Pergamon

Tetrahedron Letters 42 (2001) 1479–1481

TETRAHEDRON LETTERS

Novel phthalimide nucleosides for the specific recognition of a CG Watson–Crick base pair

David Lengeler and Klaus Weisz*

Institut für Chemie der Freien Universität Berlin, Takustraße 3, D-14195 Berlin, Germany

Received 15 November 2000; accepted 4 December 2000

Abstract—In an effort to extend the triple helix recognition code we have synthesized various substituted phthalimide-derived nucleosides that can recognize a CG Watson–Crick base pair. NMR experiments performed on the free nucleosides in methylene chloride at lowered temperatures indicate the strength and extent of H-bonding to the cytosine amino group of a CG base pair which strongly depend on the substituent of the phthalimide nucleobase. Consistent with the formation of an additional hydrogen bond to N7 of guanine, ureido-substituted nucleoside analogs show a higher overall affinity as compared to the 3-aminophthalimide nucleoside.

Triple helix formation through selective binding of a single-stranded oligonucleotide to regions of double-stranded DNA has long been recognized as a powerful tool for various applications. Examples include the site-specific inhibition of transcription or the development of sequence-specific artificial nucleases. Third strand binding is based on the formation of specific hydrogen bonds between nucleobases of the single strand and the purine bases of the Watson–Crick duplex in its major groove.[1] Thus, in the pyrimidine motif formation of Hoogsteen hydrogen bonds between third strand pyrimidine and Watson–Crick purine bases results in isomorphous C^+·GC and T·AT base triplets. As a consequence of this binding mode, the recognition by third strand bases without severe distortions of the oligonucleotide backbone is mostly restricted to GC and AT Watson–Crick base pairs in homopurine·homopyrimidine sequences. In contrast, CG and TA interruptions within the duplex target are not effectively recognized by natural nucleobases.

To overcome this major limitation in the sequence-specific triplex formation, much work has been devoted in the past to the design of non-natural nucleobases for extending the triple helix recognition code.[2] Here we describe the synthesis of novel nucleoside analogs which are based on the easily accessible amino-substituted phthalimide derivative **1** for the selective recognition of a CG Watson–Crick base pair. The 3-aminophthalimide riboside **1** is expected to interact with a CG base pair by forming two hydrogen bonds with the 4-amino group of C and the 6-carbonyl oxygen of G (Fig. 1). Molecular modeling studies suggest that a **1**·CG base triplet should be closely isomorphous to the canonical C^+·GC and T·AT base triplets, thus minimizing any distortions of the sugar-phosphate backbone when incorporated in a triplex-forming oligonucleotide. In addition, the phthalimide heterocycle is expected to be less prone to intercalation, a problem often encountered for oligonucleotide-directed triplex formation using more extended aromatic ring systems as base surrogates.[3]

The synthesis of **1** is outlined in Scheme 1. Due to the poor nucleophilicity of the phthalimide nitrogen, a Vorbrueggen-type coupling of sugar and base is not

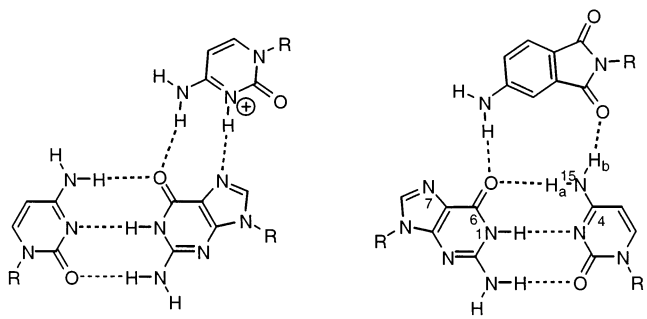
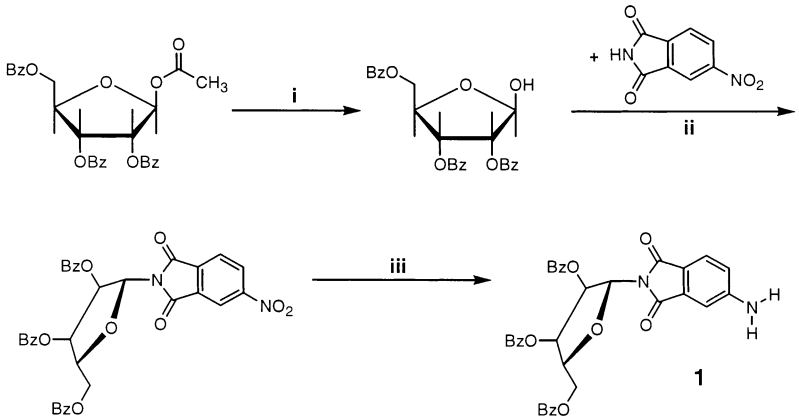


Figure 1. Canonical C^+·GC base triplet (left) and **1** binding a CG base pair to form a **1**·CG base triplet (right).

Keywords: triple helix; hydrogen bond; nucleoside; base pair.

* Corresponding author. Fax: +49-30-838-55310; e-mail: weisz@chemie.fu-berlin.de

0040-4039/01/$ - see front matter
PII: S0040-4039(00)02256-5



Scheme 1. (i) Bu_3SnOMe, $C_2H_4Cl_2$, 32%; (ii) PPh_3, DEAD, THF, 37%; (iii) $Fe_3(CO)_{12}$, toluene, MeOH, 86%.

successful. However, deacetylation of 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-β-D-ribofuranose at the 1-position with Bu_3SnOMe[4] and reaction with 3-nitrophthalimide under Mitsunobu conditions leads to *N*-(2′,3′,5′-tri-*O*-benzoyl-β-D-ribofuranosyl)-3-nitrophthalimide. Subsequent reduction of the nitro group to give **1** is accomplished under mild conditions with $Fe_3(CO)_{12}$ and methanol.[5]

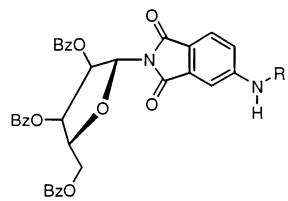
For increasing the acidity of the NH group and for extending possible hydrogen bond contacts to N7 of the guanine base we have synthesized additional derivatives **2**, **3** and **4** (Fig. 2). Reaction of **1** with benzoylchloride in THF/pyridine gives *N*-benzoylated **2** (82% yield). The urea group in compounds **3** and **4** was introduced through the reaction of **1** with chlorosulfonylisocyanate in CH_2Cl_2 and with *n*-butylisocyanate in THF in 83 and 64% yield, respectively.[6]

To study hydrogen bond interactions between our nucleoside analogs and a CG base pair, we have performed NMR measurements on a 1:1 mixture of 3′,5′-di-*O*-(triisopropylsilyl) protected 2′-deoxycytidine and 2′-deoxyguanosine in the presence of the free tri-*O*-benzoylated ribonucleosides **1**, **2**, **3** or **4** in CD_2Cl_2. Hydrogen bond formation is easily followed by a downfield shift of the corresponding proton resonance involved in hydrogen bonding. The observed chemical shift depends on: (i) the association constant for complex formation; (ii) on the limiting chemical shift of the hydrogen-bonded proton in a complex which is a measure of the geometry and strength of the hydrogen bond;[7] (iii) on the limiting chemical shift of the non-hydrogen-bonded proton in a monomer which is influenced by its individual chemical environment. In order to minimize contributions due to different proton chemical environments, we have recorded chemical shifts of the same non-Watson–Crick hydrogen-bonded cytosine amino proton NH_b in the presence of each of the four synthetic nucleosides. Clearly, this proton is always expected to be engaged in a hydrogen bond for a ternary complex.

Our strategy also makes use of low temperatures and an excess of the nucleoside surrogate to promote association. Unfortunately, observation of the cytosine amino proton NH_b is severely hampered under such conditions due to spectral overlap. However, employing specifically 4-^{15}N amino labeled cytidine enables easy detection of the amino proton even in the presence of strong overlapping proton resonances of the excess nucleoside analog by means of a ^{1}H-^{15}N filter, as implemented in a one-dimensional HMQC experiment.[8]

Due to the large association constant (K_{ass}>10^4–10^5 M^{-1} in $CHCl_3$ at 293 K) a 1:1 mixture of C and G in CD_2Cl_2 fully associates to form a CG Watson–Crick base pair that may interact with a third base by its residual hydrogen bond donor and acceptor sites. Thus, a downfield shift of the C H_b signal upon addition of a tenfold excess of **1**, **2**, **3** or **4** to a 1:1 mixture of 4-^{15}N-dC and dG in CD_2Cl_2 at 240 K indicates the formation of ternary complexes with the cytosine amino proton involved in a hydrogen bond with the phthalimide derivatives (Fig. 3).[9] In contrast, no significant changes in their chemical shift were observed for G H1 and C H_a resonances consistent with their participation in a Watson–Crick hydrogen bond that is not disrupted in the presence of the nucleoside analog.

Recently, an alkylated hexylureido phthalimide was used as a synthetic receptor for a CG base pair and found to bind via three hydrogen bonds with high affinity in chloroform ($K_{ass}\sim10^3$ M^{-1} at room temperature).[10] Assuming only minor influences of different residues and solvent, it is expected that under our low temperature conditions all of the CG base pairs have associated with the urea analogs **3** and **4** being in excess.[11] Consequently, the C H_b chemical shift should directly reflect the strength of the hydrogen bond between the cytosine amino and phthalimide carbonyl



1 : R = -H
2 : R = -CO-C_6H_5
3 : R = -CO-NH_2
4 : R = -CO-NH-C_4H_9

Figure 2. Nucleoside **1** and derivatives **2**, **3** and **4**.

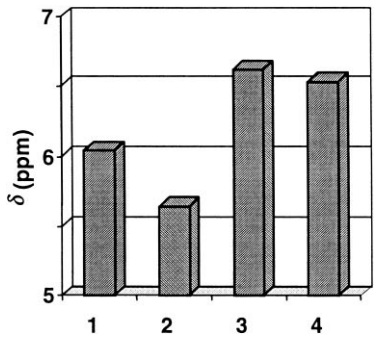


Figure 3. ^{1}H chemical shift δ for the cytosine amino proton H_b in a mixture of [4-^{15}N]-3′,5′-di-*O*-(triisopropylsilyl)-2′-deoxycytidine, 3′,5′-di-*O*-(triisopropylsilyl)-2′-deoxyguanosine and **1**–**4** (1:1:10) in CD_2Cl_2 at 240 K; $c_{CG}=5.45$ mM.

group, pointing to a slightly stronger hydrogen bond for **3** with a non-alkylated urea substituent. Compared to **3** and **4**, H_b resonances in complexes of **1** and **2** with only two potential hydrogen bonds are less downfield shifted. Excluding any cooperative effects, loss of an additional hydrogen bond contact is always expected to improve the geometry and strength of residual contacts unless there is a perfect steric fit. Thus, the available chemical shift data can only be rationalized by lower association constants with only partial CG complexation by **1** and **2** even at lower temperatures. Compared to **2**, **1** shows stronger binding towards the cytosine amino group. This might be due to steric hindrance of the bulky benzoyl group. In addition, a stronger hydrogen bond involving the more acidic amide proton of **2** might compromise the strength of the NH_b–O hydrogen bond by geometric readjustments.

In summary, we have synthesized four differently substituted phthalimide-derived nucleoside analogs that can bind a CG base pair via specific hydrogen bonds. By employing selectively ^{15}N amino labeled cytosine as a local probe, NMR experiments may not only provide rapid information on their affinity towards a CG base pair but also on the strength of individual hydrogen bonds, thus complementing future studies on the thermodynamics of binding. It is expected that such systematic investigations will yield more insight into the binding of a CG base pair with novel nucleobases and constitute a first step in understanding the complex interplay of the different interactions effective in a triple helical oligonucleotide.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft for financial support.

References

1. Thuong, N. T.; Hélène, C. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 666–690.
2. (a) Griffin, L. C.; Kiessling, L. L.; Beal, P. A.; Gillespie, P.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 7976–7982; (b) Huang, C.-Y.; Cushman, C. D.; Miller, P. S. *J. Org. Chem.* **1993**, *58*, 5048–5049; (c) Sasaki, S.; Nakashima, S.; Nagatsugi, F.; Tanaka, Y.; Hisatome, M.; Maeda, M. *Tetrahedron Lett.* **1995**, *36*, 9521–9524; (d) Zimmerman, S. C.; Schmitt, P. *J. Am. Chem. Soc.* **1995**, *117*, 10769–10770; (e) Huang, C.-Y.; Bi, G.; Miller, P. S. *Nucleic Acids Res.* **1996**, *24*, 2606–2613; (f) Lehmann, T. E.; Greenberg, W. A.; Liberles, D. A.; Wada, C. K.; Dervan, P. B. *Helv. Chim. Acta* **1997**, *80*, 2002–2022; (g) Lecubin, F.; Benhida, R.; Fourrey, J.-L.; Sun, J.-S. *Tetrahedron Lett.* **1999**, *40*, 8085–8088.
3. Koshlap, K. M.; Gillespie, P.; Dervan, P. B.; Feigon, J. *J. Am. Chem. Soc.* **1993**, *115*, 7908–7909.
4. Nudelman, A.; Herzig, J.; Gottlieb, H. E.; Keinan, E.; Sterling, J. *Carbohydr. Res.* **1987**, *162*, 145–152.
5. Landesberg, J. M.; Katz, L.; Olsen, C. *J. Org. Chem.* **1972**, *37*, 930–936.
6. Analytical and spectral data for the nucleoside analogs. Compound **1**: ^{1}H NMR (250 MHz, CD_2Cl_2): δ (ppm) 4.54–4.75 (m, 3H; H4′, H5′, H5″), 6.06 (d, 1H; H1′), 6.19 (app. t, 1H; H3′), 6.33 (dd, 1H; H2′), 7.36–8.09 (m, 18H; ArH). Anal. calcd for $C_{34}H_{26}N_2O_9$: C, 67.32; H, 4.32; N, 4.62; found: C, 66.57; H, 4.21; N, 4.38. Compound **2**: ^{1}H NMR (250 MHz, $CDCl_3$): δ (ppm) 4.56–4.81 (m, 3H; H4′, H5′, H5″), 6.09 (d, 1H; H1′), 6.17 (app. t, 1H; H3′), 6.28 (dd, 1H; H2′), 7.29–8.25 (m, 23H; ArH), 8.62 (s, 1H; NH). Anal. calcd for $C_{41}H_{30}N_2O_{10}$: C, 69.29; H, 4.25; N, 3.94; found: C, 68.35; H, 4.43; N, 3.72. Compound **3**: ^{1}H NMR (250 MHz, CD_2Cl_2): δ (ppm) 4.56–4.82 (m, 3H; H4′, H5′, H5″), 5.48 (s, br, 2H; NH_2), 6.12 (d, 1H; H1′), 6.19 (app. t, 1H; H3′), 6.29 (dd, 1H; H2′), 7.27–8.03 (m, 18H; ArH), 8.21 (s, 1H; NH). Anal. calcd for $C_{35}H_{27}N_3O_{10}$: C, 64.71; H, 4.19; N, 6.47; found: C, 63.97; H, 4.32; N, 6.36. Compound **4**: ^{1}H NMR (250 MHz, $CDCl_3$): δ (ppm) 0.87 (t, 3H; CH_3), 1.22–1.50 (m, 4H; CH_2CH_2), 3.22 (t, 2H; N-CH_2), 4.54–4.79 (m, 3H; H4′, H5′, H5″), 6.08 (d, 1H; H1′), 6.19 (app. t, 1H; H3′), 6.29 (dd, 1H; H2′), 7.22–8.09 (m, 18H; ArH). Anal. calcd for $C_{39}H_{35}N_3O_{10}$: C, 66.38; H, 5.00; N, 5.95; found: C, 66.31; H, 5.03; N, 5.56.
7. (a) Golubev, N. S.; Shenderovich, I. G.; Smirnov, S. N.; Denisov, G. S.; Limbach, H.-H. *Chem. Eur. J.* **1999**, *5*, 492–497; (b) Dingley, A. J.; Masse, J. E.; Peterson, R. D.; Barfield, M.; Feigon, J.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 6019–6027; (c) Dunger, A.; Limbach, H.-H.; Weisz, K. *J. Am. Chem. Soc.* **2000**, *122*, 10109–10114.
8. For the preparation of 4-^{15}N labeled 2′-deoxycytidine, see: Kupferschmitt, G.; Schmidt, J.; Schmidt, T.; Fera, B.; Buck, F.; Rüterjans, H. *Nucleic Acids Res.* **1987**, *15*, 8225–6241.
9. Note that care was taken to ensure that nucleoside concentration and molar ratios were equal for all samples.
10. Mertz, E.; Mattei, S.; Zimmerman, S. C. *Org. Lett.* **2000**, *2*, 2931–2934.
11. Self-association of the phthalimide nucleoside **4** in methylene chloride was found to be much smaller than the expected binding to a CG base pair ($K_{ass,self}=13$ M^{-1} at room temperature).