Research Article

Novel Method of Cell-Free In Vitro Synthesis of the Human Fibroblast Growth Factor 1 Gene

Peijun Zuo and A. Bakr M. Rabie

Orthodontics, Faculty of Dentistry, The University of Hong Kong, 2/F, Prince Philip Dental Hospital,
34 Hospital Road, SAR, Hong Kong

Correspondence should be addressed to Peijun Zuo, pzuo@hkucc.hku.hk

Recombinant DNA projects generally involve cell-based gene cloning. However, because template DNA is not always readily available, in vitro chemical synthesis of complete genes from DNA oligonucleotides is becoming the preferred method for cloning. This article describes a new, rapid procedure based on *Taq* polymerase for the precise assembly of DNA oligonucleotides to yield the complete human fibroblast growth factor 1 (FGF1) gene, which is 468 bp long and has a G+C content of 51.5%. The new method involved two steps: (1) the design of the DNA oligonucleotides to be assembled and (2) the assembly of multiple oligonucleotides by PCR to generate the whole FGF1 gene. The procedure lasted a total of only 2 days, compared with 2 weeks for the conventional procedure. This method of gene synthesis is expected to facilitate various kinds of complex genetic engineering projects that require rapid gene amplification, such as cell-free whole-DNA library construction, as well as the construction of new genes or genes that contain any mutation, restriction site, or DNA tag.

1. Introduction

Recombinant DNA projects generally involve cell-based gene cloning. However, because template DNA is not always readily available, in vitro chemical synthesis of complete genes from DNA oligonucleotides is becoming the preferred method for cloning. This approach is also simpler, more rapid, and more cost-effective. Recently, several methods of in vitro gene synthesis have been described, such as (1) ligation of overlapping single-stranded oligonucleotides comprising the entire gene followed by PCR amplification of the crude ligation mixture using two outer primers [1–3], (2) ligation of preformed duplexes of phosphorylated overlapping oligonucleotides [4, 5], (3) the Fok I method [6], (4) the LCR method [7], and (5) single-step gene assembly from multiple DNA oligonucleotides [8]. Although long oligomers of up to 300 to 600 bases can be directly synthesized, the yield is extremely low [9]. Hence, most of the oligomers in these methods are 25 to 60 nucleotides long [8, 10, 11].

The cloning of prokaryotic genes is relatively easier than that of eukaryotic genes because the former do not contain introns. In fact, eukaryotic genomic DNA cannot be cloned directly because of introns, and cloning procedures need to at least include (1) cell culture, (2) isolation of RNA, (3) reverse transcription from mRNA to cDNA, and (4) amplification of the cDNA. This conventional method has technological problems and is expensive and time-consuming, requiring about 2 weeks to perform [12]. For example, we previously used this approach to clone the open reading frame of the human fibroblast growth factor 1 (FGF1) gene, which is 468 bp long and has a G+C content of 51.5% [13]. The source of the FGF1 gene template was cultured HEK293 cells, from which total RNA was isolated and then reverse transcribed with the use of costly extraction kits. Furthermore, after the nonspecific transcription products were amplified, only one of the PCR products had the same molecular weight as FGF1, but it contained mutations. In other words, the conventional method is inefficient and not cost-effective. These shortcomings prompted us to consider developing a new PCR-directed gene synthesis method.

This new method is ligase-free, the oligonucleotides are designed to cover the complete sequence of both DNA strands, and the full-length molecule is generated in

FGF1

```
  1   ATGGCTGAAG GGGAAATCAC CACCTTCACA GCCCTGACCG AGAAGTTTAA TCTGCCTCCA GGGAATTACA AGAAGCCCAA ACTCCTCTAC TGTAGCAACG
      TACCGACTTC CCCTTTAGTG GTGGAAGTGT CGGGACTGGC TCTTCAAATT AGACGGAGGT CCCTTAATGT TCTTCGGGTT TGAGGAGATG ACATCGTTGC
```

                    BamHI   PstI

```
101   GGGGCCACTT CCTGAGGATC CTTCCGGATG GCACAGTGGA TGGGACAAGG GACAGGAGCG ACCAGCACAT TCAGCTGCAG CTCAGTGCGG AAAGCGTGGG
      CCCCGGTGAA GGACTCCTAG GAAGGCCTAC CGTGTCACCT ACCCTGTTCC CTGTCCTCGC TGGTCGTGTA AGTCGACGTC GAGTCACGCC TTTCGCACCC
```

                                   NcoI

```
201   GGAGGTGTAT ATAAAGAGTA CCGAGACTGG CCAGTACTTG GCCATGGACA CCGACGGGCT TTTATACGGC TCACAGACAC CAAATGAGGA ATGTTTGTTC
      CCTCCACATA TATTTCTCAT GGCTCTGACC GGTCATGAAC CGGTACCTGT GGCTGCCCGA AAATATGCCG AGTGTCTGTG GTTTACTCCT TACAAACAAG
```

```
301   CTGGAAAGGC TGGAGGAGAA CCATTACAAC ACCTATATAT CCAAGAAGCA TGCAGAGAAG AATTGGTTTG TTGGCCTCAA GAAGAATGGG AGCTGCAAAC
      GACCTTTCCG ACCTCCTCTT GGTAATGTTG TGGATATATA GGTTCTTCGT ACGTCTCTTC TTAACCAAAC AACCGGAGTT CTTCTTACCC TCGACGTTTG
```

```
401   GCGGTCCTCG GACTCACTAT GGCCAGAAAG CAATCTTGTT TCTCCCCCTG CCAGTCTCTT CTGATTAA
      CGCCAGGAGC CTGAGTGATA CCGGTCTTTC GTTAGAACAA AGAGGGGGAC GGTCAGAGAA GACTAATT
```

FIGURE 1: *Oligonucleotides used to synthesize FGF1*. A total of 14 oligonucleotides, 50–70 nucleotides in length, collectively encoded both strands of the FGF1 gene.

a single step by overlap extension PCR (OE-PCR). Our protocol is based on the method of Stemmer [8], which consists of synthesis of oligonucleotides, gene assembly, gene amplification, and cell-free cloning. Our method uses the same first and last stages. However, we have combined the middle two steps (gene assembly and amplification) into one step: equal volumes of each oligonucleotide solution are combined and used to generate the full-length template for gene amplification by PCR using the two outer primers.

2. Materials and Methods

2.1. Oligonucleotides. We designed and ordered 14 oligonucleotides, ranging from 50 to 70 nucleotides in length, which collectively encoded both strands of the FGF1 gene (Table 1 and Figure 1). The sense and antisense strands each consisted of 7 oligonucleotides. Complementary oligonucleotides were expected to overlap when the melting temperature (T_m) exceeds 60°C (Figure 1). We incorporated bolded sequences that overlapped with the parts of the sequence of the plasmid pcDNA3.1/V5-His-TOPO (Invitrogen, Carlsbad) to facilitate cloning into this plasmid [13]. Oligonucleotides and all reagents unless otherwise stated were purchased from Sigma-Aldrich (St Louis, MO, USA).

2.2. Polymerase Chain Reaction. We synthesized the FGF1 gene using a single-step PCR method by setting up one reaction containing 14 oligonucleotides (4 pmol of each inner oligonucleotide and 20 pmol of each outer oligonucleotide), as shown in Figure 1. After the addition of 1 U AccuPrime *Taq* High Fidelity polymerase (Invitrogen, Carlsbad), the reaction was carried out in a PCR thermocycler (Takara Bio, Shiga, Japan) for 45 cycles at the following cycling profile settings: 95°C for 3 minutes, 95°C for 30 seconds, 49°C for 1 minute, and 68°C for 1 minute. The final cycle was followed

by an additional 4 minutes at 68°C to ensure complete extension. The holding temperature was 4°C.

2.3. Assembly of Synthesized FGF1 Gene and pcDNA3.1/V5-His-TOPO Vector. Single-step DNA assembly and circularization was performed on nonpurified PCR amplification products by high-fidelity PCR, after the addition of *Taq* DNA polymerase [13]. Each 50-μL sample contained 5 μL 10 × AccuPrime PCR Buffer I, 0.5 μL of raw reaction product mix containing the linear vector backbone, 2 μL of raw reaction product mix including the FGF1 gene, and 0.2 μL AccuPrime *Taq* High Fidelity. Although the ratio of DNA insert to vector should change with the size of the fragments, maintaining the ratio at 4 : 1 for our insert range of 500 bp to 2000 bp gave satisfactory results. The PCR cycling profile (for 35 cycles) was as follows: 94°C for 2 minutes, 94°C for 30 seconds, 52°C for 30 seconds, and 68°C for 3 minutes; the holding temperature was 4°C. Because the primers included complementary directional overhangs, the correct head and tail sequences of the vector and gene were expected to anneal and assemble into the pcDNA3.1/V5-His-TOPO-FGF1 plasmid [13].

2.4. Transformation Procedure. One tube containing 50 μL of DH5α *Escherichia coli* cells was thawed on ice and 1 μL of raw reaction product mix from the combined gene assembly and plasmid insertion step was immediately added and mixed gently. The tube was then incubated on ice for 30 minutes, heated for 60 seconds in a 42°C water bath, and placed on ice for 2 minutes. The contents were added to 950 μL of prewarmed SOC medium (2% tryptone, 0.5% yeast extract, 0.05% NaCl, 2.5 mM KCl, and 20 mM glucose) in a 1.5-mL tube and shaken at 225 rpm at 37°C for 1 hour. Finally, 200 μL was spread onto a prewarmed LB agar plate containing 100 μg/mL ampicillin, and the plate was incubated overnight at 37°C.

Table 1: List of oligonucleotides.

Oligonucleotide	Sequence (5′ to 3′)
pcFGF1F	**gtccagtgtggtggaattgcccttacc**atggctgaaggggaaatcaccacc
pcFGF1R12-71	TTG TAA TTC CCT GGA GGC AGA TTA AAC TTC TCG GTC AGG GCT GTG AAG GTG GTG ATT TCC
pcFGF1F42-101	gaagtttaatctgcctccagggaattacaagaagcccaaactcctctactgtagcaacgg
pcFGF1R76-135	TGT GCC ATC CGG AAG GAT CCT CAG GAA GTG GCC CCC GTT GCTACAGTAGAGGAGTTTGGG
pcFGF1F114-173	gaggatccttccggatggcacagtggatgggacaagggacaggagcgaccagcacattca
pcFGF1R151-210	ATA CAC CTC CCC CAC GCT TTC CGC ACT GAG CTG CAG CTG AAT GTGCTGGTCGCTCCTGTC
pcFGF1F188-247	cggaaagcgtgggggaggtgtatataaagagtaccgagactggccagtacttggccatgg
pcFGF1R225-284	ACA TTC CTC ATT TGG TGT CTG TGA GCC GTA TAA AAG CCC GTC GGTGTCCATGGCCAAGTACTGGCCAGTC
pcFGF1F268-327	ggctcacagacaccaaatgaggaatgtttgttcctggaaaggctggaggagaaccattac
pcFGF1R301-360	CTT CTC TGC ATG CTT CTT GGA TAT ATA GGT GTT GTA ATG GTT CTC CTCCAGCCTTTCCAG
pcFGF1F330-389	cacctatatatccaagaagcatgcagagaagaattggtttgttggcctcaagaagaatgg
pcFGF1R365-424	GGC CAT AGT GAG TCC GAG GAC CGC GTT TGC AGC TCC CAT TCT TCTTGAGGCCAACAAACC
pcFGF1F402-461	cggtcctcggactcactatggccagaaagcaatcttgtttctccccctgccagtctcttc
pcFGF1R	**GCTGGATATCTGCAGAATTGCCCTT** TTA ATC AGA AGA GACTGGCAGGGGG

2.5. Plasmid DNA Extraction. A colony of transformed *E. coli* cells was picked and grown in 5 mL LB medium at 37°C overnight. Plasmid DNA was extracted with the Plasmid Miniprep Kit (Sigma-Aldrich, St. Louis, MO, USA).

2.6. Restriction Enzyme Digestion. Double digestion was performed on a sample of extracted plasmid DNA with *Hin*dIII and *Xho*I (20,000 U/mL and 10,000 U/mL, resp.; New England Biolabs, Ipswich, MA, USA). The two restriction sites were located just beyond either side of the gene. Digestion products were subjected to agarose gel electrophoresis to determine the size of the inserted fragment.

2.7. Sequencing. The final plasmid construct was sequenced by the Genome Research Centre, The University of Hong Kong, with the ABI BigDye sequencing method, according to the manufacturer's instructions (Applied Biosystems, Foster City, CA, USA).

3. Results

3.1. Synthesis and Assembly of the FGF1 Gene Using PCR. To clone the FGF1 gene open reading frame (468 bp) into the pcDNA3.1/V5-His-TOPO vector, we added 27 bp of extra oligonucleotides to the 5′ end of gene and 25 bp extra oligonucleotides to the 3′ end, both with sequences identical to appropriate sites of the vector. The full-length of the FGF1 fragment obtained from double digestion of the PCR reaction was 520 bp (Figure 2). The total oligonucleotides used in the method of Stemmer [8] covered both strands of the DNA to be synthesized, without any intervening gaps

between individual oligonucleotides. To minimize the cost of synthesizing oligonucleotides, we left gaps between each of the 50- to 70-mer oligonucleotides. This strategy allowed us to use only 14 oligonucleotides, compared with 18 used by Stemmer. Greater gaps and shorter overlaps resulted in a 22.22% (4/18) reduction in the cost of synthesizing the oligonucleotides.

3.2. Cloning and Sequencing of the FGF1 Gene. After the crude PCR reaction mixtures that contained the product were directly cloned into the cloning vector, we readily obtained clones that contained the correct sequences. Double digestion of extracted plasmid DNA with *Hin*dIII and *Xho*I yielded two fragments on a 1% agarose gel, of sizes 5433 bp and 561 bp, which represented the linear vector pcDNA3.1/V5-His-TOPO and the FGF1 gene, respectively (Figure 3). Because the fragment yielded by *Hin*dIII and *Xho*I digestion that contained the gene also contained some base pairs from the vector, it was 31 bp longer than the corresponding earlier PCR fragment. The DNA sequence of the FGF1 gene in the complete pcDNA3.1/V5-His-TOPO-FGF1 construct was 100% in agreement with the sequences from the NCBI gene bank (data not shown).

4. Discussion

Our goal of improving cloning technology was to reduce the cost and increase the efficiency of gene synthesis, compared with the conventional cloning method and the method of Stemmer [8]. On the basis of the data presented in this paper, we believe that our protocol of cell-free in vitro gene
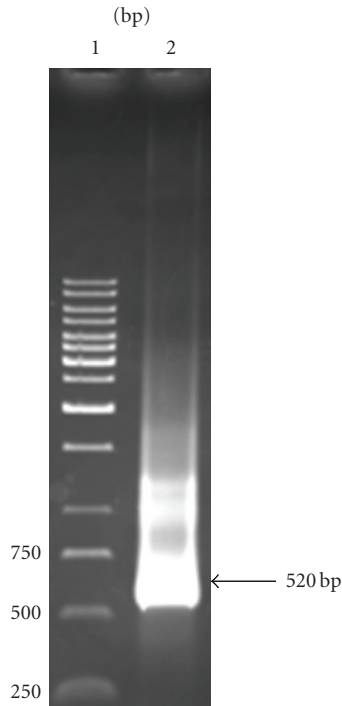
FIGURE 2: *PCR assembly of FGF1*. Lane 1: DNA molecular weight markers. Lane 2: Assembled and amplified FGF1 gene (520 bp).
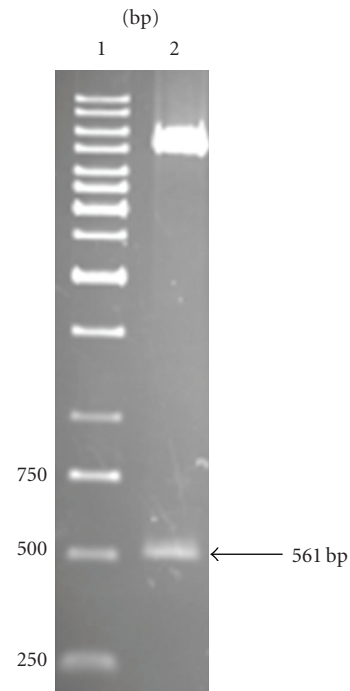

FIGURE 3: *Confirmation of assembled plasmid containing the FGF1 gene*. The pcDNA3.1/V5-His-TOPO-FGF1 plasmid was cleaved by *Hin*dIII and *Xho*I. Lane 1: DNA molecular weight markers. Lane 2: Two digestion fragments of 5433 bp and 561 bp.

synthesis offers a simple, rapid, high-fidelity, and low-cost alternative. The essential feature of our approach is that the entire gene is synthesized in a single step by PCR. It does not require the production of a template before the PCR amplification step. Instead, we combined the two steps of template formation and full-length product amplification into one step. The single-step product did not require further purification and could be directly used for cloning by our previously published method [13].

Our new method would be very useful in the field of molecular biology, especially for the cloning of eukaryotic DNA, because the genes of many eukaryotes contain introns (several per gene) [14]; for example, FGF1 contains at least 5 introns [15]. We had previously tried to clone the FGF1 gene using cultured human HEK293 cells and human stem cells and compared the 2 methods [16]. Both cell types were cultured for at least 2 weeks. Total RNA [17] was isolated from cells with an RNeasy Mini Kit (Qiagen Sciences, Germantown, MD, USA). Reverse transcription [18] of the mRNA was then performed with oligo-(dT) primers and Moloney murine leukemia virus reverse transcriptase (Applied Biosystems, Foster City, CA, USA) to generate the cDNA pool. These procedures required a total of at least 2 weeks and were costly. Our new 2-day method would greatly reduce the cost of materials, personnel, and overheads, and increase the speed by which genes could be cloned.

The new method would also be useful to readily and quickly clone genes from viruses [19], bacteria [20] or prions [21], in a way that would reduce the risk of infection. In addition, gene expression and cloning in *E. coli* is a

standard practice in biomedical research and production. Because each amino acid relates to more than one codon [22], gene sequences need to be optimized [23]. Our new method will help in codon optimization, because a series of degenerate oligonucleotide sequences can be used, and the amount of expression of foreign protein in *E. coli* can be maximised. Another potential application of our method is gene shuffling in protein research. Generally, one would have to choose a group of homologous genes, divide them into small pieces, recombine them, and select hybrids that yield improved protein products or that optimize production. However, it is very difficult to isolate DNA fragments that have fewer than 50 bp in conventional methods because most of the fragments are invisible on an agarose gel. Our cloning method would help in gene shuffling procedures because gene fragments of the desired size can be prepared and included reliably in the gene constructs.

In summary, we have developed a protocol that is a rapid and easy method for gene cloning. This method of gene synthesis is expected to facilitate various kinds of complex genetic engineering projects that require rapid gene amplification, such as cell-free whole-DNA library construction, as well as the construction of new genes or genes that contain any mutation, restriction site, or DNA tag.

Acknowledgments

References

[1] K. Jayaraman and C. J. Puccini, "A PCR-mediated gene synthesis strategy involving the assembly of oligonucleotides representing only one of the strands," *BioTechniques*, vol. 12, no. 3, pp. 392–398, 1992.

[2] K. Jayaraman, S. A. Fingar, J. Shah, and J. Fyles, "Polymerase chain reaction-mediated gene synthesis: synthesis of a gene coding for isozyme c of horseradish peroxidase," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no. 10, pp. 4084–4088, 1991.

[3] K. Jayaraman, J. Shah, and J. Fyles, "PRC mediated gene synthesis," *Nucleic Acids Research*, vol. 17, no. 11, p. 4403, 1989.

[4] R. C. Scarpulla, S. Narang, and R. Wu, "Use of a new retrieving adaptor in the cloning of a synthetic human insulin A-chain gene," *Analytical Biochemistry*, vol. 121, no. 2, pp. 356–365, 1982.

[5] N. K. Gupta, E. Ohtsuka, V. Sgaramella et al., "Studies on polynucleotides, 88. Enzymatic joining of chemically synthesized segments corresponding to the gene for alanine-tRNA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 60, no. 4, pp. 1338–1344, 1968.

[6] W. Mandecki and T. J. Bolling, "FokI method of gene synthesis," *Gene*, vol. 68, no. 1, pp. 101–107, 1988.

[7] F. Barany, "Genetic disease detection and DNA amplification using cloned thermostable ligase," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no. 1, pp. 189–193, 1991.

[8] W. P. C. Stemmer, "Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides," *Gene*, vol. 164, no. 1, pp. 49–53, 1995.

[9] R. B. Ciccarelli, P. Gunyuzlu, J. Huang, C. Scott, and F. T. Oakes, "Construction of synthetic genes using PCR after automated DNA synthesis of their entire top and bottom strands," *Nucleic Acids Research*, vol. 19, no. 21, pp. 6007–6013, 1991.

[10] A. S. Xiong, Q. H. Yao, R. H. Peng et al., "A simple, rapid, high-fidelity and cost-effective PCR-based two-step DNA synthesis method for long gene sequences," *Nucleic Acids Research*, vol. 32, no. 12, article e98, 2004.

[11] L. Young and Q. Dong, "Two-step total gene synthesis method," *Nucleic Acids Research*, vol. 32, no. 7, article e59, 2004.

[12] S. R. Thaker, S. K. Dutta, S. L. Adhya, and B. L. Mattingly-Napier, "Molecular cloning of *Ehrlichia risticii* development of a gene probe for the diagnosis of potomac horse fever," *Journal of Clinical Microbiology*, vol. 28, no. 9, pp. 1963–1967, 1990.

[13] P. Zuo and B. M. Rabie, "One-step DNA fragment assembly and circularization for gene cloning," *Current Issues in Molecular Biology*, vol. 12, no. 1, pp. 11–16, 2009.

[14] M. Csurös, I. B. Rogozin, and E. V. Koonin, "Extremely intron-rich genes in the alveolate ancestors inferred with a flexible maximum-likelihood approach," *Molecular Biology and Evolution*, vol. 25, no. 5, pp. 903–911, 2008.

[15] T. Yoneyama, H. Kasuya, H. Onda et al., "Association of positional and functional candidate genes FGF1, FBN2, and LOX on 5q31 with intracranial aneurysm," *Journal of Human Genetics*, vol. 48, no. 6, pp. 309–314, 2003.

[16] F. L. Graham, J. Smiley, W. C. Russell, and R. Nairn, "Characteristics of a human cell line transformed by DNA from human adenovirus type 5," *Journal of General Virology*, vol. 36, no. 1, pp. 59–72, 1977.

[17] J. C. Lee and R. R. Gutell, "Diversity of base-pair conformations and their occurrence in rRNA structure and RNA structural motifs," *Journal of Molecular Biology*, vol. 344, no. 5, pp. 1225–1249, 2004.

[18] A. Telesnitsky, S. Blain, and S. P. Goff, "Assays for retroviral reverse transcriptase," *Methods in Enzymology*, vol. 262, pp. 347–362, 1995.

[19] E. V. Koonin, T. G. Senkevich, and V. V. Dolja, "Compelling reasons why viruses are relevant for the origin of cells," *Nature Reviews Microbiology*, vol. 7, no. 8, p. 615, 2009.

[20] R. Rappuoli, V. Scarlato, V. Masignani, et al., "Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing," *Science*, vol. 287, no. 5459, pp. 1816–1820, 2000.

[21] A. Aguzzi, "Unraveling prion strains with cell biology and organic chemistry," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 1, pp. 11–12, 2008.

[22] C. B. Anfinsen, "The formation and stabilization of protein structure," *Biochemical Journal*, vol. 128, no. 4, pp. 737–749, 1972.

[23] M. C. Wolf, Y. Wang, A. N. Freiberg, H. C. Aguilar, M. R. Holbrook, and B. Lee, "A catalytically and genetically optimized β-lactamase-matrix based assay for sensitive, specific, and higher throughput analysis of native henipavirus entry characteristics," *Virology Journal*, vol. 6, article 119, 2009.