Real-time Analysis of RAM Amplification
Using RAM-Specific Strategies, and Methodologies Developed for PCR

Thomas P. Beals Thorne Diagnostics, Beverly, Massachusetts, USA www.ThorneDiagnostics.com

Abstract

PCR amplification efficiency and RAM amplification rate are analogous measures that must be accounted for to make accurate sample-to-sample comparisons in real-time amplification systems. PCR and RAM real-time data can be analyzed with mathematical models that provide estimates of PCR efficiency or RAM reaction rate. RAM reactions can be sampled at high density, yielding sufficient data for direct fitting of a simple exponential model.

RAM, like PCR, is an amplification method.

RAM amplification (Zhang et al., 1998) and RAM product detection are the final steps in a nucleic acid (DNA or RNA) detection process. Table 1 compares features of PCR and RAM. Figure 1 illustrates ligation of two target-specific segments of a padlock (or circularizable "C") probe on a target molecule to form a single-stranded DNA (ssDNA) circle. RAM amplification of the ligated padlock probe can be visualized as rolling-circle replication that is accompanied by two-primer mediated exponential amplification. The RAM amplification product is a ladder of linear double-stranded DNA (dsDNA).

Table 1: PCR, RAM comparison

	PCR	RAM
template	**ssDNA**	**circular ssDNA**
reaction conditions	**temperature cycling**	**isothermal**
reaction mechanism	**denature, anneal, extend**	**rolling circle replication, two-primer amplification**
data collection	**one datum per cycle**	**continuous sampling (instrumentation-limited)**
report units	**Cycle threshold (Ct)**	**Response time (Rt)**
efficiency / rate	**per-cycle doubling is efficiency maximum**	**mass per time is reaction rate**

Table 1: RAM - PCR comparison

- Template: PCR amplifies a single-stranded DNA template; the RAM-amplification template is a single-stranded circle formed by ligation of a padlock probe.
- Reaction conditions: Temperature-cycling regenerates PCR's ssDNA template; the RAM reaction is performed at a single temperature, i.e., RAM is an isothermal reaction.
- Reaction mechanism: The anneal-extend-melt PCR process is well-known. The RAM reaction builds on a long single-stranded rolling-circle amplification template with two primers in a continuous process of primer extension and strand displacement. Borrowing mathematical terminology, the RAM reaction is continuous; PCR is discrete.
- Signal collection: Ideally PCR copies all available template by the end of each extension cycle; a PCR per-cycle datum estimates the mass of dsDNA. The utility of continuous monitoring of RAM reactions is explored in more detail below.
- Report units: The time at which the RAM reaction fluorescent signal reaches a designated fluoresence level is called the response time (Rt; Burg et. al., 1995), analogous to the cycle-threshold (Ct) of PCR.
- Efficiency / rate: Ideal efficiency PCR doubles template mass at each cycle. The analogous RAM function is Rt vs. log(template number); with units of reaction-product per time, it is a measure of the reaction rate.

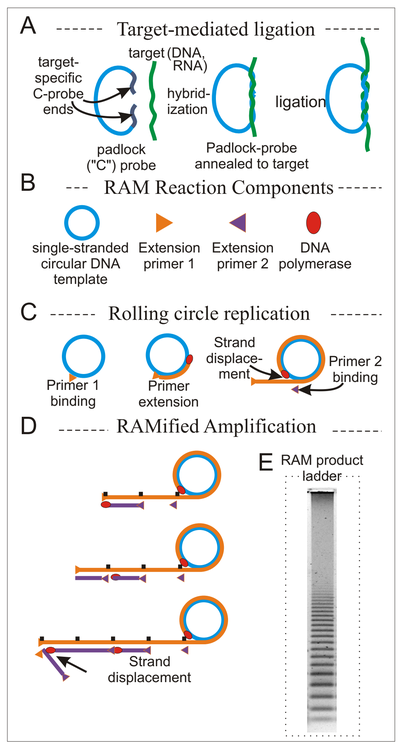

Figure 1: Target detection via RAM amplification of ligated padlock probe.

A) A padlock "C" probe detects a target nucleic acid *via* hybridization of split complementary sequences comprising the ends of the linear probe. The padlock probe ends are adjacent after target hybridization. The C-probe is an open circle after ligation.

B) Essential components of a RAM reaction.

C) Rolling circle amplification (RCA) initiates a RAM reaction.

D) An RCA template is copied into a "ladder" of double-stranded DNA products *via* primer extension and strand displacement.

E) A RAM ladder visualized in an agarose gel.

Parameterized curve fitting in PCR and RAM

The motivations for fitting parameterized curves are both theoretical and practical. Reconciliation of the simplicity of the mathematical formulation of PCR with the complexity of the reaction biochemistry has inspired a substantial literature; e.g. Pfaffl, 2001; Liu and Saint, 2002; Rutledge and Cote, 2003; Fleige et al., 2006; Rutledge and Stewart, 2008.

Parameterized models generate, from mathematical equations, curves that resemble real-time amplification signals (Figure 2A). Given a set of (time, signal) data pairs, curve-fitting software can often find a set of parameters to generate a curve that, to the eye, is an almost perfect fit to real-time signals.

The practical use of parameterized models is in determining reaction efficiency; algebraic manipulation of parameters yields a second-derivative maximum that can be used as an efficiency-defining point (Spiess et al., 2008).

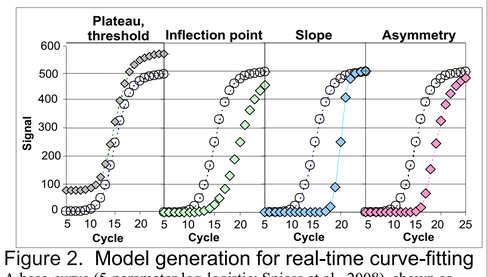

Figure 2. Model generation for real-time curve-fitting

A base-curve (5-parameter log-logistic; Spiess et al., 2008), shown as open circles in the four panels, was manipulated by changing parameters that affect five features of the curve.

Panel 1: baseline and plateau were separately adjusted to raise the curve shown as grey squares above the baseline and plateau of the base-curve.

Panel 2, green squares represent a curve with an inflection point lower than the base-curve.

Panel 3, blue squares represent a curve with a steeper slope than the base-curve.

Panel 4, red squares represent a curve with that is less symmetric than the base curve.

Parameter-fitting in PCR

To compare and contrast PCR and RAM amplification and analysis, a single target sequence was amplified by both PCR and RAM. Results from the more familiar PCR are followed by RAM results and analysis.

The amplification target, from a ribosomal-DNA (rDNA) encoding region of the "black mold" fungus *Stachybotrys chartarum*, is shown (Figure 3A) amplified via real-time PCR from a dilution series of genomic DNA. Open circles mark PCR data points; continuous lines were generated by 5-parameter logistic curve-fitting using the qpcR package in the R statistical computing environment (Ritz and Spiess, 2008; Spiess et al., 2008; R Development Core Team, 2008).

If the signal threshold is within the amplification's exponential phase (where amplified target mass is assumed to be proportional to initial amplification target) PCR efficiency for a dilution series can be determined by plotting log(dilution, or target number) vs. Ct (Figure 2-B). Since dsDNA mass doubles per cycle at ideal PCR efficiency, Ct plotted vs. $\log_2$ (dsDNA mass) has slope -1.

A substantial literature (Pfaffl, 2001; Peirson et al., 2003) and reference collection at (http://www.gene-quantification.de) documents the importance of evaluating amplification efficiency for accurate real-time PCR analysis. Figure 3C illustrates how differences in amplification efficiency would give different Ct estimates for PCR of the same initial target. Clearly, it would be desirable to estimate amplification efficiency using fewer reactions than the set shown in figure 3A. Both exponential-region determination and amplification efficiency can be determined by parametric curve fitting using data from a single amplification. Figure 3D shows three amplifications from well-separated dilutions chosen from Figure 3A. The maximum amplification region, marked on each fitted line, encompasses at most two actual PCR data points - too few points to reliably estimate an exponential model.

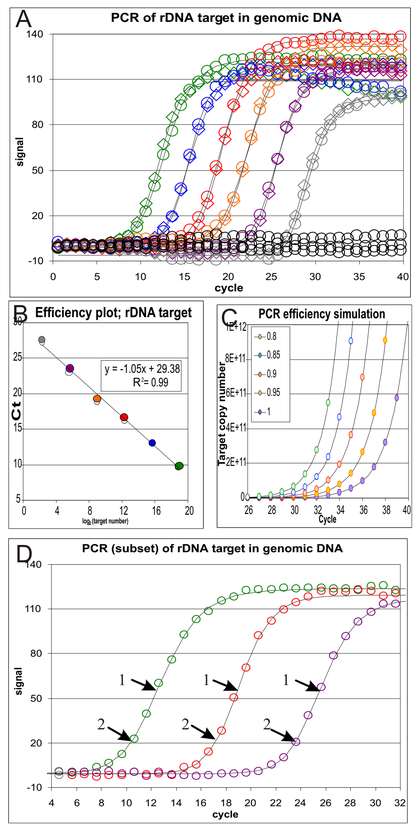

Figure 3. Parametric curves fitted to PCR data

Figure 3A. A ribosomal RNA-encoding DNA (rDNA) sequence was amplified from a dilution series of *Stachybotrys chartarum* genomic DNA. The black baseline traces are no-target PCR controls.

Figure 3B. The Ct of each amplification in Fig. 2A is plotted vs. the target number for each dilution. A target-number estimate was made after PCR to Poisson failure in the dilution series (not shown). Analysis of Ct vs. $\log_2$ (number of targets) shows slope -1.05.

Figure 3C. The effect of PCR efficiency on Ct estimates is illustrated by simulation of PCR amplification of a fixed initial number of targets. PCR efficiency as used here can range from 0 (zero) to 1; no amplification occurs at 0 efficiency; at theoretical maximum efficiency 1, PCR target copy number doubles each cycle.

Figure 3D. For visual clarity, three well-separated PCR amplifications from Figure 2A are shown. As above, open circles represent PCR data points; continuous lines are generated from parametric curve-fitting. Algebraic manipulation of parameters (Zhao and Fernald, 2005) yields fractional cycles that mark the start and end of the model curve's crossing points (arrows; 2, second derivative maximum; 1, first derivative maximum).

RAM reactions analyzed using PCR algorithms

Figure 4A shows RAM amplification of a dilution series of target circles that were ligated on the target sequence that was shown PCR-amplified in Figure 3. The RAM data points were collected using instrument-control settings that mimic PCR amplification. Data points are shown with a line fitted using the same analysis package and curve-model as the PCR reactions in Figure 3A.

Figure 4B is the RAM analog of the PCR efficiency plot in Figure 3B; response-time is plotted vs. log(target number). Like PCR, RAM generates exponential dsDNA accumulation so the response time is proportional to log(targets). The slope shows the time-rate of change of copy number, that is, the rate of the RAM reaction. There is not, for the RAM reaction, an immediate and compelling expectation of an upper bound to the reaction rate as there is for PCR. The RAM reaction rate can be varied over a wide range by manipulation of reaction conditions (unpublished data).

Figure 4C shows amplifications from well-separated dilutions chosen from Figure 4A, with RAM data points shown as open circles and a superimposed curve-fitted continuous line. The maximum amplification efficiency region is marked on each fitted line. As for the PCR example in Figure 3D, the maximum amplification efficiency region encompasses only a few data points.

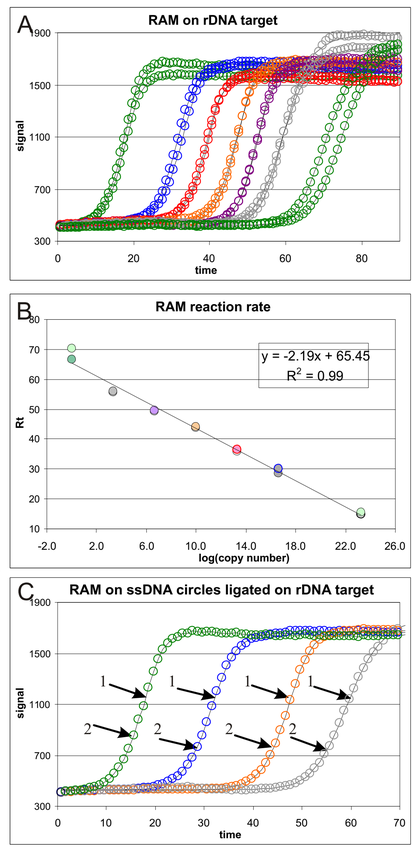

Figure 4. Parametric curves fitted to RAM data

4A. Real-time RAM reactions on a dilution series of template circles that were ligated on a *Stachybotrys chartarum* rDNA sequence.

4B: The RAM analog of a PCR efficiency plot. The Rt of each amplification in Fig. 4A is plotted vs. the target number for each dilution.

4C. Well-separated single RAM reactions; arrows mark the start and end of the model curve's maximum amplification phase derived from logistic fit parameters; 2, second derivative crossing point; 1, first derivative crossing point.

High-density sampling of RAM reactions with direct detection of exponential phase.

The real-time RAM reactions shown in Figures 4A and 4C used data that was processed as PCR data is processed by the BioRad iCycler's software: each cycle's set of data was summarized after digital filtering to give a point-estimate for that cycle.

(Why use a cycle-type program at all for RAM? A cycle-type program allows the instrument to change its per-cycle exposure-setting so that the CCD sensor does not become saturated as per-well fluorescence signal increases.)

For a real-time RAM reaction all fluorescence data can be used without filtering or other manipulation by instrumentation software. Figure 5A shows real-time RAM fluorescence data for three reactions, at well-separated response times, sampled at a maximum interval of 3.5 seconds. (The sampling interval decreases to a minimum of 1.2 seconds as the exposure time decreases.) The colored signal traces in Figure 5A are the individual data points; the data-points are sufficiently dense for the graphic symbols to appear as a line. In Figure 5B, the same data is shown with points from a 5-parameter logistic fit drawn as open circles over the data (reversing the convention used previously, where the parametric fit was shown as a continuous line).

Using a direct exponential fit to densely sampled data offers an alternative method for defining a response time. The baseline signal in real-time data is a combination of machine noise and, if used, a normalization fluor. Exponential amplification becomes visible when enough amplification products have accumulated for product-fluorescence to be significantly greater than background fluorescence. (Current technology allows only a brief period of visible exponential amplification; like icebergs, most of the amplification is below the surface.) When comparing multiple amplification reactions (e.g. all wells in a multiwell plate) it is a common practice to scale all reactions to a common baseline (e.g. by a "baseline subtraction" option in data-processing software.)

Figure 5C shows the data region that was used to generate the exponential model by overlaid graphic boxes. As in Figure 5B, arrows mark the second and first derivative maxima derived from the logistic fit parameters.

Figures 5D and 5E illustrate direct exponential fit and an alternative definition of response time. High-sampling-frequency data sets from RAM amplifications were scanned with a sliding-window to find an initial region of exponential amplification. On finding an initial exponential region, the scan-window was widened to include the greatest number of data points that fit the exponential model at a stringent quality threshold. Typically, greater than 80 data points were used, with a correlation coefficient $R^2 > 0.99$. As above, the apparently continuous line was created by showing all data points. Open circles were generated by the exponential fit equation and are shown extrapolated beyond the exponential fit data points; those points are marked by arrows in Figures 5D and 5E. Similarly, the baseline portion of the data was used to generate a linear baseline equation, and the extrapolation of the baseline equation is shown as small filled circles in Figures 5D and 5E.

The time defined by the intersection of the extrapolated baseline and extrapolated exponential curves is proposed as a data-defined response-time. Examples of those curve-intersections are shown as inset graphs in Figures 5D and 5E. Use of the baseline-exponential intersection as response-time scales to baselines without baseline subtraction and eliminates threshold choice as a variable in comparison of amplifications.

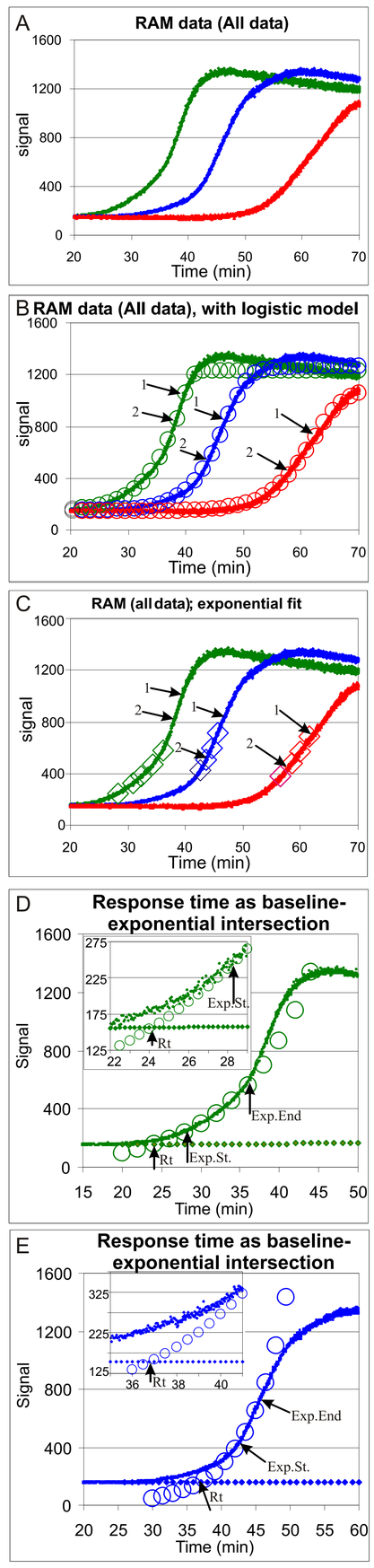

Figure 5: Direct exponential fit to real-time RAM reactions.

5A. RAM reaction fluorescence data from three amplification reactions. The data collection frequency yields sufficient data-density for the individual data points to appear as a continuous line at this graphic scale.

5B: Solid lines are the RAM data from Figure 5A. Open circles are from a 5-parameter log-logistic curve fit to the amplification. Arrows mark the start and end of the model curve's maximum amplification derived from logistic fit parameters.

5C. Graphic symbols around the data-line mark the data-region that was used to define the exponential model. Arrows mark second and first derivative crossing points as above.

5D. One set of amplification data from figure 5A is shown. An exponential growth curve that was directly fitted to a segment of the RAM data is shown as open circles; calculated points of the exponential curve are extrapolated beyond the points of best fit. The start and end of the segment of the exponential best fit to the RAM data that was shown as graphic boxes in Fig. 5C is marked by arrows (Exp.St. start of exponential fit segment; Exp.End, end of exponential fit). Points generated by a linear equation that was fit to the baseline are shown as filled circles.

The response time is marked by an arrow (Rt) indicating the intersection of the extrapolated baseline and extrapolated exponential growth curve. The inset box shows a magnified view of the intersection region.

5E. A second curve from figure 5A; depicted as described above for figure 5D.

Conclusions

Real-time RAM amplification analysis can be done with the curve-fitting tools developed for PCR.

The isothermal RAM reaction can be monitored continuously to generate effectively continuous monitoring of the reaction.

Exponential growth models can be directly fit to densely sampled RAM reactions.

A response-time can be unambiguously defined by the intersection of linear baseline and exponential amplification equations.

Discussion

RAM reaction rate, like efficiency in PCR, should be a factor in calculations for comparison of different reactions. Use of exponential curve-fit parameters for comparison of samples with different reaction rates will be demonstrated in future publications.

Detection of ligated padlock probes by a variety of techniques has been applied to applications as diverse as microarray protein profiling (Kingsmore and Patel, 2003) and SNP detection (Nilsson et al., 1997; Hardenbol et al., 2003; Krishnakumar et al., 2008). Real-time analysis of the RAM reaction may offer unique opportunities for padlock probe detection in a variety of applications.

Aknowledgements

I thank Dr. David Lane for suggesting the use of the baseline-exponential intersection point as a response time, and Mr. Raymond Nietupski, Dr. James Smith, and Dr. David Lane for reagents and helpful discussions.

Literature cited

Burg, J. L., Cahill, P. B., Kutter, M., Stefano, J. E., and Mahan, D. E. (1995). Real-time fluorescence detection of RNA amplified by Q beta replicase. Anal Biochem *230*, 263-272.

Fleige, S., Walf, V., Huch, S., Prgomet, C., Sehm, J., and Pfaffl, M. W. (2006). Comparison of relative mRNA quantification models and the impact of RNA integrity in quantitative real-time RT-PCR. Biotechnol Lett *28*, 1601-1613.

Hardenbol, P., Baner, J., Jain, M., Nilsson, M., Namsaraev, E. A., Karlin-Neumann, G. A., Fakhrai-Rad, H., Ronaghi, M., Willis, T. D., Landegren, U., and Davis, R. W. (2003). Multiplexed genotyping with sequence-tagged molecular inversion probes. Nat Biotechnol *21*, 673-678.

Kingsmore, S. F., and Patel, D. D. (2003). Multiplexed protein profiling on antibody-based microarrays by rolling circle amplification. Curr Opin Biotechnol *14*, 74-81.

Krishnakumar, S., Zheng, J., Wilhelmy, J., Faham, M., Mindrinos, M., and Davis, R. (2008). A comprehensive assay for targeted multiplex amplification of human DNA sequences. Proc Natl Acad Sci U S A *105*, 9296-9301.

Liu, W., and Saint, D. A. (2002). A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics. Anal Biochem *302*, 52-59.

Nilsson, M., Krejci, K., Koch, J., Kwiatkowski, M., Gustavsson, P., and Landegren, U. (1997). Padlock probes reveal single-nucleotide differences, parent of origin and in situ distribution of centromeric sequences in human chromosomes 13 and 21. Nat Genet *16*, 252-255.

Nilsson, M., Malmgren, H., Samiotaki, M., Kwiatkowski, M., Chowdhary, B. P., and Landegren, U. (1994). Padlock probes: circularizing oligonucleotides for localized DNA detection. Science *265*, 2085-2088.

Peirson, S. N., Butler, J. N., and Foster, R. G. (2003). Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. Nucleic Acids Res *31*, e73.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res *29*, e45.

Ritz, C., and Spiess, A. N. (2008). qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. Bioinformatics *24*, 1549-1551.

Rutledge, R. G., and Cote, C. (2003). Mathematics of quantitative kinetic PCR and the application of standard curves. Nucleic Acids Res *31*, e93.

Rutledge, R. G., and Stewart, D. (2008). A kinetic-based sigmoidal model for the polymerase chain reaction and its application to high-capacity absolute quantitative real-time PCR. BMC Biotechnol *8*, 47.

Spiess, A. N., Feig, C., and Ritz, C. (2008). Highly accurate sigmoidal fitting of real-time PCR data by introducing a parameter for asymmetry. BMC Bioinformatics *9*, 221.

R Development Core Team (2008). R: A language and environment for statistical computing. (Vienna, Austria, R Foundation for Statistical Computing; http://www.R-project.org).

Zhang, D. Y., Brandwein, M., Hsuih, T. C., and Li, H. (1998). Amplification of target-specific, ligation-dependent circular probe. Gene *211*, 277-285.

Zhao, S., and Fernald, R. D. (2005). Comprehensive algorithm for quantitative real-time polymerase chain reaction. J Comput Biol *12*, 1047-1064.