The effects of internal primer–template mismatches on RT-PCR: HIV-1 model studies

Cindy Christopherson*, John Sninsky and Shirley Kwok

Roche Molecular Systems Inc., 1145 Atlantic Avenue, Alameda, CA 94501, USA

Received August 2, 1996; Revised and Accepted November 22, 1996

ABSTRACT

We investigated the effects of internal primer–template mismatches on the efficiency of reverse transcription and PCR amplification. As models, RNA transcripts representative of different HIV-1 group M subtypes were evaluated with a previously described gag primer pair system. We observed that the presence of two to four mismatches in the primer–template duplexes did not have a significant effect on RT-PCR. However, the presence of five and six mismatches with the 28 and 30 base primers reduced PCR product yield by ~22- and 100-fold respectively, relative to the homologous template. The amount of reduction was reproducible from experiment to experiment and was independent of the initial copy number input. Under the conditions used, viral RNA measurements of the more divergent HIV-1 subtypes (A and E) would be underestimated, while isolates of subtypes B, C, D and F–H are expected to be efficiently amplified and accurately measured. The reduced amplification efficiency for targets similar to HIV subtypes A and E can be improved 4- to 10-fold by lowering the annealing temperature and implementing a reverse transcription step that gradually increases in temperature. The additional substitution of either 5-methylcytosine for cytosine throughout or the substitution of inosine at positions of variable bases resulted in a <4-fold difference in product yield between the homologous and most divergent templates.

INTRODUCTION

Nucleic acid-based assays such as the polymerase chain reaction (PCR), the ligase chain reaction (LCR), nucleic acid sequence-based amplification (NASBA) and branched chain DNA (bDNA) rely on the efficient hybridization of oligonucleotides to the targeted sequence. Mismatches between the oligonucleotides and the targeted nucleic acid can affect duplex stability and may compromise the ability of a system to amplify and/or detect the targeted sequence. The effects of mismatches depend on numerous factors, including the length of the oligonucleotide, the nature and position of the mismatches, the temperature of hybridization, the presence of co-solvents and the concentrations of oligonucleotides as well as monovalent and divalent cations (1).

The sequence heterogeneity of human immunodeficiency virus type 1 (HIV-1) challenges efficient detection with nucleic acid-based assays. HIV-1 is divided into groups M and O (2–5). There are presently eight known subtypes within Group M, designated A–H. Subtype A is predominant in Central Africa, B in North America, Europe, South America and Thailand, C in South Africa and India, D in Central Africa, E in Thailand, India and Central Africa and subtypes F–H have been reported in Central Africa. Group O, not yet divided into subtypes, is considered to be more divergent than Group M and to date has been found in a few infected individuals from Cameroon, Gabon and France (4,5). In the US, only six non-B subtypes have been reported so far; one subtype D isolate from a Zairian student (6) and three subtype E, one subtype D and one subtype A from five US servicemen (7).

Using HIV-1 as a model system, we previously reported on the effects of single 3′-terminal mismatches on PCR product yield (8). We found that primers that terminated in a T allowed significant amplification even when mismatched with C, G or T. In this study we systematically examined the effects of multiple internal primer–template mismatches on RT-PCR using templates that represent various HIV-1 subtypes. The amount of product generated with each template was determined with a quantitative RT-PCR assay (9).

MATERIALS AND METHODS

Model system

HIV-1 primers SK462, d(AGTTGGAGGACATCAAGCAGCCT-A TGCAAAT), and SK431, d(TGCTATGTCAATGTCCTTTCTTG- GTTCTCT), and probe SK102, d(GAGACCA TCAA TGAG- GAAGCTGCAGAA TGGGA T), were used in this study. The primers were specifically designed to be longer than usual in order to better accommodate mismatches. These primers amplify a 142 bp HIV-1 gag region which is highly conserved among the subtype B isolates. However, in non-subtype B isolates, as many as six mutations have been observed in the upstream primer binding region of subtype B isolates and five in the downstream primer binding region of subtype F. To systematically evaluate the effect of mismatches on RT-PCR, a series of templates were constructed that harbored mutations representative of different subtypes (Fig. 1). The templates (designated M1–M8) were engineered to contain mismatches to either the SK462 (M1–M4) or SK431 (M5–M8) primer binding regions; the internal sequences, including the probe binding region were identical in all constructs. A template that is completely homologous to both SK462 and

*To whom correspondence should be addressed. Tel: +1 510 814 2891; Fax: +1 510 522 1285; Email: cindy.christopherson@roche.com
Construction of mutant and control templates

The mutant and control templates were generated by amplification of an HIV-1-containing recombinant plasmid, pSYC1857 (10), with mutagenic primers. Two sets of primers were used to construct the templates (see Fig. 2). To generate templates with various mutations in the SK462 region, common primer D was coupled with mutagenic C primers for the first round of amplification. The second round of amplification was performed with common primers A and B. To generate templates with mutations in the SK431 region, mutagenic D primers were coupled with common primer C for the first round of amplification and common primers A and B for the second round of amplification. Primer A overlaps with primer C by 18 bases; primer B overlaps primer D by 17 bases. Mutations in the SK462 and SK431 regions were selected based on known mutations found in different HIV-1 subtypes (4). A control template that is completely homologous to both primers was similarly constructed.

To construct the templates, 10⁴ copies of plasmid pSYC1857 were amplified with 50 pmol of the first set of primers under conditions that minimize nucleotide misincorporation. Specifically, 100 µM each of dATP, dGTP, dCTP and dTTP, 1.5 mM MgCl₂ and 2.5 U AmpliTaq DNA polymerase (Perkin Elmer) were used. Stringent cycling conditions were also used: 95 °C for 10 s, 50 or 60 °C for 10 s and 72 °C for 15 s for 30 cycles. The product from the first round was purified by gel electrophoresis and 50 µl of a 10⁶ dilution of the product from the first round of amplification was reamplified with 50 µM of the second set of primers under the high fidelity conditions listed above. In order to generate templates that mimic the native sequence, dTTP rather than dUTP was used.

RNA transcripts

RNA transcripts were prepared directly from the amplified products using a MEGAscript™ transcription kit (Ambion, Austin, TX). Residual DNA was removed from the RNA transcripts by treating with DNase. Proteins were removed by extraction with phenol:chloroform and the RNA was precipitated with ethanol. After resuspension, the transcripts were purified over an Oligo-tex-dT™ column (Qiagen) as recommended by the manufacturer.

Normalization of templates

The RNA transcripts were normalized to ensure that the same copy number of each test template was used in the evaluation of primer–template mismatches. To normalize the input, the templates were amplified with primers that were completely homologous to all eight templates. Following amplification, the products were quantified on microwell plates coated with SK102 probe as previously described (9). Either 10⁴ or 10⁵ copies of each template were subsequently used in the mismatch study.

Quantification of the effects of mismatches

To test the effect of mismatches on RT-PCR, the standard cycling conditions recommended for this primer pair were first used: 2 min at 50 °C (for uracil N-glycosylase cleavage of any potential carryover of dUMP-containing PCR product from previous reactions) (11); 60 °C for 30 min (for reverse transcription); four cycles consisting of a denaturation step (95 °C) for 10 s, annealing (50 or 55 °C) for 10 s and extension (72 °C) for 10 s; followed by 24 cycles of 90, 60 and 72 °C for 10 s each. Amplified products were serially diluted and quantified on microwell plates coated with bovine serum albumin-conjugated probe SK102 as previously described (9). Modifications to the standard cycling conditions were subsequently tested. These included lowering the annealing temperature to 50 °C and introducing a gradual ‘ramp-up’ RT step (50–55 °C for 5 min, 55–60 °C for 15 min, followed by 60 °C for 10 min).
Confirmation of template sequence

To ensure that the synthetic templates harbored the expected mutations, each template was purified by gel electrophoresis, cloned into a plasmid vector and the sequence was verified using a Taq DyeDeoxy™ Terminator Cycle Sequencing Kit (Applied Biosystems).

RESULTS

SK462-431 under standard RT-PCR conditions

Under the standard RT-PCR conditions, the presence of up to four mismatches in either primer had little or no effect on PCR product yield (Table 1). The presence of five mutations in SK462 reduced PCR product yield by 22-fold relative to the homologous template; six mutations by 80- to 100-fold. In SK431, the largest difference (13-fold reduction) was observed with the template that had five mutations. Mismatch tolerance was not affected by the initial copy number input; the differences observed were identical regardless of whether 10^3 or 10^6 copies of the template were amplified (data not shown). Furthermore, the differences were consistent from run to run.

Table 1. The effect of different mutations on PCR product yield

<table>
<thead>
<tr>
<th>Template</th>
<th>No. of mutations</th>
<th>Difference relative to homologous template</th>
</tr>
</thead>
<tbody>
<tr>
<td>SK462</td>
<td>SK431</td>
<td>Standard conditions</td>
</tr>
<tr>
<td>M0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>M1</td>
<td>6</td>
<td>108</td>
</tr>
<tr>
<td>M2</td>
<td>6</td>
<td>79</td>
</tr>
<tr>
<td>M3</td>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>M4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>M5</td>
<td>3</td>
<td>1.5</td>
</tr>
<tr>
<td>M6</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>M7</td>
<td>2</td>
<td>0.9</td>
</tr>
<tr>
<td>M8</td>
<td>1</td>
<td>1.9</td>
</tr>
</tbody>
</table>

The templates were amplified under standard and modified conditions as described in Materials and Methods. The results are expressed as fold differences relative to the homologous template (M0).

Effect of reducing the reverse transcription and annealing temperatures on mismatch tolerance

Since the presence of a large number of mutations in either primer affects amplification, two different approaches were taken to alleviate the reduced efficiency. First, the annealing temperature during PCR was reduced from 55 to 50°C to decrease the hybridization stringency. Second, to enhance hybridization, the temperature of the reverse transcription reaction was slowly raised from 50 to 60°C, as described in Materials and Methods. Lowering the annealing temperature to 50°C for the entire 28 cycles allowed for greater tolerance of the mismatches, particularly in those templates that contained mutations in the SK462 binding region. However, the lower annealing temperature by itself was insufficient, as a 9-fold (versus 13-fold) reduction in yield was observed with the five mutations to SK431 (data not shown). When both a 'ramp-up' reverse transcription and a 50°C annealing temperature were applied to all templates, the effect of mutations was substantially reduced in all of the templates (Table 1). The largest difference was a 10-fold reduction in PCR product yield with the templates that contained six mutations in the SK462 region, while a 4-fold reduction was seen with the template that contained five mutations in the SK431 binding region.

Alternate primers

Given the adverse effects of multiple mismatches, we further investigated whether modified bases would allow greater mismatch tolerance. Base analogs such as 5-methylcytosine (5-MeC) have been shown to influence not only the binding equilibrium but also increase the dissociation temperature (T_d) of oligonucleotides by changing the base stacking pattern that permits intimate contact between the halogen atom and the adjacent base (12,13). Modified versions of SK462 and SK431 were synthesized containing 5-MeC in place of cytosine (Fig. 3). By using a 50°C annealing temperature, mismatch tolerance was significantly improved. Less than a 5-fold reduction in PCR product yield was observed in templates with the most mismatches (data not shown). With the exception of one position in SK462, the 5-MeC substitutions were not at positions of mismatches.

Derivatives of SK462 and SK431 were synthesized in which the variable bases were replaced with deoxyinosine. Deoxyinosine preferentially pairs with dC and to a lesser extent dA (14). Two inosine-containing derivatives of SK462 were tested. In SK462 Ino 1, six of the variable positions were replaced with inosine, four of which resulted in a dI:dC pairing (Fig. 3). SK462 Ino 2 is similar to SK462 Ino 1 but has two fewer inosine substitutions at the 5’-terminus. Derivatives of SK462 and SK431 were synthesized in which the variable bases were replaced with deoxyinosine. Deoxyinosine preferentially pairs with dC and to a lesser extent dA (14). Two inosine-containing derivatives of SK462 were tested. In SK462 Ino 1, six of the variable positions were replaced with inosine, four of which resulted in a dI:dC pairing (Fig. 3). SK462 Ino 2 is similar to SK462 Ino 1 but has two fewer inosine substitutions at the 5’-terminus. Derivatives of SK462 that contained propynyl dC and dU analogs were also evaluated (Fig. 3). Propynyl analogs have been shown to enhance binding of oligonucleotides to an RNA sequence (15).

Since the inosine substitutions were made only to the SK462 primer, only templates M0–M4 were evaluated. Reverse transcription and PCR were performed under the standard conditions. The results indicate that mismatches were better accommodated by SK462 Ino 1, which had inosine substituted at six positions, than by SK462 Ino 2, with four substitutions (Table 2). A <2-fold difference in product yield was detected between the homologous and most divergent templates. Even with inosine substituted at four positions, amplification efficiency with the variant templates was markedly improved. The base pairing of dI with dC and dI with dA likely accounts for the improved mismatch tolerance by creating a more stable duplex. The propynyl derivatives did not perform as well as inosine in accommodating mismatches.

A derivative of SK462, SK145, was also evaluated. Oligonucleotide SK145 differs from SK462 at two positions (Fig. 3) and has fewer mismatches at the 5’-terminus with the majority of isolates. SK462 was designed to eliminate the six consecutive dG nucleotides in SK145 to allow more efficient oligonucleotide

![Image](339x705 to 532x771)
The extent of mismatch tolerance is dependent on numerous factors, including the number and position of mismatches, the nature of the mismatches, the length (which influences the $T_m$ value) of the oligonucleotides and the stringency of the amplification conditions used. Mismatches between the template and the reverse transcribing primer binding region, SK431, is more conserved, with a maximum of five mismatches to a subtype F isolate. We infer from the results of this study that isolates of subtype B, which are predominant in the US and Europe, will be efficiently amplified and detected given that they harbor at most three mismatches with SK462 and two mismatches with SK431. Similarly, isolates of subtypes C and D are expected to be efficiently amplified. On the other hand, the viral RNA titers of some subtype A isolates will be underestimated due to the presence of as many as six mismatches to the SK462 region.

To improve amplification of mismatched templates, numerous changes can be implemented. In this study we examined the effect of different annealing temperatures and various modified bases in the primers on mismatch tolerance. A comparison of the results from 50 and 55°C annealing indicates that mismatches can be better accommodated when a lower annealing temperature is used. Modifications in the reverse transcription step also improved mismatch tolerance in the RT primer. The gradual temperature increase presumably facilitates the priming and extension of mismatched templates. The substitution of modified bases such as 5-mC further stabilized binding of the primer to the template.

The templates were amplified under the standard conditions used for the assay.

Quantitative assays have been extensively used in monitoring the effect of anti-retroviral drugs and for use as a prognostic indicator (21–24). The clinical efficacy of these drugs was established based upon the ability of the drugs to significantly reduce viral load as determined by quantitative RNA assays. Given the heterogeneity of the HIV genome, we embarked on this study to determine the extent to which primer–template mismatches can affect amplification efficiency. The templates used in this study were intentionally constructed to represent worst case scenarios. For example, mutant templates of SK462 represent the most divergent of HIV isolates, the A subtypes. The downstream reverse transcribing primer binding region, SK431, is more conserved, with a maximum of five mismatches to a subtype F isolate.

We infer from the results of this study that isolates of subtype B, which are predominant in the US and Europe, will be efficiently amplified and detected given that they harbor at most three mismatches with SK462 and two mismatches with SK431. Similarly, isolates of subtypes C and D are expected to be efficiently amplified. On the other hand, the viral RNA titers of some subtype A isolates will be underestimated due to the presence of as many as six mismatches to the SK462 region.

**DISCUSSION**

The effect of 3′-terminal mismatches has been extensively reported (17–20). In this study we systematically evaluated the effects of varying numbers of internal mismatches on RT-PCR. The extent of mismatch tolerance is dependent on numerous factors, including the number and position of mismatches, the nature of the mismatches, the length (which influences the $T_m$ value) of the oligonucleotides and the stringency of the amplification conditions used. Mismatches between the template and the reverse transcribing (downstream) primer should affect only the reverse transcription step; mismatches with the upstream primer are expected to affect only the first cycle of amplification. Once a mismatched template is copied, the primers are fully comple-

**Table 2. The effect of modified bases in the upstream primer on product yield**

<table>
<thead>
<tr>
<th>Template</th>
<th>No. of mutations</th>
<th>Difference relative to homologous template</th>
</tr>
</thead>
<tbody>
<tr>
<td>SK462</td>
<td>Ino 1</td>
<td>propynyl C</td>
</tr>
<tr>
<td>M0</td>
<td>2</td>
<td>0.9</td>
</tr>
<tr>
<td>M1</td>
<td>4</td>
<td>1.4</td>
</tr>
<tr>
<td>M2</td>
<td>4</td>
<td>2.7</td>
</tr>
<tr>
<td>M3</td>
<td>3</td>
<td>2.5</td>
</tr>
<tr>
<td>M4</td>
<td>2</td>
<td>1.3</td>
</tr>
</tbody>
</table>

**Table 3. The effect of SK145 on product yield**

<table>
<thead>
<tr>
<th>Template</th>
<th>No. of mutations to SK145</th>
<th>Difference relative to 145M0</th>
</tr>
</thead>
<tbody>
<tr>
<td>145M0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>M0</td>
<td>2</td>
<td>0.9</td>
</tr>
<tr>
<td>M1</td>
<td>4</td>
<td>1.4</td>
</tr>
<tr>
<td>M2</td>
<td>4</td>
<td>2.7</td>
</tr>
<tr>
<td>M3</td>
<td>3</td>
<td>2.5</td>
</tr>
<tr>
<td>M4</td>
<td>2</td>
<td>1.3</td>
</tr>
</tbody>
</table>

The templates were amplified under the standard conditions used for the assay.
Subtype E isolates, although distinct from subtype A over the env region, are indistinguishable from subtype A in the gag gene (3). Consequently, amplification of some subtype E isolates will also be underestimated. The number of available sequences for subtypes F–H are limited, but based on the sequences available to date, are expected to be amplified efficiently with SK462-431. It is important to note that although the absolute copy number determinations for subtypes A and E may be compromised, the reduction in amplification efficiency was consistent from experiment to experiment, suggesting that the assay can still be used to monitor viral load in an individual over time. These studies indicate that replacing SK462 with SK145 would significantly improve quantification of the more divergent isolates. Furthermore, the incorporation of modified bases into the primers would further enhance mismatch tolerance.

This study extends earlier work on the effects of primer–template mismatches on PCR amplification. We have demonstrated that mismatch tolerance can be improved by reducing the number of primer–template mismatches, reducing the annealing temperature, using a slow temperature ramp during reverse transcription and using modified bases to minimize destabilization of mismatches. The results of this experimental study should serve as guidelines in the design of primers for other systems. Our future plans include continuing these studies to clinical specimens of known HIV-1 subtypes.

ACKNOWLEDGEMENTS

We thank Tom White and Tom Myers for critical review of the manuscript, Sheng-Yung Chang for technical advice, Fred Reichert and Agnes Cavalli for sequencing the cloned templates, Corey Levenson, Tomas Martinez, Laura Jung and Olga Budker for providing oligonucleotides and Annie Yoon for manuscript preparation.

REFERENCES