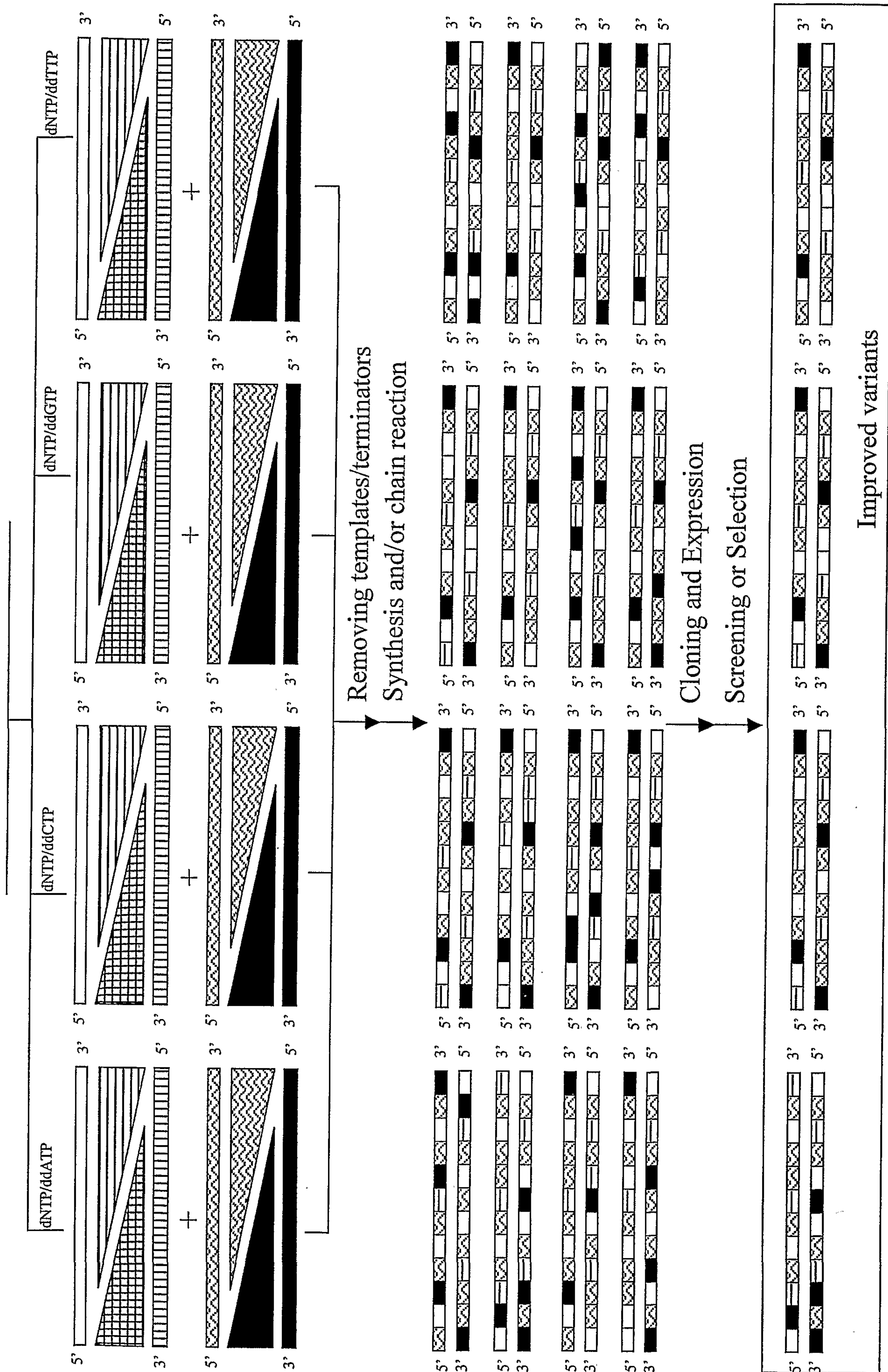


Figure 2:

Primer arrangement along human placental alkaline phosphatase (hpap) and calf intestinal alkaline phosphatase (ciap) genes

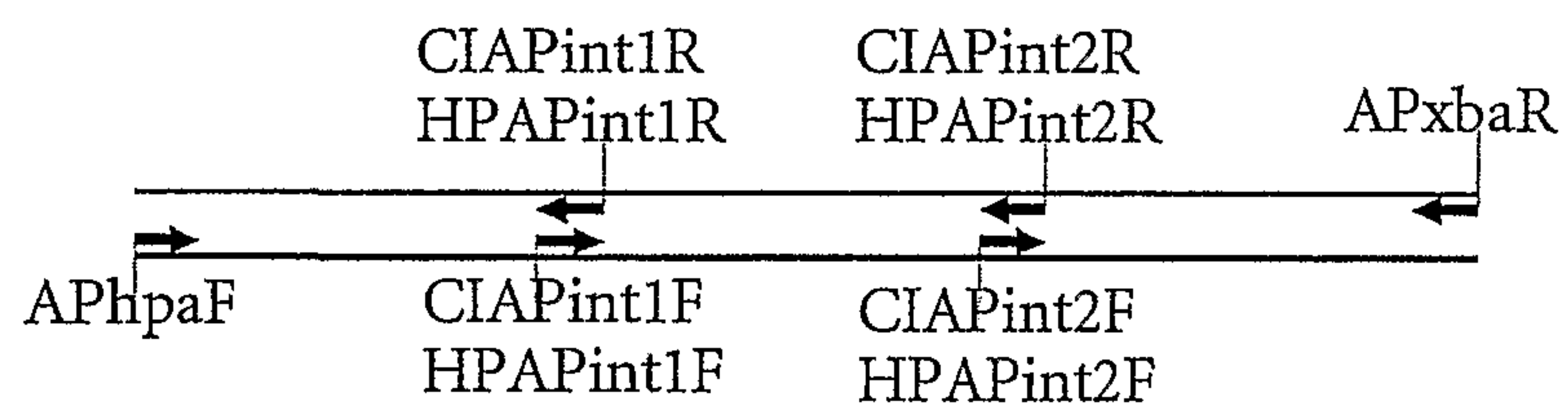


Figure 3 1/4

The N-terminal sequence alignment of parental ciap and hpap genes and their recombination variants (AP01, AP03, AP05, AP06, AP11 and AP15).

	10	20	30	40	50	
CIAP :	CTTCGGCGTTCAGTAACACGCGTTA	CGTGCTGAACAGCCGGGCATT	TT	:	50	
HPAP :	CTTCGGCGTTCAGTAACACGCGTTA	CGTGCTGAACAGCCGGGCATT	TT	:	50	
AP01 :	CTTCGGCGTTCAGTAACACGCGTTA	CGTGCTGAACAGCCGGGCATT	CTT	:	50	
AP03 :	CTTCGGCGTTCAGTAACACGCGTTA	CGTGCTGAACAGCCGGGCATT	TT	:	50	
AP05 :	CTTCGGCGTTCAGTAACACGCGTTA	CGTGCTGAACAGCCGGGCATT	TT	:	50	
AP06 :	CTTCGGCGTTCAGTAACACGCGTTA	CGTGCTGAACAGCCGGGCATT	TT	:	50	
AP11 :	CTTCGGCGTTCAGTAACACGCGTTA	CGTGCTGAACAGCCGGGCATT	TT	:	50	
AP15 :	CTTCGGCGTTCAGTAACACGCGTTA	CGTGCTGAACAGCCGGGCATT	TT	:	50	

	60	70	80	90	100	
CIAP :	TTACGCTATACCCTACATAATAAAACCGGAGCTACCATGAATAAGAAGGT	:	100			
HPAP :	TTACGCTATACCCTACATAATAAAACCGGAGCTACCATGAATAAGAAGGT	:	100			
AP01 :	TTACGCTATACCCTACATAATAAAACCGGAGCTACCATGAATAAGAAGGT	:	100			
AP03 :	TTACGCTATACCCTACATAATAAAACCGGAGCTACCATGAATAAGAAGGT	:	100			
AP05 :	TTACGCTATACCCTACATAATAAAACCGGAGCTACCATGAATAAGAAGGT	:	100			
AP06 :	TTACGCTATACCCTACATAATAAAACCGGAGCTACCATGAATAAGAAGGT	:	100			
AP11 :	TTACGCTATACCCTACATAATAAAACCGGAGCTACCATGAATAAGAAGGT	:	100			
AP15 :	TTACGCTATACCCTACATAATAAAACCGGAGCTACCATGAATAAGAAGGT	:	100			

	110	120	130	140	150	
CIAP :	ACTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTTCGGGGGCC	CACGCGC	:	150		
HPAP :	ACTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTTCGGGGGCC	CACGCGC	:	150		
AP01 :	ACTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTTCGGGGGCC	CACGCGC	:	150		
AP03 :	ACTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTTCGGGGGCC	CACGCGC	:	150		
AP05 :	ACTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTTCGGGGGCC	CACGCGC	:	150		
AP06 :	ACTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTTCGGGGGCC	CACGCGC	:	150		
AP11 :	ATTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTTCGGGGGCC	CACGCGC	:	150		
AP15 :	ACTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTTCGGGGGCC	CACGCGC	:	150		

	160	170	180	190	200	
CIAP :	ACGCGGCGATCCCAGCTGAGGAGGAGAAACCCCGC	CTTCTGGAACCGCCAG	:	200		
HPAP :	ACGCGGCGATCCCAGCTGAGGAGGAGAAACCCCGC	CTTCTGGAACCGCCAG	:	200		
AP01 :	ACGCGGCGATCCCAGCTGAGGAGGAGAAACCCCGC	CTTCTGGAACCGCCAG	:	200		
AP03 :	ACGCGGCGATCCCAGCTGAGGAGGAGAAACCCCGC	CTTCTGGAACCGCCAG	:	200		
AP05 :	ACGCGGCGATCCCAGCTGAGGAGGAGAAACCCCGC	CTTCTGGAACCGCCAG	:	200		
AP06 :	ACGCGGCGATCCCAGCTGAGGAGGAGAAACCCCGC	CTTCTGGAACCGCCAG	:	200		
AP11 :	ACGCGGCGATCCCAGCTGAGGAGGAGAAACCCCGC	CTTCTGGAACCGCCAG	:	200		
AP15 :	ACGCGGCGATCCCAGCTGAGGAGGAGAAACCCCGC	CTTCTGGAACCGCCAG	:	200		

Figure 3 2/4

	210	220	230	240	250	
CIAP :	GCAGCC	CAGGCCCTTG	ATGTAGCCAAGAAGT	TGCAGCCGATC	CAGACAGC	: 250
HPAP :	GCAGCC	CAGGCCCTG	GTGCCCAAGAAGC	TGCAGCCGATC	CAGACAGC	: 250
AP01 :	GCAGCC	CAGGCCCTG	GTGCCCAAGAAGC	TGCAGCCGATC	CAGACAGC	: 250
AP03 :	GCAGCC	CAGGCCCTG	GTGCCCAAGAAGC	TGCAGCCGATC	CAGACAGC	: 250
AP05 :	GCAGCC	CAGGCCCTG	GTGCCCAAGAAGC	TGCAGCCGATC	CAGACAGC	: 250
AP06 :	GCAGCC	CAGGCCCTG	GTGCCCAAGAAGC	TGCAGCCGATC	CAGACAGC	: 250
AP11 :	GCAGCC	CAGGCCCTG	GTGCCCAAGAAGC	TGCAGCCGATC	CAGACAGC	: 250
AP15 :	GCAGCC	CAGGCCCTTG	ATGTAGCCAAGAAGT	TGCAGCCGATC	CAGACAGC	: 250

	260	270	280	290	300	
CIAP :	TGCCAAGAATGTC	ATCCTCTTCT	TGGGGG	GATGGGATGGGGGTG	CTACGG	: 300
HPAP :	TGCCAAGAACTC	ATCCTCTTCT	TGGGGG	GATGGGATGGGGGTG	CTACGG	: 300
AP01 :	TGCCAAGAACTC	ATCCTCTTCT	TGGGGG	GATGGGATGGGGGTG	CTACGG	: 300
AP03 :	TGCCAAGAACTC	ATCCTCTTCT	TGGGGG	GATGGGATGGGGGTG	CTACGG	: 300
AP05 :	TGCCAAGAACTC	ATCCTCTTCT	TGGGGG	GATGGGATGGGGGTG	CTACGG	: 300
AP06 :	TGCCAAGAACTC	ATCCTCTTCT	TGGGGG	GATGGGATGGGGGTG	CTACGG	: 300
AP11 :	TGCCAAGAACTC	ATCCTCTTCT	TGGGGG	GATGGGATGGGGGTG	CTACGG	: 300
AP15 :	TGCCAAGAATGTC	ATCCTCTTCT	TGGGGG	GATGGGATGGGGGTG	CTACGG	: 300

	310	320	330	340	350	
CIAP :	TGACAGCCACTC	GGATCCTAAAG	GGGCAGATGAAT	TGGCAAACCTGG	CACT	: 350
HPAP :	TGACAGCCACTC	GGATCCTAAA	GGGCAGAGAA	TGGCAAACCTGG	CACT	: 350
AP01 :	TGACAGCCACTC	GGATCCTAAA	GGGCAGAGAA	TGGCAAACCTGG	CACT	: 350
AP03 :	TGACAGCCACTC	GGATCCTAAA	GGGCAGAGAA	TGGCAAACCTGG	CACT	: 350
AP05 :	TGACAGCCACTC	GGATCCTAAA	GGGCAGAGAA	TGGCAAACCTGG	CACT	: 350
AP06 :	TGACAGCCACTC	GGATCCTAAA	GGGCAGAGAA	TGGCAAACCTGG	CACT	: 350
AP11 :	TGACAGCCACTC	GGATCCTAAA	GGGCAGATGAAT	TGGCAAACCTGG	CACT	: 350
AP15 :	TGACAGCCACTC	GGATCCTAAA	GGGCAGAGAA	TGGCAAACCTGG	CACT	: 350

	360	370	380	390	400	
CIAP :	GAGACACCCCTGGCCATGGACCAG	TTCCCATA	GTGGCTCTGTCCAAGAC			: 400
HPAP :	GAGACACCCCTGGCCATGGACCAG	TTCCCATA	GTGGCTCTGTCCAAGAC			: 400
AP01 :	GAGACACCCCTGGCCATGGACCAG	TTCCCATA	GTGGCTCTGTCCAAGAC			: 400
AP03 :	GAGACACCCCTGGCCATGGACCAG	TTCCCATA	GTGGCTCTGTCCAAGAC			: 400
AP05 :	GAGACACCCCTGGCCATGGACCAG	TTCCCATA	GTGGCTCTGTCCAAGAC			: 400
AP06 :	GAGACACCCCTGGCCATGGACCAG	TTCCCATA	GTGGCTCTGTCCAAGAC			: 400
AP11 :	GAGACACCCCTGGCCATGGACCAG	TTCCCATA	GTGGCTCTGTCCAAGAC			: 400
AP15 :	GAGACACCCCTGGCCATGGACCAG	TTCCCATA	GTGGCTCTGTCCAAGAC			: 400

	410	420	430	440	450	
CIAP :	ATACAACTGTG	GACAGACAGC	TGCCAGACAGC	GCAGGCACTGCCACTGCCT		: 450
HPAP :	ATACAACTGTG	GACAAACATC	TGCCAGACAGC	GCAGGCACTGCCACTGCCT		: 450
AP01 :	ATACAACTGTG	GACAAACATC	TGCCAGACAGC	GCAGGCACTGCCACTGCCT		: 450
AP03 :	ATACAACTGTG	GACAAACATC	TGCCAGACAGC	GCAGGCACTGCCACTGCCT		: 450

Figure 3 3/4

```

AP05 : ATACAACGTGGACAGACAGCTGCCAGACAGTGCAGCACAGCCACCGCCT : 450
AP06 : ATACAATGTAGACAAACATCTGCCAGACAGTGCAGCACAGCCACCGCCT : 450
AP11 : ATACAATGTAGACAAACATCTGCCAGACAGTGCAGCACAGCCACCGCCT : 450
AP15 : ATACAATGTAGACAAACATATGCCAGACAGTGCAGCACAGCCACCGCCT : 450

```

```

          460      470      480      490      500
CIAP : ACCTGTGTGGGGTCAAAGGCGAAGTTGACAGACCAATTCCTGTAACTCCAGCC : 500
HPAP : ACCTGTGTGGGGTCAAAGGCGAAGTTGACAGACCAATTCCTGTAACTCCAGCC : 500
AP01 : ACCTGTGTGGGGTCAAAGGCGAAGTTGACAGACCAATTCCTGTAACTCCAGCC : 500
AP03 : ACCTGTGTGGGGTCAA----- : 466
AP05 : ACCTGTGTGGGGTCAAAGGCGAAGTTGACAGACCAATTCCTGTAACTCCAGCC : 500
AP06 : ACCTGTGTGGGGTCAAAGGCGAAGTTGACAGACCAATTCCTGTAACTCCAGCC : 500
AP11 : ACCTGTGTGGGGTCAAAGGCGAAGTTGACAGACCAATTCCTGTAACTCCAGCC : 500
AP15 : ACCTGTGTGGGGTCAAAGGCGAAGTTGACAGACCAATTCCTGTAACTCCAGCC : 500

```

```

          510      520      530      540      550
CIAP : CCGCGCTTAAACCACTGCAAGACAGGACAGCTCGGAATCAGGTCACGTCGTCT : 550
HPAP : CCGCGCTTAAACCACTGCAAGACAGGACAGCTCGGAATCAGGTCACGTCGTCT : 550
AP01 : CCGCGCTTAAACCACTGCAAGACAGGACAGCTCGGAATCAGGTCACGTCGTCT : 530
AP03 : ----- : -
AP05 : CCGCGCTTAAACCACTGCAAGACAGGACAGCTCGGAATCAGGTCACGTCGTCT : 550
AP06 : CCGCGCTTAAACCACTGCAAGACAGGACAGCTCGGAATCAGGTCACGTCGTCT : 550
AP11 : CCGCGCTTAAACCACTGCAAGACAGGACAGCTCGGAATCAGGTCACGTCGTCT : 550
AP15 : CCGCGCTTAAACCACTGCAAGACAGGACAGCTCGGAATCAGGTCACGTCGTCT : 549

```

```

          560      570      580      590      600
CIAP : GATCAACCGGGCCAAGAAAGCAGGGAAGGCCGTGGGAGTGGTGACCACCA : 600
HPAP : GATGAATCGGGCCAAGAAAGCAGGGAAGTCAGTGGGAGTGGTAACCACCA : 600
AP01 : ----- : -
AP03 : ----- : -
AP05 : GATGAATCGGGCCAAGAAAGCAGGGAAGTCAGTGGGAGTGGTAACCACCA : 600
AP06 : GATGAACCGGGCCAAGAAAGCAGG----- : 574
AP11 : GATGAATCGGGCCAAGAAAGCAGGGAAGTCAGTGGGAGTGGTAACCACCA : 600
AP15 : ----- : -

```

```

          610      620      630      640      650
CIAP : CCAGGGTGCAGCATGCCTCCCCAGCCGGGGCCTACGCGCACACGGTGAAC : 650
HPAP : CACGAGTGCAGCACGCCTCGCCAGCCGGCACCTACGCCCACACGGTGAAC : 650
AP01 : ----- : -
AP03 : ----- : -
AP05 : CACGAGTGCAGCACGCCTCGCCAGCCGGCACCTACGCCCACACGG----- : 645
AP06 : ----- : -
AP11 : CACGAGTGCAGCACGCCTCG----- : 620
AP15 : ----- : -

```

Figure 3 4/4

	660	670	680	
CIAP	: CGAAACTGGTACTCAGACGCCGACCTGCCTGC	:	682	
HPAP	: CGCAACTGGTACTCGGACGCCGACGTGCCTGC	:	682	
AP01	: -----	:	-	
AP03	: -----	:	-	
AP05	: -----	:	-	
AP06	: -----	:	-	
AP11	: -----	:	-	
AP15	: -----	:	-	

CIAP:	SEQ	ID	NO:	14
HPAP:	SEQ	ID	NO:	13
AP01:	SEQ	ID	NO:	15
AP03:	SEQ	ID	NO:	16
AP05:	SEQ	ID	NO:	17
AP06:	SEQ	ID	NO:	18
AP11:	SEQ	ID	NO:	19
AP15:	SEQ	ID	NO:	20

Figure 4 1/3

The C-terminal sequence alignment of parental parental ciap and hpap genes and their recombination variants (AP2, AP9 and AP13)

	960	970	980	990	1000	
CIAP	: ATGGGCCTCTTTGAGCCGGCAGACATGAAGTATAATGTTT				CAGCAAGACCA	: 1000
HPAP	: ATGGGTCTCTTTGAGCCTGGAGACATGAAATACGAGATCC				CAACCAGACTTC	: 1000
AP2	: -----				GAGATCC	CAACCAGACTTC : 17
AP9	: -----				CAACCAGACTTC	: 11
AP13	: -----				CAACCAGACTTC	: 10
	1010	1020	1030	1040	1050	
CIAP	: CACCAAGGACCCGACCCCTGGCGGAGATGACGGAGGCGGCCCTGCAAGTGC					: 1050
HPAP	: CACACTGGACCCGACCCCTGGATGGAGATGACAGAGGCTGCCCTGCGGCTGC					: 1050
AP2	: CACACTGGACCCGACCCCTGGATGGAGATGACAGAGGCTGCCCTGCGGCTGC					: 67
AP9	: CACACTGGACCCGACCCCTGGATGGAGATGACAGAGGCTGCCCTGCGGCTGC					: 61
AP13	: CACACTGGACCCGACCCCTGGATGGAGATGACAGAGGCTGCCCTGCGGCTGC					: 60
	1060	1070	1080	1090	1100	
CIAP	: TGAACAGGAACCCCCGGGGCTTCTACCTCTTCGTGGAGGGAGGCGCGCATTC					: 1100
HPAP	: TGAACAGGAACCCCCGGGGCTTCTTCCTCTTCGTGGAGGGTGGTCGCATTC					: 1100
AP2	: TGAACAGGAACCCCCGGGGCTTCTTCCTCTTCGTGGAGGGTGGTCGCATTC					: 117
AP9	: TGAACAGGAACCCCCGGGGCTTCTTCCTCTTCGTGGAGGGTGGTCGCATTC					: 111
AP13	: TGAACAGGAACCCCCGGGGCTTCTTCCTCTTCGTGGAGGGTGGTCGCATTC					: 110
	1110	1120	1130	1140	1150	
CIAP	: GACCACGGTCAACCATGACGGCAAAAGCTTATATGGCACTGACTGAGGCGGAT					: 1150
HPAP	: GACCAAGGTCAACCATGAAGGCAAGGCTTAAGGGCACTGACTGAGAGCGAT					: 1150
AP2	: GACCAAGGTCAACCATGAAGGCAAGGCTTAAGGGCACTGACTGAGAGCGAT					: 167
AP9	: GACCAAGGTCAACCATGAAGGCAAGGCTTAAGGGCACTGACTGAGAGCGAT					: 161
AP13	: GACCAAGGTCAACCATGAAGGCAAGGCTTAAGGGCACTGACTGAGAGCGAT					: 160
	1160	1170	1180	1190	1200	
CIAP	: CATGTTTGACCAATGACATCGCCAAAGGCTAACGAGCTCACTAGCGAACTGG					: 1200
HPAP	: CATGTTTGACCAATGACATCGAGAGGGCGGGCAGCTCACAGCGAGGAGG					: 1200
AP2	: CATGTTTGACCAATGACATCGAGAGGGCGGGCAGCTCACAGCGAGGAGG					: 217
AP9	: CATGTTTGACCAATGACATCGAGAGGGCGGGCAGCTCACAGCGAGGAGG					: 211
AP13	: CATGTTTGACCAATGACATCGAGAGGGCGGGCAGCTCACAGCGAGGAGG					: 210

Figure 4 2/3

	1210	1220	1230	1240	1250	
CIAP :	ACACGCTGAT	CCTT	GTCAC	TGCA	GACCACTCCCAT	GTCTTCTCCTTTGGT : 1250
HPAP :	ACACGCTGAC	CCTT	GTCAC	TGCA	GACCACTCCCA	GTCTTCTCCTTTGGA : 1250
AP2 :	ACACGCTGAC	CCTT	GTCAC	TGCA	GACCACTCCCA	GTCTTCTCCTTTGGA : 267
AP9 :	ACACGCTGAC	CCTT	GTCAC	TGCA	GACCACTCCCA	GTCTTCTCCTTTGGA : 261
AP13 :	ACACGCTGAC	CCTT	GTCAC	TGCA	GACCACTCCCA	GTCTTCTCCTTTGGA : 260
	1260	1270	1280	1290	1300	
CIAP :	GGCTACAC	CACTGCGT	GGGAC	CTCCAT	TTTCGGT	CTGGCCCCCGGCAAGGC : 1300
HPAP :	GGCTAC	CCCTGCGA	GGGAC	CTCCAT	TTTCGG	CTGGCCCCCGGCAAGGC : 1300
AP2 :	GGCTAC	CCCTGCGA	GGGAC	CTCCAT	TTTCGG	CTGGCCCCCGGCAAGGC : 317
AP9 :	GGCTAC	CCCTGCGA	GGGAC	CTCCAT	TTTCGG	CTGGCCCCCGGCAAGGC : 311
AP13 :	GGCTAC	CCCTGCGA	GGGAC	CTCCAT	TTTCGG	CTGGCCCCCGGCAAGGC : 310
	1310	1320	1330	1340	1350	
CIAP :	CTTAGACAG	CAAGT	CCTACAC	CTC	CATCCTCTAT	GGCAATGGCCAGGCT : 1350
HPAP :	CCCCGACAG	CAAG	CCTACAC	CTC	CTCCTATA	GGAAAGGCTCCAGGCT : 1350
AP2 :	CCCCGACAG	CAAG	CCTACAC	CTC	CTCCTATA	GGAAAGGCTCCAGGCT : 367
AP9 :	CCCCGACAG	CAAG	CCTACAC	CTC	CTCCTATA	GGAAAGGCTCCAGGCT : 361
AP13 :	CCCCGACAG	CAAG	CCTACAC	CTC	CTCCTATA	GGAAAGGCTCCAGGCT : 360
	1360	1370	1380	1390	1400	
CIAP :	ATGCGCTT	TGGCGGG	GGCTCG	GAGGCC	CGATGTTAAT	TGGCAGCACAAGCGAG : 1400
HPAP :	ATGTGCT	CAAGGAC	GGCCCC	GGCC	GATGTTA	CGGAGAGCGAGCGCG : 1400
AP2 :	ATGTGCT	CAAGGAC	GGCCCC	GGCC	GATGTTA	CGGAGAGCGAGCGCG : 417
AP9 :	ATGTGCT	CAAGGAC	GGCCCC	GGCC	GATGTTA	CGGAGAGCGAGCGCG : 411
AP13 :	ATGTGCT	CAAGGAC	GGCCCC	GGCC	GATGTTA	CGGAGAGCGAGCGCG : 410
	1410	1420	1430	1440	1450	
CIAP :	GAA	CCCTCGTAC	CGGCAGCAG	GCGGCC	GTGCCCCTGG	CTAGCGAGACCCA : 1450
HPAP :	AGC	CCCAGTA	CGGCAGCAG	TCAGCA	GTGCCCCTGG	AGGAAAGAGACCCA : 1450
AP2 :	AGC	CCCAGTA	CGGCAGCAG	TCAGCA	GTGCCCCTGG	AGGAAAGAGACCCA : 467
AP9 :	AGC	CCCAGTA	CGGCAGCAG	TCAGCA	GTGCCCCTGG	AGGAAAGAGACCCA : 461
AP13 :	AGC	CCCAGTA	CGGCAGCAG	TCAGCA	GTGCCCCTGG	AGGAAAGAGACCCA : 460

Figure 4 3/3

	1460	1470	1480	1490	1500	
CIAP :	CGGGGGCGAA	GACGTGGCGGTGTT	CGCGCG	GGCCCCGCAGGCGCACCTGG		: 1500
HPAP :	CGGGGGCGAA	GACGTGGCGGTGTT	CGCGCG	GGCCCCGCAGGCGCACCTGG		: 1500
AP2 :	CGGGGGCGAA	GACGTGGCGGTGTT	CGCGCG	GGCCCCGCAGGCGCACCTGG		: 517
AP9 :	CGGGGGCGAA	GACGTGGCGGTGTT	CGCGCG	GGCCCCGCAGGCGCACCTGG		: 511
AP13 :	CGGGGGCGAA	GACGTGGCGGTGTT	CGCGCG	GGCCCCGCAGGCGCACCTGG		: 510

	1510	1520	1530	1540	1550	
CIAP :	TGCACGGCGTGCAGGAG	GAGACCTTCGT	GGCGCACATCATGGCCTTT	GGCG		: 1550
HPAP :	TGCACGGCGTGCAGGAG	GAGACCTTCAT	AGCGCACCTCATGGCCTTT	GGCG		: 1550
AP2 :	TGCACGGCGTGCAGGAG	GAGACCTTCGT	GGCGCACATCATGGCCTTT	GGCG		: 567
AP9 :	TGCACGGCGTGCAGGAG	GAGACCTTCAT	AGCGCACCTCATGGCCTTT	GGCG		: 561
AP13 :	TGCACGGCGTGCAGGAG	GAGACCTTCAT	AGCGCACCTCATGGCCTTT	GGCG		: 560

	1560	1570	1580	1590	1600	
CIAP :	GGCTGCGTGGAGCCCTACACCG	ACTGCAATCTGC	CCAGCCCCCTCCACCGC			: 1600
HPAP :	GGCTGCGTGGAGCCCTACACCG	ACTGCGACCTGCG	CCCCCGCCGGCAC			: 1600
AP2 :	GGCTGCGTGGAGCCCTACACCG	ACTGCAATCTGC	CCAGCCCCCTCCACCGC			: 617
AP9 :	GGCTGCGTGGAGCCCTACACCG	ACTGCGACCTGCG	CCCCCGCCGGCAC			: 611
AP13 :	GGCTGCGTGGAGCCCTACACCG	ACTGCGACCTGCG	CCCCCGCCGGCAC			: 610

	1610	1620	1630	1640	1650	
CIAP :	CACCGACATCCCCGACTGATAA	TCTAGAGGTCGAAATTCACCTCGAAAGC				: 1650
HPAP :	CACCGACTGATAAAAGCTTGG	TCTAGAGGTCGAAATTCACCTCGAAAGC				: 1649
AP2 :	CACCGACATCCCCGACTGATAA	TCTAGAGGTCGAAATTCACCTCGAAAGC				: 667
AP9 :	CACCGACTGATAAAAGCTTGG	TCTAGAGGTCGAAATTCACCTCGAAAGC				: 661
AP13 :	CACCGACTGATAAAAGCTTGG	TCTAGAGGTCGAAATTCACCTCGAAAGC				: 660

CIAP: SEQ ID NO: 22

HPAP: SEQ ID NO: 21

AP2: SEQ ID NO: 23

AP9: SEQ ID NO: 24

AP13: SEQ ID NO: 25

Figure 5: 1/3

Sequences of hpap, ciap, AP01, AP03, AP05, AP06, AP11, AP15, AP2, AP9 and AP13.

ciap (SEQ ID NO: 26)

CTTCGGCGTTTCAGTAACACGCGTTAACGTGCTGAACAGCCGGGCATTTTTTTTACGCTATACCCTACATAATAAAACCGGA
GCTACCATGAATAAGAAGGTACTGACCCCTTTCTGCCGTGATGGCAAGTCTGTTATTTCGGGGGCCACGCGCACGCGGCGAT
CCCAGCTGAGGAGGAAAACCCCGCCTTCTGGAACCGCCAGGCAGCCAGGCCCTTGATGTAGCCAAGAAGTTGCAGCCGA
TCCAGACAGCTGCCAAGAATGTCATCCTCTTCTTGGGGGATGGGATGGGGGTGCCTACGGTGACAGCCACTCGGATCCTA
AAGGGGCAGATGAATGGCAAACCTGGGACCTGAGACACCCCTGGCCATGGACCAGTTCCCATACGTGGCTCTGTCCAAGAC
ATACAACGTGGACAGACAGGTGCCAGACAGCGCAGGCACTGCCACTGCCTACCTGTGTGGGGTCAAGGGCAACTACAGAA
CCATCGGTGTAAGTGCAGCCGCCCGCTACAATCAGTGCAACACGACACGTGGGAATGAGGTACAGTCTGTGATCAACCGG
GCCAAGAAAGCAGGGAAGGCCGTGGGAGTGGTGACCACCACCAGGGTGCAGCATGCCTCCCCAGCCGGGGCCTACGCGCA
CACGGTGAACCGAACTGGTACTCAGACGCCGACCTGCCTGCTGATGCACAGAAGAATGGCTGCCAGGACATCGCCGCAC
AGCTGGTCTACAACATGGATATTGACGTGATCCTGGGTGGAGGCCGAATGTACATGTTTCTTGAGGGGACCCAGACCCT
GAATACCCAGATGATGCCAGTGTGAATGGAGTCCGGAAGGACAAGCAGAACCTGGTGCAGGAATGGCAGGCCAAGCACCA
GGGAGCCCAGTATGTGTGGAACCGCACTGCGCTCCTTCAGGCGGCCGATGACTCCAGTGTAACACACCTCATGGGCCTCT
TTGAGCCGGCAGACATGAAGTATAATGTTTCAGCAAGACCACACCAAGGACCCGACCCCTGGCGGAGATGACGGAGGCGGCC
CTGCAAGTGCTGAGCAGGAACCCCGGGGCTTCTACCTCTTCTGAGGGGAGGCCGCATTGACCACGGTCACCATGACGG
CAAAGCTTATATGGCACTGACTGAGGCGATCATGTTTGACAATGCCATCGCCAAGGCTAACGAGCTCAC'TAGCGAACTGG
ACACGCTGATCCTTGTCACTGCAGACCACTCCCATGTCTTCTCCTTTGGTGGCTACACACTGCGTGGGACCTCCATTTTC
GGTCTGGCCCCCGGCAAGGCCCTTAGACAGCAAGTCTACACCTCCATCCTCTATGGCAATGGCCCAGGCTATGCGCTTGG
CGGGGGCTCGAGGCCCGATGTTAATGGCAGCACAAAGCGAGGAACCCCTCGTACCGGCAGCAGGCGGCCGTGCCCTGGCTA
GCGAGACCCACGGGGGCGAAGACGTGGCGGTGTTGCGCGAGGCCCGCAGGCGCACCTGGTGCACGGCGTGCAGGAGGAG
ACCTTCGTGGCGCACATCATGGCCTTTGCGGGCTGCGTGGAGCCCTACACCGACTGCAATCTGCCAGCCCCCTCCACCGC
CACCAGCATCCCCGACTGATAATCTAGAGGTGGAATTCACCTCGAAAGC

Hpap (SEQ ID NO: 27)

CTTCGGCGTTTCAGTAACACGCGTTAACGTGCTGAACAGCCGGGCATTTTTTTTACGCTATACCCTACATAATAAAACCGGA
GCTACCATGAATAAGAAGGTACTGACCCCTTTCTGCCGTGATGGCAAGTCTGTTATTTCGGCGCGCACGCGCACGCGGCGAT
CCCAGTTGAGGAGGAGAACCCGGACTTCTGGAACCGCGAGGCAGCCAGGCCCTGGGTGCCGCCAAGAAGCTGCAGCCTG
CACAGACAGCCGCCAAGAACCTCATCATCTTCTGGGCGATGGGATGGGGGTGTCTACGGTGACAGCTGCCAGGATCCTA
AAAGGGCAGAAGAAGGACAAACTGGGGCCTGAGATACCCCTGGCCATGGACCCTTCCCATATGTGGCTCTGTCCAAGAC
ATACAATGTAGACAAACATGTGCCAGACAGTGGAGCCACAGCCACGGCCTACCTGTGCGGGGTCAAGGGCAACTTCCAGA
CCATTGGCTTGAGTGCAGCCGCCCGCTTTAACCAGTGCAACACGACACGCGGCAACGAGGTCTCTCCGTGATGAATCGG
GCCAAGAAAGCAGGGAAGTCAGTGGGAGTGGTAACCACCACAGAGTGCAGCACGCCTCGCCAGCCGGCACCTACGCCCA
CACGGTGAACCGCAACTGGTACTCGGACGCCGACGTGCCTGCCTCGGCCCGCCAGGAGGGGTGCCAGGACATCGCTACGC
AGCTCATCTCCAACATGGACATTGACGTGATCCTAGGTGGAGGCCGAAAGTACATGTTTCCCATGGGAAACCCAGACCCT
GAGTACCCAGATGACTACAGCCAAGGTGGGACCAGGCTGGACGGGAAGAATCTGGTGCAGGAATGGCTGGCGAAGCGCCA
GGGTGCCCGGTATGTGTGGAACCGCACTGAGCTCATGCAGGCTTCCCTGGACCCGTCTGTGACCCATCTCATGGGTCTCT
TTGAGCCTGGAGACATGAAATACGAGATCCACCGAGACTCCACACTGGACCCCTCCCTGATGGAGATGACAGAGGCTGCC
CTGCGCCTGCTGAGCAGGAACCCCGCGGCTTCTTCTCTTCTGAGGGGTGGTTCGCATCGACCATGGTTCATCATGAAAG
CAGGGCTTACCGGGCACTGACTGAGACGATCATGTTTCGACGACGCCATTGAGAGGGCGGGCCAGCTCAC'CAGCGAGGAGG
ACACGCTGAGCCTCGTCACTGCCGACCACTCCACGTCTTCTCCTTCGGAGGCTACCCCTGCGAGGGAGCTCCATCTTC
GGGCTGGCCCCCTGGCAAGGCCCGGGACAGGAAGGCCCTACACGGTCTCTCTATACGGAAACGGTCCAGGCTATGTGCTCAA
GGACGGCGCCCGGCCGGATGTTACCGAGAGCGAGAGCGGGAGCCCCGAGTATCGGCAGCAGTCAGCAGTGCCCCCTGGACG

Figure 5: 2/3

AAGAGACCCACGCAGGCGAGGACGTGGCGGTGTTTCGCGCGCGGCCCCGCAGGCGCACCTGGTTTCACGGCGTGCAGGAGCAG
ACCTTCATAGCGCACGTTCATGGCCTTCGCCGCCTGCCTGGAGCCCTACACCGCCTGCGACCTGGCGCCCCCGCGGCAC
CACCGACTGATAAAAGCTTGGTCTAGAGGTCGAAATTCACCTCGAAAGC

AP01 (SEQ ID NO: 28)

CTTCGGCGTTCAGTAACACGCGTTAACGTGCTGAACAGCCGGGCATTCTTTTACGCTATAACCCTACATAATAAAACCGGA
GCTACCATGAATAAGAAGGTACTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTCGGGGCCACGCGCACGCGGCGAT
CCCAGTTGAGGAGGAGAACCCGGACTTCTGGAACCGCGAGGCAGCCAAGGCCCTGGGTGCCGCCAAGAAGCTGCAGCCTG
CACAGACAGCCGCCAAGAACCTCATCATCTTCCTGGGCGATGGGATGGGGGTGTCTACGGTGACAGCTGCCAGGATCCTA
AAAGGGCAGAAGAAGGACAAACTGGGGCCTGAGATACCCCTGGCCATGGACCGCTTCCCATATGTGGCTCTGTCCAAGAC
ATACAATGTAGACAAACATGTGCCAGACAGTGGAGCCACAGCCACGGCCTACCTGTGCGGGGTCAAGGGCAACTTCCAGA
CCATTGGCTTGAGTGCAGCCGCCCGCTTTAACAGTGCAACACGACACGC

AP03 (SEQ ID NO: 29)

CTTCGGCGTTCAGTAACACGCGTTAACGTGCTGAACAGCCGGGCATTTTTTTTACGCTATAACCCTACATAATAAAACCGGA
GCTACCATGAATAAGAAGGTACTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTCGGCGCGCACGCGCACGCGGCGAT
CCCAGTTGAGGAGGAGAACCCGGACTTCTGGAACCGCGAGGCAGCCAAGGCCCTGGGTGCCGCCAAGAAGCTGCAGCCTG
CACAGACAGCCGCCAAGAACCTCATCATCTTCCTGGGCGATGGGATGGGGGTGTCTACGGTGACAGCTGCCAGGATCCTA
AAAGGGCAGAAGAAGGACAAACTGGGGCCTGAGATACCCCTGGCCATGGACCGCTTCCCATATGTGGCTCTGTCCAAGAC
ATACAATGTAGACAAACATGTGCCAGACAGTGGAGCCACAGCCACGGCCTACCTGTGTGGGGTCAA

AP05 (SEQ ID NO: 30)

CTTCGGCGTTCAGTAACACGCGTTAACGTGCTGAACAGCCGGGCATTTTTTTTACGCTATAACCCTACATAATAAAACCGGA
GCTACCATGAATAAGAAGGTACTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTCGGCGCGCACGCGCACGCGGCGAT
CCCAGTTGAGGAGGAGAACCCGGACTTCTGGAACCGCGAGGCAGCCAAGGCCCTGGGTGCCGCCAAGAAGCTGCAGCCTG
CACAGACAGCCGCCAAGAACCTCGTCATCTTCCTGGGCGATGGGATGGGGGTGCCTACGGTGACAGCCACTCGGACCCTA
AAAGGGCAGAAGAAGGACAAACTGGAGCCTGAGATACCCCTGGCCATGGACCGCTTCCCATATGTGGCTCTGTCCAAGAC
ATACAACGTGGACAGACAGGTGCCAGACAGTGGAGCCACAGCCACGGCCTACCTGTGCGGGGTCAAGGGCAACTTCCAGA
CCATTGGCTTGAGTGCAGCCGCCCGCTTTAACAGTGCAACACGACACGCGGCAACGAGGTCATCTCCGTGATGAATCGG
GCCAAGAAAGCAGGGAAGTCAGTGGGAGTGGTAACCACCACAGAGTGCAGCACGCCTCGCCAGCCGGCACCTACGCCCA
CACGG

AP06 (SEQ ID NO: 31)

CTTCGGCGTTCAGTAACACGCGTTAACGTGCTGAACAGCCGGGCATTTTTTTTACGCTATAACCCTACATAATAAAACCGGA
GCTACCATGAATAAGAAGGTACTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTCGGCGCGCACGCGCACGCGGCGAT
CCCAGTTGAGGAGGAGAACCCGGACTTCTGGAACCGCGAGGCAGCCAAGGCCCTGGGTGCCGCCAAGAAGCTGCAGCCTG
CACAGACAGCCGCCAAGAACCTCATCATCTTCCTGGGCGATGGGATGGGGGTGTCTACGGTGACAGCTGCCAGGATCCTA
AAAGGGCAGAAGAAGGACAAACTGGGGCCTGAGATACCCCTGGCCATGGACCGCTTCCCATATGTGGCTCTGTCCAAGAC
ATACAATGTAGACAAACATGTGCCAGACAGTGGAGCCACAGCCACGGCCTACCTGTGCGGGGTCAAGGGCAACTTCCAGA
CCATTGGCTTGAGTGCAGCCGCCCGCTTTAACAGTGCAACACGACACGCGGCAGCGAGGTCATCTCCGTGATGAACCGG
GCCAAGAAAGCAGG

AP11 (SEQ ID NO: 32)

CTTCGGCGTTCAGTAACACGCGTTAACGTGCTGAACAGCCGGGCATTTTTTTTACGCTATAACCCTACATAATAAAACCGGA
GCTACCATGAATAAGAAGGTATTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTCGGCGCGCACGCGCACGCGGCGAT
CCCAGTTGAGGAGGAGAACCCGGACTTCTGGAACCGCGAGGCAGCCAAGGCCCTGGGTGCCGCCAAGAAGCTGCAGCCTG
CACAGACAGCCGCCAAGAACCTCATCATCTTCCTGGGCGATGGGATGGGGGTGTCTACGGTGACAGCTGCCAGGATCCTA

Figure 5: 3/3

AAAGGGCAGATGAATGGCAAACCTGGGACCTGAGACACCCCTGGCCATGGACCAGTTCCCATACGTGGCTCTGTCCAAGAC
ATACAATGTAGACAAACATGTGCCAGACAGTGGAGCCACAGCCACGGCCTACCTGTGCGGGGTCAAGGGCAACTTCCAGA
CCATTGGCTTGAGTGCAGCCGCCCGCTTTAACCAGTGCAACACGACACGCGGCAACGAGGTCATCTCCGTGATGAATCGG
GCCAAGAAAGCAGGGAAGTCAGTGGGAGTGGTAACCACCACACGAGTGCAGCACGCCTCG

AP15 (SEQ ID NO: 33)

CTTCGGCGTTTACGTAACACGCGTTACCGTGCTGAACAGCCGGGCATTTTTTTTACGCTATACCTACATAATAAAACCGGA
GCTACCATGAATAAGAAGGTACTGACCCTTTCTGCCGTGATGGCAAGTCTGTTATTCGGGGCCCACGCGCACGCGGCGAT
CCCAGCTGAGGAGGAAAACCCCGCCTTCTGGAACCGCCAGGCAGCCAGGCCCTTGATGTAGCCAAGAAGTTGCAGCCGA
TCCAGACAGCTGCCAAGAATGTCATCCTCTTCTTGGGGGATGGGATGGGGGTGTCTACGGTGACAGCTGCCAGGATCCTA
AAAGGGCAGAAGAAGGACAAACTGGGGCCTGAGATACCCCTGGCCATGGACCGCTTCCCATATGTGGCTCTGTCCAAGAC
ATACAATGTAGACAAACATATGCCAGACAGTGGAGCCACAGCCACGGCCTACCTGTGCGGGGTCAAGGGCAACTTCCAGA
CCATTGGCTTGAGTGCAGCCGCCCGCTTTAACCAGTGCAACACGACACGCGGCAACGAGGTCATCTCCG

AP2 (SEQ ID NO: 34)

GAGATCCACCGAGACTCCACACTGGACCCCTCCCTGATGGAGATGACAGAGGCTGCCCTGCGCCTGCTGAGCAGGAACCC
CCGCGGCTTCTTCCTCTTCGTGGAGGGTGGTTCGCATCGACCATGGTCATCATGAAAGCAGGGCTTACCGGGCACTGACTG
GGACGATCATGTTTCGACGACGACATTGAGAGGGCGGGCCAGCTCACCAGCGAGGAGGACACGCTGAGCCTCGTCACTGCC
GACCACTCCCACGTCTTCTCCTTCGGAGGCTACCCCTGCGAGGGAGCTCCATCTTCGGGCTGGCCCCCTGGCAAGGCCCG
GGACAGGAAGGCCCTACACGGTCCTCCTATACGGAACGGTCCAGGCTATGTGCTCAAGGACGGCGCCCGGCCGGATGTTA
CCGAGGGCGAGAGCGGGAGCCCCGAGTATCGGCAGCAGTCAGCAGTGCCCTGGACGAAGAGACCCACGCAGGCGAGGAC
GTGGCGGTGTTTCGCGCGCGGCCCGCAGGCGCACCTGGTGCACGGCGTGCAGGAGGAGACCTTCGTTGGCGCACATCATGGC
CTTTGCGGGCTGCGTGGAGCCCTACACCGACTGCAATCTGCCAGCCCCCTCCACCGCCACCAGCATCCCCGACTGATAAT
CTAGAGGTCGAAATTCACCTCGAAAGC

>AP9 (SEQ ID NO: 35)

CACCGAGACTCCACACTGGACCCCTCCCTGATGGAGATGACAGAGGCTGCCCTGCGCCTGCTGAACAGGAACCCCCGCGG
CTTCTTCCTCTCCGTGGAGGGTGGTTCGCATCGACCATGGTCATCATGAAAGCAGGGCTTACCGGGCACTGACTGAGACGA
TCATGTTTCGACGACGCCATTGAGAGGGCGGGCCAGCTCACCAGCGAGGAGGACACGCTGAGCCTCGTCACTGCCGACCAC
TCCCACGTCTTCTCCTTCGGAGGCTACCCCTGCGAGGGAGCTCCATCTTCGGGCTGGCCCCCTGGCAAGGCCCGGGACAG
GAAGGCCCTACACGGTCCTCCTATACGGAACGGTCCAGGCTATGTGCTCAAGGACGGCGCCCGGCCGGATGTTACCGAGA
GCGAGAGCGGGAGCCCCGAGTATCGGCAGCAGTCAGCAGTGCCCTGGACGAAGAGACCCACGCAGGCGAGGACGTGGCG
GTGTTTCGCGCGCGGCCCGCAGGCGCACCTGGTTCACGGCGTGCAGGAGCAGACCTTCATAGCGCACGTCAATGGCCTTCGC
CGCTGCTGGAGCCCTACACCGCCTGCGACCTGGCGCCCCCGCCGGCACCACCGACTGATAAAAGTTTGGCTCTAGAG
GTCGAAATTCACCTCGAAAGC

>AP13 (SEQ ID NO: 36)

ACCGAGACTCCACACTGGACCCCTCCCTGATGGAGATGACAGAGGCTGCCCTGCGCCTGCTGAGCAGGAACCCCCGCGGC
TTCTTCTCTTCGTGGAGGGTGGTTCGCATCGACCATGGTCATCATGAAAGCAGGGCTTACCGGGCACTGACTGAGACGAT
CATGTTTCGACGACGCCATCGAGAGGGCGGGCCAGCTCACCAGCGAGGAGGACACGCTGAGCCTCGTCACTGCCGACCAC
CCCACGTCTTCTCCTTCGGAGGCTACCCCTGCGAGGGAGCTCCATCTTCGGGCTGGCCCCCTGGCAAGGCCCGGGACAGG
AAGGCCTACACGGTCCTCCTATACGGAACGGTCCAGGCTATGTGCTCAAGGACGGCGCCCGGCCGGATGTTACCGAGAG
CGAGAGCGGGAGCCCCGAGTATCGGCAGCAGCCAGCAGTGCCCTGGACGAAGAGACCCACGCAGGCGAGGACGTGGCGG
TGTTTCGCGCGCGGCCCGCAGGCGCACCTGGTTCACGGCGTGCAGGAGCAGACCTTCATAGCGCACGTCAATGGCCTTCGCC
GCCTGCTGGAGCCCTACACCGCCTGCGACCTGGCGCCCCCGCCGGCACCACCGACTGATAAAAGCTTGGCTCTAGAGG
TCGAAATTCACCTCGAAAGC

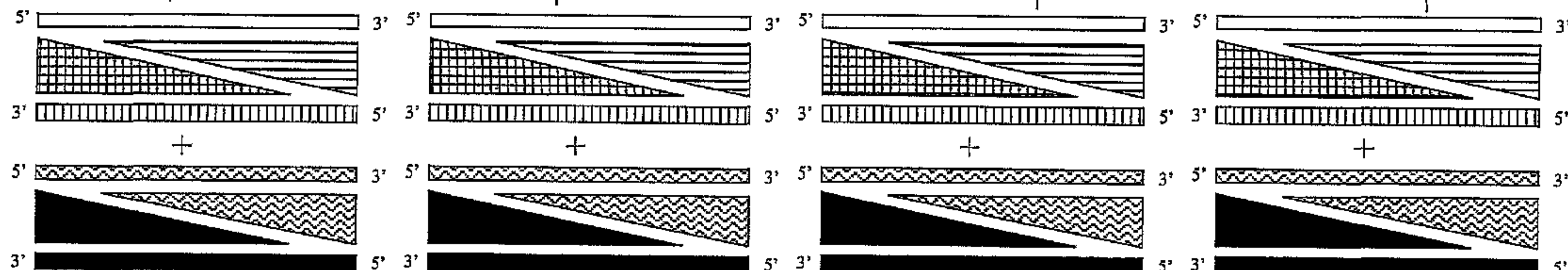
Synthesis and/or chain reaction

dNTP/ddATP

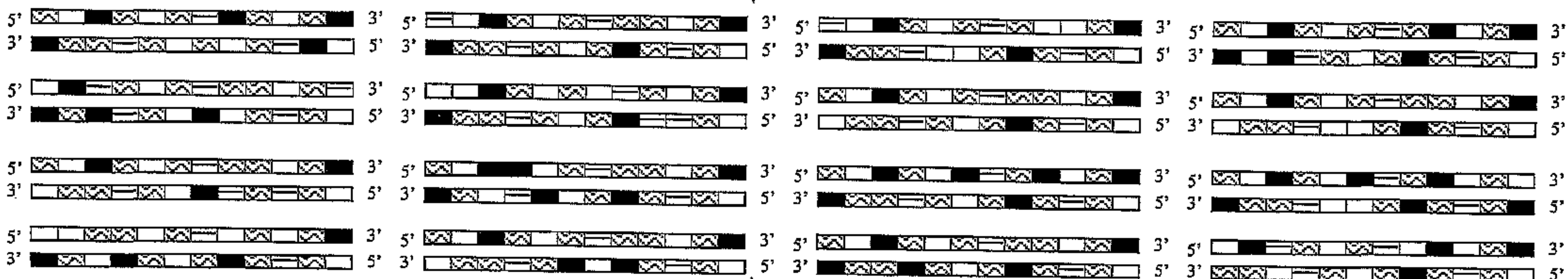
dNTP/ddCTP

dNTP/ddGTP

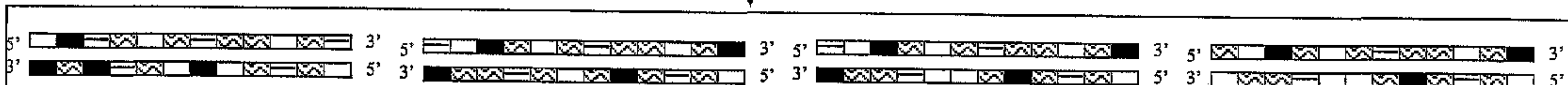
dNTP/ddTTP



Removing templates/terminators
Synthesis and/or chain reaction



Cloning and Expression
Screening or Selection



Improved variants