2 Overview

2.1 Principles of RACE

Rapid Amplification of cDNA Ends (RACE) is a procedure for amplification of nucleic acid sequences from a messenger RNA template between a defined internal site and either the 3´ or the 5´ end of the mRNA (1). This methodology of amplification with single-sided specificity has been described as “one-sided” PCR (2) or “anchored” PCR (3). PCR requires two sequence-specific primers that flank the sequence to be amplified (4,5). However, to amplify and characterize regions of unknown sequences, this requirement imposes a limitation (3).

3´ RACE takes advantage of the natural poly(A) tail found in mRNA as a generic priming site for PCR. In this procedure, mRNAs are converted into cDNA using reverse transcriptase (RT) and an oligo-dT adapter primer. Specific cDNA is then amplified by PCR using a gene-specific primer (GSP) that anneals to a region of known exon sequences and an adapter primer that targets the poly(A) tail region. This permits the capture of unknown 3´-mRNA sequences that lie between the exon and the poly(A) tail.

5´ RACE uses an antisense gene specific primer for the synthesis of specific cDNA by reverse transcriptase. Prior to PCR, a TdT-tailing step attatches an adapter sequence to the unknown 5´ sequences of the cDNA. Specific cDNA is then amplified by PCR using a GSP that anneals in a region of known exon sequences and an adapter primer that targets the 5´ terminus.

RACE has been used for amplification and cloning of rare mRNAs (6) and may be applied to existing cDNA libraries (7). Additionally, RACE products can be directly sequenced without any intermediate cloning steps (8,9), or the products may be used to prepare probes (10). Products generated by the 3´ and 5´ RACE procedures may be combined to generate full-length cDNAs (10,11). Lastly, the RACE procedures may be utilized in conjunction with exon-trapping methods (12) to enable amplification and subsequent characterization of unknown coding sequences.

2.2 Summary of the 3´ RACE System

The 3´ RACE procedure is summarized in figure 1. First strand cDNA synthesis is initiated at the poly(A) tail of mRNA using the adapter primer (AP). After first strand cDNA synthesis, the original mRNA template is destroyed with RNase H, which is specific for RNA:DNA heteroduplex molecules. Amplification is performed, without intermediate organic extractions or ethanol precipitations, using two primers: one is a user-designed GSP that anneals to a site located within the cDNA molecule; the other is a universal amplification primer that targets the mRNA of the cDNA complementary to the 3´ end of the mRNA. Two universal amplification primers are provided with the system. The universal amplification primer (UAP) is designed for the rapid and efficient cloning of RACE products using the uracil DNA glycosylase (UDG) cloning method (13–16). The abridged universal amplification primer (AUAP) is homologous to the adapter sequence used to prime first strand cDNA synthesis.

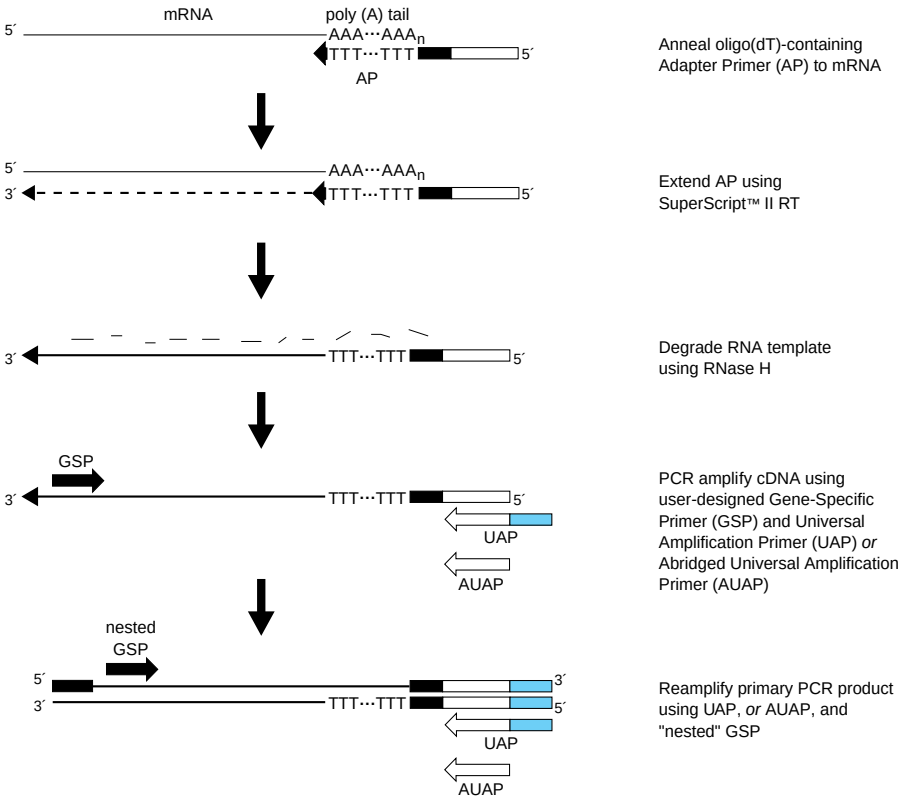

Figure 1. Summary of the 3´ RACE System procedure.

Since the 3´ RACE System utilizes the poly(A) tail region as an initial priming site, multiple amplification products may be synthesized, depending on the degree of specificity conferred by the GSP. To generate a specific amplification product, the user may find it advantageous to design a second "nested" GSP, as recommended by Frohman *et al.* (10) and reamplify the RACE products; this procedure is discussed in greater detail at the end of this chapter.

2.3 Isolation of Total RNA

One of the most important factors preceding the synthesis of substantially full-length cDNA is the isolation of intact RNA. The quality of the RNA dictates the maximum amount of sequence information that can be converted into cDNA. Thus, it is important to optimize the isolation of RNA from a given biological source and to prevent adventitious introduction of RNases (17) and inhibitors of reverse transcriptase such as guanidinium salts, SDS and EDTA (18). RNA can be isolated using a variety of methods. The recommended method for 3´ RACE is the guanidine isothiocyanate/acid-phenol method originally described by Chomzynski and Sacchi (19). The TRIzol® Reagent method is an improvement of the original single-step method of Chomczynski and Sacchi (20) and can be used for the preparation of RNA from as little as 10^3 cells or milligram quantities of tissue (21). Total RNA isolated with TRIzol® Reagent is undegraded and essentially free of protein and DNA contamination. The TRIzol® RNA isolation protocol is described in Section 6.2. For the isolation of RNA from small quantities of sample (<10^6 cells or <10 mg tissue) without using phenol, the GlassMAX RNA Microisolation Spin Cartridge System is recommended (22).

Total RNA isolated by these methods may contain small amounts of genomic DNA that may subsequently be amplified along with the target cDNA. The presence of this DNA is not likely to cause problems because it lacks the poly(A) region present in the mRNA analyte. As a precaution, however, we recommend performing a control experiment without reverse transcriptase to determine whether a given fragment is of genomic DNA or of cDNA origin. Products generated in the absence of RT are of genomic origin. If your application requires removal of all genomic DNA from your RNA preparation, refer Section 5.3, *DNase I Digestion of RNA Preparation.*

2.4 First Strand cDNA Synthesis from Total RNA

The first strand cDNA synthesis reaction is catalyzed by SuperScript™ II RT. This enzyme is a mutant of M-MLV RT that has been engineered to reduce RNase H activity, resulting in greater yields and more full-length synthesis (23,24,25). The enzyme exhibits increased thermal stability and may be used at temperatures up to 50°C. In addition, SuperScript™ II RT is not inhibited significantly by ribosomal and transfer RNA and may be used to synthesize first strand cDNA from a total RNA preparation. The RNA template is removed from the cDNA:RNA hybrid molecule by digestion with RNase H after cDNA synthesis to increase the sensitivity of PCR (26).

The AP (see figure 2), which primes first strand cDNA synthesis, has been engineered to contain three restriction endonuclease sites and a *Not* I half-site. Inclusion of these sequences in the primer may facilitate post-amplification cloning using either a restriction endonuclease-based (27) or a T4 DNA polymerase-based (28) method. Because the AP initiates cDNA synthesis at the poly(A) region of the mRNA, it effectively selects for polyadenylated mRNAs; thus, oligo(dT)-selection for poly(A)$^+$ RNA is typically not necessary although incorporating this step may facilitate the detection of rare mRNA transcripts.

2.5 Amplification of a Target cDNA

Amplification of a target cDNA requires priming with two oligonucleotides and *Taq* DNA polymerase. The sense amplification primer is the user-provided GSP,

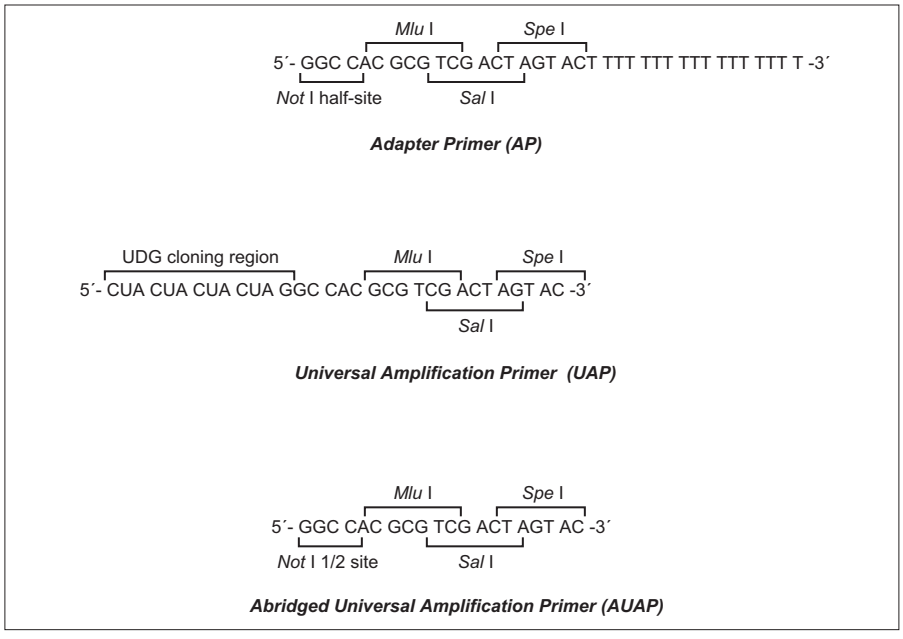

Figure 2. AP, UAP, and AUAP primer sequences.

which is specific for the particular gene or sequence of interest and may be designed to include sequence elements that facilitate subsequent cloning steps. The antisense amplification primer is one of the two universal amplification primers (see figure 2) provided with the system. The AUAP contains a restriction endonuclease site sequence (adapter region) homologous to the adapter region of the AP. The UAP is composed of the same adapter region plus a dUMP-containing sequence at the 5´ end of the primer required for UDG-mediated cloning. The UAP should not be used to prime DNA synthesis with any archaeobacterial polymerase (*e.g., Pyrococcus furiosus, Pyrococcus woesei,* etc.) or any long PCR enzyme mixture (*e.g.,* Elongase® Enzyme Mix) that contains one of these enzymes due to the inhibition of polymerase activity by dUMP-containing DNA. Both the AUAP and the UAP will function in PCR at annealing temperatures up to 68°C.

2.6 Design of the Gene-Specific Primer

Efficient and specific PCR amplification is highly dependent on primer design. This is especially true for RACE applications since the PCR is carried out with only a single GSP. In general, effective primers form stable duplexes with their target sequences, are highly specific for their target sequences, and are free of secondary structure such as hairpin loops and dimers (29–31). Additionally, the complementarity of primer 3´-termini must be minimized since primer-dimer artifacts may significantly reduce PCR efficiency. Therefore, dimer formation with the AUAP or UAP primer, as well as itself, should be reduced. Computer algorithms that have been developed (32–35) and are commercially available often facilitate this analysis. Discussion of primer design for RACE applications may be found in Frohman (11) and Loh (6). It should be noted that in cases where only limited peptide sequence information is available, a degenerate GSP may be prepared.

The AUAP and UAP included with the system have been engineered to function at PCR annealing temperatures up to 68°C and to facilitate the cloning step. The user-defined GSPs need to be compatible with the cloning method. Add the following to the 5' end of the GSP:

for UDG cloning:	5´–CAU CAU CAU CAU–3´	(use with UAP)
for T4 DNA polymerase cloning:	5´–CGA–3´	(use with AUAP)

2.7 Nested Amplification

The AP is designed to synthesize first strand cDNA from all polyadenylated mRNAs. The sequence specificity in the amplification reaction is therefore derived solely from the GSP. Often, a second "nested" GSP may be utilized in conjunction with the AUAP or UAP in a second amplification reaction to give the 3´ RACE procedure the specificity of a second primer (9). The nested GSP can anneal immediately adjacent to the first GSP or at sequences within the cDNA further downstream. The nested amplification reaction may be conveniently conducted using a plug of agarose from the gel analysis of the initial 3´ RACE reaction (see Section 5.5, *Nested Amplification from an Agarose Plug*). Ultimately, the 3´ RACE procedure should produce a single, prominent band on an agarose gel. When performing 3´ RACE with a nested primer, sequences specific for subsequent cloning manipulations (see Section 2.6, *Design of the Gene-Specific Primer*) must be designed into the nested GSP.

2.8 Cloning of Amplification Products

Conventional cloning methods that typically involve end-repair and blunt-end cloning can be problematic for amplified products (36–38). An alternative is a rapid and efficient method involving the use of UDG (13–16). This method requires that the user design a GSP containing containing a 5'-$(CAU)_4$ sequence. Incorporation of dUMP into the GSP may be accomplished on most automated synthesizers or with Invitrogen Custom Primers (see Section 2.6, *Design of the Gene-Specific Primer*). The product of the 3´ RACE reaction primed with the UAP and the dUMP-containing GSP is treated with UDG, which converts dUMP residues to abasic sites (39,40), to generate 3´ overhangs. The directional nature of the UDG cloning process can be exploited to lend an added level of specificity to the RACE procedure. Only amplification product that results from priming by both the UAP and the appropriately designed GSP are efficient substrates for UDG cloning.

Another alternative to conventional cloning methods uses the 3´ to 5´ exonuclease activity of T4 DNA polymerase as the basis for cloning as described by Stoker (28). In this procedure, the AUAP is used in the amplification reaction, and the 3´ RACE products are treated with T4 DNA polymerase to generate a *Not* I 5´ overhang. Similarly, the user may design a site into the GSP (see Section 2.6, *Design of the Gene-Specific Primer*).

Another approach to cloning is to digest the 3´ RACE product using one of the restriction endonuclease sites designed into the AUAP (figure 2; ref. 1). The user may also design unique restriction sites into the GSP, exploit a site present in the cDNA sequence or end-repair the 3´ RACE product prior to restriction endonuclease digestion (37).