General method for amplifying regions of very high G + C content

Charyl M. Dutton, Christine Paynton and Steve S. Sommer*

Department of Biochemistry and Molecular Biology, Mayo Clinic/Foundation, Rochester, MN 55905, USA

Received April 22, 1993; Accepted May 25, 1993

The thermal cycler can be used with maximal efficiency by multiple individuals with different experiments when all PCR amplifications are optimized for one cycling program. More than 150 regions of genomic DNA from human and other species were successfully amplified in our laboratory by using one basic protocol that includes *Taq* DNA polymerase, primers containing 14 to 17 bp of genomic sequence, and standard cycling conditions of one minute at 94°C, two minutes at 50°C, and three minutes at 72°C.

Application of the standard three-step cycling protocol was unsuccessful for regions of the dopamine D4 receptor gene (D4DR), a gene with a very high G + C content (75%). In an effort to retain the standard cycling conditions, *Taq*, *Tth*, *Pfu*, *Vent*, and *Deep Vent* thermophilic polymerases were tried under a variety of conditions. DMSO (1), formamide (2), 7-deaza-2'-deoxy GTP (3), inosine, large PCR primers (up to 47 bp) nested amplification (4), amplification of short segments, T4 gene 32 protein (5), reduced potassium, 'touchdown' PCR (6), and Tricine buffer (7) were used unsuccessfully with *Taq* polymerases (data not shown). Many of the procedures were also tried with the other polymerases. Multiple spurious bands were generally seen and direct genomic sequencing with a nested primer indicated that the sequence of interest was not detectable. Two step cycling parameters, as initially described by (8) (denaturation at 94°C for 1 minute, annealing/elongation at 70°C or 75°C for two to five minutes) also were unsuccessful.

After further experiments, it was found that robust and specific amplification could be obtained with a high G + C (hGC) PCR protocol that included (i) high temperature denaturation (98°C for one minute), (ii) five minutes of combined annealing/elongation at 70°C, (iii) very heat stable thermal polymerases: *Vent* (New England Biolabs) and native *Pfu* (Stratagene), [recombinant *Pfu* (Stratagene) did not work consistently] and (iv) oligonucleotides of Wallace temperatures 80°C or 120°C [$4^{\circ}\text{C} \times (\text{G} + \text{C}) + 2^{\circ}\text{C} \times (\text{A} + \text{T})$] (Figure 1). These four components of the hGC protocol each seem important because alteration of any one parameter eliminated or markedly reduced specific amplification (e.g., reducing denaturation temperature to 94°C; reducing the annealing/elongation time to two minutes; using less thermally stable polymerases such as *Taq* and *Tth*; or using 14 to 17 bp primers that have a Wallace temperature of 48°C to 50°C) (data not shown).

Successful amplification was achieved for each of seven segments (309 to 690 bp in length) in four regions that included different exons of the D4DR gene. In addition, the method was

Figure 1. Sample comparisons of the standard three-step PCR protocol with the high G + C (hGC) PCR protocol. **A.** Four regions of the D4DR gene were amplified with *Taq* polymerase using the standard protocol as described below (lanes 1 through 4) and *Pfu* polymerase using the hGC protocol (lanes 6 through 9). From left to right, the expected sizes of the regions are 402, 469, 321, and 505 bp; respectively. Lane 5—a positive control in which *Taq* polymerase amplifies a 321 bp low G + C region of the factor IX gene. Lane 10—No DNA control. Lane S—size markers composed of *Hae*III digestions of ϕ X174. The doublet in lane 6 may be related to the polymorphic tandem repeat sequences found in this region (9). **B.** A 207 bp region of the POMC gene which contains tandem repeats exhibiting size variation was amplified with standard and hGC protocols using different thermophilic polymerases. Lane 1—a positive control (see lane 5 above). Lanes 2 through 5—amplification of the POMC gene with the standard protocol using *Taq* polymerase and no additions, 5% DMSO, 10% DMSO, or dITP, respectively. Lanes 6 through 9—amplification of the region with *Tth*, *Pfu*, recombinant *Pfu*, and *Vent* polymerase, respectively, with the standard protocol. Lanes 10 through 14—amplification of the POMC gene with *Taq*, *Tth*, *Pfu*, recombinant *Pfu*, and *Vent* polymerases, respectively, with the hGC protocol. **PROTOCOLS:** Standard three-step protocol using *Taq* polymerase: 1 μ M oligonucleotide, 200 μ M dNTP, 250 ng DNA template, 50 mM KCl, 10 mM Tris—Cl pH 8.3, 1.5 mM MgCl₂, 0.5 unit *Taq* DNA polymerase, in 25 μ l volume was cycled 30 times under oil at 94°C for one minute, 50°C for two minutes, and 72°C for three minutes. Amplifications with *Pfu*, recombinant *Pfu*, *Tth*, and *Vent* polymerases were as recommended by the manufacturer. **hGC Protocol using *Pfu* polymerase:** the concentrations of the oligonucleotides, DNA templates, and dNTPs are as described above. Three units of *Pfu* DNA polymerase was added to 20 mM Tris—Cl pH 8.2, 10 mM KCl, 6 mM ammonium sulfate (NH₄)₂SO₄, 2 mM MgCl₂, 0.1% Triton X-100, 10 ng/ μ l nuclease free BSA, and cycled 30 times at 98°C for one minute and 70°C for five minutes.

* To whom correspondence should be addressed

successful for a 207 bp segment of the proopiomelanocortin gene (G + C = 71%) (Figure 1B). Direct genomic sequencing verified that the regions of interest had indeed been amplified (data not shown).

In conclusion, these data suggest that three problems contribute to the difficulty in amplifying high G + C segments with the standard three step cycling protocol: (i) poor denaturation at 94°C, (ii) poor access of the PCR primer due to template secondary structure, and (iii) slow elongation due to template secondary structure. These problems are circumvented by hGC (see legend to Figure 1). This protocol was successful in each of eight segments of very high G + C content (71–75%) that were attempted. Additionally, the hGC protocol also helped to increase the specificity and yield of a 698 bp segment of the D5 dopamine receptor gene which had a *moderately* high G + C content (57%) (data not shown). However, for segments with G + C content significantly greater than 75%, it may be necessary to either increase the denaturation and the annealing/elongation temperatures or to add denaturants such as DMSO or formamide.

ACKNOWLEDGEMENTS

We thank Tammy Lind and Bobbie Gostout, M.D. for performing some of the amplifications. The work was supported by RO1 MH44276.

REFERENCES

1. Pomp, D. and Medrano, J.F. (1991) *BioTechniques* **10**, 58–59.
2. Sarkar, G., Kapelner, S. and Sommer, S.S. (1990) *Nucleic Acids Res.* **18**, 7465.
3. Innis, M.A. (1990) In Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. (eds.), *PCR Protocols: A Guide to Methods and Applications*. Academic Press, Inc., San Diego, pp. 54–59.
4. Haqqi, T.M., Sarkar, G., David, C.S. and Sommer, S.S. (1988) *Nucleic Acids Res.* **16**, 11844.
5. Schwarz, K., Hansen-Hagge, T. and Bartram, C. (1990) *Nucleic Acids Res.* **18**, 1079.
6. Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K. and Mattick, J.S. (1991) *Nucleic Acids Res.* **19**, 4008.
7. Ponce, M.R. and Micol, J.L. (1992) *Nucleic Acids Res.* **20**, 623.
8. Kim, H.-S. and Smithies, O. (1988) *Nucleic Acids Res.* **16**, 8887–8903.
9. Van Tol, H.H.M., Wu, C.M., Guan, H.C., Ohara, K., Bunzow, J.R., Civelli, O., Kennedy, J., Seeman, P., Niznik, H.B. and Jovanovic, V. (1992) *Nature* **358**, 149–152.