3-1-2010

Beyond Affymetrix Arrays: Expanding the Set of Known Hybridization Isotherms and Observing Pre-Wash Signal Intensities

Alexander Pozhitkov
University of Southern Mississippi, alexander.pozhitkov@evolbio.mpg.de

Idrissa Boube
University of Southern Mississippi

Marius Brouwer
University of Southern Mississippi

Peter A. Noble
University of Washington, panoble@washington.edu

Published online 6 December 2009

Nucleic Acids Research, 2010, Vol. 38, No. 5 ***e28***
doi:10.1093/nar/gkp1122

Beyond Affymetrix arrays: expanding the set of known hybridization isotherms and observing pre-wash signal intensities

Alex E. Pozhitkov[1,2], Idrissa Boube[1], Marius H. Brouwer[1] and Peter A. Noble[3,4,5,*]

[1]Gulf Coast Research Laboratory, University of Southern Mississippi, 703 E Beach Dr, Ocean Springs, MS 39564, USA, [2]Max-Planck-Institut fuer Evolutionsbiologie, August-Thienemannstrasse 2, 24306 Ploen, Germany, [3]Civil and Environmental Engineering, [4]Department of Periodontics, University of Washington, Seattle, WA 98195 and [5]PhD Microbiology Program, Department of Biological Sciences, Alabama State University, 1627 Hall St. Montgomery, AL 36104, USA

Received April 17, 2009; Revised November 11, 2009; Accepted November 13, 2009

ABSTRACT

Microarray hybridization studies have attributed the nonlinearity of hybridization isotherms to probe saturation and post-hybridization washing. Both processes are thought to distort 'true' target abundance because immobilized probes are saturated with excess target and stringent washing removes loosely bound targets. Yet the paucity of studies aimed at understanding hybridization and dissociation makes it difficult to align physicochemical theory to microarray results. To fill the void, we first examined hybridization isotherms generated on different microarray platforms using a ribosomal RNA target and then investigated hybridization signals at equilibrium and after stringent wash. Hybridization signal at equilibrium was achieved by treating the microarray with isopropanol, which prevents nucleic acids from dissolving into solution. Our results suggest that (i) the shape of hybridization isotherms varied by microarray platform with some being hyperbolic or linear, and others following a power-law; (ii) at equilibrium, fluorescent signal of different probes hybridized to the same target were not similar even with excess of target and (iii) the amount of target removed by stringent washing depended upon the hybridization time, the probe sequence and the presence/absence of nonspecific targets. Possible physicochemical interpretations of the results and future studies are discussed.

INTRODUCTION

Oligonucleotide microarrays have been widely used to determine the relative abundance of transcripts in gene expression studies (1,2) and the composition of rRNA targets in microbial populations (3). The advantage of microarrays over other technologies, such as those based on PCR amplification, is that microarrays allow the simultaneous detection of multiple targets within the same sample (4,5). This high throughput is due to hundreds of thousands of different immobilized oligonucleotide probes; each one acting as an individual sensor with its own specificity and sensitivity to different nucleic acid targets in solution (6). Interpreting probe signal is problematic, however, because the underlying physicochemistry of microarray target hybridization and dissociation has not been adequately established (7) and few studies have successfully linked experimental results to physicochemical models. Systematic studies aimed at thoroughly understanding the behavior of targets hybridized to probes would significantly improve the interpretation of signals from oligonucleotide microarrays.

One area of discrepancy is the lack of agreement of physicochemical theory to experimental results for probe–target saturation (8). At low target abundance, the signal intensity of an immobilized probe is linearly proportional to the concentration of complementary targets. But, at moderate to high target abundance, signal intensity is nonlinearly proportional to the concentration of complementary targets in solution because it is thought that there are few binding sites available for targets due to probe saturation. In other words, a doubling of a highly abundant target does not result in a similar increase (i.e. a doubling) of fluorescent

*To whom correspondence should be addressed. Tel: +1 206 409 6664 or +1 334 229 8451; Fax: +1 334 229 6709; Email: panoble@washington.edu

signal intensity. With infinite amount of target, the fluorescent signal should approach a maximum value (i.e., a horizontal asymptote). The relationship between the target abundance and signal intensity is commonly referred to as a 'hybridization isotherm' and it is thought to follow Langmuir's adsorption model (9)

$$\theta \equiv \frac{[PT]}{P_0} = \frac{[T]}{K+[T]} + b \quad \mathbf{1}$$

where $[PT]$ is the surface concentration of probe–target duplexes; P_0 is the surface concentration of probes before binding; θ is the fraction of target-occupied probes; $[T]$ is the free target concentration in solution; K is the equilibrium constant, i.e. $[P][T]/[PT]$, with $[P]$ being the surface concentration of unoccupied probes and b is the background intensity. According to this model, at infinite target abundance $[T]$, all probes should saturate to the same level [i.e. in Equation (1), $\theta \to 1$, regardless of K]. However, experimental evidence suggests otherwise—some probes saturate with target at lower than expected signal intensities [e.g. see Figure 1 in (10)]. Burden *et al*. (8) and Held *et al*. (11) reasoned that this discrepancy was due to the dissociation of target molecules during stringent wash. But, with exception to Skvortsov *et al*. (10), no microarray experiments have been conducted to thoroughly test this hypothesis. In addition, most of the current understanding of microarray hybridization and dissociation is based on physicochemical studies that used the Affymetrix datasets (9,11–15). The analysis of other microarray platforms would substantially broaden our understanding of microarray hybridization and dissociation in general.

A compelling model explaining the variation in probe hybridization isotherms has been proposed by Skvortsov *et al*. (10). The model is based on the idea that there is competitive hybridization among different targets for the same probe and that nontargets (i.e. targets that are not complementary to a probe) partially inhibit target hybridization due to steric hindrances. The notion that sufficient hybridization time is needed to achieve equilibrium distinguishes the Skvortsov *et al*. model from the rest (e.g. 9,14–18). Hybridization time therefore plays a key role in explaining the discrepancies among probe hybridization isotherms. For example, when a microarray is hybridized for a short period of time, the following is taking place: (i) equilibrium is not attained for all probe–target duplexes; as a consequence, (ii) there is a population of bound targets, a population of loosely bound targets and a population of loosely bound nontargets. The population of loosely bound targets cannot completely bind to complementary probes due to the nontargets sterically inhibiting their binding. Therefore, the loosely bound targets are removed from the microarray by the stringent wash along with the nontarget sequences, resulting in lower than expected probe signal intensities. In contrast, when a microarray is hybridized for an extended period of time: (i) equilibrium is achieved for all probe–target duplexes; and as a consequence, (ii) only nontargets are removed by the stringent wash, resulting in probe signal that is proportional to the specific targets in solution. Hence, the Skvortsov *et al*. (10) model suggests that it is the removal of loosely bound targets by the stringent wash that contributes to the observed variation in probe saturation. Furthermore, the variation in probe hybridization isotherms in the Affymetrix data sets is a result of insufficient hybridization time and the presence of nontarget sequences, which hindered the hybridization of specific targets.

While the Skvortsov *et al*. (10) study provided substantially new insights into the discrepancy between physicochemistry theory and experimental microarray results, initial probe signals were recorded after a low stringency wash and therefore do not reflect signal at equilibrium. In this study, probe signal at equilibrium was recorded after a microarray was hybridized for 4 h or 24 h and then treated with isopropanol (i.e. no stringent wash). In pure isopropanol, nucleic acids targets are not soluble; therefore, targets cannot dissolve into the solution phase. Hence, the purpose of an isopropanol treatment is to mechanically remove the hybridization solution. All targets that were bound to the microarray probes will presumably remain bound; therefore, the isopropanol treatment reveals a 'snapshot' of signal intensities attained at equilibrium. The probe signal, obtained at equilibrium, offers a baseline for thoroughly examining the effects of wash stringency on microarray probe signal. Our results will show that even a low stringent wash can significantly affect probe signal, resulting in signals that are lower than expected.

Preliminary experiments in our laboratory suggested that hybridization signals on two microarray platforms (Erie Scientific and VWR) were different using the same probes and target. The objectives of this study were therefore two-fold: (i) to compare the hybridization isotherms of nucleic acid targets hybridized to the same probes on the different microarray platforms for the purpose of determining the shape and slopes (K) of the isotherms, and (ii) to determine the effects of stringent wash on probe signal by establishing a baseline of 'true' probe signals at equilibrium. Given that the K of a hybridization isotherm represents the probe–target binding affinity, we also investigated the feasibility of developing a microarray-based nearest-neighbor (NN) model to predict probe–target signal based on nucleic acid sequence.

It should be noted that in contrast to previous studies, this study uses scanner calibration to account for the nonlinear response of scanners in order to obtain true surface concentrations of the probe–target duplexes.

MATERIALS AND METHODS

Ribosomal RNA target

The following primers and annealing temperatures were used to amplify 28S rRNA of *Cyprinodon variegates*: 5′-CTGGTTGATCCTGCCAGT-3′ (forward) and 5′-CYGCAGGTTCACCTACRG-3′ (reverse), 63.1 to 66.0°C. The PCR products were purified using a Promega Wizard SV Gel and PCR Clean-Up System, and ligated into pGEM-T Easy vector (Invitrogen

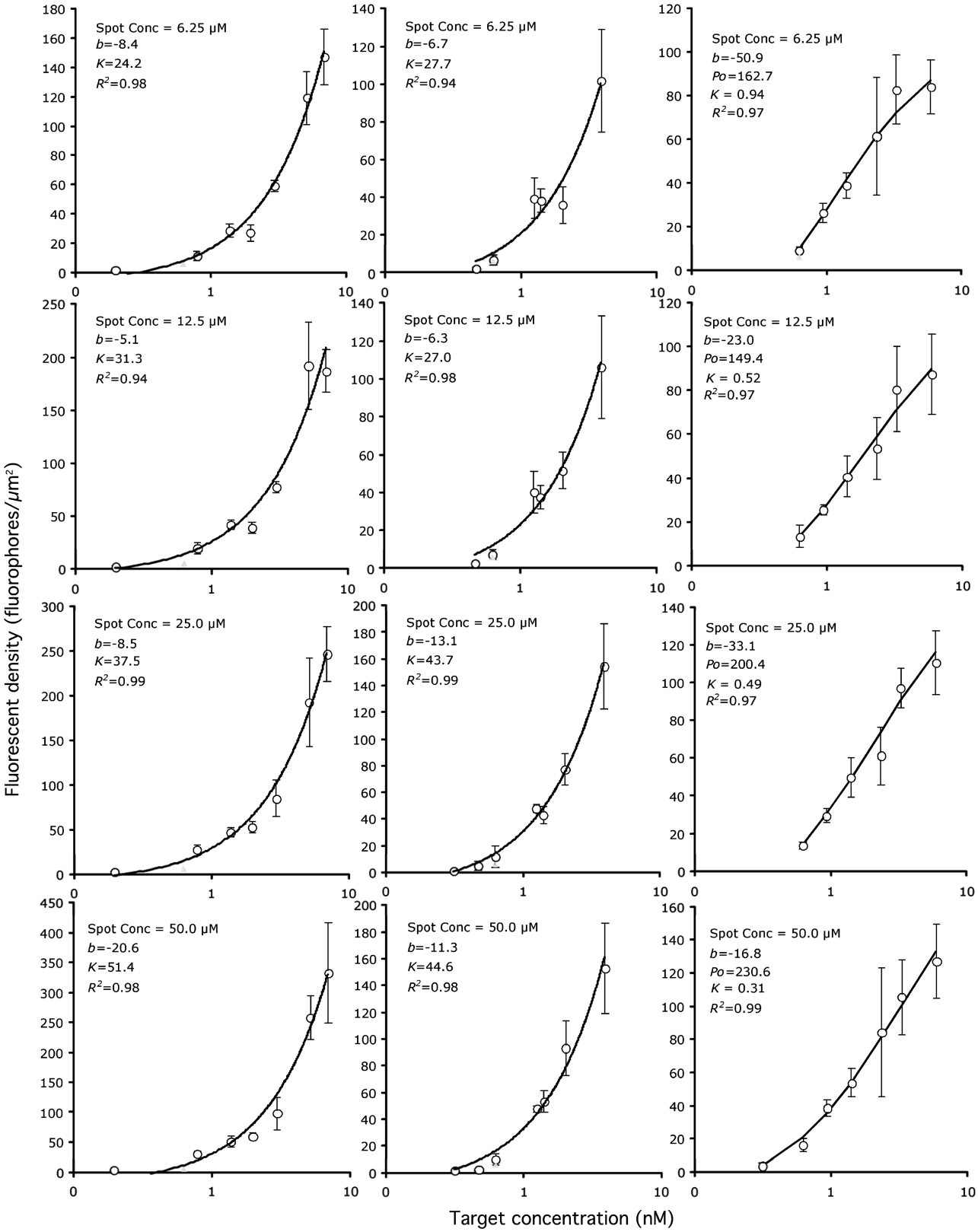

Figure 1. Hybridization isotherms obtained on two different microarray platforms for rRNA target and Probe 596. Left and middle panels represent replicated Erie microarrays while the right panels represent VWR microarrays. Different rows represent distinct spotted probe concentrations. Each data point is the mean of at least six replicated spots. Error is measured as standard deviation. Isotherms from Erie microarrays are linear (but plotted on a log scale) while VWR microarrays followed the Langmuir model.

Corporation, USA). The ligation products were ethanol precipitated and transformed into ElectroMAX DH10B T1 Phage Resistant Cells (Invitrogen Corporation, USA) using a MicroPulser Electroporation Apparatus (Bio-Rad Laboratories, USA). Cells were plated and single colonies picked for further enrichment. Plasmid DNA was isolated using a Wizard Plus SV Minipreps DNA Purification System (Promega Corporation, USA) and sequenced using M13 forward and reverse primers on a CEQ8000 Genetic Analysis System (Beckman Coulter, Inc., USA). The colonies yielded identical sequences and only one colony was chosen for further studies. This colony was sequenced with multiple overlapping primers and a full-length sequence (1919 nt) was submitted to NCBI under accession number EF431912.

Prior to *in vitro* RNA synthesis, a plasmid containing the 18S rRNA insert was restricted with *Sal I* endonuclease (New England Biolabs, USA) in the following reaction mixture: plasmid (340 ng/μl) 25 μl, *Sal I* (20 units/μl) 10 μl, NEBuffer 3 (10×) 5 μl, BSA (100×) 0.5 μl, H_2O 9.5 μl. The reaction mix was incubated at 37°C for 8 h. After the reaction, DNA was extracted with phenol and precipitated with ethanol. RNA was *in vitro* synthesized by using a RiboMAX Large Scale RNA Production System T7 (Promega Corporation, USA). The linearized plasmid was used as the template. The quality of the RNA product was assessed using a Bioanalyzer (Agilent Technologies, Inc., USA). The RNA product was then labeled with AlexaFluor 546 using a ULYSIS Nucleic Acids Labeling kit (Invitrogen Corporation, USA). The only deviation from the manufacturer's protocol was that the concentration of the dye was doubled to increase the amount of labeling. Unincorporated dye was removed by using Micro Bio-Spin 30 RNase-free columns (Bio-Rad Laboratories, USA).

In experiments examining the effects of nonspecific binding, nontarget sequences (unlabeled *Escherichia coli* tRNA, 74–95 nt; Sigma) were added at equimolar concentrations to the rRNA target.

Oligonucleotide probes design

Oligonucleotide probes for the microarrays were complementary to their corresponding target. The probes (20 nt) were generated *in silico* by tiling along the targets using a one base-pair shift. The actual number of oligonucleotides produced and the number of replicates for each probe were determined by the amount of available space on the microarrays.

Fabrication of the microarrays

The microarrays consisted of epoxysilane slides and were produced by two different companies: VWR (USA) and Erie Scientific Company (Portsmouth, NH). Spotting of the probes was accomplished using VersArray ChipWriter Compact System (Bio-Rad Laboratories, USA). Oligonucleotide probes were modified at 5′ with Amine-C6 group (Invitrogen Corporation, USA). Oligonucleotide stock solutions (200 uM) were mixed with water and Micro Spotting Solution Plus 2× (VWR, USA) to achieve various dilutions of the probes.

Information on the custom-designed CombiMatrix and NimbleGen microarrays that was used to back up results of this study is found in Supplementary Data.

Hybridization conditions and washing

The hybridization buffer (7× PBS, BSA 0.2 mg/ml, 0.01% Tween) and 25% v/v of diluted target was hybridized to the microarray at 42°C for 4 or 24 h depending on the requirements of the experiment. Pre-hybridization, hybridization and all washing steps were performed in a-Hyb hybridization station (Miltenyi Biotec, Inc., USA). Printed microarray slides were pre-hybridized with BlockIt (TeleChem, USA) blocking buffer for 1 h at 25°C, and then washed five times with 0.1% sodium sarkosylate at 25°C for 2 min and finally five times with H_2O at 25°C for 2 min. Purified labeled RNA was prepared at various dilutions in 7× PBS buffer with the addition of BSA at a final concentration of 0.2 mg/ml. The hybridization was conducted at 45°C for 4 h or 24 h (depending on the experiment). Unless otherwise specified, the slides were then washed by 2× SSC with 0.1% SDS, two times (2 min each) with a final wash of 0.2× SSC (1 min). For the isopropanol experiments, the slides were first washed with isopropanol (1 min), and an image of the microarray was recorded.

Microarray image acquisition and processing

The VersArray ChipReader 10 μm System scanner (Bio-Rad Laboratories, USA) was calibrated using a microarray scanner calibration slide from Full Moon Biosystems, and optimal PMT settings were chosen for scanning Erie and VWR microarrays. Images were stored as 16-bit TIFF files and processed using ImaGene software (BioDiscovery, Inc., USA). Local background of each spot was not subtracted.

Data analysis

An explicit calibration of the array scanner was performed because there is a nonlinear response between measured signal intensity and the abundance of labeled molecules in a microarray spot (i.e. fluorophore density) (19). Raw signal intensities were converted to actual fluorophore densities using the calibration curve generated with the Full Moon BioSystems calibration slide (Sunnyvale, CA) fitted to a straight line using a weighted regression.

Hybridization isotherms were fitted to linear, or Langmuir model, depending on the shape of the isotherms. The slope of the line for each probe was considered as the equilibrium binding constant, K. Energy predictions based on the existing parameters were conducted using *mfold* (http://dinamelt.bioinfo.rpi.edu/twostate.php).

A system of equations was used to examine the relationship between K and nearest-neighbor (NN) counts of probes and the relationship between fluorescent densities and NN counts at fixed probe concentrations. In matrix form, the system can be re-written as $\mathbf{N} \cdot \mathbf{X} = \mathbf{S}$ where $\mathbf{N}$ is the matrix of the NN counts arranged as column-vectors

(design matrix); **X** is the sought vector of the weights and **S** is the vector of fluorescent densities that corresponds to the NN counts to be solved. This system can be solved analytically by a common method $\mathbf{X} = (\mathbf{N^T} \cdot \mathbf{N})^{-1} \cdot \mathbf{N^T} \cdot \mathbf{S}$ (20). To assess the goodness of fit, one will calculate the R^2 of the numeric solution, which is the squared correlation coefficient between **NX** and **S** vectors, where **X** was determined as a solution of the system. Diagonal elements of the scaled covariance matrix are in fact variances (squared standard deviations) of each solution in the vector **X**. Therefore, these diagonal elements can be used as a measure of an error associated with each solution. Values in the covariance matrix as well as solution are only meaningful if there is a good correlation between **S** and **NX**, where **X** is the computed solution.

RESULTS

All hybridization isotherm experiments were based on pure full-length 28S rRNA-target sequence that was hybridized for 4 h, and then washed using standard stringent washing procedures (see Materials and methods). The reason for using full-length 18S RNA is that most of the physical-chemistry-oriented papers dealing with microarrays try to predict hybridization thermodynamics by taking into account secondary structure of probes and targets, and the secondary structure is usually predicted using *mfold* or some analogous program. We have previously shown that there is no correlation between *mfold* predictions and signal intensity (6). To eliminate the potential problem of secondary structure predictions, we used a molecule that has known secondary structure, which was determined experimentally. We did not fragment this molecule because fragmentation would most likely cause secondary structure rearrangements.

Hybridization isotherms on low-density microarray platforms

Isotherms were generated by hybridizing rRNA target to microarrays at concentrations ranging from 0.15 to 6.83 nM. Both Erie and VWR microarrays contained 95 probes (20 nt) that were spotted at concentrations ranging from 6.25 to 50 μM. Therefore, hybridization isotherms were generated for the same probes spotted at different concentrations.

Figure 1 shows the hybridization isotherms obtained for one of the 95 probes. The left and middle panels represent the isotherms obtained using Erie microarrays, while the right panel represents the isotherms obtained using VWR microarrays. Different rows of panels in Figure 1 represent distinct spotted probe concentrations. Note the difference in the shape of the hybridization isotherms by microarray platform. All probes on Erie microarrays followed a linear hybridization model with high R^2 (the lines look curved due to the log scaling of the *x*-axis). Replication of this experiment using a different batch of Erie microarrays yielded similar results (i.e. compare the left panels to the middle panels). The Pearson correlation (r) of the *K*s for the two Erie microarray experiments was $r = 0.91$, indicating high concordance ($n = 309$ samples based on 95 probes × 4 probe concentrations; ~18% were excluded because the R^2 was ≤ 0.85 for the linear regression between signal and target concentration). For VWR microarrays, the isotherms closely followed the Langmuir adsorption model with high R^2. Comparison of the *K*s for the two different platforms revealed that the *K*s for VWR microarrays (average $K = 0.33$) were more than a 100-fold smaller than those obtained for Erie microarrays (average $K = 58.9$). The *K*s of the probes on VWR microarrays were not correlated to the *K*s of the same probes on Erie microarrays.

Figure 1 and Figure 2 show a general trend that increased spotted probe concentration resulted in increased values of *K* for Probe 596 on Erie microarrays. This trend was consistent for all probes on Erie microarrays (Figure 3, top and middle panels). Although increased spotted probe concentration slightly decreased *K* for VWR microarrays (Figure 2), the slope was close to zero, suggesting that increased probe concentration does not affect the signal for this probe. Similarly, the distribution of *K* values for all VWR probes did not change with the increased probe concentration (Figure 3, lower panel), which follows the Langmuir model. The probe concentration therefore affects the distribution and value of *K* on Erie microarrays, but not VWR microarrays.

Consistent with other studies (e.g. (10)), the horizontal asymptote (P_0) of the hybridization isotherms for the VWR microarray varied for different probes, as reflected in the standard deviation of the mean (449.1 ± 349.5 fluorophores/μm^2, $n = 93$ probes; the data were collapsed by probe concentration). Clearly, some probes saturated at lower than expected signal intensities.

The results demonstrate that the shapes of hybridization isotherms are dependent on the platform and that the concentration of the spotted probes affects the distribution and value of *K* for Erie microarrays, but not VWR microarrays.

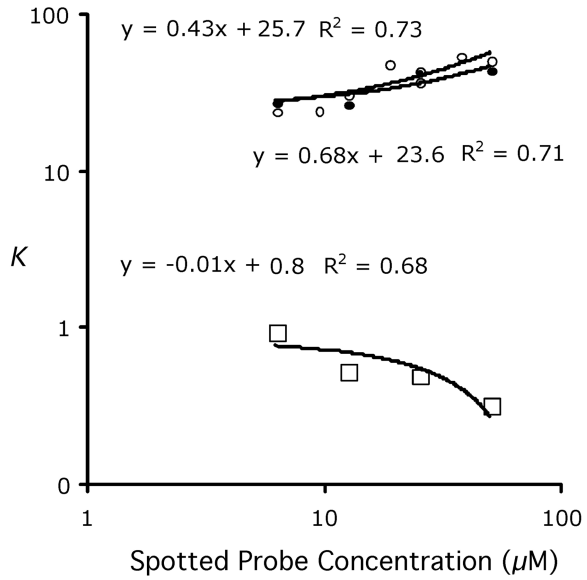

Figure 2. *K* versus spotted probe concentration for Probe 596. Open circles, August experiment, Erie microarrays; Closed circles, April experiment, Erie microarrays; Open squares, VWR microarrays.

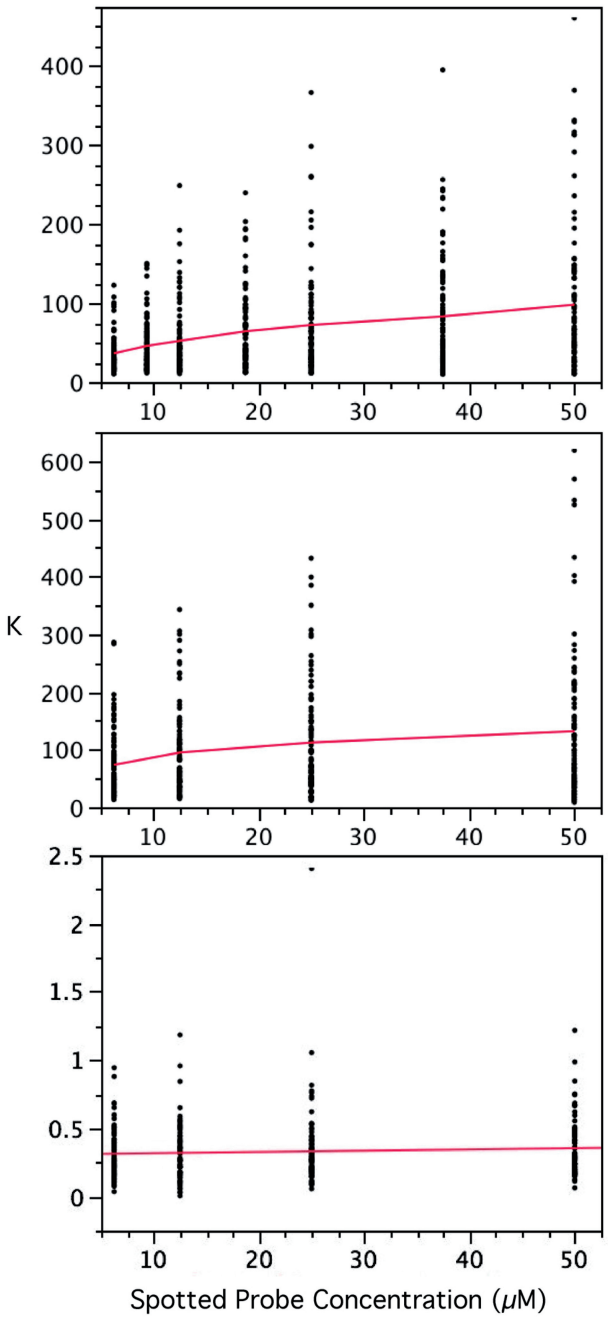

Figure 3. *K* versus spotted probe concentration for the pure rRNA target and the same probes ($n = 95$) but different microarray platforms. Top and middle panels, replicated Erie microarrays; Bottom panel, VWR microarray. Red line shows the general trend based on the mean.

Hybridization isotherms on high-density microarray platforms

In light of the finding that hybridization isotherms are dependent on microarray platform, we investigated the isotherms produced on NimbleGen and CombiMatrix microarrays using a random 70-mer DNA target that had no secondary structure (see Supplementary material for details). We chose this target because there has been many studies published using rRNA and we wanted to eliminate secondary structure effects.

We found that the hybridization isotherms of all probes on NimbleGen microarrays ($n = 3420$) closely followed a Langmuir model with high R^2 (Table S1, left panel), while those on CombiMatrix microarrays ($n = 553$) followed a power-law with high R^2 (Table S1, right panel). These results are consistent with the notion that the two platforms yielded differently shaped hybridization isotherms for the same probes and targets.

Nearest-neighbor model and Gibbs free energies by microarray platform

The *Ks* of the probe hybridization isotherms were used to determine the relationships with nearest-neighbor (NN) counts and Gibbs free energies (ΔG°s). In theory, ΔG°s obtained using the NN model (21,22) and *Ks* obtained from the hybridization isotherms should be highly correlated through the following equation:

$$K = e^{-\frac{\Delta G^0}{RT}}, \quad \textbf{2}$$

where R is the universal gas constant and T is the absolute temperature. According to Equation (3), a decrease in ΔG° should result in an increase of K.

Pearson's correlation analysis of ΔG° and K indicated that these variables were not correlated for either platforms ($r < 0.05$ for Erie and VWR microarrays). We also calculated ΔG°_{total} by first folding rRNA into its secondary structure and then calculating ΔG° for the reaction involving partial rRNA unfolding and binding with the probe (OligoWalk software). In this case, the highest $r = 0.59$ (probe concentration = 37.5 nM, $n = 83$), with the average r for all probe concentrations of 0.28 ($n = 1484$). Therefore, solution-based ΔG° calculations were not sufficient for determining the K of immobilized probes.

To determine microarray-based ΔG°, we applied the same approach used to make the original solution-based NN model (21,22). Specifically, this involved counting the NNs for each probe and then attempting to establish a relationship between NN counts and *Ks* using a system of over-defined equations (see Materials and methods).

The R^2 of the numeric solutions were not sufficiently high to build an accurate microarray-based NN model for either microarray platform (maximum $R^2 \sim 0.42$) (Table S1). We slightly improved the R^2 for Erie microarrays by analyzing each probe concentration individually, rather than examining all probe concentrations at once, or incorporating probe concentration as an additional factor in the system of equations. For the two replicated Erie microarray experiments, the R^2 of the numeric solution improved from 0.36 and 0.42 (for all probes) to 0.53 and 0.52, respectively, for those probes spotted at 50 μM.

These results suggest that the microarray-specific NN model does not sufficiently explain the variability in the data and that other (yet to be defined) factors should be considered to further improve the model, which is part of our ongoing research.

Determination of fluorescent density of immobilized probes at equilibrium

The purpose of treating the microarray with isopropanol immediately after hybridization was to mechanically remove hybridization solution in order to record probe signal at equilibrium. Ethanol could have been used for this purpose as well, but we chose isopropanol over ethanol because preliminary experiments indicated that ethanol resulted in a higher background signal than isopropanol. All isopropanol/stringent washing experiments were conducted on Erie microarrays.

Table 1 shows the distribution of fluorescent densities for 96 perfect match (PM) probes that were hybridized to rRNA target for 4 h. At equilibrium, the fluorescent densities widely varied as reflected by the high standard deviations. Similar results were obtained for repeated experiments (i.e. using another Erie microarray; Table S2). Increased hybridization time (4 h versus 24 h) did not change the distribution about the mean (compare Table S2 to Table S3). These results suggest that signal from probes hybridized to the same target varied in fluorescent density even before stringent washing.

In theory, probe signals at equilibrium should be correlated to ΔG° (21,22). However, no correlation could be established with solution-based ΔG°. These findings reiterate the fact that solution-based ΔG° are not sufficient for predicting signal of immobilized probes. It should be noted that we also tried to build a microarray-based NN model, similar to the one above. Of the seven different probe concentrations considered, the best linear regression was obtained for probes at 25 μM concentration ($R^2 = 0.37$; $n = 215$ probes). This low R^2 value suggests that, even at equilibrium, our NN model does not adequately explain the signal produced from a single target hybridized to microarray probes.

Stringent wash on microarray probe signal

To monitor the change of probe signal as a result of stringent washing, the fluorescent densities of the probes were recorded at equilibrium, the microarrays were washed with a stringent buffer and the probe fluorescent densities were rerecorded. Visual inspection of the microarrays revealed that more fluorescent debris was associated with the microarray at equilibrium than after a stringent wash (Figure 4). Stringent washing therefore dissolves fluorescent debris on the microarray surface as well as removes some of the loosely bound target. From this observation, stringent washing might also remove bound target due to the establishment of nonequilibrium conditions.

To imitate actual hybridizations of biological samples, we added nontarget sequences (i.e. tRNAs), which could interfere with target–probe binding. The reason we chose prokaryotic tRNAs as nontarget sequences rather than other types of sequences (e.g. eukaryotic rRNAs or tRNAs) was because the pure target was of eukaryotic origin, and having both an eukaryotic target and nontargets might impart a bias to the experimental results. Although prokaryotic tRNAs differ in terms of their nucleotide composition, they are uniform in terms of their size (74–95 nucleotides). As our study will show (below), the prokaryotic tRNA sequences were ideal for investigating the interactions of immobilized probes with nontarget sequences because the tRNAs *did* cross-hybridize with probes, which is consistent with another study that showed cross-hybridization of nontargets to immobilized probes cannot be predicted (23).

Figure 5 is a composite of washing experiments; each panel in the composite represents a comparison of probe signals recorded at equilibrium to those recorded after stringent washing. Three variables were studied in our experiments: (i) hybridization time (4 h versus 24 h), (ii) signals recorded before versus after stringent washing and (iii) presence versus absence of nontargets. Table 2 is a summary of these experiments.

For pure rRNA target, hybridization time had a significant effect on the extent of signal intensity reduction after stringent washing. Specifically, 51–54% of the target was washed off the 4 h hybridized microarrays, while only 24–34% of the target was washed off the 24 h hybridized microarrays. This finding suggests that increased hybridization time resulted in fewer targets washed off the microarray.

For the mixture of rRNA and nontargets (tRNA), hybridization time had a similar effect on a mixture for the microarrays hybridized for 4 h, with 49–56%. However, no target was washed off the 24 h hybridized microarrays. These findings are consistent with the notion that increased hybridization time resulted in more binding of targets to probes (compare Tables S2 and S3), which might be expected given the size of the rRNA target used in this study (1919 nt). Presumably, portions of the bound target (e.g. tails) that are far away from the probe–target binding site serve as an enthalpic barrier, 'entrapping' target and nontarget sequences and therefore minimizing the removal of loosely bound target and nontargets during the stringent wash.

It should be noted that the slope recorded for the first stringent wash for the 24 h hybridized target with nontarget was >1 (Figure 5, Panels in row D). This finding suggests that probe signal was greater after stringent washing than before and was confirmed in a replicated experiment.

Table 1. Distribution of fluorescence density (F_d, fluorophores/μm^2) at equilibrium by probe concentration on Erie microarrays hybridized for 4 h. Same probes and target were used for each probe concentration

Probe concentration (μM)	Mean F_d	± SD	Number of probes
6.25	76	76	96
8.75	92	70	96
12.5	83	61	96
17.5	108	90	96
25	105	85	96
35	109	90	96
50	106	85	96

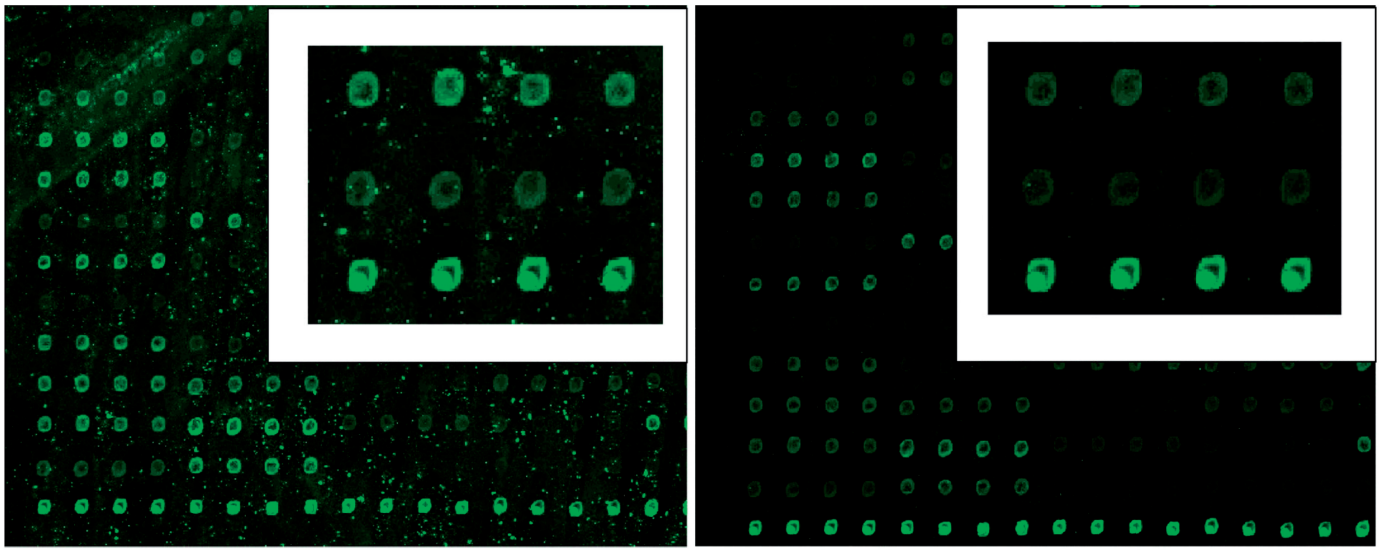

Figure 4. Fluorescent signal before (left panel) and after stringent wash (right panel). Left inset is a zoomed image of 12 probes in lower left corner of left panel. Right inset is a zoomed image of 12 probes in lower left corner of right panel. Note that the image collected before stringent wash contains more fluorescent particles than the image collected after stringent wash.

In summary, hybridization time and the presence/absence of nontarget sequences had a substantial influence on probe signal recorded after stringent washing.

Effects of the probe concentration on signal after stringent washing

To determine if the effects of stringent washing are dependent on the spotted probe concentration, we compared the slopes of the regression lines of signal intensities recorded before and after washing with pure targets.

Table 3 reveals a general trend that probes spotted at higher concentrations (e.g. 35 μM) yielded higher slopes (e.g. $m = 0.72$) than those spotted at lower concentrations (e.g. 6.25–25 μM; $m = 0.51$–0.63). However, the highest spotted probe concentration (i.e. 50 μM) did not yield the highest slope, as one might have expected, presumably due to steric hindrances (24–27).

Effects of probe sequence on signal after stringent wash

Based on the slopes of the linear regressions for the pure rRNA target (Figure 5, Panels in rows A and C), one might erroneously conclude that the fluorescent densities change after stringent wash in a consistent manner for all probes. In fact, when each of the 96 probes was individually examined to determine possible probe sequence effects, no obvious pattern could be resolved for two-thirds of the probes (i.e. $n = 63$). It is important to note that these probes had both high and low signal intensities, which rules out 'noise' as the sole source of variation. However, the remaining third of probes had slopes that were ≥ 0.20 with $R^2 > 0.67$ (Table 4). Figure 6 shows the distribution of the fluorescent densities, the equations and the slopes for three of the 33 probes listed in Table 4. Examination of the probe sequences, their corresponding Gibbs free binding energies, the number of points used to make the regression line and the maximum fluorescent density value for each regression revealed that no one factor could explain why these 33 probes were different from the other 63 probes. In summary, some probes (~34%) were more affected by stringent washing than others.

DISCUSSION

Prior to this study, the number of reported hybridization isotherms was very small (28), particularly with regard to isotherms on microarray platforms other than Affymetrix. Probe hybridization isotherms are important to microarray studies because, once an isotherm has been established for a probe, target concentrations can be estimated from signal intensity. If isotherms were similar, or at least partially correlated, for the same probe–target duplexes on different platforms, cross-platform analysis of microarray results would be possible. Hence, the first question we sought to address was 'are hybridization isotherms similar for the same target hybridized to the same probe on different platforms?'.

Also, prior to this study, the effects of stringent washing on probe signal were not well known. By determining probe–target signals at equilibrium, we reasoned that it might be possible to develop a model that compensates for stringent washing effects. The development of such a model could ultimately improve the estimation of target abundances. Hence, the second question we sought to address was 'is it possible to model the effects of stringent washes on probe–target duplexes?'.

Since this study involved several platforms and many different experiments, Figure 7 provides an overview of experiments conducted and results obtained.

Hybridization isotherms and microarray platforms

The probe hybridization isotherms on VWR microarrays followed the Langmuir adsorption model (e.g. Figure 1,

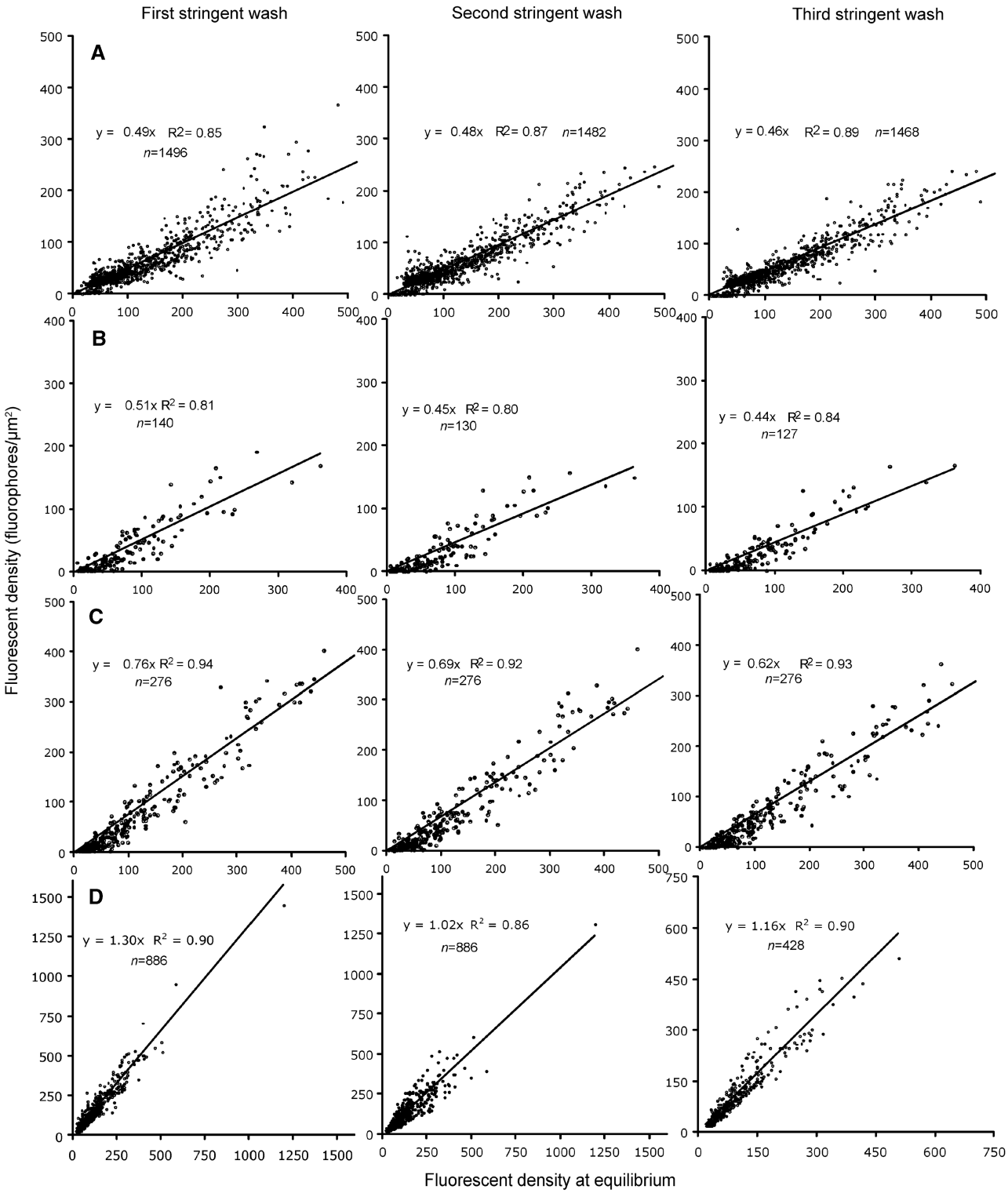

Figure 5. The effect of successive stringent washes on probe fluorescent densities recorded at equilibrium. Samples were hybridized for 4 h (Panels in rows A and B) or 24 h (Panels C and D in rows) with active mixing and were either pure rRNA (Panels in rows A and C) or rRNA mixed with unlabeled tRNA (Panels in rows B and D). These results are based on at least two independent microarray experiments. Note that the number of points is not the same for different washes because converting signal intensity to fluorescent density sometimes generated numbers that were nonphysical.

Table 2. Summary of slopes of probe signal before and after stringent washing of pure rRNA target and mixed targets with unlabeled tRNA (+tRNA)

Hybridization		First wash			Second wash			Third wash		
Time	Target	m	N	R^2	m	n	R^2	m	n	R^2
4 h	Pure	0.49	1496	0.85	0.48	1482	0.87	0.46	1458	0.89
4 h	+tRNA	0.51	140	0.81	0.45	130	0.80	0.44	127	0.84
24 h	Pure	0.76	276	0.94	0.69	276	0.92	0.62	276	0.93
24 h	+tRNA	1.30	886	0.90	1.02	886	0.86	1.16	428	0.90

Data were taken from Figure 5.

right panels), which is consistent with previous Affymetrix studies (9,11–15). However, probe hybridization isotherms on Erie microarray platforms were linear in shape (Figure 1, left and middle panels) and could not be accurately modeled using the Langmuir adsorption equation. Moreover, the isotherms from Erie microarrays were affected by probe concentration (see Figure 2 and Figure 3), which violates the Langmuir adsorption model (28). These findings suggest that the underlying assumptions of hybridization isotherms are not justified, especially in terms of probe saturation, which will become evident below. A possible explanation for the isotherms not following the Langmuir curve is that as more target is bound, the energetics for hybridization changes, which goes against the basic assumption of the Langmuir model.

Since the microarray experiments, for our first question posited, were performed using the same target, probes, buffers and scanner, the only factor that could possibly explain the difference in the shape of the hybridization isotherms was the microarray platforms themselves. Potential differences in the platforms include: the composition of synthesized probes and linker molecules that attach probes to the microarray surface, and the electrostatic charge of the microarray surface. We ruled out differences in the composition of synthesized probes and linker molecules because the same synthesized probes and linkers were used to attach to the epoxysilane surface of the glass microarrays. We also ruled out the way the probes were attached because the same arrayer, buffer conditions and probe concentrations were used.

The electrostatic interactions on the surface of a microarray could be affected by the way the microarray was treated during the manufacturing process. An examination of the manufacturer's information sheet accompanying one of the microarrays revealed a disparity in the way the glass surfaces were treated. Specifically, the epoxy moiety was attached to the surface of Erie microarrays using a vapor-coating process, which yields a uniform surface free of impurities. Other microarray manufacturers (e.g. VWR) attach the epoxy moiety to the microarray by dipping the slide in a 1% solution of epoxysilane for various lengths of time. It is possible that the epoxy layer afforded by the vapor-coating process affects the electrostatic interactions on a microarray surface by screening the charge associated with the glass, while the 'dipped' slides are less effective at screening this charge. Hence, the way the surfaces of the microarrays were treated prior to attachment of the probes presumably explains the variation in shape of the hybridization isotherms and the differences in K for the same probe–target duplexes (Figure 1).

Electrostatic interactions and hybridization isotherms

It is widely believed that the hyperbolic shape of hybridization isotherms is due to probe–target saturation (8,11) and that the lack of agreement in the height of the horizontal asymptote for different probes is the result of stringent washing, insufficient hybridization time and/or the presence of nontarget sequences, which sterically hinder the hybridization of specific targets (10). Results presented in this study do *not* support the idea that the hyperbolic shape is due to probe–target saturation because neither Erie nor CombiMatrix microarrays produced a horizontal asymptote at high target abundance. One could argue that the target used for these experiments was not sufficiently abundant to produce a horizontal asymptote. But, the same target abundances were used for both VWR and Erie microarray experiments and the VWR microarrays produced an asymptote while the Erie microarrays did not. Similarly, when the same target abundances were used for both NimbleGen and CombiMatrix microarrays, the NimbleGen microarrays produced a horizontal asymptote while the CombiMatrix microarrays did not. Therefore, if the shape of the hybridization isotherms was not due to probe saturation *what then explains the differences in the shapes?*

It can be argued that the variation in isotherm shapes could be due to different quantities of probes per spot. We ruled out probe–density effects because variations in probe concentrations of Erie microarrays did not yield differently shaped isotherms. For the high-density microarrays (CombiMatrix and NimbleGen), one is less sure about the density of probes directly synthesized on the microarray surface because this information is proprietary. This synthesis leads us to the following working hypothesis: electrostatic interactions associated with the microarray surface, at least partially, explain the difference in the shapes of hybridization isotherms among the platforms.

Electrostatic interactions play a key role in probe–target hybridization because nucleic acid targets, the ions in solution and the microarray surface have an electrostatic

charge. In a neutral-pH buffer, all nucleic acid molecules have a negatively charged phosphate backbone that is surrounded by a cloud of counter-ions (e.g. cations such as Na^+) (29). The density of this counter-ion cloud plays an important role in hybridization because it reduces the electrostatic repulsion between a probe and a target, which allows hydrogen bonding to take place and subsequent formation of a duplex. For glass microarrays, the counter-ions in solution also balance the charge on the glass surface (30). In a 1.0 M salt buffer used for hybridization, for example, the glass surface of a microarray has a charge of −35 mV (30). But when the same microarray is washed with a 0.1 M salt buffer used for stringent washing, the glass surface has a charge of −80 mV (30). The presumed reason for this difference is that the 0.1 M salt buffer contains fewer cations to balance the charge on the glass surface of the microarray. In the case of CombiMatrix microarrays, target molecules are not affected by the surface because it is composed of polysaccharide, which has neutral charge (published patent information) and is estimated to be 1–5 monolayers thick, or 1–5 nm (10–50 Å) (CombiMatrix, personal communication).

We have emphasized electrostatic interactions on the surface of the microarrays because all hybridization isotherms generated in our study were obtained after stringent washing and microarrays with charged surfaces would have been more affected by stringent washing

Table 3. Relationship of probe concentration and the linear regressions of fluorescent density (F_d, fluorophores/μm^2) before and after stringent wash on Erie microarrays

Probe concentration (μM)	*m*	*b*	R^2	Number of probes
6.25	0.54	−2.4	0.79	228
8.75	0.52	−0.2	0.87	231
12.5	0.51	0.3	0.78	235
17.5	0.55	−2.4	0.81	256
25	0.63	−8.1	0.85	262
35	0.72	−17.9	0.87	279
50	0.61	−9.1	0.86	281
All	0.62	−8.3	0.85	1772

Each linear regression is shown in the standard $y = mx + b$ format, with m representing the slope, x representing F_d before stringent wash, y representing the predicted F_d after the first stringent wash and b representing the intercept. In general, the slope increases with the increased spotted probe concentration. Based on two replicated experiments.

Table 4. Probes yielding linear relationships in fluorescent density (F_d, fluorophores/μm^2) before and after stringent wash with $R^2 > 0.66$

Probe number	Sequence	*m*	*B*	*n*	R^2	Max F_d	ΔG°
1366	ACCAACTAAGAACGGCCATG	0.20	14.70	14	0.68	140	−32.1
1373	TCGCTCCACCAACTAAGAAC	0.33	3.80	20	0.67	130	−32.8
123	TTCACTGTACAGCCCGTGTG	0.34	13.30	27	0.68	200	−34.4
1513	GCATCACAGACCTGTTATTG	0.36	10.60	20	0.68	225	−30.2
1288	GGCCGGGTGAGGTTTCCCGT	0.38	9.70	14	0.76	160	−39.8
1656	GAATTCCTCGTTCATGGGAA	0.39	0.60	24	0.78	175	−31.0
262	GGTTTTGGATCTGATAAATG	0.40	4.90	14	0.84	100	−26.1
1507	CAGACCTGTTATTGCTCAAT	0.45	7.20	24	0.85	200	−28.9
1106	GCATCGTTTATGGTCGGAAC	0.46	−4.10	25	0.8	250	−31.5
978	TCACCTCTAGCGGCACAATA	0.47	1.00	23	0.86	200	−33.8
639	GGCTGCTGGCACCAGACTTG	0.48	−4.90	16	0.82	250	−38.0
1839	GACTTTTACTTCCTCTAGAT	0.49	−11.20	30	0.81	350	−27.4
891	TCCTATTCCATTATTCCTAG	0.50	−3.20	21	0.71	110	−27.1
1667	GCGCTTACTGGGAATTCCTC	0.51	0.20	17	0.89	80	−34.0
645	TACCGCGGCTGCTGGCACCA	0.52	−4.20	18	0.92	200	−41.0
1073	GATCGTCTTCGAACCTCCGA	0.52	4.90	28	0.78	400	−34.1
1214	CCTTTAAGTTTCAGCTTTGC	0.54	−10.60	26	0.92	275	−27.8
512	AGTGGGTAATTTGCGCGCCT	0.54	−1.10	14	0.69	75	−34.7
235	CCCCGGGGGTCAGCGCTCGT	0.57	−6.50	15	0.82	150	−43.0
636	TGCTGGCACCAGACTTGCCC	0.58	−7.10	24	0.91	200	−39.2
953	GCCCCCGGCCGTCCCTCTTA	0.59	8.50	28	0.67	225	−42.0
1006	TCGTCCGTCTTGCGCCGGTC	0.60	4.80	32	0.82	400	−39.3
537	TTCGTCACTACCTCCCCGAG	0.61	−0.50	39	0.75	450	−36.0
1531	AGCCCCGGACATCTAAGGGC	0.63	−11.60	16	0.76	100	−38.6
1059	CTCCGACTTTCGTTCTTGAT	0.64	−27.10	37	0.87	500	−29.9
1067	CTTCGAACCTCCGACTTTCG	0.65	−9.30	35	0.83	400	−32.1
1235	GGTGGTGCCCTTCCGTCAAT	0.65	8.60	33	0.78	450	−36.5
879	ATTCCTAGCTGCGGTATTCA	0.68	−22.30	37	0.87	600	−31.7
1850	AACCTTGTTACGACTTTTAC	0.70	−47.80	36	0.9	900	−26.4
1510	TCACAGACCTGTTATTGCTC	0.73	−32.80	38	0.91	500	−31.5
494	CTGCTGCCTTCCTTGGATGT	0.78	−25.80	39	0.78	700	−34.8
1076	TCTGATCGTCTTCGAACCTC	0.82	−28.70	34	0.78	350	−32.7
1228	CCCTTCCGTCAATTCCTTTA	0.85	−55.70	39	0.88	750	−30.2

Each linear regression is shown in the standard $y = mx + b$ format as explained in Table 3. Max F_d, maximum F_d in the linear regression; ΔG°, Gibbs free energy of binding; probe number reflects the position the probe relative to the target. The table is ordered by slope. No relationship could be established between position of probe relative to the target, the slope, the number of data points (n) used for the regression line, the max F_d or the ΔG° value.

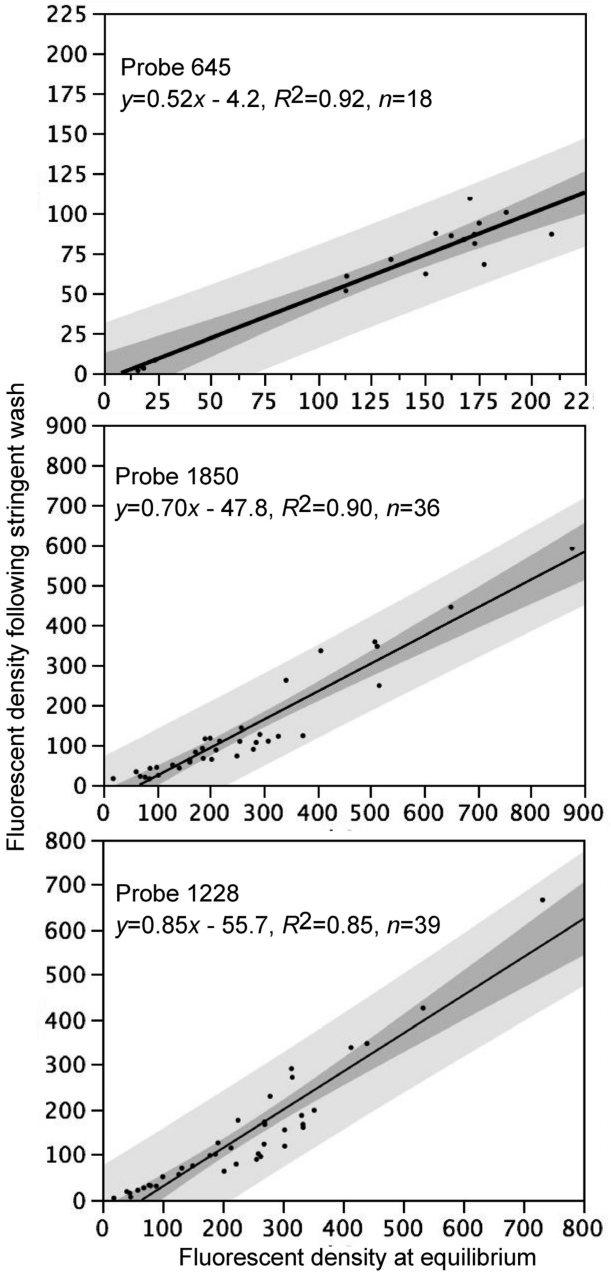

Figure 6. Three probes from Table 4 that are affected by stringent washing. Note the difference in the slopes of the three probes. More targets were washed off Probe 645 ($m = 0.52$) than Probe 1228 ($m = 0.85$). Shaded regions represent the 99% individual and slope confidence levels. Fluorescent density is in fluorophores/μm^2.

than microarrays with a neutrally charged surface. One would expect more loosely bound target to be washed off a negatively charged surface than a neutral-charged surface because the negatively charged surface de-stabilizes and repels nucleic acids away from the surface in a low salt buffer (30,31) while the neutrally charged microarray does not. Therefore, although the exact physicochemical mechanisms have not been worked out and are part of an ongoing study, it is the differences in the electrostatic interactions on the microarray surface that presumably explain the variation in shapes of hybridization isotherms on different microarray platforms.

We suggest that the difference in the shape of isotherms from Erie and VWR microarrays, which both have glass surfaces, is due to the way the surfaces are treated prior to probe spotting. It has been shown that solution versus vapor deposition has a significant effect on the morphology of silane molecules on the glass surface (32 and references therein). Apparently, the vapor coating of the Erie microarrays yields a different texture of an epoxysilane layer than that on VWR microarrays.

In the case of CombiMatrix microarrays, the electrostatic effects on the microarray surface are negated because of the surface has a neutral charge. The neutral charge surface explains the reason why hybridization isotherms on CombiMatrix microarrays are so different in shape compared to those produced on other platforms.

In summary, and in answer to the first question posited ('are hybridization isotherms similar for the same target hybridized to the same probe on different platforms?'), hybridization isotherms of the same probe–target duplexes vary in shape on different microarray platforms due to electrostatic interactions on the surface of the microarrays. VWR, NimbleGen and Affymetrix microarrays produced hyperbolic isotherms because these microarrays have negatively charged glass surfaces, which increase electrostatic interactions relative to neutrally charged surfaces. Erie microarrays produce linear-shaped hybridization isotherms because the epoxy layer presumably screens the negative charge associated with the glass surface. CombiMatrix microarrays yield isotherms that follow a power-law because the surface is neutrally charged. Isotherms generated from CombiMatrix microarrays are similar to those from Erie microarrays because they do not produce a horizontal asymptote, indicative of probe saturation. Certainly more research is needed to physically investigate the differences between the microarray surfaces.

Nearest-neighbor model for microarrays

There is some confusion in the literature regarding NN models and the ability to accurately predict signal on an oligonucleotide microarray. It has recently been reported that thermodynamic behavior of short oligonucleotide microarrays can be described using the NN model (33). This is only true for a few microarray platforms that are composed of three-dimensional matrices (34; e.g. gel-pad, gel-drop or CodeLink microarrays) because the oligonucleotide probes are not attached to any surface and behave like probes in solution. For most microarrays, i.e. those with surface-attached oligonucleotides, there is no correlation between ΔG° and signal (6), which is consistent with the results found in this study.

Given the lack of correlation between ΔG° and signal, we attempted to develop a microarray-based NN method. By relating the K of probe hybridization isotherms to NN parameters for both PM and MM probe–target duplexes, we were able to account for up to 53% of the variability in the model (Table S1). However, the low to moderate R^2 of

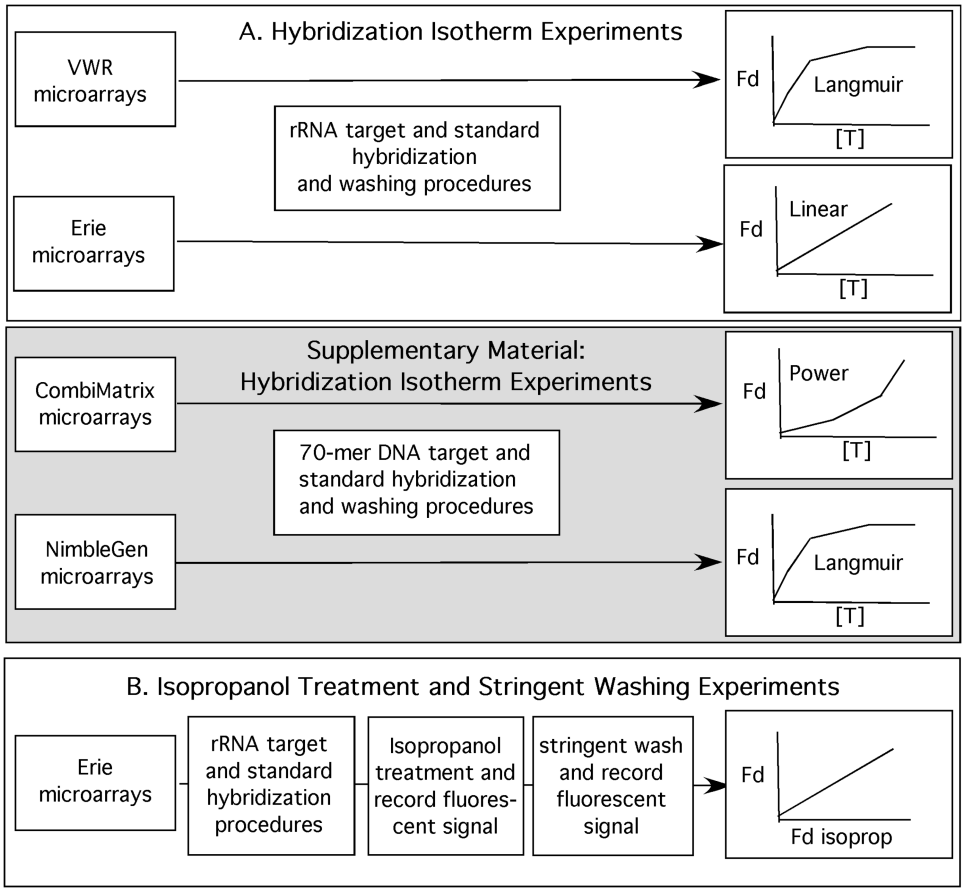

Figure 7. Schematic representation of the study. (**A**) Hybridization isotherms experiments, which include isotherms in the Supplementary Data. (**B**) Isopropanol treatment experiments and effects of stringent washing.

the regression line suggested that the model does not adequately account for all factors affecting *K*. One factor that might have contributed to the variability of *K* was that all microarrays were stringently washed. To test if stringent washing contributed to this variability, we developed the isopropanol treatment approach, which provides a 'snapshot' of signal intensities attained at equilibrium. We found that neither the solution-based NN model nor our own microarray-based NN model sufficiently explained the variation of signal intensities before (using isopropanol treatment) or after stringent wash.

Stringent washing and probe–target equilibrium

The purpose of the stringent wash is to improve the signal-to-noise ratio by washing away unbound labeled target while minimizing removal of PM targets. The procedure typically involves rinsing the microarray with the low-salt buffer at constant temperature. Once the stringent wash has been performed, it is believed that nontarget duplexes are washed away and the observed signal represents the 'true' signal of the specific target duplexes. This belief is not supported in this study or in previous studies (35,36). In addition, it appears that the probes were not saturated even before stringent washing (i.e. at equilibrium), suggesting that stringent washing alone is not the cause of this phenomenon.

We examined the effects of stringent washing on Erie microarrays rather than VWR microarrays because Erie microarrays yielded linear hybridization isotherms with high R^2 and the *K*s were shown to be consistent for probes in replicated experiments, as shown in Figure 1 (left and middle panels). In order to perform clean experiments, only one rRNA target (as opposed to multiple targets) was used in the first set of experiments. To investigate the effects of nontargets on stringent washing, tRNAs was included with the rRNA target in a second set of experiments. In both sets of experiments, hybridization time was varied from 4 h to 24 h in order to generate different populations of loosely bound and (completely) bound targets. The reason 4 h hybridization time was chosen as opposed to some other time is that 4 h was proven to be sufficient by the manufacturer of our active-mixing hybridization station (aHyb, Miltenyi Biotec) and also this amount of time was used in our previous microbiology experiments (5). We anticipated that the signal of 4 h hybridized microarrays to be composed of both loosely bound and bound targets and that of 24 h hybridized microarrays to be mostly composed of bound targets.

An important finding of our study was that stringent washing distorts the 'true' signal of specific probe–target duplexes as reflected in the inconsistent differences of

probe fluorescent densities before and after stringent wash (Figure 5). Apparently, the distortion varies by hybridization time (Table 2), probe concentration (Table 3) and probe sequence (Table 4). Hybridization time affected the amount of bound target retained on the microarray after the stringent wash. Extended hybridization times resulted in more bound target and less target removed from the microarray due to stringent washing. This finding is consistent with the model described by Skvortsov *et al*. (10). Stringent washing was found to differentially affect probe signal as a function of probe concentration, which was not previously recognized, nor particularly well understood. Stringent washing was also found to differentially effect signal as a function of probe sequence, which has been reported in previous studies (35,36).

In answer to the second question posited ('is it possible to model the effects of stringent washes on probe–target duplexes?'), it is very challenging to model the effects of stringent washing because stringent washing affects probes in many different ways (e.g. hybridization time, probe sequence, probe concentration) that are poorly understood in terms of physicochemistry. Clearly, more systematic studies are needed to better understand the physicochemistry of stringent washing on immobilized probes.

Isopropanol treatment as an alternative for stringent washing

We have initiated our exploration of the pre-wash signal intensities using the Erie microarrays because (i) they yielded linear isotherms, which were easy to interpret and (ii) washing experiments required a very high number of arrays, which means that other platforms like Combimatrix or NimbleGen were too expensive to try at that stage of our research.

Isopropanol treatment appears to be a robust approach for removing hybridization solution while not effecting bound targets because controlled experiments that repeatedly treated the microarrays with isopropanol revealed that signal from bound pure target did not change substantially with repeated treatments (i.e. the data closely followed the line of equality, Figure S2).

An advantage of isopropanol treatment over stringent washing is that isopropanol provides a snapshot of probe–target duplexes at equilibrium. In contrast to stringent washing, isopropanol prevents the dissociation of nucleic acids from taking place, whereas a stringent wash is a definitive nonequilibrium process and therefore has an additional effect on nucleic acid targets. Future studies are needed to examine if other solvents (e.g. acetone) have more desirable properties for removing unbound targets than isopropanol.

Probe signal variation at equilibrium

A key finding of this study is that, at equilibrium, fluorescent densities varied for microarray probes hybridized to the same pure target. Previously, it was thought that all microarray probes should yield a constant signal at equilibrium because the probes saturate in the presence of excess target in solution (11). Apparently, excess target was not enough to saturate all probes in the same way, as shown by the high standard deviation of the mean fluorescent density in Tables 1, S2 and S3. As mentioned in the discussion of hybridization isotherms, probe signals vary for the same target at equilibrium, presumably due to electrostatic interactions.

Other factors potentially contributing to the variation in probe signals at equilibrium include the same factors affecting hybridization isotherms (e.g. 10,24,25,27,28, 37–40). In this study, we did not investigate which of these factors contributed to the variation in probe signal at equilibrium because that would go beyond our stated objectives.

Final remarks

Our finding that there was no obvious relationship between solution-based ΔG and measured ΔG, or between solution-based ΔG and fluorescent density ($\sim$signal intensity), seems to contradict other studies (31,41). For example, Poulsen *et al*. reported [in Figure 2C in (31)] that the best R^2 for the linear regression between solution-based ΔG and signal intensity was 0.58, which means that $<60\%$ of the variability was explained. This result is consistent with the present study, as well as the Pozhitkov *et al*. (6) study, where one (i.e. *Plectus minimus*) of nine rRNA targets investigated yielded $R^2 = 0.40$ (see Table 4 in (6)) (note that the other eight rRNA targets yielded lower R^2). Another example is the Wei *et al*. study, which also showed a trend in the relationship between ΔG and signal intensity [Figure 1 of (41)]. This trend only appears when a very large set of probes was examined and when the probe signal intensities were binned. Close examination of the spread of the data revealed large deviations about the average values. Hence, in reviewing the studies of Poulsen *et al*., Pozhitkov *et al*., Wei *et al*. and this study, one is left to conclude that the theoretical foundation for the design of microarray oligonucleotide probes using ΔG is nowhere near that of other instruments implemented in other fields (e.g. the Nernst equation perfectly describes the behavior of a pH probe or any other ion-selective probe). Therefore, studies advocating the utilization of ΔG as a predictive measure of signal intensity of an immobilized probe should be viewed with caution.

In light of the findings in this study, there is yet another belief in the microarray field that needs to be addressed: that 'stringency conditions are very important for attaining a good correlation between ΔG and signal intensity' (31). From a theoretical point of view, ΔG characterizes the equilibrium attained between targets and a probe. On the other hand, stringent washing is a fundamentally nonequilibrium process used to wash off unbound target. Therefore, the argument that stringency (a nonequilibrium process) is important to obtaining a good correlation between ΔG and signal intensity (an equilibrium process) is 'contradictory'. In retrospect, stringency conditions have no relevance for determining the correlation between ΔG and signal intensity.

We have emphasized the distinction between equilibrium and nonequilibrium processes because potentially, an optimally stringent buffer can remove nonspecifically bound targets, which empirically might result in a better correlation between ΔG and signal intensity for solutions containing nontargets. Since the section in our study dealing with ΔG and signal intensity (i.e. fluorescent density) involved pure targets only, the argument that our experiments were not conducted using the 'correct' stringency is irrelevant because nonspecific targets were not present.

To further demonstrate the point, we investigated two different post-hybridization conditions (i.e. one condition examined signal at equilibrium (isopropanol treatment) and the other condition examined the signal following nonequilibrium (stringent wash using $0.2 \times$ SSC—which is a standard microarray buffer). Regardless of the post-hybridization condition, we found a poor relationship between predicted and directly measured ΔG and between signal intensity and directly measured ΔG. These results bring us back to the notion that theoretical predictions have little merit for designing immobilized probes.

The take-home message of this study is that there is a considerable misinformation in the literature, and reiterations of studies that mix theoretical predictions with empirical probe optimizations (which are often highly cited) further add confusion to microarray research. Also, there is a multitude of papers reanalyzing the decade-old Affymetrix datasets, which brought about the notion that all microarray hybridizations follow a Langmuir isotherm. We assert that in order to more fully understand hybridization (an equilibrium process) and dissociation (a nonequilibrium process) on oligonucleotide microarrays, the fundamentals of the microarray physicochemistry need to be established and that these fundamentals must be coupled with experimental results on different microarray platforms.

CONCLUSION

Apparently, the shape of hybridization isotherms was determined by the microarray platform. Our study discovered that the Langmuir adsorption model is not the only model to explain microarray hybridizations. We also discovered that probes do not always saturate at equilibrium, even in excess of target in solution.

Hybridization time had a substantial effect on the amount of target washed off during stringent washes, with fewer targets being washed off with extended hybridization times. An alternative to stringent washing is to treat the microarray with isopropanol. By performing the isopropanol treatment, the unbound targets are removed from the microarray surface, which simplifies microarray processing because the problems associated with stringent washes (i.e. nonequilibrium conditions) are eliminated and washing is no longer necessary.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Michael Lodes and CombiMatrix Corp. who generously provided the microarrays and performed microarray experiments, and Shuzhao Li who performed preliminary experiments. They also thank Rebecca A. Rule for her critical comments on an earlier version of the manuscript and to the support provided by Diethard Tautz and Monika A. Leutenegger.

FUNDING

Funding for open access charge: National Science Foundation, CREST (HRD#0734232); Deutsche Forschungsgemeinschaft (DFG Ta99/21-2); National Oceanic and Atmospheric Administration, Coastal Ocean Program (NA05NOS4261163, NA06NOS42600 117); University of Washington Royalty Research Fund; University of Washington Provost Bridge Funding Program (to P.A.N.).

Conflict of interest statement. None declared.

REFERENCES

1. Brown,P.O. and Botstein,D. (1999) Exploring the new world of the genome with DNA microarrays. *Nat. Genet.*, **21**, 33–37.
2. Lipshutz,R.J., Fodor,S.P., Gingeras,T.R. and Lockhart,D.J. (1999) High-density synthetic oligonucleotide arrays. *Nat. Genet.*, **21**, 20–24.
3. Peplies,J., Lachmund,C., Glockner,F.O. and Manz,W. (2006) A DNA microarray platform based on direct detection of rRNA for characterization of freshwater sediment-related prokaryotic communities. *Appl. Environ. Microbiol.*, **72**, 4829–4838.
4. Pozhitkov,A., Bailey,K.D. and Noble,P.A. (2007) Development of a statistically robust quantification method for microorganisms in mixtures using oligonucleotide microarrays. *J. Microbiol. Methods*, **70**, 292–300.
5. Pozhitkov,A.E., Nies,G., Kleinhenz,B., Tautz,D. and Noble,P.A. (2008) Simultaneous quantification of multiple nucleic acids in target mixtures using high density microarrays. *J. Microbiol. Methods*, **75**, 92–102.
6. Pozhitkov,A., Noble,P.A., Domazet-Loso,T., Staehler,P., Beier,M. and Tautz,D. (2006) Tests of rRNA hybridization to microarrays suggest that hybridization characteristics of oligonucleotide probes for species discrimination cannot be predicted. *Nucleic Acids Res.*, **34**, e66.
7. Pozhitkov,A., Tautz,D. and Noble,P.A. (2007) Oligonucleotide arrays: widely applied–poorly understood. *Brief. Funct. Genom. Proteom.*, **6**, 141–148.
8. Burden,C.J., Pittelkow,Y. and Wilson,S.R. (2006) Absorption models of hybridization and post-hybridization behavior on oligonucleotide microarrays. *J. Phys. Condens. Matter*, **18**, 5545–5565.
9. Hekstra,D., Taussig,A.R., Magnasco,M. and Naef,F. (2003) Absolute mRNA concentrations from sequence-specific calibration of oligonucleotide arrays. *Nucleic Acids Res.*, **31**, 1962–1968.
10. Skvortsov,D., Abdueva,D., Curtis,C., Schaub,B. and Tavare,S. (2007) Explaining differences in saturation levels for Affymetrix GeneChip arrays. *Nucleic Acids Res.*, **35**, 4154–4163.
11. Held,G.A., Grinstein,G. and Tu,Y. (2006) Relationship between gene expression and observed intensities in DNA microarrays—a modeling study. *Nucleic Acid Res.*, **34**, e70.
12. Li,S., Pozhitkov,A. and Brouwer,M. (2008) A competitive hybridization model predicts signal intensity on high density DNA microarrays. *Nucleic Acids Res.*, **20**, 6585–6591.

13. Burden,C.J. (2008) Understanding the physics of oligonucleotide microarrays: the Affymetrix spike-in data reanalysed. *Phys. Biol.*, **5**, 16004.
14. Held,G.A., Grinstein,G. and Tu,Y. (2003) Modeling of DNA microarray data by using physical properties of hybridization. *Proc. Natl Acad. Sci. USA*, **100**, 7575–7580.
15. Carlon,E. and Heim,T. (2006) Thermodynamics of RNA/DNA hybridization in high-density oligonucleotide microarrays. *Physica A*, **362**, 433–449.
16. Zhang,L., Miles,M.F. and Aldape,K.D. (2003) A model of molecular interactions on short oligonucleotide microarrays. *Nat. Biotechnol.*, **21**, 818–821.
17. Mei,R.E., Hubbell,E., Bekiranov,S., Mittmann,M., Christians,F.C., Shen,M.M., Lu,G., Fang,J., Liu,W.M., Ryder,T. *et al.* (2003) Probe selection for high-density oligonucleotide arrays. *Proc. Natl Acad. Sci. USA*, **100**, 11237–11242.
18. Ono,N., Suzuki,S., Furusawa,C., Agata,T., Kashiwagi,A., Shimizu,H. and Yomo,T. (2008) An improved physicochemical model of hybridization on high density oligonucleotide microarrays. *Bioinformatics*, **24**, 1278–1285.
19. Shi,L., Tong,W., Su,Z., Han,T., Han,J., Puri,R.K., Fang,H., Frueh,F.W., Goodsaid,F.M., Guo,L. *et al.* (2005) Microarray scanner calibration curves: characteristics and implications. *BMC Bioinform.*, **6**, S11.
20. Press,W.H., Flannery,B.P., Teukolsky,S.A. and Vetterling,W.T. (1992) *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, New York, NY, USA.
21. SantaLucia,J., Allawi,H.T. and Seneviratne,A. (1996) Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry*, **35**, 3555–3562.
22. SantaLucia,J. and Hicks,D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, **33**, 415–440.
23. Rule,R.A., Pozhitkov,A.E. and Noble,P.A. (2009) Use of hidden correlations in short oligonucleotide array data are insufficient for accurate quantification of nucleic acid targets in complex target mixtures. *J. Microbiol. Methods*, **76**, 188–195.
24. Halperin,A., Buhot,A. and Zhulina,E.B. (2005) Brush effects on DNA chips: thermodynamics, kinetics, and design guidelines. *Biophys. J.*, **89**, 796–811.
25. Halperin,A., Buhot,A. and Zhulina,E.B. (2006) Hybridization at a surface: the role of spacers in DNA microarrays. *Langmuir*, **22**, 11290–11304.
26. Watterson,J.H., Piunno,P.A.E., Wust,C.C. and Krull,U.J. (2000) Effects of oligonucleotide immobilization density on selectivity of quantitative transduction of hybridization of immobilized DNA. *Langmuir*, **16**, 4984–4992.
27. Peterson,A.W., Heaton,R.J. and Georgiadis,R.M. (2001) The effect of surface probe density on DNA hybridization. *Nucleic Acids Res.*, **29**, 5163–5168.
28. Halperin,A., Buhot,A. and Zhulina,E.B. (2006) On the hybridization isotherms of DNA microarrays: the Langmuir model and its extensions. *J. Phys. Condens. Matter*, **18**, S463–S490.
29. Fritz,J., Cooper,E.B., Gaudet,S., Sorger,P.K. and Manalis,S.R. (2002) Electronic detection of DNA by it intrinsic molecular charge. *Proc. Natl Acad. Sci. USA*, **99**, 14142–14146.
30. Vainrub,A. and Pettitt,B.M. (2003) Surface electrostatic effects in oligonucleotide microarrays: control and optimization of binding thermodynamics. *Biopolymer*, **68**, 265–270.
31. Poulsen,L., Soe,M.J., Snakenborg,D., Moller,L.B. and Dufva,M. (2008) Multi-stringency wash of hybridized 60-mer probes reveals that the stringency along the probe decreases with distance from the microarray surface. *Nucleic Acids Res.*, **36**, e132.
32. Wang,W. and Vaughn,M.W. (2008) Morphology and amine accessibility of (3-aminopropyl) triethoxysilane films on glass surfaces. *Scanning*, **30**, 65–77.
33. Weckx,S., Carlon,E., Vuyst,L.D. and Hummelen,P.V. (2007) Thermodynamic behavior of short oligonucleotides in microarray hybridizations can be described using Gibbs free energy in a nearest-neighbor model. *J. Phys. Chem. B*, **111**, 13583–13590.
34. Pozhitkov,A., Chernov,B., Yershov,G. and Noble,P.A. (2005) Evaluation of gel-pad oligonucleotide microarray technology using artificial neural networks. *Appl. Environ. Microbiol.*, **71**, 8663–8676.
35. Pozhitkov,A., Rule,R.A., Stedtfeld,R.G., Hashsham,S.A. and Noble,P.A. (2008) Concentration-dependency of nonequilibrium thermal dissociation curves in complex target samples. *J. Microbiol. Methods*, **74**, 82–88.
36. Pozhitkov,A., Stedtfeld,R.D., Hashsham,S.A. and 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.
37. Binder,H., Kristen,T., Loeffler,M. and Stadler,P.D. (2004) Sensitivity of microarray oligonucleotide probes: variability and effect of base composition. *J. Phys. Chem.*, **10**, 18003–18014.
38. Halperin,A., Buhot,A. and Zhulina,E.B. (2004) Sensitivity, specificity and the hybridization isotherms of DNA chips. *Biophys. J.*, **86**, 718–730.
39. Zhang,L., Hurek,T. and Reinhold-Hurek,B. (2005) Position of the fluorescent label is a crucial factor determining signal intensity in microarray hybridizations. *Nucl. Acids Res.*, **33**, e166.
40. Naef,F. and 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.
41. Wei,H., Kuan,P.F., Tian,S., Yang,C., Nie,J., Sengupta,S., Ruotti,V., Jonsdottir,G.A., Keles,S., Thomson,J.A. *et al.* (2008) A study of the relationships between oligonucleotide properties and hybridization signal intensities from NimbleGen microarray datasets. *Nucleic Acids Res.*, **36**, 2926–2938.