Physicochemical perspectives on DNA microarray and biosensor technologies

Rastislav Levicky\(^1\) and Adrian Horgan\(^2\)

\(^1\)Department of Chemical Engineering, Columbia University, 500 West 120th Street, New York, NY, USA
\(^2\)Solexa Ltd, Chesterford Research Park, Little Chesterford, Nr Saffron Walden, Essex, CB101XLT, UK

Detection and sequence-identification of nucleic acid molecules is often performed by binding, or hybridization, of specimen ‘target’ strands to immobilized, complementary ‘probe’ strands. A familiar example is provided by DNA microarrays used to carry out thousands of solid-phase hybridization reactions simultaneously to determine gene expression patterns or to identify genotypes. The underlying molecular process, namely sequence-specific recognition between complementary probe and target molecules, is fairly well understood in bulk solution. However, this knowledge proves insufficient to adequately understand solid-phase hybridization. For example, equilibrium binding constants for solid-phase hybridization can differ by many orders of magnitude relative to solution values. Kinetics of probe–target binding are affected. Surface interactions, electrostatics and polymer phenomena manifest themselves in ways not experienced by hybridizing strands in bulk solution. The emerging fundamental understanding provides important insights into application of DNA microarray and biosensor technologies.

Introduction

Solid-phase hybridization, in which nucleic acid strands tethered to a solid support bind DNA or RNA molecules from solution, underpins the modern microarray and biosensor biotechnologies now widely used for genotyping, studying gene expression and for biological detection [1–6]. Applications of these tools have increased tremendously as is evident, for example, in the number of scientific publications in this area (Figure 1). Microarrays, in which thousands of hybridization reactions are carried out in parallel, are commonly used for addressing fundamental questions in biology and for sample characterization up to whole-genome scale. Biosensors tend to be dedicated to specific tasks, such as detection of a small number of analyte sequences, with data acquisition often performed in real-time.

In solid-phase or ‘surface’ hybridization, association of immobilized strands, referred to as ‘probes’, with ‘target’ sequences from solution occurs at a solid–liquid interface. The interfacial environment is distinct from the bulk solution; thus, surface hybridization might be anticipated to deviate from expectations solely based on knowledge of solution hybridization. These differences are crucial to guiding device design, formulating assay protocols and developing data analysis tools. Because of space constraints, discussion of surface modification chemistries, detection methods and alternate hybridization schemes (such as those employing hydrogel matrices [7,8]) has been omitted. Much of this additional information is available in other recent reviews [9–11].

Background

As with all natural processes, solid-phase hybridization is subject to thermodynamic and kinetic constraints [12]. Thermodynamics define limits on discrimination of different target sequences and detection of low copy targets. Kinetics determine how quickly equilibrium is approached. Kinetic considerations also include mass transport, which is well understood within standard transport theory. The following discussion therefore focuses on the hybridization reaction itself.

Kinetics

In the simplest model, hybridization is treated as a one-step reversible reaction with rate constants \(k_f\) and \(k_r\) (\(P\): probe; \(T\): target; \(D\): duplex):

\[
P + T \xrightleftharpoons[k_f]{k_r} D
\]  

(Equation 1)

The forward reaction is assumed to proceed at a rate \(k_f(1-f)P_0C_T\), whereas the reverse rate is \(k_rP_0\). \(P_0\) is the total surface density (number of strands/unit of area) of probes of which a fraction \(f\) is hybridized; thus \((1-f)P_0\) is the surface density of unhybridized probes. \(C_T\) is the bulk concentration of target molecules. Equilibrium is characterized by the equality of forward and reverse rates:

\[
k_fC_T(1-f)P_0 = k_rfP_0 \text{ OR } f/(C_T(1-f)) = k_f/k_r = K_E
\]  

(Equation 2)

where \(K_E\) is the equilibrium constant. At the outset of hybridization \(f=0\), the forward rate is maximum and the reverse rate is zero. As hybridization proceeds \(f\) increases and \(C_T\) decreases; these changes result in a decrease of the forward rate \(k_f\) and an increase in the reverse rate. Within this basic picture, Figure 2 plots the forward and reverse rates as a function of time for two values of \(k_r\), assuming \(k_fC_T\) is fixed. The case characterized by smaller \(k_r\) approaches equilibrium more slowly. Because \(k_f\) varies strongly with sequence, such considerations directly impact on achievement of equilibrium in microarray...
assays in which numerous probe–target sequences hybridize simultaneously [13]. The simplicity of this one-step assays in which numerous probe–target sequences hybridize simultaneously [13]. The simplicity of this one-step assays in which numerous probe–target sequences hybridize simultaneously [13]. The simplicity of this one-step assays in which numerous probe–target sequences hybridize simultaneously [13]. The simplicity of this one-step assays in which numerous probe–target sequences hybridize simultaneously [13]. The simplicity of this one-step assays in which numerous probe–target sequences hybridize simultaneously [13]. The simplicity of this one-step assays in which numerous probe–target sequences hybridize simultaneously [13]. The simplicity of this one-step assays in which numerous probe–target sequences hybridize simultaneously [13].

Thermodynamics
Translating understanding of solution hybridization to the surface case is central to designing probe sets for microarrays and for interpreting experimental data [16–20]. Hybridization thermodynamics in bulk solution are sufficiently well understood to allow prediction of standard free energies ($\Delta G^\circ$) and equilibrium association constants $K_E = \exp(-\Delta G^\circ/RT)$ ($R$: gas constant; $T$: absolute temperature) from nearest-neighbor models [21]. Figure 3 illustrates the thermodynamic relationship between solution and solid-phase hybridization. In solution hybridization (Figure 3a), probe strands $P$ and target strands $T$ combine to form duplex $D$ with a free energy change $\Delta G_D^-$. By contrast, in solid-phase hybridization (Figure 3b) a strand $T$ must penetrate into a layer of $P$ strands to hybridize, with an overall free energy change $\Delta G_{Tot}^-$. Applying the thermodynamic tenet that free energy changes depend only on initial and final states, surface hybridization can be broken down into three hypothetical steps as in Figure 3c: (i) release of a $P$ strand into solution; (ii) hybridization with a $T$ strand under solution conditions; and (iii) immobilization of the duplex $D$ formed. The associated change in free energy is $\Delta G_{Tot}^- = \Delta G_1^- + \Delta G_2^- + \Delta G_3^-$. And the equilibrium binding or association constant for solid-phase hybridization $K_{ES}$ is:

$$K_{ES} = K_{EB} \exp(-\Delta G_2^-/RT)$$  \hspace{1cm} (Equation 3)

where $K_{EB}$ is the bulk solution equilibrium constant. The factor $\exp(-\Delta G_2^-/RT)$ ‘corrects’ $K_{EB}$ so that the product describes the situation of solid-phase hybridization. If $\Delta G_2^- = -\Delta G_3^-$ then bulk and surface hybridization would be thermodynamically equivalent. In general, this equivalence will not be obeyed because the extraction of a single-stranded $P$ (step 1) is not simply a reversal of an insertion of a double-stranded $D$ (step 3) because single- and double-stranded molecules are expected to interact differently with the probe film environment owing to differences in size, charge, rigidity or other molecular properties.

Interpretation of experimental data on solid-phase hybridization can pose significant challenges. For example, although end-attachment of probes is often the intended scenario, it can be difficult to confirm and some ostensibly end-immobilization chemistries might lead to cross-linking at internal sites, resulting in multipoint attachment that can hinder association with target molecules. Caution is also warranted with regard to achievement of hybridization equilibrium. Equilibrium can be convincingly demonstrated if the same final state is reached from different initial states; such controls, however, are usually not available. This article is based on studies representing well-defined situations; nonetheless, the original reports remain the best source for ascertaining the experimental conditions of a particular study.

Comparison of solid-phase ($K_{ES}$) and bulk ($K_{EB}$) equilibrium constants
Figure 4 compares experimental values of $K_{ES}$ with $K_{EB}$ values calculated for the same sequences and conditions as outlined in [21]. The $K_{ES}$ data represent four different studies employing end-tethered probe strands 10–30 nucleotides long hybridizing to fully complementary, same-length targets [22–25]. Four different immobilization chemistries, ionic strengths from 0.1–1 M, and temperatures from 10–70 °C were used. $K_{ES}$ was derived either from measurements of the hybridized fraction $f$ [22,23,25] or from the ratio of forward and reverse rate constants [24] (see Equation 2). Measurements were carried out in situ
so that perturbation owing to washing of hybridized surfaces was avoided. It is noteworthy that, in this spectrum of conditions, $K_{ES}$ falls in a narrow range of $\sim 5$ decades whereas $K_{EB}$ varies over $\sim 30$ decades. This indicates that bulk solution and surface thermodynamics are not equivalent; that is, $\Delta G_{1}^{o} \neq -\Delta G_{3}^{o}$. Usually $K_{ES}$ is suppressed relative to $K_{EB}$ (right of dashed line in Figure 4), although data also indicate that hybridization on a surface can be more thermodynamically favored than in solution (left of dashed line). Suppression of $K_{ES}$ relative to $K_{EB}$ might reflect electrostatic and steric penalties associated with penetration of a target strand into a probe layer. On the other hand, a high local concentration of probes might stabilize binding of targets in states not possible under solution conditions, leading to $K_{ES} > K_{EB}$ [25]. Analysis of data obtained from commercial microarrays suggests that solid-phase hybridization is less thermodynamically favored than hybridization of the same sequences in solution [20].

That thermodynamics of solution and surface hybridization differ is perhaps not surprising. Typically, oligonucleotide probe films are characterized by a coverage of $10^{12}$ to $10^{13}$ probes/cm$^2$ and a layer thickness of several nanometers [26]. These values translate to a local concentration of $\sim 0.1$–$1$ M in nucleotides, much higher than when hybridization is carried out in solution, implying a very different local environment for solid-phase hybridization.

**Impact of probe layer structure**

Experiments indicate that surface hybridization is suppressed when the coverage of probe molecules is too high [27–32]. For end-tethered 25-mer (i.e., 25 nucleotides long) probes Steel et al. reported close to 100% hybridization to complementary targets under 1 M ionic strength when probe coverage was below $\sim 4 \times 10^{12}$ cm$^{-2}$, with a sharp fall off in hybridization extents at higher coverages [27]. An alternate measure of probe–target affinity is melting temperature $T_{M}$, defined as that temperature at which...
half of hybridizable probes remain single-stranded while in equilibrium with target solution of fixed composition [39–35]. A $T_M$ decrease of nearly 10 °C, corresponding to decreased stability of hybridized probe–target pairs, was reported when the density of 20mer probes increased from $\sim 1 \times 10^{11}$ cm$^{-2}$ to $\sim 4 \times 10^{12}$ cm$^{-2}$ [34,35]. Thus changes in both $T_M$ and hybridization extents indicate that solid-phase hybridization is disfavored at high probe coverages.

Because nucleic acids carry a minus one charge per nucleotide, electrostatic effects might be expected to be partially responsible for destabilization of probe–target binding. A theory developed by Vainrub and Pettitt [36] derives an electrostatic correction factor corresponding to \( \text{exp}\left(-\left(\Delta G_f^0 + \Delta G_s^0\right)/RT\right) \) (compare with Equation 3). This theory models the probe layer as a plane with a smeared charge density $\sigma(f) (\text{charge/area})$ where $f$ denotes extent of hybridization as before. Suppression of $T_M$ is also predicted. A subsequent analysis of the impact of competitive hybridization on solid-phase assays by Halperin et al. treats the probe film as three-dimensional with a smeared and uniform charge density $\rho(f) = \sigma(f)/H$, where $H$ is the film thickness [37]. Both analyses find that the electrostatic part of $\Delta G_f^0 + \Delta G_s^0$ increases linearly with probe coverage (i.e. density of immobilized charge); however, a different dependence on concentration of bulk salt is predicted.

Probe coverage also affects hybridization kinetics, as evident from a decrease in initial rates of hybridization at higher probe densities [28,38]. Henry et al. reported that $k_f$ for hybridization of end-tethered 22mers in 0.11 M ionic strength was five times slower than in solution when probe coverage was $1.4 \times 10^{12}$ cm$^{-2}$, and ten times slower when probe coverage increased to $2.8 \times 10^{12}$ cm$^{-2}$ [38]. Peterson et al. also observed a slowdown in hybridization with increasing probe coverage [28]. Hagan and Chakraborty developed a theory describing kinetics of solid-phase hybridization. These authors adapted methods from polymer science [39,40] to model solid-phase hybridization subject to configurational statistics of polymer chains [41]. The forward hybridization rate was predicted to decrease with probe coverage $P_0$ as $k_f \sim P_0^{1.8}$ in reasonable agreement with experimental trends [41]. Understanding how polymer phenomena impact solid-phase hybridization is not yet complete. One study investigating assembly of probe films suggested that polymer configurational effects became important once probe length exceeded 20 nucleotides [42].

In the limits of sparse films, when probes are too far apart to come into contact, patches of bare surface will be accessible to adsorption of target strands. In this limit targets might first adsorb and then diffuse along the solid support before hybridizing to a probe [43]. The additional pathway provided by such ‘two-dimensional’ (2D) hybridization is predicted to enhance hybridization rates [43,44]. Comparison of theoretical expectations to experimental data showed evidence of 2D enhancement for hybridization of 20mers at coverages of $\sim 7 \times 10^{10}$ cm$^{-2}$ [44,45]. The impact of 2D enhancement would be expected to diminish when complex target mixtures are assayed because much of the accessible surface around a probe will become occupied by other (non-complementary) target sequences, and also if surface treatments to suppress nonspecific adsorption are employed.

The influence of strand length was investigated for end-tethered 10mer, 20mer, and 30mer probe–target pairs at a fixed probe coverage of $1.2 \times 10^{13}$ cm$^{-2}$ [24]. The forward hybridization rate $k_f$ increased with strand length, as might be expected if longer probe-target pairs present more sites for initiation of duplex formation. Intriguingly, experimental dependence of $k_f$ on length $L$ was significantly stronger ($k_f \sim L^{-0.5}$) than suggested by theoretical arguments ($k_f \sim L^{1.5}$) [41]. Using polymeric (116–2057 nt) probes and targets, Stillman and Tonkinson also observed increased hybridization rates with probe size [46]. Graves and coworkers investigated hybridization of polymeric (157–864 nt) targets to oligonucleotide probes [47]. A delay in onset of hybridization was observed, increasing from 1 h to 8 h with target length at a fixed target concentration of $4 \times 10^{-5}$ M. By contrast, a 20mer target hybridized in less than 5 min. The delay was postulated to reflect initial formation of relatively weakly surface-adsorbed target states, which only slowly rearranged into stable probe–target duplexes capable of withstanding the washing protocol.

The structure of a probe layer is asymmetric in that it is bounded by a solid support from one side. This asymmetry suggests that the location of a particular base along a probe strand might influence its impact on hybridization. Using surface plasmon resonance, Georgiadis and coworkers monitored hybridization between a 25mer probe and two 18mer targets that hybridized at different locations on the probe [23]. Solution thermodynamics of both target–probe pairs were equivalent. Target 18low hybridized to the first 18 bases at the tethered end of the probe, whereas 18hi hybridized to the last 18 bases at the free end. At a sparse coverage of $1.5 \times 10^{12}$ cm$^{-2}$, both targets hybridized at the same rate. By contrast, when probe coverage increased to $3 \times 10^{12}$ cm$^{-2}$, 18hi hybridized about three times faster. These results suggest presence of an activation barrier associated with target penetration into the probe film, which is also expected on theoretical grounds [41]. In their study of 60mer oligonucleotide arrays, Hughes et al. [48] found that base mismatches further away from the surface exerted a stronger effect on hybridization. These results similarly suggest that probe accessibility to hybridization improves with increased distance from the surface, consistent with earlier studies in which incorporation of molecular spacers to displace probes from the solid support led to higher hybridization yields [49,50].

Proximity to a solid support also implies that surface interactions will be present. This issue is somewhat intractable in that it is specific to the modification chemistry employed. Using single-molecule spectroscopy, Osborne et al. observed multiple fluorescence behaviors from isolated DNA probes immobilized through an amine endgroup to 3′-glycidoxypropyltrimethoxysilane modified silica [51]. The fluorescence data suggest that the labeled probes can exist in different arrangements on the surface. Just $\sim 5\%$ of the probes were active toward targets, indicating that surface chemistry was prominent in governing hybridization activity. Suppressed hybridization in the presence of strong probe–surface interactions has been also reported by others [26]. In some instances surface
interactions can be tuned to achieve a specific effect. On conductive supports application of electric fields can enhance mismatch discrimination [52], alter orientation of immobilized DNA molecules [53] and accelerate hybridization kinetics [47]. Vainrub and Pettitt have recently extended their theoretical model [36] to discuss the role of surface electrostatics [54].

When present, intramolecular structure, whether in probe or target species, can strongly influence hybridization yields. Using a tRNA target, Mir and Southern observed high hybridization yields when the new duplex region was optimally accommodated within the native tRNA structure, for example, through stacking with native double-stranded regions. [55]

Effect of temperature and ionic strength
Higher temperatures are often reported to increase the forward rate of solid-phase hybridization [38,45,56]. On the other hand, using 10mer probes at a coverage of $1.2 \times 10^{13} \text{cm}^{-2}$, Okahata et al. observed that $k_f$ did not strongly depend on temperature although $k_r$ increased significantly so that equilibrium was still reached ~60 times faster at 30 °C than at 15 °C [24] (see Figure 2). Increase in ionic strength from 0.1 to 0.5 M NaCl increased $k_f$ about six-fold while decreasing $k_r$ by an order of magnitude [24]. These trends qualitatively agree with solution behavior. [57] In solution, higher ionic strength stabilizes double-stranded structure leading to faster $k_f$ and slower $k_r$. Higher temperature causes $k_r$ to increase whereas $k_f$ can increase, stay the same or even decrease. The varied response of $k_f$ is attributed to temperature dependence of forming a hybridization nucleus [57].

Impact of mismatches on solid-phase hybridization
A crucial consideration for applications is the ability to discriminate mismatched from fully complementary targets. Using 25mer end-tethered probes and 25mer targets, Peterson et al. found that mismatches at one or two base positions slowed down approach to equilibrium at probe coverages of $3 \times 10^{12} \text{cm}^{-2}$, accompanied by features in hybridization vs time curves suggestive of multiple structural rearrangements [23]. By contrast, at a more sparse probe coverage of $1.5 \times 10^{12} \text{cm}^{-2}$ the rate of hybridization was similar for matched and mismatched targets. This suggests that probe density is a key factor governing the influence of mismatches. Forman et al. observed that a central mismatch in a 20mer probe–target pair did not influence hybridization kinetics [58]. Okahata et al. studied hybridization of 20mers at a probe coverage of $1.2 \times 10^{13} \text{cm}^{-2}$, finding that $k_f$ decreased and $k_r$ increased for mismatched sequences [24]. The overall effect was that approach to equilibrium was faster for mismatched than for complementary targets. The impact of mismatches on hybridization between 60mer probes and synthetic mRNA polymers was investigated by Dai et al. [13]. In contrast to the other studies, samples containing a complex mixture of target sequences were used. Under these conditions, mismatched targets achieved equilibrium faster than fully complementary ones. Usage of kinetic data to better identify signals arising from mismatched targets was suggested [13].

The diversity of experimental observations regarding influence of mismatches could in part reflect formation of structures more complex than one-to-one hybridization. For example, a target molecule can bridge and hybridize across multiple probes (Figure 5b). Compared with a complementary target, the presence of a mismatch might facilitate bridging by destabilizing duplex formation at the location of the bridge (Figure 5c). The progression through bridging or other structures during the approach to equilibrium would be expected to depend on details of sequence, including mismatches, on the probe spacing, and on probe and target lengths.

Bhanot et al. developed a computer model of hybridization in a microarray type experiment, accounting for a complex sample background. [19] It was shown that cross-hybridization, involving binding of mismatched probe–target pairs in competition with perfect complements, can profoundly slow down approach to equilibrium. Poor specificity of the probe set, and thus increased susceptibility to cross-hybridization, translated to increased ‘concentration bias’ in which true concentration of rarer targets was underestimated and that of more abundant ones was overestimated. This is an important consideration for gene expression studies if levels of gene activity are to be estimated from concentrations of mRNA targets.

Modeling of data from microarrays
Application of physical models to microarray data has been undertaken in several recent studies [18,20,59]. The relationship of solution target concentration $C_T$ to the extent of hybridization $I$, as monitored by fluorescence intensity $I$, was shown to exhibit nonlinear behavior owing to saturation of probe sites with targets at high $C_T$. The nonlinear response agreed with a Langmuir isotherm form [20,59]. Held et al. showed that $I$ increased with more favorable free energy of hybridization $\Delta G^e_2$, between fully complementary probe-target pairs, with the consequence that strongly binding sequences exhibited chemical saturation at lower $C_T$. [20] This is an important consideration for selection of microarray probe sequences [18]. Moreover, background signal from mismatched targets scaled exponentially with $\Delta G^e_2$, indicating that high affinity

Figure 5. (a) One-to-one hybridization. (b) A target bridge across two probe strands. (c) A target bridge facilitated by a mismatch (cross).
probes are also more susceptible to association with targets that are only partially complementary [18,20]. The impact of dye label incorporation into target strands on thermodynamics of probe-target association has also been theoretically modeled. [60] Significantly, most microarray experiments involve a washing step to improve stringency of sequence discrimination; the perturbation this represents to hybridization equilibria is an important question yet to be addressed.

Conclusions

Solid-phase hybridization, as practiced in DNA biosensor and microarray technologies, is a more complex process than solution hybridization. Although recent years brought significant progress in physical understanding of solid-phase hybridization, many aspects crucial both to applications and to fundamental insight, remain unresolved. Among these, perhaps the foremost questions include identifying the origin of the difference between $K_E$ and $K_{EB}$ (see Figure 4), establishing how solution conditions and probe layer structure affect discrimination of mismatches and understanding key factors influencing hybridization kinetics. Resolution of these issues will require close interaction between experimentalists and theorists, and between researchers who aim to understand solid-phase hybridization under precisely defined conditions (e.g. equilibrium, characterized surfaces and target solutions) on one hand and technology practitioners faced with complex sample mixtures and experimental protocols on the other.

References


Elsevier.com – Dynamic New Site Links Scientists to New Research & Thinking

Elsevier.com has had a makeover, inside and out. Designed for scientists’ information needs, the new site, launched in January, is powered by the latest technology with customer-focused navigation and an intuitive architecture for an improved user experience and greater productivity.

Elsevier.com’s easy-to-use navigational tools and structure connect scientists with vital information – all from one entry point. Users can perform rapid and precise searches with our advanced search functionality, using the FAST technology of Scirus.com, the free science search engine. For example, users can define their searches by any number of criteria to pinpoint information and resources. Search by a specific author or editor, book publication date, subject area – life sciences, health sciences, physical sciences and social sciences – or by product type. Elsevier’s portfolio includes more than 1800 Elsevier journals, 2200 new books per year, and a range of innovative electronic products. In addition, tailored content for authors, editors and librarians provides up-to-the-minute news, updates on functionality and new products, e-alerts and services, as well as relevant events.

Elsevier is proud to be a partner with the scientific and medical community. Find out more about who we are in the About section: our mission and values and how we support the STM community worldwide through partnerships with libraries and other publishers, and grant awards from The Elsevier Foundation.

As a world-leading publisher of scientific, technical and health information, Elsevier is dedicated to linking researchers and professionals to the best thinking in their fields. We offer the widest and deepest coverage in a range of media types to enhance cross-pollination of information, breakthroughs in research and discovery, and the sharing and preservation of knowledge. Visit us at Elsevier.com.