Quarterly Reviews of Biophysics **38**, 4 (2005), pp. 311–320. © 2006 Cambridge University Press
doi:10.1017/S0033583506004197 Printed in the United Kingdom
First published online 31 May 2006

ESSAY

An extra dimension in nucleic acid sequence recognition

Keith R. Fox[1] and Tom Brown[2]*

[1] School of Biological Sciences, University of Southampton, Southampton, UK
[2] School of Chemistry, University of Southampton, Highfield, Southampton, UK

Abstract. Watson–Crick base pairing is a natural molecular recognition process that has been exploited in molecular biology and universally adopted in many fields. An additional mode of nucleic acid sequence recognition that could be used in combination with normal base pairing would add an exta dimension to nucleic acid interactions and open up many new applications. In principle the triplex approach could provide this if developed to recognize any DNA sequence. To this end modified nucleosides have been incorporated into triple-helix-forming oligonucleotides (TFOs) and used to recognize mixed sequence DNA with high selectivity and affinity at neutral pH. Continuing developments are directed towards improving TFO affinity at high pH and increasing triplex association kinetics. A number of applications of triplexes are currently being explored.

Introduction 312

Triple helices in DNA 312

Chemically modified TFOs 313

Further development 316
Recognition of GC base pairs 316
Recognition of TA base pairs 316
Recognition of AT base pairs 317
Recognition of CG base pairs 317

RNA triplexes 317

Kinetics of triplex formation 318

Practical applications of triplexes 318

Conclusions 319

References 319

* Author for correspondence: Tom Brown, School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, UK.
Tel.: +44 (0)2380 592974; Fax: +44 (0)2380 592991; E-mail: tb2@soton.ac.uk

https://doi.org/10.1017/S0033583506004197 Published online by Cambridge University Press

Introduction

Nucleic acid sequence recognition is a fundamental natural process upon which all living systems depend. It forms the basis of the scientific disciplines of molecular biology and genetics and underpins genomics, DNA diagnostics and modern aspects of forensic science. It involves the highly specific interaction of two strands of DNA by means of Watson–Crick base pairing (G-C and A-T). Other (mismatched) base pairs destabilize the duplex, thus ensuring that only complementary DNA strands combine. Mismatches maintain hydrogen-bonding interactions but perturb and weaken base-stacking interactions in the duplex [1–3]. In Nature, additional sequence-specificity is provided by DNA polymerases and proofreading enzymes which detect incorrect base pairs and remove them. Thus, Watson–Crick base pairing lies at the heart of an exquisitely sensitive and robust set of processes that ensures fidelity of replication and maintains the genetic integrity of all living systems. In practical applications such as DNA sequencing, SNP analysis and mutation detection, short synthetic pieces of single-stranded DNA (oligonucleotide primers and probes) are used in the same molecular recognition process. They are labelled, typically with fluorescent dyes and quenchers, for subsequent detection by suitable spectroscopic or spectrometric methods. In such applications it is necessary to separate the two strands of the target DNA duplex so that the oligonucleotide primer or probe can interact with one of these strands. The necessity for strand separation usually necessitates extra manipulations. Moreover, when the oligonucleotide probe binds to its DNA target there is a tendency for it to be displaced by the other DNA strand since the DNA:DNA duplex is thermodynamically more stable than the oligonucleotide:DNA duplex. This is a limiting factor in many applications such as fluorogenic methods of genetic analysis [4–6]. In therapeutics the problem of target accessibility is even more severe as DNA is only transiently single stranded (e.g. during transcription), and it is not possible to introduce a denaturation step *in vivo*. Methods of recognizing the sequences of *intact DNA duplexes* are, therefore, of great interest. Minor groove-binding agents have been explored in this context but recognition of extended sequences is challenging [7]. Moreover, the majority of proteins that bind to DNA do so in the major groove, so there may be advantages in targeting this region in order to interfere with primary metabolic processes. A readily accessible methodology for reading the sequence of an intact duplex and/or blocking the DNA major groove in a sequence-specific manner would open up exciting new areas of genetics, biotechnology, medicine and nanotechnology. It would also add an extra dimension to DNA sequence recognition which could then be extended to the RNA world (RNA intra- and inter-molecular duplexes and RNA:DNA hybrids).

Triple helices in DNA

Each Watson–Crick base pair presents a unique profile in the major groove of the DNA duplex, so it should be possible to specifically recognize (and 'read') any DNA sequence from this groove. It is well established that single-stranded oligonucleotides can bind in the major groove provided that one of the two strands of the duplex is purine-rich. The third strand (triplex-forming oligonucleotide, TFO) can bind in a parallel or antiparallel orientation to the purine strand of the duplex and parallel triplexes are intrinsically more stable, particularly at low pH (Fig. 1).

https://doi.org/10.1017/S0033583506004197 Published online by Cambridge University Press

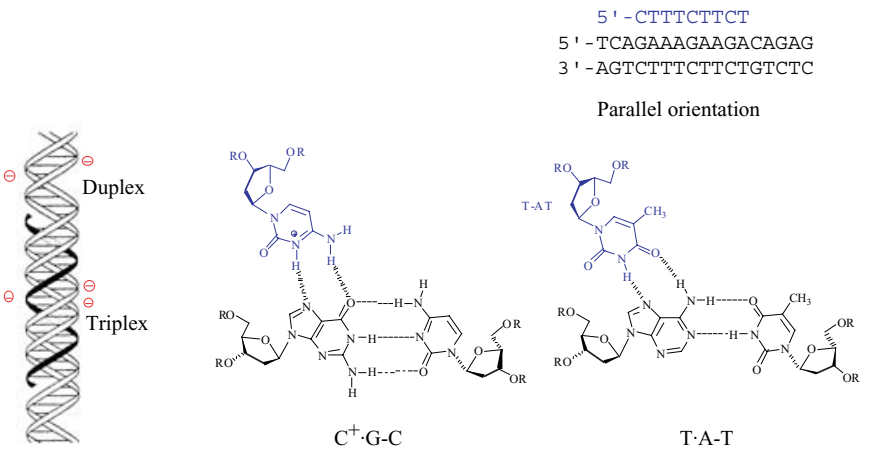


Fig. 1. Parallel triplets involving normal bases.

In parallel TFOs thymine recognizes A-T and cytosine recognizes G-C. There are two clear reasons for the instability of parallel triplexes at physiological pH:

(i) The $C^+ \cdot$ G-C triplet requires protonation which is unfavourable at physiological pH (the pK_a of cytosine is 4·5) [8].
(ii) The juxtaposition of three polyanionic DNA strands is destabilizing.

In addition to the low affinity of the TFO, there is another major limitation. TFOs that contain natural bases cannot bind to mixed-sequence DNA and only interact with poly-purine:poly-pyrimidine duplexes. There are relatively few DNA sequences of this kind in Nature.

Chemically modified TFOs

From the above limitations it is clear that natural nucleosides cannot be used as building blocks in TFOs for general DNA sequence recognition. Consequently TFOs containing altered bases, sugars and phosphates have been synthesized in attempts to produce stable triplexes at physiological pH. Although some compounds have been shown to be effective in isolation there is a dearth of literature on the use of *combinations* of modified nucleosides. Nevertheless, this provided a starting point for studies in our laboratory on DNA sequence recognition by chemically modified TFOs.

The nucleoside analogues in Fig. 2 constituted the 'first-generation' triplex monomers that have been studied by various biophysical techniques including DNase-1 footprinting and fluorescence melting. The pyrimidine analogue 3-methyl-2-aminopyridine (M, 2a) specifically recognizes G-C [9–12]. The pK_a of N(1) of this C-nucleoside is $\sim$6·2, which is much higher than that of cytosine. The thymidine analogue 5-(3-amino-1-propynyl)-2′-aminoethoxyuridine (B, 2b) has two primary amino groups that act co-operatively to neutralize the phosphate charges in the triplex [13, 14]. The propynyl moiety increases base stacking and as a consequence of these factors B has a much greater affinity than thymidine for A · T. The nucleoside analogue S (2C) binds to T · A base pairs [15, 16] and the *N*-methylpyrrolopyrimidine nucleoside P (2d) interacts selectively with CG base pairs [17].

https://doi.org/10.1017/S0033583506004197 Published online by Cambridge University Press

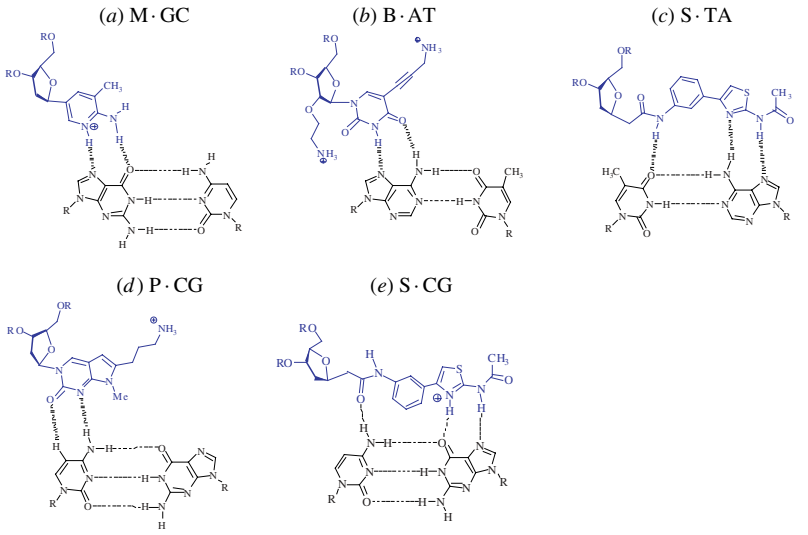


Fig. 2. Modified nucleosides used in parallel triplexes.

Triplex stability studies with TFOs containing the individual analogues in Fig. 2 provided information from which a number of hypotheses were made:

- The B · AT triad forms far more stable triplexes than T · AT and stability remains good up to pH 8, suggesting that protonation of the third strand promotes triplex formation. The propynyl moiety is important in stabilizing triplexes by improved base stacking [18]. The B · G-C 'mismatched' triplet is much less stable than B · A-T, and the B · C-G and B · T-A mismatched triplets are very unstable. Therefore, B has excellent selectivity for A-T and the degree of selectivity may be greater than for Watson–Crick base pairing [14, 19].
- M is specific for G-C base pairs and above pH 5·5 forms triplexes which are significantly more stable than those containing C · G-C. However, the M · G-C triad exhibits decreased affinity above pH 6. This is because the hydrogen bond between the N(1)-H of M and N(7) of G is only likely to exist in the lower pH range and its contribution to affinity and specificity will be lost at higher pHs. Protonation of the 3-methyl-2-aminopyridine base is likely to be an important factor in triplex stabilization by partly neutralizing the negative charge on neighbouring phosphodiesters. The 3-methyl group of aminopyridine contributes to triplex stability, probably through improved base stacking, although it also has a minor beneficial effect on the pK_a of N(1) [9–12].
- S binds to T · A with good affinity [15, 16] but also binds to C · G and to a lesser extent to A · T and G · C [20]. This may be a consequence of multiple conformations in this non-rigid system, and alternative binding modes of this nucleoside (e.g. groove binding or intercalation). A putative S · C-G triplet is shown in Fig. 2*e*.
- P binds preferentially to C · G [17], but there is scope to design analogues with increased triplet stability, particularly if the guanine base of the C-G base pair is involved in the recognition process.

The stability of nucleic acid duplexes, triplexes and higher order structures depends on hydrogen bonding and electrostatic, steric and hydrophobic interactions between bases, sugars and phosphates. These forces cannot currently be accurately calculated or predicted for unnatural

https://doi.org/10.1017/S0033583506004197 Published online by Cambridge University Press

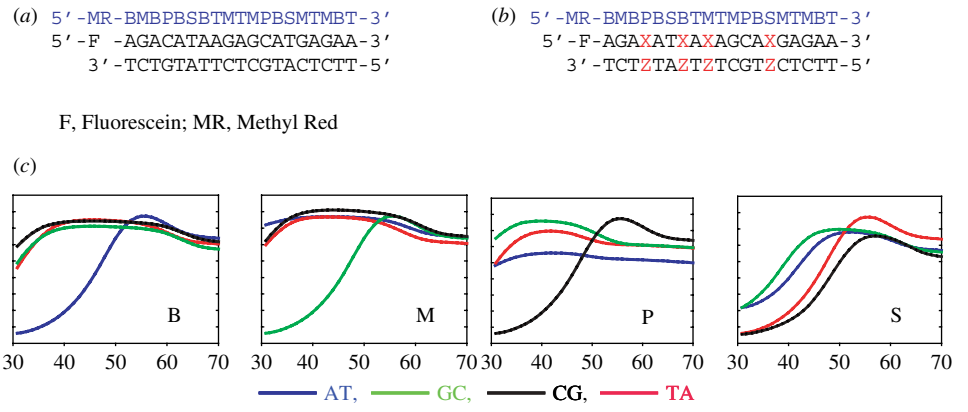


Fig. 3. First example of four-base recognition using triplexes at neutral pH [fluorescence *versus* temperature (°C)].

nucleosides in an aqueous salt environment, so at this stage in the design process it was important to adopt an empirical approach and incorporate all four 'first- generation' nucleosides in a TFO to study interactions with a series of mixed-sequence DNA duplexes. Accordingly, the phosphoramidite monomers of B, M, P and S were incorporated into a single TFO using standard solid-phase methods (sequence in blue in Fig. 3). Oligonucleotide assembly was efficient and HPLC purification gave one major product. The TFO was combined with 13 different duplexes, only one of which was the correct target (Fig. 3*a*). The other 12 had a single mismatched triplet (Fig. 3*b* where XZ are base pairs other than those in Fig. 3*a*). Fluorescence melting studies [21] at pH 7·0 showed that the triplex formed with the correct target was the most stable, and the triplets formed between the modified nucleosides B, M, and P in the TFO and incorrect Watson–Crick base pairs were extremely unstable (Fig. 3*c*). This is an important observation, demonstrating that triplex stability can be extremely sensitive to the presence of a single incorrect triplet. The results of footprinting experiments provided confirmation of these results [22].

Results from combining these first-generation analogues highlighted some important areas for future development.

- As indicated in previous studies [15, 16], S discriminated clearly between pyrimidine-purine and purine-pyrimidine base pairs but failed to discriminate between TA and CG. For many applications this degree of selectivity is probably sufficient but there is scope for an improved nucleoside analogue to recognise T-A base pairs.
- The T_m of the fully matched triplex was 49 °C at pH 7·0, 58 °C at pH 6·5, 62 °C at pH 6·0 and 64 °C at pH 5·5. This pH dependence can be attributed largely to the 3-methyl-2-aminopyridine · G-C triplets, indicating that the pK_a of M is too low for optimum triplex stability at elevated pHs. It would be desirable to have an alternative nucleoside analogue to recognize G-C base pairs at higher pH.

The above results were obtained from fluorescence melting studies [21] on a Roche LightCycler (Roche Diagnostics Ltd, Lewes, E. Sussex, UK), an instrument designed for genetic analysis. The technique of fluorescence melting is particularly valuable for triplex studies as it enables the triplex melting process to be clearly differentiated from duplex melting. This is not straightforward by the traditional ultraviolet melting method, particularly when very stable triplexes are

https://doi.org/10.1017/S0033583506004197 Published online by Cambridge University Press

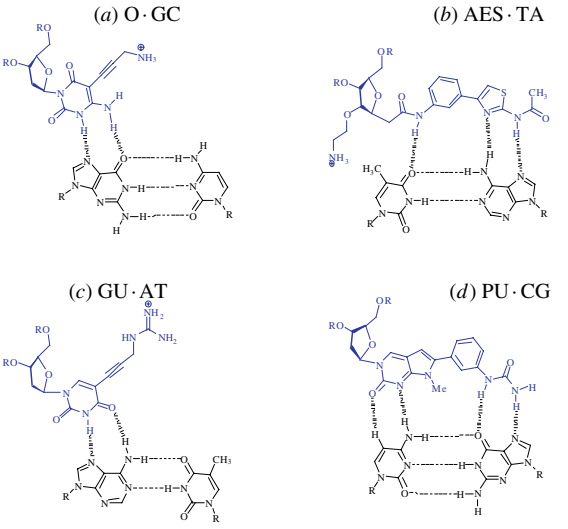


Fig. 4. Second-generation nucleosides for use in parallel triplexes.

being analysed. An added advantage of the fluorescence method is the high throughput of the LightCycler instrument on which 32 simultaneous melting curves can be run. This facilitates the determination of triplex melting under a wide range of conditions (e.g. dependence on salt concentration, divalent cations, pH). In triplex melting experiments in is important that slow heating and cooling rates are used, as the kinetics of triplex formation and dissociation are much slower than for duplexes.

Further development

Recognition of GC base pairs

The pH problem in G-C recognition might be eliminated by means of a nucleoside analogue such as 5-(3-amino-1-propynyl)-6-oxodeoxycytidine (O, Fig. 4*a*). The aminopropynyl group should increase base stacking and also introduce a site of protonation. This modification has been shown to be effective in propargylamino dU [23, 24] and bis-amino U [14, 19] (Fig. 2*b*). 6-Oxo cytidine analogues without this substituent do not significantly stabilize triplexes [25, 26].

Recognition of TA base pairs

An improvement in TA recognition has recently been achieved by synthesis of the 2′-aminoethoxy analogue of S (Fig. 4*b*) [20]. This analogue is more selective than S, forming more stable triplets with T-A than with C-G. The origin of the improved selectivity is as yet unclear. The AES · T-A triplet is slightly more stable than S · T-A, particularly when multiple TA inversions are present in the target duplex. This is consistent with previous observations that the 2′-aminoethoxy group increases triplex stability (B · A-T is more stable than 5-(3-amino-1-propynyl) dU · A-T [13, 14] and 2′-aminoethoxy T · A-T is more stable than T · A-T [27, 28]). It has recently been demonstrated that TFOs in which all the sugars are 2′-aminoethoxyribose form very stable triplexes [29].

https://doi.org/10.1017/S0033583506004197 Published online by Cambridge University Press

Recognition of AT base pairs

The selectivity and stability of the B · A-T triplet do not require any improvement. However, there are cases when it might be necessary to conjugate TFOs to crosslinking agents such as psoralen for *in vivo* gene modification [30–34]. This can be done by reacting a terminal aminoalkyl group on the TFO with an active ester of psoralen. If the TFO contains bis amino U, the psoralen would be expected to react at several amino groups and the oligonucleotide would become multiply labelled. Another strategy would be to incorporate psoralen as a phosphoramidite during oligonucleotide synthesis. However, the conditions required to deprotect the 2′-phthalimidoethyl group of B (aqueous methylamine) are too harsh and would cause decomposition of the psoralen. The propynylguanidinium deoxyuridine nucleoside (Fig. 4*c*) [35] could be used to circumvent these problems as the protonated guanidine would remain chemically unreactive during the psoralen-labelling reaction. The guanidyl nucleoside might also be useful in stabilizing triplexes above pH 8.

Recognition of CG base pairs

The recognition of CG base pairs by P (Fig. 2) is reasonably selective [17] but it would be desirable to produce a monomer that could form more stable triplets than P. The interaction between P and C-G is relatively weak due to the fact that there are only two inter-base hydrogen bonds, one of which is a C-H···O bond. Moreover, hydrogen bonding is exclusively between P and cytosine and does not involve guanine. The analogue (PU, 4d) is designed to recognize both bases of the C-G base pair and should form four hydrogen bonds, producing a true triad. In theory this should produce a more stable and selective triple helix.

RNA triplexes

RNA is of critical importance in cellular and viral life cycles. Many large RNAs are composed of complex arrangements of short single- and double-stranded regions joined by structures such as loops, bulges, and pseudoknots that are recognized in nature by proteins, nucleic acids, and small molecules. There is considerable interest in the design of molecules that interfere with RNA function, not only to inhibit the transcription and translation of messenger RNAs but also to block the functions of catalytic and regulatory RNAs [36]. Therefore, studies on RNA triplexes are of great interest [37, 38]. There is strong evidence to suggest that TFOs can interact with RNA duplexes to form stable triplexes and RNA:DNA hybrids can also form triplexes [39–44]. A study on the rev responsive element of HIV-1 has shown that RNA can be targeted by chimeric oligonucleotides that contain two separate linked sequences: one that interacts with a single-stranded region of RNA to form a duplex and another that forms a triplex with a nearby double-stranded region of the rev responsive element (RRE) [39]. The advantage of this approach is that duplex formation is kinetically very fast, but as only short stretches of RNA are single stranded, the ability to form a triplex with an adjacent double-stranded RNA greatly increases the thermodynamic stability of the complex. These studies were carried out using TFO composed of natural nucleosides and the binding of modified TFOs to RNA is not well studied. The modified nucleosides discussed above may form stable triplexes with RNA and the 2′-aminoethoxy substituent should confer further triplex stabilization by imposing an A-like conformation on the TFO [29]. Selective recognition of the UA base pair of RNA by the TFO presents a different problem from the

https://doi.org/10.1017/S0033583506004197 Published online by Cambridge University Press

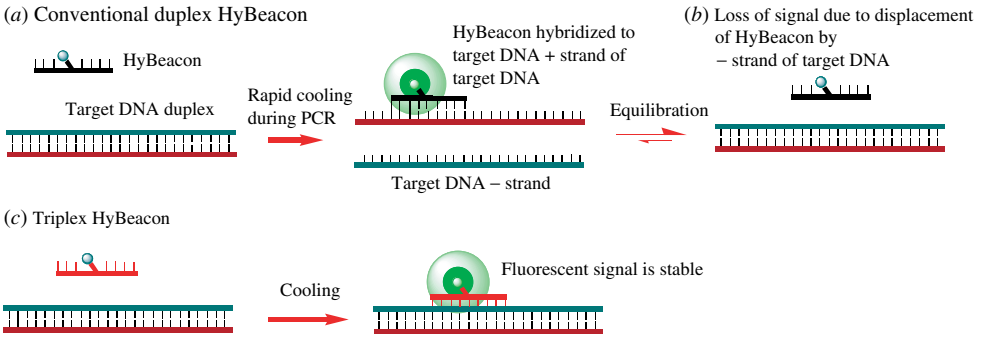


Fig. 5. Conventional HyBeacons and Triplex HyBeacons.

DNA counterpart, as uracil has a hydrogen atom in the 5-position as opposed to the methyl group of thymine.

Kinetics of triplex formation

The kinetics of triplex association is slow compared to duplex formation [39, 45–48] and this will be a limitation in some applications. The slow association of the third strand with the duplex could be due to conformational effects, i.e. the TFO might be required to adopt a particular shape and this may not be readily accessible. Repulsive electrostatic forces between the anionic third strand and duplex might also be a factor. Triplex dissociation is slower for some chemically modified TFOs, particularly those with amino groups that can partly neutralize the charge on the phosphodiester backbones. DNA-binding moieties (e.g. intercalators) can be attached to TFOs to speed up the initial binding process but this could lead to a reduction in sequence selectivity. The kinetics of association and dissociation of chemically modified triplexes are currently under investigation in our laboratory and it will be a challenge to design TFOs with favourable selectivity, thermodynamic and kinetic properties.

Practical applications of triplexes

A key aim of this research is to develop triplexes that can be used routinely to recognize any mixed-sequence DNA duplex at physiological pH so that the technology is available for exploitation by the scientific community. In addition to their use as anti-gene agents, there are several potential biochemical applications that could benefit from the slow dissociation kinetics of triplexes. For example, PCR products could be isolated by binding to a biotinylated triplex-forming oligonucleotide and then capturing the triplex using magnetic beads. The impurities (together with primers, triphosphates, enzymes, etc.) could then be washed off provided that the triplex dissociation kinetics is slow at the washing temperature. Heating to a specific temperature or increasing the pH would then liberate the pure duplex but leave the TFO attached to the beads. This ‘triplex affinity capture’ approach has been previously demonstrated, but has been limited to oligopurine tracts at low pHs [49]. In a related context triplexes could also be used in genomic analysis on DNA arrays, either in an immobilized form to capture intact duplexes of specific sequence, or as fluorescent probes for immobilized double-stranded PCR products. For the triplex approach to be generally useful in the above applications, four-base-pair recognition is an absolute requirement.

https://doi.org/10.1017/S0033583506004197 Published online by Cambridge University Press

Triplex-forming oligonucleotides could also be used in fluorescence-based genetic analysis. The approach will be particularly suitable for 'end-point' assays, for example in triplex HyBeacons. A HyBeacon is an oligonucleotide probe with a single internal fluorescent label attached to a base [50–53]. The mode of action of normal HyBeacons is shown in Fig. 5*a*. The unhybridized HyBeacon is weakly fluorescent due to quenching from the heterocyclic bases. On hybridization to a target sequence, the spatial relationship between the fluorescent dye and the bases is changed, resulting in greatly reduced quenching and a significant increase in fluorescence.

One problem with all commonly used methods of genetic analysis is illustrated in Fig. 5*b*. The probe/target hybrid is formed during the rapid cooling step of PCR as it is kinetically favoured. However, it is thermodynamically less stable than the much longer double-stranded target duplex so the probe becomes displaced, leading to a rapid decrease in fluorescent signal. The effect is more pronounced with molecular beacons which are sequestered into a stable hairpin loop when displaced [5]. HyBeacons have no intramolecular secondary structure, but signal decay is still too fast for convenient and reliable endpoint analysis. Asymmetric PCR has been used to reduce the concentration of the opposite strand of the target but this is an inefficient process. The problem might be eliminated by the use of Triplex HyBeacons (Fig. 5*c*) for which the intact duplex is the target, so no displacement of the probe should occur. This is an interesting concept, as end-point genetic analysis is currently being developed for rapid point of care diagnostics in general practitioners' surgeries. The aim is to provide a 'while you wait' diagnostic service, which if successful will represent a major advance in patient care.

Conclusions

This is a problem for which there is not a single solution. There are many approaches to nucleic acid sequence recognition and in recent years a variety of unnatural nucleic acid analogues have been produced that can bind to the major groove of DNA. There has been significant progress in the synthesis of chemically modified triplex-forming oligonucleotides and associated biophysical studies have produced encouraging results. General acceptance of the triplex approach, and the conversion of basic science to important and useful technology depends upon the ready availability of TFOs with favourable thermodynamic and kinetic properties at physiological pH. This remains a significant challenge and its successful implementation could have far-reaching consequences in biology, medicine, diagnostics and nanotechnology.

References

1. Brown, T., Kennard, O., Kneale, G. & Rabinovich, D. (1985). *Nature* **315**, 604–606.
2. Hunter, W. N., Brown, T., Anand, N. N. & Kennard, O. (1986). *Nature* **320**, 552–555.
3. Brown, T., Hunter, W. N., Kneale, G. G. & Kennard, O. (1986). *Proc. Natl. Acad. Sci. USA* **83**, 2402–2406.
4. Holland, P. M., Abramson, R. D., Watson, R. & Gelfand, D. H. (1991). *Proc. Natl. Acad. Sci. USA* **88**, 7276–7280.
5. Tyagi, S. & Kramer, F. R. (1996). *Nature Biotech.* **14**, 303.
6. Whitcombe, D., Theaker, J., Guy, S. P., Brown, T. & Little, S. (1999). *Nature Biotech.* **17**, 804–807.
7. Urbach, A. R. & Dervan, P. B. (2001). *Proc. Natl. Acad. Sci. USA* **98**, 4343–4348.
8. James, P. L., Brown, T. & Fox, K. R. (2003). *Nucleic Acids Res.* **31**, 5598–5606.
9. Bates, P. J., Laughton, C. A., Jenkins, T. C., Capaldi, D. C., Roselt, P. D., Reese, C. B. & Neidle, S. (1996). *Nucleic Acids Res.* **24**, 4176–4184.
10. Hildbrand, S., Blaser, A., Parel, S. P. & Leumann, C. J. (1997). *J. Am. Chem. Soc.* **119**, 5499–5511.

https://doi.org/10.1017/S0033583506004197 Published online by Cambridge University Press

11. Chen, D. L. & McLaughlin, L. W. (2000). *J. Org. Chem.* **65**, 7468–7474.
12. Cassidy, S. A., Slickers, P., Trent, J. O., Capaldi, D. C., Roselt, P. D., Reese, C. B., Neidle, S. & Fox, K. R. (1997). *Nucleic Acids Res.* **25**, 4891–4898.
13. Sollogoub, M., Dominguez, B., Fox, K. R. & Brown, T. (2000). *Chem. Commun.* (23), 2315–2316.
14. Sollogoub, M., Darby, R. A. J., Cuenoud, B., Brown, T. & Fox, K. R. (2002). *Biochemistry* **41**, 7224–7231.
15. Guianvarc'h, D., Benhida, R., Fourrey, J. L., Maurisse, R. & Sun, J. S. (2001). *Chem. Commun.* (18), 1814–1815.
16. Guianvarc'h, D., Fourrey, J. L., Maurisse, R., Sun, H. S. & Benhida, R. (2003). *Bioorg. Med. Chem.* **11**, 2751–2759.
17. Ranasinghe, R. T., Rusling, D. A., Powers, V. E. C., Fox, K. R. & Brown, T. (2005). *Chem. Commum.* (20), 2555–2557.
18. Lacroix, L., Lacoste, J., Reddoch, J. F., Mergny, J. L., Levy, D. D., Seidman, M. M., Matteucci, M. D. & Glazer, P. M. (1999). *Biochemistry* **38**, 1893–1901.
19. Osborne, S. D., Powers, V. E. C., Rusling, D. A., Lack, O., Fox, K. R. & Brown, T. (2004). *Nucleic Acids Res.* **32**, 4439–4447.
20. Wang, Y., Rusling, D. A., Powers, V. E. C., Lack, O., Osborne, S. D., Fox, K. R. & Brown, T. (2005). *Biochemistry* **44**, 5884–5892.
21. Darby, R. A. J., Sollogoub, M., McKeen, C., Brown, L., Risitano, A., Brown, N., Barton, C., Brown, T. & Fox, K. R. (2002). *Nucleic Acids Res.* **30**, art. no.-e39.
22. Rusling, D. A., Powers, V. E. C., Ranasinghe, R. T., Wang, Y., Osborne, S. D., Brown, T. & Fox, K. R. (2005). *Nucleic Acids Res.* **33**, 3025–3032.
23. Bijapur, J., Keppler, M. D., Bergqvist, S., Brown, T. & Fox, K. R. (1999). *Nucleic Acids Res.* **27**, 1802–1809.
24. Gowers, D. M., Bijapur, J., Brown, T. & Fox, K. R. (1999). *Biochemistry* **38**, 13747–13758.
25. Berressem, R. & Engels, J. W. (1995). *Nucleic Acids Res.* **23**, 3465–3472.
26. Parsch, U. & Engels, J. W. (2000). *Chem. Eur. J.* **6**, 2409–2424.
27. Blommers, M. J. J., Natt, F., Jahnke, W. & Cuenoud, B. (1998). *Biochemistry* **37**, 17714–17725.
28. Cuenoud, B., Casset, F., Husken, D., Natt, F., Wolf, R. M., Altmann, K.-H., Martin, P. & Moser, H. E. (1998). *Angew. Chem. – Int. Ed.* **37**, 1288–1291.
29. Buchini, S. & Leumann, C. J. (2004). *Angew. Chem. – Int. Ed.* **43**, 3925–3928.
30. Barre, F. X., Giovannangeli, C., Helene, C. & Harel-Bellan, A. (1999). *Nucleic Acids Res.* **27**, 743–749.
31. Duval-Valentin, G., Takasugi, M., Helene, C. & Sage, E. (1998). *J. Mol. Biol.* **278**, 815–825.
32. Guieysse, A. L., Praseuth, D., Giovannangeli, C., Asseline, U. & Helene, C. (2000). *J. Mol. Biol.* **296**, 373–383.
33. Raha, M., Lacroix, L. & Glazer, P. M. (1998). *Photochem. Photobiol.* **67**, 289–294.
34. Takasugi, M., Guendouz, A., Chassignol, M., Decout, J. L., Lhomme, J., Thuong, N. T. & Helene, C. (1991). *Proc. Natl. Acad. Sci. USA* **88**, 5602–5606.
35. Roig, V. & Asseline, U. (2003). *J. Am. Chem. Soc.* **125**, 4416–4417.
36. Vuyisich, M. & Beal, P. A. (2000). *Nucleic Acids Res.* **28**, 2369–2374.
37. Roberts, R. W. & Crothers, D. M. (1992). *Science* **258**, 1463–1466.
38. Han, H. & Dervan, P. B. (1993). *Proc. Natl. Acad. Sci. USA* **90**, 3806–3810.
39. Moses, A. C. & Schepartz, A. (1997). *J. Am. Chem. Soc.* **119**, 11591–11597.
40. Escude, C., Sun, J. S., Rougee, M., Garestier, T. & Helene, C. (1992). *Comptes Rendus Acad. Sci. Ser. III – Sci. Vie-Life Sci.* **315**, 521–525.
41. Escude, C., Francois, J. C., Sun, J. S., Ott, G., Sprinzl, M., Garestier, T. & Helene, C. (1993). *Nucleic Acids Res.* **21**, 5547–5553.
42. Holland, J. A. & Hoffman, D. W. (1996). *Nucleic Acids Res.* **24**, 2841–2848.
43. vanDongen, M. J. P., Heus, H. A., Wymenga, S. S., vanderMarel, G. A., vanBoom, J. H. & Hilbers, C. W. (1996). *Biochemistry* **35**, 1733–1739.
44. Raghunathan, G., Miles, H. T. & Sasisekharan, V. (1995). *Biopolymers* **36**, 333–343.
45. Bernal-Mendez, E. & Leumann, C. J. (2002). *Biochemistry* **41**, 12343–12349.
46. Ellouse, C., Piot, F. & Takahashi, M. (1997). *J. Biochem.* **121**, 521–526.
47. Paes, H. M. & Fox, K. R. (1997). *Nucleic Acids Res.* **25**, 3269–3274.
48. Fox, K. R. (1995). *FEBS Lett.* **357**, 312–316.
49. Ito, T., Smith, C. L. & Cantor, C. R. (1992). *Nucleic Acids Res.* **20**, 3524–3524.
50. McDowell, D., French, D. & Brown, T. (2000). *J. Med. Genet.* **37**, SP12.
51. French, D. J., Archard, C. L., Brown, T. & McDowell, D. G. (2001). *Mol. Cell Probes* **15**, 363–374.
52. French, D. J., Archard, C. L., Andersen, M. T. & McDowell, D. G. (2002). *Mol. Cell Probes* **16**, 319–326.
53. Dobson, N., McDowell, D. G., French, D. J., Brown, L. J., Mellor, J. M. & Brown, T. (2003). *Chem. Commun.* (11), 1234–1235.

https://doi.org/10.1017/S0033583506004197 Published online by Cambridge University Press