Journal of Microbiological Methods 74 (2008) 82–88

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

# Concentration dependency of nonequilibrium thermal dissociation curves in complex target samples

Alex E. Pozhitkov [b], Rebecca A. Rule [a], Robert D. Stedtfeld [c], Syed A. Hashsham [c], Peter A. Noble [a,*]

[a] 201 More Hall, Civil and Environmental Engineering, University of Washington, Seattle, WA 98195, USA
[b] Gulf Coast Research Laboratory, University of Southern Mississippi, 703 East Beach Drive, Ocean Springs, MS 39564, USA
[c] Center for Microbial Ecology, Department of Civil and Environmental Engineering, Michigan State University, East Lansing, MI, USA

ARTICLE INFO

*Article history:*
Received 21 February 2008
Received in revised form 26 March 2008
Accepted 28 March 2008
Available online 4 April 2008

*Keywords:*
Dissociation
Nonequilibrium
Nonspecific hybridizations
Oligonucleotide arrays

ABSTRACT

The nonequilibrium thermal dissociation (NTD) methodology has been proposed to provide a superior discrimination between specific and nonspecific hybridizations than the commonly used array techniques involving hybridization followed by a single stringent wash. Multiple studies have used this method on gel-pad, planar, and nylon membrane arrays to identify specific microbial targets in complex target mixtures. A recent physicochemical study revealed several problems, particularly when the method was used to examine complex target samples. In the present study, we investigated the effect of target concentration on NTD of complex target samples obtained from an anaerobic bioreactor. Our purpose was to experimentally demonstrate that variation in the concentrations of both specific and nonspecific targets determines the course of dissociation, which was not evaluated in initial microbiological studies. We also present an approach for analyzing the dissociation curves that is less error prone compared to those used in the previous studies. Our results show that: (i) a specific target in a mixture, at a certain concentration, may have a higher dissociation temperature/time than that of the same pure target, and (ii) the concentration dependence of the dissociation precludes usage of reference curves for identifying a target. Contrary to the previous studies, an explicit calibration is required, which makes the NTD approach impractical for high throughput analysis.


## 1. Introduction

For microbial identification, nonequilibrium thermal dissociation (NTD) has been suggested to offer better discrimination between target (specific) and nontarget (nonspecific) nucleic acids than measuring signal at an appropriate wash stringency (DeLosReyes et al., 1997, 1998; ElFantroussi et al., 2003; Eyers et al., 2006; Hansen et al., 1999; Kelly et al., 2005; Koizumi et al., 2002; Li et al., 2004; Liu et al., 2001; Loy et al., 2002; McMahon et al., 1998; Mobarry et al., 1996; Siripong et al., 2006; Urakawa et al., 2002, 2003; Zheng et al., 1996). The NTD approach was first developed for membrane arrays (DeLosReyes et al., 1997; DeLosReyes et al., 1998; Hansen et al., 1999; Koizumi et al., 2002; McMahon et al., 1998; Mobarry et al., 1996; Raskin et al., 1994a,b; Zheng et al., 1996), and later adapted for gel-pad (ElFantroussi et al., 2003; Eyers et al., 2006; Kelly et al., 2005; Liu et al., 2001; Siripong et al., 2006; Urakawa et al., 2002, 2003) and planar (Li et al., 2004) arrays. The rationale for this approach is that, while probe signal intensities may vary, dissociation behavior is supposedly dependent upon whether or not the binding resembles that of a perfectly matching duplex. The NTD approach is performed on an array by increasing the temperature of a buffer solution and recording the signal, which is interpreted in the form of a dissociation curve (Liu et al., 2001; Pozhitkov et al., 2005). A conceptual example of the NTD is presented in Fig. 1. The approach relies on two assumptions: nonspecific duplexes dissociate faster than specific ones, and a dissociation curve is unique to a probe–target duplex (*i.e.*, a target in a mixture can be identified by curve matching).

At face value, the NTD approach seems quite appealing, since it is promising to alleviate the difficulties associated with the interpretation of microarray signals obtained at one stringent wash condition. Specifically, these problems arise from the fact that probes naturally have different binding energies (Pozhitkov et al., 2006), and targets occur at different concentrations in a complex target sample. In a mixed target sample, for example, it is not possible to determine *a priori* if the signal intensity of a probe is due to differences in the binding energies of hybridized targets or to differences in their concentrations. On the other hand, comparison of NTD curves, if proven valid, would provide an attractive alternative for ensuring the specificity of a probe to a target. For example, comparison of a curve of a reference pure target to that of a sample could reveal whether or not the reference target was present in an environmental sample, which has been alluded to in previous studies (*e.g.*, ElFantroussi et al., 2003; Raskin et al., 1994a,b). At that time, little was known about how a target, or a mixture of targets, dissociates from probes in solution, or from probes immobilized on the surface of a

* Corresponding author. Tel.: +1 206 685 7583; fax: +1 206 685 3836.
*E-mail addresses:* alexander.pozhitkov@usm.edu (A.E. Pozhitkov), panoble@washington.edu (P.A. Noble).

0167-7012/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.mimet.2008.03.010

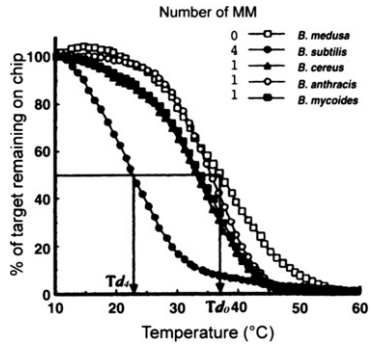


**Fig. 1.** Normalized signal intensity of targets to a probe specific for *Bacillus medusa* based on gel-pad arrays. Shown is the number of mismatches and symbol. Adapted from Liu et al. (2001).

microarray. Only recently has extensive practical and theoretical studies of the hybridization and dissociation of duplexes on a microarray been conducted, *i.e.*, hybridization thermodynamics (Held et al., 2006; Mei et al., 2003; Naef and Magnasco, 2003; Pozhitkov et al., 2006; Wu and Irizarry, 2004; Zhang et al., 2003).

To more thoroughly investigate the NTD approach, we previously conducted an extensive experimental study using a large dataset of curves recorded on gel-pad microarrays (Pozhitkov et al., 2005). To our surprise, we discovered major problems associated with how the signal intensities from an array were processed by an image acquisition and processing system, as well as multiple problems associated with complicating physicochemical factors, such as the diffusivity of the target in the gel pad and the surrounding solution. These multiple overlapping problems have a substantial impact on the interpretation of signal in gel-pad studies. Beyond the technological problems, we then asked ourselves if the NTD approach had ever been theoretically and experimentally validated. Using a different (and less complicated) platform and signal processing algorithm (avoiding all the problems previously identified; Pozhitkov et al., 2005; Pozhitkov and Noble, 2007), we began unraveling nonequilibrium dissociation curves of perfectly matching duplexes from a physicochemical perspective (Pozhitkov et al., 2007). We found that, contrary to popular belief, nonspecific duplexes do not always dissociate faster than specific ones. Furthermore, interpretation of NTD curves was found to be complicated by the fact that the intensity of an array spot is a composite of all specific and nonspecific targets bound to the same probe (Zhang et al., 2005).

This study explicitly addresses the issue of identification of targets in complex target mixtures, which should be of concern to microbiologists that use, or have used, the NTD approach to identify microbes. Our study is unique compared to all previous NTD studies because: (i) it provides a method for comparing NTD curves that is a significant improvement from previous studies (*i.e.*, it is an objective measurement based on physical theory), (ii) we show that when a specific target is present at 1%, 5% or 10% of the total amount of DNA, the dissociation curves tend to shift to lower temperatures as the concentration of the specific target decreases, (iii) dissociation of the pure target is situated between 5% and 10% mixtures, and (iv), comparison of the 100% specific target to 0% (*i.e.*, mix of all targets but the specific one) revealed that a greater number of nonspecific curves actually dissociate at higher temperatures than the specific ones.

## 2. Materials and methods

### 2.1. Microarray and raw data

The 16S ribosomal RNA gene from *Burkholderia xenovorans* strain LB400 (accession number U86373) was used to design 220 perfect-match probes. The probes were synthesized *in situ* (in quadruplicate) on Xeotron (Invitrogen) microfluidic arrays. Hybridization protocols, solutions, and the way the dissociation curves were recorded have been previously described (Pozhitkov et al., 2007; Wick et al., 2006). After scanning, the microarray was washed for 2.2 min at 22 °C and scanned again. Washing and scanning cycles were repeated up to 70 °C. The data generated can be downloaded at http://staff.washington.edu/pozhit/default.htm. To minimize the effects of background noise, probes that had initial signal intensities of <200 a.u. before dissociation were excluded from all analyses. In addition, in some experiments at one temperature, there was a sudden drop of signal intensity followed by a continuation of the normal dissociation course. This drop was an apparent outlier and the corresponding temperature was excluded from the analysis for every probe on the microarray.

### 2.2. Targets

A fragment of the 16S rRNA genes (1466 bp) was amplified from a pure culture of the *B. xenovorans* strain LB400. A mixture of unknown microbial targets (that did not contain *B. xenovorans*) was obtained by amplifying rRNA genes from an anaerobic bioreactor. Amplicons tested on all batches were generated from the same pools of gDNA. Sample mixtures included 100, 10, 5, 1 and 0% pure culture in anaerobic bioreactor, and were mixed based on the mass of amplicons (*e.g.*, 2.5 ng of pure culture amplicons and 247.5 ng of bioreactor amplicons for 1% mixture). The targets were labeled with Cy3 as previously described (Wick et al., 2006).

### 2.3. Comparison of NTD curves

Concentration of a single bound nucleic acid target, s1, within the microarray spot after the first washing step at temperature $T_1$ for a fixed time period $\Delta t$ is given by a combination of the first order kinetics and Arrhenius equations:

$$C^{s1}_{T_1} = C^1_0 \cdot \exp\left(-A^{s1} \cdot \exp\left(-\frac{E^{s1}_a}{RT_1}\right) \cdot \Delta t\right) \quad (1)$$

where $C^{s1}_0$ is the concentration of the nucleic acid s1 before the washing, $E^{s1}_a$ is the activation energy, $A^{s1}$ is the preexponential coefficient, and $R$, the universal gas constant. The same equation can be used to calculate the concentration following the next washing step except that $C^{s1}_0$ was changed to $C^{s1}_{T_1}$, and $T_1$ in the exponent was changed to $T_2$. All subsequent washing steps can be described analogously. Thus, the concentration of the nucleic acid following the $n$th washing step can be expressed as follows:

$$C^{s1}_{T_n} = C^{s1}_0 \cdot \exp\left(-A^{s1} \cdot \Delta t \cdot \sum_{i \le n} \exp\left(-\frac{E^{s1}_a}{RT_i}\right)\right). \quad (2)$$

In order to use the equation above for comparing one dissociation curve to another, one has to log transform this equation. Since the signal intensity is proportional to the surface concentration, $C$ can be substituted with $I$ (signal intensity):

$$\log\left(I^{s1}_{T_n}\right) = \log\left(I^{s1}_0\right) - A^{s1} \cdot \Delta t \cdot \sum_{i \le n} \exp\left(-\frac{E^{s1}_a}{RT_i}\right). \quad (3)$$

Considering another target, s2, whose dissociation profile is being compared to s1, the same formalism holds true. If we assume that $E^{s1}_a \approx E^{s2}_a$, which is reasonable in accordance with Ikuta et al. (1987) and our previous study (Pozhitkov et al., 2007), we can express log values of s2 profile via log values of s1 profile:

$$\log\left(I^{s2}_{T_n}\right) = \frac{A^{s2}}{A^{s1}} \log\left(I^{s1}_{T_n}\right) - \frac{A^{s2}}{A^{s1}} \log\left(I^{s1}_0\right) + \log\left(I^{s2}_0\right). \quad (4)$$

Hence, the profile s2 is expressed via s1 as a straight line with the slope related to the difference in dissociation kinetics (the last two

summands are constants). If the slope is >1, therefore $A^{s2}$ is higher than $A^{s1}$, which means that the curve being compared (s2) corresponds to faster kinetics than s1.

For actual analysis, finding the slope was done by linear regression analysis of the log2 transformed signal intensities for the curves being compared. The $R^2$ was used to assess the goodness of fit and reliability of slope calculations. We only considered probes that had $R^2 > 0.90$ for the regression line.

One might argue that the abovementioned approach should not be applied to mixtures of targets because the observed signal intensity of the mixture is a sum of signal intensities of various target and nontarget duplexes, with each duplex having different kinetics. Log transformation of the sum might result in nonlinearity. Nonetheless, we observed a linear relationship, indicating that there is a dominating duplex that determines most of the observed intensity or else the mixture of duplexes behaves like an average duplex.

## 3. Results

### 3.1. Characterization of the bioreactor sample

A bioreactor sample was subjected to 454 pyrosequencing using universal amplicons (coverage of 93% of sequenced 16S rRNA genes) flanking a hypervariable region (V4) of the 16S rRNA gene. The RDP classifier (Wang et al., 2007) was then used to assign 16S rRNA sequences to the taxonomical hierarchy with a 50% confidence level. Out of approximately 10,000 sequences examined, none were classified as *Burkholderia*. Therefore, *B. xenovorans*, or any closely related species within this genus, was not experimentally detected in the bioreactor sample.

### 3.2. Reproducibility of kinetics

A physical model for the NTD curves based on immobilized probes was evaluated in our recent study (Pozhitkov et al., 2007) and the equation for the relationship between signal intensity and temperature (and time) is presented in the Materials and methods section (Eq. (2)). It can be assumed that any significant experimental deviation from the model would suggest technical problems with the dissociation and/or the microarrays. Because the current study used a different batch of arrays than our previous study (Pozhitkov et al., 2007), we had to determine whether or not the slides from the new batch followed the same kinetic model as the previous batch, otherwise they could not be legitimately compared. We used two approaches to test for potential "batch effects". In the first approach, we determined the dissociation kinetics (activation energy, $E_a$, and preexponential coefficients, $A$) of the pure reference targets, as previously described (Pozhitkov et al., 2007). We then substituted the values of $E_a$ and $A$ into the Arrhenius equation (to model the data), and used the determined model to compare with the experimental data. In theory, the model should closely follow the experimental data. In the second approach, we compared the slopes of the log2 transformed intensity values for each probe; the Materials and methods section explains the approach. In theory, the slope of the same probe by different batch should be close to 1.

First, we examined the actual data to see if they followed the theoretical model. Fig. 2 (panels A and B) is representative of the probes examined, indicating good agreement between actual and modeled values for both batches.

Second, using the first 11 intensity values (*i.e.* 20 °C to 40 °C), we examined the slopes of log2–log2 plots. The 20 to 40 °C range

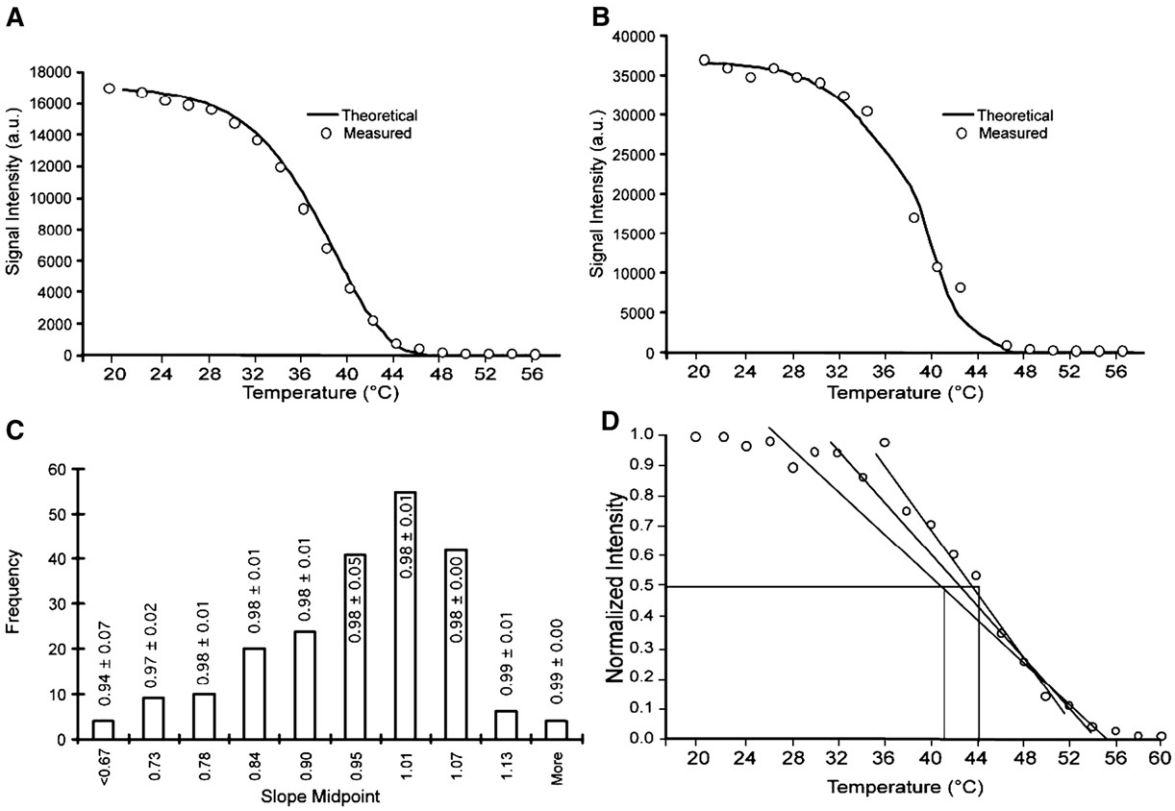


**Fig. 2.** Comparison of observed intensity values versus those predicted by the model for probe LB32 and pure target *B. xenovorans* hybridized to two different array batches (panel A, batch used in previous study (Pozhitkov et al., 2007); panel B, batch used in this study). The frequency histogram of the slopes of pure targets by array batch is shown in panel C. The slope of the batches was defined as $y/x$; where $x$ is one batch and $y$ is the other. The average $R^2$ (±SD) of the slope regression lines for each bar is shown. Note that in general, higher variability in $R^2$ occurs when the slope values were less than 1 ($n$=214 probes) (see text). An example of the difficulty in determining the $T_d$ for probe LB10 is shown in panel D. The choice of the optimal regression line is rather arbitrary, since the $T_d$ ranged from 40 °C to 44 °C.

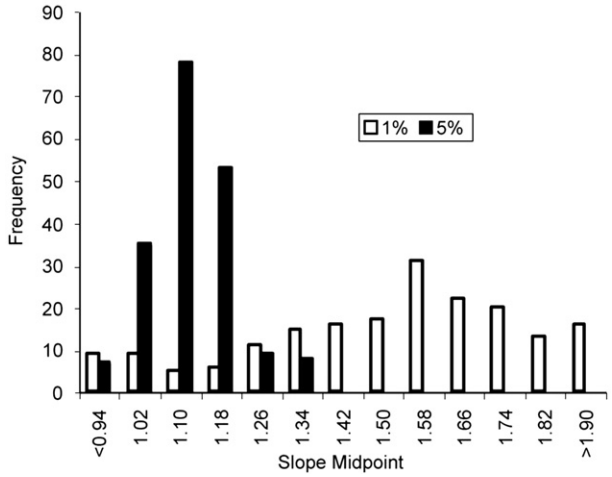


**Fig. 3.** Frequency distribution of slopes by target mixture ($n$=190 probes) using the 10% sample as the reference. Three different proportions of target *B. xenovorans* were added to a mixture composed of unknown targets from an anaerobic bioreactor sample: 5% (black bars), 1% (white bars). Each slope represents the log intensity values of 11 points in the range of 20 to 40 °C, relative to those of 10% mixture. Specifically, the slope of the 1% mixture was defined as $y/x$; where $y$=10%, $x$=1% and the slope of the 5% mixture was defined as $y/x$, where $y$=10%, $x$=5%. Decreasing the concentrations of specific *B. xenovorans* target in the mixtures significantly increased the slope of intensity values (paired Student *T*-test), which resulted in a shift in the dissociation curves before that of the 10% reference.

corresponds to the majority of the curve data points (see Fig. 2, panels A and B). Importantly, previous studies such as Liu et al. (2001) (Fig. 1) also showed that most of the dissociation occurred in that range. Beyond 40 °C, the signal to noise ratio is low since intensity of the signal approaches the detection limits of the system (see Pozhitkov et al., 2005). The log2–log2 slopes had a median value (±SD) of 0.95±0.11 a.u. for all probes (see Fig. 2, panel C). The Shapiro–Wilk test revealed that the slope data was *not* normally distributed ($W$=0.94, $P$<0.001, $n$=215 probes), but rather slightly negatively skewed ($_1 G_1$=−1.1) with high kurtosis ($G_2$=2.63). The presumptive reason for the skew was that approximately 20% of the probes (42/215) had slopes that were less than 0.90. Careful inspection of the dissociation data for these probes suggested that low initial signal intensity values and/or high signal to noise ratios were responsible for the deviation, particularly at high temperatures, *i.e.*, beyond 40 °C. The intensity values of these probes might have been affected by the detection limits of the system, as revealed by the $R^2$ of the linear regression (which was not close to 1). Nevertheless, high kurtosis in the slope distribution indicated that a majority of the slopes fell within two standard deviations of the median (approximately 94% of all probes). One can therefore conclude from these two independent tests that the "batch effect" was not significant because dissociations of duplexes from the two batches of slides followed the same physical kinetic model, and the slope of most probes was close to 1.

### 3.3. Concentration-induced shift

Physicochemical simulations conducted in our previous study (Pozhitkov et al., 2007, see Fig. 6) revealed that decreasing concentrations of a specific target in a mixture should shift the dissociation curves to the left for the hypothetical case where the nonspecific targets dissociate before specific ones. In this study, we compared mixtures of targets containing 1%, 5%, and 10% of the specific target. Using the log–log comparison approach, we set the curves of probes of 10% mixtures as the reference and compared them to curves of the 1% and 5% mixtures. As predicted by theory (Pozhitkov et al., 2007), decreasing the concentration of the reference target in the mixtures did, in fact, shift the dissociation curves to the left, *i.e.*, towards lower temperatures as indicated by the slopes >1 (Fig. 3). Paired Student *T*-tests revealed significant differences between the slopes for the 1% and 5% mixtures. In general, the lower the concentration of the reference target, the greater the shift.

### 3.4. Dissociation of the pure target

One of the assumptions of the NTD approach is that a match between a curve recorded from a mixture and that from a pure target suggests a positive identification of the target in the mixture. Having compared the dissociations of the 1% and 5% mixtures to those from a 10% mixture, which was used as a reference, we sought to compare the dissociations of the 10% mixture to those of a pure target. Fig. 4 (panel A) shows that in general, the pure target actually dissociated before the 10% mixture. Comparison of Fig. 4 (panel B) to Fig. 3 shows that on average, the dissociation rates of the pure target

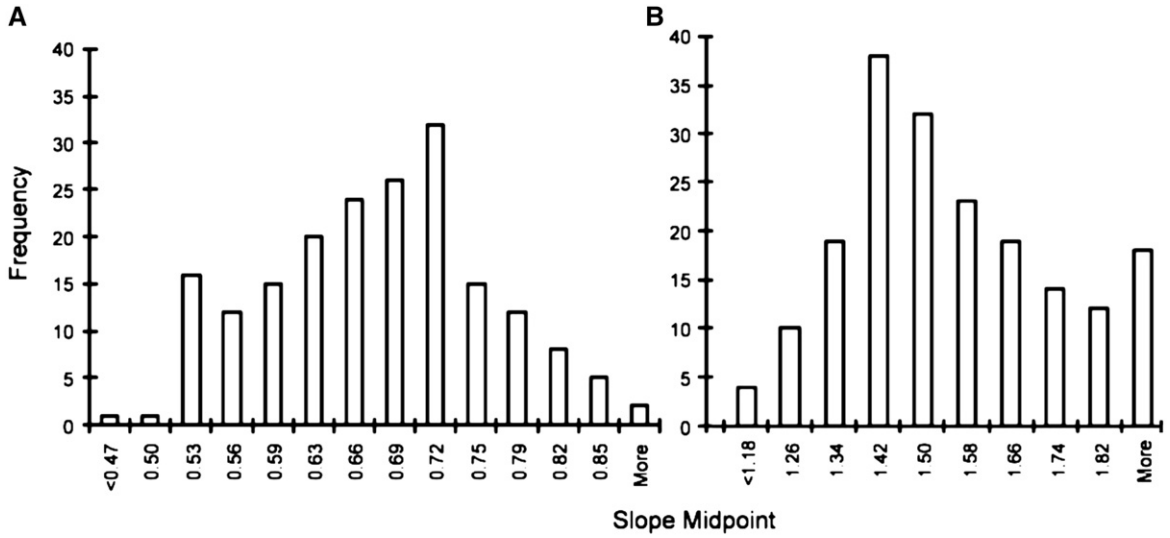


**Fig. 4.** Comparison of 10% mixture curves using the pure target as a reference ($n$=189 probes). The slope of the pure target in panel A was defined as $y/x$, where $y$=pure target and $x$=10% mixture, and the slope of panel B was defined as $y$=10% mixture and $x$=pure target, in order to compare with Fig. 3.

were somewhere between those of the 1% and 5% mixtures. As discussed below, this finding indicates that the dissociation of a pure target has little or no relevance to that of mixtures.

### 3.5. Dissociation of complex background targets

The previous experiments (above) revealed that some curves of the pure target dissociated before curves of the mixtures containing the target at different concentrations (1%, 5%, and 10%). Logically, this also means that many bioreactor targets in the mixtures, presumably most of them mismatched (nonspecific) to the probes, dissociated after the pure target. We compared the dissociations of the pure target with that of the bioreactor. Note, in this case, the bioreactor nucleic acids did not contain any of the *B. xenovorans*, *i.e.*, 0% mixture. We found that only a small fraction of the 0% mixture curves dissociated before the pure target, with a majority of the 0% mixture curves dissociating after the pure target (Fig. 5). This finding is consistent with our previous study (Pozhitkov et al., 2007); specifically, a perfectly matched target does not necessarily dissociate after targets containing mismatches to a probe.

In the absence of the *B. xenovorans* target, the observed distribution of the log–log slopes (Fig. 5) might be due to some probes binding to universal regions of the bioreactor targets. However, since the composition of the targets is not known, it is difficult to assess whether certain probes bind more targets than others, and/or if some probes are binding targets that have similar sequence homology to *B. xenovorans*. The "universality" of a probe can be estimated by examining the number of times it perfectly matches rRNA gene sequences in the RDP database (Cole et al., 2007). A "more universal" probe should match more RDP sequences than otherwise.

When we determined the number of times each of the 59 probes perfectly matched target sequences in the RDP database, we found that the number of hits ranged from 0 (negative control) to 303,045 (out of 451,545 possible sequences). Fig. 6 shows the relationship between the numbers of microbial targets having rRNA genes perfectly matching a given probe (*i.e.*, hits), with the log–log dissociation slope of that probe. We also conducted these experiments allowing for one to three mismatches. Evidently, no defined relationship could be resolved. Hence, even considering the complementary of the probes to sequences in the RDP database, we can draw no solid conclusions on whether or not the observed log–log dissociation slope is due to "universal" probes, but it seems likely that it is not. The Discussion section further indicates the importance of the 0% dissociation experiment, because the bioreactor targets alone (without the *B. xenovorans*) act as a negative control for the NTD calibration.

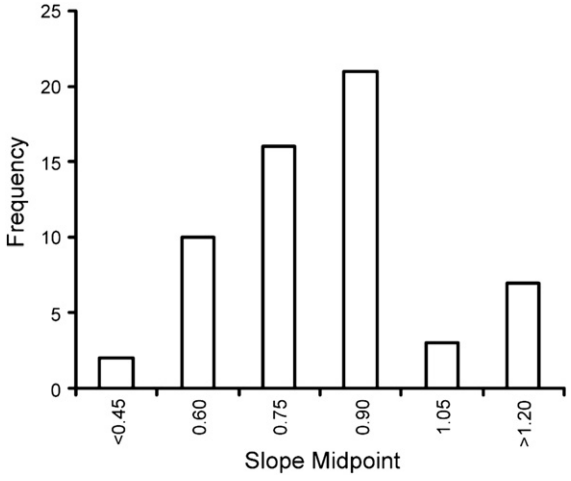


**Fig. 5.** Frequency distribution of the slopes for bioreactor targets (not containing *B. xenovorans*) and pure *B. xenovorans* target, used as the reference ($n$=59 probes). The slope of bioreactor targets was defined by $y/x$, where $y$ is the bioreactor targets and $x$ is the pure target. Note that in general, the slopes are less than 1, indicating that bioreactor targets are dissociating after the pure target.

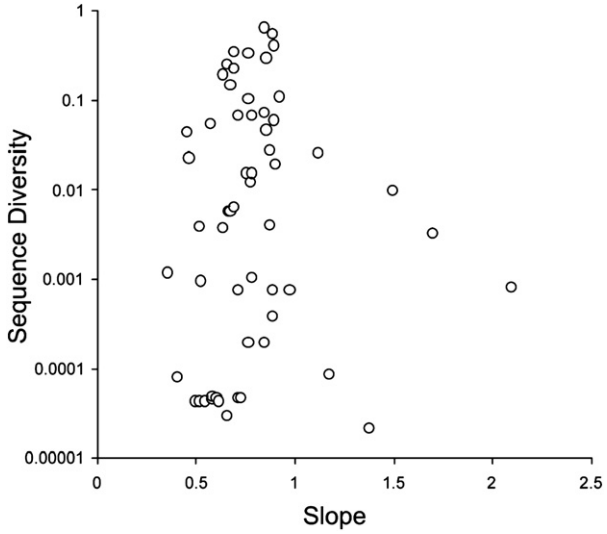


**Fig. 6.** Relationship between sequence diversity and slope of bioreactor targets (not containing *B. xenovorans*) and pure *B. xenovorans* target, used as the reference ($n$=59 probes) (same slopes as used in Fig. 5). The sequence diversity was based on log10 ratio of number of hits for a probe sequence in the RDP database project II (Cole et al., 2007) divided by the number of possible targets in RDP ($n$=451,545). Therefore, a probe having a low hit ratio has a low diversity, while a probe that has a high hit ratio (*i.e.*, it is universal) will have high sequence diversity.

## 4. Discussion

### 4.1. Variability of dissociation curves

The dissociation kinetics of replicated NTD experiments did not follow a normal distribution (Fig. 2, panel C), though most of the slopes of the log–log plots were close to 1. To our surprise, several curves had differences in their dissociation kinetics between replicates. These results imply that the "50% melting temperatures" ($T_d$s) of some probes in the two replicated experiments may be different from one another. This is an important finding, because the width of the distribution (see Fig. 2, panel C) determines the systematic variability that should be assessed before using the NTD approach for identifications. To our knowledge, none of the previous studies (DeLosReyes et al., 1997, 1998; ElFantroussi et al., 2003; Eyers et al., 2006; Hansen et al., 1999; Kelly et al., 2005; Koizumi et al., 2002; Li et al., 2004; Liu et al., 2001; Loy et al., 2002; McMahon et al., 1998; Mobarry et al., 1996; Siripong et al., 2006; Urakawa et al., 2002, 2003; Zheng et al., 1996) explicitly assessed this variability.

### 4.2. Physical comparison of NTD curves

Many previous studies compared NTD curves by normalizing them (*e.g.*, $I_{new\_value}=I_{old\_value}-I_{min}/I_{max}-I_{min}$, where $I$ is the signal intensity) and overlaying the resulting curves onto one plot (*e.g.*, ElFantroussi et al., 2003). Also, rather than visually inspecting the curves, various indices have been developed to quantitatively assess differences between curves such as the $T_d$ or discrimination index (DI) (see Eyers et al., 2006; Li et al., 2004; Urakawa et al., 2002, 2003; Zheng et al., 1996). There are several difficulties associated with $T_d$. For example, assigning the $T_d$ is rather arbitrary, as clearly shown in Fig. 2, panel D. Moreover, comparison of two curves by $T_d$ has several sources of error, which introduce uncertainty to interpreting curve differences. For example, there are errors associated with normalizing each curve because intensity minimum and maximum are determined by two

points, and in some cases, the maximum value is uncertain (as discussed in Urakawa et al., 2002) because initial signal intensity values decrease rapidly (*i.e.*, there is no plateau) with increasing temperature due to nonequilibrium conditions. Also, there is an error associated in picking the optimal regression line for $T_d$ determination (Fig. 2, panel D). Other methods that directly compare curves suffer from similar artifacts (*e.g.*, Li et al., 2004; Urakawa et al., 2003) because they are mostly *ad hoc* procedures rather than thoroughly verified physical models.

Given the current trend in the development of accurate physical models for quantifying nucleic acid targets using microarrays (*e.g.*, see Halperin et al., 2006), it is critical that microbial identification studies follow the same path, because a truly analytical tool will eventually be developed that will drastically improve the robustness of existing microarray studies. In fact, in contrast to most of the previous microbial identification studies, the array platform used in this study has previously been shown to yield accurate physical modeling of dissociation curves (Pozhitkov et al., 2007). Based on this physical model, it is easy to compare two curves by plotting log2 transformed signal intensities of one curve against another (see Materials and methods for details). The resulting plot is a regression line with its slope indicating whether one curve is dissociating before or after another. If the slope is > 1, then the log2 transformed data on the *y*-axis is dissociating before the one on the *x*-axis. An additional advantage of this approach is that error in calculating the slope of the dissociations can be determined by the $R^2$ of the regression line.

### 4.3. "Unpredictable" probes

It is important to recognize that the goal of this study was *not* to detect *B. xenovorans* but rather to study the dissociation of various probe–target duplexes. Therefore, no steps were taken to optimize the probes on the arrays. We mention probe optimization because several microbiological publications have alluded to the need to use a set of optimized probes (e.g., Loy and Bodrossy, 2006). Probe optimization involves removing "unpredictable" probes that have low signal intensity and that cross-hybridize to nontarget molecules. The "unpredictable" probes are determined by hybridizing with a sample containing sequences closely related to the targets being detected. Probes yielding similar signal intensities and/or identical NTD curves between the closely related samples are later removed from the microarray design. From a physicochemical perspective, there are no unpredictable probes — it is merely an interaction of nucleic acids. We intentionally did not remove any probes from the pool of probes designed to be complimentary to the *B. xenovorans* target because we expected cross-hybridization to occur to some degree for all probes on the microarrays.

Interestingly, we were able to identify probes that would be deemed as "unpredictable" from the perspective of most of the microbiological studies. These probes are shown on Fig. 5. Specifically, the probes having slopes less than 1 correspond to the situation when the pure target dissociates faster than the targets from the bioreactor (Pozhitkov et al., 2007). In fact, most of the previous microbiological studies have shown only the opposite situation, such as shown in Fig. 1. We believe that the reason for the incompleteness of the picture is due to this subjective notion of "bad" or "unpredictable" probes.

### 4.4. Practical challenges of the NTD approach

Based on these findings, the NTD does not seem to be a practical approach for identifying microbes especially for high throughput systems because it needs individualized optimization of each probe. Specifically, the NTD approach requires: (i) calibrating each probe to determine its unique concentration dependency with different targets, and (ii) conducting dissociations with a negative control, *i.e.*, a biomedical or environmental sample not containing targets in question. Probes that have slopes close to 1 (Fig. 5) would not be optimal for identifying the target (*e.g.*, *B. xenovorans*) because there is no difference between the curves obtained from the pure target and those obtained from the negative control (*i.e.* the bioreactor sample not containing *B. xenovorans* LB400). In other words, these probes would not be able to distinguish between the two samples. It is important to recognize that that this second reason alone is not sufficient in view of the abovementioned concentration dependencies (*i.e.*, the first reason).

In conclusion, a method was developed for comparing NTD curves that was less error prone than conventional methods, and based on the physicochemistry of dissociation. The method showed that there was a significant concentration dependency of the NTD curves. This implies that the rationale for microbial identification using NTD in complex target mixtures is compromised, and that the NTD approach is not practical for probe evaluation in high throughput systems because it needs individualized optimization of each probe. Also, contrary to the previous studies, and consistent with our earlier study (Pozhitkov et al., 2007), it was found that some nonspecific targets dissociate at higher temperatures than the specific ones.

## Acknowledgements

We thank Dieter M. Tourlousse for performing the 454 sequencing analysis of the bioreactor sample. This research was supported in part by a grant from the US National Oceanic and Atmospheric Administration (NA05NOS4261163), National Institutes of Health (Grants R01 RR018625-03 and 1U01DE014955-01), US Environmental Protection Agency (RD83301001), and the University of Washington Provost Bridge and Royalty funds.

## References

Cole, J.R., Chai, B., Farris, R.J., Wang, Q., Kulam-Syed-Mohideen, A.S., McGarrell, D.M., Bandela, A.M., Cardenas, E., Garrity, G.M., Tiedje, J.M., 2007. The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. Nucleic Acids Res. 35, D169–D172.

DeLosReyes, F.L., Ritter, W., Raskin, L., 1997. Group-specific small-subunit rRNA hybridization probes to characterize filamentous foaming in activated sludge systems. Appl. Environ. Microbiol. 63, 1107–1117.

DeLosReyes, M.F., DeLosReyes, F.L., Hernandez, M., Raskin, L., 1998. Quantification of *Gordona amarae* strains in foaming activated sludge and anaerobic digester systems with oligonucleotide hybridization probes. Appl. Environ. Microbiol. 64, 2503–2512.

ElFantroussi, S., Urakawa, H., Bernhard, A.E., Noble, P.A., Kelly, J.J., Stahl, D.A., 2003. Direct profiling of environmental microbial populations by thermal dissociation analysis of native Ribosomal RNAs hybridized to oligonucleotide microarrays. Appl Environ. Microbiol. 69, 2377–2382.

Eyers, L., Smoot, J.C., Smoot, L.M., Bugli, C., Urakawa, H., McMurry, Z., Siripong, S., ElFantrousii, S., Lambert, P., Agothos, S.N., Stahl, D.A., 2006. Discrimination of shifts in a soil microbial community associated with TNT-contamination using a functional ANOVA of 16S rRNA hybridized to oligonucleotide microarrays. Environ. Sci. Technol. 40, 5867–5873.

Halperin, A., Buhot, A., Zhulina, E.B., 2006. On the hybridization isotherms of DNA microarrays: the Langmuir model and its extensions. J. Phys. Condens. Matter. 18, S463–S490.

Hansen, K.H., Ahring, B.K., Raskin, L., 1999. Quantification of syntrophic fatty acid-B-oxidizing bacteria in a mesophilic biogas reactor by oligonucleotide probe hybridization. Appl. Environ. Microbiol. 65, 4767–4774.

Held, G.A., Grinstein, G., Tu, Y., 2006. Relationship between gene expression and observed intensities in DNA microarrays—a modeling study. Nucleic Acids Res. 34, e70.

Ikuta, S., Takagi, K., Wallace, R.B., Itakura, K., 1987. Dissociation kinetics of 19 base paired oligonucleotide-DNA duplexes containing different single mismatched base-pairs. Nucl. Acids Res. 15, 797–811.

Kelly, J.J., Siripong, S., McCormack, J., Janus, L.R., Urakawa, H., ElFantroussi, S., Noble, P.A., Sappelsa, L., Rittmann, B.E., Stahl, D.A., 2005. DNA microarray detection of nitrifying bacterial 16S rRNA in wastewater treatment plant samples. Water Res. 39, 3229–3238.

Koizumi, Y., Kelly, J.J., Nakagawa, T., Urakawa, H., ElFantroussi, S., Muzaini, A., Fukui, M., Urushigawa, Y., Stahl, D.A., 2002. Parallel characterization of anaerobic toluene- and ethylbenzene-degrading microbial consortia by PCR-denaturing gradient gel electrophoresis, RNA-DNA membrane hybridization, and DNA microarray technology. Appl. Environ. Microbiol. 68, 3215–3225.

Li, E.S.Y., Ng, J.K.K., Wu, J.-H., Liu, W.T., 2004. Evaluating single-base-pair discriminating capability of planar oligonucleotide microchips using a nonequilibrium dissociation approach. Environ. Microbiol. 6, 1197–1202.

Liu, W.T., Mirzabekov, A.D., Stahl, D.A., 2001. Optimization of an oligonucleotide microchip for microbial identification studies: a non-equilibrium dissociation approach. Environ Microbiol. 3, 619–629.

Loy, A., Bodrossy, L., 2006. Highly parallel microbial diagnostics using oligonucleotide microarrays. Clin. Chim. Acta. 363, 106–119.

Loy, A., Lehner, A., Lee, N., Adamczyk, J., Meier, H., Ernst, J., Schleifer, K.H., Wagner, M., 2002. Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. Appl. Environ. Microbiol. 68, 5064–5081.

McMahon, K.D., Stahl, D.A., Raskin, L., 1998. A comparison of the use of *in vitro*-transcribed and native rRNA for the quantification of microorganisms in the environment. Microb. Ecol. 36, 362–371.

Mei, R., Hubbell, E., Bekiranov, S., Mittmann, M., Christians, F.C., Shen, M.M., Lu, G., Fang, J., Liu, J., 2003. Probe selection for high-density oligonucleotide arrays. Proc. Natl. Acad. Sci. USA. 100, 11237–11242.

Mobarry, B.K., Wagner, M., Urbain, V., Rittmann, B.E., Stahl, D.A., 1996. Phylogenetic probes for analyzing abundance and spatial organization of nitrifying bacteria. Appl. Environ. Microbiol. 62, 2156–2162.

Naef, F., Magnasco, M.O., 2003. Solving the riddle of the bright mismatches: labeling and effective binding in oligonucleotide arrays. Phys. Rev. E. 68 Art. No. 011906 Part 1.

Pozhitkov, A., Noble, P.A., 2007. Comment on "Discrimination of shifts in soil microbial communities using nonequilibrium thermal dissociation and gel pad array technology". Environ. Sci. Tech. 41, 1797–1798.

Pozhitkov, A., Chernov, B., Yershov, G., Noble, P.A., 2005. Evaluation of gel-pad oligonucleotide microarray technology using artificial neural networks. Appl. Environ. Microbiol. 71, 8663–8676.

Pozhitkov, A., Noble, P.A., Domazet-Loso, T., Staehler, P., Beier, M., Tautz, D., 2006. Tests of rRNA hybridization to microarrays suggest that hybridization characteristics of oligonucleotide probes for species discrimination cannot be predicted. Nucl. Acids Res. 34, e66.

Pozhitkov, A.E., Stedtfeld, R.D., Hashsham, S.A., Noble, P.A., 2007. Revision of the nonequilibrium thermal dissociation and stringent washing approaches for identification of mixed nucleic acid targets by microarrays. Nucl. Acids Res. 35, e70.

Raskin, L., Poulsen, L.K., Noguerea, D.R., Rittman, B.E., Stahl, D.A., 1994a. Quantification of methanogenic groups in anaerobic biological reactors by oligonucleotide probe hybridization. Appl. Environ. Microbiol. 60, 1241–1248.

Raskin, L., Stromley, J.M., Rittman, B.E., Stahl, D.A., 1994b. Group-specific 16S rRNA hybridization probes to describe natural communities of methanogens. Appl. Environ. Microbiol. 60, 1232–1240.

Siripong, S., Kelly, J.J., Stahl, D.A., Rittmann, B.E., 2006. Impact of prehybridization PCR amplification on microarray detection of nitrifying bacteria in wastewater treatment plant samples. Environ. Microbiol. 8, 1564–1574.

Urakawa, H., Noble, P.A., ElFantroussi, S., Kelly, J.J., Stahl, D.A., 2002. Single-base-pair discrimination of terminal mismatches by using oligonucleotide microarrays and neural network analyses. Appl. Environ. Microbiol. 68, 235–244.

Urakawa, H., ElFantroussi, S., Smidt, H., Smoot, J.C., Tribou, E.H., Kelly, J.J., Noble, P.A., Stahl, D.A., 2003. Optimization of single-base-pair mismatch discrimination in oligonucleotide microarrays. Appl. Environ. Microbiol. 69, 2848–2856.

Wang, Q.G., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol. 73, 5261–5267.

Wick, L.M., Rouillard, J.M., Whittam, T.S., Gulari, E., Tiedje, J.M., Hashsham, S.A., 2006. On-chip nonequilibrium dissociation curves and dissociation rate constants as methods to assess specificity of oligonucleotide probes. Nucl. Acids Res. 34, e26.

Wu, Z., Irizarry, R.A., 2004. Stochastic models inspired by hybridization theory for short oligonucleotide microarrays. Proceedings of RECOMB'04, San Diego, CA.

Zhang, L., Miles, M.F., Aldape, K.D., 2003. A model of molecular interactions on short oligonucleotide microarrays. Nat. Biotechnol. 21, 818–821.

Zhang, Y., Hammer, D.A., Graves, D.J., 2005. Competitive hybridization kinetics reveals unexpected behavior patterns. Biophys. J. 89, 2950–2959.

Zheng, D., Alm, E.W., Stahl, D.A., Raskin, L., 1996. Characterization of universal small-unit rRNA hybridization probes for quantitative molecular ecology studies. Appl. Environ. Microbiol. 62, 4504–4513.