Nucleic Acids Research, 1993, Vol. 21, No. 20 ***4663–4669***

A study of oligonucleotide reassociation using large arrays of oligonucleotides synthesised on a glass support

Uwe Maskos* and Edwin M.Southern
Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

Received August 13, 1993; Accepted September 6, 1993

ABSTRACT

An extensive analysis of oligonucleotide interactions was carried out by hybridising a synthetic pool of 256 10mers, *A(C,T)$_8$A*, representing all oligopyrimidine octamer sequences to an array of four copies of all 256 different octapurine sequences. The resulting 256 duplexes were quantified by phosphorimaging and analysed to determine the dependence of duplex formation on base composition, sequence, and salt concentration. The results show that the base composition dependence of duplex formation can be reduced by high concentrations of tetramethylammonium chloride. This chaotropic solvent also increases duplex yield by up to fiftyfold.

INTRODUCTION

We have recently reported the development of a novel method of analysing nucleic acid sequences [1, 2, 3, 4] that differs in principle from established methods [5, 6]: a large number of oligonucleotides (up to several thousand) are synthesised on a glass plate and the sequence to be analysed is hybridised to the array. The sequence of the probe can be reconstructed from the obtained hybridisation pattern using computer-based methods [3, 7].

These large arrays of oligonucleotides also allow the simultaneous analysis of physicochemical properties of large numbers of oligonucleotides in a way not possible with standard solution methods (e.g. [8]). The dependence of oligonucleotide duplex formation on oligonucleotide length, sequence, base composition, concentration, base mismatches, hybridisation solvent, and from these measurements the kinetic and thermodynamic parameters of oligonucleotide hybridisation, e.g. association and dissociation rate constants, and melting temperatures, can be studied on a large set of oligonucleotides.

In the following we present analyses of the dependence of formation of oligopurine:oligopyrimidine duplexes on base composition, sequence, and hybridisation solvent. We have made on a single glass plate an array comprising four copies of all 256 octapurines, eight copies of all 128 heptapurines, and four copies of all 64 hexapurines (Fig. 1) [3] and hybridised it with all 256 octapyrimidines. This choice of sequences allows the analysis of duplexes ranging from 0 to 100% *G+C* content in a relatively small set because of the purine-pyrimidine base pairing rule. A further advantage of this set is that no secondary structure formation by intrastrand Watson–Crick pairing is possible; such structures would complicate the interpretation of the data.

MATERIALS AND METHODS

Plate preparation, oligonucleotide synthesis, and probe synthesis was carried out as described [3]. Fig. 1 shows the sequence organisation on the plate, and Fig. 2 the arrangement of individual octapurine sequences.

Probe synthesis

A pool of 256 different oligonucleotides of the formula 5′-*A(C,T)$_8$A* was synthesised on an Applied Biosystems Model 381A. After deprotection in conc. ammonia at 55°C overnight the pool was purified using an Applied Biosystems Oligonucleotide Purification (OPC) Cartridge [9]. An *A* residue at the 3′ end was chosen so that a single oligonucleotide synthesis column could be used to generate the whole pool. The *A* nucleotide at the 5′ end would offer the same substrate to the polynucleotide kinase provided by these *A*'s thus eliminating differences in the rate or extent of labelling. The 'dangling' ends provided by the *A* residues would also provide additional stabilisation of duplexes [10] and mimic the behaviour of a longer sequence.

Analysis of the oligonucleotide mixture

Analysis of the labelled mixture by standard denaturing polyacrylamide gel electrophoresis showed the expected rather broad band since *T* and *C* residues are not equivalent in their electrophoretic behaviour, with *C* rich oligonucleotides migrating faster than *T* rich oligonucleotides of the same length [11].

The base composition of the mixture was 44% *C* and 56% *T*, as analysed by HPLC after digestion of the mixture with snake venom phosphodiesterase and bacterial alkaline phosphatase [12].

The oligonucleotide composition of the probe mixture was also analysed by electrophoresis in an *acidic* denaturing polyacrylamide gel [13] that allowed the separation of the oligonucleotides according to *C* and *T* content; the values for base compositions ranging from one to seven *T*'s were in good agreement with the

* To whom correspondence should be addressed at: Laboratory of Molecular Biology, National Institute of Neurological Disorder and Stroke, National Institutes of Health, Bethesda, MD 20892, USA

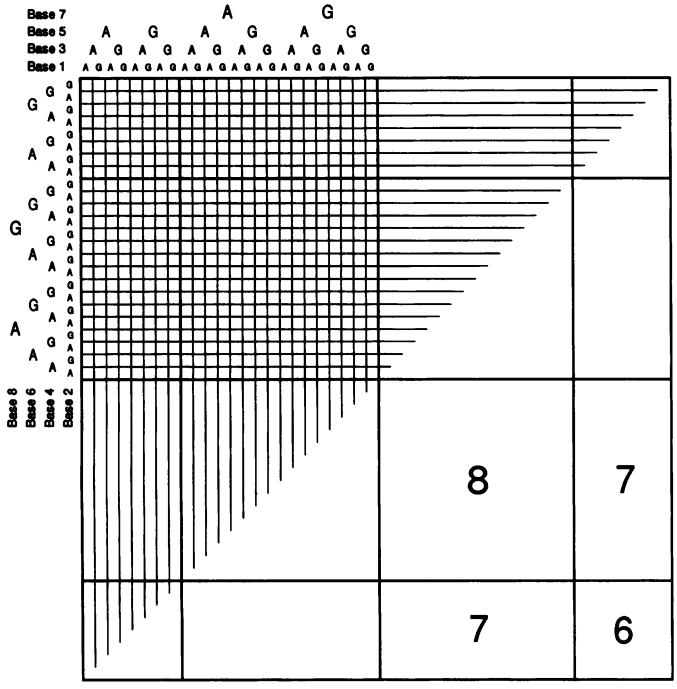

Figure 1. An outline of the derivatisation of the plate used for the experiments. On the plate four copies of all 256 octapurines (indicated **8**), eight copies of all 128 heptapurines (indicated **7**), and four copies of all 64 hexapurines (indicated **6**) were synthesised. The small squares in the drawing outline an individual 'cell' containing the oligonucleotide. For the synthesis, precursor solutions were injected into channels in the order set out in the drawing. Thus in the first step ('1st base') *A* and *G* were introduced into alternating channels, then the mask turned 90°, and the process repeated ('2nd base'). In steps 3 and 4 two neighbouring channels at a time were filled with the same solution, in steps 5 and 6 four, and in steps 7 and 8 eight at a time.

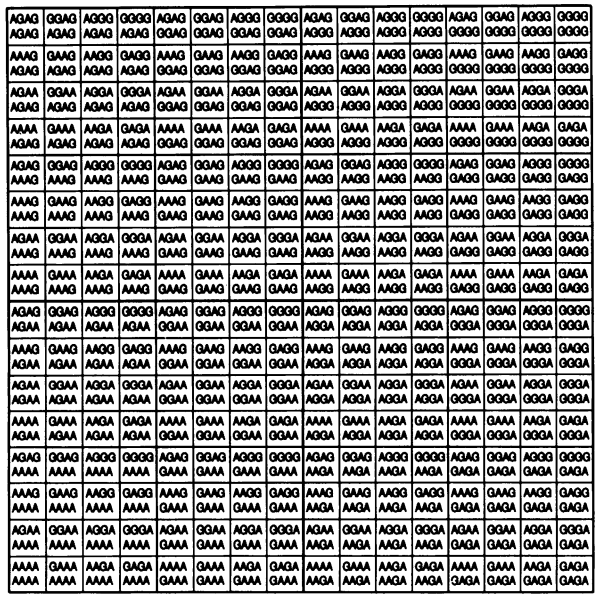

Figure 2. Positions of oligonucleotides present in an octamer set. The oligonucleotides present in one array of octamers on the oligopurine plate of Fig. 1 have been listed. Sequences are written in the 3′ to 5′ directon. Oligonucleotides are arranged such that sequences richest in *A* can be found in the bottom left, and those richest in *G* in the top right corner.

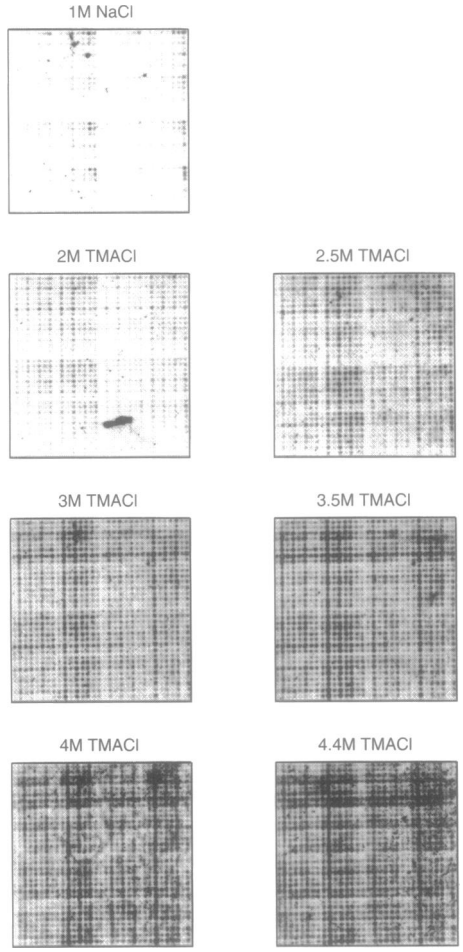

Figure 3. The hybridisation pattern of a synthetic pool of 256 10mers $A(C,T)_8A$ in different solvents. In the analysis only a few (~two dozen) 'hot-spots' were discovered among the 7168 with octapurines analysed, where unspecific adsorption of probe material had occurred. They were detected because of their 'off-center' location in a cell, their shape, their rather high values after integration; they were eliminated by median filtering. This indicates the general cleanliness of experiments. No spot was missing. The patch of non-specific adsorption visible in the 2M experiment was due to physical damage to the glass plate that had occurred before that experiment was carried out. It was the last hybridisation in the series.

HPLC data. This bias towards *T* rich oligonucleotides was taken into account in the analysis of the data.

Labelling and hybridisation

After end labelling with polynucleotide kinase and γ-^{32}P-ATP [14] an aliquot was used in a hybridisation to the oligopurine plate for 6 to 8 hours at 4°C in the set-up previously described [3]. 4°C was chosen as the hybridisation temperature because at room temperature the degree of duplex formation was not sufficient for reliable quantitation with the PhosphorImager (Molecular Dynamics) due to the low concentration of each oligonucleotide in the probe.

In a series of experiments hybridisation solvents were 1M NaCl (containing 10mM Tris·HCl pH 7.5, 1mM EDTA, 7% sarcosine), 2M, 2.5M, 3M, 3.5M, 4M, 4.4M, 5M and 5.5M TMACl (all containing 50 mM Tris·HCl pH 8.0, 2 mM EDTA, SDS at less than 0.04 mg/ml). The plate was rinsed for 10 minutes at 4°C in the hybridisation solvent to remove non-hybridised

molecules, sealed in a plastic bag and exposed to a PhosphorImager storage phosphor screen (Fuji, STIII) [15] at 4°C, overnight in the dark.

The plates were scanned, the PhosphorImager files converted using specially developed imaging software and printed (e.g. Fig. 3).

In a typical experiment, between 10 and 30 c.p.s. could be typically detected with a hand-held minimonitor held over the surface of the plate. Data were analysed using the ImageQuant (Molecular Dynamics) software that allowed the calculation of hybridisation signal of each oligonucleotide on the plate by superimposing grids on the array, and integrating the individual pixel values for every cell. These data were transferred to a Microsoft Excel spreadsheet for further analysis. The data were also analysed with our own imaging software and median filtering was carried out as described [3, 7].

The 5M and 5.5M TMACl experiments could not be analysed accurately. Presumably due to insolubility of oligonucleotides under those conditions (i.e. very high salt concentration, low temperature) the values obtained were close to background and large areas of the plate were covered with a smear of radioactive material apparently precipitated onto the surface.

EXPERIMENTS AND RESULTS

There is extensive experimental evidence [16, 17] that base and sequence dependent contributions to DNA stability are averaged out for long double-stranded duplexes in chaotropic solvents, such as TMACl. However, the conclusion is based on a limited set of sequences and may not be relevant to short oligonucleotides. Only two studies [8, 18] have focussed on hybridising oligonucleotides in chaotropic solvents, with undecamers as the shortest sequences. The number of probes was again rather limited. Wood *et al.* [18] measured the stability of duplexes of different lengths but *similar GC* content; Jacobs *et al.* [8] also varied base composition and found similar stabilities for oligonucleotides of the same length down to 16mers in 3M TMACl. But the five 14mers analysed differ by 7°C in their T_d (dissociation temperature, a term used when melting experiments are carried out under conditions far from thermodynamic equilibrium), with the richest in $G+C$ showing the *lowest* stability. This can be taken to indicate that for duplexes shorter than 15 base pairs, stabilities do become highly dependent on base composition and possibly sequence.

Effects of base composition and cation on duplex yield

Oligonucleotide sequences were grouped according to base composition and an average value for the hybridisation to the corresponding cells was calculated.

Averages were normalised (highest ≡ 100%) and plotted against base composition in the form of histograms (Fig. 4). A conservative estimate of the variation associated with the average yield, and a measure of the sequence dependent differences between oligonucleotides with the same base composition can be obtained by calculating the standard deviation σ associated with the mean according to

$$\sigma = \left(\frac{\sum x_i^2 - \frac{(\Sigma x_i)^2}{n}}{n} \right) 1/2,$$

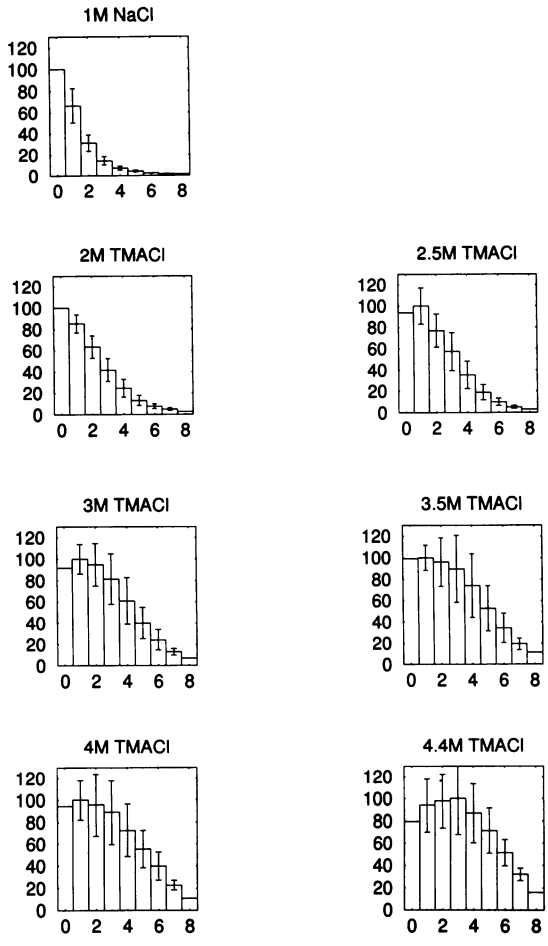

Figure 4. The dependence of duplex yield on base composition and hybridisation solvent. The four values corresponding to the same sequence in the four octamer quadrants were median filtered. An average was calculated for oligonucleotides with the same base composition. Those data were plotted as histograms. The abscissa indicates base composition (number of *A*'s in the sequence), the ordinate relative yield with the highest defined as 100%. The average standard deviation for the intensity values of the four identical sequences was 21.5% in the seven experiments, a reasonably low figure that allows to interpret the data with some confidence. This standard deviation ('empirical deviation') was calculated using the formula

$$\sigma = \left(\frac{\sum x_i^2 - \frac{(\Sigma x_i)^2}{n}}{n-1} \right) 1/2$$

with n = 4. The average difference between the lowest and the highest value for the same sequence was 36%.

Table 1. Numbers of octapurine sequences with the same base composition

Base composition/number of *A*'s in the sequence	0	1	2	3	4	5	6	7	8
Number n of different sequences	1	8	28	56	70	56	28	8	1

where n equals the number of oligonucleotides with the same base composition (Table 1), and the calculation was carried out for all x_i, i.e. the median filtered values from the four quadrants corresponding to the same sequence, for a given base composition.

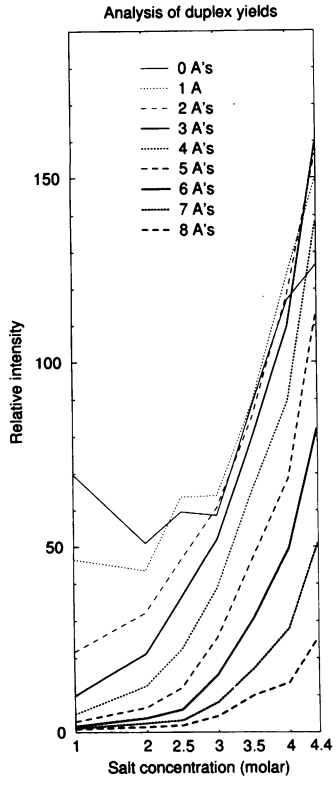

Figure 5. The dependence of overall yield on cation and concentration. Average normalised values of duplex yields for the same base composition have been plotted against concentration of solvent used in the hybridisation. Intensity values were calculated as follows: The four values for oligonucleotides of the same sequence were median filtered. Then a normalisation with respect to hybridisation time, hybridisation volume, number of c.p.m. of the probe used and exposure time to the PhosporImager storage phosphor screen was carried out, as detailed in the text. The concentrations plotted along the abscissa are in mol/l. A 1M concentration refers to hybridisation in 1M NaCl, other concentrations refer to TMACl solutions.

Deviations were highest for intermediate base compositions, as may be expected given the larger variety of sequence contexts for those base compositions. In 1M NaCl there was a steep increase in the amount of hybrid formation with $G+C$ content, as expected from the greater stability of $G:C$ vs. $A:T$ base pairs.

With increasing concentrations of TMACl, sequences with greatest yield shift to those with higher $A+T$ content: in 2.5M TMACl, the eight oligonucleotides with seven $G:C$ base pairs give the highest yield; and at the highest TMACl concentration that could be used, molecules with three $A:T$ base pairs become most stable—on average.

Because of variation in the derivatisation of the plate, only those values that are averages over a larger number of molecules are reliable. However, sequences within the two to six $A:T$ base pair range include 93% of all sequences.

The smallest spread in hybridisation yields for two to six $A:T$ base pairs occurs at 4.4M TMACl with a 47% difference between the highest and lowest average intensity, and with only a 28% difference between the values for three to five $A:T$ base pairs, encompassing 71% of all sequences. The corresponding values in 3.5M TMACl are 64% and 40%, respectively. This has to be compared to the 93% difference obtained in 1M NaCl, and 75% in 3M TMACl, the solvent commonly used to eliminate effects of base composition in high molecular weight DNA.

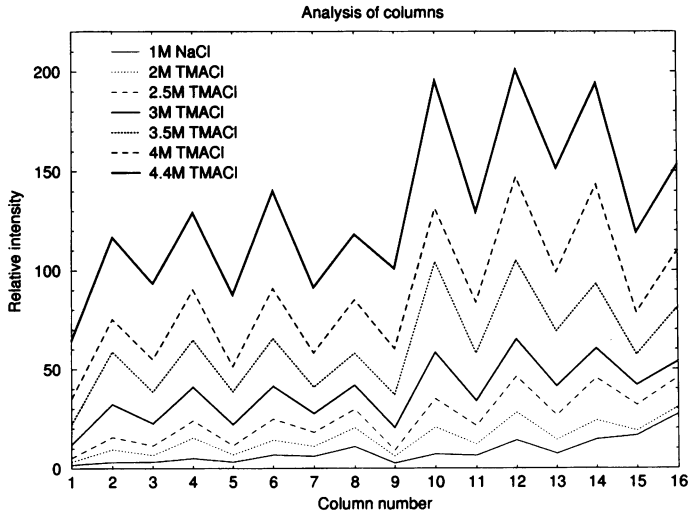

Figure 6. The dependence of duplex yield on the nucleotide at the 3′ end. The median filtered values for the 16 oligonucleotides in every single column of the four octamer sets were added, divided by the number of different sequences in a column (16), and normalised with respect to hybridisation time, volume of solution, oligonucleotide input c.p.m. and exposure time, as in Fig. 5. These average values for every column were plotted against column number for the results obtained in the seven different hybridisation solutions. The sequences in every odd numbered column differ from those in the neighbouring even numbered column by an A instead of a G residue at their 3′ end, cf. Fig. 2. Differences in the plotted average values therefore give an indication of the contributions of different bases at oligonucleotide ends to duplex yield. Note that yields for even numbered columns (oligonucleotides with a 3′G) are consistently higher in all solvents.

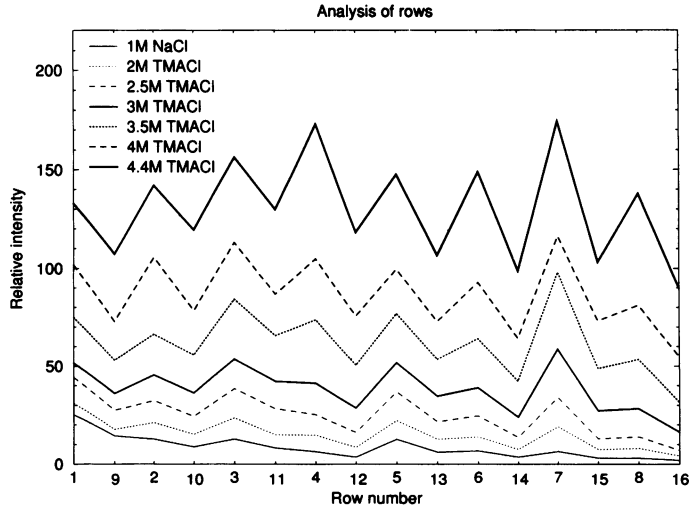

Figure 7. The dependence of duplex yield on the nucleotide at the 5′ end. The intensity values for the 16 oligonucleotides in every single row of the four octamer patches were added, normalised in the same way as detailed in Fig. 6 and plotted against the row number. The sequences in rows 1, 2, ..., differ from those in rows 9, 10, ..., by a G instead of an A residue at their 5′ end, cf. Fig. 2. The same general trend is observed as in Fig. 6, i.e. those oligonucleotides with a G at the 5′ end show higher yields.

Effects of TMACl concentration on overall duplex yield

It was noted that duplex yields increased with TMACl concentration for all base compositions. To measure this effect the data for different experiments were normalised by dividing the counts

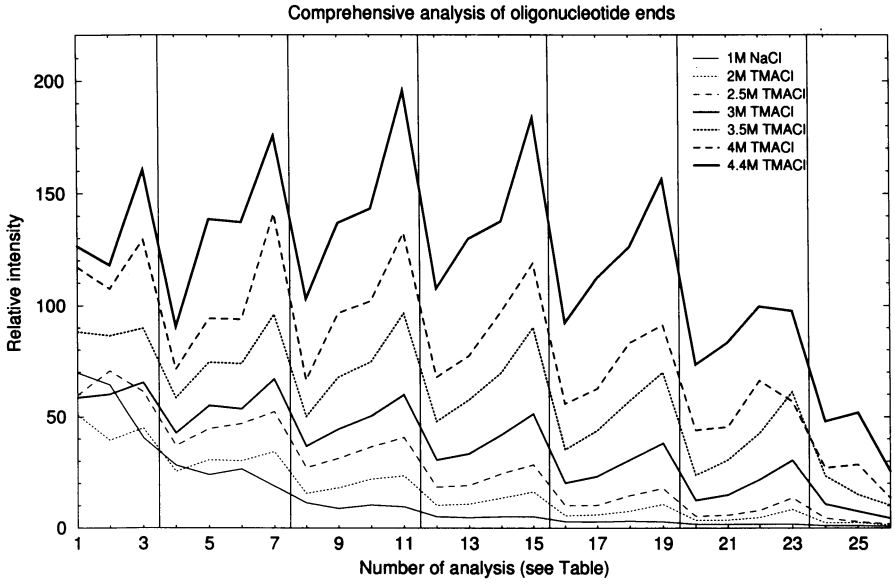

Figure 8. A comprehensive analysis of the dependence of duplex yield on the bases at oligonucleotide ends. For the analysis oligonucleotides were grouped according to base composition and then further subdivided according to whether they have an *A* residue at each end; an *A* at the 3′ and a *G* at the 5′ end; a *G* at the 3′ and an *A* at the 5′ end; and a *G* at each end. The following table defines the 'number of analysis' that corresponds to a given base composition and pattern of residues at the ends of the octapurines on the plate, together with an indication of the number of different sequences in each subgroup.

Number of analysis	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Number of *G*'s in octapurine	8	7	7	6	6	6	6	5	5	5	5	4	4	4	4	3	3	3	3	2	2	2	2	1	1	0
Base at 3′ end	G	*	*	A	A	G	G	A	A	G	G	A	A	G	G	A	A	G	G	A	A	G	G	*	*	A
Base at 5′ end	G	*	*	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	*	*	A
Number of sequences	1	2	6	1	6	6	15	6	15	15	20	15	20	20	15	20	15	15	6	15	6	6	1	2	6	1

For the entries marked with a '*' a different scheme applies. They correspond to analyses 2, 3 and 24, 25, for octapurines with 7 *G*'s and 7 *A*'s, respectively. In analysis 2 the sequences with an *A* at one or the other end are included, in analysis 24 the sequences with a *G* at one or the other end. For analyses 3 and 25 averaging was carried out for those sequences with, respectively, a *G* or *A* at each end. In the figure the datapoints belonging to the same base composition are separated by vertical lines. Normalisation of PhosphorImager counts was carried out in the same way as detailed in Figs. 6 and 7 and in the text. Data for all seven experiments are included.

by hybridisation time, exposure time, number of input c.p.m. used for the hybridisation and the reciprocal of hybridisation volume (Fig. 5).

There is a marked increase in yield with TMACl concentration for all base compositions. As expected, the increase in yield is greater for sequences of high *A*+*T* content: the intensity of eight *G:C* base pairs increases twofold, and that for six *A:T* base pairs increases 48-fold in yield over that seen in 1M NaCl; all other base compositions show intermediate behaviour.

Effects of sequence on duplex yield

That factors other than base composition determine duplex yield can be seen in the differences in intensities of rows and columns containing closely related sequences. The only difference between an odd numbered column (Fig. 2) and its neighbouring even numbered column is the *G* instead of an *A* residue at the 3′ end. Fig. 6 shows the result of summing values for each column.

For rows, the only difference is a *G* instead of an *A* residue at the 5′ end when comparing rows 1 and 9, 2 and 10, and so forth. Sequences with *G* at either end give up to 50% higher duplex yields (Fig. 7). For duplexes of the same base composition, those with an *A* at both ends show lowest intensity, those with a *G* at one end intermediate, and those with a *G* at either end highest yields (Fig. 8).

DISCUSSION

The analyses of the reassociation data yielded a number of unexpected results. 4.4M TMACl gave the smallest difference in yield for sequences of different base compositions, as we found in a less extensive study [2] commonly used 3M TMACl gave much larger differences. Some of the variation could be due to the small number of oligonucleotides at the extremes of base composition, where differences in stepwise yield during synthesis of oligonucleotides on the plate will be a major factor determining duplex yield. Such imperfections are averaged out for intermediate base compositions, where many different sequences contribute to one base composition (Table 1).

Effects of TMACl concentration on duplex yield

Comparison of intensities in different experiments with different cation concentrations was possible because the pseudo first order

kinetics of reassociation determined in experiments to be presented elsewhere [1, 4, 19] and the linear response of the PhosphorImager storage phosphor plate to radioactive decay [15] allow normalisation with respect to hybridisation time and concentration of probe.

The main effects were an increased yield of duplex with TMACl concentration and the expected stabilisation of $A+T$ rich duplexes [16, 20].

The increase in yield was an unexpected benefit of using TMACl as this was not predicted from previous work in which the melting behaviour of polynucleotides was studied. Hamaguchi and Geiduschek [21] found that polynucleotides in 15.8 molal TMACl and in 0.15M NaCl showed the same thermodynamic stability, and Melchior and von Hippel [16] found an increase in stability for solutions containing up to 3.28M TMACl, but a decrease at higher concentrations, a result confirmed by Orosz and Wetmur [17]. For different salts, e.g. LiCl, CsCl, a similar behaviour was observed [21, 22, 23]. But it is probably not justified to compare rigorous thermodynamic melting experiments with the present set of experiments because of the rather different conditions and experimental parameters.

A factor that should be taken into account in the interpretation of our experiments is the possibility of triple-strand formation, which extensive studies have shown to be complete at 15°C at a Na^+ concentration of 0.2M [24]. The requirements are an oligopurine and two oligopyrimidine strands that are mirror images of each other, which is certainly the case for every octapyrimidine in the pool. The additional condition of a pH of preferably <7, necessary for the protonation of C residues, is not met, but very high salt concentrations could stabilise the triplexes that form with the small equilibrium concentration of oligonucleotide decamers containing protonated C residues. And extrapolation of the measurements of Plum *et al.* [24] of the dependence of T_m on pH leads to the prediction of a melting temperature for the triplex of approx. 5°C in solutions of pH 7.6, values very close to the present series of experiments.

Possible contributions of mismatches to duplex yield

The average signals from heptamers (ca. 40−50% of the octamers) and hexamers (ca. 4−22%) were considerably higher than expected. These yields suggest that mismatched octamer pairs might contribute substantially and complicate the interpretation.

For any single immobilised hexamer there are 12, and for a heptamer four different fully matched decamers in the mixed probe. For the octamers the hybridisation pattern could be more complicated, because not only decamers with sub-terminal $G\cdot T$ and $A\cdot C$ mismatches will have to be taken into account, but there might be a contribution from terminal $A\cdot A$ and $A\cdot G$ mismatches.

Despite this potentially complex hybridisation behaviour, the general conclusions from the experiments, i.e. the relative intensity changes with TMACl concentration, are nevertheless valid, because for every octamer on the plate *the same number of related decamers* will contribute.

Control experiments (data not shown) were carried out by hybridising single oligonucleotides, e.g. AC_8A, AT_8A, AC_4T_4A, under the same stringency to the array. The relative intensities of hybridisation, e.g. AC_8A vs. AT_8A, were in good agreement with those determined from the extensive pool hybridisation. Absolute hybridisation phosphorimaging counts were within 50% of those measured in the hybridisations shown.

Effects of base sequence

A possible explanation for the bias towards higher intensities for octamers with G at the ends, and also for the somewhat higher concentration of TMACl that had to be chosen to equalise the $G{:}C$ versus $A{:}T$ difference, is that if the octamer possesses a G residue at one or both ends, the mismatched decamers are expected to show a higher yield, because both a $G\cdot T$ and a possible $A\cdot G$ mismatch (because of the A's added to the probes) have been reported to be less destabilising than the respective $A\cdot C$ and $A\cdot A$ mismatches [25]. This would be consistent with the results in Figs. 4 and 8, because a higher TMACl concentration has to be chosen to stabilise $A+T$ base pairs sufficiently to overcome this additional contribution to duplex yield.

As has been pointed out already, the NaCl results differ in that the overall $G+C$ content represents the main determinant of yield, with differences between different end bases smaller than in the other solvents, and, by contrast, A residues at both ends most stable, i.e. for three to six $G{:}C$ base pairs.

In 1M NaCl, comparatively stringent hybridisation conditions, the overall yield of hybrid is low and possibly not much mismatching occurs, i.e. hybridisation is more specific, or that, alternatively, TMA^+ ions selectively stabilise $G\cdot T$ and $A\cdot G$ mismatches. The completely different picture in 2M TMACl, with comparable average intensities, actually lends support to the second possibility. 1M NaCl certainly allows the best discrimination between hexamers and longer sequences, with a 4% average hexamer yield as compared to octamers, and an average 50% for heptamers.

That G's at the ends of the tethered oligonucleotides increase duplex yields stands out as the major sequence specific effect seen in this study. All sequences are affected in TMACl solutions, but not in NaCl, which suggests that it is not a systematic deficiency in yields of oligonucleotides synthesised on the plate. One possible explanation is that terminal GC base pairs contribute more to stability of the duplex than do internal GC pairs, as has been observed by Breslauer *et al.* [26]. Such an effect could also influence the rate of duplex formation if the first step were formation of a single base pair and the ends were more likely to take part in this than internal bases, but in this case the effect would be more pronounced in sodium than TMA salts and the opposite is found. An alternative could be increased opportunities for mispairing with G as compared with A as discussed.

These studies can now be extended to address the effects of different temperatures and hybridisation times on duplex yield and specificity of hybridisation.

CONCLUSIONS

A wealth of data can be generated in a single experiment in which large pools of oligonucleotides are hybridised to a large array. Calculations based on the larger number of values provide a reliable basis for estimating factors affecting duplex yield and stability.

ACKNOWLEDGEMENTS

We are grateful to J.K.Elder for data analysis, software and help with computers. We wish to thank Dawn Ryder for running acidic gels and Tony Willis for carrying out the HPLC analysis. This work was supported by the MRC.

REFERENCES

1. Maskos, U. and Southern, E. M. (1992) *Nucleic Acids Res.* **20**, 1679–1684.
2. Maskos, U. and Southern, E. M. (1992) *Nucleic Acids Res.* **20**, 1675–1678.
3. Southern, E. M., Maskos, U., and Elder, J. K. (1992) *Genomics* **13**, 1008–1017.
4. Maskos, U. (1991) 'A novel method of nucleic acid sequence analysis', D. Phil. thesis, Oxford University.
5. Maxam, A. M. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 560–564.
6. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467.
7. Elder, J. K. (1993) Analysis of DNA oligonucleotide hybridization data by maximum entropy. In Mohammad-Djafari, A. and Demoment, G. (eds.), Maximum Entropy and Bayesian Methods, Paris 1992. Kluwer, Dordrecht, pp. 363–371.
8. Jacobs, K. A., Rudersdorf, R., Neill, S. D., Dougherty, J. P., Brown, E. L., and Fritsch, E. F. (1988) *Nucleic Acids Res.* **16**, 4637–4650.
9. McBride, L. J., McCollum, C., Davidson, S., Efcavitch, J. W., Andrus, A., and Lombardi, S. J. (1988) *BioTechniques* **6**, 362–367.
10. Mellema, J.-R., van der Woerd, R., van der Marel, G. A., van Boom, J. H., and Altona, C. (1984) *Nucleic Acids Res.* **12**, 5061–5078.
11. Frank, R. and Köster, H. (1979) *Nucleic Acids Res.* **6**, 2069–2087.
12. McLaughlin, L. W. and Piel, N. (1984) 'Chromatographic purification of synthetic oligonucleotides.' In Gait, M. J. (ed.), Oligonucleotide Synthesis—A Practical Approach. IRL Press, Oxford, pp. 117–133.
13. de Wachter, R. and Fiers, W. (1972) *Anal Biochem.* **49**, 184–197.
14. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. CSH Laboratory Press, Cold Spring Harbor, New York.
15. Johnston, R. F., Pickett, S. C., and Barker, D. L. (1990) *Electrophoresis* **11**, 355–360.
16. Melchior, W. B., Jr., and von Hippel, P. H. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 298–302.
17. Orosz, J. M. and Wetmur, J. G. (1977) *Biopolymers* **16**, 1183–1199.
18. Wood, W. I., Gitschier, J., Laskey, L. A., and Lawn, R. M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1585–1588.
19. Maskos, U. and Southern, E. M. (manuscript in preparation).
20. Shapiro, J. T., Stannard, B. S., and Felsenfeld, G. (1969) *Biochemistry* **8**, 3233–3241.
21. Hamaguchi, K. and Geiduschek, E. P. (1962) *J. Am. Chem. Soc.* **84**, 1329–1338.
22. Gruenwedel, D. W. and Hsu, C.-H. (1969) *Biopolymers* **7**, 557–570.
23. Gruenwedel, D. W., Hsu, C.-H., and Lu, D. S. (1971) *Biopolymers* **10**, 47–68.
24. Plum, G. E., Park, Y.-W., Singleton, S. S., Dervan, P. B., and Breslauer, K. J. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9436–9440.
25. Aboul-ela, F., Koh, D., Tinoco, I., Jr., and Martin, F. H. (1985) *Nucleic Acids Res.* **13**, 4811–4824.
26. Breslauer, K. J., Frank, R., Blöcker, H., and Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3746–3750.