NOTES & TIPS

Elimination of Primer–Dimer Artifacts and Genomic Coamplification Using a Two-Step SYBR Green I Real-Time RT-PCR[1]

Jo Vandesompele, Anne De Paepe, and Frank Speleman[2]

Center for Medical Genetics, Ghent University Hospital, B-9000 Ghent, Belgium

Received September 5, 2001; published online

Gene expression analysis plays an increasingly important role in many fields of biological research. The recently developed real-time PCR quantification method has many advantages over the conventional quantifications in terms of accuracy, sensitivity, dynamic range, high-throughput capacity, and absence of post-PCR manipulations (1, 2). Sequence-specific fluorescence-labeled probes (e.g., TaqMan) have been considered as a standard detection format in many diagnostic and research applications (3, 4), but are not very well suited for quantification of a large number of different sequences, because a new and relatively expensive probe is generally required for each amplicon under investigation. We have therefore optimized and validated a reverse transcriptase PCR (RT-PCR) assay for accurate expression profiling using the double stranded DNA-binding dye SYBR green I, which is a much more economical alternative to quantify any given transcript in a reaction. Using such a generic dye, different PCR amplicons and/or nonspecific amplification products could accurately be distinguished by the generation of so-called DNA melting curves and first-derivative melting peaks (5).

During initial one-step RT-PCR reactions, we observed extensive accumulation of primer–dimers (PD)[3] in no-template-control (NTC) tubes. This might obscure the true result in quantitative assays using generic DNA dyes, especially in samples where the gene of interest is of low abundance and PD are readily formed. A two-step RT-PCR protocol was therefore introduced and eliminated this problem. Further validation of the real-time PCR illustrated the prerequisite and efficacy of DNase treatment of RNA samples prior to cDNA synthesis. This step resulted in a significantly facilitated primer design for RT-PCR as the position of the primers no longer has to be considered to control for genomic contamination. We finally addressed important issues of reproducibility by determination of the intra- and interassay variation of our established method. It was shown that DNase treatment, cDNA synthesis, and RT-PCR are very reproducible. In conclusion, a two-step real-time quantitative RT-PCR based on SYBR Green I detection chemistry and DNase-treated RNA samples is the method of choice for sensitive, reproducible, and large-scale measurements of gene expression levels.

In accordance with previous successful results obtained in our laboratory with one-step quantitative RT-PCR assays using a TaqMan probe as detection format, it was initially decided to apply a one-step RT-PCR for gene expression measurements using SYBR green I (i.e., cDNA synthesis using gene-specific primers and real-time PCR amplification are sequentially performed in the same tube). All primers were designed with PrimerExpress 1.0 software (Applied Biosystems) using the default TaqMan parameters, with modified minimum amplicon length requirements (75 bp). An additional requirement was a maximum GC content of 40% for the five last 3′ end nucleotides. Primer sequences can be obtained from the authors on request. To determine the minimum primer concentrations giving the lowest threshold cycle (C_t) and maximum amplification efficiency while minimizing nonspecific amplification, nine different combinations of reverse and forward primer concentrations were tested (two-by-two combinations of 100, 300, or 900 nM). For each combination, a no-template control was included. SYBR green I amplification mixtures (25 μl) contained 25 ng total RNA, 2× SYBR Green I Master Mix buffer (12.5 μl) (Applied Biosystems), 0.125 μl MuLV-RT enzyme (50 U/μl) (Applied Biosystems), 0.125 μl RNasin (40 U/μl) (Promega), and varying concentrations of for-

[1] This study was supported by a grant from the Flemish Institute for the Promotion of Scientific Technological Research in Industry (IWT), GOA Grant 12051397, FWO Grant G.0028.00, and BOF Grant 011B4300.

[2] To whom correspondence and reprint requests should be addressed at Center for Medical Genetics, Ghent University Hospital, 1K5 De Pintelaan 185, B-9000 Ghent, Belgium. Fax: +32-9-2404970. E-mail: franki.speleman@rug.ac.be.

[3] Abbreviations used: PD, primer–dimers; NTC, no-template controls.

doi:10.1006/abio.2001.5564
0003-2697/02 $35.00

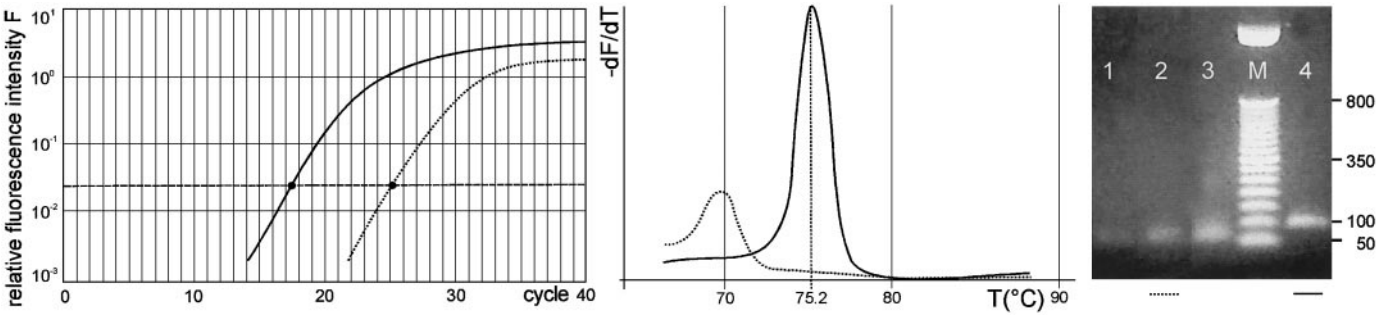

FIG. 1. Primer–dimer accumulation during one-step RT-PCR. Left, amplification plots of no-template control (dotted line, loaded in gel lane 2) and 50 ng input of total RNA (solid line, gel lane 4) using 300 nM of each primer. Middle, melting peaks of same products as in left panel ($-dF/dT$, negative first derivative of fluorescence intensity with respect to temperature T). Right, agarose gel electrophoresis; lanes 1–3, no-template controls (increasing *MME* primer concentrations, respectively, 100, 300, and 900 nM each); lane M, 50-bp DNA marker; and lane 4, 98-bp *MME* amplicon.

ward and reverse primer. Reactions were run on a ABI Prism 5700 sequence detector (Applied Biosystems). The cycling conditions comprised a cDNA synthesis step at 48°C for 30 min, 10 min of polymerase activation at 95°C, and 40 cycles at 95°C for 15 s and 60° C for 1 min. Upon analysis, amplification products were readily observed in all NTCs. DNA melting curve analysis (using the built-in feature in the software for automatic fluorescence data capturing during gradual temperature increase from 60 to 95°C) and agarose gel electrophoresis of PCR products demonstrated extensive PD accumulation in NTC tubes during one-step RT-PCR. This phenomenon was found for five genes tested and was more pronounced with increasing primer concentrations (Fig. 1). To the best of our knowledge, extensive PD formation during a quantitative one-step RT-PCR in NTCs has not been reported yet, probably because most assays use a TaqMan probe that does not detect this kind of nonspecific amplification. Although no general mechanism for the formation of PDs is described, primers appear to be tail-to-tail oriented in sequenced PDs (6).

In an attempt to reduce PD formation, we introduced a two-step protocol, i.e., cDNA synthesis and PCR amplification in separate tubes. First-strand cDNA was synthesized from 1 μg total RNA using random hexamers (AP Biotech) and SuperScriptII reverse transcriptase (RTase) according to the manufacturer (Invitrogen). After synthesis, cDNA was diluted with 60 μl RNase-free water (Sigma) to obtain a concentration 12.5 ng/μl. The one-step amplification protocol was altered by leaving out the 30-min incubation step at 48°C and the RTase, and RNA was substituted for by 25 ng of cDNA. In the two-step NTC reactions, no amplification of PD was observed, in contrast to presence of PDs in the one-step NTCs that were run in parallel as controls (data not shown). It was further shown that for all tested genes, combinations of 300 nM forward and reverse primer resulted in optimal amplification. Besides elimination of PDs, a two-step protocol with oligo(dT) or random hexamer primers for cDNA synthesis has additional advantages. It allows amplification of multiple targets from the same cDNA pool while variation in RT efficiency other than sample-to-sample variation is controlled for as the cDNA mixture is diluted and split to quantify the different genes of interest. Furthermore, the two-step approach eliminates the need to perform repeated housekeeper amplifications for normalization, which is a prerequisite in the one-step method, in the absence of any other controls.

To investigate the nature of the observed PD accumulation, the MuLV RTase and the RNase inhibitor RNasin were omitted separately from NTC reactions during one-step RT-PCR. PDs were only absent in the tubes without the RTase. Two other widely used RTases (SuperScriptII, Invitrogen; AMV, Promega) were also tested for this phenomenon. According to the manufacturer's guidelines, the temperature of the cDNA synthesis step was altered from 48 to 42°C. PD accumulation was also clearly observed for these two RTases (data not shown). These observations suggest that the reported DNA-dependent polymerase activity of RTases on DNA templates and RNA:DNA hybrids—although not significant—accounts for the initial formation of PDs that are subsequently and efficiently amplified by the DNA polymerase during PCR.

Another critical aspect of quantitative RT-PCR with respect to accurate results is the presence of trace amounts of genomic DNA within the RNA samples to be tested. Virtually no extraction method can guarantee the isolation of DNA-free RNA. PCR primers spanning an exon–exon boundary are often designed to control for genomic contamination. However, this strategy can no longer be used if processed pseudogenes are present in the genome. Furthermore, intron–exon boundary information is not available (yet) for the majority of the genes, nor is it often known in advance

if putative pseudogenes of the gene under investigation reside in the genome. Initial interpretation of the sequence data from human chromosomes 21 and 22 indicated that about 20% of the coding sequences turned out to be pseudogenes (7, 8). Therefore, RNase-free DNase treatment of RNA samples is a prerequisite for accurate RT-PCR results. To illustrate the requirement and efficacy of a DNase treatment of the RNA samples prior to cDNA synthesis, real-time SYBR green I PCR reactions were tested for 25 genes in reaction tubes with RNA (−RT controls) or cDNA (+RT controls) as template (both before and after DNase treatment of RNA) and no-template controls. RNA samples from various tissues (RNeasy extraction, Qiagen) were treated with the RQ1 RNase-free DNase according to the manufacturer (Promega). Treated RNA samples were desalted prior to cDNA synthesis using Microcon-100 spin columns (Millipore). DNase treatment of RNA resulted in complete elimination of amplification in tubes with treated RNA as template, while preserving specific amplification in the cDNA controls (Fig. 2). Together with improved accuracy for RT-PCR, primer design is facilitated (as exon–exon spanning primers are no longer required) and the choice of a housekeeping gene for normalization is no longer restricted to those genes without any known pseudogenes. So far, we have analyzed 10 different housekeeping genes belonging to different functional and abundance classes in various tissues and cells. We clearly observed that no gene is really constantly expressed and multiple housekeeping genes are required for an accurate normalization (Vandesompele *et al.*, submitted for publication).

One-step RT-PCR reactions are often preferred over two-step reactions, because of a presumed minimization of experimental variation, as all enzymatic steps occur in the same tube under controlled thermocycler

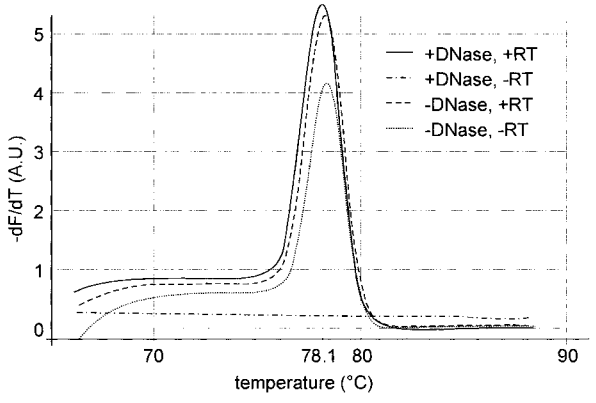

FIG. 2. Efficacy of removal of contaminating DNA in RNA preparations as evidenced by melting peak analysis of generated PCR products, here shown for a representative 91-bp *RPS25* amplicon (+DNase, RNA was DNase treated; −DNase, no DNase treatment; +RT, RT was added for cDNA synthesis; −RT, RT was omitted).

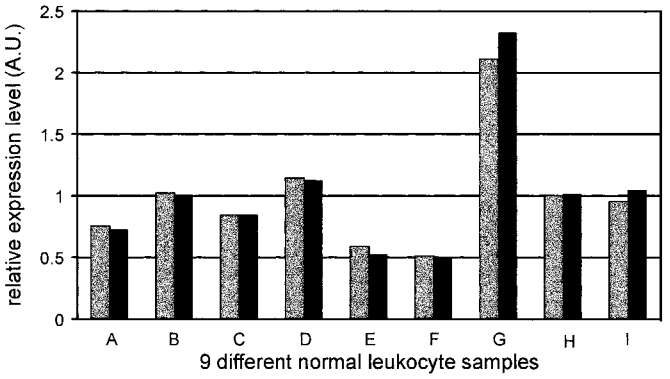

FIG. 3. Relative *HMBS* gene expression levels in nine normal leukocyte samples of which RNA was subjected to two parallel rounds (black and gray) of DNase treatment and cDNA synthesis.

conditions. To address the issue of reproducibility, we determined the interassay variation of our established two-step RT-PCR protocol. We have therefore subjected nine different RNA samples to two parallel rounds of DNase treatment and cDNA synthesis and quantified the relative expression levels of each of four housekeeping genes (*HMBS*, *UBC*, *ACTB*, and *HPRT1*) in all samples using the comparative C_t method. The two-step protocol is highly reproducible with Pearson correlation coefficients ranging from 0.974 to 0.988 between the expression levels of the two parallel series of cDNA samples for the four tested genes, of which one is shown in Fig. 3. The interassay variation was calculated as the median coefficient of variation (standard deviation divided by the arithmetic mean) for the paired relative quantities obtained in the two parallel series for the four genes and amount to 7.7%. This figure is only slightly higher than the observed median intraassay variation of 5.8% (reflecting the variation between duplicated reactions in the same PCR run; based on 25 different genes tested on 30 samples, data not shown) and demonstrates the reproducibility of the two-step approach.

To summarize, SYBR green I is the detection format of choice for accurate and reproducible real-time transcript abundance measurements of a large series of genes. Obscuring PD observed in one-step RT-PCR reactions were eliminated by the use of a two-step and DNase treatment of RNA prior to cDNA synthesis was shown to be a prerequisite for accurate RT-PCR and facilitated primer design.

Acknowledgments. We thank Johanna Iso-Oja and Filip Pattyn for their excellent technical assistance.

REFERENCES

1. Higuchi, R., Fockler, C., Dollinger, G., and Watson, R. (1993) Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology (N. Y.)* **11,** 1026–1030.

2. Heid, C. A., Stevens, J., Livak, K. J., and Williams, P. M. (1996) Real time quantitative PCR. *Genome Res* **6,** 986–994.
3. Bustin, S. A. (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J. Mol. Endocrinol.* **25,** 169–193.
4. Lie, Y. S., and Petropoulos, C. J. (1998) Advances in quantitative PCR technology: 5′ nuclease assays. *Curr. Opin. Biotechnol.* **9,** 43–48.
5. Ririe, K. M., Rasmussen, R. P., and Wittwer, C. T. (1997) Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. *Anal. Biochem.* **245,** 154–160.
6. Brownie, J., Shawcross, S., Theaker, J., Whitcombe, D., Ferrie, R., Newton, C., and Little, S. (1997) The elimination of primer–dimer accumulation in PCR. *Nucleic Acids Res.* **25,** 3235–3241.
7. Dunham, I., Shimizu, N., Roe, B. A., Chissoe, S., Hunt, A. R., Collins, J. E., Bruskiewich, R., Beare, D. M., Clamp, M., Smink, L. J., Ainscough, R., Almeida, J. P., Babbage, A., Bagguley, C., Bailey, J., Barlow, K., Bates, K. N., Beasley, O., Bird, C. P., Blakey, S., Bridgeman, A. M., Buck, D., Burgess, J., Burrill, W. D., O'Brien, K. P., *et al.* (1999) The DNA sequence of human chromosome 22. *Nature* **402,** 489–495.
8. Hattori, M., Fujiyama, A., Taylor, T. D., Watanabe, H., Yada, T., Park, H. S., Toyoda, A., Ishii, K., Totoki, Y., Choi, D. K., Soeda, E., Ohki, M., Takagi, T., Sakaki, Y., Taudien, S., Blechschmidt, K., Polley, A., Menzel, U., Delabar, J., Kumpf, K., Lehmann, R., Patterson, D., Reichwald, K., Rump, A., Schillhabel, M., and Schudy, A. (2000) The DNA sequence of human chromosome 21: The chromosome 21 mapping and sequencing consortium. *Nature* **405,** 311–319.