Recent highlights in modified oligonucleotide chemistry

Alexander J. A. Cobb*

Received 27th June 2007
First published as an Advance Article on the web 6th September 2007
DOI: 10.1039/b709797m

The synthesis of modified nucleic acids has been the subject of much study ever since the structure of DNA was elucidated by Watson and Crick at Cambridge and Wilkins and Franklin at King's College over half a century ago. This review describes recent developments in the synthesis and application of these artificial nucleic acids, predominantly the phosphoramidites which allow for easy inclusion into oligonucleotides, and is divided into three separate sections. Firstly, modifications to the base portion will be discussed followed secondly by modifications to the sugar portion. Finally, changes in the type of nucleic acid linker will be discussed in the third section. Peptide Nucleic Acids (PNAs) are not discussed in this review as they represent a separate and large area of nucleic acid mimics.

1. Introduction

Modified nucleic acids continue to generate a large amount of interest to the synthetic and medicinal chemist. The reasons for this are numerous, but perhaps most topical are the recent developments in RNA interference (RNAi) technology for which Andrew Fire and Craig Mello were awarded the Nobel Prize for medicine in 2006.[1] This technology, discovered in 1998, has rapidly become one of the most powerful tools available to molecular biologists.[2] RNAi uses short double-stranded RNA (dsRNA) molecules to switch genes on and off. The beauty of this approach is that the RNA can be designed to be specific to the gene sequence of interest by binding to the mRNA coded by that gene, hence repressing any protein synthesis from that mRNA. When the dsRNA enters a cell, it is cleaved into shorter segments of roughly 22 nucleotides, known as short interfering RNAs (siRNAs), and it is these segments which bind to the mRNA.

Related to this is the concept of antisense oligonucleotides, which directly target the mRNA through the complimentary single-stranded RNA (ssRNA) sequence. The mRNA is then either physically blocked as in RNAi, or the enzyme RNase H digests the RNA strand of the DNA–RNA hybrid.[3]

The role of *modified* nucleic acids in these technologies is to provide an oligonucleotide mimic which can not only bind to the mRNA in a complimentary fashion, but will also be far less susceptible to degradation pathways (*e.g.* endonucleases), which are one of the major stumbling blocks for RNAi and antisense technologies. Unnatural oligonucleotides have therefore emerged, where the linker or (deoxy)ribose portion has been modified in some way to be resilient to these exogenous processes.

Another reason for modification of the nucleic acids is the desire to expand the genetic code, as well as to understand the scope and limits of Watson–Crick base-pairing. Changes to the sugar portion (*e.g.* use of a hexose instead of a pentose) or to the nucleobase portion (through expanding the size of the naturally occurring bases) go a long way to answering these questions.

School of Pharmacy, University of Reading, Whiteknights, Reading, Berks, RG6 6AD, UK. E-mail: a.j.a.cobb@reading.ac.uk; Fax: +44 (0)118 378 4638; Tel: +44 (0)118 378 4644

Andre Cobb

Andre Cobb was born in London, UK, in 1975 and obtained his BSc in Chemistry from King's College London in 1997. He moved to UCL for his research degree under the guidance of Professor Charles M. Marson, where he worked in the area of zinc-mediated asymmetric processes. After three years of postdoctoral study with Professor Steven V. Ley at the University of Cambridge, he was appointed in 2005 to the faculty in the School of Pharmacy at the University of Reading, where he is currently a lecturer. Cobb's research interests are in the design and synthesis of modified nucleic acids as probes for enzyme function, natural product synthesis, organocatalysis and medicinal chemistry.

2. Modifications to the nucleobase

(a) Probing nucleobase pairing

Perhaps the most impressive work in changing the base portion of the nucleic acids in deoxyribose nucleic acids has been undertaken by Kool and co-workers[4] over the years, who have synthesised a range of expanded nucleobases, primarily to see how these modifications affect the Watson–Crick pairing, but ultimately to expand the genetic code. They have synthesised several variations of these base-expanded analogues through the use of phosphoramidite chemistry and have termed these modified sequences xDNA ('expanded DNA'),[5] yDNA ('wide DNA')[6] and yyDNA ('double wide DNA').[7]

In the case of xDNA, each base has been expanded by the width of a phenyl ring—around 2.4 Å (Fig. 1). The expanded *C*-linked cytidine (dxC) phosphoramidite **11** was obtained from the expanded thymidine (dxT) **5**, which itself was synthesised through the Heck-type coupling of aryl iodide **1** to the glycal **2** using a $Pd(OAc)_2/AsPh_3$ combination to give product **3**. Removal of the *tert*-butyldiphenylsilyl protecting group gave ketone **4**, which was stereoselectively reduced to deoxyribose **5** (Scheme 1).[5d] This established[8] method of *C*-nucleoside formation was also used by McLaughlin and co-workers in their synthesis of dC and dU analogues, which could be used to probe enzyme–substrate interactions.[9]

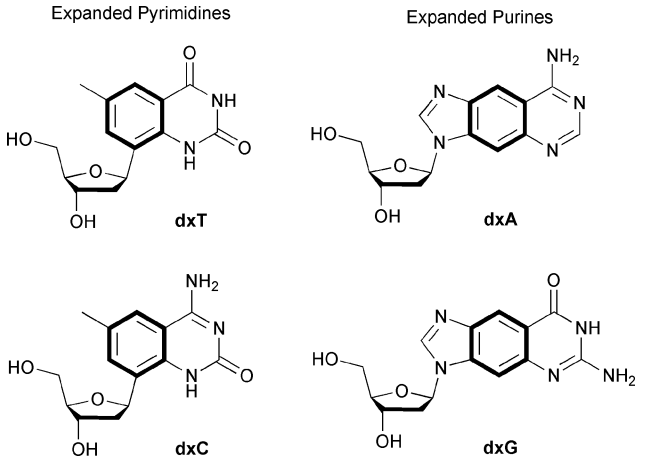


Fig. 1 Kool's expanded xDNA set: each of these xDNA bases is a homologation of the naturally occurring bases with the exception of dxC, which is strictly an expansion of 5-MeC. Synthetically, however, it was easier to access owing to its shared starting materials with dxT.

The expanded *N*-linked nucleobases of dxA[5f] and dxG[5d] were accessed through nucleobases **12** and **13** respectively (Scheme 2). These glycosidations were effected through the two different methods shown in Scheme 3. dxA was synthesised through the use of Hoffer's α-chlorosugar **14** (Scheme 2a)—a method which due to the lack of stereocontrol generated four products; two regioisomers resulting from both nitrogens on the imidazole ring of **11** being able to undergo glycosidation, and the two anomers associated with each of these regioisomers. Kool and co-workers go some way to addressing this problem with their synthesis of dxG, which utilises neighbouring group participation to form exclusively the desired β-anomer, although two regioisomers still form, and it is necessary to separate them at a later stage. Furthermore, as it is a deoxyribose that is ultimately required, the alcohol in the 2′-position needs to be removed later on, and this is achieved *via* a Barton–McCombie-like deoxygenation. Subsequent removal of the silicon protecting group exposes the 3′ and 5′ alcohol functionalities, which can be converted to the phosphoramidite through conventional methods for incorporation into an oligonucleotide. These four new expanded nucleobases thus allowed for the investigation of an eight-base genetic code comprising the four natural DNA bases and the four size-expanded ones.

Using dxA as a dA analogue to form Watson–Crick pairs with dT, it was shown that the self-complimentary 10-mer, 5′-d(xATxAxATxATTxAT), formed a more thermodynamically stable duplex with itself than it did with the corresponding DNA sequence, 5′-d(ATAATATTAT). The main implication of this is that there is a greater helical diameter, and this was shown clearly by CD analysis. Remarkably, solution characterisation by NMR showed that there were many similarities with DNA. The overall structure retained its B-DNA characteristics, having both right-handedness and the same Watson–Crick pairings as DNA, the glycosidic bond conformations were *anti*, and all deoxysugar conformations were 2′-*endo*.[5e]

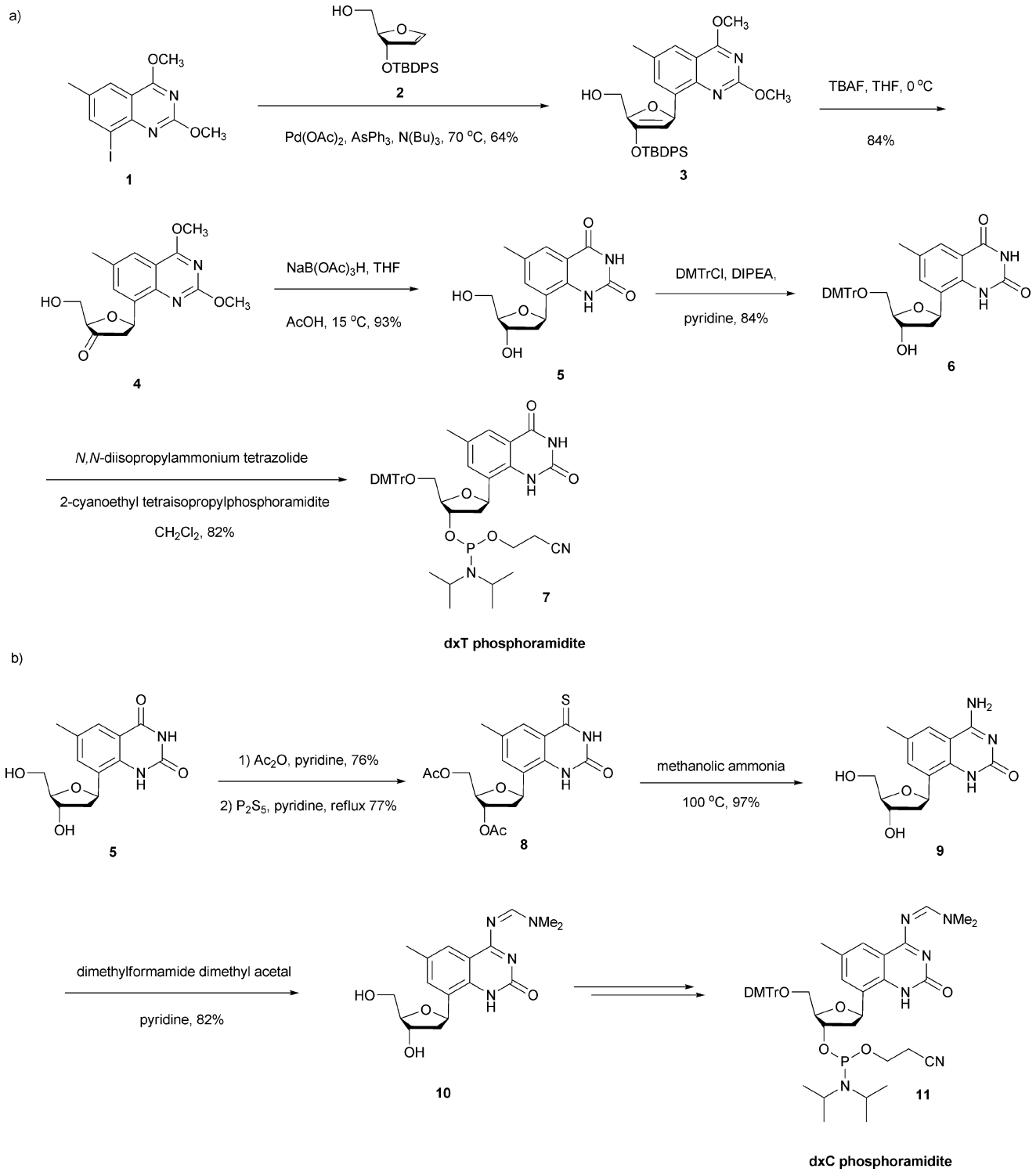
Furthermore, examination of the thermal stability of all these expanded bases and all their possible base-pairings showed that they exhibited sequence recognition. Indeed, all four xDNA bases were shown to have similar mismatching selectivity preferences as natural DNA. In addition to generating new Watson–Crick pairs, these expanded nucleobases were all found to be fluorescent, giving them the potential to be used as biophysical tools.[5b]

yDNA differs from xDNA by virtue of its changed geometry around the benzohomologated system (Fig. 2). When yC (which is actually an analogue of 5-MeC) was incorporated into a self-complimentary sequence, the resulting complex was found to be more thermally stable than the DNA controls. yT was also investigated, and this too showed dramatic stabilisation compared to the DNA controls, with a melting temperature 43 °C higher and an estimated free energy that was more favourable by 11 kcal mol^{-1}. The stability of self-complimentary sequences that contained yC and yT with G and A partners was also measured, and these too showed a high degree of stability compared to the controls. Interestingly, the difference in melting temperatures between an oligonucleotide containing yT and a *mismatched* sequence was greater than for the DNA control, suggesting that the selectivity for the yDNA is higher than that for DNA. The ordering of the mismatch stability was the same in both cases ($A \gg G > T \geq C$). Furthermore, yDNA was also shown to have a preference for antiparallel binding.

The utility of this innovation has been beautifully demonstrated by a natural DNA octameric sequence being recognised using an oligonucleotide entirely composed of yDNA. The subsequent complex had a melting temperature 24 °C higher and an estimated free energy >4 kcal mol^{-1} more favourable than the natural complimentary DNA oligonucleotide.[6a]

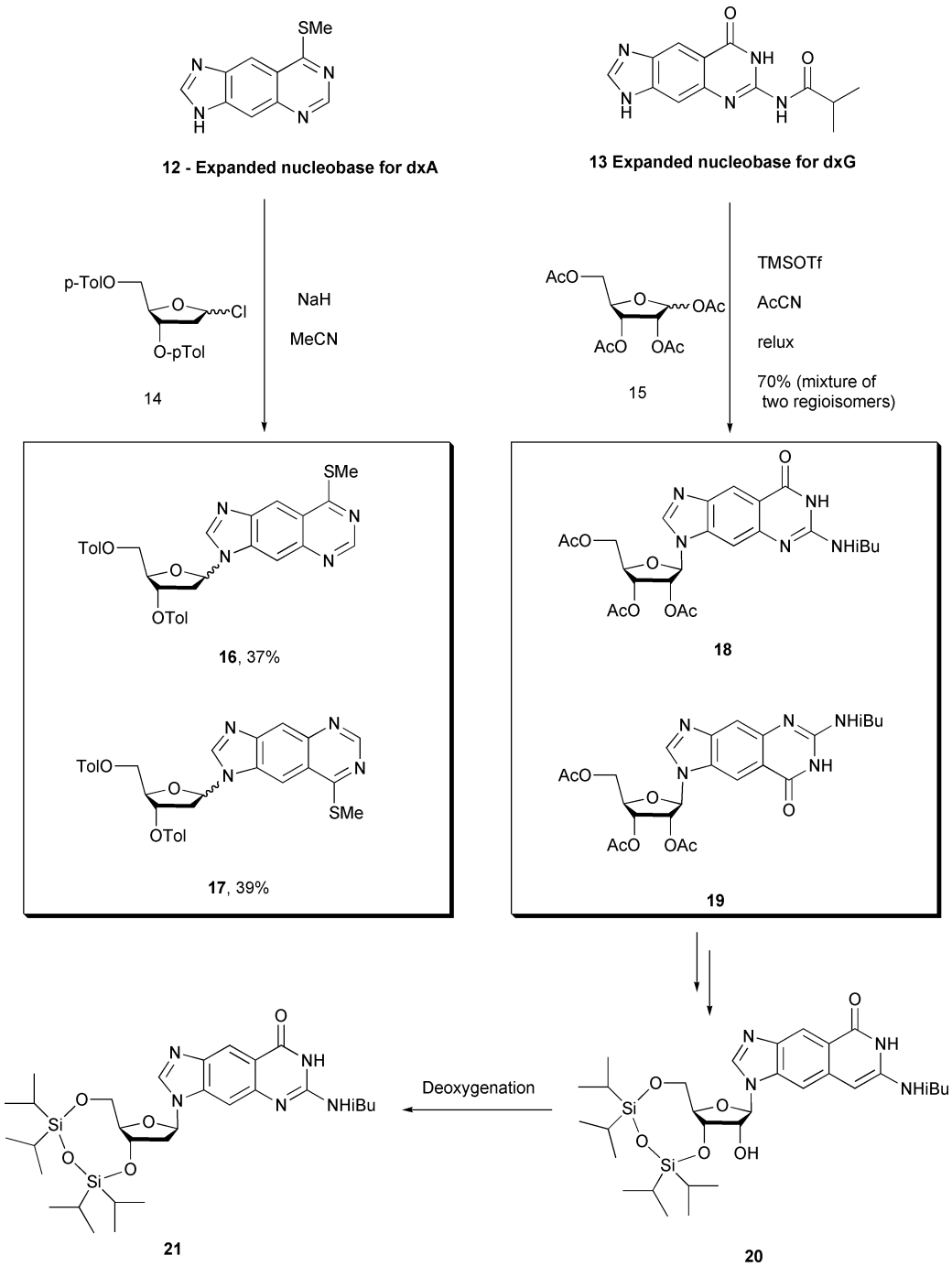
yyDNA represents the set of naptho-homologated DNA bases and pairs (specifically dyyT and dyyC, Fig. 3), and explores the limits for the size of DNA-like molecules. This has demonstrated that they have an unusually strong stacking ability and, as expected, were fluorescent. Perhaps more impressive is the fact that these size-expanded nucleosides also form stable helices.[7]

Triplex-forming oligonucleotides (TFOs) have also been the subject of much study, since an oligonucleotide that can recognise and bind to a sequence-specific portion of dsDNA would allow transcriptional activation and deactivation, as well a range of other molecular biology techniques such as targeted mutagenesis and targeted recombination.[10] Generally, the use of triplex targeting has been limited because the TFO can only recognise and bind to purine bases (*via* Hoogsteen pairs) in dsDNA. This then limits the formation of triple-stranded structures to regions of dsDNA which are furnished with homopurine sequences (*i.e.* A or G),



Scheme 1 Kool's synthesis of expanded pyrimidine bases dxT and dxC.

because the TFO cannot traverse back and forth between the two strands of dsDNA—not only is there the issue of having to span approximately 10 Å, but the polaritiy of the TFO would also have to be reversed (Fig. 4). In a series of papers, Gold and co-workers have developed *C*-nucleosides that are designed to address these issues and to specifically recognise T:A, A:T, G:C or C:G base pairs in dsDNA, thus participating in triplex formation with them.[11–14] They achieved this through the design of *C*-nucleosides which can easily change their conformation to bind with the purine base regardless of which strand it is on. Since the glycosidic bond is so close to the centre of the major groove, the distance that the TFO has to traverse is much shorter and the problem of changing strand



Scheme 2 Kool's synthesis of extended purine analogues.

polarity is removed because Gold's TFO reads purine information on either strand in a single direction (Fig. 5).

The most recent of Gold's syntheses is that of the fluorinated antiAT phosphoramidite precursor **25** which was able to be incorporated into an oligonucleotide, unlike the non-fluorinated version.[11] The synthesis was achieved by coupling the 4-fluoroaniline salt **22** with ethyl cyanoacetate to generate **23**. This was converted to bromide **24** and glycosidated *via* a Heck coupling and reduction as described previously (Scheme 3).

Brown and co-workers[15] have also synthesised a triplex-forming nucleoside analogue, which, this time, was capable of recognising CG inversions. Mimics **26–28** (Fig. 6) were synthesised *via* the 3*H*-furo[2,3-*d*]pyrimidin-2-one ring **32** (Scheme 4). These furopyrimidinones have been found to be unstable to oligonucleotide deprotection conditions, with the oxygen being replaced with the NH of ammonia. Brown and co-workers exploit this in order to access derivatives **26–28** by allowing it to react with methylamine so as to insert NMe into the ring. They propose that the triplet

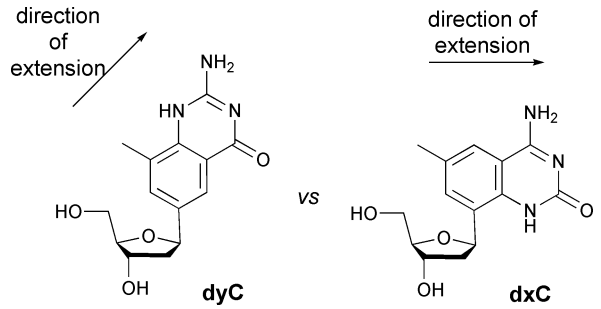


Fig. 2 Comparison of a yDNA base and an xDNA base.

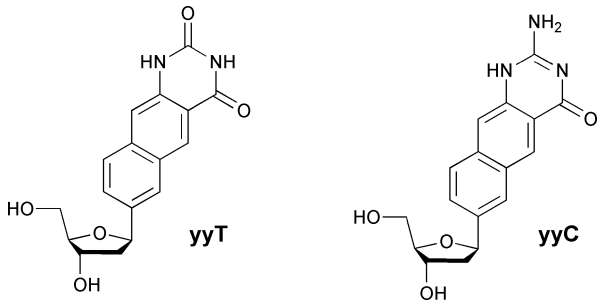
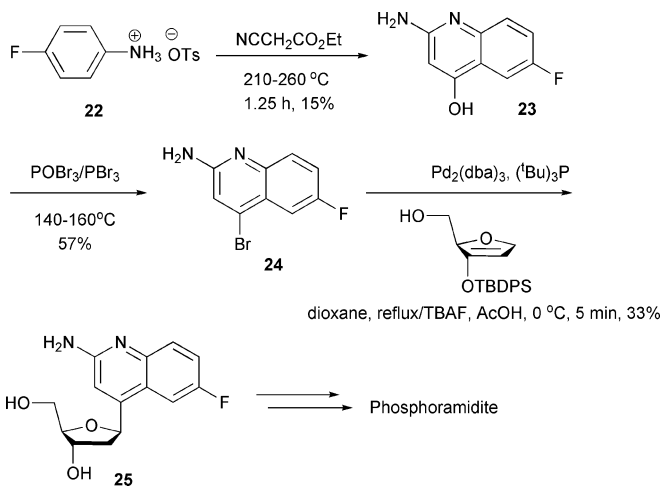


Fig. 3 Kool's yyDNA bases.



Scheme 3 Gold's synthesis of a *C*-nucleoside designed to participate in triplex formation with A:T base pairs in DNA.

shown in Fig. 7 is the one that is formed between their nucleobase and a CG pair.

(b) Fluorescent nucleobases

Fluorescent DNA bases play an important role in biochemistry and molecular biology, commonly as reporters in structural studies of DNA and RNA as well as in processes such as DNA repair and replication.[16–23] As might be expected, this is a large topic, which has been covered in an excellent recent review by Wilson and Kool.[16]

Selected highlights of recent work not included in this review, however, include the work of Greco and Tor, who have designed a simple fluorescent pyrimidine analogue **33** (Fig. 8) to detect the presence of DNA abasic sites.[17]

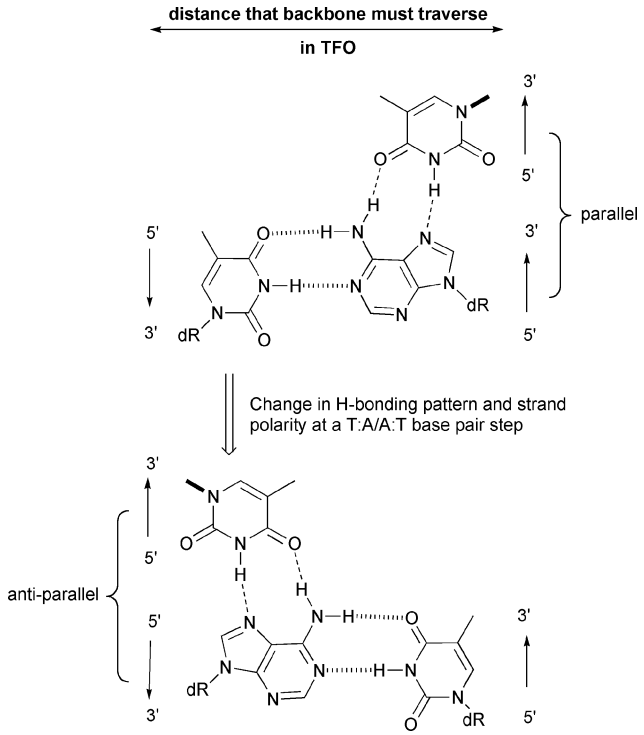


Fig. 4 Required change in strand polarity and movement of a pyrimidine motif TFO backbone between complementary strands in duplex DNA at a 5′-T:A/A:T base pair step (darkened bonds represent the attachment to the sugar–phosphate backbone).[14]

Abasic sites are generated spontaneously or through the enzymic cleavage of damaged nucleosides (in some cases such as that of antibody diversification, these damaged nucleosides are deliberately formed[18]). The ability to detect where these abasic sites occur can lead to a greater understanding of various biochemical mechanisms.

Ju and co-workers have reported a 3′-OH-capped fluorescent nucleotide that can act as a reversible terminator.[19] Such a nucleotide would help in the efficient sequencing by synthesis (given the acronym SBS) of DNA, as the terminal fluorescent nucleotide would enable the length of the nucleotide to be easily measured. The uncapping and cleavage of the fluorescent groups is achieved in a palladium-mediated deallylation reaction (Scheme 5).

Brown and co-workers have recently made an interesting contribution to nucleic acid fluorescence chemistry through the development of non-fluorescent quenchers.[20] When these non-fluorescent nucleotides are contained in a duplex and are complementing a fluorescent nucleotide, they quench that fluorescence until such a time that the duplex is denatured by heating (Fig. 9).

(c) Novel synthetic approaches

Inouye and co-workers have performed the stereoselective synthesis of alkynyl *C*-2-deoxy-β-D-ribofuranosides *via* an intramolecular Nicholas reaction. In this sequence, a 3,5-protected ribofuranose **37** was treated with an alkynyllithium reagent to afford a diastereomeric mixture of the alkynyldiol **38**. Treatment of this

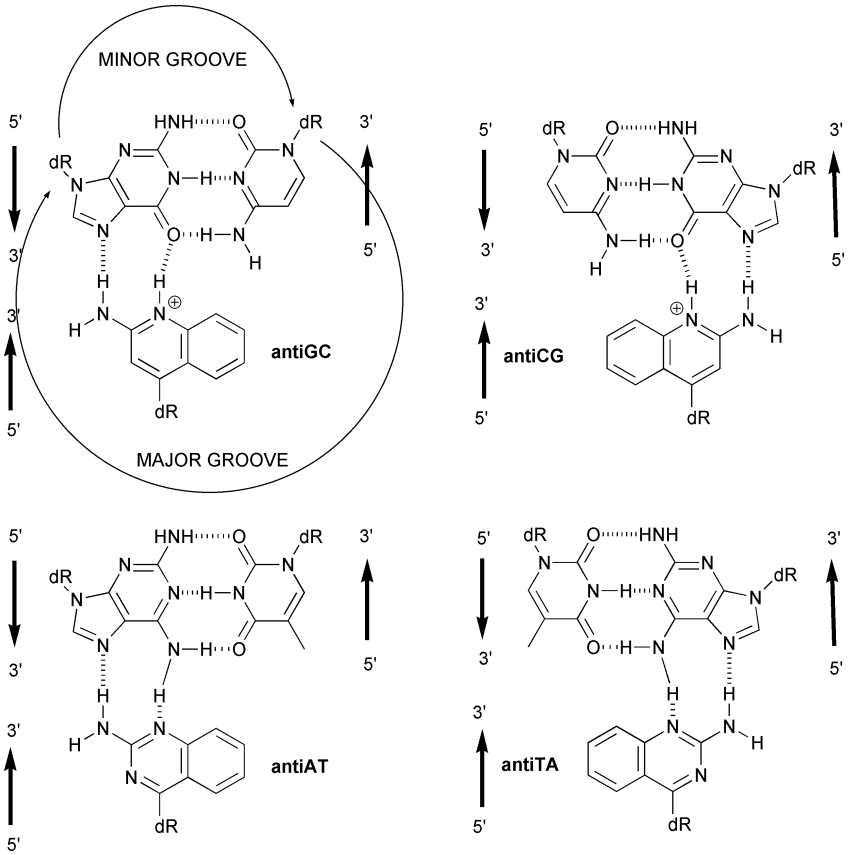


Fig. 5 Gold's nucleobase is capable of spanning across the dsDNA ,as it can rotate through the glycosidic bond to reach both bases.

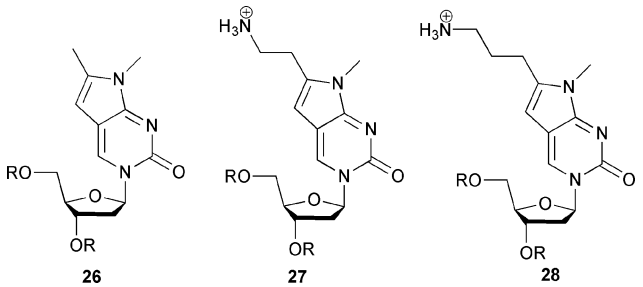


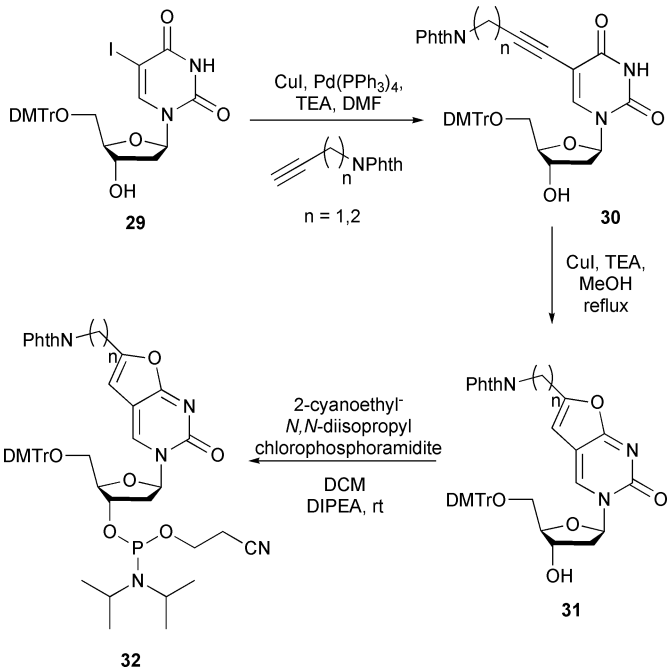
Fig. 6 Brown's CG-recognising nucleosides.

with $Co_2(CO)_8$, catalytic triflic acid, triethylamine and iodine in one pot gave the alkynyl ribofuranoside **40** with high β-selectivity *via* the transition state shown (Scheme 6).[24]

More recently, Hocek and co-workers have utilised the alkynyl moiety in a transition metal catalysed [2 + 2 + 2]-cyclotrimerisation (Scheme 7).[25]

Kobayashi and co-workers have also devised a new method to access *C*-glycosides through the use of a surfactant-type Brønsted acid, where the β-anomer predominated by a factor of >20 : 1 using water as a solvent (Scheme 8).[26]

Lakshman and co-workers have devised a method for the synthesis of a range of inosine-derived ribose and deoxyribose nucleosides through substitution of the C6 position, critically in the presence of BOP **46** (Scheme 9). This generates the benzotriazole intermediate **49**, which has been exposed to a range of nitrogen, oxygen and sulfur nucleophiles.[27]



Scheme 4 Brown's synthesis of CG-recognising nucleoside precursor phosphoramidite **32**. Reaction of the nucleobase with methylamine was shown to replace the O7 unit with an NMe unit.

Robins and co-workers have also investigated the synthesis of 6-substituted inosines by the Suzuki coupling of purine

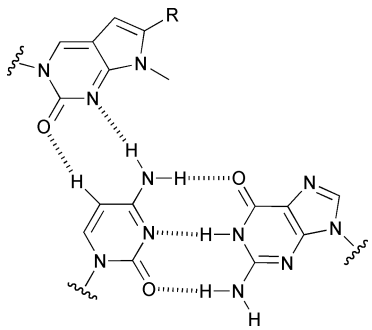


Fig. 7 Proposed model for CG recognition by Brown's nucleosides.

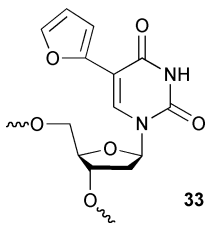
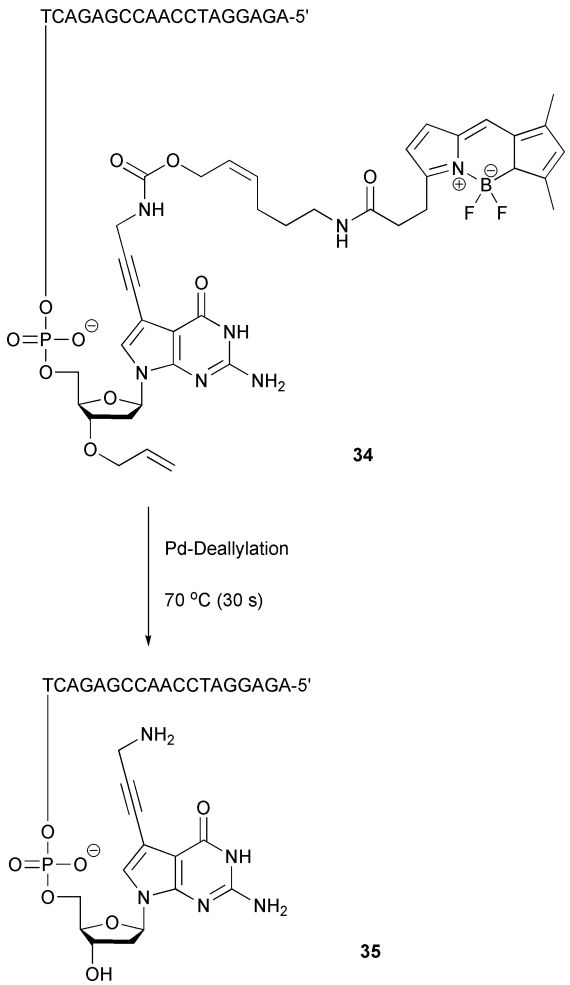
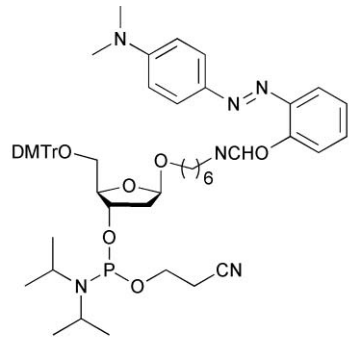


Fig. 8 Tor's abasic site-recognising nucleoside. When part of a duplex and pairing with an abasic site, it displays significantly increased emission.



Scheme 5 Ju's 3′-OH-capped fluorescent nucleoside and its subsequent uncapping.



36 : Methyl red phosphoramidite monomer

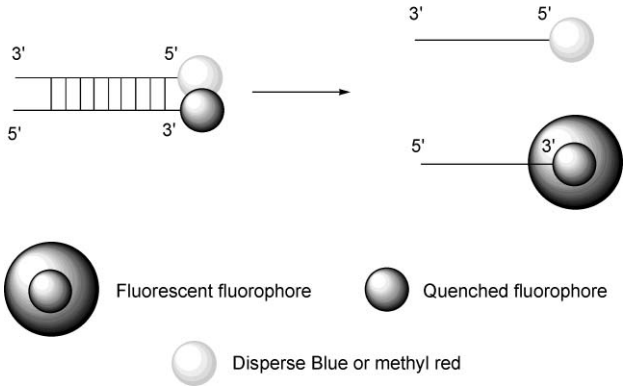
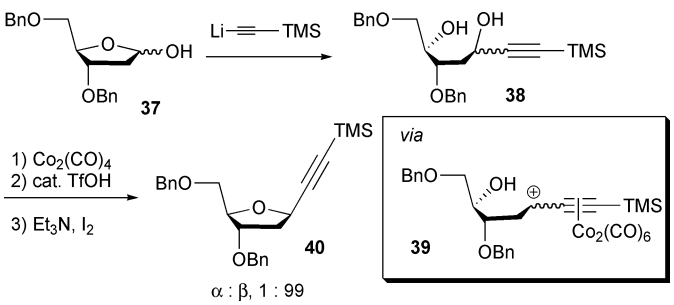
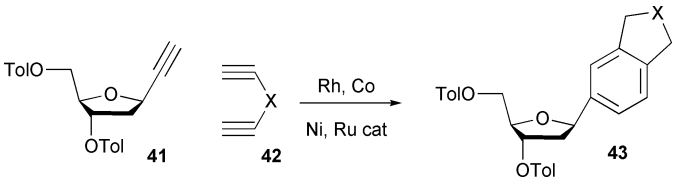


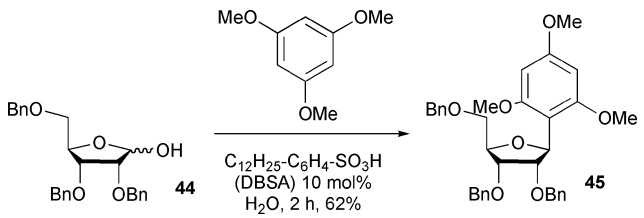
Fig. 9 Brown's non-fluorescent quencher.



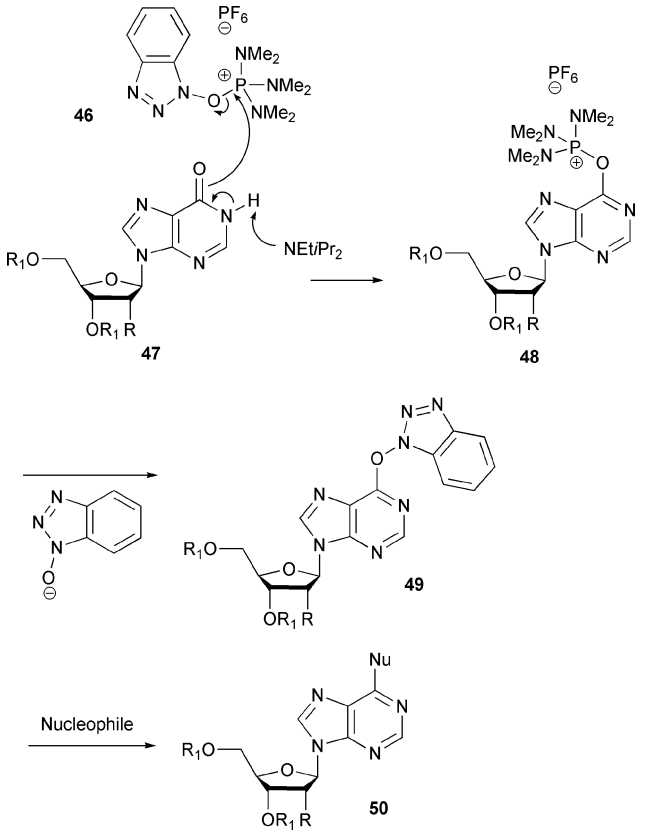
Scheme 6 Inouye's synthesis of alkynyl *C*-2-deoxy-β-D-ribofuranosides by an intramolecular Nicholas reaction.



Scheme 7 Hocek's [2 + 2 + 2]-cyclotrimerisation.

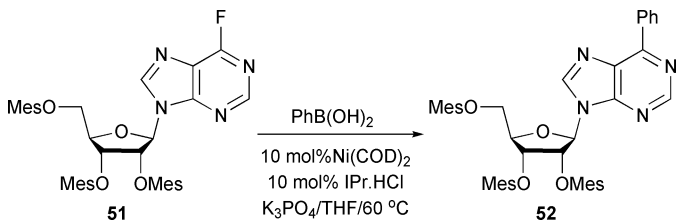


Scheme 8 Kobayashi's *C*-glycosidation.



Scheme 9 Lakshman's access to C6-substituted inosine nucleosides.

2′-deoxynucleosides and nucleosides with a range of boronic acids using a nickel- and palladium-based system with imidazolium–carbene ligands (Scheme 10). It is envisioned that the products of this reaction could be used as cytostatic agents, although they still have the potential to be incorporated into oligonucleotides.[28]



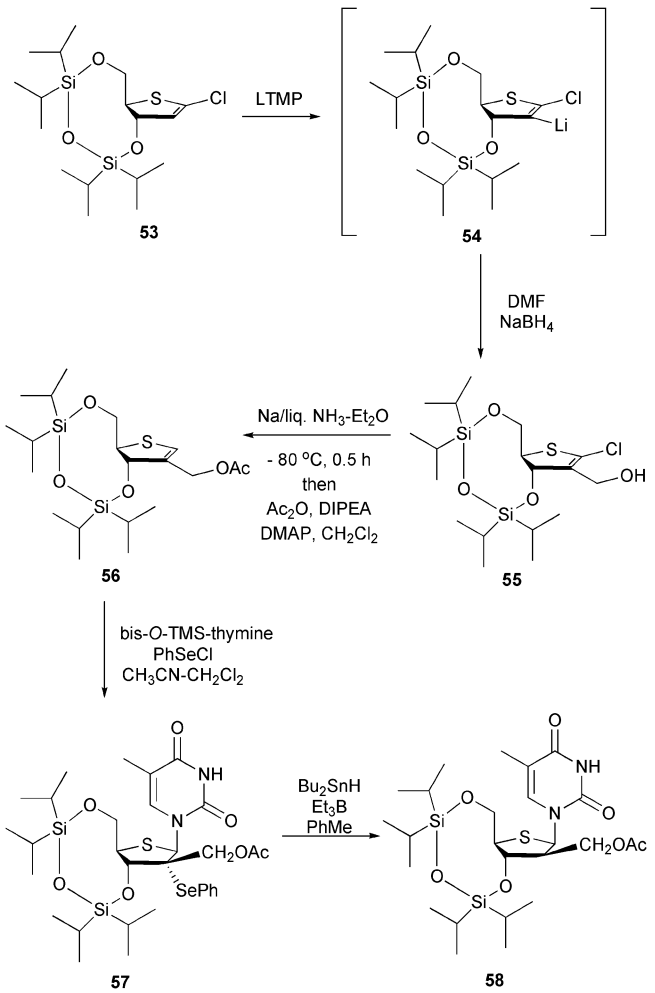
Scheme 10 Robins' access to C6-substituted inosine nucleosides.

2. Modifications to the sugar portion

There are several reasons as to why the (deoxy)ribose part of a nucleic acid might be manipulated. One which has become apparent in recent years is the desire to synthesise oligomers that are stable to enzymic degradation,[10d,29–31] and can therefore be used effectively in RNA interference and antisense technologies. However, other reasons can range from the desire to add functionality to plasmid DNA[32] and triplex formation[33] to the need for improving nucleic acid–protein binding.[34]

(a) Substitution of the ribose oxygen

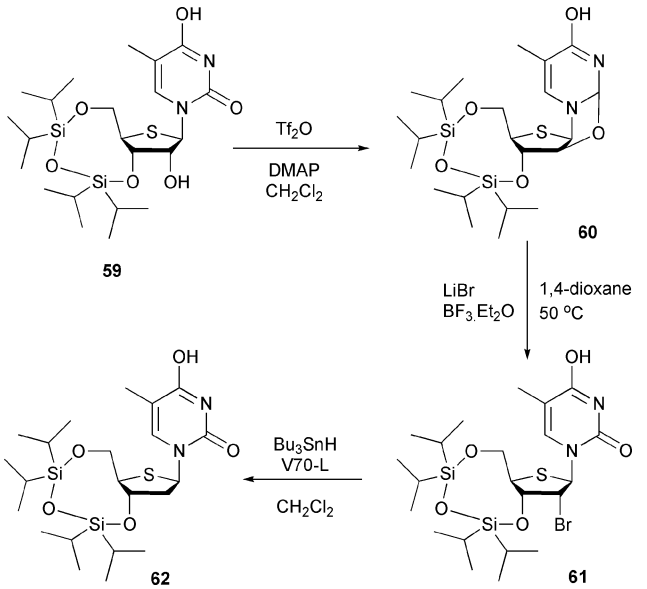
Substitution of the oxygen in the ribose ring for atoms such as sulfur and nitrogen (as in the case of many PNAs which are not the subject of this review[35]) can also lead to nuclease-resistant properties. Although not intended for incorporation into an oligonucleotide, Haraguchi and co-workers have devised the ingenious synthesis of a range of 4′-thionucleosides in the investigation of their antiviral and antitumour activities.[36] These thionucleosides were accessed by the PhSeCl-mediated electrophilic glycosidation of 4-thiofuranoid glycals with a carbon substituent at the C2-position. These glycals were synthesised by the C2-lithiation of 1-chloro-4-thiofuranoid glycal **53** with LTMP. Birch reduction of the chlorine atom followed by electrophilic glycosidation using the nucleobase and PhSeCl generated intermediate **57** exclusively as the β-anomer. Removal of the selenium provided the desired 4′-thionucleoside **58** (Scheme 11).



Scheme 11 Haraguchi's synthesis of a 4′-thionucleoside.

Matsuda and Minakawa have also developed a practical synthesis of 4′-thioribonucleosides, specifically 2′-deoxy-4′-thioribonucleosides.[37] Problems associated with the Barton radical deoxygenation of the 2′-OH, where the sulfur participated to generate a major side product, were overcome by replacing the 2′-OH with a bromine atom, and subsequent radical reduction at

this position (Scheme 12). Matsuda and Minakawa found that oligonucleotides containing these 4′-thioDNAs showed higher resistance to 3′-exonuclease cleavage. Interestingly, they also showed that the 4′-thioDNA:4′-thioDNA homoduplex exhibits RNA-like characteristics in that it adopts the A-form in aqueous buffer at neutral pH and moderated salt conditions.[37b]



Scheme 12 Access to 2′-deoxy-4′-thionucleosides by radical reduction of the 2′-bromonucleoside.

In their investigation of the degradation of DNA with reactive oxygen species, Carell and co-workers have synthesised the carbocyclic nucleoside analogue of guanine **64**.[38] One of the most common DNA lesions formed is the 2′-deoxyguanisine ring-opened lesion dubbed FaPydG **63**, but the problem with studying these types of lesion is that the β-FaPydG lesion readily anomerises to give the α-FaPydG lesion. This then makes any study into base-pairing effects of these lesions difficult to assess. The carbocyclic analogue dcG **64** made by Carell is, of course, unable to anomerise (Fig. 10). The effect of replacing the oxygen with carbon on base-pairing was investigated, and it was found that dcG had similar base pairing capabilities to dG, thus allowing confidence in any subsequent lesion study.

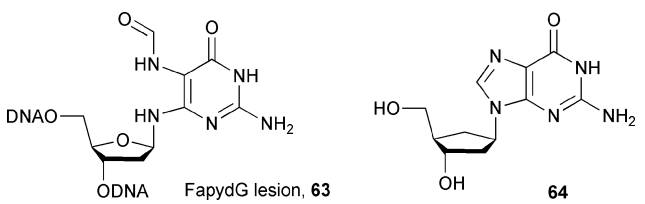
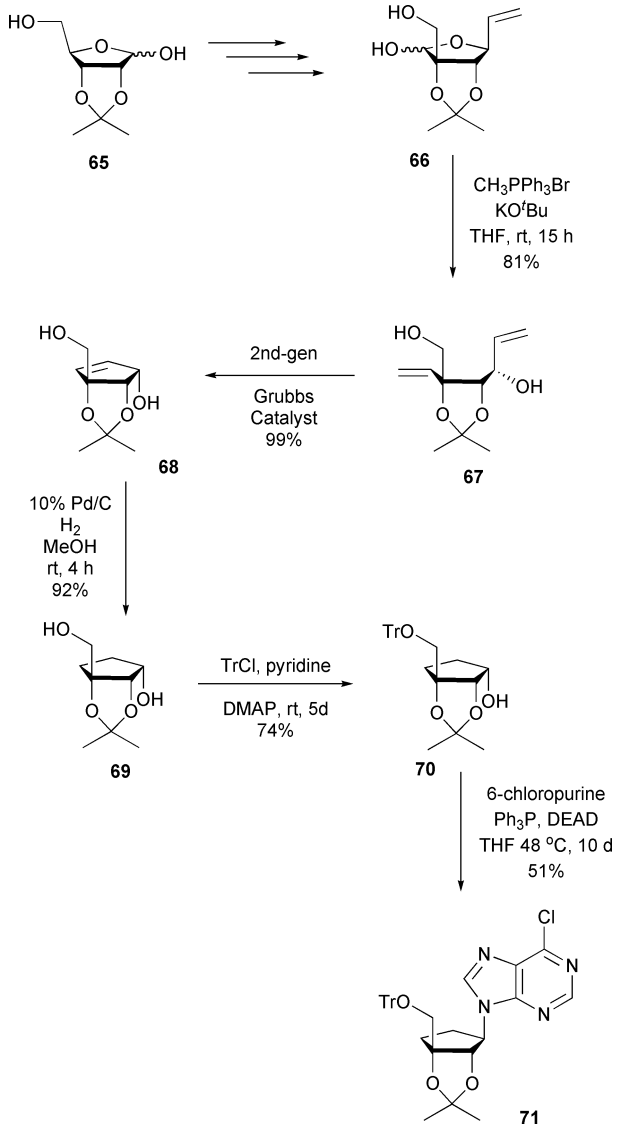


Fig. 10 Carell's carbocyclic guanine.

Jeong, Moon and co-workers have also synthesised novel carbocyclic nucleoside analogues as selective A_3 adenosine receptor agonists.[34] Although they do not incorporate this into an oligonucleotide, their method still represents a useful way of accessing modified *C*-nucleosides. Starting from 2,3-isopropylidene-D-ribose **65**, they obtained C2-hydroxymethyl lactol **66** in 3 steps. They then incorporated a methylene unit by a Wittig reaction and utilised ring-closing metathesis to generate the core carbocycle. Reduction of the resulting double bond followed by Mitsunobu incorporation of the nucleobase gave them the desired product **71** (Scheme 13).



Scheme 13 Jeong and Moon's synthesis of a carbocyclic nucleoside analogue.

(b) Addition of functionality

Nielsen and co-workers have synthesised a range of riboses which carry *two* nucleobases (Fig. 11).[39,40] Such modified nucleic acids could stabilise secondary structures or three-way junctions and could also potentially take part in a double-coding system. In particular, they showed that the incorporation of (5′*S*)-*C*-(thymine-1-ylmethyl)thymidine **73** into duplexes revealed the (5′*S*)-C-position as ideal for the orientation of the extra nucleobase into the minor groove. The authors speculate that this might lead to the development of self-assembling nucleic acid nanostructures.[40]

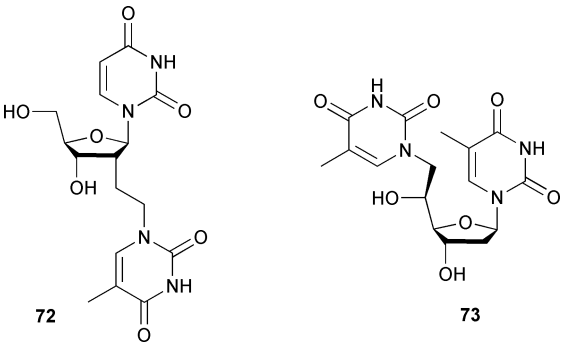
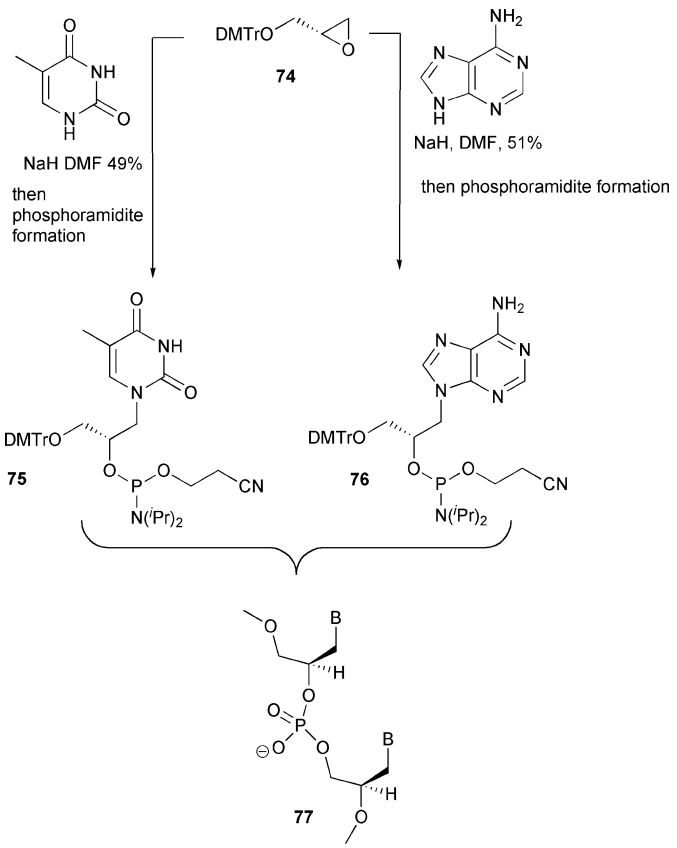


Fig. 11 Nielsen's bifunctionalised nucleosides.

(c) Replacement of the ribose

Several reports have detailed modifications to the sugar structure, and one of the motivations for this is to explore the limits of the conditions necessary to invoke Watson–Crick pairing. Meggers and co-workers have developed a simplified nucleic acid based on a glycol framework.[41] They utilised (*R*)- or (*S*)-glycidol **74** as a template to introduce thymine or adenine nucleobases (Scheme 14). They found that this glycol nucleic acid forms highly stable antiparallel duplex structures using Watson–Crick pairing, and this prompts them to suggest that this might have been a possible predecessor of RNA as the genetic material for early life on Earth.



Scheme 14 Megger's simplified oligonucleotide.

Eschenmoser and co-workers have spent many years investigating the effect of a hexose as opposed to a pentose sugar within the nucleic acid framework.[42] They have recently investigated the crystal structure of a homo-DNA octameric duplex, and have shown that it has a weakly twisted right-handed duplex and that the geometries of individual base-pair steps varies considerably.[43] These insights have also allowed them to conclude that for fully hydroxylated hexopyranose sugars **78–80** (Fig. 12), base pairing is prevented because the bulkiness of the sugar prevents the bases from adopting the orientation required for stacking.

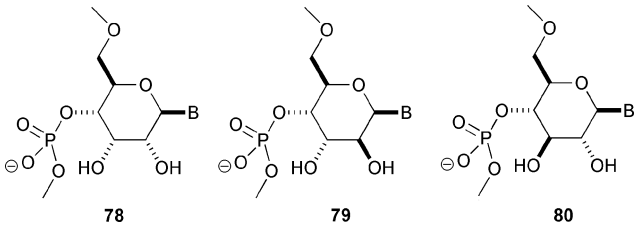


Fig. 12 Eschenmoser's hexose nucleoside equivalents.

(d) Conformationally restricted nucleosides

Naturally occurring DNA and RNA exist in an equilibrium between two major conformers known as the North (*N*) **81** and South (*S*) **82** types (Fig. 13).

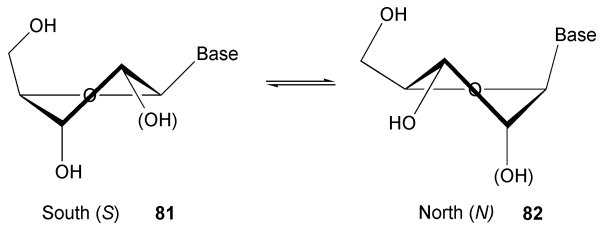


Fig. 13 Equilibrium between North (*N*) and South (*S*) conformations.

The conformer which is adopted dictates the nature of the double-helix which is formed, namely the A- and B-forms. The *N*-type conformer is preferred by ribonucleosides that give rise to the A-form, and the *S*-type conformer is preferred by the 2′-deoxyribonucleosides, which adopt the B-form. This conformational equilibrium is thought to be important for a variety of biological functions.[44] Indeed, modifications to the nucleic acid which maintain and reinforce the A-type geometry of RNA perform well in RNAi.

The restriction of this equilibrium using bi- or tricyclic carbohydrate moieties has therefore been a very useful tool in nucleic acid chemical biology,[45] as well as in the fields of diagnostics and therapeutics.[46]

Locked nucleic acids (LNA) were introduced in 1998 by Wengel and co-workers in order to generate monomers which will be favourably locked in the *N*-type conformation.[47] Their excellent ability to recognize complimentary DNA and RNA strands has made them the subject of much study, particularly as already mentioned, owing to their stability to enzymic degradation. Various analogues of these original LNAs have been synthesised (Fig. 14).[48,49]

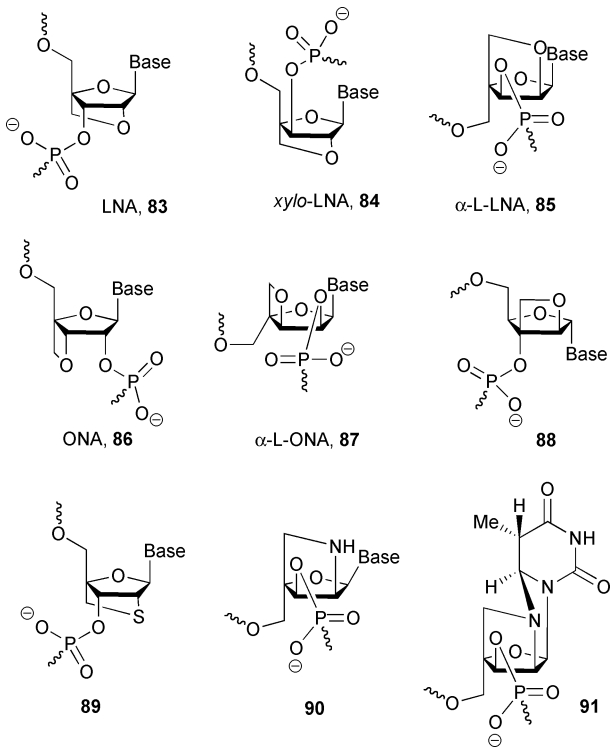


Fig. 14 Conformationally restricted nucleosides.

Most recently, Hrdlicka and co-workers have developed a range of 2′-amino-α-L-LNAs, which, although only imparting modest changes in thermal stability with complimentary DNA, give significant increases in the thermal stability observed with RNA compliments and excellent Watson–Crick discrimination,[49] such qualities being useful as nucleic acid probes in medicinal and biotechnological applications. Two interesting products from this work were the tetracyclic 'locked' LNAs **96** and **97**, which were synthesised from the 2′-amino-α-L-LNA **92** using KOAc, 18-crown-6 and dioxane, and formed through an intramolecular aza-Michael reaction (Scheme 15). Phosphoramidites of both tetracyclic nucleosides were made and incorporated into an oligonucleotide, although **97** underwent retro-aza-Michael addition to form **91**. Thermal denaturation studies showed that this monomer showed increased thermal stability with both DNA and RNA and superior Watson–Crick selectivity. The tetracyclic nucleoside **89**, however, destabilised the duplex in all test sequences and also appeared to destabilise neighbouring base pairs, which the authors attribute to loss of aromaticity and subsequent increased steric bulk.

Another example of a conformationally restricted nucleoside comes from the laboratories of Chattopadhyaya and co-workers, who synthesised a series of 1′,2′-azetidine-fused bicyclic pyrimidine nucleosides of T, U, C and 5-Me-C.[50] These modified nucleosides were shown to be locked into a 'North-east' conformation. The key step in the synthesis of these new nucleotides was the use of the nucleobase to invoke the correct stereochemistry in the 2′-position for efficient formation of the azetidine (Scheme 16). When incorporated into a 15-mer oligonucleotide by standard phosphoramidite chemistry, these sequences were shown to bind to the complimentary RNA strand with improved binding efficiencies over the analogous oxetane set of oligonucleotides and, of course, this improved binding to RNA is particularly interesting owing to its potential application in antisense technologies.

Recently, Wengel and co-workers have extended their LNA systems to the formation of triplexes (as discussed in Section 1).[33b] They tested a range of LNA-modified compounds against a 33-mer DNA sequence using a 16-mer sequence that had between 1 and 15 nucleobases replaced with the LNA variant (Fig. 15).

They found that the sequence which formed the most stable sequences incorporated the LNA monomer **107** (Fig. 16). This combines two favourable qualities for triplex formation: (i) that of 'dual recognition', whereby the nucleobase or sugar part of the monomer is linked to a unit that contains one or more amines that can be (at least partially) protonated under physiological conditions, and (ii) that the ribose is locked in the *N*-conformation.

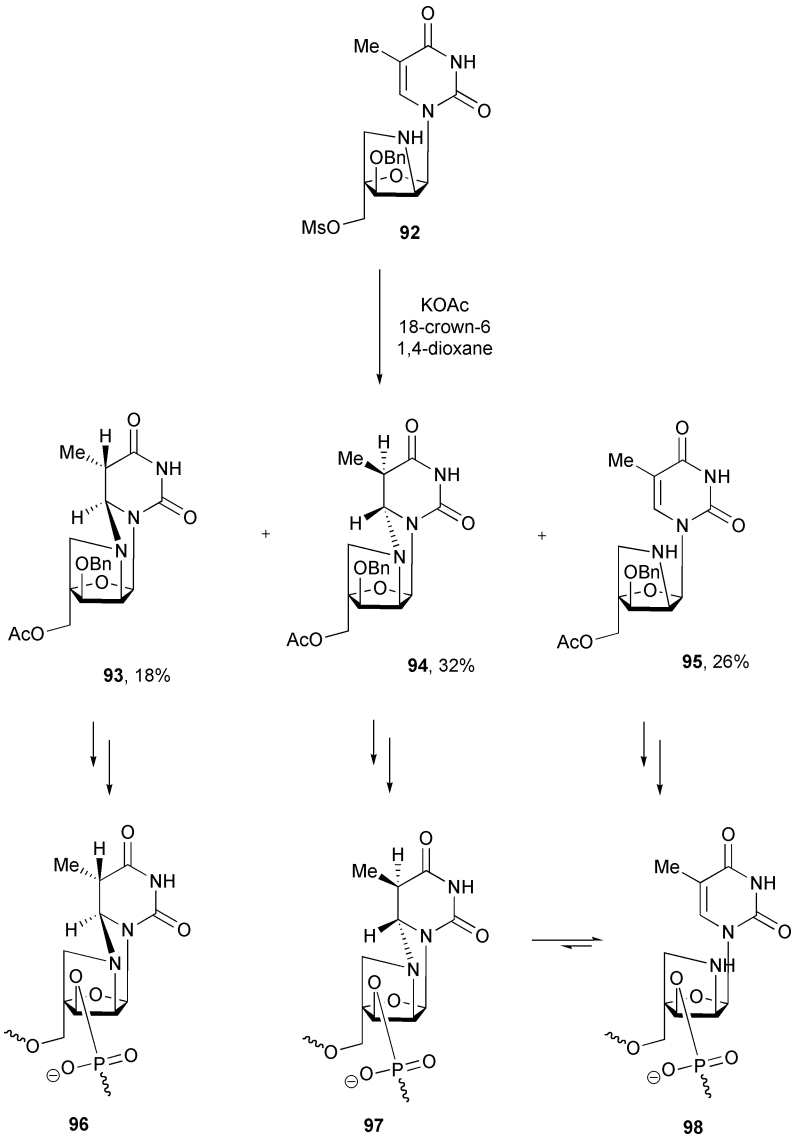
Paquette has combined the use of a different heteroatom within the ribose with conformational rigidity to generate a class of ribonucleosides where conformational rotation around the 5′-hydroxymethyl substituent is inhibited.[51] These types of compound represent a new class of nucleoside mimic.[52] The key step in this synthesis is the generation of the sulfonium intermediate **112**. This is accessed through the oxidation of the sulfur to the sulfone using the Davis oxaziridine **109** followed by triethylamine-induced elimination. Neighbouring group participation of the dimethoxybenzoyl functionality at the 2′-position then ensures that the nucleobase is delivered to the β-face (although some elimination of the dimethoxybenzoyl group also occurs; see Scheme 17).

3. Modifications to the nucleic acid linker

Once again, one of the principal aims behind the design of modified internucleoside linkages is to allow antisense and RNAi strategies to become more effective through improved enzymic stability, cellular uptake, biodistribution and minimisation of toxicity. A common modification is the use of phosphorothioate linkages where one of the non-bridging oxygens is formally replaced by a sulfur atom, which not only invoke this stability but also increases the binding to serum proteins and leads to reduced clearance rate and improved bioavailability. There have been many reports on the synthesis of these phosphorothioate linkages by the use of sulfur transfer reagents such as 3*H*-1,2-benzodithiol-3-one-1,1-dioxide (Beaucage reagent),[53] dimethylthiarum disulfide (DTD)[54] and dibenzoyl tetrasulfide.[55] One of the newer reagents, phenylacetyl

```
3'  G G T  G A A  A A A  T T T  T C T  T T T  C C C  C C C  T
G A  C C
5'  C C A  C T T  T T T  A A A  A G A  A A A  G G G  G G G  A
C T  G G
                          x x x  x x x  x x x  x x x  x x x  x
                    5'  T T T  T C T  T T T  C C C  C C C  T
```

Fig. 15 The sequence of the target duplex and the unmodified TFO strand which Wengel has furnished with modified LNAs.



Scheme 15 Hrdlicka's 2′-amino-α-L-LNAs.

disulfide (PADS) **116** has been developed by Ravikumar and co-workers.[56] When an oligonucleotide was exposed to a solution of this reagent, the phosphorothioate was generated, and the authors suggest this proceeds by the mechanism shown in Scheme 18. Remarkably, they note that addition of water to the PADS solution does not lead to increased levels of the hydrolysed product **117**, and speculate that an alternative mechanism might be possible in this case.

Stec and co-workers have designed the related phosphoroselenoate internucleotide linker **123** (Scheme 19)[57] which could potentially be used to probe structure and function by X-ray crystallography because of multiwavelength anomalous dispersion (MAD).[58] These selenolinkers have a predetermined sense of *P*-chirality.[59]

Recent studies by Rozners and co-workers have focused on the synthesis and utility of a range of oligoribonucleotides with amide linkages (Fig. 17).[60] Their most recent report focused on

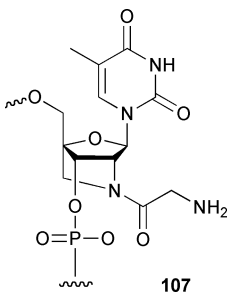
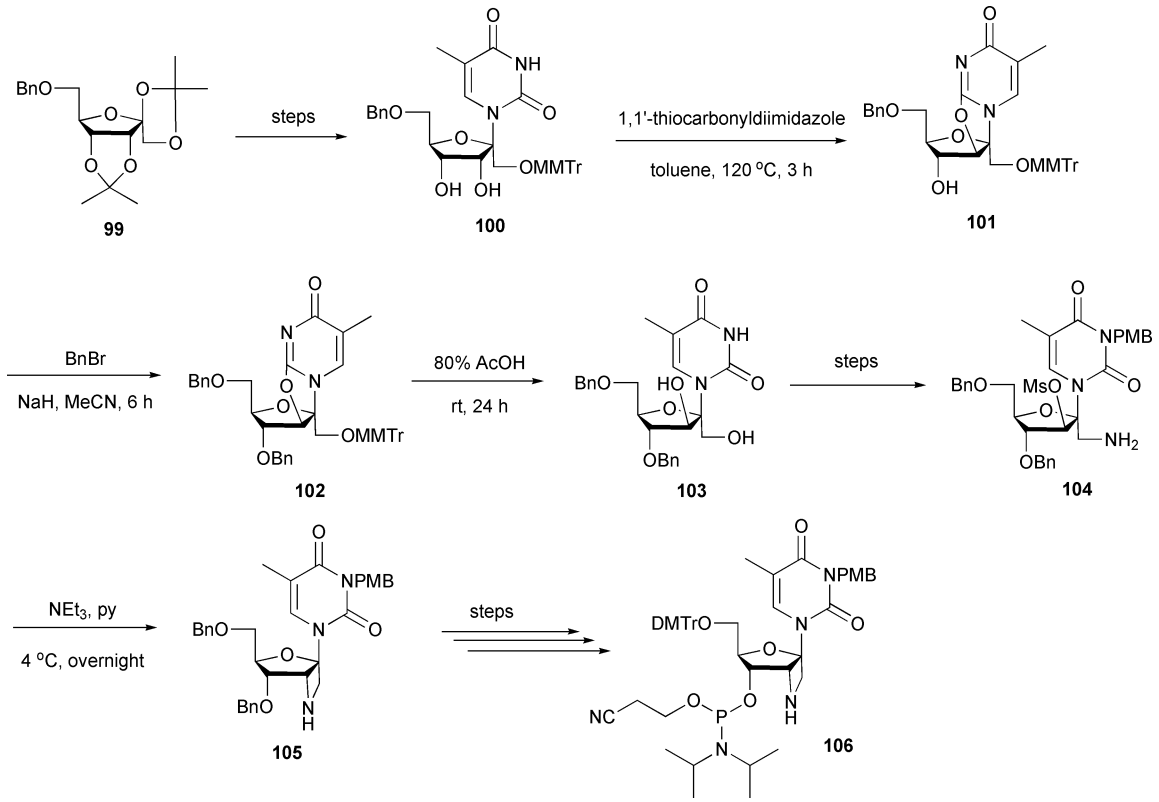


Fig. 16 Example of an LNA TFO designed by Wengel and co-workers.

monomer **129**. Two syntheses were devised (each with their own advantages and disadvantages), and the route shown (Scheme 20) has the advantage that a variety of heterocyclic bases can be introduced towards the end of the synthesis, allowing for late-stage diversification.[60a]



Scheme 16 Chattopadhyaya's 1′,2′-azetidine mimic of thymine.

Caruthers and co-workers have developed oligodeoxynucleotides with phosphonoformate linkers, again with the main purpose of forming oligonucleotides capable of participating in antisense binding without being degraded by cellular nucleases.[61] By using phosphinylformic acid ester **133** (synthesised from **130**) to form phosphoramidite **135**, they were able to introduce this modified linker. The one issue that did rise in testing the ability of these modified phosphoramidites to DNA synthesis was the lability of the C–P bond to nucleophilic attack, which meant that standard amine protecting groups on the exocyclic amines of cytosine, adenine and guanine could not be used. However, the Fmoc protecting group was found to be compatible with the new linker, as it could be removed under the same conditions as the DPSE group (Scheme 21).

Nielsen and co-workers have utilised ring-closing metathesis to create a range of internucleotide linkages in addition to the phosphate linker.[62] Most recently, they have prepared a carbon linker between the 5′-positions of adjacent thymine residues using ring-closing metathesis (Scheme 22).[62a] Such modifications can lead to the stabilisation of secondary nucleic acid structures such as bulges or three-way junctions. This, the authors suggest, could be used to modulate RNA function for future therapeutic uses and nanobiotechnology.

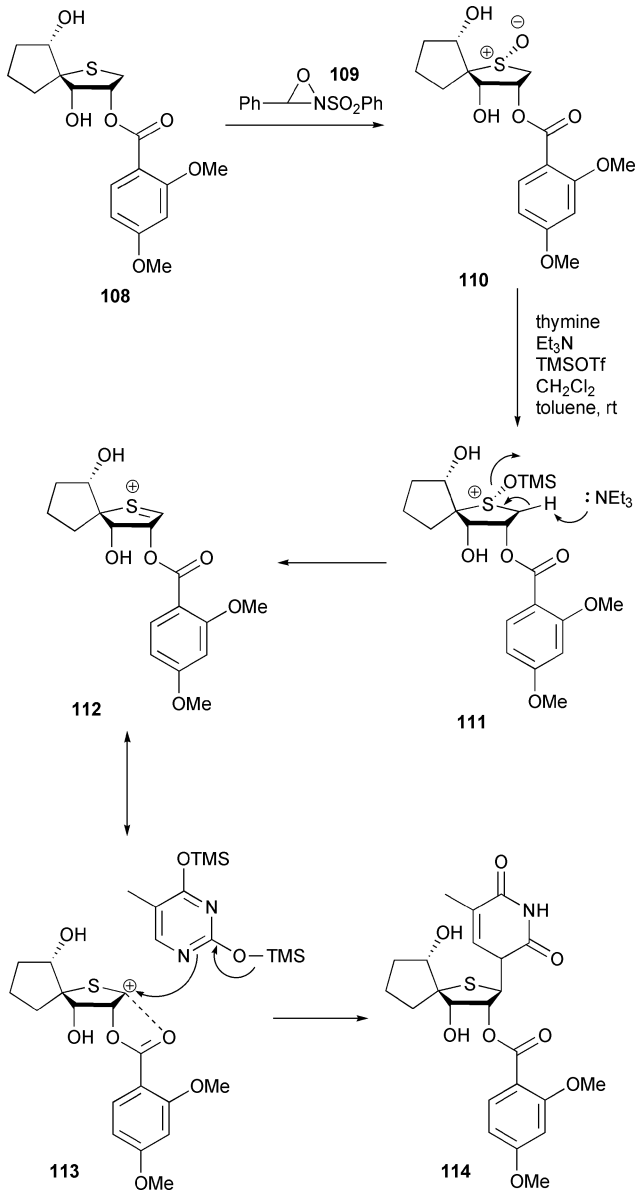
Conclusions

This review has attempted to highlight recent developments in modified nucleotides, and specifically their incorporation into novel oligonucleotides. There has been a notable resurgence in the development of nucleic acid mimics, and this is arguably a direct result of developments in RNAi (and to a certain extent antisense) technologies. The design of these mimics aims to address the problem of RNA degradation, which is a major problem in this methodology. Locked nucleic acids and oligonucleotides with modified linkers have been directed towards this purpose. Overcoming such obstacles will lead to the real possibility that RNAi technology will be able to act against cancer through the targeting of oncogenes, as well as against viral infection through the targeting of viral genes. Of course there are a great many other obstacles, such as delivery, which must be overcome before this becomes a reality, but the discovery of an oligonucleotide mimic which is resilient to endonuclease degradation would be a giant leap forward.

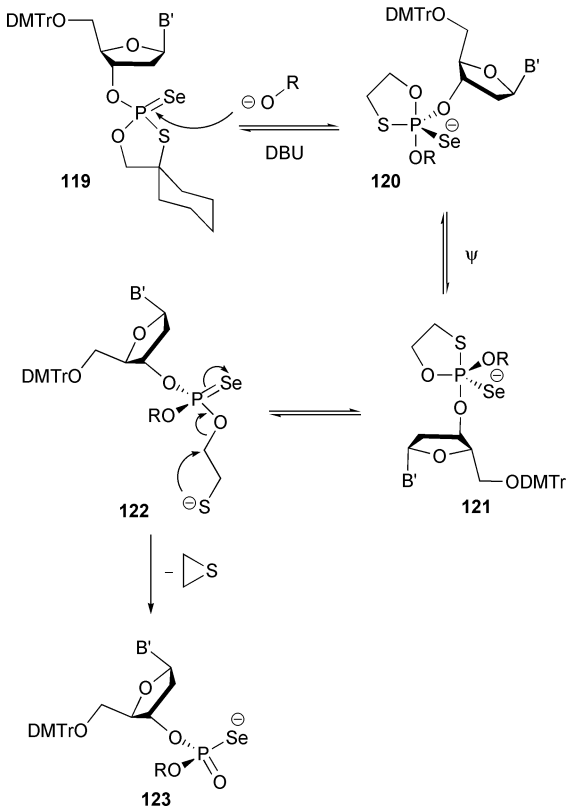
In addition to this, there have been exciting developments in oligonucleotides that show improved binding to double helices to form triplexes. This will allow for improved transcriptional activation or deactivation, and this too could lead to important therapeutic developments.

The use of fluorescent nucleosides as probes in molecular biology applications has become an important field of study and an indispensable tool in nucleic acid biophysics.

Modified nucleic acids have also been developed to answer the more fundamental questions about Nature and why it has chosen the genetic coding system that it has. Such insights will no doubt shed light onto the origins of life and how it evolved into the complex genetic system it is today.



Scheme 17 Paquette's 4′-thioribonucleoside.



Scheme 19 Stec's predetermined *P*-chirality

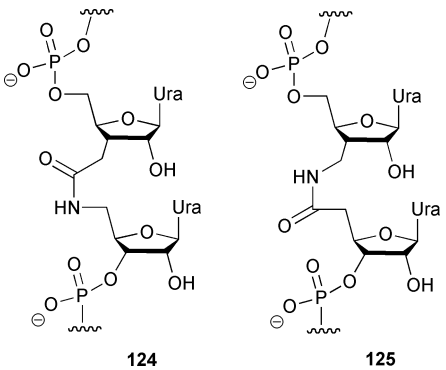
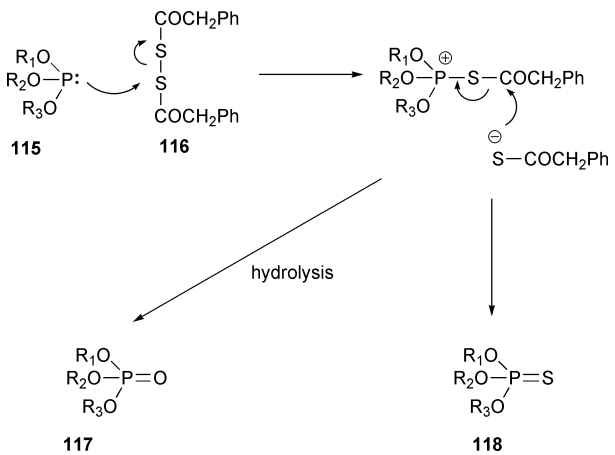
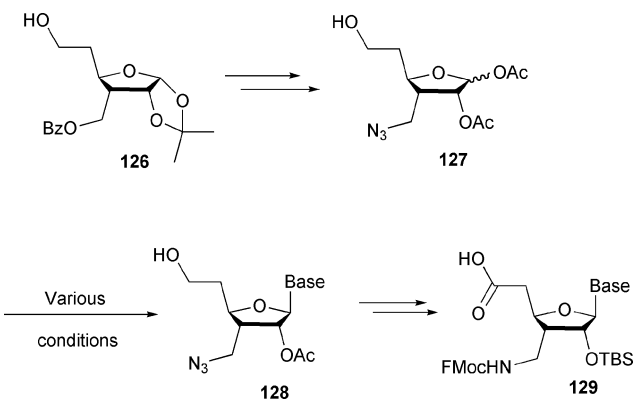


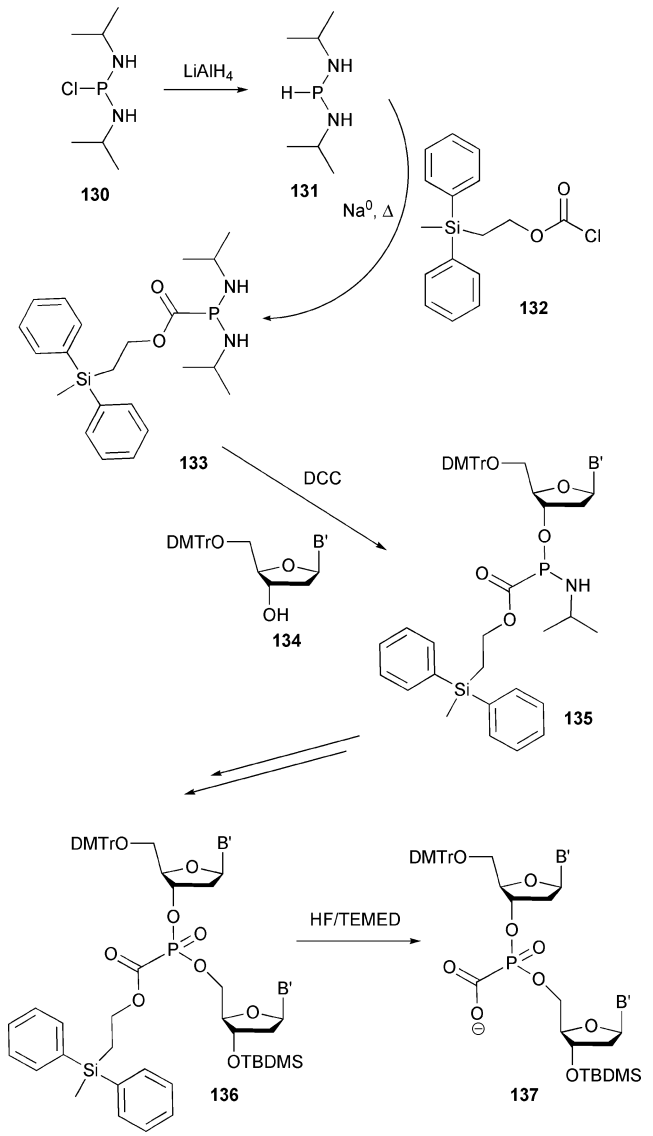
Fig. 17 Rozner's amide linkages.



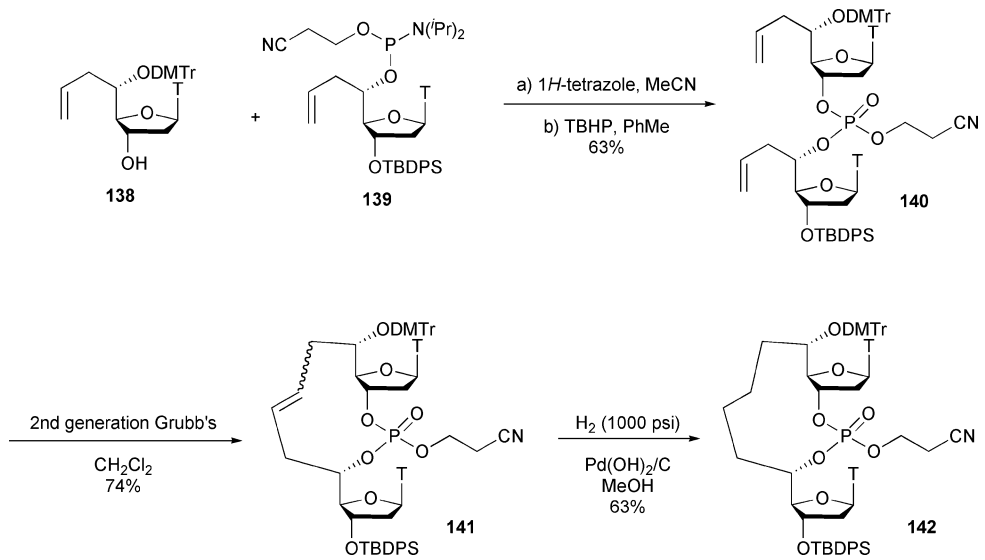
Scheme 18 Ravikumar's sulfur transfer reagent.



Scheme 20 Rozner's monomer synthesis.



Scheme 21 Caruthers' phosphonoformate linkage.



Scheme 22 Nielsen's ring-closing metathesis.

Acknowledgements

The author wishes to thank Cancer Research UK and the Reading Endowment Trust Fund (RETF) for funding work associated with this review, and Professor Helen Osborn for useful discussions.

References

1 A. Fire, S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver and C. C. Mello, *Nature*, 1998, **391**, 806.

2 (*a*) R. Micura, *Angew. Chem., Int. Ed.*, 2002, **41**, 2265; (*b*) D. H. Kim and J. J. Rossi, *Nat. Rev. Genet.*, 2007, **3**, 173; (*c*) A. Gaither and V. Iourgenko, *Curr. Opin. Oncol.*, 2007, **1**, 50. For a good overview, see: (*d*) P. G. Rychahou, L. N. Jackson, B. J. Farrow and B. M. Evers, *Surgery*, 2006, **5**, 719.

3 (*a*) N. Dias and C. A. Stein, *Mol. Cancer Ther.*, 2002, **1**, 347; (*b*) A. Dove, *Nat. Biotechnol.*, 2002, **20**, 121; (*c*) S. T. Crooke, *Curr. Mol. Med.*, 2004, **4**, 465.

4 (*a*) A. T. Krueger, H. Lu, A. H. F. Lee and E. T. Kool, *Acc. Chem. Res.*, 2007, **40**, 141; (*b*) E. T. Kool, *Acc. Chem. Res.*, 2002, **35**, 936.

5 (*a*) S. R. Lynch, H. B. Liu, J. M. Gao and E. T. Kool, *J. Am. Chem. Soc.*, 2006, **128**, 14704; (*b*) J. Gao, H. Liu and E. T. Kool, *Angew. Chem., Int. Ed.*, 2005, **44**, 3118; (*c*) H. B. Liu, J. M. Gao and E. T. Kool, *J. Am. Chem. Soc.*, 2005, **127**, 1396; (*d*) H. Liu, J. Gao and E. T. Kool, *J. Org. Chem.*, 2005, **70**, 639; (*e*) H. Liu, S. R. Lynch and E. T. Kool, *J. Am. Chem. Soc.*, 2004, **126**, 6900; (*f*) H. Liu, J. Gao, L. Maynard, Y. D. Saito and E. T. Kool, *J. Am. Chem. Soc.*, 2004, **126**, 1102.

6 (*a*) A. H. F. Lee and E. T. Kool, *J. Am. Chem. Soc.*, 2005, **127**, 3332; (*b*) A. H. F. Lee and E. T. Kool, *J. Org. Chem.*, 2005, **70**, 132; (*c*) H. Lu, K. He and E. T. Kool, *Angew. Chem., Int. Ed.*, 2004, **43**, 5834.

7 A. H. F. Lee and E. T. Kool, *J. Am. Chem. Soc.*, 2006, **128**, 9219.

8 (*a*) K. W. Wellington and S. A. Benner, *Nucleosides, Nucleotides Nucleic Acids*, 2006, **25**, 1309; (*b*) R. N. Farr and G. D. Daves, *J. Carbohydr. Chem.*, 1990, **9**, 653.

9 Z. Sun, S. Ahmed and L. W. McLaughlin, *J. Org. Chem.*, 2006, **71**, 2922.

10 (*a*) R. V. Guntaka, B. R. Varma and K. T. Weber, *Int. J. Biochem. Cell Biol.*, 2003, **35**, 22; (*b*) M. M. Seidman and P. M. Glazer, *J. Clin. Invest.*, 2003, **112**, 487; (*c*) H. E. Moser and P. B. Dervan, *Science*, 1997, **238**, 645; (*d*) S. Buchini and C. J. Leumann, *Curr. Opin. Chem. Biol.*, 2003, **7**, 717; (*e*) S. Buchini and C. J. Leumann, *Angew. Chem., Int. Ed.*, 2004, **43**, 3925.

11 J.-S. Li and B. Gold, *J. Org. Chem.*, 2005, **70**, 8764.

12 J.-S. Li, F.-X. Chen, R. Shikiya, L. A. Marky and B. Gold, *J. Am. Chem. Soc.*, 2005, **127**, 12657.

13 J.-S. Li, R. Shikiya, L. A. Marky and B. Gold, *Biochemistry*, 2004, **43**, 1440.
14 J.-S. Li, Y.-H. Fan, Y. Zhang, L. A. Marky and B. Gold, *J. Am. Chem. Soc.*, 2003, **125**, 2084.
15 R. T. Ranasinghe, D. A. Rusling, V. E. C. Powers, K. R. Fox and T. Brown, *Chem. Commun.*, 2005, 2555.
16 (*a*) J. N. Wilson and E. T. Kool, *Org. Biomol. Chem.*, 2006, **4**, 4265; (*b*) A. P. Silverman and E. T. Kool, *Chem. Rev.*, 2006, **106**, 3775.
17 N. J. Greco and Y. Tor, *J. Am. Chem. Soc.*, 2005, **127**, 10784.
18 M. Samaranayake, J. M. Bujnicki, M. Carpenter and A. S. Bhagwat, *Chem. Rev.*, 2006, **106**, 700.
19 L. Bi, D. H. Kim and J. Ju, *J. Am. Chem. Soc.*, 2006, **128**, 2542.
20 J. P. May, L. J. Brown, I. van Delft, N. Thelwell, K. Harley and T. Brown, *Org. Biomol. Chem.*, 2005, **3**, 2534.
21 S.-i. Nakano, Y. Uotani, K. Uenishi, M. Fujii and N. Sugimoto, *J. Am. Chem. Soc.*, 2005, **127**, 518.
22 S. G. Srivatsan and Y. Tor, *J. Am. Chem. Soc.*, 2007, **129**, 2044.
23 K. Tainaka, K. Tanaka, S. Ikeda, K.-i. Nishiza, T. Unzai, Y. Fujiwara, I. Saito and A. Okamoto, *J. Am. Chem. Soc.*, 2007, **129**, 4776.
24 M. Takase, T. Morikawa, H. Abe and M. Inouye, *Org. Lett.*, 2003, **5**, 625.
25 P. Novák, R. Pohl, M. Kotora and M. Hocek, *Org. Lett.*, 2006, **8**, 2051.
26 S. Shirakawa and S. Kobayashi, *Org. Lett.*, 2007, **9**, 311.
27 S. Bae and M. K. Lakshman, *J. Am. Chem. Soc.*, 2007, **129**, 782.
28 J. Liu and M. J. Robins, *Org. Lett.*, 2005, **7**, 1149.
29 J. Kurreck, *Eur. J. Biochem.*, 2003, **270**, 1628.
30 D. A. Braasch and D. R. Corey, *Biochemistry*, 2002, **41**, 4503.
31 M. Manoharan, *Curr. Opin. Chem. Biol.*, 2004, **8**, 570.
32 K. M. L. Hertoghs, J. H. Ellis and I. R. Catchpole, *Nucleic Acids Res.*, 2003, **31**, 5817.
33 (*a*) N. Kumar, K. E. Nielsen, S. Maiti and M. Petersen, *J. Am. Chem. Soc.*, 2006, **128**, 14; (*b*) T. Højland, S. Kumar, B. R. Babu, T. Umemoto, N. Albæk, P. K. Sharma, P. Nielsen and J. Wengel, *Org. Biomol. Chem.*, 2007, **5**, 2375.
34 J. A. Lee, H. R. Moon, H. O. Kim, K. R. Kim, K. M. Lee, B. T. Kim, K. J. Hwang, M. W. Chun, K. A. Jacobsen and L. S. Jeong, *J. Org. Chem.*, 2005, **70**, 5006.
35 For good reviews of this area see: (*a*) S. Karkare and D. Bhatnagar, *Appl. Microbiol. Biotechnol.*, 2006, **71**, 575; (*b*) S. Shakeel, S. Karim and A. Ali, *J. Chem. Technol. Biotechnol.*, 2006, **81**, 892; (*c*) Z. V. Zhilina, A. J. Ziemba and S. W. Ebbinghaus, *Curr. Top. Med. Chem.*, 2005, **5**, 1119.
36 K. Haraguchi, N. Shiina, Y. Yoshimura, H. Shimada, K. Hashimoto and H. Tanaka, *Org. Lett.*, 2004, **6**, 2645.
37 (*a*) N. Inoue, D. Kaga, N. Minakawa and A. Matsuda, *J. Org. Chem.*, 2005, **70**, 8597; (*b*) N. Inoue, N. Minakawa and A. Matsuda, *Nucleic Acids Res.*, 2006, **34**, 3476.
38 (*a*) M. Ober, H. Müller, C. Pieck, J. Gierlich and T. Carell, *J. Am. Chem. Soc.*, 2005, **127**, 18143; (*b*) M. Ober, U. Linne, J. Gierlich and T. Carell, *Angew. Chem., Int. Ed.*, 2003, **42**, 4947.
39 M. S. Christensen, C. M. Madsen and P. Nielsen, *Org. Biomol. Chem.*, 2007, **5**, 1586.
40 S. L. Pedersen and P. Nielsen, *Org. Biomol. Chem.*, 2005, **3**, 3570.
41 (*a*) L. L. Zhang, A. Peritz and E. Meggers, *J. Am. Chem. Soc.*, 2005, **127**, 4174; (*b*) L. L. Zhang, A. E. Peritz, P. J. Carroll and E. Meggers, *Synthesis*, 2006, 645.
42 A. Eschenmoser, *Chimia*, 2005, **59**, 836.
43 M. Egli, P. S. Pallan, R. Pattanayek, C. J. Wilds, P. Lubini, G. Minasov, M. Dobler, C. J. Leumann and A. Eschenmoser, *J. Am. Chem. Soc.*, 2006, **128**, 10847.
44 (*a*) For definitions, nomenclature and conformation behaviour of nuclosides and nucleotides, see: W. Saenger, *Principles of Nucleic Acid Structure*, Springer, New York, 1984; (*b*) H. Ford, Jr., F. Dai, L. Mu, M. A. Siddiqui, M. C. Nicklaus, L. Anderson, V. E. Marquez and J. J. Barchi, Jr., *Biochemistry*, 2000, **39**, 2581.
45 (*a*) P. K. Sharma, M. Petersen and P. Nielsen, *J. Org. Chem.*, 2005, **70**, 4918, which also cites the following articles on the subject: (*b*) E. T. Kool, *Chem. Rev.*, 1997, **97**, 1473; (*c*) P. Herdewijn, *Biochim. Biophys. Acta*, 1999, **1489**, 167; (*d*) M. Meldgaard and J. Wengel, *J. Chem. Soc., Perkin Trans. 1*, 2000, 3539; (*e*) C. J. Leumann, *Bioorg. Med. Chem.*, 2002, **10**, 841; (*f*) J. Kurreck, *Eur. J. Biochem.*, 2003, **270**, 1628.
46 (*a*) D. A. Braasch and D. R. Corey, *Chem. Biol.*, 2001, **8**(1), 1–7; (*b*) D. A. Braasch, Y. H. Liu and D. R. Corey, *Nucleic Acids Res.*, 2002, **30**, 5160; (*c*) J. L. Childs, M. D. Disney and D. H. Turner, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 11091.
47 S. K. Singh, P. Nielsen, A. A. Koshkin and J. Wengel, *Chem. Commun.*, 1998, 456.
48 For example: (*a*) V. K. Rajwanshi, R. Kumar, M. Kofod-Hansen and J. Wengel, *J. Chem. Soc., Perkin Trans. 1*, 1999, 1407; (*b*) S. Obika, K.-i. Morio, D. Nanbu, Y. Hari, H. Itoh and T. Imanishi, *Tetrahedron*, 2002, **58**, 3039; (*c*) D. S. Pedersen, C. Rosenbohm and T. Koch, *Synthesis*, 2002, 802; (*d*) M. D. Sørensen, L. Kværnø, T. Bryld, A. E. Håkansson, B. Verbeure, G. Gaubert, P. Herewijn and J. Wengel, *J. Am. Chem. Soc.*, 2002, **124**, 2164; (*e*) D. S. Pedersen and T. Koch, *Synthesis*, 2004, 578; (*f*) N. K. Christensen, T. Bryld, M. D. Sørensen, K. Arar, J. Wengel and P. Nielsen, *Chem. Commun.*, 2004, 282; (*g*) Raunak, B. R. Babu, V. S. Parmar, N. H. Harrit and J. Wengel, *Org. Biomol. Chem.*, 2004, **2**, 80; (*h*) M. Raunkjær, M. D. Sørensen and J. Wengel, *Org. Biomol. Chem.*, 2005, **3**, 130; (*i*) G. Gaubert, B. R. Babu, S. Vogel, T. Bryld, B. Vester and J. Wengel, *Tetrahedron*, 2006, **62**, 2278.
49 T. S. Kumar, A. S. Madsen, J. Wengel and P. J. Hrdlicka, *J. Org. Chem.*, 2006, **71**, 4188.
50 D. Honcharenko, O. P. Varghese, O. Plashkevych, J. Barman and J. Chattopadhyaya, *J. Org. Chem.*, 2006, **71**, 299.
51 (*a*) L. A. Paquette and S. Dong, *J. Org. Chem.*, 2005, **70**, 5655. For the synthesis of the ribose equivalent, see: (*b*) S. Dong and L. A. Paquette, *J. Org. Chem.*, 2006, **71**, 1647.
52 L. A. Paquette, *Aust. J. Chem.*, 2004, **57**, 7.
53 (*a*) R. P. Iyer, W. Egan, J. B. Regan and S. L. Beaucage, *J. Am. Chem. Soc.*, 1990, **112**, 1253; (*b*) R. P. Iyer, L. R. Phillips, W. Egan, J. B. Regan and S. L. Beaucage, *J. Org. Chem.*, 1990, **55**, 4693.
54 Q. Song, Z. Wang and Y. S. Sanghvi, *Nucleosides, Nucleotides Nucleic Acids*, 2003, **22**, 629.
55 M. V. Rao, C. B. Reese and Z. Y. Zhao, *Tetrahedron Lett.*, 1992, **33**, 4839.
56 V. T. Ravikumar, M. Andrade, R. L. Carty, A. Dan and S. Barone, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 2513.
57 (*a*) P. Guga, A. Maciaszek and W. J. Stec, *Org. Lett.*, 2005, **7**, 3901. For the synthesis of an H-phosphonselenoate monoester, see: (*b*) M. Kullberg and J. Stawinski, *Synthesis*, 2005, 1668.
58 Q. Du, N. Carrasco, M. Teplova, C. J. Wilds, M. Egli and Z. Huang, *J. Am. Chem. Soc.*, 2002, **124**, 24.
59 For another example of predetermined P-chirality, see: B. Nawrot, B. Rebowska, K. Cieślińska and W. J. Stec, *Tetrahedron Lett.*, 2005, **46**, 6641.
60 (*a*) Q. Xu, D. Katkevica and E. Rozners, *J. Org. Chem.*, 2006, **71**, 5906; (*b*) Y. Liu and E. Rozners, *J. Org. Chem.*, 2005, **70**, 9841; (*c*) E. Rozners, D. Katkevica, E. Bizdena and R. Stromberg, *J. Am. Chem. Soc.*, 2003, **125**, 12125; (*d*) Y. Liu and E. Rozners, *Org. Lett.*, 2003, **5**, 181; (*e*) E. Rozners and R. Stromberg, *Nucleosides Nucleotides*, 1997, **16**, 967.
61 C. M. Yamada, D. J. Dellinger and M. H. Caruthers, *J. Am. Chem. Soc.*, 2006, **128**, 5251.
62 (*a*) P. K. Sharma, B. H. Mikkelsen, M. S. Christensen, K. E. Nielsen, C. Kirchhoff, S. L. Pedersen, A. M. Sørensen, K. Østergaard, M. Petersen and P. Nielsen, *Org. Biomol. Chem.*, 2006, **4**, 2433; (*b*) P. Børsting, M. S. Christensen, S. I. Steffansen and P. Nielsen, *Tetrahedron*, 2006, **62**, 1139; (*c*) P. Børsting, K. E. Nielsen and P. Nielsen, *Org. Biomol. Chem.*, 2005, **3**, 2183; (*d*) P. Børsting, M. Freitag and P. Nielsen, *Tetrahedron*, 2004, **60**, 10955; (*e*) C. Kirchhoff and P. Nielsen, *Tetrahedron Lett.*, 2003, **44**, 6475; (*f*) P. Børsting and P. Nielsen, *Chem. Commun.*, 2002, 2140; (*g*) P. Børsting, A. M. Sørensen and P. Nielsen, *Synthesis*, 2002, 797; (*h*) A. M. Sørensen, K. E. Nielsen, B. Vogg, J. P. Jacobsen and P. Nielsen, *Tetrahedron*, 2001, **57**, 10191; (*i*) A. M. Sørensen and P. Nielsen, *Org. Lett.*, 2000, **2**, 4217.