Chemistry for the synthesis of nucleobase-modified peptide nucleic acid*

R. H. E. Hudson[‡], R. D. Viirre, Y. H. Liu, F. Wojciechowski, and A. K. Dambenieks

Department of Chemistry, the University of Western Ontario, London, Ontario N6A 5B7, Canada

Abstract: Peptide nucleic acid (PNA) presents a versatile scaffold for chemical modifications that may benefit its solubility and hybridization properties, conjugation chemistry, cell membrane permeability, and so forth. We have employed straightforward chemical methods for the synthesis of PNA monomers containing C5- or C6-modified pyrimidines. The C5-modified pyrimidines are based on 5-hydroxymethyl- uracil or cytosine, or are achieved by cross-coupling from the 5-iodonucleobase derivatives, while C6 modifications are accessed from 6-carboxyuracil (orotic acid). We have also developed an on-resin activation/conversion of uracil-containing PNA oligomers to N^4-cytosine-containing derivatives.

INTRODUCTION

A renaissance of oligonucleotide analog chemistry was sparked over the past decade by reports from Nielsen and coworkers on peptide nucleic acid (PNA) [1]. Although formally neither a peptide nor a nucleic acid, PNA embodies structural features from each of these classes of biomolecules. The original design of PNA is based on a 2-aminoethylglycine repeat unit to which pendant nucleobases are attached, thus preserving the basic geometric features of natural nucleic acids (Fig. 1). PNA hybridizes avidly and with high sequence selectivity to complementary nucleic acids [2]. These properties combined with its biological and chemical stability, as compared to unmodified DNA or RNA, has attracted attention for applications ranging from in vitro diagnostics to antisense agents [3].
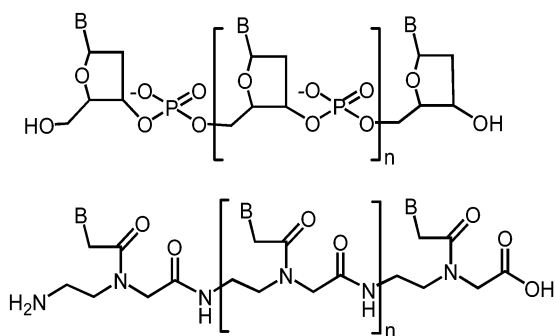
Fig. 1 Comparison of the generalized structures of DNA (top) and PNA (bottom). B = nucleobase.


Many second-generation PNA-like molecules have been designed, synthesized, and evaluated for interesting or improved properties. The focus of much of this research has been on changes to the polyamide backbone. For example, charged or hydrophilic modification may be introduced by incorporation of amino acids other than glycine [4], or by replacement of the backbone with an aliphatic heterocycle [5]. Other work has been concerned with introducing ring structures into the backbone with the goal of increasing the selectivity or strength of interaction with natural nucleic acids [4]. These types of modifications introduce stereocenters into the polymer which present additional complexity in both the synthesis and hybridization properties.

NUCLEOBASE-MODIFIED PNA

A generous selection of nonstandard nucleobases has been incorporated into PNA, most often with the goal of controlling some aspect of hybridization. For example, modified nucleobases for triple helix formation have been designed as well as modified bases to avoid PNA:PNA duplex formation in complementary sequences [4,6,7]. Conversely, our work has concentrated on modification of the non-Watson/Crick base-pairing face of the pyrimidines as an alternative position for derivatization with charged, hydrophilic, reactive or potential helix-stabilizing substituents. Employing straightforward methods, we have prepared derivatives to evaluate the effect of introduction of substituents at the C4, C5, and C6 positions. The advantage of making such derivatives is that these positions direct the substituent into the major groove and are generally very tolerant of modification (esp. C5), Fig. 2. Additionally, these positions do not directly affect the base-pairing face (esp. C5 and C6) and do not introduce stereocenters, thus simplifying some aspects of the synthesis.


(a)   **(b)**
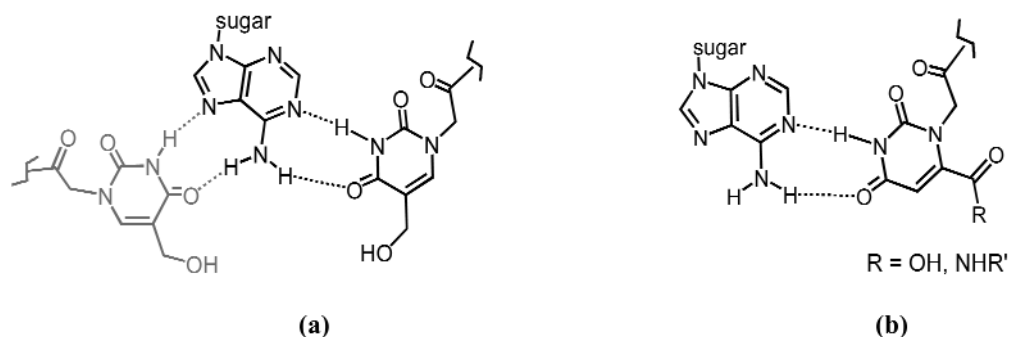
Fig. 2 (a) Expected base-pairing of deoxyadenosine with C5-hydroxymethyluracil-containing PNA in duplex formation (black) and triplex formation (black and gray). (b) Expected base-pairing between deoxyadenosine (DNA strand) and orotic acid derivatives (PNA strand).

Using the chemical approaches described herein, functional groups similar to those that have been placed in the backbone (hydrophilic or charged) have been incorporated into the nucleobase. As well, unprecedented derivatives with potentially new and useful properties have been synthesized. The chemistry that we have used for the derivatization of pyrimidines is reviewed below.

## C5-Alkynylpyrimidine and related derivatives

Our initial foray into nucleobase-modified PNA utilized chemistry that was proven for the nucleosides: Pd-catalyzed cross-coupling of aryl iodides with terminal alkynes [8,9]. The resulting C5-alkynyl modification has been found useful in oligonucleotide chemistry as a helix-stabilizing and antisense-potentiating substitution [10]. A known side reaction during cross-coupling is the formation of the annulated

product furanouracil [11]. This heterocycle is of interest as a fluorescent nucleobase analog, but its hydrogen-bonding properties are not preserved from the parent heterocycle, which limits its use (Fig. 3).
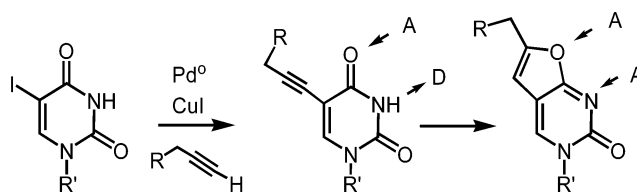
Fig. 3 Conversion of 5-iodouracil derivatives to 5-alkynyl- and subsequently to furanouracil derivatives and the effect on hydrogen bond donor (D) and acceptor (A) sites.

While extending the cross-coupling chemistry to 5-iodocytosine, we rediscovered a chemical switch that controls which product is formed. Iodocytosine derivatives cleanly give the 5-alkynyl product in high yield while N^4-benzoyl-5-iodocytosine derivatives lead to the annulated product with concomitant loss of the *N*-acyl group, as shown in Fig. 4 [12–14]. Importantly, the pyrrolocytosine derivatives are fluorescent and possess the same H-bonding pattern as the parent heterocycle [15,16]. Together with the observation that the fluorescence of pyrrolocytosine is quenched during hybrid formation makes it useful as a reporter group [16]. As well, the pyrrolocytosine nucleobase is expected to be a helix-stabilizing substitution on the basis of increased π-stacking interactions as compared to cytosine.
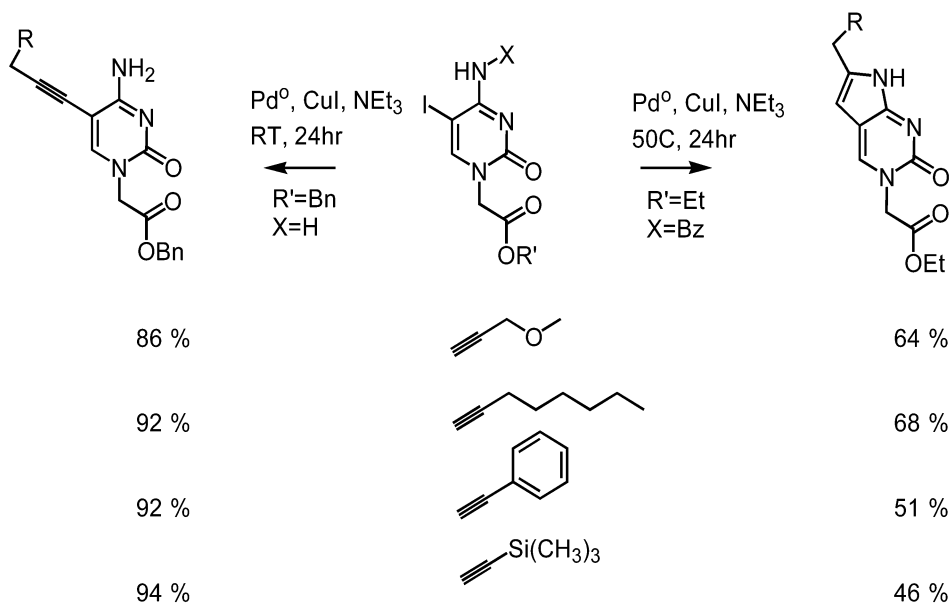

**Fig. 4** N^4-protecting group influence on the partitioning of products.

We have also prepared PNA monomers compatible with Boc oligomerization chemistry for the 5-alkynyl- and pyrrolocytosine products derived from reaction with methyl 2-propynyl ether, which illustrates that the pyrrolocytosine heterocycle need not necessarily be blocked during amide bond formation and is stable to ester hydrolysis conditions. Subsequent studies are directed toward investigating the effects of incorporating these modifications into PNA.

C5-Hydroxymethylpyrimidines and related derivatives

The reaction of formaldehyde with uracil to yield 5-hydroxymethyluracil is well known [17]. We envision access to a variety of substituted uracils by exploiting this chemistry [18]. For example, we have prepared PNA monomers derived from compounds **1**, **2**, and **4** thus far, which represent hydrophilic modifications (**1–3**) or charge modification (**4**). The chemistry leading to PNA monomers based on **1** is detailed below.
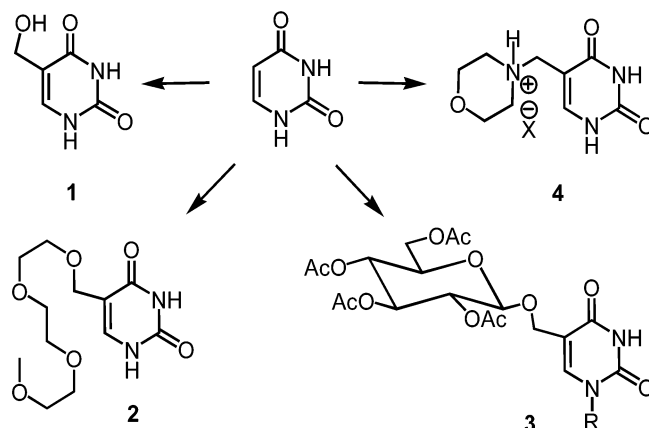
A minimalist approach to the modification of pyrimidines led us to design hydroxymethylated pyrimidine PNA [hm(Y) PNA], Fig. 5. This molecule results from the replacement of a hydrogen on the methyl group of thymine and 5-methylcytosine with a hydroxyl group. This type of oligomer would be accessed via the respective hydroxymethyluracil (hmU) and hydroxymethylcytosine (hmC) monomers.
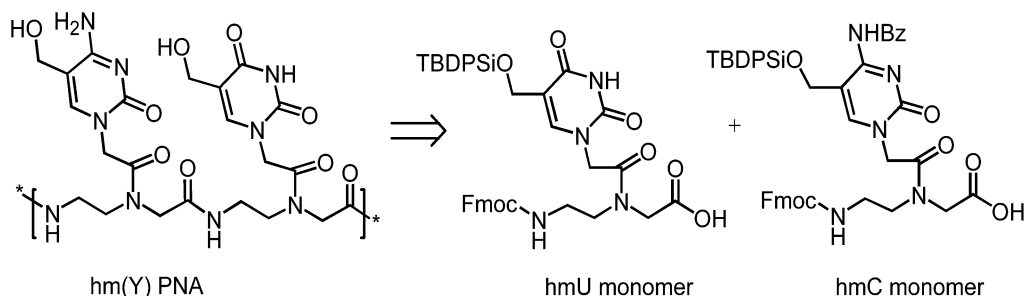

Fig. 5 Hydroxymethylated(Y) PNA and the monomers suitable for its synthesis.

This particular substitution is expected to be well tolerated based on examination of molecular models built from existing structural information for PNA:DNA duplex [19] and PNA:DNA:PNA triplex [20]. As earlier illustrated in Fig. 2, the hydroxyl group should be solvent-exposed in the major groove and sterically compatible with either duplex (PNA:DNA) or triplex (PNA:DNA:PNA) formation. This modification should promote the water solubility of PNA oligomers and also provides a versatile route to other derivatives, vide infra. Other potentially interesting derivatives include: cationic groups to benefit DNA or RNA binding and strand invasion; conjugation to oligo(ethylene glycol) to benefit membrane penetration [20]; and carbohydrates for potential modulation of pharmacokinetic properties [21].

The synthetic route to the hmU and hmC PNA monomers are very similar and efficient. For example, reaction of uracil with aqueous formaldehyde gave 5-hydroxymethyluracil, which was then regioselectively alkylated and 5-O-silylated. After hydrolysis, the nucleobase acetic acid was afforded in >70 % chemical yield over four steps. For hydroxymethylcytosine, the exocyclic amine was protected as the benzamide. The hmU/C PNA monomers were produced after standard condensation with the backbone submonomer (not shown). Current efforts are directed to oligomerization of these monomers.

N4-Cytosine derivatives

While pursuing chemistry not described herein, we prepared an N^3-protected thymine PNA monomer [23]. We reasoned that without an acidic proton on N3, the thymine should resist conversion to cytosine derivatives.

X = -Cl, -OSO$_2$Ar, -triazole
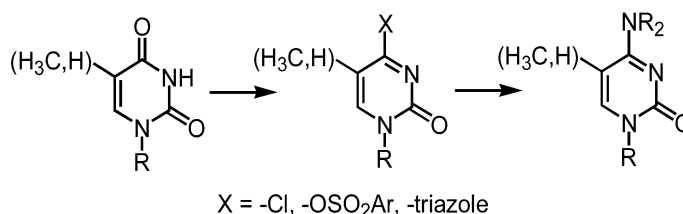
To test this hypothesis, we prepared a model system which consisted of a resin-bound PNA dimer comprised of a uracil and paramethoxybenzyl (PMB) protected thymine. When treated with a solution of a primary amine (e.g., n-BuNH$_2$), the dimer is released from the resin as the C-terminal butylamide. When the resin-bound dimer is first treated with a ten-fold excess of a sulfonyl chloride followed by aminolysis, the only product isolated by RP-HPLC was the cytosine-PMB(thymine) C-terminal amide (Fig. 6).

This protocol has been extended to PNA oligomers and a selection of amines. These results will be reported separately [24].
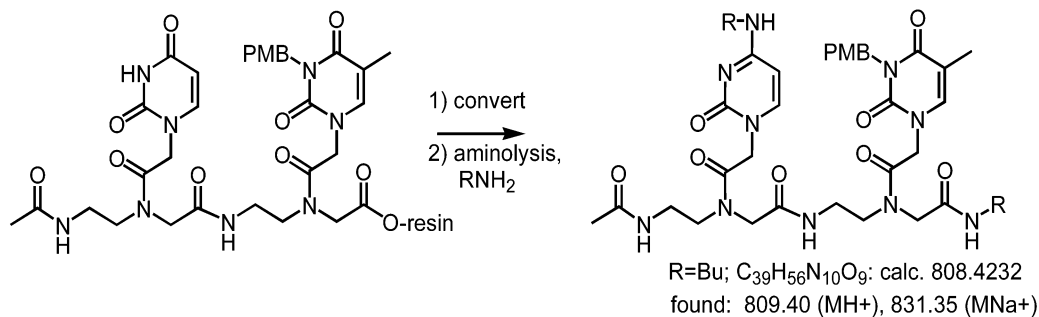

R=Bu; C$_{39}$H$_{56}$N$_{10}$O$_9$: calc. 808.4232
found: 809.40 (MH+), 831.35 (MNa+)

Fig. 6 Chemistry for a convert-and-release strategy for preparing N^4-cytosine derivatives.

C6-Derivatives: Methyl orotate and orotamides

To our knowledge, there has been no report of orotic acid or any orotamide incorporated into oligodeoxynucleotides. It is tempting to speculate that the 6-carboxy or 6-carboxamide group would disfavor the anti-glycosidic bond conformer and subsequently disfavor Watson/Crick duplex formation. However, the PNA backbone is potentially less sterically demanding and more flexible than the backbone of DNA and may tolerate this substitution better. Nonetheless, substitution at C6 is expected to be more sensitive to sterics since it is closer to the PNA backbone, as compared to C5, Fig. 2b. Orotic acid

permits facile entry into this chemistry and will permit comparisons to be made on the effect of substitution at C5 vs. C6.

Our first approach to orotamide-containing PNA focused on the synthesis of individual monomers for discrete incorporation into an oligomer. This route started with orotic acid (**5**) and furnished methyl orotate (**6**) in excellent yield by Fischer esterification. Treatment of methyl orotate with an excess of a small selection of *N*-based nucleophiles lead to the amide or hydrazone (**7a–d**, Fig. 7). However, in all but one instance, it was not possible to alkylate the N1 and further elaborate the heterocycle into a PNA monomer.
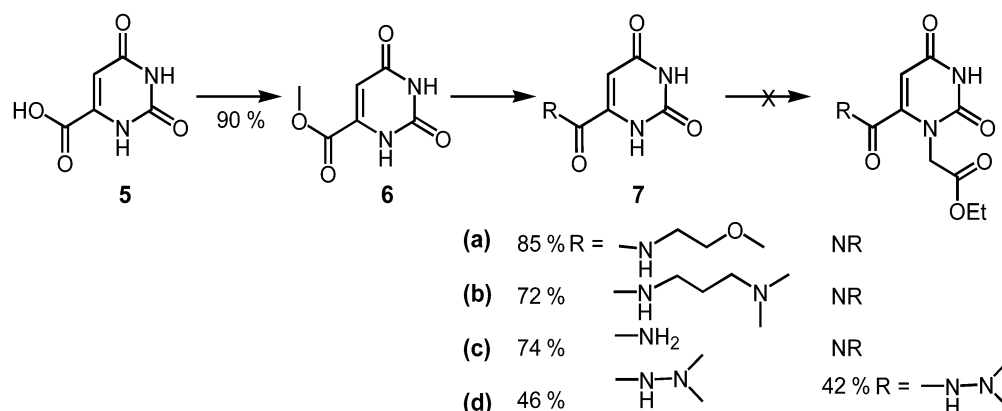
Fig. 7 Initial route to orotamide-containing PNA.

A second approach that has proven successful is the synthesis of a methyl orotate PNA monomer. This monomer will be incorporated into PNA oligomers by conventional solid-phase peptide synthesis (SPPS) and may be converted to orotic acid or an orotamide depending on the conditions used for conversion and release from the resin, Fig. 8.
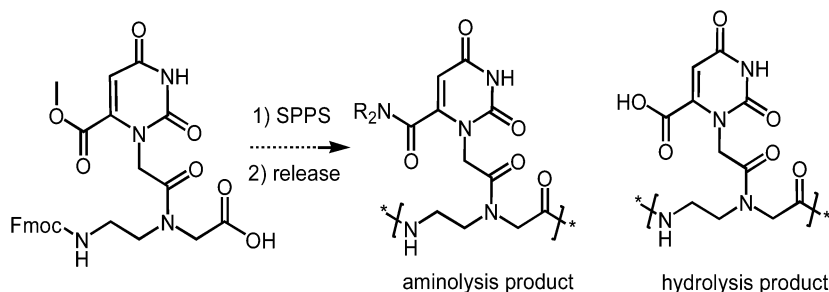
Fig. 8 A convert-and-release strategy for the synthesis of orotamide or orotic acid-containing PNA.

SUMMARY AND OUTLOOK

In this report, we have summarized our current efforts in the development of synthetic approaches to afford nucleobase-modified PNA. Generally, the nucleobase is a convenient site of modification with accessible chemistry leading to a wide variety of possible substitutions. We have demonstrated the synthesis of pyrimidines bearing hydrophilic, hydrophobic, and charged groups. Efforts are underway to evaluate the properties of oligomers incorporating such modified monomers. As well, the pyrrolocytosine derivatives are being investigated for hybridization state sensitive probes.

ACKNOWLEDGMENTS

REFERENCES

1. (a) P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt. *Science* **254**, 1497 (1991); (b) M. Egholm, O. Buchardt, P. E. Nielsen, R. H. Berg. *J. Am. Chem. Soc.* **114**, 1895 (1992); (c) M. Egholm, P. E. Nielsen, O. Buchardt, R. H. Berg. *J. Am. Chem. Soc.* **114**, 9677 (1992); (d) M. Egholm, C. Behrens, L. Christensen, R. H. Berg, P. E. Nielsen, O. Buchardt. *Chem. Commun.* 800 (1993).
2. M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, P. E. Nielsen. *Nature* **365**, 566 (1993).
3. (a) D. Braasch and D. R. Corey. *Biochemistry* **41**, 4503 (2002); (b) F. Beck and P. E. Nielsen. In *Artificial DNA: Methods and Applications*, Y. E. Khudyakov and H. A. Fields (Eds.), pp. 91–114, CRC Press, Boca Raton, FL (2003).
4. Reviewed in: K. Ganesh and P. E. Nielsen. *Curr. Org. Chem.* **4**, 931 (2000).
5. (a) M. D'Costa, V. A. Kumar, K. N. Ganesh. *Org. Lett.* **1**, 1513 (1999); (b) A. Puschl, T. Tedeschi, P. E. Nielsen. *Org. Lett.* **2**, 4161 (2000); (c) V. A. Kumar *Eur. J. Org. Chem.* 2021 (2002).
6. For example: M. Egholm, L. Christensen, K. Dueholm, O. Buchardt, J. Coull, P. E. Nielsen. *Nucleic Acids Res.* **23**, 217 (1995).
7. J. Lohse, O. Dahl, P. E. Nielsen. *Proc. Natl. Acad. Sci. USA* **96**, 11804 (1999).
8. K. Sonogashira, Y. Tohda, N. Hagihara. *Tetrahedron Lett.* **16**, 4467 (1975).
9. R. H. E. Hudson, G. Li, J. Tse. *Tetrahedron Lett.* **43**, 1381 (2002).
10. R. W. Wagner, M. D. Matteucci, J. G. Lewis, A. J. Gutierrez, C. Moulds, B. C. Froehler. *Science* **260**, 1510 (1993).
11. M. J. Robins and P. J. Barr. *Tetrahedron Lett.* **22**, 421 (1981).
12. R. H. E. Hudson, R. D. Viirre, N. McCourt, J. Tse. *Nucleosides Nucleotides* **22**, 1029 (2003).
13. Tandem cross-coupling/annulation reactions has been exploited in heterocycle synthesis: R. C. Larock. *J. Organomet. Chem.* **576**, 111 (1999).
14. Subsequent to our "discovery", we became aware of an untranslated report of the cross-couple/annulate reaction applied to deoxy-5-iodocytidine: H. Inoue, A. Imura, E. Ohtsuka. *Nippon Kagaku Kaishi* **7**, 1214 (1987).
15. J. Woo, R. B. Meyer, H. B. Gamper, Jr. *Nucleic Acids Res.* **24**, 2470 (1996).
16. C. Liu and C. T. Martin. *J. Mol. Biol.* **308**, 465 (2001).
17. R. E. Cline, R. M. Fink, K. Fink. *J. Am. Chem. Soc.* **81**, 2521 (1959).
18. (a) Etherification of hmU: S. Tardy-Planechaud, J. Fujimoto, S. S. Lin, L. C. Sowers. *Nucleic Acids Res.* **25**, 553 (1997); (b) Glycosylated hmU nucleoside has been reported: R. Tona, R. Bertolini, J. Hunziker. *Org. Lett.* **2**, 1693 (2000); (c) Mannich reaction of uracil with morpholine: J. H. Burckhalter, R. J. Seiwald, H. C. Scarborough. *J. Am. Chem. Soc.* **82**, 991 (1960).
19. M. Eriksson and P. E. Nielsen. *Nat. Struct. Biol.* **3**, 410 (1996).
20. One example of high-molecular-weight PEG-assisting cellular uptake: M. Ballico, S. Drioli, F. Morvan, L. Xodo, G. M. Bonora. *Bioconjugate Chem.* **12**, 719 (2001).
21. R. Hamzavi, F. Dolle, B. Tavitian, O. Dahl, P. E. Nielsen. *Bioconjugate Chem.* **14**, 941 (2003).
22. L. Betts, J. A. Josey, J. M. Veal, S. R. Jordan. *Science* **270**, 1838 (1995).
23. (a) R. D. Viirre, R. H. E. Hudson. *Org. Lett.* **3**, 3931 (2002); (b) R. D. Viirre and R. H. E. Hudson. *J. Org. Chem.* **68**, 1630 (2003).

24. Convertible nucleobases have been incorporated into PNA, but this report is the first example of on-resin activation with simultaneous conversion and release. (a) Uhlmann and coworkers were unable to convert the 4-triazolyl derivative of thymine: E. Uhlmann, A. Peyman, G. Breipohl, D. W. Will. *Angew. Chem., Int. Ed.* **37**, 2796 (1998); (b) E. Ferrer and R. Eritja. *Lett. Pep. Sci.* **7**, 195 (2000); (c) B. G. de la Torre and R. Eritja. *Biorg. Med. Chem. Lett.* **13**, 391 (2003).