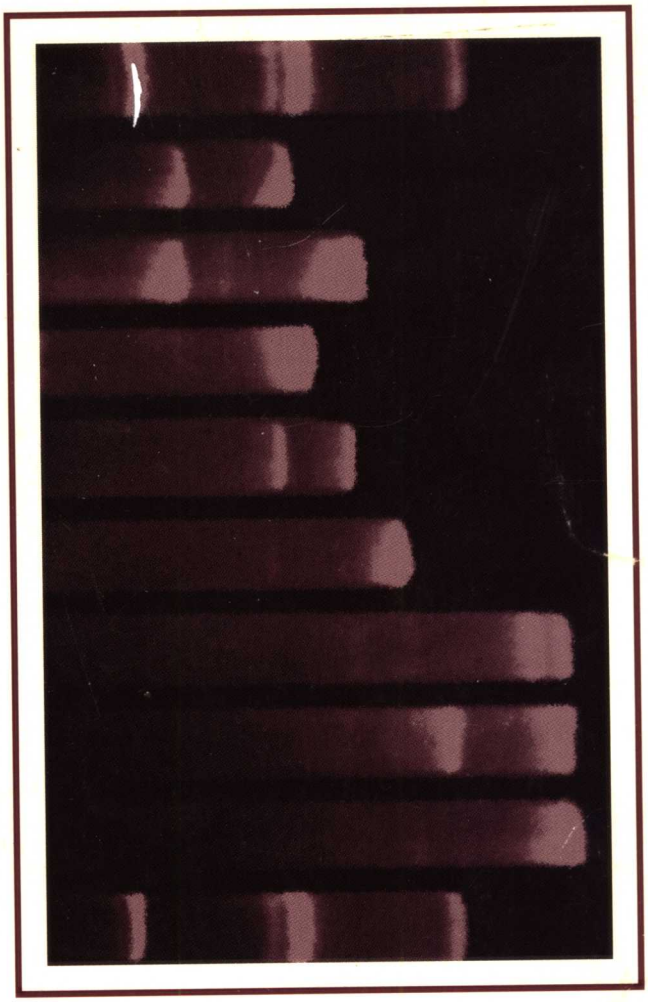

PCR TECHNOLOGY
Current Innovations

Edited by

Hugh G. Griffin
Annette M. Griffin

PCR TECHNOLOGY
Current Innovations

Edited by

Hugh G. Griffin

CRC Press
Boca Raton Ann Arbor London Tokyo

Library of Congress Cataloging-in-Publication Data

PCR technology : current innovations / edited by Hugh G. Griffin, Annette M. Griffin.
p. cm.
Includes bibliographical references and index.
ISBN 0-8493-8674-8
1. Polymerase chain reaction—Methodology. I. Griffin, Hugh G. II. Griffin, Annette M.
QP606.D46P362 1994
574.87'3282—dc20 93-37561
CIP

No claim to original U.S. Government works
International Standard Book Number 0-8493-8674-8
Library of Congress Card Number 93-37561
Printed in the United States of America 4 5 6 7 8 9 0
Printed on acid-free paper

PREFACE

The technique of polymerase chain reaction (PCR)* was first described in 1985. This ingenious tool has had an enormous impact on biological research that can probably be compared to the development of recombinant DNA technology in the 1970s. One of the most important features of the PCR technique is its simplicity. Thus, it can be used to generate meaningful results by scientists who have little or no familiarity with molecular biology techniques; for instance, scientists working in the fields of zoology, botany, environmental science, and forensic science.

Following the appreciation of the true potential of PCR, an explosion of applications of this technique soon occurred. This book brings together a selection of the most widely used applications including the generation and detection of genetic mutations, diagnosis of clinical disease, detection of food-borne pathogens, and the determination of genetic relatedness of plant and animal species. Other chapters deal more closely with the needs of the scientist involved in basic research tasks and cover topics such as the purification and cloning of PCR products, sequencing PCR products, primer design, generation of labeled probes, and screening of lambda and cosmid libraries. To aid the beginner, chapters on primer design, choice of polymerase, and precautions necessary to avoid false positives in PCR are included.

This essential reference book should serve both as the foundation of basic instruction for the scientist new to PCR and a source of updated applications for those already familiar with the basic method.

H. G. Griffin
A. M. Griffin

* The polymerase chain reaction (PCR) is covered by patents owned by Hoffmann-La Roche, Inc. A license is required to use the PCR process.

THE EDITORS

Dr. Hugh G. Griffin and **Dr. Annette M. Griffin** are both senior scientists at the B.B.S.R.C. Institute of Food Research, Norwich Research Park, Colney, Norwich England.

Dr. Hugh Griffin received his training at Trinity College, University of Dublin obtaining a BA (mod) degree in Microbiology in 1983 and a Ph.D. in 1986. He has worked at Washington University, St. Louis, MO, the Babraham Institute, Cambridge, England and the Institute for Animal Health, Huntingdon, Cambridgeshire. His current major research interests relate to the molecular biology of Lactic Acid Bacteria.

Dr. Annette Griffin graduated in 1982 from University College Cork, Cork, Ireland with a B.Sc. (hons) degree in Biochemistry. She obtained her Ph.D. degree in 1986 from Trinity College, Dublin, Ireland for her work on Semliki Forest virus. She has studied the molecular biology of infectious laryngotracheitis virus at the Institute for Animal Health, Huntingdon, Cambridgeshire. Her current research interests are in the molecular biology of bacterial exopolysaccharide biosynthesis.

Both editors have published extensively on various aspects of molecular biology and have acted as editors on a number of books related to DNA sequencing and computer analysis of sequence data. They are married and have three children.

CONTRIBUTORS

Robert P. Adams
Plant Biotechnology Center
Baylor University
Waco, TX

S. F. An
University of Oxford
Nuffield Department of Pathology and Bacteriology
John Radcliffe Hospital
Oxford, U.K.

Asim K. Bej
Department of Biology
University of Alabama at Birmingham
Birmingham, AL

Bruce Budowle
FBI Laboratory
Washington, D.C.

Jean-Paul Charlieu
Institut de Biologie
Montpellier, France

Ernesto d'Aloja
Immunohematology Laboratory
Department of Forensic Medicine
Catholic University of Sacred Heart
Rome, Italy

Tigst Demeke
Agriculture Canada
Lethbridge Research Station
Lethbridge, Alberta, Canada

Marina Dobosz
Immunohematology Laboratory
Department of Forensic Medicine
Catholic University of Sacred Heart
Rome, Italy

K. A. Fleming
University of Oxford
Nuffield Department of Pathology and Bacteriology
John Radcliffe Hospital
Oxford, U.K.

Hugh Griffin
Genetics and Microbiology Department
BBSRC Institute of Food Research
Norwich Research Park
Colney, Norwich, U.K.

Barbara Grubinska
Department of Anatomy
School of Medicine
West Virginia University
Morgantown, WV

Daniel D. Jones
Department of Biology
University of Alabama at Birmingham
Birmingham, AL

Gregory W. Konat
Department of Anatomy
School of Medicine
West Virginia University
Morgantown, WV

Iwona Laszkiewicz
Department of Anatomy
School of Medicine
West Virginia University
Morgantown, WV

K. C. Patrick Lee
Department of Chemical Engineering
Purdue University
West Lafayette, IN

Andrew M. Lew
The Walter and Eliza Hall Institute of Medical Research
Melbourne, Victoria, Australia

Y-M. D. Lo
University of Oxford
Nuffield Department of Pathology and Bacteriology
John Radcliffe Hospital
Oxford, U.K.

Meena H. Mahbubani
Department of Biology
University of Alabama at Birmingham
Birmingham, AL

Vikki M. Marshall
The Walter and Eliza Hall Institute of Medical Research
Melbourne, Victoria, Australia

Michael McClelland
California Institute for Biological Research
La Jolla, CA

J. O'D. McGee
University of Oxford
Nuffield Department of Pathology and Bacteriology
John Radcliffe Hospital
Oxford, U.K.

Louis M. Mezei
Promega Corporation
Madison, WI

Michael Panaccio
Victoria Institute of Animal Sciences
Attwood, Victoria, Australia

Vincenzo L. Pascali
Immunohematology Laboratory
Department of Forensic Medicine
Catholic University of Sacred Heart
Rome, Italy

K. Peter Pauls
Crop Science Department
University of Guelph
Guelph, Ontario, Canada

Leena Peltonen
Department of Human Molecular Genetics
National Public Health Institute
Helsinki, Finland

Lawrence A. Presley
FBI Laboratory
Washington, D.C.

David Ralph
California Institute for Biological Research
La Jolla, CA

Antti Sajantila
Department of Human Molecular Genetics
National Public Health Institute
Helsinki, Finland

Andrew D. Sharrocks
Department of Biochemistry and Genetics
The Medical School
University of Newcastle upon Tyne
Newcastle, U.K.

Anand K. Srivastava
Department of Molecular Microbiology
Washington University School of Medicine
St. Louis, MO

Jörg Stappert
Max-Planck Institute for Immunobiology
Freiburg, Germany

Douglas R. Storts
Promega Corporation
Madison, WI

Bernard Y. Tao
Biochemical and Food Process Engineering
Department of Agricultural Engineering
Purdue University
West Lafayette, IN

Anthony B. Troutt
Immunix Research and Development Corporation
Seattle, WA

Andrey B. Vartapetian
Belozersky Institute of Physico-Chemical Biology
Moscow State University
Moscow, Russia

Richard C. Wiggins
Department of Anatomy
School of Medicine
West Virginia University
Morgantown, WV

E. P. H. Yap
University of Oxford
Nuffield Department of Pathology and Bacteriology
John Radcliffe Hospital
Oxford, U.K.

Kangfu Yu
Department of Molecular Biology and Genetics
University of Guelph
Guelph, Ontario, Canada

TABLE OF CONTENTS

Chapter 1

PCR AS A TECHNIQUE USED DAILY IN MOLECULAR BIOLOGY

Jean-Paul Charlieu

The polymerase chain reaction (PCR) is a very powerful technique in molecular biology and is widely used today for an increasing number of applications. Several are presented in this volume, however many of these were developed for a specific purpose. In this chapter, PCR approaches for more general uses will be presented. Through a few examples, a PCR approach to a problem will be compared to more "classical" approaches in order to show that the former is often easier and faster to perform than the latter. Since PCR generally reduces the number of steps of an analysis and therefore the number of products and enzymes required, the economic character is also taken into account.

The characterization of cloned DNA fragments can sometimes be very time consuming, despite the availability of an increasing number of kits allowing a more rapid utilization of some techniques of molecular biology (e.g., labeling of DNA, molecular hybridization, purification of plasmids).

The determination of the size of DNA fragments inserted into plasmids, for example, requires the growth of bacterial clones followed by the preparation, purification, and enzymatic hydrolysis of plasmid DNAs. In addition, problems may occur at each step of such an analysis: (1) bacteria can grow insufficiently to obtain enough material; (2) the prepared DNA may not be pure enough; or (3) contamination with chemicals (e.g., phenol, chloroform) may inhibit the activity of restriction enzymes resulting in partial hydrolysis of the DNA. When double digestion is necessary to release the cloned fragment, it may be necessary to change the incubation buffer of enzymes that are not compatible. PCR provides a simple way to determine the size of DNA fragments inserted into plasmid vectors. PCR primers can be designed from the vector sequence on both side of the cloning site (Figure 1A). For plasmids of the pUC series or derived from it, the M13 "universal" and "reverse" primers can be used (M13 primer = 5′ GTAAAACGACGGCCAGT 3′; reverse primer = 5′ AACAGCTATGACCATG 3′). A one-step PCR study can be achieved according to the scheme of Figure 1B. The analysis of PCR products in an agarose gel (Figure 1C) allows the direct determination of the insert sizes. These PCR products can also be purified from the gel and sequenced, without the need of producing single-stranded templates from the bacterial strains.[1]

The screening of a library for a given sequence can be performed either by hybridization or by PCR. Here again, the PCR approach is cheaper, easier, and faster. For screening by molecular hybridization, clones have to be first plated and grown (usually overnight), then transferred onto a solid support (nitrocellulose or nylon). In case of bacterial or yeast clones, colonies are then lysed and the DNA bound to the membrane is denatured. A probe must be labeled either radioactively or using biotinylated or digoxigenin-linked nucleotides by nick translation or oligolabeling. After hybridization and washing, the positive clones are detected by autoradiography or enzymatic immunodetection. All these steps must then be repeated at least once in order to obtain pure clones. The PCR strategy consists of amplifying a DNA fragment known as "sequence tagged site" (STS),[2] which is characterized and localized in the genome using pools of clones as templates.[3] The secondary PCR screening is then performed on individual clones of the positive pools. In this PCR approach again, it is not necessary to prepare the DNA templates. Note that the two PCR applications described above can be easily automated.

0-8493-8674-8/94/$0.00+$.50

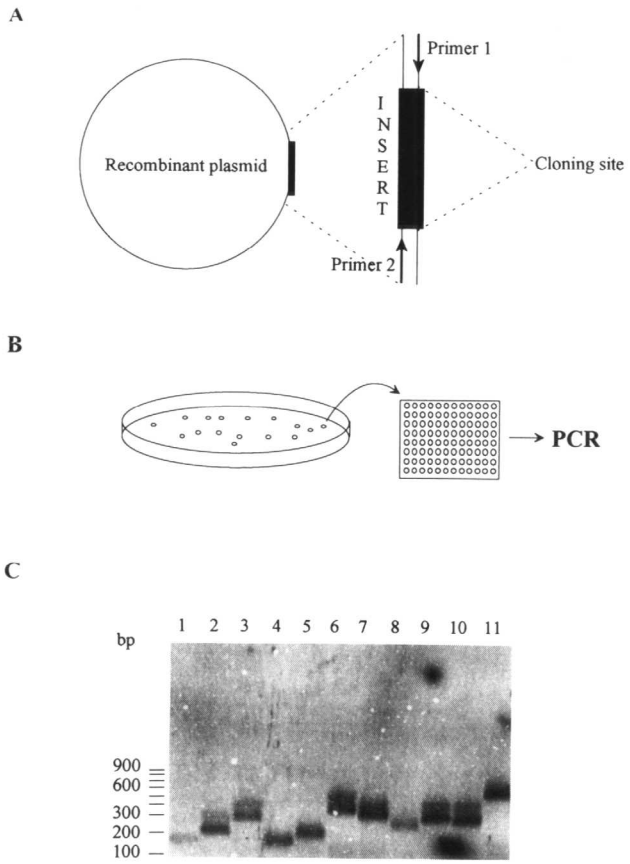

Figure 1. Determination of the insert size of recombinant plasmids by PCR. (A) Position of PCR primers designed from the vector sequence, on both sides of the cloning site. (B) A single colony of each clone is picked up (with a sterile toothpick) from the plate, inoculated in 10 µl of PCR mix containing the Taq DNA polymerase buffer, 20 pmol of each primer, 0.8 m*M* of dNTP (total), and 1 U of Taq DNA polymerase, and 30 cycles of PCR consisting in a denaturation step at 92°C for 10 s, an annealing step at 60°C (for M13 and reverse primers) for 30 s, and an elongation step at 72°C for 1 min are performed. (C) 2 to 5 µl of PCR products are electrophoresed in a 1.5% agarose gel and detected by ethydium bromide staining. PCR amplification of the plasmid vector without insert (lane 1) provides a control of the basic size.

STSs can also serve to characterize overlapping clones, without the establishment and comparison of restriction maps of each clone. This PCR approach is described in Figure 2.

Nelson et al.[4] have developed a method to PCR amplify DNA fragments that are comprised between two Alu repeats in the human genome ("Alu PCR", Figure 3). Alu PCR can be performed with yeast artificial chromosomes (YACs) or cosmids to obtain in a simple and rapid way "Alu fingerprints", which are very useful for the construction of contigs.[5]

Alu PCR can also be performed for other purposes. One of them is the characterization of somatic hybrid cell lines. Cytogenetic identification of the human chromosome(s) present on a rodent genomic background is the classical approach to this problem but is quite long and difficult. In addition, small fragments of chromosomes can escape cytogenetic detection. The Alu PCR pattern specific for each human chromosome allows a molecular characterization of the human DNA in rodent/human somatic hybrid cells.[6]

Probes for *in situ* and Southern blot hybridization are generally obtained by labeling a cloned DNA fragment by nick translation or oligolabeling. PCR is an alternative method used

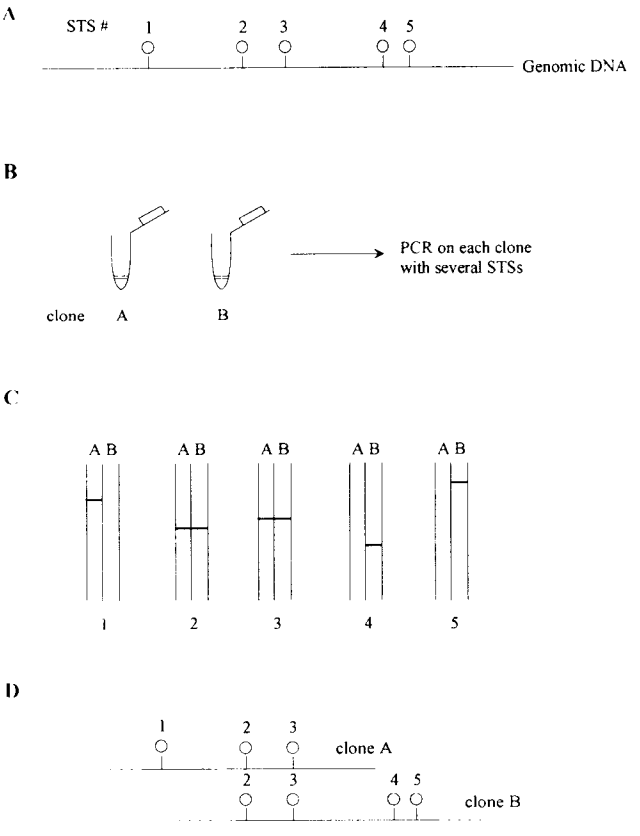

Figure 2. Construction of a contig with STSs. (A) Contiguous STSs are chosen in a given region of the genome. (B) Using PCR, YACs or cosmids are tested for the presence of each STS. If the PCR products are of different sizes, multiplex PCR can be performed. (C) The PCR products are analyzed (presence or absence of the amplified band in each clone) by the appropriate method. (D) In this example, STSs #2 and #3 are common to clones A and B. These clones therefore overlap in the region containing these STSs.

to produce a fragment to be used as a probe from genomic DNA. An improvement to this method is to incorporate a labeled precursor in the *in vitro* synthesized DNA fragment. In this case, the nucleotide mix for PCR should contain a 1:10 molar ratio between the unlabeled nucleotide (0.02 m*M*) and the labeled precursor (0.2 m*M*). In place of a radioactive precursor, it is possible to use cold labeled nucleotides such as biotin-16-dUTP or DIG-11-dUTP. These are particularly useful as *in situ* probes since they are stable for up to one year and can be stored for several experiments.

Alu PCR is also another way to produce probes for *in situ*[7,8] or Southern blot[9] hybridization from a YAC clone or from somatic hybrid cells (painting probes for *in situ* hybridization).[10-12]

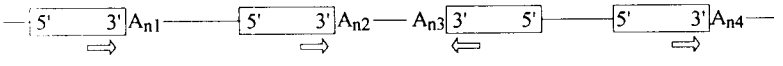

Figure 3. Alu PCR. Alu sequences are found interspersed throughout the human genome.[13] These repeats are polarized (the 3′ end contains a [dA]-rich extension)[14] and can be oriented in any sense.[15] A primer (represented by an arrow) directed toward the 3′ end of Alu sequences will allow the *in vitro* amplification of DNA fragments contained between two "tail-to-tail" Alu repeats.

This can be performed directly from a yeast colony or cultured cells without the DNA preparation step.

The PCR approach cannot substitute for all others currently used in molecular biology, but it provides a good alternative in many cases, and can be used for the study of most problems. Note also that PCR does not need heavy equipment and that the thermocycler can be used for purposes other than PCR, such as incubation of samples at the appropriate temperature for enzymatic reactions and sequencing using the chain-termination method.

REFERENCES

1. **Casanova, J.-L., Pannetier, C., Jaulin, C., and Kourilsky, P.,** Optimal conditions for directly sequencing double-stranded PCR products with sequenase, *Nucleic Acids Res.*, 18, 4028, 1990.
2. **Olson, M., Hood, L., Cantor, C. R., and Botstein, D.,** A common language for physical mapping of the human genome, *Science*, 245, 1434, 1989.
3. **Kwiatkowski T. J., Jr., Zoghbi, H., Ledbetter, S. A., Ellison, K. A., and Chinault, A. C.,** Rapid identification of yeast artificial chromosome clones by matrix pooling and crude lysate PCR, *Nucleic Acids Res.*, 18, 7191, 1990.
4. **Nelson, D. L., Ledbetter, S. A., Corbo, L., Victoria, M. F., Ramirezsolis, R., Webster, T. D., Ledbetter, D. H., and Caskey, C. T.,** Alu polymerase chain reaction — a method for rapid isolation of human specific sequences from complex DNA sources, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 6686, 1989.
5. **Chumakov, I., Rigault, P., Guillou, S., Ougen, P., Billaut, A., Guasconi, G., Gervy, P., LeGall, I., Soularue, P., Grinas, L., Bougueleret, L., Bellanné-Chantelot, C., Lacroix, B., Barillot, E., Gesnoin, P., Pook, S., Vaysseix, G., Frelat, G., Schmitz, A., Sambucy, J.-L., Bosch, A., Estivill, X., Weissenbach, J., Vignal, A., Riethman, H., Cox, D., Patterson, D., Gardiner, K., Hattori, M., Sakaki, Y., Ichikawa, H., Ohki, M., LePaslier, D., Heilig, R., Antonorakis, S., and Cohen, D.,** Continuum of overlapping clones spanning the entire human chromosome 21q, *Nature*, 359, 380, 1992.
6. **Ledbetter, S. A., Garcia-Heras, J., and Ledbetter, D. H.,** "PCR caryotype" of human chromosomes in somatic cell hybrids, *Genomics*, 8, 614, 1990.
7. **Lengauer, C., Green, E., and Cremer, T.,** Fluorescence *in situ* hybridization of YAC clones after Alu-PCR amplification, *Genomics*, 13, 826, 1992.
8. **Baldini, A., Ross, M., Nizetic, D., Vatcheva, R., Lindsay, E., Lehrach, H., and Siniscalco, M.,** Chromosomal assignment of human YAC clones by fluorescent *in situ* hybridization: use of single-yeast-colony PCR and multiple labeling, *Genomics*, 14, 181, 1992.
9. **Charlieu, J.-P., Laurent, A.-M., Orti, R., Viegas-Péquignot, E., Bellis, M., and Roizès, G.,** A 37-kilobases DNA fragment common to the pericentromeric region of chromosomes 13 and 21 and to the inactive centromere on chromosome 2, *Genomics*, 15, 576, 1992.
10. **Desmaze, C., Zucman, J., Delattre, O., Thomas, G., and Aurias, A.,** *In situ* hybridization of PCR amplified inter-Alu sequences from a hybrid cell line, *Hum. Genet.*, 88, 541, 1992.
11. **Lichter, P., Ledbetter, S. A., Ledbetter, D. H., and Ward, D. C.,** Fluorescent *in situ* hybridization with Alu and L1 polymerase chain reaction probes for rapid characterization of human chromosomes,
12. **Cotter, F. E., Hampton, G. M., Nasipuri, S., Bodmer, W. F., and Young, B. D.,** Rapid isolation of human chromosome-specific DNA probes from a somatic cell hybrid, *Genomics*, 7, 257, 1990.
13. **Deininger, P. L.,** SINEs Short interspersed repeated DNA elements in higher eucaryotes, in *Mobile DNA*, Howe, M. and Berg, D., Eds., ASM Press, Washington, DC, 619–636.
14. **Weiner, A. M., Deininger, P. L., and Efstratiadis, A.,** Nonviral transposon: genes, pseudogenes and transposable elements generated by the reverse flow of genetic information. *Annu. Rev. Biochem.*, 55, 631, 1986.
15. **Slagel, V., Flemmington, E., Traina-Dorge, V., Bradshaw, H., and Deininger, P. L.,** Clusturing and subfamily relationships of the Alu family in the human genome, *Mol. Biol. Evol.*, 4, 19, 1987.

Chapter 2

THE DESIGN OF PRIMERS FOR PCR

Andrew D. Sharrocks

TABLE OF CONTENTS

I. INTRODUCTION

The development of the polymerase chain reaction (PCR) has revolutionized the field of molecular biology. PCR has been used for a plethora of applications, many of which are covered in this book. These applications involve both novel procedures (e.g., gene amplification from nanograms of genomic DNA[1,2]) and modifications of existing methods (e.g., site-directed mutagenesis[3,4]). Although many variables need to be optimized in the design of PCR-based procedures for each of these various applications, the most critical parameter in all cases is the correct designing of PCR primers. Indeed, the correct choice of PCR primers often dictates the success or failure of the PCR amplification. Careful design of primers can therefore save valuable research time, and in addition, can lead to significant savings in costs as the primers usually represent the most expensive component in a PCR.

One of the principal considerations in the design of a PCR protocol is to obtain unique, specific products as dictated by the selected primers. The first step in PCR primer design is to ensure this specificity. However, after specificity has been assured, further manipulation of the PCR primer design is possible. This allows the introduction of novel genetic information into the product. Such alterations range from single point mutations to the tagging of products with new coding sequences or regulatory elements. By careful and thoughtful primer design, specific products can be produced from a PCR with a multitude of possible engineered features.

Most of the rules for primer design are empirical with no guarantee of success. However, careful adherence to these rules will significantly increase the probability of a successful PCR. Computer programs are especially useful in the assessment of the basic parameters governing primer design. Even so, some primer pairs that fulfill all known criteria still fail to work for obscure reasons. This chapter provides a guide to reducing the possibility of an unsuccessful PCR but simultaneously demonstrates the flexibility that can be incorporated into primer design for product manipulation.

0-8493-8674-8/94/$0.00+$.50

II. GENERAL RULES FOR PRIMER DESIGN

Successful primer design and hence successful PCRs rely on the unique annealing of the two primers to the template with both high specificity and high efficiency. This ensures that only the desired product is synthesized. The problem of non-specific amplifications is intensified during the early rounds, when amplifications are performed on very small quantities of target DNA, which is often immersed in an excess of non-specific sequences. Correct annealing at this stage is imperative or errors will be compounded throughout the ensuing PCR.

Several parameters must be carefully considered in order to ensure correct annealing (Table 1). The first of these is to choose primers that have a sequence unique within the region to be amplified. The most important region to check is at the 3′ end of the primer as this is where synthesis of the PCR product begins. Such a procedure is tedious when executed manually but is handled easily by computer programs (see below). The second parameter to consider is the inclusion of a G/C residue at the 3′ end of the primer. This "G.C clamp" helps to ensure correct annealing at the 3′ end due to the strong hydrogen bonding utilized by G/C base pairs.

Primers should also be designed with no self-homology. Such self-homology can lead to partially double-stranded "snap back" structures that render the primer incapable of hybridizing to the template. A general rule of thumb is that no self-homology involving four contiguous base pairs should be present in the primer. A related parameter is that primers should show no homology to their antisense counterparts. Formation of partial hybrids between primer pairs can lead to the formation of "primer-dimers" in the ensuing PCR. Elimination of this artifact is essential as this by-product can easily swamp a PCR. Particular care should again be taken in removing complementarity between the 3′ ends of the two primers.

The primer base composition should also be closely monitored. In general a G/C content between 45 and 55% should be selected to direct specific binding yet allow efficient melting during the PCR. Efforts should also be made to keep the base composition close to that exhibited by the amplified region. In addition, the base distribution of the primers should be random, with polypurine and polypyrimidine tracts avoided. Nucleotide sequence repeats should also be avoided in primer design. The target sequence-specific part of primers should ideally be between 18 and 25 bases long. It is also important to have primer pairs with similar melting temperatures (T_m). This can be accurately calculated using the nearest-neighbor method with the formula: $T_m^{primer} = \Delta H/[\Delta S + R \ln (c/4)] - 273.15°C + 16.6 \log_{10} [K^+]$ where ΔH and ΔS are the enthalpy and entropy for helix formation, respectively, R is the molar gas constant, and c is the concentration of probe.[5] However, approximate T_ms can be calculated manually using the simpler formula $T_m = 2AT + 4GC$.[6] An equity in primer T_ms ensures simultaneous annealing of the primers. The calculated T_m can then be incorporated into the PCR protocol to optimize specific binding. This can be exploited when two sets of primer pairs with different matched T_ms are used in a single PCR to amplify different specific fragments.

An additional parameter that can be incorporated into primer design is to ensure that the T_m of the amplified region between the primers is low enough to ensure 100% melting at 92°C. This can be calculated using the formula $T_m = 81.5 + 16.6 (\log_{10}[K^+]) + 0.41 (\%G + C) - 675/\text{length}$.[7] This reduces to $T_m = 59.9 + 0.41 (\%G + C) - 675/\text{length}$ at standard PCR conditions containing 50 m*M* KCl. Inefficient melting will ultimately lead to a reduced yield from the PCR. Finally it is useful to design primers whose annealing sites are spaced between 100 and 600 bp. This distance allows efficient synthesis of product during the PCR.[8]

Adherence to the above parameters in designing primers for a PCR helps ensure specificity of the product. Such specificity is essential in applications where the non-target DNA is in great excess over the target DNA. Many of these PCR applications, covered in this volume, include procedures involving amplifications from genomic DNA, gene libraries, and whole cells.