SURVEY AND SUMMARY

Microarrays for identifying binding sites and probing structure of RNAs

Ryszard Kierzek¹, Douglas H. Turner² and Elzbiet’a Kierzek¹,*

1Institute of Bioorganic Chemistry Polish Academy of Sciences, 61-704 Poznan, Noskowskiego 12/14, Poland and
2Department of Chemistry, University of Rochester, Rochester, NY 14627, USA

Received October 21, 2014; Revised November 27, 2014; Accepted November 28, 2014

ABSTRACT

Oligonucleotide microarrays are widely used in various biological studies. In this review, application of oligonucleotide microarrays for identifying binding sites and probing structure of RNAs is described. Deep sequencing allows fast determination of DNA and RNA sequence. High-throughput methods for determination of secondary structures of RNAs have also been developed. Those methods, however, do not reveal binding sites for oligonucleotides. In contrast, microarrays directly determine binding sites while also providing structural insights. Microarray mapping can be used over a wide range of experimental conditions, including temperature, pH, various cations at different concentrations and the presence of other molecules. Moreover, it is possible to make universal microarrays suitable for investigations of many different RNAs, and readout of results is rapid.

INTRODUCTION

Microarrays contain a library of oligonucleotides or polynucleotides with each spotted in a defined location. It is difficult to say unambiguously when microarrays appeared on the scene (1–5). Arrays with DNA attached to cellulose or another support and used for probing DNAs were precursors of currently used microarrays (6). In the 90’s, high-throughput preparation of microarrays was made possible by employment of glass as support (2) and by methods for preparation of probes, including chemical synthesis on solid support (7). Currently, microarrays are widely used to study many biological processes and to perform high-throughput analysis (8–11).

This review describes investigations in which DNA or RNA microarrays were used for identifying oligonucleotide binding sites and probing structure. Deep sequencing allows fast determination of DNA and RNA sequence. High-throughput methods for the determination of secondary structures of RNAs have been developed (12–17). These methods, however, do not reveal binding sites for oligonucleotides. In contrast, microarrays directly determine binding sites while also providing structural insights. Microarray mapping can be used over a wide range of experimental conditions, including temperature, pH, various cations at different concentrations and the presence of other molecules. Moreover, it is possible to make universal microarrays suitable for investigations of many different RNAs, and readout of results is rapid.

EARLY-STAGE HISTORY

Studies of RNA structure with oligonucleotide microarrays were started by E. Southern (8,18–26). To generate DNA microarrays, DNA probes of 1–20 nucleotides were synthesized on long functionalized glass plates that were moved after each phosphoramidite coupling reaction (Figure 1) (24,27–29). Hybridization of radioactively labeled RNA to microarrays revealed binding of DNA probes to target RNA. Moreover, based on the studies of interactions of model oligonucleotides, the Southern group demonstrated that (i) probes bind to single-stranded regions of target RNAs and the presence of mismatches between probe and RNA sequence reduces binding intensity, (ii) strong hybridization

*To whom correspondence should be addressed. Tel: +48 61 853 8503; Fax: +48 61 852 0532; Email: elzbieta.kierzek@ibch.poznan.pl

© The Author(s) 2014. Published by Oxford University Press on behalf of Nucleic Acids Research. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
can be observed when only a part of long probes canonically base-pairs to target RNA and (iii) the presence of secondary structural motifs adjacent to single-stranded regions influences hybridization.

The Southern group intensively studied hybridization to tRNA^{Phe} (21). Hybridization to tRNA^{Phe} was performed in the presence of 3.5 M tetramethylammonium chloride (TMACl) which reduces the sequence dependence of association constants relative to 1 M NaCl and 10 mM MgCl₂. The results provided the following conclusions: (i) only a few of 65 dodecamer probes bind to tRNA^{Phe}, (ii) the hybridization results depend on salt (3.5 M TMACl versus 1 M NaCl/10 mM MgCl₂), (iii) only four regions of tRNA^{Phe} (acceptor stem, D stem-loop, variable loop/Psi1 C stem and variable loop/anticodon stem) strongly bind to the microarray probes, (iv) strength of hybridization covers two to three orders of magnitude, (v) hybridization strongly depends on RNA binding region, sometimes shifting binding region by one position results in changing hybridization yield up to 15-fold, (vi) hybridization yield is not always proportional to length of DNA probes, (vii) hybridization is the same for probes attached at the 5’ or 3’ end to the glass support. The authors observed hybridization of probes to single-stranded loops and often to adjacent helical regions. They postulated that even a short single-stranded region is sufficient for duplex nucleation and subsequent invasion of an adjacent stem, especially when the single-stranded region has A-form character and stacks on the stem. The authors concluded that ‘major determinants of hybridization lie in the structure of RNA’ (21).

The Southern group also studied application of microarrays to transcripts of 175, 336, 512, 1291 and 1389 nucleotides of the mRNA for cyclin B5 (18,19,30). Probe sequences were designed to scan only a particular fragment of 111 nucleotides. Hybridization of the transcripts to microarrays containing up to 21-nucleotide-long DNA probes showed similar patterns in certain regions of the 111 nucleotides but also different patterns characteristic for each length of transcript. Often hybridization yield of the same DNA probes was more efficient for shorter transcripts than for longer ones. They suggested that this results from more tertiary folding of longer transcripts and concluded that hy-
bridization results can provide insight into both short- and long-range interactions. The latter can be indicated by loss of hybridization to some regions when transcript length is increased.

Results collected by the Southern group were mostly used to select antisense oligonucleotides to achieve efficient gene silencing. For example, they studied three *Xenopus laevis* B1, B4 and B5 cyclins mRNAs ranging in size between 420 and 1400 nucleotides (19). They scanned abilities of 21-nucleotide-long DNA probes to hybridize to the first 120 nucleotides of those mRNAs. Often probes bound to specific regions in all three mRNAs, but strong hybridization was not correlated with probe base composition, presumably because of secondary structure in the target RNA. Predicted secondary structures of target mRNAs indicated that probes often bind intensively to helical regions. The microarray results provided optimal antisense oligonucleotides to the first 120-nt region. Selected antisense oligonucleotides were then assayed for their effect on translation of endogenous cyclin mRNAs in *Xenopus* egg extracts and their ability to promote RNase H cleavage of cyclin mRNAs in oocytes. DNA oligonucleotides identified as strong binders in microarray experiments inhibited translation of their cognate targets in *Xenopus* egg extract and strongly reduced synthesis of the targeted cyclins, whereas those oligonucleotides with weak binding to microarrays produced no or little inhibition. Generally, the authors demonstrated the relationship of target site and oligonucleotide sequence with antisense activity. The authors suggested that many factors could be responsible for behavior, including sequence composition and length of DNA probes, secondary structures of target RNAs and less favorable thermodynamic stability of DNA/RNA duplexes over intramolecular RNA/RNA helices.

**DEVELOPING RNA-LIKE AND ISOENERGETIC MICROARRAYS FOR PROBING BINDING**

The Kierzek and Turner groups developed a new approach to study RNAs with oligonucleotide microarrays (31–37). In this approach: (i) the probes are 5- or 6-nucleotides long and based on the natural 2'-O-methylated backbone because single-stranded regions within natural RNAs are short and 2'-O-methylRNA/RNA duplexes are more stable (38) than RNA/RNA (39) or DNA/RNA (40) duplexes, and (ii) the probes are modified to make hybridization to unstructured RNA relatively independent of base sequence, i.e. probes are isoenergetic.

Short isoenergetic probes simplify the interpretation of probe binding, which depends on differences between the free energies of breaking self-structure of probe and target RNA and formation of duplex between probe and target. Penta/hexanucleotide probes can be assumed to have no self-structure. Moreover, short probes facilitate differentiating between strong binding complementary and weaker binding mismatched probes by changing hybridization temperature. Due to smaller enthalpy change for short duplexes, the difference in melting temperature (Tm) between complementary and mismatched probes is larger than for long probes.

Standard microarrays and structural studies of 5S rRNA–short oligonucleotides can give sufficient binding

Initial experiments were performed on microarrays with 2'-O-methylated (2'OMe) heptanucleotide probes (37). Ribosomal 5S RNA from *Escherichia coli* was selected as target because computer programs such as Mfold (41) and RNAstructure (42) predict only 27% of the known canonical base pairs of the native structure. Hybridizations of 32P-labeled 5S RNA were performed in various conditions and results used as constraints in folding of target RNA with the RNAstructure program. In particular, the 5S RNA nucleotide complementary to the middle nucleotide of a strongly binding heptanucleotide probe was constrained not to be canonically base-paired. The best hybridization results were obtained in buffer containing 1 M NaCl, 4 mM MgCl2 and 10 mM Tris-HCl, pH 7.4 at room temperature (RT) and at 4°C, where probe binding was significantly enhanced relative to 0.15 M NaCl, 4 mM MgCl2 and 0.04 M NaCl, 10 mM MgCl2. Strong hybridization was reported to loops B and C but not loop E, which is known to be unusually stabilized by Mg2+ (43). Thus, microarrays can distinguish between loops with strong and weak local structure.

In the absence of Mg2+, *E. coli* 5S rRNA forms a different secondary structure (44,45). When 5S rRNA was mapped with chemicals and microarrays at 1 M NaCl, the patterns differed dramatically from those observed in the presence of Mg2+ (37).

The results with 5S rRNA revealed advantages and disadvantages of the microarray approach. Advantages of short probes include limited interference with target RNA secondary structure and ability to have probes representing all possible sequences. Another advantage is that probe binding often does not completely overlap with chemical modification, so that complementary information is generated. A disadvantage is that a short probe may have more than one binding site. This is particularly problematic because A-U and G-U pairs have similar stability (46–48). Replacing uridine with 2-thiouridine (s2U), however, improves discrimination between A-U and G-U because A-s2U is more stable than A-U, whereas G-s2U is less stable than G-U (37,49,50). Another disadvantage is that the binding constant for unmodified probes is very sequence dependent, which complicates interpretation. Moreover, binding constants are low for probes with high A/U content. For example, thermodynamic stabilities (free energy, ΔG°37) of 5'(AmUm)3/3'(UA)3 and 5'(GmCm)3/3'(CG)3 duplexes are −2.8 and −12.2 kcal/mol, respectively, whereas melting temperatures (in 1 M NaCl) are 11 and 71°C, respectively (calculated based on (51), (http://rnachemlab.ibch.poznan.pl/calculator2.php)). This disadvantage can be eliminated, however, by using modified probes as described below.

Isoenergetic microarrays – modifications provide roughly sequence-independent binding to simplify interpretation

Interpretation of binding is greatly simplified if all probes have the same binding constant to an unstructured RNA. One approach to equalizing binding constants to unstructured RNA is to replace some 2'-O-methylated nucleotides with locked nucleic acid (LNA) residues (51–53). Thermodynamic studies found that isolated LNA residues at

**Nucleic Acids Research, 2015, Vol. 43, No. 1 3**
internal positions enhance stabilities (ΔAG° excited) of LNA-2′OMeRNA/RNA duplexes by ca. 1.5 ± 0.5 kcal/mol per single LNA nucleotide, which increases binding constant by roughly 10-fold at 37°C. Little or no additional substitution is achieved if an adjacent LNA is added. This phenomenon may be caused by C3′-endo conformation of LNA nucleotides forcing 3′-adjacent nucleotide also to adopt C3′-endo conformation and thus facilitate formation of A-form helix (54). At terminal positions the contribution of LNA residues on stabilities is ca. 0.3–1.4 kcal/mol per LNA nucleotide. Internal LNA–RNA mismatches destabilize LNA-2′OMeRNA/RNA duplexes between 1.5 and 7 kcal/mol depending on type of mismatch and position within duplex. Destabilizations are less at terminal positions. The results revealed that it is optimal to place LNA residues at every second or even third nucleotide within a strand and to avoid placing LNA residues at terminal positions. As described below, an exception to the latter rule is that a terminal LNA-G is often added as an extra nucleotide to increase stability.

A second approach to improve thermodynamic stabilities is to replace A with 2,6-diaminopurine riboside as either a 2′-O-methyl or LNA nucleotide (55). The 2,6-diaminopurine residue interacts with uridine via three hydrogen bonds and improves stability by ca. 1 kcal/mol relative to A.

Based on many 2′OMeRNA/RNA microarray studies, the free energy change at 37°C (ΔAG°57) has to be more favorable than −6 kcal/mol to observe clearly detectable hybridization signals. For hybridization to target sequences with high A/U content it is often difficult to get ΔAG°57 more favorable than −6 kcal/mol. For this reason, a third approach for designing isoenergetic probes is to add LNA-G to the 3′-end. Model studies showed that 3′-terminal mismatches G31-A, G31-U and G31-G stabilize 2′OMeRNA/RNA duplexes by 1.6, 1.5 and 1.4 kcal/mol, respectively, at 37°C (56). Formation of 3′-terminal G31-C pairs, however, stabilizes the same model LNA-2′OMeRNA/RNA duplex by 3.4 kcal/mol, which adds ambiguity to the identification of binding site. Figure 2 shows an example of differences made when probes designed to be strongly binding and isoenergetic. When unmodified DNA probes are used, ΔAG°57 for binding to unstructured RNA is predicted to oscillate around −3.5 ± 2.5 kcal/mol. In contrast, 2′OMeRNA probes modified with LNA and 2,6-diaminopurine riboside give binding predicted to oscillate around −9 ± 1 kcal/mol (Figure 2).

There are 1024 different pentamer sequences. On the basis of thermodynamic predictions, it was possible to synthesize around 850 penta/hexanucleotide LNA-2′OMeRNA/RNA relatively isoenergetic probes with ΔAG°57 values favorable enough for use in microarrays. Of the 850 probes, around half were with 3′-terminal LNA G. Probes contained a 5′-end aminohexyl linker allowing them to covalently bind to oxidized agarose-coated microscope slides (57,58). Oligonucleotide spots with diameters of ca. 0.15 mm and separated by 0.75 mm are automatically printed on a functionalized microscope slide. Thus, ~2000 spots can be spotted on a single microarray. Usually spots are printed in triplicate, including several controls and reference probes. Therefore, two microarrays can be designed to probe essentially any RNA.

Studies of retrotransposon R2 5′RNAs—large RNAs can be studied. The first structural studies using isoenergetic microarrays were performed on the roughly 330-nt 5′-regions of R2 retrotransposon RNA from Bombyx mori (R2Bm), Samia cynthia (R2Sc), Coscinocera hercules (R2Ch), Callosamia promethea (R2Cpr) and Saturnia pyri (R2Spy) (Figure 3) (34,36). Initially, only the B. mori sequence was available and the secondary structure was unknown. Hybridizations were performed in the presence of 3 mM MgCl2 with 0.2 M or 1 M of NaCl and at 4°C or RT. Optimal results were obtained at 0.2 M NaCl, 5 mM MgCl2, 10 mM Tris-HCl, pH 8, RT. In principle, mapping an RNA containing ‘n’ nucleotides with shifting of probes by one position requires n−4 probes. In reality, the number of probes is lower because only probes with free energies of hybridization more favorable than −6 kcal/mol are synthesized. For example, studies of R2Bm that contains 323 nucleotides used 232 instead of 319 probes. The average predicted free energy (ΔAG°57) of binding of the modified probe library to complementary sites of the R2Bm 5′ RNA if the sites were single stranded was −9.8 ± 1.2 kcal/mol. Based on microarray mapping, a secondary structure of R2Bm 5′ RNA was predicted (36). After omitting 27 probes with alternative binding sites, only seven probes with unambiguous binding sites were used as constraints. The same target RNA was chemically mapped with dimethyl sulfate (DMS), 1-cyclohexyl-(2-morpholinoethyl)carbodiimide metho-p-toluene sulfonate (CMCT), 1,1-Dihydroxy-3-ethoxy-2-butane (kethoxal) and N-methylisatoic anhydride (NMIA). The same RNA secondary structure was predicted by RNAstructure from thermodynamics with microarray constraints alone as with constraints from both microarray and chemical mapping data (36). Interestingly, fragment 50–123 of R2Bm 5′RNA was not hit strongly by chemical reagents even though there were many nucleotides predicted to be single stranded. This suggested alternative folding to a pseudoknot. At the time, pseudoknots were not allowed in algorithms predicting secondary structure on the basis of thermodynamics and experimental constraints. Nuclear magnetic resonance spectra, however, proved the pseudoknot folding (59). The pseudoknot and most of the rest of the secondary structure was confirmed by sequence comparison when four additional sequences became available (34). One test of a structure is whether it has reasonable binding sites for strongly binding probes not used as constraints because there are more than one potential binding site. With exception of probes 227 and 228, all 26 probes with multiple potential binding sites have at least one site in the proposed R2Bm 5′RNA secondary structure that would obviously be a good binding site. Probes 227 and 228 would be expected to bind if the top of helix P7b can shift up or down by one nucleotide to form an internal C-A pair. Equivalent studies of the other four R2 5′RNAs revealed conserved structures of hairpins P1, P6, P7 and P8, and the pseudoknot containing P2, P3, P4 and P5.

Binding of oligonucleotide probes identifies regions of weak or no intramolecular base pairing in RNA. Interestingly, of 14 probes with unambiguous or very probable binding sites in R2Bm 5′RNA, only one of the central target
nucleotides is strongly modified during chemical mapping. It demonstrates again that microarray mapping can provide information complementary to that available from chemical mapping.

Studies of models for RNA structural motifs—invansion of terminal helixes. Studies of model systems with hairpins, internal loops, bulges, 3′- and 5′-dangling ends and pseudoknots revealed some factors that influence binding (32). The models had short helices because natural RNA helices are typically short, containing ca. 6–8 base pairs. All model RNAs were UV-melted to determine thermodynamic stability to assure that each model RNA folded exclusively into one defined structure under chosen hybridization conditions. The models were hybridized to isoenergetic microarrays in various conditions, including different concentrations of NaCl, MgCl₂ and temperature.

Binding to hairpin loops was observed, as expected. However, binding to terminal helixes was also observed often, usually by extension of binding to nucleotides not canonically paired (Figure 4A and B). Interestingly, helix invasion was more likely from a 5′ than a 3′ trinucleotide dangling end. This trend could reflect the larger helix stabilization provided by a 3′-dangling end relative to a 5′-
Figure 4. Binding of isoenergetic oligonucleotide probes to model hairpins and pseudoknot (32). In hairpin IL1 (A) loop b terminates helix I, whereas in hairpin IL2 (B) loop c closes helix II. Inversion of the element terminating a helix correlates with inverting binding ability to the adjacent helical region. (C) Secondary structure and hybridization results for 74-pseudoknot (Pk2), a part of B. mori R2 5′RNA. Hybridization was performed in 200 mM NaCl, 5 mM MgCl2 and 10 mM Tris-HCl, pH 8.0 at room temperature. Pk2 was hybridized to a microarray with all possible isoenergetic oligonucleotide probes complementary to Pk2. Filled in squares indicate site of strong binding to isoenergetic probe (middle nucleotide of region binding to pentamer probe), open squares indicate site of medium binding to isoenergetic probe. See the legend for meaning of other symbols.

dangling end. For example, a 5′-unpaired UCU is predicted to enhance free energies of a helix (ΔG37°) by 0.1–0.4 kcal/mol at 37°C, whereas a 3′-unpaired UUC stabilized by 1.0–1.9 kcal/mol (60,61). Two pseudoknots were also studied, one of the 74 nucleotides from the B. mori R2 retrotransposon (Figure 4C) and a second of 32 nucleotides composed of two RNA hairpins (32,34). In both cases probe binding was observed to single-stranded regions. Surprisingly, the isolated R2 pseudoknot bound 11 probes (Figure 4C) compared to only one in the 323 fragment (Figure 3). It was also much more reactive to chemicals.

The analysis of RNA structural motifs binding to microarray probes demonstrates the following features: (i) bulges, internal loops and dangling ends bind strongly to microarray probes; (ii) a terminal stem can also bind strongly, presumably due to invasion of stem by the modified oligonucleotide; (iii) internal stems do not typically bind strongly. Along with further results described below, the data demonstrate that the location of the structural motif in the context of the whole target structure and both secondary and tertiary interactions influence the accessibilities of RNA to probe binding (32,33,62,63).

Studies of RNase P RNA—binding to an internal helix but not to a large internal loop. Microarray studies (62) on the 154-nucleotide construct of the specificity domain of RNase P RNA from Bacillus subtilis (RNrspBs) that was used to solve the crystal structure (64) revealed unexpected probe binding (Figure 5). While three of the four hairpin loops at most bind probes as expected, a terminal (P7) and an internal (P8) helix also bind strongly. Most surprisingly, a large internal loop at most possibly bound only two ambiguous probes. This RNA has a complex crystal structure and chemical mapping in solution is consistent with this structure. There are several potential reasons for the unexpected binding, including oligonucleotide trapping of minor species (i.e. structural rearrangement (65)) and helix invasion stabilized by probe modifications and/or triple helix formation.

Many approaches are being developed to target RNA with oligonucleotide mimics as therapeutics. It would be ad-
vantageous to develop approaches that use short oligonucleotides. The results for RNA spBs show that even a crystal structure does not allow obvious design of suitable oligonucleotides. Microarray experiments, however, rapidly reveal binding sites in target RNAs. The results suggest also that short oligonucleotides may induce refolding of RNA into an alternative conformation that could abrogate function.

Studies of influenza RNA—identifying potential sites for chemical genetics and therapeutics. Isoenergetic microarrays are being used to probe sequences in influenza A that are predicted to fold into stable, conserved structures (66–69). In the presence of Mg\(^{2+}\), the 3′-splice site of influenza A segment 7 mRNA is in equilibrium between a hairpin and a pseudoknot conformation. In the two conformations, the splice site and other functional elements exist in different structural environments, suggesting that splicing can be controlled by a conformational switch that exposes or hides the splice site similarly as postulated for HIV-1 RNA (65,70). Only the hairpin is present in the absence of Mg\(^{2+}\) and only the pseudoknot is present when Mg\(^{2+}\) is replaced with Co(NH\(_3\))\(_6\)\(^{3+}\). Microarray probing of the two conformations revealed different patterns of binding, suggesting that splicing could be regulated by oligonucleotides. Because both proteins encoded in segment 7 are essential, the splice site is a potential therapeutic target.

The functions of two other structures predicted by bioinformatics (71) are unknown. Microarray results identify sites of oligonucleotide binding, suggesting that a ‘chemical genetics’ approach with modified oligonucleotides could help reveal the function of these structures.

Model studies of tRNAs—small molecules but a lot of information. Microarray mapping of tRNAs revealed the ability of microarrays to detect positions of modifications that affect base pairing and also subtle differences in local structure.

In nature, only initiator tRNA\(^{\text{Met}}\) can denote the first amino acid in polypeptide synthesis and only the elongator tRNA\(^{\text{Met}}\) allows extension of polypeptide. Studies on initiator tRNA\(^{\text{Met}}\) and elongator tRNA\(^{\text{Met}}\) from *Lupinus luteus* showed that differences in structure of those tRNAs produce different patterns of probe binding (Figure 6A and B) (33). One difference is accessibility of the D loop in tRNA\(^{\text{Met}}\) and lack of binding to the D loop of tRNA\(^{\text{Met}}\). This reflects the lack of poorly pairing dihydroU nucleotides in the D loop of tRNA\(^{\text{Met}}\).

A second difference concerns binding to the anticodon loop. The binding sites of tRNA\(^{\text{Met}}\) are 35 and 38, whereas in tRNA\(^{\text{Met}}\) binding sites are 32 and 34. Generally, the 5′-side of the anticodon loop of tRNA\(^{\text{Met}}\) is accessible for binding, whereas in tRNA\(^{\text{Met}}\) only the central and 3′-side are accessible. Structural changes induced by MgCl\(_2\) also change microarray binding. For example, high concentration of NaCl (1 M) and 3 mM MgCl\(_2\) in hybridization buffer is required for binding at sites 52 and 56 of the tRNA\(^{\text{Met}}\) TΨC loop. The tRNA\(^{\text{Met}}\) TΨC loop is more accessible for probes.

Studies on *Saccharomyces cerevisiae* tRNA\(^{\text{Phe}}\) and its unmodified RNA transcript provide another demonstration of microarrays revealing positions of modifications (Figure 6C) (63). Differences in binding were found in each hairpin loop, all of which are modified in natural tRNA\(^{\text{Phe}}\). Generally, fewer binding sites and weaker binding were observed in natural tRNA\(^{\text{Phe}}\) than its unmodified transcript. RNA modifications such as dihydropyrimidine (D), N\(^2\)-methylguanosine (m\(^2\)G) and N\(^2\),N\(^2\)-dimethylguanosine (m\(^2\)G\(^2\)) in internal positions influenced thermodynamic stabilities of hybridization duplexes (72,73). Not every change in probe binding, however, can be explained by reduced duplex stability. Part of the differences could originate from tertiary interactions between D and TΨC loops in tRNA\(^{\text{Phe}}\) and different interactions in transcript tRNA\(^{\text{Phe}}\). Comparison of tRNA\(^{\text{Phe}}\) mapping with chemical and enzymatic methods provides similar structural conclusions (74–76).

Studies of RNA within binary and tertiary complexes

Microarray mapping can also be used to study RNA in complex with protein and/or other RNAs (77). The first models for such studies were non-coding DsrA RNA and OxyS RNA bound to Hfq protein. In principle, only single-stranded fragments of RNAs not involved in any type of interactions will bind to probes (Figure 7). Formation of complexes can thus change probe binding patterns in a microarray.

Formation and stabilities of complexes were first analyzed by gel electrophoresis. Then hybridizations were performed with DsrA alone and with Hfq. The microarray mapping of isolated DsrA confirmed its published secondary structure (78,79). As shown in Figure 7, microarray mapping in the presence of protein reduced microarray binding to sequences expected to bind protein. Similar experiments with 25- and 140-nt fragments of rpoS mRNA (rpoS25 and rpoS140) and Hfq revealed an Hfq-induced change in rpoS secondary structure (Figure 7) (77).

Inside cells, DsrA RNA regulates translation of many mRNAs, including rpoS mRNA. It was postulated that the regulation is achieved by interactions of DsrA RNA and the 5′UTR of rpoS mRNA (80). Hybridization of the rpoS140/DsrA RNA complex to probes on isoenergetic microarrays revealed binding sites and demonstrated that the interaction region of DsrA RNA and rpoS mRNA is longer than previously reported. Microarray studies of the DsrA RNA/Hfq/rpoS 140 RNA revealed that some probes no longer bind and that the main Hfq binding site is probably displaced to another site of DsrA RNA and that Hfq interacts with rpoS mRNA outside the region of complementarity to DsrA RNA.

Other approaches to generating isoenergetic microarrays

Isostable DNA microarrays. The Richert group in 2007 developed DNA modifications that could be used to make ‘isostable DNA microarrays’ (81). To make ‘isostable DNA’, they introduced pyrene or anthraquinone derivatives into DNA probes and in some cases trimethoxystilbene at the 5′-terminus. Moreover, LNA, particularly LNA-thymidine and LNA-adenosine were incorporated at terminal positions. To reduce thermodynamic stability of G-C pairs, 2′-deoxyguanosine or 2′-deoxycytidine was replaced with 2′-deoxyinosine or 2′-deoxy-N4-ethylcytidine, respectively.
**Figure 6.** Isoenergetic microarray mapping results for tRNA\textsuperscript{Met} \textsuperscript{A} and tRNA\textsuperscript{Met} \textsuperscript{B} in 1 M NaCl, 20 mM sodium cacodylate, 0.5 mM Na\textsubscript{2}EDTA pH 7.5 (yellow bars), 1 M NaCl, 0.5 mM MgCl\textsubscript{2}, 10 mM Tris-HCl pH 7.5 (red bars) and 1 M NaCl, 3.0 mM MgCl\textsubscript{2}, 10 mM Tris-HCl pH 7.5 (blue bars). The height of the bar corresponds to strength of probe binding. Isoenergetic microarray mapping results for tRNA\textsuperscript{Phe} \textsuperscript{C} (red bars) and unmodified transcript (blue bars) in 1 M NaCl, 0.5 mM MgCl\textsubscript{2}, 10 mM Tris-HCl, pH 7.5 at 4°C. The height of the bar corresponds to strength of probe binding (33,63).

**Figure 7.** (A) Isoenergetic microarrays mapping results for DsrA sRNA in complex with Hfq. On the left side, microarrays with hybridized DsrA RNA and complex of DsrA RNA/Hfq. On the right side, the bars are related to intensity of binding to probes (at selected positions) of DsrA RNA in the absence and presence of Hfq. On secondary structure of DsrA RNA red squares represent sites of binding that have the same intensity in the absence and presence of Hfq. (B) On the left side, mapping with isoenergetic microarrays DsrA RNA alone, complex DsrA RNA/rpoS mRNA and complex DsrA RNA/rpoS mRNA/Hfq. On the right side, schematic presentation of the interactions of DsrA RNA, rpoS mRNA and Hfq protein. The bar graph demonstrates intensity of binding to probes (at selected positions) of DsrA RNA in the absence and presence of Hfq. Various patterns hybridization of DsrA RNA is related with different positioning probes (77).
Presumably, these DNA microarrays could also be used to probe RNA, although different design rules would be necessary because the sequence dependence of DNA/RNA and DNA/DNA duplexes differs (40,82). Moreover, DNA probes are less likely to maintain RNA conformations.

Unmodified 2′-O-methylated microarrays. The Verdine group (83) used microarrays fabricated by photolithography and consisting of all possible unmodified 2′-O-methylated tetramers to octamers. They tested their method on the 451-nucleotide fluorescently labeled human telomerase RNA (hTR) and on shorter structures including two pseudoknots. Consistent with previous studies (21,22,24,37,62), probes that bound to single-stranded regions could invade adjacent helices, especially those formed between the 5′ and 3′ ends of an RNA.

SUMMARY OF LESSONS FROM PUBLISHED STUDIES

Isoenergetic oligonucleotide microarray mapping is based on interactions of oligonucleotide probes with target RNAs. Hybridizations can be performed over a wide range of buffer, cations, temperature, time and pH conditions. This contrasts with enzymatic mapping that requires conditions favorable for enzyme activity, e.g. the presence of specific divalent cations and limited pH. The same concerns chemical mapping since to react with target RNA the chemical reagents (DMS, CMCT, kethoxal, NMIA, 1M7) require conditions that could influence the structure of RNA. For example, some reagents used for chemical mapping require presence in reaction mixture of ethanol (for CMCT and kethoxal) or dimethylsulfoxide (for NMIA).

Moreover, microarray mapping is simple and fast. The 850 penta/hexanucleotide probes that form hybridization duplexes thermodynamically stable enough for mapping can be multiply printed on a large number of single microscope slides that can then be used to map any RNA.

Of course, isoenergetic microarrays have limitations. Pentanucleotide probes can have multiple complementary binding sites, particularly in large RNAs. Moreover, U binds to A and G with roughly equal stability (47). Additionally, probes can occasionally bind to canonically paired helices, especially those formed by the 5′ and 3′ ends of the RNA (37,62,83). All those effects complicate interpretation. Based on published experiments, microarray mapping is particularly useful for RNAs containing up to 300–500 nucleotides. Unfortunately, isoenergetic microarrays are not yet commercially available.

FUTURE IMPROVEMENTS AND POTENTIAL

Isoenergetic oligonucleotide microarrays can now quickly identify regions accessible to bind oligonucleotides, which are therefore potential targets for therapeutics or chemical genetics experiments. In concert with other methods, they can also reveal regions protected by strong secondary, tertiary and quaternary interactions, including structures and interactions in complexes with proteins and RNAs. Potentially, comparisons between isoenergetic microarray results from in vitro transcribed RNA and RNA isolated from natural sources could be used to search for natural modified nucleotides within single-stranded regions of RNA. It should be particularly easy to detect modifications of functional groups involved in base pairing (such as N-alkyl, thio-modified nucleotides). Comparisons between microarray binding and secondary structures predicted by thermodynamics guided by chemical mapping can reveal unusual folding. For various reasons, an RNA may fold into multiple structures or its folding may be determined by kinetics rather than thermodynamics. In such cases, comparison of chemical mapping and microarray binding could discriminate between very different potential structures. For example, it has been proposed that encapsidated Satellite Tobacco Mosaic Virus (STMV) has a single structure (84) or an ensemble of structures (85,86). Very different binding to a microarray is predicted for each possibility.

Of course, there is a lot of room for improvements, including: (i) application of oligonucleotide probes containing 2-thiouridine residues to improve selectivity of base pairing with A in comparison to G (37,49,50), (ii) use of 3′-pyrene terminated oligonucleotide probes. The presence of pyrene enhances stability of LNA-2′MeRNA(probe)/RNA(target) duplexes by 2.3 ± 0.1 kcal/mol, independent of nucleotide in the opposite strand and the sequence of adjacent base pairs (56), (iii) development of computer programs for automatic analysis of hybridization results, including application of pseudo-ΔG’ restraints and evaluation of alternative binding sites and (iv) use of methods not requiring radioactivity to detect hybridization of target RNA to microarrays. These include fluorescence quenching (31), other label free methods (87) or mass spectrometry (88).

Applications of microarrays could be expanded to include probing more complex folding. For example: (i) tertiary interactions within RNA could be probed with spacer-linked oligonucleotides simultaneously complementary to different fragments of target RNA, and (ii) sequences allowing formation of bimolecular triplexes, quadruplexes and parallel duplexes could be identified.

FUNDING

National Science Center [UMO-2013/08/M/NZ1/01062 to E.K.; UMO-2011/03/B/NZ1/00576, UMO-2011/03/B/ST5/01098, UMO-2013/08/A/ST5/00295 to R.K.]; National Institutes of Health [RO3TW008739 to E.K., D.H.T.; GM22939 to D.H.T.]; Founding for open access charge: National Science Center.

Conflict of interest statement. None declared.

REFERENCES

1. Pirrung,M.C. and Southern,E.M. (2014) The genesis of microarrays. Biochem. Mol. Biol. Educ., 42, 106–113.
2. Maskos,U. and Southern,E.M. (1992) Oligonucleotide hybridizations on glass supports—a novel linker for oligonucleotide synthesis and hybridization properties of oligonucleotides synthesized in situ. Nucleic Acids Res., 20, 1679–1684.
3. Augenlicht,L.H. and Kobrin,D. (1982) Cloning and screening of sequences expressed in a mouse colon-tumor. Cancer Res., 42, 1085–1093.
4. Kulesh,D.A., Clive,D.R., Zarlinga,D.S. and Greene,J.J. (1987) Identification of interferon-modulated proliferation-related cDNA sequences. Proc. Natl. Acad. Sci. U.S.A., 84, 8453–8457.
87. Pan, S.L. and Rothberg, L.J. (2003) Interferometric sensing of biomolecular binding using nanoporous aluminum oxide templates. *Nano Lett.*, 3, 811–814.

88. Zhang, X.Y., Zhu, S.C., Deng, C.H. and Zhang, X.M. (2012) An aptamer based on-plate microarray for high-throughput insulin detection by MALDI-TOF MS. *Chem. Commun.*, 48, 2689–2691.