Pergamon

Tetrahedron Letters 43 (2002) 61–64

TETRAHEDRON LETTERS

Nucleosides derived from urocanic acid: potential ligands for CG base pairs

Maria G. M. Purwanto, David Lengeler and Klaus Weisz*

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

Received 29 September 2001; accepted 24 October 2001

Abstract—A nucleoside analog based on imidazole-4-acrylamide (urocanamide) was synthesized and studied for its use as a specific ligand for a cytosine–guanosine Watson–Crick base pair. One- and two-dimensional ^{1}H NMR experiments in methylene chloride at ambient and low temperatures not only indicate the strength of association but also confirm specific binding of the novel nucleoside to the base pair through the formation of two hydrogen bonds.

The use of natural or non-natural nucleosides as specific receptors for a given Watson–Crick base pair is prerequisite for various applications in biotechnology and medicine. In particular, the so-called antigene approach of targeting regions of double-stranded DNA with a single-stranded oligonucleotide to form a triple helix makes use of such interactions. Binding of the third strand is based on the formation of specific hydrogen bonds between nucleosides of the single strand and the purine bases of the Watson–Crick duplex in its major groove.[1] However, only GC and AT Watson–Crick base pairs in homopurine·homopyrimidine sequences are effectively recognized by natural nucleobases. In contrast, CG and TA interruptions within the duplex target limit triple helix formation due to their lack of specific and favorable interactions with bases of the third strand. This major restriction in the sequence-specific triplex formation has prompted the development of non-natural nucleobases with suitable arrangement of hydrogen bond donor and acceptor

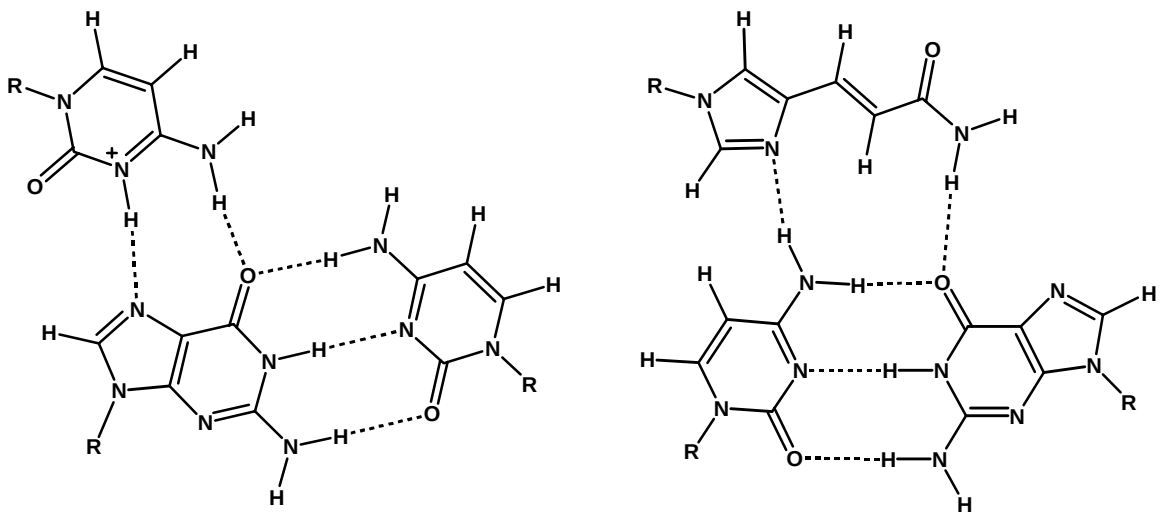


Figure 1. C^+·GC (left) and **1**·CG base triad (right) formed through specific binding of a protonated cytosine and the urocanamide nucleoside **1** to a GC and CG Watson–Crick base pair, respectively.

Keywords: nucleoside; urocanic acid; triple helix; hydrogen bond.

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

0040-4039/02/$ - see front matter
PII: S0040-4039(01)02046-9

functionalities for extending the triple helix recognition code.[2]

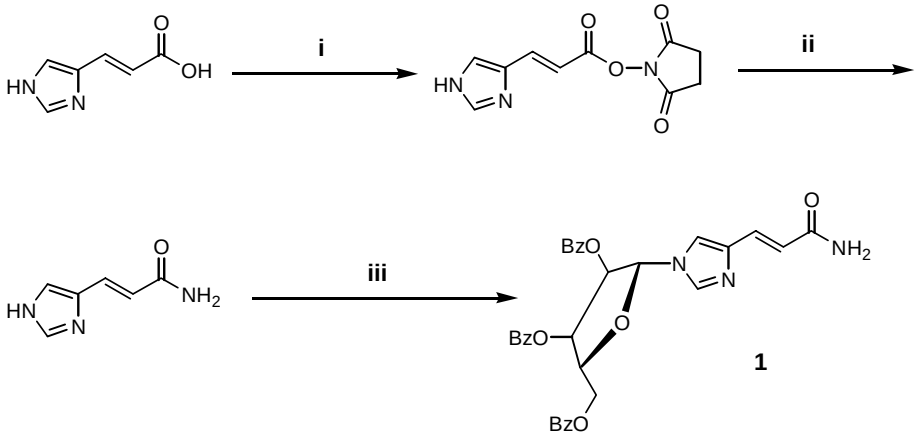
Various compounds derived from *trans*-urocanic acid have been synthesized and tested in view of being potent histamine H_2 receptor agonists or H_3 receptor antagonists in the past.[3] However, due to the structural features of the imidazole-containing carboxylic acid, corresponding nucleosides of such compounds are also expected to target CG Watson–Crick base pairs of double-helical DNA through the formation of specific hydrogen bonds. In particular, the amide of urocanic acid **1** seems to be a promising CG targeting surrogate given the arrangement of hydrogen bond donor and acceptor sites (Fig. 1). Moreover, molecular models show that nucleosides based on the urocanamide may form base triples with CG base pairs closely isomorphous to the canonical C^+·GC and T·AT base triads, prerequisite for their selective binding in triple-helical structures without excessive backbone distortions.

Condensation of the imidazole sodium salt with a 1-chloro sugar derivative has found wide applications in the stereo- and regioselective synthesis of imidazole nucleosides.[4] Here we employed the one-step–one-pot Silyl–Hilbert–Johnson procedure in the presence of a Friedel–Crafts catalyst for the glycosylation of the urocanamide. This Vorbrüggen reaction is commonly employed in pyrimidine and purine nucleoside synthesis but has found less widespread application for the preparation of imidazole nucleosides.[5] The synthesis of 3-[1-(2′,3′,5′-tri-*O*-benzoyl-β-ribofuranosyl)imidazol-4-yl]acrylamide **1** is outlined in Scheme 1. Starting with *trans*-urocanic acid (imidazole-4-acrylic acid) its activated *N*-hydroxysuccinimide ester was prepared by addition of *N*-hydroxysuccinimide (NHS) and *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride (DCI) in DMF under an inert atmosphere in the dark. Reaction with ammonia yielded the amide of urocanic acid which was glycosylated to the nucleoside with 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-β-D-ribofuranose under Vorbrüggen conditions.

The regio- and stereospecificity of the nucleoside synthesis was confirmed by NMR experiments in methylene chloride solution. A reliable method based on proton cross-ring coupling constants was earlier reported for the differentiation between 1,4- and 1,5-disubstituted imidazoles.[6] Accordingly, $J_{2,5}$ is larger than $J_{2,4}$ and measures in the range 1.1–1.5 Hz while $J_{2,4}$ measures in the range 0.9–1.0 Hz. With a scalar coupling of 1.3 Hz determined for **1**, 4-substitution is clearly indicated.[7]

Complex formation of the novel nucleoside **1** with a 1:1 mixture of TIPS-protected cytidine and guanosine was studied by 2D NOE experiments in methylene chloride at 213 K.[8] Measurements at low temperatures benefit from an enhanced association and from slower exchange processes. A portion of the homonuclear 2D NOE spectrum is shown in Fig. 2. Typical NOE connectivities observed between amino and imino protons of cytidine and guanosine indicate the formation of an intact CG Watson–Crick base pair that is not disrupted in the presence of **1** (not shown). In addition, intranucleotide NOE contacts of both imidazole H2 and H5 protons to H1′, H2′, and H3′ sugar protons again identify a 4-substituted imidazole nucleoside. Correspondingly, no NOE cross-peaks are observed between both vinyl protons and their own sugar protons as expected for a 4-substituent pointing away from the sugar moiety (Fig. 2). Most notably, however, intermolecular contacts between base protons H2 of **1** and cytosine H5 and NH_{2a} amino protons as well as between αH of **1** and the cytosine amino proton are easily identified (Fig. 2). These results clearly indicate formation of a base triple with **1** specifically bound through two hydrogen bonds to the CG base pair.

In agreement with the proposed binding mode and its position opposite the guanine carbonyl functionality, the αH vinyl resonance of **1** undergoes a downfield shift when titrating the nucleoside analog (2.45 mM) with a 1:1 mixture of the TIPS-protected cytidine and guanosine (0–12.92 mM) in CD_2Cl_2. From a least-squares fit of these data with a non-linear equation



Scheme 1. *Reagents and conditions*: (i) NHS, DCI, DMF; (ii) NH_3, 90%; (iii) 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-β-D-ribofuranose, K-nonaflate, HMDS, TCS, acetonitrile, 37%.

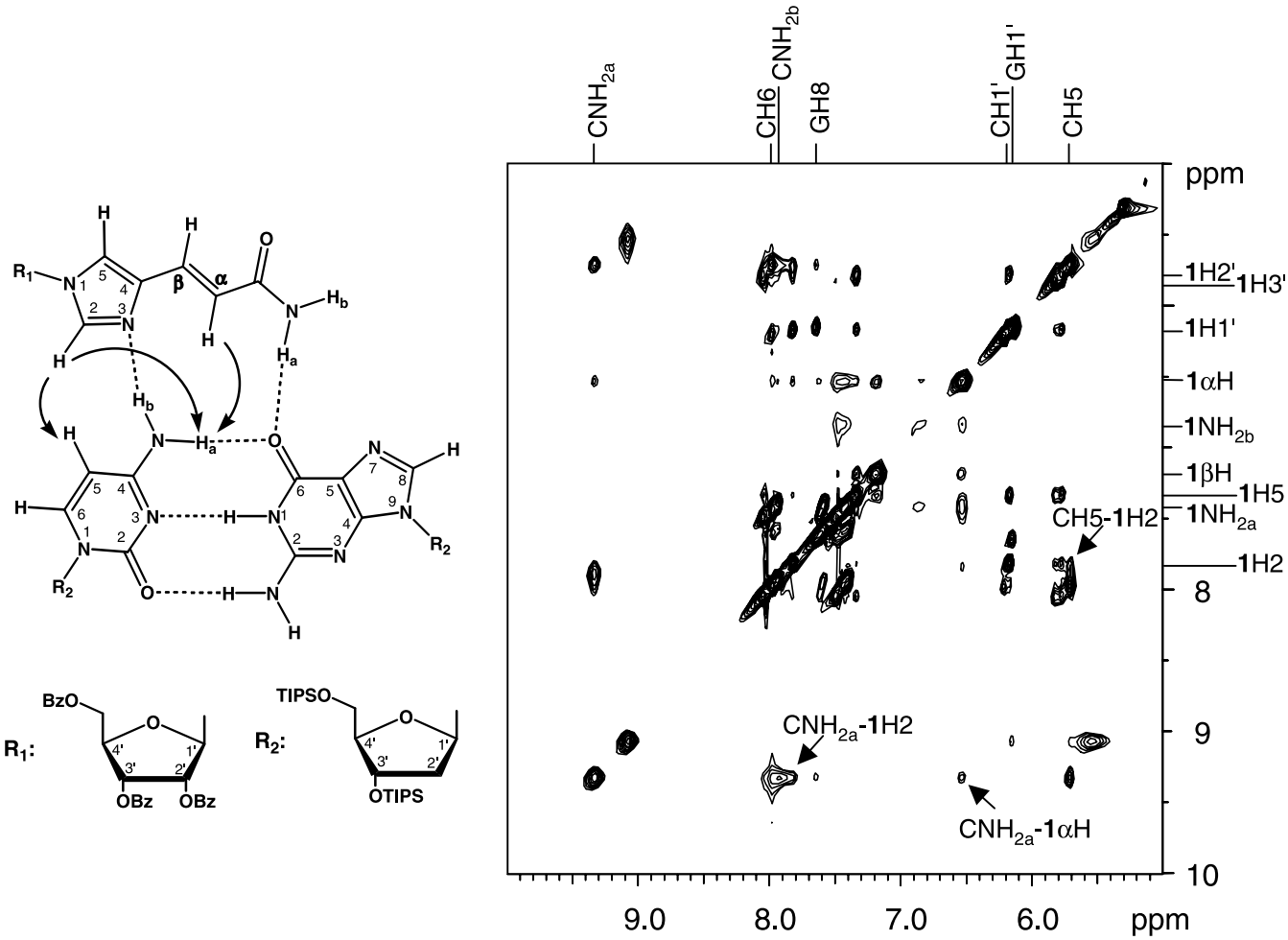


Figure 2. Portion of a 2D NOE spectrum of a 1:1:1 mixture of **1** and TIPS-protected cytidine and guanosine in CD_2Cl_2 at 213 K (right). Resonance assignments for CG and **1** are given at the top and to the right, respectively. Intermolecular cross-peaks found between CG and **1** are indicated in the structure of the complex (left).

describing a simple 1:1 association model, an association constant of 37 (±5) M^{-1} could be determined at 293 K. Given an association constant of 70 M^{-1} for the adenosine–uridine base pair forming two cooperative cyclic hydrogen bonds for each of four possible configurations under comparable conditions,[9] this value is within the expected range for a single complex involving only two weakly coupled hydrogen bonds. In contrast, self-association of **1** was found to be negligible (K_{ass} ~1 M^{-1} at 293 K).

In summary, novel urocanamide derived nucleosides can be prepared by a simple synthetic scheme. Their specific binding to a CG Watson–Crick base pair as confirmed by 2D NOE measurements makes them promising tools for targeting such base pairs in nucleic acids. Also, containing only one imidazole heterocycle should make them less prone to intercalation when bound to double-helical DNA, a problem often encountered with nucleobases having more extended ring systems.[10] In a next step, structural modifications at the imidazole side chain will be introduced and compared with respect to their specificity and affinity towards the base pair. Such detailed studies on the free nucleosides involving the determination of association constants as well as enthalpic and entropic contributions to the free energy of base triple formation will be expected to yield more information about the predominant factors involved in the hydrogen bond mediated complex formation.

References

1. Thuong, N. T.; Hélène, C. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 666–690.
2. (a) Huang, C.-Y.; Cushman, C. D.; Miller, P. S. *J. Org. Chem.* **1993**, *58*, 5048–5049; (b) Sasaki, S.; Nakashima, S.; Nagatsugi, F.; Tanaka, Y.; Hisatome, M.; Maeda, M. *Tetrahedron Lett.* **1995**, *36*, 9521–9524; (c) Zimmerman, S. C.; Schmitt, P. *J. Am. Chem. Soc.* **1995**, *117*, 10769–10770; (d) Huang, C.-Y.; Bi, G.; Miller, P. S. *Nucleic Acids Res.* **1996**, *24*, 2606–2613; (e) Lehmann, T. E.; Greenberg, W. A.; Liberles, D. A.; Wada, C. K.; Dervan, P. B. *Helv. Chim. Acta* **1997**, *80*, 2002–2022; (f) Lecubin, F.; Benhida, R.; Fourrey, J.-L.; Sun, J.-S. *Tetrahedron Lett.* **1999**, *40*, 8085–8088; (g) Lengeler, D.; Weisz, K. *Tetrahedron Lett.* **2001**, *42*, 1479–1481.

3. (a) Durant, G. J.; Ganellin, C. R.; Hills, D. W.; Miles, P. D.; Parsons, M. E.; Pepper, E. S.; White, G. R. *J. Med. Chem.* **1985**, *28*, 1414–1422; (b) Sellier, C.; Buschauer, A.; Elz, S.; Schunack, W. *Liebigs Ann. Chem.* **1992**, 317–323; (c) Stark, H.; Purand, K.; Hüls, A.; Ligneau, X.; Garbarg, M.; Schwartz, J.-C.; Schunack, W. *J. Med. Chem.* **1996**, *39*, 1220–1226.
4. (a) Kazimierczuk, Z.; Seela, F. *Liebigs Ann. Chem.* **1991**, 695–697; (b) Griffin, L. C.; Kiessling, L. L.; Beal, P. A.; Gillespie, P.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 7976–7982; (c) Durland, R. H.; Rao, T. S.; Bodepudi, V.; Seth, D. M.; Jayaraman, K.; Revankar, G. R. *Nucleic Acids Res.* **1995**, *23*, 647–653.
5. (a) Vorbrüggen, H.; Bennua, B. *Tetrahedron Lett.* **1978**, 1339–1342; (b) Vorbrüggen, H.; Bennua, B. *Chem. Ber.* **1981**, *114*, 1279–1286.
6. Matthews, H. R.; Rapoport, H. *J. Am. Chem. Soc.* **1973**, *95*, 2297–2303.
7. Analytical and spectral data for **1**: ^{1}H NMR (500 MHz, 288 K, CD_2Cl_2): δ (ppm) = 4.67 (m, 1H; H5″), 4.84 (m, 2H; H4′, H5′), 5.53 (br, 2H; NH_2), 5.80 (t, J = 5.4 Hz, 1H; H2′), 5.86 (dd, J = 5.4 Hz, J = 4.5 Hz, 1H; H3′), 6.11 (d, J = 5.4 Hz, 1H; H1′), 6.54 (d, J = 15.3 Hz, 1H; C=C–H), 7.30 (d, J = 1.3 Hz, 1H; ImH), 7.30 (d, J = 15.3 Hz, 1H; C=C–H), 7.40–8.11 (m, 16H; ArH, ImH). MS (FAB$^+$) calcd for $C_{32}H_{28}N_3O_8$: 582.1876; found: 582.1892.
8. Due to the large association constant even at ambient temperatures ($K_{ass} \sim 10^5$ M^{-1}), a 1:1 mixture of C and G can be treated as a fully associated CG Watson–Crick base pair.
9. Dunger, A.; Limbach, H.-H.; Weisz, K. *J. Am. Chem. Soc.* **2000**, *122*, 10109–10114.
10. Koshlap, K. M.; Gillespie, P.; Dervan, P. B.; Feigon, J. *J. Am. Chem. Soc.* **1993**, *115*, 7908–7909.