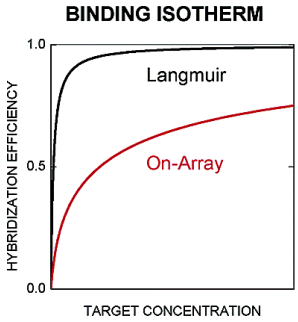
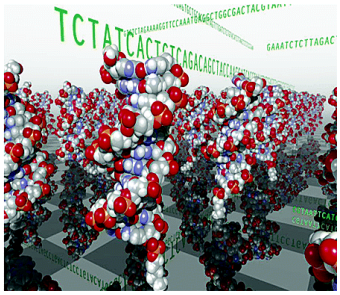
Communication

Sensitive Quantitative Nucleic Acid Detection Using Oligonucleotide Microarrays

Arnold Vainrub, and B. Montgomery Pettitt

J. Am. Chem. Soc., **2003**, 125 (26), 7798-7799• DOI: 10.1021/ja035020q • Publication Date (Web): 06 June 2003





More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 7 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 06/06/2003

Sensitive Quantitative Nucleic Acid Detection Using Oligonucleotide Microarrays

Arnold Vainrub and B. Montgomery Pettitt*

Department of Chemistry, University of Houston, Houston, Texas 77204-5003

Received March 6, 2003; E-mail: pettitt@uh.edu

DNA microarrays with surface tethered oligonucleotide probes have become an important tool in biomedical research[1] and drug discovery.[2] We wish to understand on-surface hybridization mechanisms, which play a significant role in the sensitivity, specificity, and reliability of the detection as well as for accurately quantifying the assayed nucleic acids targets when using microarrays. In this Communication, we present simple new criteria to optimize detection. Our recent theoretical analysis of on-array hybridization has shown the major role of the interface electrostatic interactions appearing as a result of the high negative charge of nucleic acid oligomers.[3−5] In particular, the repulsion between the single strand nucleic acid target and the layer of surface mounted oligonucleotide probes gives rise to the Coulomb blockage of the binding and specific hybridization isotherm, eq 10 in our Communication.[4] Typically, this mechanism overrides the electrostatic interaction with the array's substrate,[3,5] which depends on the material and its surface charge. A theoretical isotherm was compared and found to be in accord[4] with a set of experiments including low on-surface hybridization efficiency,[6,7] melting curve broadening and melting temperature decrease,[8,9] and surface probe density effects.[10−12]

Here, we present a more general framework to quantify and optimize microarray gene expression assays, where the number of transcribed mRNAs can be measured over a suitably broad concentration range. We generalize our isotherm[4] to account for the target depletion and use it to consider how to maximize the sensitivity and detection dynamic range.

The surface hybridization isotherm[4] relates the equilibrium hybridization efficiency θ ($0 \le \theta \le 1$) with the assayed nucleic acid target concentration C_0 (expressed in mol/L)

$$C_0 = \frac{\theta}{1-\theta}\exp\left(\frac{\Delta H_0 - T\Delta S_0}{RT}\right)\exp\left[\frac{wn_PZ_P(Z_P + Z_T\theta)}{RT}\right] \quad (1)$$

Here, ΔH_0 and ΔS_0 are the enthalpy in J mol^{-1} and entropy in J mol^{-1} K^{-1}, respectively, for hybridization in solution, T is the temperature in K, $R = 8.315$ J mol^{-1} K^{-1} is the gas constant, Z_P and Z_T are the length of probe and target oligonucleotides in the number of bases, n_P is the surface density of probes in m^{-2}, and $w = 4 \times 10^{-16}$ J m^2 mol^{-1}.[4] Typically, the concentration C_0 and volume V of the hybridization solution are low because of a usually small available amount of mRNA. Therefore, the target's depletion to the concentration, $C_0 - n_PS\theta/N_AV$, occurs during the course of hybridization, and eq 1 more generally has the form

$$C_0 = \frac{n_PS\theta}{N_AV} + \frac{\theta}{1-\theta}\exp\left(\frac{\Delta H_0 - T\Delta S_0}{RT}\right)\exp\left[\frac{wn_PZ_P(Z_P + Z_T\theta)}{RT}\right] \quad (2)$$

where S is the probe spot area in m^2, and $N_A = 6.02 \times 10^{23}$ mol^{-1} is Avogadro's number. Finally, from eq 2 we obtain a useful relation for the number of duplexes formed, $N_D = n_PS\theta$

$$T = \frac{\Delta H_0 + wn_PZ_P(Z_P + Z_TN_D/n_PS)}{\Delta S_0 + R\ln[(n_PS/N_D - 1)(C_0 - N_D/N_AV)]} \quad (3)$$

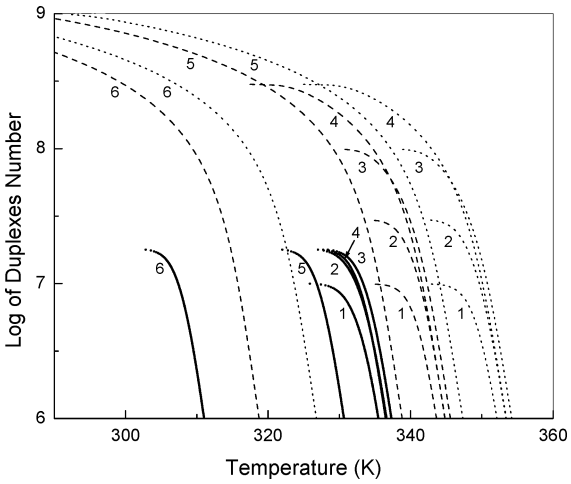


Figure 1. Number of hybrids vs temperature for the assayed target concentrations 10^{-12} (solid), 10^{-9} (dashed), and 10^{-6} mol/L (dotted). The numbers denote the probe surface density: (1) 10^{11}, (2) 3×10^{11}, (3) 10^{12}, (4) 3×10^{12}, (5) 10^{13}, and (6) 3×10^{13} cm^{-2}.

First, we consider the strength of the hybridization signal important in detecting low-expressed genes over nonspecific binding noise. Figure 1 shows the melting curves, that is, the number of duplexes N_D as a function of the temperature T for the target concentrations $C_0 = 10^{-12}$, 10^{-9}, and 10^{-6} mol/L. In Figure 1, T was calculated and plotted as a function of N_D from eq 3 and was properly oriented. The results are demonstrated for 25-mer probes of sequence GTCCGATAAGCCTGTGTCCAATAAC with a perfectly matched (complimentary) target of length 100 bases, $\Delta H_0 = -819.3$ kJ mol^{-1} and $\Delta S_0 = -2.229$ kJ mol^{-1} K^{-1} at 1 M added NaCl,[13] the probe spot area $S = 0.01$ mm^2, the hybridization solution volume $V = 30$ μL, and room temperature $T = 298$ K. At the lowest considered $C_0 = 10^{-12}$ mol/L, the low-temperature saturation of N_D is due to depletion of targets, whereas at higher C_0 it corresponds to complete hybridization ($\theta = 1$) of the probes with the surface density n_p. The important result in Figure 1 is that the hybridization signal increases with n_p ($10^{11} < n_p < 10^{12}$ cm^{-2}) and achieves a substantial peak of sensitivity near $n_p = 10^{12}$ cm^{-2} because of strong Coulomb repulsion (blockage) of hybridization at higher probe surface density (Figure 2). Notice at signal maximum the mean interprobe distance is 10 nm, and thus no significant steric effect is expected for 2 nm diameter DNA double

10.1021/ja035020q CCC: $25.00

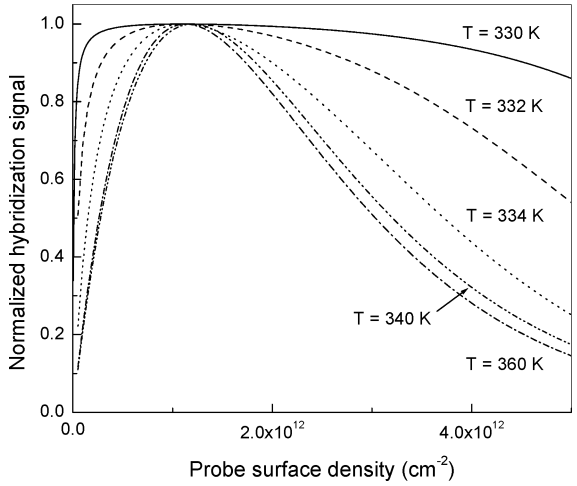


Figure 2. Normalized hybridization signal vs the probe surface density.

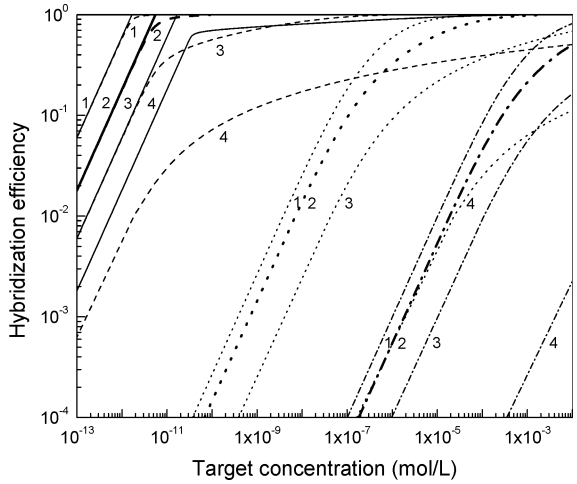


Figure 3. Hybridization isotherms at temperatures of 298 K (solid), 328 K (dashed), 348 K (dotted), and 358 K (dashed−dotted). The numbers denote the probe surface density: (1) 3×10^{11}, (2) 10^{12}, (3) 3×10^{12}, and (4) 10^{13} cm^{-2}.

helixes. Optimizing analytically to find the maximum of N_D as a function of n_p in eq 3, we obtain

$$n_p = \frac{RT}{wZ_p^2} \tag{4}$$

The predicted strong probe length dependence as $1/Z_p^2$ shows that for common 70-mer probe microarrays the optimal surface probe density is as low as 1.3×10^{11} cm^{-2}.

Next, we consider the dynamic range of detection. It should be broad enough (3−4 orders of magnitude) to measure lowly and highly expressed genes in a single assay. Moreover, the linear part is of special interest in connection with standard two-color dye labeled mRNA relative abundance array-based assays.[1] Figure 3 displays the hybridization isotherms at different temperatures calculated from eq 2 for the above specified array parameters and hybridization conditions. The bold curves are for the probe density $n_p = 10^{12}$ cm^{-2} corresponding to the strongest hybridization signal. As shown, the broad linear detection range up to almost complete hybridization ($\theta = 1$) can occur at room temperature, but is limited to $\theta \ll 1$ for somewhat elevated temperatures. However, hybridization is typically done above room temperature to avoid contributions of mismatched targets with lower melting temperatures and also to melt the target's possible secondary structure. For any specific hybridization temperature, Figure 3 can be used to find the concentration range below saturation and its appropriate linear region. Alternatively, from eq 2 the nonlinear error δ is

$$\delta = \left[\frac{wn_pZ_pZ_T\theta(1-\theta) + RT}{(1-\theta)^2RT}\exp\left(\frac{wn_pZ_pZ_T\theta}{RT}\right) - 1\right] \times \exp\left(\frac{\Delta H_0 - T\Delta S_0 + wn_pZ_p^2}{RT}\right) \tag{5}$$

The ratio of the stronger to weaker hybridization signals in two-color measurement provides a lower estimate for the ratio of concentrations with the relative error below δ. For an acceptable error δ, the plot δ versus θ from eq 5 can be used to find the upper limit of θ. For accurate quantitative measurements in nonlinear range at higher θ, the isotherm from eq 2 allows one to find the target concentration C_0.

In conclusion, on the basis of the theoretical binding isotherm confirmed by published experiments, we developed a model of on-array hybridization. We derived simple, analytical relations for the sensitivity maximum and linear dynamic range. We demonstrated the use of these equations for the rational design of gene expression assays. Current microarrays evaluate only relative mRNA abundance, but the absolute number is required for data comparison. Importantly, the mRNA number can be obtained using the isotherm if the hybridization signal is calibrated to the number of duplexes. Our approach can also be applied to optimize other microarray applications including genotyping, single nucleotide polymorphisms, and sequencing.

Acknowledgment. We thank Lloyd Smith and Michael Shortreed for stimulating discussions. This work was partially supported by a grant from NIH and the R. A. Welch foundation.

References

(1) Lockhart, D. J.; Winzeler, E. A. *Nature* **2000**, *405*, 827−836.
(2) Roses, A. D. *Nat. Rev. Drug Discovery* **2002**, *1*, 541−549.
(3) Vainrub, A.; Pettitt, B. M. *Chem. Phys. Lett.* **2000**, *323*, 160−166.
(4) Vainrub, A.; Pettitt, B. M. *Phys. Rev. E* **2002**, *66*, art. no.-041905.
(5) Vainrub, A.; Pettitt, B. M. *Biopolymers* **2003**, *68*, 265−270.
(6) Guo, Z.; Guilfoyle, R. A.; Thiel, A. J.; Wang, R.; Smith, L. M. *Nucleic Acids Res.* **1994**, *22*, 5456−5465.
(7) Shchepinov, M. S.; Case-Green, S. C.; Southern, E. M. *Nucleic Acids Res.* **1995**, *25*, 1155−1161.
(8) Forman, J. E.; Walton, I. D.; Stern, D.; Rava, R. P.; Trulson, M. O. *ACS Symposium Series* **1998**, *682*, 206−228.
(9) Lu, M. C.; Hall, J. G.; Shortreed, M. R.; Wang, L. M.; Berggren, W. T.; Stevens, P. W.; Kelso, D. M.; Lyamichev, V.; Neri, B.; Skinner, J. L.; Smith, L. M. *J. Am. Chem. Soc.* **2002**, *124*, 7924−7931.
(10) Steel, A. B.; Herne, T. M.; Tarlov, M. J. *Anal. Chem.* **1998**, *70*, 4670−4677.
(11) Watterson, J. H.; Piunno, P. A.; Wust, C. C.; Krull, U. J. *Langmuir* **2000**, *16*, 4984−4992.
(12) Peterson, A. W.; Heaton, R. J.; Georgiadis, R. M. *Nucleic Acids Res.* **2001**, *29*, 5163−5168.
(13) SantaLucia, J.; Allawi, H. T.; Seneviratne, P. A. *Biochemistry* **1996**, *35*, 3555−3562.

JA035020Q