Nucleic Acids Research, 1994, Vol. 22, No. 8 1365–1367

Studies of oligonucleotide interactions by hybridisation to arrays: the influence of dangling ends on duplex yield

Jennifer C.Williams, Stephen C.Case-Green, Kalim U.Mir and Edwin M.Southern*
Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

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

ABSTRACT

Effects of dangling ends on duplex yield have been assessed by hybridisation of oligonucleotides to an array of oligonucleotides synthesised on the surface of a solid support. The array consists of decanucleotides and shorter sequences. One of the decanucleotides in the array was fully complementary to the decanucleotide used as solution target. Others were complementary over seven to nine bases, with overhangs of one to three bases. Duplexes involving different decanucleotides had different overhangs at the 3′ and 5′ ends. Some duplexes involving shorter oligonucleotides had the same regions of complementarity as these decanucleotides, but with fewer overhanging bases. This analysis allows simultaneous assessment of the effects of differing bases at both 5′ and 3′ ends of the oligonucleotide in duplexes formed under identical reaction conditions. The results indicate that a 5′ overhang is more stabilising than a 3′ overhang, which is consistent with previous results obtained with DNA overhangs. However, it is not clear whether this is due to the orientation of the overhang or to the effect of specific bases.

INTRODUCTION

Duplex formation between an oligonucleotide and a target molecule forms the basis of many genetic methods. Analytical methods include: the PCR; sequence determination by primer extension; mutant analysis by dot-blots or by reverse dot-blots. Experimental methods include gene ablation by antisense oligonucleotides and ribozymes. Nucleic acid interactions are also important in processes such as interactions of mRNAs with spliceosomes and ribosomes, and intramolecular interactions are responsible for the folded structure of single stranded molecules.

It is important to understand the molecular basis for these interactions. The central approach has been to use statistical mechanical calculations of duplex stability, based initially on the contributions of Watson–Crick base pairing, which are well understood (1, 2). More recent programs incorporate data for nearest neighbour interactions (3), acknowledging the major contribution to duplex stability of base stacking interactions between adjacent ring systems, and for loops and bulges.

However, data on these and other factors which may make substantial contributions to stability are not adequate for accurate predictions. Structures in motifs such as dangling ends, bulges and loops are not fully characterised and the vast range of possible structures will necessitate a commensurate number of analyses to collect the relevant data. The conventional approach of measuring interactions one at a time in the liquid phase, while producing accurate measurements of the thermodynamic properties needed for stability calculations, are time consuming and difficult to apply on a large scale. This paper describes an alternative experimental approach which gives a measure of the effect of different structures on the rate of formation of duplexes under a defined set of conditions. Though these are not readily transformed into the free energy values used to calculate duplex stabilities, they provide a simple empirical approach to what is often required—a measure of sequence dependent effects on duplex yield for a large number of sequences.

The experiment described in this paper was part of a larger study of the hybridisation behaviour of a region within a 528 base transcript of exon 10 of the CFTR gene. For that study, we made an array comprising a set of oligonucleotides complementary to bases 273–320 of the transcript. Hybridisation of the full transcript to the array produces around 150 different interactions with oligonucleotides. In the study described here, we focus on the hybridisation behaviour of a short region, a synthetic deoxydecanucleotide corresponding to bases 288–298 of the transcript, and show how the approach can be used to provide data on basic interactions, in this case the effects on duplex yield of unpaired bases at the ends of the duplexes.

EXPERIMENTAL

The array was synthesised on the surface of a glass plate (165×50 mm) as described in detail in the accompanying paper (4 and ref. 5). The circular template had a diameter of 30 mm, and was moved in steps of 3 mm, resulting in an array with decamers in the central lenticular window (the lens shaped region). Subsequences of the full length decamer are represented within the outer lattice pattern (see Fig. 1 of accompanying paper (4) for further explanation).

The target sequence, 5′ GAAAATATCA 3′ was made in an Applied Biosystems 381A oligonucleotide synthesiser using conventional phosphoramidite chemistry, and labelled using γ-^{33}P-ATP (Amersham International, 300 Ci/mM) in the presence of polynucleotide kinase (Sigma). The product was purified in a spun column, and diluted in 1 ml of 4.5M TMACl

*To whom correspondence should be addressed

(final concentration = 3 million cpm/ml). The hybridisation was performed by injecting ca. 500 μl of the oligonucleotide solution between the array and a plain glass plate placed against its surface. Hybridisation was carried out in a moist chamber at 4°C for 15 hours. The plate was washed briefly in 4.5M TMACl at 4°C, wrapped in Clingfilm, and exposed on a pre-cooled phosphor storage screen (Fuji, ST III) for 3 hours at 4°C. The concentration of the solution oligonucleotide is ca. 1 pmol/ml and the solution phase is spread over a total area of ca. 10 000 mm^2; the density of oligonucleotides on the surface of the array is around 1 pmol/mm^2. Thus the solid phase oligonucleotides are in ca. 10 000-fold excess over those in the solution immediately above them. We know that the occupancy of the oligonucleotides in the array can be greater than 1/1000. It is likely that the rate of reaction is affected by the rate of diffusion of oligonucleotides from adjacent regions in areas where there is a high forward rate. Other experiments, to be reported elsewhere (Case-Green and Southern, in preparation), show that the reverse reaction is very slow as compared with duplex formation under the conditions used in this study.

RESULTS

The target oligonucleotide was the perfect complement of one of the full length decamers in the array, but the design of the array (Fig. 1) permits analysis of many other related interactions. As these are carried out in a single experiment, subtle differences in yield are readily detected and measured. Furthermore, the symmetrical structure of the array provides duplicate results for the shorter oligonucleotides, and stripping and rehybridisation using a separately labelled target produced an identical pattern. The technique is therefore highly reproducible, and we have found that the arrays may be reused many times.

The target hybridises to its perfect 10 base complement, represented by the starred lenticular window (Fig. 1). It also hybridises to 9, 8 and 7 base regions of complementarity as expected under the non-stringent conditions used in the experiment. The hybridisation pattern shows that some duplexes with the same region of sequence complementarity produced a different yield, whereas others produced the same yield (Fig. 1 and Table 1).

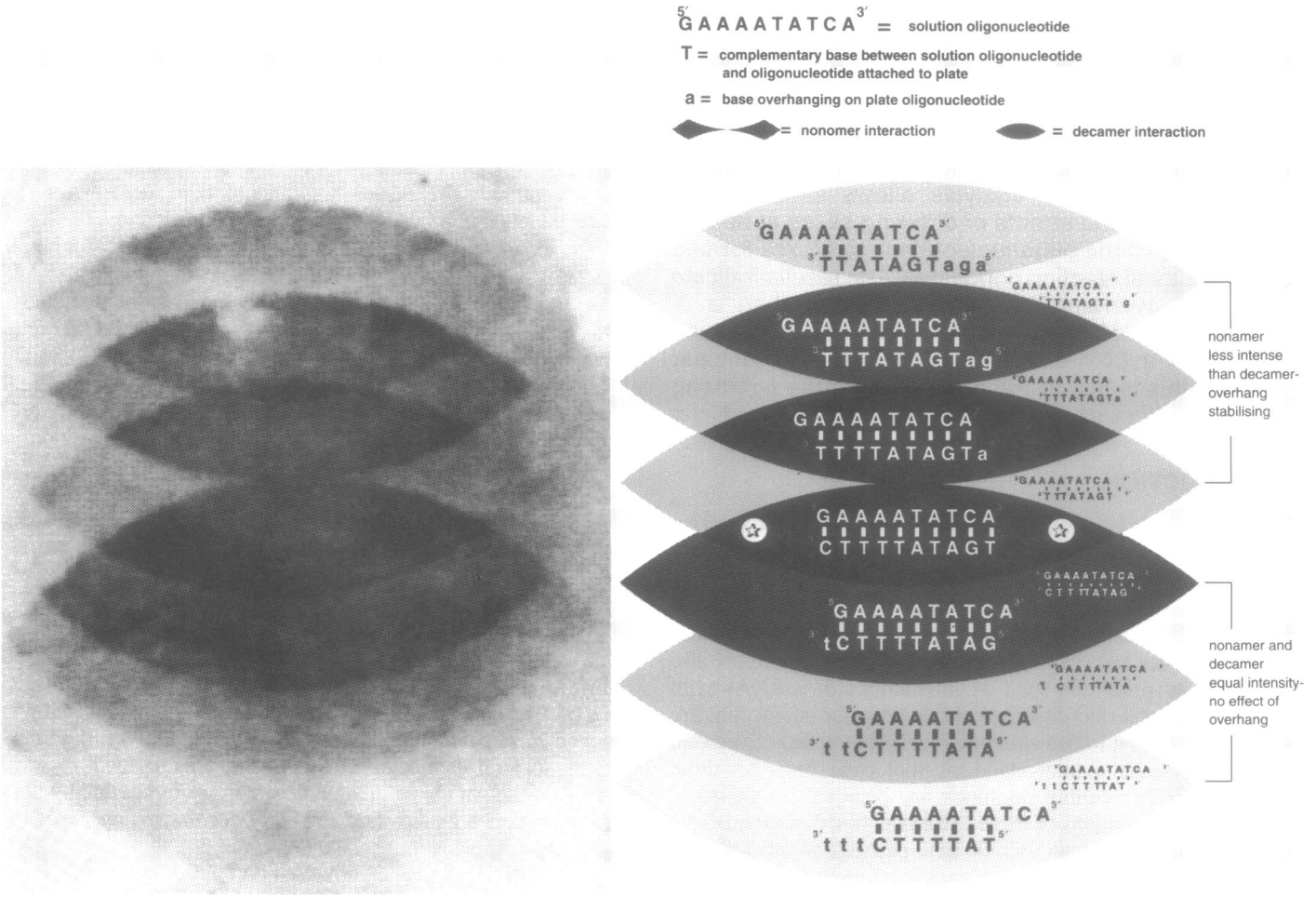

Figure 1. On the left is the hybridisation pattern of the decanucleotide pGAAAATATCA-3′ to a scanning array representing complements of the CFTR exon 10. The target is homologous to bases 288 to 298 of the CFTR transcript. On the right is a representation of the hybridisation pattern with a key to the interacting oligonucleotides. The perfect complement of the target, GAAAATATCA-3′, is marked with asterisks; oligonucleotides surrounding this decamer give different overhangs, as illustrated. Note that the uniform intensity of hybridisation along the three arcs at the bottom of the array shows that adding Ts to the 3′ end of the tethered oligonucleotide, producing an overhang at the 5′ end of the solution target, has no effect on yield. By contrast, the three arcs above the asterisks show a sharp difference in intensity between the central decanucleotide and the corresponding nonanucleotide, indicating that adding As and a G to form a 5′ overhang, produces a marked increase in yield.

Adding one to three overhanging bases to the 5′ end of the tethered oligonucleotide increased the yield of duplex relative to shorter oligonucleotides without overhangs. The effect of adding the first base is large; the second and third have smaller effects. By contrast, addition of one, two or three overhanging bases at the 3′ end had no effect on duplex yield.

DISCUSSION

The two ends of the duplex differ in a number of respects. First, the 3′ end of the tethered oligonucleotide is attached to the glass surface through a twenty atom aliphatic oligoether. It is unlikely that the linker causes the effects described above, because the overhanging bases *extend* the distance between the duplex and the linker. If the linker interfered with duplex formation, we would expect to see an increase in yield where we see none.

A second difference between the two ends is that the overhanging bases on the 3′ end are pyrimidines whereas those at the 5′ end are purines. It is known that purines produce more stabilising overhangs than pyrimidines, and this additional stability could explain the difference we see in yields (6, 7). However, there are reports of increased duplex stability due to overhanging Ts (8, 9). Furthermore, the terminal base of the duplex is A at the 3′ end, and G at the 5′ end; these could interact in different ways with the first base in the dangling ends (6). In addition, the 5′ end of the solution oligonucleotide is phosphorylated whereas that of the tethered oligonucleotide end is not. Phosphorylation is likely to affect the local stability of a duplex, but seems unlikely to have effects on overhangs longer than a single base. Thus it is unlikely that phosphorylation accounts for the present observations.

Finally, the orientation of the nucleotides differs in the two overhangs. Senior *et al.* (9) using temperature dependent UV spectroscopy, circular dichroism and NMR, have shown that 5′ overhanging bases are more stabilising than 3′ overhangs in DNA. This is in agreement with our results. (In contrast, in RNA, 3′ overhangs are more stable than 5′ overhangs probably as a result of the difference between the A and B forms of the double helix (6).

Thus there are a number of possible contributions to the effects on yield that we describe here, and other experiments would be needed to separate them. Nevertheless the present work, carried out using arrays designed for other purposes, show how the approach can be used to investigate effects of interactions other than Watson–Crick base pairing. Arrays can be designed specifically to address particular problems, for example, we could extend the present investigation by varying the sequences in the dangling ends: all trinucleotides could be represented in 64 sequences. The implications of single strand effects in many important aspects of nucleic acid structure are increasingly recognised; they also have implications for the design of oligonucleotides used in many experimental procedures.

CONCLUSION

It is clear that the simple rules currently used are inadequate for predicting hybridisation potential of long nucleic acid molecules, and even short molecules show unpredictable behaviour. There is a need for measurement of effects beyond those of nearest neighbouring bases. The data set needed for a comprehensive description of all possible interactions becomes very large even for relatively short-range interactions. In addition to the effects of dangling ends and other features discussed in this paper, there are important effects of intramolecular base pairing. Arrays of the kind used in this study and larger arrays comprising all sequences of a given length (10) offer the opportunity to study these effects systematically.

Table 1. The digitised image (Fig. 1) was analysed using xvseq, (Wang and Elder, in preparation, and see accompanying paper).

Decamer	Mean pixel value	Nonamer	Mean pixel value
TTATAGTaga	724	TTATAGTag	580
TTTATAGTag	1019	TTTATAGTa	890
TTTTATAGTa	1210	TTTTATAGT	801
CTTTTATAGT	1392		
tCTTTTATAG	1429	CTTTTATAG	1247
ttCTTTTATA	840	tCTTTTATA	839
tttCTTTTAT	330	ttCTTTTAT	298

A template corresponding to the set of overlapping circles used to create the array was laid over the image. Areas under the sections corresponding to all the oligonucleotides in the array were integrated and displayed as histograms. Each oligonucleotide provided two values, one for each side of the array. The two values were in good agreement except for a few areas where there is obvious high background. These last values were omitted; for the other oligonucleotides, the average of the two values is given. Each nonamer is paired with the decamer which differs by having a single extra base in the overhang, indicated by lower case letters. The underlined decamer is the perfect complement.This quantitative analysis confirms the visual impression that the 3′ overhangs make little difference to duplex yield.

ACKNOWLEDGEMENTS

This work was supported by Beckman Instruments and the Medical Research Council.

REFERENCES

1. Tinoco, I, Uhlenbeck, O.C., and Levine, M.D. (1971). *Nature*, **230**: 362–367.
2. Breslauer, K.J., Frank, R., Blöcker, H. and Marky, L. (1986). *Proc. Natl. Acad. Sci. USA*, **83**: 3746–3750.
3. Zuker M. (1989). In *Methods in Enzymology* Grossman and Moldave, eds. Acad. Press; London and New York. **180**: 262–288.
4. Southern E.M., Case-Green S.C., Elder J.K., Johnson M., Mir K.U., Wang L. and Williams J.C. (1994) *Nucleic Acids Research* **22**: 1368–1373.
5. Maskos U., and Southern E.M. (1992) *Nucleic Acids Research*, **20**: 1679–1684.
6. Freier S.M., Burger B.J., Alkema D., Neilson T. and Turner D.H. (1983). *Biochemistry* **22**: 6198–6206.
7. Petersheim M. and Turner D.H. (1983) *Biochemistry* **22**: 256–263.
8. Lane A., Martin S.R., Ebel S. and Brown T. (1992) *Biochemistry* **31**: 12080–12095.
9. Senior M., Jones R.A. and Breslauer K.J. (1988) *Biochemistry* **27**: 3879–3885. %8
10. Southern E.M., Maskos U., and Elder J.K. (1992) *Genomics*, **13**: 1008–1017.