Published in final edited form as:

Mol Biotechnol. 2013 March ; 53(3): 345–350. doi:10.1007/s12033-012-9559-y.

Amplification of GC-rich DNA for high-throughput family based genetic studies

Sadaf Naz^{1,*} and Amara Fatima^{1,2}

Sadaf Naz: naz.sbs@pu.edu.pk; Amara Fatima: amara_lhr@yahoo.com

¹School of Biological Sciences, University of the Punjab, Lahore, Pakistan

Abstract

Researchers face a significant problem in PCR amplification of DNA fragments with high GC contents. Analysis of these regions is of importance since many regulatory regions of different genes and their first exons are GC-rich. There are a large number of protocols for amplification of GC-rich DNA, some of which perform well but are costly. Most of the economical protocols fail to perform consistently, especially on products with >80% GC contents and a size of >300 base pairs. One of these protocols requires multiple additions of DNA polymerase during thermal cycling which therefore rules out its utility if a large number of samples have to be amplified. We have established a method for simultaneous amplification of specific PCR products from a large number of human DNA samples using general laboratory reagents. These amplicons have GC contents ranging from 65–85% and sizes up to 870 base pairs. The protocol uses a PCR buffer containing co-solvents including 2-mercaptoethanol and bovine serum albumin for amplification of DNA. A specific thermal cycling profile is also used which incorporates a high annealing temperature in the first 7 cycles of the reactions. The PCR products are suitable for different molecular biology applications including sequencing.

Keywords

GC-rich; PCR; First exons; DNA amplification; Formamide; DMSO; Annealing

Introduction

Polymerase chain reaction is widely used for different molecular biology applications including mutation detection in genes causing human disorders. Many amplicons may have to be screened during candidate and positional cloning efforts in order to identify genes involved in pathogenesis of disorders. Molecular characterization of inherited disorders in large pedigrees involves mutational screening of DNA from several individuals in the family to establish co-segregation of a mutation with the genotypes. Hundreds of normal controls may also have to be analyzed by PCR in order to eliminate polymorphisms or determine allele frequencies for newly identified mutations. This can be problematic if the mutation lies in a GC-rich sequence.

Many protocols referenced by Mamedov et al. and Wei et al. (1,2) and proprietary reagents are available for amplification of GC-rich DNA. Some protocols are inexpensive (3,4), but

*Correspondence to Dr. Sadaf Naz, School of Biological Sciences, University of the Punjab, Quaid-i-Azam Campus, Lahore, 54590, Pakistan. naz.sbs@pu.edu.pk.

²Equal contribution

Competing Interests:

The authors declare no conflict of interest.

do not perform consistently. Other protocols involve purchase of unusual chemicals or expensive enzymes (5–8).. The highly effective heat pulse extension PCR protocol requires the use of thermal cyclers with fast ramp rates in addition to the use of an expensive enzyme for amplification of GC-rich DNA (7). An inexpensive method has been described to amplify GC-rich DNA which involves the use of primers with very high melting temperatures of 80 °C–90 °C (9). The primers have to be designed manually with great care to many parameters. Moreover, the highest GC content products (80–84%) amplified by this protocol range in size from 180–283 base pairs which are generally too small to be useful when a larger amplicon with identical GC content has to be amplified. SAFE (satisfactory, adaptable, fast and efficient PCR), seems to be the best comparatively low-cost protocol for amplification of a larger sized GC-rich DNA fragment, but it requires multiple additions of *Taq* DNA polymerase during the course of a single experiment (2) and therefore cannot be used for large-scale projects. We have evolved a strategy to amplify GC-rich DNA using a specific PCR buffer with additives and custom designed thermal cycling parameters which include a high initial annealing temperature. The protocol is scalable for amplification of GC-rich templates from a large number of DNA samples.

Materials and Methods

Primer design

Twenty primer pairs were designed to amplify first exons and flanking regions of different genes for sequencing and to detect mutations in control samples by allele specific PCR or size fractionation of the products (Table 1). The primers were designed to have a calculated melting temperature (T_m) of 70 °C–84 °C with Primer3 (v. 0.4.0) software (<http://frodo.wi.mit.edu/>) by manually changing the parameters of maximum primer length to 30, and accepted minimum and maximum primer melting temperatures. However, 4 primer sets had a T_m in the range of 62 °C–65 °C. Primers for allele specific PCR; Tetra primer amplification refractory mutation system (ARMS) (sets 5 and 6) were designed similarly using a free online software (10) with additional adjustments of parameters of “Maximum inner product size”, “Maximum relative size difference of two inner products” and “Minimum relative size difference of two inner products” to 400 bp, 1.9 and 1.5 respectively. Primer sets 18 and 19 were selected since these either gave no PCR product or yielded poor amplification, respectively, with three reported economical methods for amplification of GC-rich DNA (4). An identical product to that amplified by the SAFE protocol with a GC content of 84% was amplified using primer set 20 which was specifically designed for this study to have high T_m using Primer3 software.

DNA extraction and PCR

Human genomic DNA was extracted from blood using a non-organic method (11,12) and dissolved in 10 mM Tris-HCl pH 8.0 and 0.2 mM EDTA pH 8.0 buffer. Reactions were set up in final volumes of 50 μ L containing 5 μ L of a 10X PCR buffer (13) [450 mM Tris-HCl pH 9, 110 mM $(\text{NH}_4)_2\text{SO}_4$, 67 mM 2-mercaptoethanol, 45 μ M EDTA, 1100 μ g/mL BSA, 45 mM MgCl_2], 200 μ M each dNTP, 0.24 μ M each primer, 50 ng DNA, 5% DMSO, 1.25% formamide and 1.2 U of *Taq* DNA polymerase (laboratory prepared or from Fermentas, Glen Burnie, MD). Two PCR buffers supplied with *Taq* DNA polymerase (Fermentas) supplemented with the same final 1X concentration of 4.5 mM MgCl_2 were also used to perform reactions with primers 1, 2, 12, 13, 14 and 20 for comparison with the laboratory prepared buffer. The volume of PCR reaction was reduced to 10 μ L for Tetra primer ARMS PCR and those for size fractionation by capillary electrophoresis.

DNA samples were added first to each 0.2 ml PCR tube and heated for 3 minutes at 94 °C. The PCR mix was added afterwards in a window of 2 minutes so that the DNA had a total of

5 minutes for denaturation. PCR was carried out for a total of 40 cycles, in which the first 7 cycles were performed at 70 °C. The thermal cycling profile is illustrated (Fig. 1). Cycling was carried out on Mycycler (Bio-Rad, Hercules, CA) or 2720 thermal cycler (Applied Biosystems, Foster City, CA, USA). 10 µl of each product was electrophoresed on a 1.5% agarose gel to check for amplification of the DNA and absence of non-specific bands.

Sequencing

Unincorporated dNTP and primers were removed from the PCR products by adding ethanol to a final concentration of 62% to each PCR reaction and centrifugation. The air-dried precipitates were dissolved in 20 µl of water. Big Dye Terminators (BDT) v3.1 (Applied Biosystems) was used to perform sequencing reactions. The reactions were set up in a final volume of 10 µl. Around 3–6 µl of the sample, 0.5–1 µl of the BDT and 1.5 µl of a 5X buffer (400 mM Tris pH 9, 10 mM MgCl₂) were mixed, together with 4 pico moles of either the forward or the reverse primer. DMSO was also added to a final concentration of 10% in order to sequence products of primers 2, 9, 12, 13 and 14. After an initial denaturation at 96 °C for 1 minute, the sequencing reactions were denatured at 96 °C for 30 seconds, the primers were annealed at 50 °C for 30 seconds and extension was performed at 60 °C for 4 minutes for a total of 36 cycles. The sequencing reactions were precipitated with 30 µl of 95% ethanol, 1 µl of 3M CH₃COONa and 1 µl of 125 mM Na₂EDTA. The precipitates were washed with 70% ethanol and the air-dried pellets were dissolved in 10 µl of Hi-Di formamide for loading on ABI 310 capillary electrophoresis system (Applied Biosystems).

Size fractionation for *TPRN*

A primer with the identical sequence as to that of the forward oligonucleotide of set 12 was fluorescently tagged with 6-FAM and the same reverse primer was used to set up the reactions. PCR was performed in a final volume of 10 µl as described above. 2 µl of each product was diluted to 40 µl with water and 1 µl of this was mixed with 10 µl of Hi-di formamide containing LIZ-500 size standard. Electrophoresis was performed on ABI 310 (Applied Biosystems). Traces were inspected in GeneMapperIDV3.2 and the size of the fragment for each sample was observed.

Results and Discussion

We obtained products from all 20 sets of primers, including those that had previously failed to amplify DNA with the addition of many inexpensive enhancers including betaine and ethylene glycol (specifically primer sets 18 and 19), or with the GC-Rich PCR system by Roche Applied Systems (Fig. 2A and data not shown). Although primers 13 and 14 were designed according to all parameters suggested by Li et al (9) while that of 20 had a high melting temperature, these yielded none, or non-specific products when used according to the recommended (9) protocol (data not shown). Amplification was only possible when primers 13, 14 and 20 were used for PCR in conditions described in this manuscript (Fig. 2A).

We chose primer pairs 1, 2, 12, 13, 14 and 20 for examining effects of different PCR parameters on amplification of GC-rich DNA. No products or nonspecific amplification was observed if the PCR buffer was replaced by the two buffers supplied by the vendor (2B, C). Amplification failed for all primers if DMSO was omitted from the reaction (Fig. 2D). The absence of formamide from the PCR reaction led to good amplification with primer pairs 1 and 2, (Fig. 2E). The amplifications from primers 12, 13, 14 and 20 were adversely affected due to omission of formamide (Fig. 2E). The incorporation of initial seven cycles at 70 °C was important for specific amplification (Fig. 2F, G). It was crucial for reproducible success of reactions to add PCR mix after DNA had denatured for 3 minutes (Fig. 2H). For the

amplicons with the highest GC contents (85% and 84%-primer pairs 14 and 20 respectively), primers with melting temperatures of ~70 °C failed to yield products (data not shown). Products were only obtained after using primers with melting temperatures of ~80 °C (Fig. 2A).

We obtained good quality sequence data from all PCR products (Fig. 3). Products of primer pairs 5 and 6 were amplified for allele specific PCR of *SANS* while those of 18–20, were PCR amplified for comparisons with other established protocols and were therefore not sequenced. Using this protocol for amplification and after sequencing, we detected a fifteen nucleotide deletion (14) at the first exon-intron boundary of *SANS* (product GC content 69%) and an eleven nucleotide deletion in *TPRN* (product GC content 80%, Bashir et al. manuscript submitted). Tetra Primer allele specific PCR for the mutation in *SANS* was performed across a region of 68% GC content for one hundred control DNA samples (14) while the DNA from 100 normal individuals were screened for the deletion in *TPRN* by size fractionation of the 80% GC-rich 433 base pair amplified region (Bashir et al, manuscript submitted).

The role of DMSO and formamide in enhancing amplification from GC-rich DNA, as well as in increasing specificity of the amplification has been described previously (15,16). However, results are not consistently reproducible (5). The 10X PCR buffer used in our experiments was formulated to amplify low quantities of DNA (13). Specifically, BSA can increase enzyme stability while 2-mercaptoethanol may serve in reducing secondary DNA structures (17). Withholding reaction mixture till DNA has denatured provides a minimal “hot start” to PCR reactions. This has previously been shown to be important for amplification of GC-rich DNA (3). The use of initial high primer annealing temperature of 70 °C may increase specificity of PCR amplifications by inhibiting mis-priming. Potential for non-specific priming exists when primers have high GC compositions (3).

Our results demonstrate that the combination of the ions in the PCR buffer together with DMSO and the use of a hot start with high initial annealing temperature are important for the success of this protocol. Formamide may play a less critical role if GC contents are lower than 80%. Additionally, primers with calculated high melting temperatures are crucial for success of PCR reactions when product GC contents are between 81 to 85%.

The described protocol can be used to amplify genomic sequences for mutation detection in GC-rich regions of candidate genes. It can also be utilized for screening of a large number of controls in order to determine allele frequencies if a mutation is identified in any of these amplicons. Furthermore, it is noteworthy that currently massively parallel sequencing technologies do not capture GC-rich regions efficiently (18). Therefore, if bona fide disease causing mutations are not identified in the targeted regions, the GC-rich regions can be analyzed by PCR amplification of the corresponding amplicons followed by Sanger sequencing (Our unpublished results).

In summary, we describe a reliable method for amplification of GC-rich DNA which is also cost-effective. Commercial enhancers and systems for amplification of GC-rich DNA are not economical for large-scale projects as the cost may run into thousands of dollars for PCR of a few hundred samples. All PCR components used in the protocol described here are readily available in general molecular biology laboratories and are relatively inexpensive. GC-rich templates can therefore be amplified with no substantial differences in cost from that of general PCR reactions.

Acknowledgments

Some of the primers for amplification of *GNAS* and those of *MCART1* were designed for our collaborative project with University of Lübeck, Germany and were kindly provided for this study by Dr. Christine Klein. We are grateful to Ms. Rasheeda Bashir for independently verifying this protocol with primers for *TPRN*. We thank Dr. Edward Wilcox for his valuable comments on the manuscript. The research was supported by Higher Education Commission, Pakistan, (Grant #836) and Fogarty International Center and National Institute on Deafness and other Communication Disorders, National Institutes of Health, USA, (R01TW007608).

References

1. Mamedov TG, Pienaar E, Whitney SE, TerMaat JR, Carvill G, Goliath R, Subramanian A, Viljoen HJ. A fundamental study of the PCR amplification of GC-rich DNA templates. *Comput Biol Chem*. 2008; 32:452–457. [PubMed: 18760969]
2. Wei M, Deng J, Feng K, Yu B, Chen Y. Universal method facilitating the amplification of extremely GC-rich DNA fragments from genomic DNA. *Anal Chem*. 2010; 82:6303–6307. [PubMed: 20565067]
3. Frey UH, Bachmann HS, Peters J, Siffert W. PCR-amplification of GC-rich regions: ‘slowdown PCR’. *Nat Protoc*. 2008; 3:1312–1317. [PubMed: 18714299]
4. Zhang Z, Yang X, Meng L, Liu F, Shen C, Yang W. Enhanced amplification of GC-rich DNA with two organic reagents. *Biotechniques*. 2009; 47:775–779. [PubMed: 19852763]
5. Baskaran N, Kandpal RP, Bhargava AK, Glynn MW, Bale A, Weissman SM. Uniform amplification of a mixture of deoxyribonucleic acids with varying GC content. *Genome Res*. 1996; 6:633–638. [PubMed: 8796351]
6. Deng J, Wei M, Yu B, Chen Y. Efficient amplification of genes involved in microbial secondary metabolism by an improved genome walking method. *Appl Microbiol Biotechnol*. 2010; 87:757–764. [PubMed: 20376630]
7. Orpana AK, Ho TH, Stenman J. Multiple Heat Pulses during PCR Extension Enabling Amplification of GC-Rich Sequences and Reducing Amplification Bias. *Anal Chem*. 2012
8. Ralser M, Querfurth R, Warnatz HJ, Lehrach H, Yaspo ML, Krobisch S. An efficient and economic enhancer mix for PCR. *Biochem Biophys Res Commun*. 2006; 347:747–751. [PubMed: 16842759]
9. Li LY, Li Q, Yu YH, Zhong M, Yang L, Wu QH, Qiu YR, Luo SQ. A primer design strategy for PCR amplification of GC-rich DNA sequences. *Clin Biochem*. 2011; 44:692–698. [PubMed: 21315705]
10. Ye S, Dhillon S, Ke X, Collins AR, Day IN. An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res*. 2001; 29:E88–88. [PubMed: 11522844]
11. Grimberg J, Nawoschik S, Belluscio L, McKee R, Turck A, Eisenberg A. A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. *Nucleic Acids Res*. 1989; 17:8390. [PubMed: 2813076]
12. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res*. 1988; 16:1215. [PubMed: 3344216]
13. Jeffreys AJ, Neumann R, Wilson V. Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. *Cell*. 1990; 60:473–485. [PubMed: 2406022]
14. Bashir R, Fatima A, Naz S. A frameshift mutation in *SANS* results in atypical Usher syndrome. *Clinical Genetics*. 2010; 78:601–603. [PubMed: 21044053]
15. Sarkar G, Kapelner S, Sommer SS. Formamide can dramatically improve the specificity of PCR. *Nucleic Acids Res*. 1990; 18:7465. [PubMed: 2259646]
16. Winship PR. An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. *Nucleic Acids Res*. 1989; 17:1266. [PubMed: 2922271]
17. Nagai M, Yoshida A, Sato N. Additive effects of bovine serum albumin, dithiothreitol, and glycerol on PCR. *Biochem Mol Biol Int*. 1998; 44:157–163. [PubMed: 9503159]
18. Sirmaci A, Edwards YJ, Akay H, Tekin M. Challenges in whole exome sequencing: an example from hereditary deafness. *PLoS One*. 2012; 7:e32000. [PubMed: 22363784]

Figure 1. PCR profile for amplification of GC-rich DNA

Cycling parameters for performing PCR reactions reported in this paper. Initially, only DNA was denatured for 3 minutes. Later, the PCR reaction mix was added to DNA. This results in five minutes denaturation for DNA and also constitutes a minimal hot start. For samples in 2H, DNA and PCR mix were added together at the beginning of this step. *Annealing temperature of 65 °C, 60 °C or 55 °C (adjusted according to the calculated T_m of the primers) was used instead of 41°C when formamide was omitted from the PCR mix for samples in 2E and 2F.

Figure 2. Electrophoresis analyses of GC-rich templates

DNA amplified with (A)= 10X PCR buffer+Formamide+DMSO, (B)=Fermentas 10X PCR buffer containing KCl+Formamide+DMSO (C) Fermentas 10X PCR buffer containing $(\text{NH}_4)_2\text{SO}_4$ + Formamide+DMSO, (D)=10X PCR buffer +Formamide (No DMSO), (E)=10X PCR buffer +DMSO (No Formamide), (F)=10X PCR buffer +DMSO (No Formamide) and omitting the seven cycles at 70 °C, (G)=10X PCR buffer+Formamide +DMSO and omitting the seven cycles at 70 °C, (H) =10X PCR buffer+Formamide+DMSO and PCR mix heated together with DNA at 94 °C for 5 minutes. PCR products were resolved on a 1.5% agarose gel. 10 μ l of each product was loaded on the gel. DNA was visualized by ethidium bromide staining. M is a DNA size marker, while “-” is a negative control. Numbers above each lane correspond to the primer as designated in Table 1. Each PCR reaction was performed in duplicate and repeated at least twice. Specific amplification

is obtained in the presence of 10X PCR buffer+DMSO with formamide and in 2 cases without formamide. No results are obtained on omission of DMSO or substitution of 10X PCR buffer with vendor's supplied buffers. The numbers in "bp", (base pairs) denote the DNA size.

Figure 3. Sequencing results of GC-rich DNA

Partial sequence traces of PCR products of primer sets 2, 8, 9 and 11 with GC contents of 76%, 72%, 81%, and 71% respectively. All sequences shown here were generated with the forward primers.

Table 1

Information for primers and their products

Name	Gene	Primers	T _m	GC% (primer)	Product size (bp)	GC% (product)
1	<i>GCHI</i>	CGGCTCGAGTGTGATCTAAGCAGGT AGAGGTCGCTGCCACCCAGGAAG	71 72	58 65	697	70
2	<i>GJB2</i>	AGATCGGGACCTCGAAGGGGACTTG GCCCAAGGACGTGTGTGGTCCAG	72 73	60 63	484	76
3	<i>SANS</i>	GGGTGTTTAGGACCCAGGGGAAGA GCTTGGGTCCCTCAACATGTGTC	69 70	58 58	597	69
4	<i>SANS</i>	ATCTCAGGCCACGGTCCCTTTG AGTCGGTGTGACTCCGAGTGG	71 71	61 65	865	66
5	<i>SANS</i>	CCTAATACCGCGCCCTCCCTGC GGCACTGGGTGGGGCTACTCAACC	77 77	71 68	260	68
6	<i>SANS</i>	CCTAATACCGCGCCCTCCCTGC ACATGTGTCTCTCCCAAAACGCCAGGG	77 77	71 57	375	66
7	<i>GNAS</i>	TACACCAC'TACCGAAAAATTGGGAGG ATGGCAGGAGTCTGTTACCTCAGGC	70 71	50 56	689	73
8	<i>GNAS</i>	CGGAGGAGGAAAAAGTACCC CCATCGTCGGACTCGTCTC	62 62	60 63	649	72
9	<i>GNAS</i>	ACCGACACCTCCCTTCCCGC AGAGAGAGACACTGACGGGCGAGGC	75.1 74.5	73 65	606	81
10	<i>MCART1</i>	TCAAC AAGGCTCCACAGAGGGTACTTTG GAGGCTCAGAACTGATTCGGCCCG	70 73	50 63	644	71
11	<i>GPSM2</i>	AACTTCAGTCTGAGGGTTGAGGACACG TCTCAAGCCGAGTTCTGTTCTTG	69 70	52 48	680	71
12	<i>TPRN</i>	GCGGGCTTTCCCAATGGTC CCGGGCACCGTCTCGATGAT	70 70	65 65	433	80
13	<i>TPRN</i>	GAGCTGGGCCCCGCTGCGCGAGAACC GCGGCGGGGGTTCGAACTCTCCAGT	84 84	77 74	282	81
14	<i>TPRN</i>	TCCGCGCCGCCGAGGTGCTGGTG GGCTCAGGGCCGAGTCCCGTTGGAGA	84 83	78 71	337	85
15	<i>TPRN</i>	CAGCGACTTCTCCAGAAAGACC CCCAAGGGATGGTCTCCATA	65 65	59 57	360	72
16	<i>TPRN</i>	CCAGCGCAGTGCCTCTCC GCAGGGTCCGGGAGGTAGGG	70 70	74 75	591	67
17	<i>TPRN</i>	CAGAGCCGTCTCACCCG CCCACGGCCCAAGTCAGGG	69 70	74 74	684	65
18	<i>BIAP3</i>	AGTGCATGGAGGCGGACC GCCAAGAAAGCCCCCTTGTGAG	65 64	67 60	788	65

Name	Gene	Primers	T _m	GC% (primer)	Product size (bp)	GC% (product)
19 [¶]	NA	CAGCAGGGTCCAGGATGG CCCCGTTACCGAAGAGGC	63 63	67 67	734	67
20	GNAS	GGCGCGGCTCCCTCCCTTCC GCTGCCCTTCTCTCGTGGCTGGT	80 80	82 67	845	84

The sizes of amplicons are given in base pairs (bp) and the GC contents of the primers and products are also indicated.
NA, not applicable.

[¶]Oligos^{*} sequences correspond to the previously published primer sequences 4 and 76 reported in the work of Zhang et al. 2009 (reference #4).