Mol. Cells, Vol. 10, No. 2, pp. 236–240

Communication

Molecules and Cells

A Novel Technique for the Effective Production of Short Peptide Analogs from Concatameric Short Peptide Multimers

Sang Jun Lee*, Jong Hee Lee, Hyun Joo Jin, Jeong Ho Lee, Ho Young Ryu, Yoon Kim, In Soo Kong[1], and Kyu Won Kim[2]

Biotechnology Division, National Fisheries Research & Development Institute, Pusan 619-900, Korea;
[1] Department of Biotechnology & Bioengineering, Pukyong National University, Pusan 608-737, Korea;
[2] Department of Molecular Biology, College of Natural Sciences, Pusan National University, Pusan 609-735, Korea.

(Received on January 7, 2000)

We designed a basic unit of the modified chicken gonadotropin releasing hormone II (cGnRH-II) peptide containing a trypsin cleavable linker peptide at both ends of the original peptide. We made a synthetic DNA coding for the modified cGnRH-II peptide with asymmetric and complementary cohesive ends of linker nucleotides. A tandemly repeated DNA cassette for the expression of concatameric short peptide multimers was constructed by ligating the basic units. The expressed peptide multimers were purified and subject to amino-terminal sequence analysis, which displayed the amino acid sequences expected from the designed nucleotides of the expression cassette. The monomeric cGnRH-II peptide analogs were generated after trypsin digestion. The present results showed that the technique developed for the production of the concatameric peptide multimers with cleavable linker peptides can be generally applicable to the production of short peptide analogs.

Keywords: Cleavable Linker Peptide; Concatameric Peptide Multimers; Peptide Analogs; Tandemly Repeated DNA Cassette.

Introduction

Short peptides that are used widely as bioactive peptides for experimental or clinical purposes are generally produced by chemical synthesis. However, the cost of chemical synthesis for short peptides is high for small scale production. Genetic engineering techniques have successfully produced many proteins. So far, protein production by genetic engineering techniques has been conducted generally by isolating the required genes, constructing expression vectors, and producing the proteins in *Escherichia coli* host. It is a very effective method for the production of the proteins of bigger size than short peptides, but this method has hardly been applied to produce short peptides because of a low expression level.

Addressing these problems, many scientists have studied intensively to develop cloning techniques for tandem DNA multimers to produce proteins or DNA. Rare asymmetric and complementary cohesive ends of the restriction cleavages of *Ava*I (Hartley and Gregori, 1981) and class-IIS (Kim and Szybalski, 1988; Lee *et al*., 1996) were developed. The other methods for generating directional DNA sequences of tandem repeat were obtained by attaching the synthetic directional adaptors to the DNA fragment ends to establish the production of direct repeats (Taylor and Hagerman, 1987), by attaching synthetic directional adaptors to the terminals of the gene fragment (Pan *et al*., 1993), and by using 2 bp overhangs to direct orientation and random incorporation of linkers containing restriction sites (Jiang *et al*., 1996). However, general methods of cloning for the tandemly repeated DNA multimers for an effective and quantitative expression of the concatameric short peptides have not been developed yet. In this study, we developed a novel technique for producing short peptide analogs from the concatameric short peptide multimers that were produced in *E. coli*.

* To whom correspondence should be addressed.
Tel: 82-51-720-2450; Fax: 82-51-720-2456

Abbreviations: cGnRH-II, chicken gonadotropin releasing hormone II; cGnRH-II(M), modified cGnRH-II; DTT, 1,4-dithiotreitol; FPLC, fast-protein liquid chromatography; IPTG, isopropyl-β-D-thiogalactopyranoside; OD, optical density; PVDF, polyvinylidene difluoride; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Materials and Methods

Bacterial strain, plasmids and phage *E. coli* strain of JM101 (Messing, 1983), JM109 and a plasmid pUC19 (Yanish-Peron *et al*., 1985) were used, and a *recA* strain of *E. coli* TOP10F′, the plasmid vector pRSET-B and M13/T7 phage purchased from Invitrogen Co. were used.

Reagents and molecular techniques All enzymes for DNA manipulations were purchased from Boehringer Mannheim (USA) and used under the conditions recommended by the supplier. Metaphor agarose and marker proteins were purchased from FMC Co. and Bio-Rad Co., respectively. All the molecular techniques were conducted by molecular cloning (Sambrook *et al*., 1989). Nucleotide sequencing using the dideoxy chain termination method (Sanger *et al*., 1977) was done with the Sequenase 2.0 kit (United States Biochemical). A computer program, DNASIS^{TM} (Hitachi, Japan) was used to analyze RNA and protein structure.

Oligonucleotides Oligonucleotides were synthesized, phosphorylated, and purified by reversed high pressure liquid chromatography in Bioneer Co. (Korea). For the left adaptor synthetic DNA, oligonucleotides of 5′-A ATT CAA GGA TCC CCC GGG GGG AAG AGA-3′ and 5′-CCC GGG GGA TCC TTG-3′ were used, for the right adaptor synthetic DNA, oligonucleotides of 5′-CTC GAG AAG CTT ACG-3′ and 5′-T CGA CGT AAG CTT CTC GAG TCT CTT CCC-3′ were used, and for cGnRH-II(M) synthetic DNA, oligonucleotides of 5′-GAG CAC TGG TCC CAC GGG TGG TAC CCA GGC GGG AAG AGA-3′ and 5′-GCC TGG GTA CCA CCC GTG GGA CCA GTG CTC TCT CTT CCC-3′ were used (see, Fig. 2). For the stop codon synthetic DNA, a self-complementary single oligonucleotide of 5′-AG CTT TGA CAG CTG TCA A-3′ was used after self-annealing. For Met-linker peptide-cGnRH-II portion synthetic DNA, oligonucleotides of 5′-T ATG GGC AAA CGA GAG CAT TGG TCA CAC GGG TGG TAC C-3′ and 5′-CC GGG GTA CCA CCC GTG TGA CCA ATG CTC TCG TTT GCC CA-3′ were used (see, Fig. 4).

Protein expression Transformants of *E. coli* TOP10F′ were cultured in SOB media (Tryptone, 20 g; yeast extract, 5.0 g; NaCl, 0.5 g; KCl, 1.86 mg/L) overnight in the presence of 50 μg/ml ampicillin. The overnight culture was diluted 200 times with SOB media. The diluted culture was incubated to reach OD600 of 0.3 at which IPTG was added at 1 mM final concentration and infected with M13/T7 phage after 1 h incubation. The phage infected culture was incubated a further for 5 h expression. 1 ml of the culture was centrifuged and the obtained cell pellet was resuspended in 200 ml of sample buffer (0.05 M Tris·HCl, pH 6.8, 0.1 M DTT, 2% SDS, 1% glycerol, 0.1% bromophenol blue). The resuspension was heated for 5 min at 90°C and analyzed by 16% SDS-PAGE (Laemmli, 1970).

Purification and trypsin digest of the cGnRH-II multimers Cells were harvested by centrifugation at 2000 × *g* for 20 min at 4°C. The harvested cells were washed twice with cold 20 mM Tris·HCl buffer (pH 8.0), resuspended in a 1/10 of the original culture, and ruptured by sonication. Anion-exchange chromatography was performed with the fast-protein liquid chromatography (FPLC) (Pharmacia LKB). A protein sample was loaded into a Mono-Q 5/5 anion-exchange column that had been equilibrated with a 20 mM Tris·HCl buffer at pH 8.0. A gradient was established starting from 100% of buffer A (20 mM Tris·HCl at pH 8.0) and ending with 100% of buffer B (20 mM Tris·HCl at pH 8.0 with 1.0 M NaCl). The pooled sample was desalted, lyophilized, and resuspended. The resuspended protein was digested with trypsin in 20 mM Tris·HCl (pH 8.0) for 3 h at room temperature. The cleaved protein was analyzed by 10–20% Tricine gel (Novex).

Amino acid sequencing The corresponding fragment of the concatameric multimer protein and the cleaved monomer protein were electroblotted on polyvinylidene difluoride (PVDF) membrane (Matsudaria, 1987) and analyzed for the N-terminal amino acid sequence with a protein sequencer (Applied Biosystems, model 476A) in a pulse-liquid mode at Korea Basic Science Institute (KBSI), Taejon, Korea.

Results

Design of the modified cGnRH-II synthetic DNA and cloning of the tandem repeated cGnRH-II(M) synthetic DNA cassette We used chicken gonadotropin releasing hormone II (cGnRH-II) peptide (Miyamoto *et al*., 1984) as a model. We designed a basic unit of the modified cGnRH-II [see below as cGnRH-II(M)] peptide by attaching a trypsin-cleavable linker peptide sequence of Gly-Lys-Arg at both ends of the original peptide. After deciphering the nucleotide of the basic unit of cGnRH-II(M), we designed the basic unit of cGnRH-II(M) synthetic DNA containing a single strand of 3′ overhang cohesive end (Fig. 1). The oligonucleotides were synthesized and annealed to make for the left and right adaptors as well as cGnRH-II(M) synthetic DNA (Fig. 1). The basic unit of the 3′ overhang cohesive end of cGnRH-II(M) synthetic DNA was self-ligated with T4 DNA ligase to make a tandemly repeated cGnRH-II(M) synthetic DNA cassette and then adaptors were added to make a clonable adaptor-ligated cGnRH-II(M) DNA cassette by a second ligation. The adaptor-ligated cGnRH-II(M) synthetic DNA cassette was cloned into the *Eco*RI/*Sal*I site of pUC19, then transformed into the *E. coli* strain of JM101 or JM109, and a few plasmids were screened from the white colonies by restriction enzyme analysis (Sambrook *et al*., 1989) on a 3% metaphor agarose gel. We isolated one plasmid containing an insert of three repeats of cGnRH-II(M) DNAs [as left adaptor-cGnRH-II(M) DNA insert (n = 3)-right adaptor], which was designated as pUC19-cGnRH-II(M)-3R(repeat).

Construction of the expression vector containing four repetitive cGnRH-II(M) synthetic DNAs The cGnRH-II(M)-3R cassette from pUC19-cGnRH-II(M)-3R was isolated by cutting with *Bam*HI/*Hin*dIII, and the fragment was transferred into the *Bam*HI/*Hin*dIII site of pRSET-B vector to generate pRSET-B-cGnRH-II(M)-3R. Since this

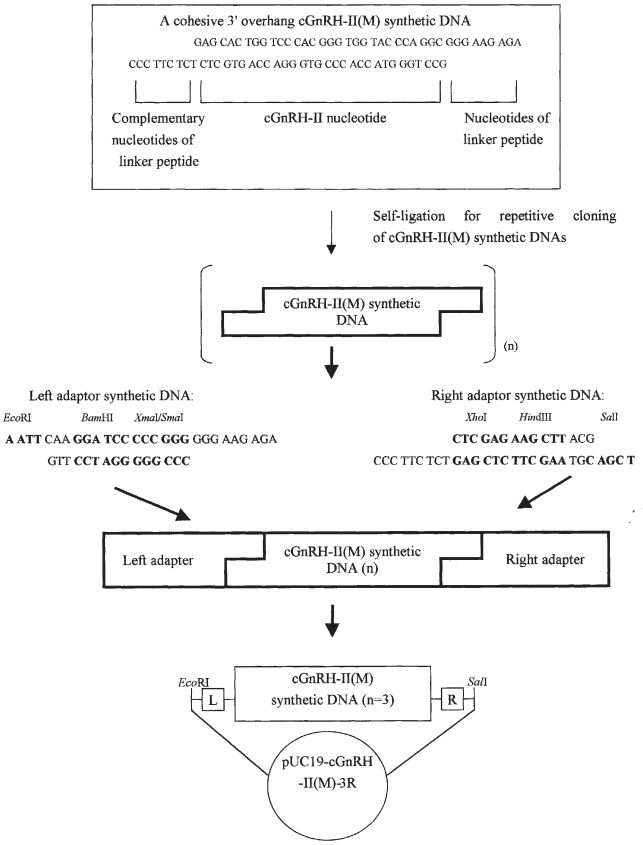

Fig. 1. Strategy for repetitive cloning of the cGnRH-II(M) synthetic DNAs. Oligonucleotides were synthesized and annealed to make the left and right adaptor synthetic DNAs, and cGnRH-II(M) synthetic DNA. A cohesive 3′ overhang cGnRH-II (M) synthetic DNA was first self-ligated and then ligated with adapter synthetic DNAs to make a clonable cassette. The adapter ligated cassette mixture was cloned into the *Eco*RI/*Sal*I site of pUC19. Abbreviations: L, left adaptor synthetic DNA; R, right adaptor synthetic DNA.

plasmid has no stop codon down stream of the frame, we inserted a synthetic DNA fragment containing the stop codon, constructed with a self-complementary single oligonucleotide by self-annealing, at *Hin*dIII site resulting pRSET-B-cGnRH-II(M)-3R-STOP (Fig. 2). We also replaced the leader peptide region of the *Nde*I/*Xma*I DNA fragment of pRSET-B-cGnRH-II(M)-3R-STOP with a designed synthetic DNA for Met-linker peptide-cGnRH-II peptide portion with a lower ΔG value for RNA secondary structure upstream of the cassette and named it pRSET-B-Δ G-cGnRH-II(M)-4R-STOP (Fig. 2). Finally, this expression vector was confirmed to contain four repetitive basic units of cGnRH-II(M) synthetic DNA with 183 bp as an open reading frame after DNA sequencing (Sanger *et al*., 1977) (Fig. 3). The deduced concatameric peptide multimers consist of four repetitions of cleavable linker peptide-cGnRH-II peptide sequences attached with a final cleavable linker peptide.

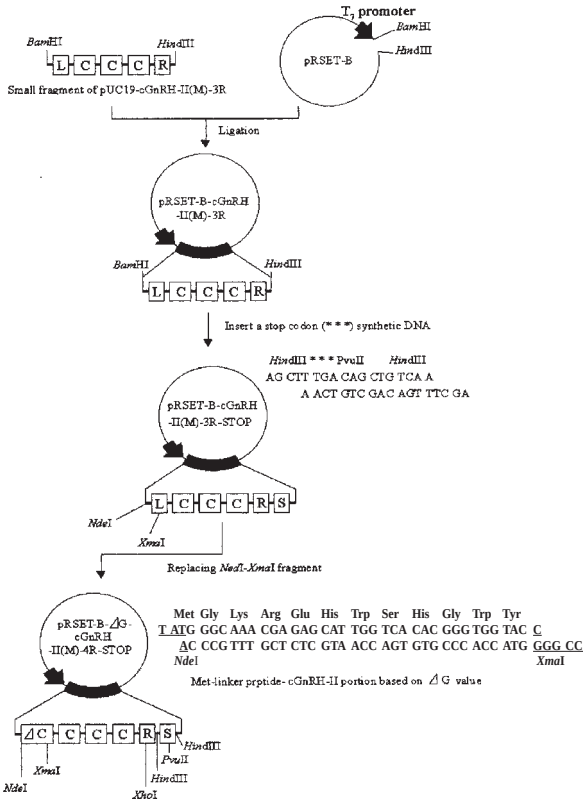

Fig. 2. Construction of the expression vector containing four repetitive basic units of the cGnRH-II(M) synthetic DNAs. Abbreviations: C, cGnRH-II(M) synthetic DNA; ΔC, Met-linker peptide-cGnRH-II portion synthetic DNA; S, stop codon synthetic DNA; L, and R are the same as in Fig. 1.

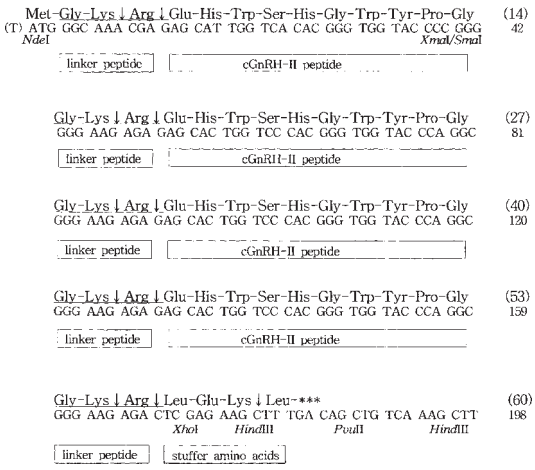

Fig. 3. Nucleotide and deduced amino acid sequence of four repetitive basic units of the cGnRH-II(M) synthetic DNAs in the expression vector, pRSET-B-Δ G-cGnRH-II(M)-4R-STOP. Abbreviations: ***, stop codon; Gly-Lys-Arg, linker peptide sequence; ↓, trypsin cleavage. The bp numbers and amino acid numbers are indicated on the right margin as small and large plain numerals in the brackets, respectively. Linker peptide and cGnRH-II peptide are marked in the box.

Expression and identification of the concatameric cGnRH-II peptide multimers, and generating the cleaved monomeric cGnRH-II analogs We transformed pRSET-B-Δ G-cGnRH-II(M)-4R-STOP plasmid into an *E. coli recA* strain of TOP10F′; then we expressed and analyzed the protein on SDS-PAGE (Fig. 4A). The expressed protein was calculated to be about 8.0 kDa, very close to the deduced protein of 7.3 kDa based on the designed nucleotide sequence (DNASIS™ computer program) and the protein was identified as cGnRH-II peptide multimers after sequencing the N-terminal sequence with PVDF membrane technique. We purified the cGnRH-II peptide multimers with FPLC, cleaved it with trypsin, and analyzed the cleaved peptide on Tricine gel. The cleaved fragment was very close to the expected 1.4 kDa cGnRH-II analog (Fig. 4B). We transferred the cleaved fragment to the PVDF membrane, analyzed the amino acid sequence and obtained the expected amino acid sequence of cGnRH-II analog (Glu-His-Trp-Ser-His-Gly-Trp-Tyr-Pro-Gly-Gly-Lys), cGnRH-II-(Gly11, Lys12) as shown in Fig. 3.

Discussion

In this study, we investigated the effective production of short peptide monomers using genetic engineering techniques. We completed the following tasks; 1) designed the basic unit of the modified peptide by linking a cleavable linker peptide on both ends of the original peptide; 2) designed and synthesized the synthetic DNA with cohesive ends for the modified peptide; 3) constructed tandemly repeated synthetic DNA cassette by repetitive cloning; 4) produced a concatameric short peptide multimers; 5) produced short peptide monomer analogs effectively. The essence of the technique developed in this study are attaching the cleavable linker peptide at both ends of the original peptide to make the modified short peptide as a unit and designing the synthetic DNA of the modified short peptide with cohesive ends to make tandemly repeated synthetic DNA cassette. To design a proper linker peptide, a proper proteolytic enzyme should be chosen first, which has no cleavage site in the sequence of short peptide. Secondly, one of the best linker peptides should be designed, which will decide the amino acid sequence of the final monomeric short peptide analog and also affect the production of the concatameric short peptide multimers by charge and hydrophobicity. We strongly believe that we can effectively increase the expression level of other short peptide analogs with this technique.

Through these experiments we confirmed that the designed 3′ cohesive overhang cGnRH-II(M) synthetic DNA was ligated easily to produce three repetitions and

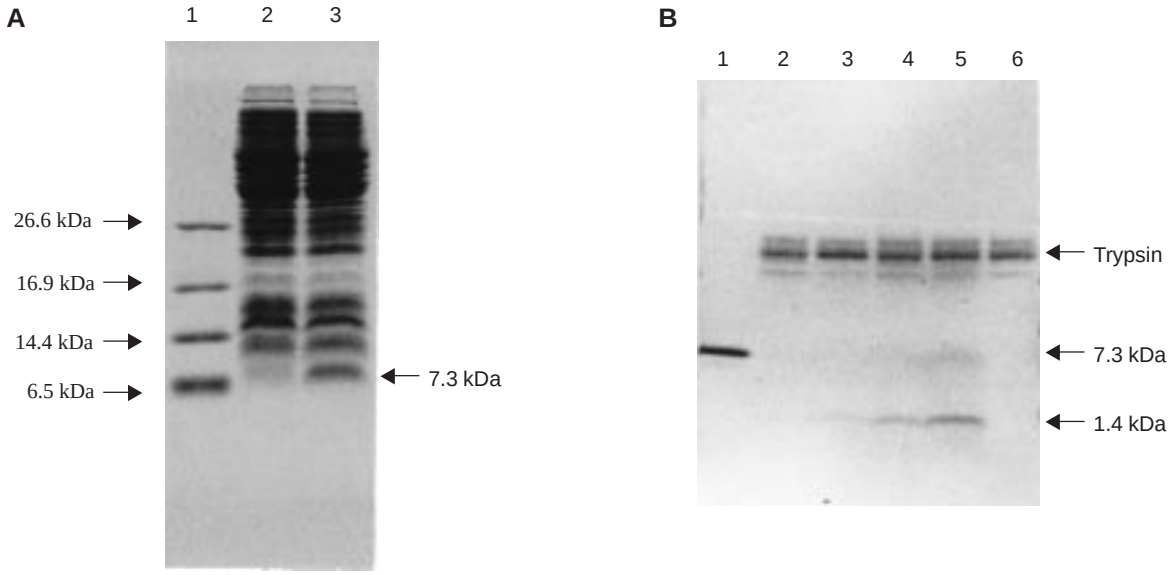

Fig. 4. Expression of the cGnRH-II peptide multimers and generating the cleaved monomeric cGnRH-II peptide analog. **A.** SDS-PAGE analysis of the crude cGnRH-II peptide multimers containing four repeats of the cleavable linker peptide-cGnRH-II peptide sequence with a final cleavable linker peptide. The molecular masses of the standard polypeptides are indicated on the left and the 7.3 kDa cGnRH-II peptide multimers indicated by an arrow on the right. Lane 1, Marker protein; Lane 2, pRSET-B-cGnRH-II(M)-3R-STOP control; Lane 3, pRSET-B-Δ G-cGnRH-II(M)-4R-STOP. **B.** Tricine gel analysis of the cleaved monomeric cGnRH-II peptide analogs. The FPLC purified cGnRH-II peptide multimers was digested with trypsin and the cleaved cGnRH-II peptide analog monomers was analyzed. Arrows indicate trypsin, the purified 7.3 kDa cGnRH-II multimers, and the cleaved 1.4 kDa cGnRH-II peptide analog monomers, respectively, on the right. Lane 1, 2 μg of cGnRH-II peptide multimers control; Lanes 2–5, various amounts of cGnRH-II peptide multimers (lane 2, 2 μg; lane 3, 4 μg; lane 4, 6 μg; lane 5, 10 μg) were digested with 1 μg of trypsin; Lane 6, 1 μg of trypsin control.

remained stable in the *recA*+ strain of JM101 during cloning. Even the constructed expression vector containing four repetitive cGnRH-II(M) synthetic DNA sequences was stable in the *recA* strain of *E. coli* TOP10F′ throughout our research. To increase the efficiency of expression, we replaced the leader peptide region of the expression vector with the nucleotide of the Met-linker peptide-cGnRH-II portion based on the ΔG value of RNA secondary structure whose cGnRH-II nucleotide is slightly different from the originally designed nucleotide of cGnRH-II of cGnRH-II(M)-3R cassette, which would reduce the internal homology among the basic units of the cGnRH-II(M) nucleotides. After introducing the designed nucleotide sequence of Met-linker peptide-cGnRH-II portion with low ΔG value, the 7.3 kDa cGnRH-II peptide multimers were highly expressed as shown in Fig. 4A.

Concatameric expression of the four repetitions of cleavable linker peptide-cGnRH-II peptide sequences attached with a final cleavable linker peptide would be one of the first successes in the high yield production of short peptides. Concatameric peptide multimers can be converted into the monomeric short peptide analogs quantitatively by cleaving the linker peptide with proper enzymes or chemical reagent as shown in Fig. 4B. In this study we used the trypsin cleavable linker peptide sequence of Gly-Lys-Arg, which is also considered to serve as a signal for proteolytic processing and C-terminal amidation (Nikolics *et al*., 1988). If we choose Lys or Arg as the linker peptide, the resultant cGnRH-II analog will receive Lys or Arg at the C-terminal. Therefore the used linker peptide will decide the stuffer amino acid(s) of the cleaved short peptide analog.

We conclude that our technique is one of the most simple and effective methods for the production of short peptide multimers repetitively by cloning the designed basic unit of the modified short peptide synthetic DNA repetitively. The produced short peptide multimers can be collected as original peptides or their analogs after cleaving the linker peptide by the established method of protein chemistry. The method can be easily applied for producing many short peptide analogs and bioactive peptide analogs including hormones, enzymes, immunochemical proteins, antibacterial proteins, and other proteins for laboratory research and industrial purposes.

Acknowledgments We are grateful to Dr. Seon Mi Park for PVDF membrane work and to Dr. David Bastin for reading this manuscript. This work was supported by the MAFF-SGRF (Ministery of Agriculture, Forestry and Fisheries-Special Grants Research Program) in Korea.

References

Hartly, J. L. and Gregori, T. J. (1981) Cloning multiple copies of a DNA segment. *Gene* **13**, 347–353.

Jiang, S. W., Trujillo, M. A., and Eberhardt, N. L. (1996) An efficient method for generation and subcloning of tandemly repeated DNA sequences with defined length, orientation and spacing. *Nucleic Acids Res*. **15**, 3278–3279.

Kim, S. C. and Szybalski, W. (1988) Amplification of cloned DNA as tandem multimers using BspMI-generated asymmetric cohesive ends. *Gene* **15**, 1–8.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

Lee, J. H., Skowron, P. M., Rutkowska, S. M., Hong, S. S., and Kim, S. C. (1996) Sequential amplification of cloned DNA as tandem multimers using class-IIS restriction enzymes. *Genet. Anal*. **13**, 139–145.

Matsudaria, P. (1987) Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem*. **262**, 10035–10038.

Messing, J. (1983) New M13 vectors for cloning. *Methods Enzymol*. **101**, 20–98.

Miyamoto, K., Hasegawa, Y., Nomura, M., Igarashi, M., Kangawa, K., and Matsuo, H. (1984) Identification of the second gonadotropin-releasing hormone in chicken hypothalamus: evidence that gonadotropin secretion is probably controlled by two distinct gonadotropin-releasing hormones in avian species. *Proc. Natl. Acad. Sci. USA* **81**, 3874–3878.

Nikolics, K., Seeburg, P. H., and Ramachandran, J. (1988) The biosynthetic precusor of gonadotropin-releasing hormone; in *Frontiers in Neuroendocrinology*, Martini, L. and Ganong, W. F. (eds.), Vol. 10, pp. 153–166, Raven Press, New York.

Pan, A., Tie, F., Yang, M., Luo, J., Wang, Z., Ding, X., Li, L., Chen, Z., and Ru, B. (1993) Construction of multiple copy of alpha-domain gene fragment of human liver metallothionein IA in tandem arrays and its expression in transgenic tobacco plants. *Protein Eng*. **6**, 755–762.

Sambrook, J., Fritch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.

Taylor, W. H. and Hagerman, P. J. (1987) A general method for cloning DNA fragments in multiple copies. *Gene* **53**, 139–144.

Yanish-Peron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.