1368–1373 Nucleic Acids Research, 1994, Vol. 22, No. 8 © 1994 Oxford University Press

Arrays of complementary oligonucleotides for analysing the hybridisation behaviour of nucleic acids

E.M.Southern*, S.C.Case-Green, J.K.Elder, M.Johnson, K.U.Mir, L.Wang and J.C.Williams
Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

Received February 8, 1994; Revised and Accepted March 9, 1994

ABSTRACT

Arrays of oligonucleotides corresponding to a full set of complements of a known sequence can be made in a single series of base couplings in which each base in the complement is added in turn. Coupling is carried out on the surface of a solid support such as a glass plate, using a device which applies reagents in a defined area. The device is displaced by a fixed movement after each coupling reaction so that consecutive couplings overlap only a portion of previous ones. The shape and size of the device and the amount by which it is displaced at each step determines the length of the oligonucleotides. Certain shapes create arrays of oligonucleotides from mononucleotides up to a given length in a single series of couplings. The array is used in a hybridisation reaction to a labelled target sequence, and shows the hybridisation behaviour of every oligonucleotide in the target sequence with its complement in the array. Applications include sequence comparison to test for mutation, analysis of secondary structure, and optimisation of PCR primer and antisense oligonucleotide design.

INTRODUCTION

Arrays of synthetic oligonucleotides can be used in a number of ways to analyse a target sequence by hybridisation. We have previously described two different kinds of array: complete sets of oligonucleotides which can be used to obtain sequence information when nothing is known about the target sequence (1); and dedicated arrays comprising lines of allele-specific oligonucleotides (ASOs) which can be used to analyse several targets against several ASOs simultaneously (2).

We now describe a third kind of array, which can be used to analyse the hybridisation behaviour of a sequence to a complete set of complementary oligonucleotides. We describe how the arrays of complementary oligonucleotides are made, how they are used to analyse hybridisation behaviour, software for qualitative and quantitative analysis, and some typical results.

BASIS OF THE METHOD

Structure of the arrays

The arrays comprise the oligonucleotides complementary to the sequence of a target to be analysed, for example, a region of a gene to be scanned for mutations. The series of oligonucleotides is made by coupling the bases to a solid surface in the order in which they occur in the complement of the target sequence. The reagents used to synthesise the oligonucleotide are applied in a patch to the surface of the solid subtrate. (We use glass plates which have been coated with a linker derived from hexaethylene glycol (3).) If the template used to create the patch were kept in the same place during all the synthetic cycles, the result would be a complete complement of the target sequence. If, however, the template is moved along the surface after each coupling, the result is a series of oligonucleotides, each one complementary to a region of the target sequence (Fig 1). The length of oligonucleotides synthesised depends on the ratio of the diameter of the template to the displacement at each coupling step. For example, a template of length 30 mm will produce 10-mers, 12-mers or 15-mers using offsets of 3 mm, 2.5 mm or 2 mm respectively. The length of the oligonucleotides is the number of bases that have been coupled at the point where the back of the template passes the front. As a result, it is possible to create arrays comprising sets of oligonucleotides of all sizes from monomers up to the largest in a single series of couplings, by using a template of appropriate shape. For example, a diamond shaped template creates a series of smaller diamond shaped patches (Fig. 1, *a* to *c*); at the centre, the longest oligonucleotides will be made, corresponding in length to the diagonal divided by the offset. At the edge the diamonds will contain monomers. Between the centre and the edge will be all lengths from the longest to monomers. A circular template also creates arrays of all lengths up to the longest, equal to the diameter divided by the offset. However, the cells created by a circle differ in shape: at the centre line, they are lenticular, but off this line, they form

*To whom correspondence should be addressed

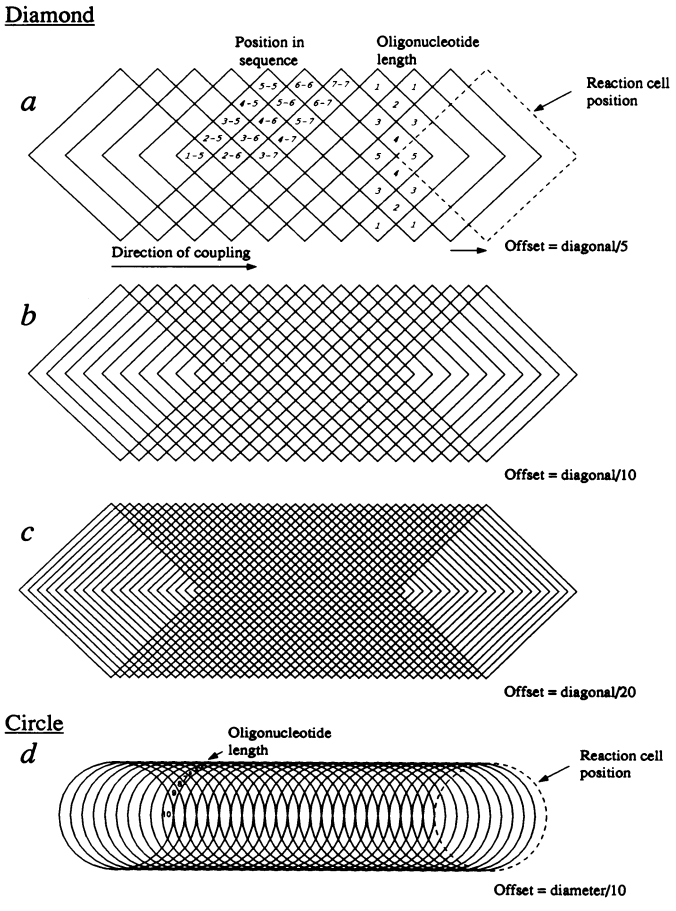

Figure 1. Construction of scanning arrays. Two template shapes are illustrated here: a 'diamond' shape created by turning a square through 45° and a circle. In either case, a cell is created by pressing a seal of the desired shape against the substrate on which synthesis takes place. Reagents, washing solutions and gas are introduced at the bottom of the cell, and removed from the top or the bottom. These processes can be carried out automatically by coupling the cell to an automatic synthesiser such as the ABI 381A. After each coupling, the cell is moved along the substrate by a predetermined offset. For a cell of fixed dimension, longer oligonucleotides in smaller areas are made by reducing the amount of the offset. A further effect is to shorten the overall length of the plate covered by the array (Compare *a*, *b* and *c*, where seven 5-mers, 12 10-mers or 23 20-mers are made in the same length.) Along the centre line, the length of oligonucleotide that is synthesised is equal to the diameter of the cell divided by the offset. Oligonucleotides of successively shorter lengths, down to mononucleotides, are created in the cells flanking the centre line.

a four-cornered 'spearhead' that diminishes in scale towards the edge (Fig. 1*d*).

MATERIALS AND METHODS

Apparatus for making arrays

The device used to form the cell must make a good seal against the substrate on which synthesis takes place. We use glass for the substrate. Teflon is the only material we have found which makes a good seal to glass and withstands the solvents and chemicals used in oligonucleotide synthesis. The surface of the cell must be finished before machining the walls that create the reaction chamber, and we have been unable to make a diamond shaped template that would create a good seal when pressed against a glass surface because tool marks on the surface run across the edge of the template wall. However, a circular template can be cut in a lathe in such a way that the surface tooling runs in the same lines as the cuts used to form the wall, enabling the formation of an effective seal when it is pressed against the glass surface (Fig. 2*a*). The depth of the cell is ca. 0.5 mm. Inlet and outlet ports were made by drilling 1.0 mm diameter holes at the top and bottom of the circle and fitting sawn-off 19 gauge (1.1 mm o.d.) syringe needles through from the back. The Teflon cell was mounted on a rail fitted with a lead screw that was used to displace the glass plate relative to the Teflon reaction cell (Fig. 2*b*). A 'G' clamp, fitted with a dished polyethylene cushion, was fixed to the rail to apply pressure to the glass plate and form the seal to the reaction cell. The rail was mounted on the front of the frame of an ABI 381A oligonucleotide synthesiser so that the delivery lines normally connected to the column could be connected to the reaction cell. Our present system is semi-automatic. The ABI 381A is used to deliver reagents automatically during each coupling cycle, but the glass plate is shifted along by turning the lead screw by hand (Fig. 2*c*). It is important to have a reference point to know the position of each oligonucleotide in the array. We start the coupling with the wall of the cell at one end of the glass plate. All other positions can then be related to this point.

Making arrays

Glass plates (50×220×3 mm) were first coated with a covalently attached linker (3). Plates were immersed in a mixture of glycidoxypropyl trimethoxysilane/ diisopropylethylamine/ xylene (17.8 : 1 : 69, by volume), heated to 80°C and held at this temperature for 9 h, and then washed in ethanol and ether. In a second step, the plates were heated in neat hexaethylene glycol, containing a catalytic amount of sulphuric acid, at 80°C for 10 h, washed with ethanol and ether, air dried and stored at −20°C. Oligonucleotide synthesis used standard reagents for phosphoramidite chemistry, omitting the capping step. The ABI 381A was programmed to couple bases in the order corresponding to the complement of the target sequence, with an interrupt after deprotection. The scale was for 0.2 mmol synthesis, adjusted slightly to provide volumes that would just fill the reaction chamber. Some steps were shortened to speed up the procedure (Table 1).

Final deprotection in 30% ammonia was carried out in a specially constructed bomb, comprising a chamber (230×230×20 mm) cut into a Nylon block (300×300×30 mm), sealed by a sheet of silicone rubber (3 mm thick), compressed against the rim of the chamber by clamping the whole assembly between two mild steel plates (6 mm thick) using four bolts along each side of the square. After 5−8 h at 55°C the bomb was cooled to 4°C before opening. The plate was then washed in ethanol followed by Tris/EDTA (0.01M, pH 7.8, 0.1% SDS) and ethanol and then dried in an air stream.

HYBRIDISATION REACTIONS

We have used a variety of target molecules in experiments with scanning arrays: synthetic oligonucleotides labelled using polynucleotide kinase with γ-^{32}P, γ-^{33}P or γ-^{35}S-ATP to tag the 5′ end; RNAs labelled at the 3′ end using RNA ligase with 5′-^{32}P cytosine-3′,5′-diphosphate; or transcripts of DNA fragments made from PCR amplified fragments using T7 or SP6 polymerase to incorporate α-^{32}P or α-^{35}S UTP. All of these

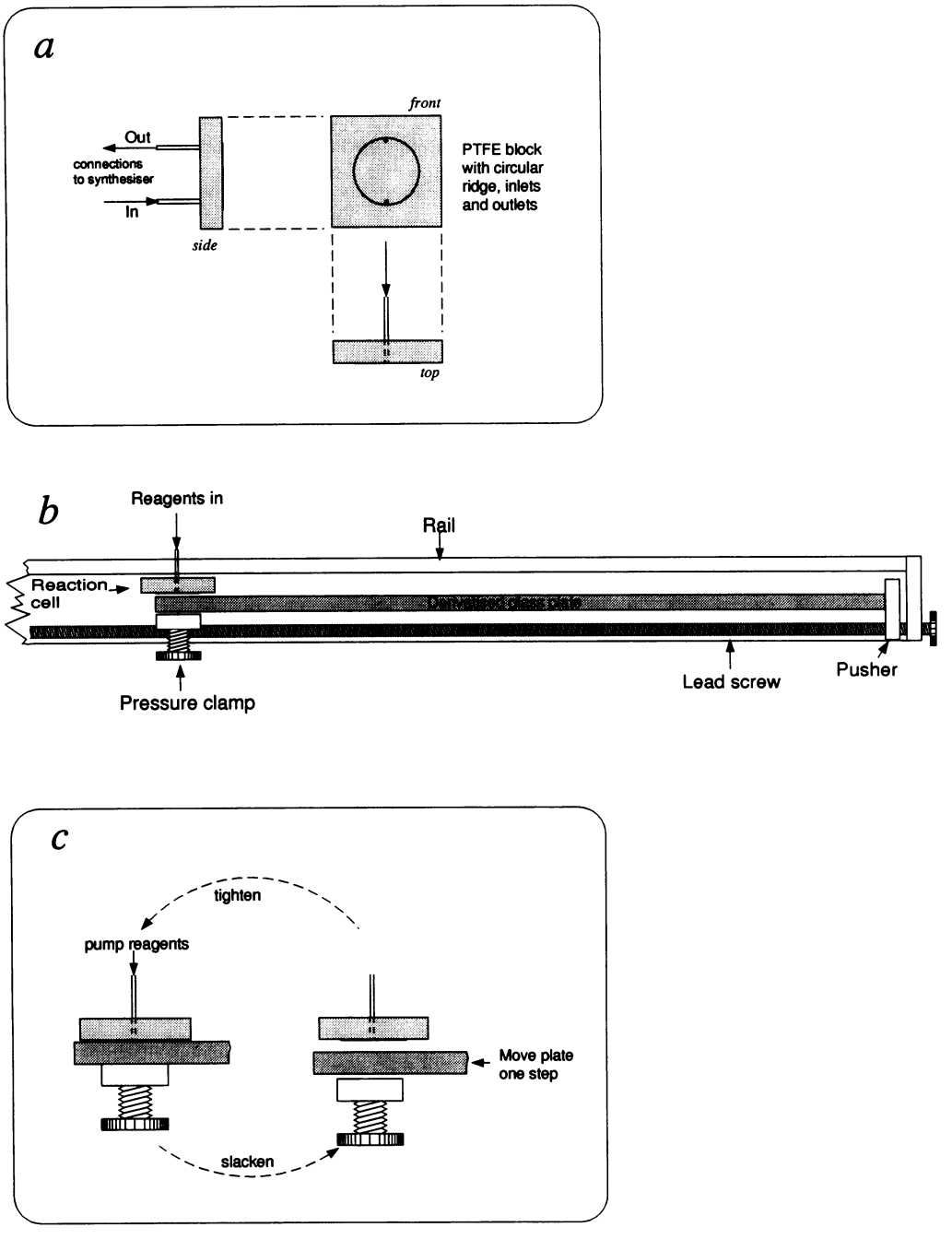

Figure 2. Apparatus used for applying reagents to glass plate. *a*. The Teflon reaction cell consists of a block, 50×50 mm, with a circular ridge, 0.5 mm height×30 mm diameter. Two holes, 1.0 mm diameter are drilled inside the top and bottom of the circle to take the shortened 19G syringe needles used to connect the reaction cell to the oligonucleotide synthesiser. The assembled cell is fixed to the rail which carries the lead screw shown in *b*. *b*. The apparatus is mounted on a rail, an 'L' section bar, that is fixed to the front of the ABI 381A oligonucleotide synthesiser. The lead screw, 1 mm pitch, is fitted with a 'pusher' that drives the glass plate across the front of the reaction cell. Its knob is marked at half turns, to make it easy to drive the plate forward in half millimeter steps. The pressure clamp is a modified carpenter's 'G' clamp, fixed to the back of the rail, and with a polyethylene cushion mounted on its pressure pad. Moderate hand tightening is enough to seal the glass plate against the reaction cell without breaking it. *c*. One coupling cycle comprises: clamping the plate up to the reaction cell; activating the synthesiser to go through the preprogrammed cycle to couple the appropriate nucleotide; slackening the clamp and pushing the plate one offset by turning the lead screw.

make good hybridisation probes. Most hybridisation reactions were carried out at 4–25°C, in solutions containing 3–4.5M TMACl or 1.0M NaCl. After hybridisation, the plate was rinsed in the hybridisation solvent and exposed through Clingfilm to a storage phosphor screen (Fuji STIII) which was then scanned in a Molecular Dynamics 400A PhosphorImager.

EXPERIMENTS AND RESULTS

We have tested the 'scanning' arrays in several applications. We find that the arrays are relatively easy to fabricate using the simple apparatus described. Each coupling step takes about 8 min, and at each step the operator spends about 1 min moving and clamping

Table 1. Program for ABI381A to deliver reagents for one coupling cycle. At the interrupt, the operator moves the plate one offset, and restarts the program.

Step Number	Function Number	Function Name	Step
1	10	#18 to waste	5
2	9	#18 to column	30
3	2	Reverse Flush	10
4	1	Block Flush	5
5	28	Phos Prep.	3
6	90	Tet to Column	5
5	28	Phos Prep.	3
6	90	Tet to Column	5
7	19	B+Tet to column	3
8	90	Tet to column	3
9	19	B+Tet to column	3
10	90	Tet to column	3
11	19	B+Tet to column	3
12	90	Tet to column	3
13	9	#18 to column	2
14	4	Wait	
15	10	#18 to waste	5
16	2	Reverse flush	5
17	1	Block flush	5
18	81	#15 to waste	5
19	13	#15 to column	23
20	10	#18 to waste	5
21	4	Wait	
22	2	Reverse flush	10
23	1	Block flush	5
24	10	#18 to waste	5
25	9	#18 to column	15
26	2	Reverse flush	5
27	9	#18 to column	15
28	2	Reverse flush	5
29	9	#18 to column	15
30	2	Reverse flush	5
31	9	#18 to column	15
32	2	Reverse flush	5
33	1	Block flush	5
34	33	Cycle entry	1
35	10	#18 to waste	5
36	9	#18 to column	20
37	2	Reverse flush	5
38	1	Block flush	5
39	5	Advance FC	1
40	6	Waste port	1
41	82	#14 to waste	5
42	14	#14 to column	40
43	10	#18 to waste	5
44	9	#18 to column	30
45	1	Block flush	5
46	7	Waste bottle	1
47	2	Reverse flush	5
48	10	#18 to waste	5
49	9	#18 to column	30
50	2	Reverse flush	5
52	1	Block flush	5
53	34	Cycle end	1

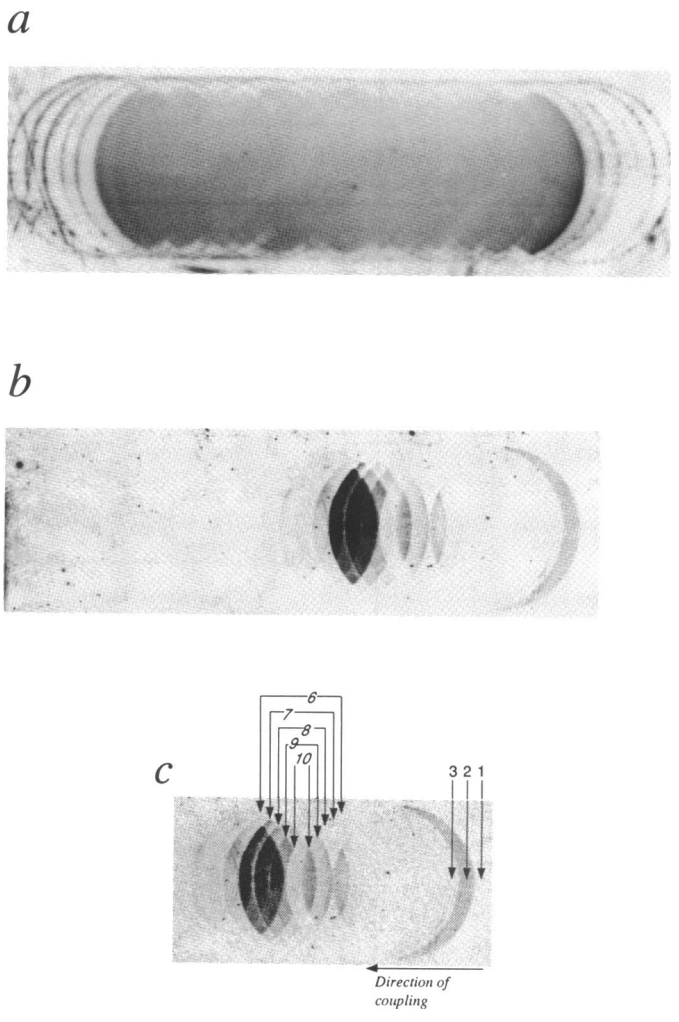

Figure 3. Hybridisation of an oligopyrimidine and a RNA to scanning arrays. *a.* Hybridisation of a sequence of pyrimidines, **5′-ctctctctttccctttccttctctctctt-tcccttccttc**, to an array of complementary oligopurines based on the sequence **3′-agagaaagggaaggaaggagagagaaagg**. *b.* Hybridisation of a 528-base transcript of exon 10 of the CFTR gene to an array representing bases 287–305, **3′-ggtaatttattttatagtagaaacca.** In both experiments, the reaction cell was 30 mm diameter and the offset 3 mm, giving rise to decanucleotides along the centre line. *c.* The hybridisation pattern shown in 3*b* has unexpected features. Coupling started on the right of the plate, so the bases in the crescent shape at position 2 are the dinucleotide GG; adding a T in the third position abolishes the 'hybridisation' of the target. Further along the plate, it can be seen that a shift of one position along the sequence can cause a fall from strong to negligible interaction, and in one position, the 7-, 8- and 9-mers all interact more strongly than the 10-mer.

the plate. Since the sequence to be constructed is typed into the synthesiser, errors in entering the sequence are unlikely. The only operator error we have recorded is a failure to shift the plate between couplings. This can be checked by recording the distance travelled by the plate and comparing it with that expected from the total number of couplings.

As the amount of reagent coupled is ca. 1 nmol, it is difficult to check the course of synthesis from the trityl yields. However, several considerations suggest that stepwise yields must be high. First, the coupling reagents are in 200-fold greater excess over the substrate as compared with conventional solid phase synthesis on a CPG column; for 1 nmol, we use quantities that would normally be used to make 0.2 mmol. Second, it is a feature of the protocol that the synthesis of any oligonucleotide on the array is dependent on several others; each nucleotide residue in a decamer for example will occur in ten other oligonucleotides. Thus successful synthesis of any one oligonucleotide, indicated by a positive hybridisation signal, will indicate successful coupling in these others. Third, we have tested efficiency of coupling by 'capping' the oligonucleotide ends either with ^{14}C labelled acetic anhydride, or with γ-^{32}P–ATP and polynucleotide kinase. Both of these tests show that coupling yields are high. Nevertheless, we cannot be sure that every coupling step has worked efficiently.

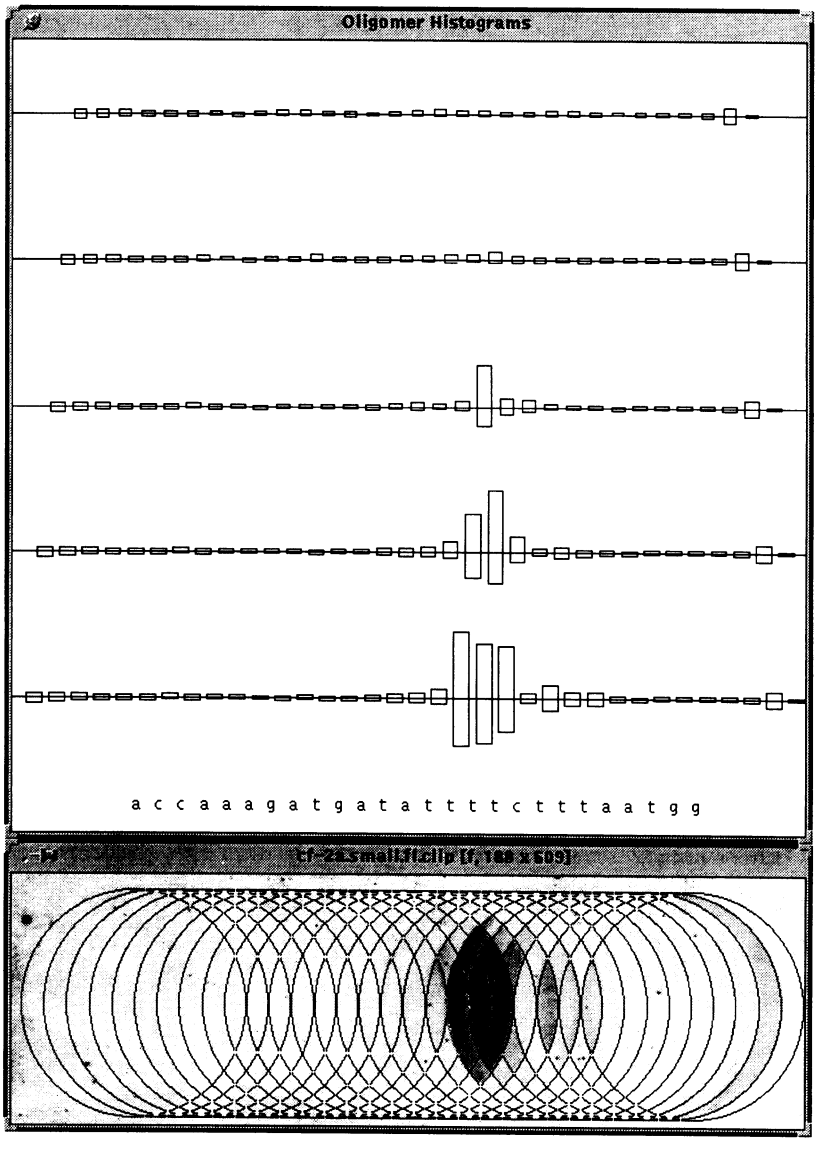

Figure 4. Display and analysis of results. The lower panel shows the image of the array already described in Fig. 3 *b.*, with a template overlay. The upper panel shows paired histograms of integrated pixel values for decanucleotides (bottom histograms) to hexanucleotides (top histograms). The positions of the bars in the histograms correspond to the positions of the oligonucleotide in the array. Values above and below the lines represent regions above and below the centre line of the array, respectively. Image, template and histograms were created in the program **xvseq** (Wang and Elder, in preparation). This program includes tools to aid the interpretation of the large amounts of data in the figure. For example, clicking on a cell in the array highlights the corresponding region in the sequence, the relevant bars in the histogram and the integrated pixel values of the two cells. Clicking on a bar in the histogram highlights the corresponding cells of the array. The figure is a screendump.

Hybridisation is easily done by introducing a solution of labelled target into a gap between the array plate and a plate of similar size held against it. The solution is drawn into the gap by capillary action. After hybridisation, excess solution is rinsed briefly at the temperature, and in the buffer, used for hybridisation. Care must be taken not to warm the plate by contact with fingers, as this would be enough to melt short duplexes.

We find some deterioration of the plates with time. This takes the form of high background in very small spots and seems to be caused by damage to the surface, exposing the glass beneath the linker. The damage may be caused by dust particles trapped against the glass surface during hybridisation or exposure. Other indications of this kind of damage can sometimes be seen as rings corresponding to the edge of the teflon reaction chamber, which must be caused by the high point pressure used to seal it against the glass (e.g. Fig. 3*a*); we also see it when we deliberately score a reference mark into the surface of the plate using a diamond pen. Radioactive target RNA or DNA binds strongly to untreated glass. However, the oligonucleotides appear to be very stable; we have used some arrays for more than a year without apparent loss of performance.

Hybridisation of a polypyrimidine sequence to a oligopurine array

We show below that intramolecular folding in nucleic acids can play a dominant role in determining the extent of hybridisation to oligonucleotides. To avoid this effect, we prepared an array comprising a set of oligopurines and hybridised it with an

oligopyrimidine. The array was made by applying the purines in the order:

3′-agagaaagggaaggaaggagagagaaagg

using a template and offset that gave decanucleotides in the centre.

The array has complements in the sequence of 40 pyrimidines:

5′-ctctctctttcccttccttcctctctctttcccttccttc

This sequence was end-labelled using γ-^{32}P−ATP and polynucleotide kinase and hybridised to the array in 3.5M TMA-Cl at 4°C. The yield of hybrid is remarkably uniform over the length of the array (Fig. 3*a*) which indicates that the yields of oligonucleotide are consistent along the array.

Hybridisation of a RNA molecule to a scanning array

By contrast with the first example, hybridisation of a 528-base transcript of exon 10 of the CFTR gene (5) to an array representing the complements of bases 287−305:

3′-ggtaatttattttatagtagaaacca

shows very complex interactions (Fig. 3*b*). First, there is a marked variation in hybridisation signal along the array. This is unlikely to be due to effects of base composition as the hybridisation was carried out under conditions which minimise this effect (3.5 M TMACl) (4). The target sequence has a high degree of structure predicted by **mfold**, (ref.6; data not shown), which is likely to interfere with interactions with complementary oligonucleotides. However, it is not easy to correlate the predicted folding pattern with the hybridisation behaviour: note particularly the fall in signal from positions 14 to 13, which is a shift of a single base along the sequence, and produces a twentyfold drop in signal. Even more puzzling is the hybridisation to the dinucleotide gg, a result confirmed in a duplicate experiment, which produces a signal strength equal to or stronger than some fully complementary decanucleotides (Fig. 3*b* and 4).

We have carried out numerous experiments using 'scanning' arrays and find that special software (Fig. 4) and some simple rules help in interpreting the complex patterns of hybridisation. First, reading the variation in intensity along any line of isostychs (oligonucleotides of the same length) shows regions of the target which are open to duplex formation; such regions are likely to have unpaired bases. However, we have found in many other cases that regions with unpaired bases are not necessarily open to duplex formation. Second, in some cases we find that the signal is uniform over a large lenticular patch extending over, say, three decanucleotides on the centre line, two nonanucleotides in the second row and one octanucleotide in the third row; we interpret this to mean that only eight bases corresponding to the octanucleotide are available for hybridisation in the target, and that the complements of these eight are all available in the oligonucleotides on the array. Third, and related to the second feature, 'arcs' or 'crescents' are often seen in which the hybridisation signal for the longest oligonucleotide is equalled by that of shorter sequences down to, say, the octanucleotides; such arcs have in common the sequences that are in the shortest one, but so do others which show lower levels of hybridisation. We take this result to indicate that only the short length of oligonucleotide is available for hybridisation in the target sequence, otherwise the signal would be stronger over the longer than the shorter oligonucleotides, but further, that those longer oligonucleotides in the array which show lower signal themselves have intramolecular structure which prevents duplex formation. In a fourth pattern, we have seen 'arcs' in which the central, longer oligonucleotide has lower intensity than the shorter members of the arc. Again, we interpret this pattern as showing secondary structure in the longer oligonucleotides. Examples of these features will be shown elsewhere.

DISCUSSION

Molecular hybridisation to oligonucleotides is a powerful analytical tool in its own right, finding applications in mutation analysis (7); it also forms the basis of important methods such as the Sanger DNA sequencing method (8) and the polymerase chain reaction (9); the use of oligonucleotides as antisense reagents also depends on molecular hybridisation. In these applications, the sequence of the target is already known and the oligonucleotides to use are normally chosen by relatively simple criteria: base composition is used to predict melting and annealing temperatures; string matching is used to eliminate oligonucleotides with internal pairing regions or which may find alternative partners in the target sequence or in a second primer in the PCR. The experiments we describe here, and others carried out with a range of nucleic acid sequences including *in vitro* RNA transcripts, natural RNAs such as tRNA and synthetic oligonucleotides, show that the dominant effect in determining the extent of interaction between an oligonucleotide and a longer target sequence is prevention of duplex formation by structure in the target sequence. However, some positive interactions such as that of the dinucleotide, and others to tri- and tetranucleotides we have found in experiments using complete sets of oligonucleotides, are not predictable by the normal rules of Watson−Crick base pairing, and may be important. The array method presented in this paper provides an empirical method for analysing the interactions of a target molecule with a complete set of complementary oligonucleotides.

ACKNOWLEDGEMENTS

This work was supported by a grant from the MRC HGMP directed programme. JCW is supported by Beckman Instruments.

REFERENCES

1. Southern, E.M., Maskos, U. and Elder, J.K. (1992). *Genomics*, **13**: 1008−1017.
2. Maskos, U. and Southern, E.M. (1992) *Nucleic Acids Research*, **20**: 1675−1678.
3. Maskos, U. and Southern, E.M. (1992) *Nucleic Acids Research*, **20**: 1679−1684.
4. Maskos, U. and Southern, E.M. (1993) *Nucleic Acids Research*, **21**: 2269−2270.
5. Zielenski, J., Rozmahel, R., Bozon, D., Kerem, B.-S., Grzelczak, Z., Riordan, J. R., Rommens, J. and Tsui, L.-C. (1991). *Genomics* **10:** 2114−228.
6. Jaeger, J. A., Turner, D. H. and Zuker, M. (1989). *Proc. Natl. Acad. Sci. USA* **86**: 7706−7710.
7. Conner, B. J., Reyes, A. A., Morin, C., Itakura, K., Teplitz, R. L. and Wallace, R. B. (1983). *Proc. Natl. Acad. Sci. USA* **80:** 278.
8. Sanger, F., Nicklen, S. and Coulson, A. R. (1977). *Proc. Natl. Acad. Sci. USA* **74:** 5463−5467.
9. Saiki, R. K., Walsh, P. S., Levenson, C. H. and Erlich, H. A. (1989). *Proc. Natl. Acad. Sci. USA* **86:** 6230−6234.
10. Breslauer, K. J., Frank, R., Blöcker, H. and Marky, L. A. (1986). Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA* **83:** 3746−3750.
11. Wetmur, J. G. (1991). DNA probes: Applications of the principles of nucleic acid hybridization. *Critical Reviews in Biochemistry and Molecular Biology* **26:** 227−259.