0273-2289/95/5001—0035$05.80

Chemical Synthesis and Cloning of Human β-Endorphin Gene in *Escherichia coli*

MYOUNG HEE KIM,* HYUN JU LEE, AND KYUNG T. MIN

Genetic Resources Center, Genetic Engineering Research Institute, KIST, Taedok Science Town, P. O. Box 17, Taejon, 305-606, Korea

Received December 27, 1993; Accepted March 21, 1994

ABSTRACT

Total synthesis of human β-endorphin gene has been designed for the expression in bacterial system. Eight individual oligonucleotides corresponding to the β-endorphin gene were chemically synthesized and joined through the enzyme-catalyzed reaction. The final yield of the 111-nucleotide-long synthetic β-endorphin gene construct was about 10% of the total oligonucleotide used. The synthetic human β-endorphin gene was cloned into the bacterium *Escherichia coli*, using pUC8 vector and shown to have the correct nucleotide sequences as designed.

Index Entries: β-endorphin; human gene; oligonucleotide synthesis; recombinant DNA; *Escherichia coli*.

INTRODUCTION

β-endorphin, consisting of a 31 amino acid residue, is a polypeptide hormone found in the pituitary gland, brain, and peripheral tissues of all vertebrates, and appears to be part of a much larger straight-chain polypeptide called pro-opiomelanocortin (POMC) (*1*). Endorphins are physiological ligands of the opiate receptors in the central nervous system (*2*).

*Author to whom all correspondence and reprint requests should be addressed.

They have been found to be useful in producing a slowed movement of the gut, relaxing smooth muscle, and preventing pain and stress, like opiates, such as morphine. The discovery of the function of β-endorphin as one of the opioid peptide has kindled the hope of finding nonaddictive analgesic drugs. Unlike rigid morphine-type compounds, peptide drugs are expected to be minimally toxic because metabolic degradation should decompose such drugs into nontoxic endogeneous amino acids.

In an effort to produce a large amount of human β-endorphin in a bacterial system, the gene consisting of rather short nucleotides was taken to be synthesized, because the advances in the recombinant DNA technology together with the advent of the automated DNA synthesizer provide a useful approach to the programmed expression of known mammalian gene sequences in bacteria (*3*). Therefore, we report here the complete synthesis of 111-bp-long double-stranded DNA coding for human β-endorphin gene through the enzymatic joining of the synthetic oligonucleotide, and the cloning of it into *Escherichia coli*.

MATERIALS AND METHODS

Bacterial Strain and Plasmid

Escherichia coli strain JM83 (*ara, Δlac-pro, strA, thi, φ80dlacZΔM15*) was obtained from KCTC (Korean Collection for Type Culture, Genetic Engineering Research Institute, KIST, Korea) and used as a recipient for transformation. Plasmid vector pUC8 was purchased from Pharmacia Biotech.

Enzymes and Chemicals

All restriction endonucleases, T4 DNA ligase, polynucleotide kinase, Klenow fragment, and other enzymes were purchased from KOSKO enzymes (Seoul, Korea), New England Biolabs, and Bethesda Research Laboratories, and the enzymatic reactions were done as recommended by the manufacturers. Radiolabeled nucleotides, α-^{32}P-dATP and γ-^{32}P-ATP, were obtained from New England Nuclear. Other dNTPs and 2′, 3′-dideoxynucleotide triphosphates were from P-L Biochemicals.

Oligonucleotide Synthesis

Eight different oligonucleotides (E1–E8) were separately synthesized through the phosphite triester method on solid support using the automated DNA synthesizer (Beckman System Plus) following the modified procedure (*4*).

Kination and the 5′-end Analysis of Oligonucleotide

Oligonucleotides were phosphorylated at the 5′-end in the presence of 10 m*M* Tris-HCl, pH 7.6, 10 m*M* $MgCl_2$, 10 m*M* DTT, 1 m*M* spermidine, 4 μ*M* ATP, 0.6 μ*M* γ-^{32}P-ATP (SA 3000 Ci/mol), and T4 DNA kinase at 37°C for 1 h, by monitoring the aliquots on a PEI plate every 30 min. The 5′-end-labeled oligonucleotides were separately desalted through the SEP-PAK column (Waters Associates), and the relative concentration of each oligonucleotide was calculated based on the radio-sepcific activity of each phosphorylated oligonucleotide.

To analyze the 5′-end nucleotide, oligonucleotide labeled at the 5′-end was incubated with snake venom phosphodiesterase at 37°C for 2 h in 2X snake venom buffer (0.1*M* Tris-H_3PO_4, pH 8.9). An aliquot of each reaction mixture was loaded on the PEI plate, developed in the 1*M* LiCl solution, and autoradiographed. The autoradiograph was analyzed by comparing with the mobilities of cold controls (*5*).

Construction of β-Endorphin Gene Through Enzyme-Catalyzed Reaction

The joining of each oligonucleotide was conducted through the block as well as shot-gun ligation methods (*5*). For the ligation, 150 pmol of each external nonphosphorylated (E1 and E8) and 100 pmol of each internal phosphorylated oligonucleotides (E2–E7) were mixed and boiled for 5 min and annealed by slow cooling to room temperature. The mixture was then incubated at 12°C for 16 h in the presence of 100 μ*M* ATP, 5 m*M* DTT, and T4 DNA ligase, and 3 h further at room temperature. The joined products were identified on 6% denaturing polyacrylamide gel by autoradiography, the band possessing 111-bp-long final products was sliced out of the gel, and the DNA was eluted in the Maxam-Gilbert extraction buffer (*6*) and desalted through SEP-PAK column.

Construction of pβen

Plasmid vector pUC8 was digested with both *Bam*HI and *Eco*RI, and the large fragment was electroeluted from 1% agarose gel. The synthetic β-endorphin gene harboring *Eco*RI and *Bam*HI single-stranded cohesive nucleotides (5 ng) at each end was inserted at the corresponding site in the vector (200 ng) in the presence of T4 DNA ligase.

Total ligation mixture was used for transformation of *Escherichia coli* JM83 following the $CaCl_2$ transformation procedure (*7*). The transformants were selected on the ampicillin plates and identified through the digestion pattern of several restriction endonucleases.

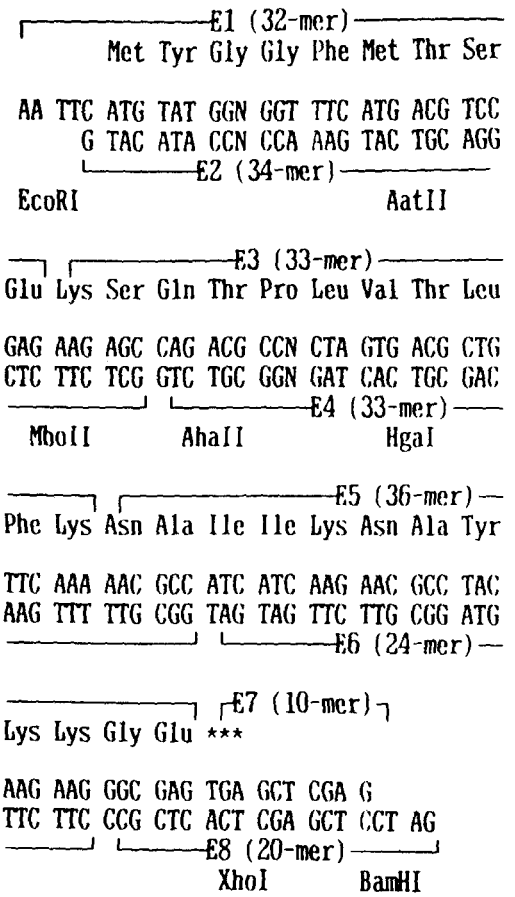

Fig. 1. Design for gene synthesis of human β-endorphin. Initiation (ATG) and termination (TGA) codons are placed at the border of endorphin structural gene for direct expression.

DNA Sequencing Analysis

Slight modification of the dideoxynucleotide sequencing method (*8*) was adopted to determine the nucleotide sequences directly on double-stranded DNA; the denaturation of the template DNA was performed through the incubation of the plasmid DNA at room temperature for 5 min in the presence of 2*M* NaOH and 2 m*M* EDTA and ethanol precipitation (*4*). As a primer, 21-mer oligonucleotide (5′GCTATGACCATGATTACGCCA) was synthesized and used.

RESULTS AND DISCUSSION

The 111-bp-long β-endorphin gene was designed for the chemical synthesis and for the expression in *Escherichia coli*, based on the nucleotide sequence of cDNA for human β-endorphin gene (*1*). With respect to the cutting site, which should not contain the palindromic sequence, eight dif-

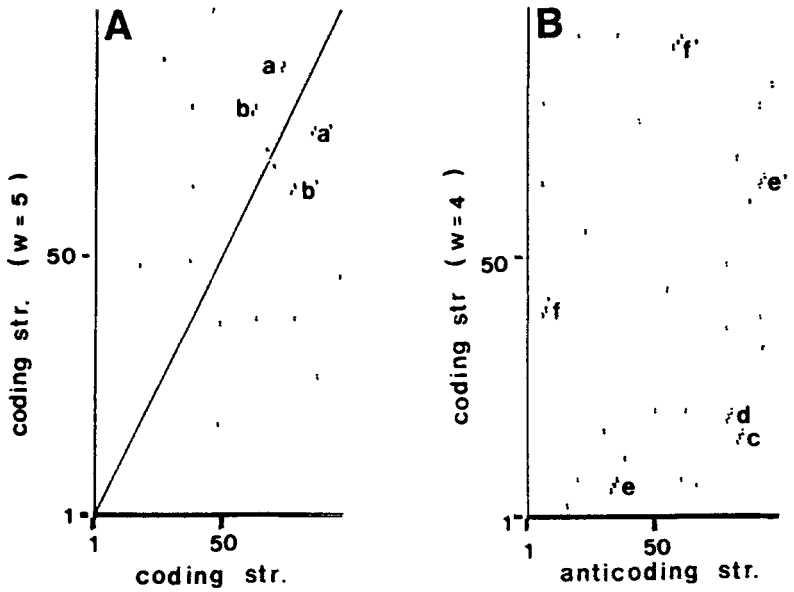

Fig. 2. Dot matrix displacing the homologous nucleic acids in β-endorphin gene. W (window) describes the minimum number of consecutive nucleotides that match between coding and coding (A), as well as coding and anticoding strands, i.e., endorphin coding and its inverted complementary strands (B) (*see text*).

ferent oligonucleotides ranging 10- to 36-mer oligonucleotides were designed to be synthesized (Fig. 1).

In order to avoid the intra- and interoligonucleotide complementarity, which hinders following gene construction, homology searches were performed with PC sequence (PCS) program written by Lagrimini and Brentano. As shown in Fig. 2, there are two pairs of a perfect match of six nucleotides within the coding strand (a/a′ and b/b′ in Fig. 2A), whereas four complete homologies—three perfect matches of six consecutive nucleotides (c, d, and e/e′ in Fig. 2B) and one pair of five consecutive nucleotides (f/f′ in Fig. 2B)—were found between coding and anticoding strands. Since two direct repeat sequences (a/a′ and b/b′) found in coding strand are located in the same oligonucleotide (E5 as well as its complementary oligonucleotide E6), it should not be harmful for annealing during gene construction. Also, two different homologous sequences (c and d) found between the endorphin coding strand and its inverted complementary strand turnd out to be palindromic stretches existing in E1 oligomer (as well as in E2), and their formation of secondary (hairpin) structure is not likely to happen during gene construction. Therefore, we considered only two pairs of homologies (e/e′ and f/f′ in Fig. 2B), and four oligonucleotide (E1, E2, E3 and E4) were chosen to be synthesized with random (i.e., any one of dA, dG, dC, or dT) nucleotides at one position, without changing its corresponding amino acid sequence. The 5′- and 3′- termini of the gene posses *Eco*RI and *Bam*HI cohesive sequences,

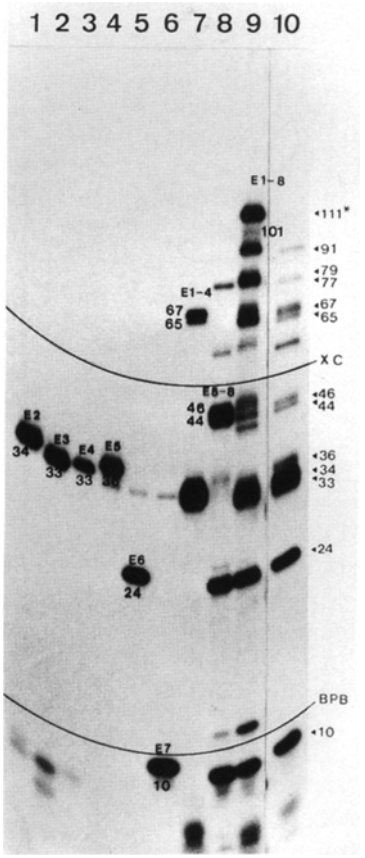

Fig. 3. Autoradiograph of 6% denaturing polyacrylamide gel showing gene construction for human β-endorphin. Lanes 1–6, 5′-phosphorylated oligonucleotides, E2 (34-mer), E3 (33-mer), E4 (33-mer), E5 (36-mer), E6 (24-mer), and E7 (10-mer), respectively; Lanes 7 and 8, half-ligation products of E1–E4, and E5–E10, respectively; Lanes 9 and 10, block and shot-gun ligation products, respectively.

respectively, which enable the constructed gene to be inserted into a cloning vector in one step. The initiation and termination codons were introduced at the border of the endorphin structural gene for direct expression in bacterial system, and the codon usage was not considered to be optimized because it has been known that it is not the absolute necessity for the expression of foreign gene inside of the bacterium *Escherichia coli* (*9*).

The individual oligonucleotide was then chemically synthesized and purified following the method previously reported (*10*). In order to confirm the nucleotide sequences before gene construction, the 5′-end analysis was done after quantitative labeling at the 5′-end with P^{32} and complete digestion with snake venom phosphodiesterase (data not shown). Since the nucleotide at the 5′-end was synthesized after the last coupling step during chemical synthesis, we presumed that the correct

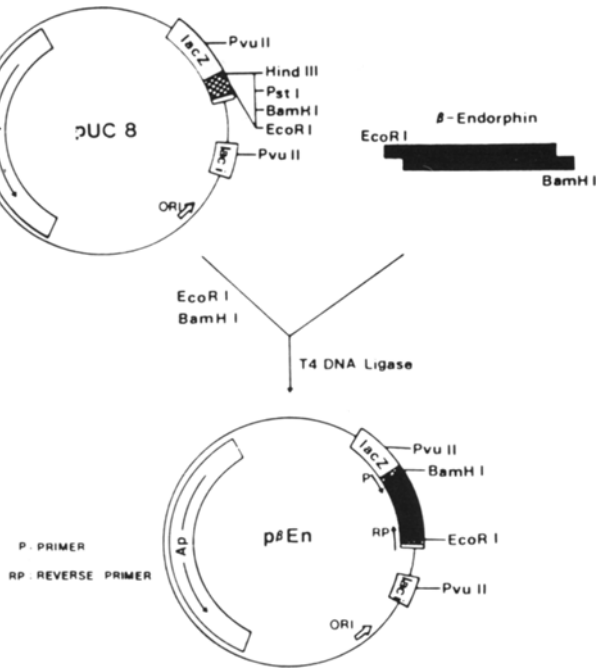

Fig. 4. Construction scheme for pβen.

nucleotide at the 5′-end together with the right gel mobility of each oligonucleotide implies the correct oligonucleotide sequences as designed.

The gene construction was done through block as well as shot-gun ligation method. All internal oligonucleotides with the same molar ratio, which was calculated based on the radio-specific activity after phosphorylation, and two different nonphosphorylated external oligonucleotides with excess molar ratio were adopted for the annealing and joining reaction, and were analyzed on denaturing polyacrylamide gel. As shown in Fig. 3, the total yield of the 111-bp-long synthetic β-endorphin gene construct was about 10% in the case of block ligation (lane 9), whereas the final yield with shot-gun ligation (lane 10) was much less. Although the yield was not as satisfactory, the amount was sufficient for cloning in the next step.

The synthetic gene designed to contain *Eco*RI and *Bam*HI cohesive ends at each end was eluted out of the gel and introduced into the pUC8 vector, which was linearized after digestion of *Eco*RI and *Bam*HI restriction endonucleases (Fig. 4). Since pUC8 was originally designed to lose its α-complementation activity of β-galactosidase when a foreign gene is inserted at the polylinker site, the recombinants were selected on the indicater plate as *Lac*$^-$ cells after the transformation of *Escherichia coli* strain JM83 with constructed plasmids. From the possible recombinant appearing as white colonies on X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside) plate, plasmid was isolated and analyzed with several restriction endonucleases as well as dideoxynucleotide sequencing (Fig. 5). The resulting recombinant

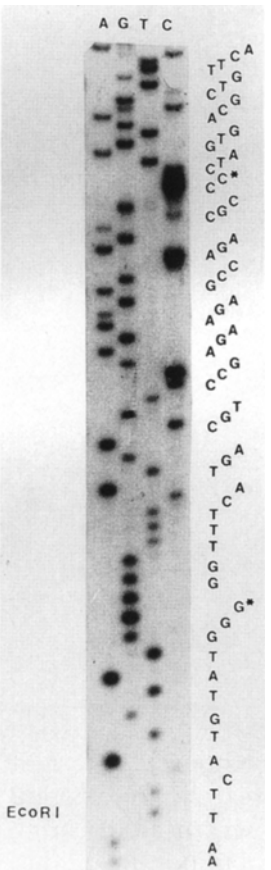

Fig. 5. Dideoxy sequencing gel pattern showing human β-endorphin gene. *Eco*RI site located at the 5′-end of the structural gene was shown. The * indicates the site where random nucleotide was used for the oligonucleotide synthesis.

plasmid, named as pβen, was confirmed to have the correct human β-endorphin gene sequences as designed. We are now under study of overproduction of human β-endorphin in bacterial system.

REFERENCES

1. Chang, A. C. Y., Cochet, M., and Cohen, S. N. (1980), *Proc. Natl. Acad. Sci. USA* **77**, 4890–4894.
2. Akil, H., Watson, S. J., Young, E., Lewis, M. E., Khachaturian, H., and Walker, J. M. (1984), *Ann. Rev. Neurosci.* **7**, 223–255.
3. Oprian, D. D., Nassal, M., Ferretti, L., Karnik, S. S., and Khorana, G. (1986), in *Protein Engineering*, Inouye, M., and Sarma, R., eds., Academic, New York, pp. 111–123.
4. Min, K. T., Kim, M. H., and Lee, D. S. (1988), *Nucleic Acids Res.* **16**, 5075–5088.

5. Lo, K.-M., Jones, S. S., Hackett, N. R., and Khorana, H. G. (1984), *Proc. Natl. Acad. Sci. USA* **81,** 2285–2289.
6. Maxam, A. M. and Gilbert, W. (1980), *Methods Enzymol.* **65,** 499–560.
7. Goodman, H. M. and MacDonald, R. J. (1979), *Methods Enzymol.* **68,** 75–90.
8. Sanger, E., Nicklen, S., and Coulson, A. R. (1977), *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
9. Ernst, J. F. and Kawashima, E. (1988), *J. Biotechnology* **7,** 1.
10. Park, P. C., Lim, H. S., Kim, M. H., and Lee, D. S. (1988), *Korean Biochem. J.* **21,** 214–220.