Clinical Biochemistry 44 (2011) 692–698

Contents lists available at ScienceDirect

Clinical Biochemistry

journal homepage: www.elsevier.com/locate/clinbiochem

A primer design strategy for PCR amplification of GC-rich DNA sequences

Li-Yan Li [a,b,1], Qiang Li [c,1], Yan-Hong Yu [b], Mei Zhong [b], Lei Yang [d], Qing-Hong Wu [e], Yu-Rong Qiu [c], Shen-Qiu Luo [a,*]

[a] Department of Cytobiology, Southern Medical University, Guangzhou 510515, PR China
[b] Department of Gynecology and Obstetrics, Nanfang Hospital, Guangzhou 510515, PR China
[c] Department of Laboratory Medicine, Nanfang Hospital, Guangzhou 510515, PR China
[d] Department of Burns, Nanfang Hospital, Guangzhou 510515, PR China
[e] Laboratory Animal Center, Southern Medical University, Guangzhou 510515, PR China

ARTICLE INFO

Article history:
Received 10 July 2010
Received in revised form 20 December 2010
Accepted 1 February 2011
Available online 18 February 2011

Keywords:
Polymerase chain reaction
GC-rich sequences
Primer design

ABSTRACT

Objectives: To establish a primer design method for amplification of GC-rich DNA sequences.

Design and methods: A group of 15 pairs of primers with higher T_m (>79.7 °C) and lower level ΔT_m (<1 °C) were designed to amplify GC-rich sequences (66.0%–84.0%). The statistical analysis of primer parameters and GC content of PCR products was performed and compared with literatures. Other control experiments were conducted using shortened primers for GC-rich PCR amplifications in this study, and the statistical analysis of shortened primer parameters and GC content of PCR products was performed compared with primers not shortened. A group of 26 pairs of primers were designed to test the applicability of this primer designing strategy in amplifications of non-GC-rich sequences (35.2%–53.5%).

Results: All the DNA sequences in this study were successfully amplified. Statistical analyses show that the T_m and ΔT_m were the main factors influencing amplifications.

Conclusions: This primer designing strategy offered a perfect tool for amplification of GC-rich sequences. It proves that the secondary structures cannot be formed at higher annealing temperature conditions (>65 °C), and we can overcome this difficulty easily by designing primers and using higher annealing temperature.

Introduction

Polymerase chain reaction (PCR), widely applied in many research fields, is one of the routine molecular biology techniques. However, PCR amplifications of GC-rich sequences involve much greater difficulties than those of non-GC-rich ones [1,2]. Although only about 3% of the human DNA sequences are GC-rich, the majority of the important regulatory domains including promoters, enhancers, and control elements consist of GC-rich sequences [3], and most housekeeping genes, tumor suppressor genes, and approximately 40% of tissue-specific genes contain GC-rich sequences in their promoter region [4]. Obviously, ineffective PCR amplifications of these GC-rich DAN sequences hamper the progress of the study into these gene sequences.

The conventional practice for cracking this hard nut of ineffective PCR amplification is adding certain organic additives such as betaine, dimethylsulfoxide, formamide, polyethylene glycerol, non-ionic detergents, 7-deaza-dGTP, dUTP and their combinations [5–16] to the reaction mixture, or jointly using such highly effective DNA polymerase as AmpliTaq™, Taq Gold™, and KOD Hot-Start polymerase [4], Optimase DNA polymerase, Platinum® Taq DNA Polymerase High-Fidelity, etc. [17], during the course of PCR amplification. In addition, techniques such as template denaturation with NaOH, hot start PCR, stepdown PCR, and slowdown PCR can also improve the PCR amplification of GC-rich sequences [18–21]; other factors including adjusting magnesium concentration, buffer pH, denaturing and annealing time and/or temperature, and PCR cycling numbers are sometimes also to be taken into account.

It is widely believed that the major cause of ineffective amplifications of GC-rich templates is the formation of secondary structures such as loops or hairpins brought about by GC-rich DNA templates and/or primers' self-complementary or other conformational features [1].

Primer is a crucial factor for successful PCR amplifications. Precise primer designing and analyzing, especially that of their secondary structures such as self-dimers, hairpins, and cross-dimers [22], before PCR amplifications, are necessary. In this study, an in-depth study was made into fifteen pairs of primers recorded in the related literature for amplifications of GC-rich templates. It was found that

* Corresponding author at: Department of Cell Biology, Southern Medical University, Guangzhou 510515, PR China.
E-mail address: Luoshen888@163.com (S.-Q. Luo).
[1] These authors contributed equally to this work.

0009-9120/$ – see front matter © 2011 The Canadian Society of Clinical Chemists. Published by Elsevier Inc. All rights reserved.
doi:10.1016/j.clinbiochem.2011.02.001

all the primers studied shared some defects such as over-low melting temperature (T_m), over-high level of T_m difference of primer sets (ΔT_m), or mismatching between the primer and multiple loci of the templates. All these defects with the primer will lead to either annealing failure or annealing of multiple loci of the DNA templates, which would further give rise to the ineffective amplifications or non-specific ones. Given that these primer defects are common reasons for invalid PCR amplifications of GC-rich sequences, it just shows that the solution to this problem lies in nothing but the effective primer designing, especially for amplifications of GC-rich DNA templates. In this study, a specific primer designing strategy featuring high T_m and low ΔT_m was established. According to this strategy, fifteen pairs of primers with specific parameters for amplifications of GC-rich DNA sequences and twenty-six more pairs were designed for testing the applicability of this strategy in amplifications of non-GC-rich sequences.

Materials and methods

DNA templates

A total of fifteen GC-rich and twenty-six non-GC-rich DNA target sequences were employed in this study. The fifteen GC-rich DNA sequences are from eight genes: HBA2 (NC_000016.9), FMR1 (NC_000023.10), APOE (NC_000019.9), HRES1 (NC_000001.9), CSTB (NC_000021.8), INSR (NC_000019.9), AR (NC_000023.10), and GATA4 (NC_000008.10), with the GC content of these primers ranging from 66.0% to 84.0%; the twenty-six non-GC-rich DNA target sequences are from twenty-six exons of F8 gene located in Xq28, whose GC content range from 35.2% to 53.5%. (All sequences of these genes were obtained from http://www.ncbi.nlm.nih.gov/nucleotide/).

Genomic DNA is isolated from human white blood cells using the conventional phenol–chloroform method.

Primer designing and optimizing

The fifteen pairs of primers for amplifications of GC-rich sequences were designed manually. In order for the primers to have the highest possible T_m (>79.7 °C) and the smallest possible ΔT_m (1.0 °C), several nucleotides were added to or removed from the 5′ and/or 3′ ends of each primer based on DNA targets, under the precondition that the primers were complementary to the DNA sequence templates. The GC contents of PCR products and the primer parameters including T_m, ΔT_m, the maximum ΔG free energy of self-dimer, cross-dimer, and hairpin were worked out by means of the software Oligo 6.64 (Molecular Biology Insights, Inc., CO).

The fifteen pairs of primers in this study were shortened by reducing two bases from primer 5′ ends as control experiments.

In addition, twenty-six other pairs of primers with higher T_m and lower ΔT_m of primer sets were designed according to the strategy mentioned above, for testing the applicability of the strategy in the amplifications of non-GC-rich sequences. All the primers in this study were synthesized by Invitrogen (Guangzhou, PR China). The parameters of the two groups of primers designed for this study and the fifteen pairs recorded in the related reference literature were detailed in Tables 1-3, respectively.

PCR system and amplification conditions

All PCRs were carried out under the following conditions: 50–200 ng of human genomic DNA, 0.2 μM of each primer, 50 μM of each dNTPs, 2.0 U of Taq DNA polymerase enzyme, and 5 μL of 10× buffer supplemented with 100 mM Tris–HCl (pH 8.3), 500 mM KCl, and 15 mM $MgCl_2$ (Takara, Dalian, PR China). The final volume was adjusted to 50 μL with water.

The reactions were performed in a Gene Amp PCR system 9700 (PE Applied Biosystems). The amplifications of the GC-rich DNA

Table 1
Parameters of fifteen pairs of primers for GC-rich DNA amplifications in this study.

Gene	Primer				Max ΔG free energy (kcal/mol) of primer			PCR products	
	Sequence (5′–3′)	Length (mer)	T_m	ΔT_m	Self-dimer	Cross-dimer	Hairpin	GC %	Size (bp)
INSR	F: CGCGGCCCCCAGCGCCTCTT	20	86.5	0.1	10.3	13.4	2.6	82.8	180
	R: GCCCCGCGGCGCCCAGTAGCA	20	86.6		16.5		1.0		
HRES1 (a)	F: GGCCCGGGCTGGCGGGGCG	19	91.3	1.0	22.2	22.7	5.5	84.0	206
	R: CCCCCGCGCCCCGCGCAC	18	90.3		10.3		5.1		
HRES1 (b)	F: GGAAACACATCCCCAGCTGAGGGC	24	81.0	0.5	10.1	6.2	2.1	68.9	1226
	R: GCAAGAGGAAACTTGAAAAGGCGGATCAC	29	80.5		5.4		1.7		
HRES1 (c)	F: GGAAACACATCCCCAGCTGAGGGCCGGGAG	30	90.1	0.1	10.1	6.7	2.9	68.9	1226
	R: GCAAGAGGAAACTTGAAAAGGCGGATCACGCCAATGC	37	90.0		9.8		5.4		
FMR1	F: GCCCCGCACTTCCACCACCAGCTC	24	85.4	0.4	6.3	6.7	0	80.9	283
	R: GGCGCTCAGCTCCGTTTCGGTTTCACTTC	29	85.8		9.8		1.5		
HBA2 (a)	F: CCCCACAGACTCAGAGAGAACCCACCA	27	81.4	0.6	3.2	5.0	2.0	66.3	885
	R: CAGGTAAACACCTCCATTGTTGGCACATTCC	31	82.0		6.0		1.5		
HBA2 (b)	F: CCCCACAGACTCAGAGAGAACCCACCA	27	81.4	0.0	3.2	5.0	2.0	68.0	764
	R: GGGGAGGCCCAAGGGGCAAG	20	81.4		9.3		4.1		
APOE (a)	F: CCCGGTGGCGGAGGAGACG	19	81.2	0.5	9.8	9.4	2.2	75.5	322
	R: GTCGCGGCCCTGTTCCACCAG	21	81.7		10.3		0.8		
APOE (b)	F: GCCTGGGGCAGGGGGAGAACAGC	23	84.6	0.2	6.6	9.3	2.1	67.4	282
	R: GCCCGACCCCGAGTAGCTCTCCTGAGAC	28	84.4		6.3		2.0		
HBA2 (c)	F: CGCGGGTTGCGGGAGGTGTAG	21	81.2	0.7	10.3	9.4	0	66.2	343
	R: CCGGGATAGAGAGAACCCAGGCACAC	26	80.5		9.8		0		
CSTB (a)	F: CGGCGCCCGGAAAGACGATAC	21	80.0	0.1	16.0	9.8	1.5	76.7	473
	R: CGGGGCCAAAGCGGCTTCTT	20	79.9		9.3		1.8		
CSTB (b)	F: GCCCCCACCCAGCCTGGAG	19	80.2	0.3	6.6	12.4	1.8	76.0	786
	R: CGGGGCCAAAGCGGCTTCTT	20	79.9		9.3		0.9		
GATA4 (a)	F: GCGGGTGCCCTCCTCCGTG	19	81.3	0.5	6.7	6.7	2.4	78.8	510
	R: CCCTCGCGCTCCTACTCACCGAGA	24	81.8		10.3		0		
GATA4 (b)	F: GAGCCTAGAGCCCTTTGCTCAATGCTG	27	79.7	0.7	6.3	9.4	1.9	66.5	500
	R: GGGGTGTAAGCGGCTCCGTCG	21	80.4		6.7		0		
AR	F: TCGGCCGCCGTCCAAGACCTAC	22	81.6	0.2	16.5	19.6	2.3	66.0	380
	R: CGGCGGCTCCAGGCTCTGG	19	81.4		6.6		1.7		