JOC*Note*

Binding of Imidazole-Derived Nucleosides to a CG Base Pair

Maria G. M. Purwanto and Klaus Weisz*

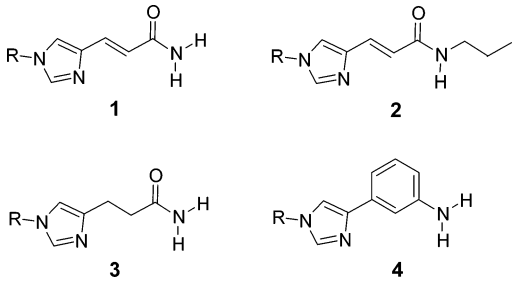
Institut für Chemie und Biochemie, Ernst-Moritz-Arndt-Universität Greifswald, Soldmannstrasse 16, D-17489 Greifswald, Germany

weisz@uni-greifswald.de

Received October 30, 2003

Abstract: Novel imidazole nucleosides with substituents of different flexibility were studied for their binding to a CG Watson−Crick base pair by ^{1}H NMR spectroscopy in an aprotic solvent. Thermodynamic data as determined by titration experiments at different temperatures reveal the influence of the substituent on the enthalpy and entropy of complex formation and thus on the strength of binding.

DNA triplex formation through Hoogsteen hydrogen bonds between third-strand bases and the bases of the duplex target is mostly limited to homopurine sequences. However, excluding the effective recognition of pyrimidine interruptions within a purine tract of the duplex generally results in triple helices of low stability, thus restricting potential applications. To overcome these limitations, various nucleobase analogues have been designed in the past as potential ligands for any given base pair, but the reliable prediction of their interactions within a triple helix has not yet been achieved.[1,2] Recently, we have reported the specific binding of a novel urocanamide derived nucleoside **1** toward a CG Watson−Crick base pair at its major groove side in an organic solvent.[3] Based on these results, we now describe binding studies involving different structural variants **1**−**4**.



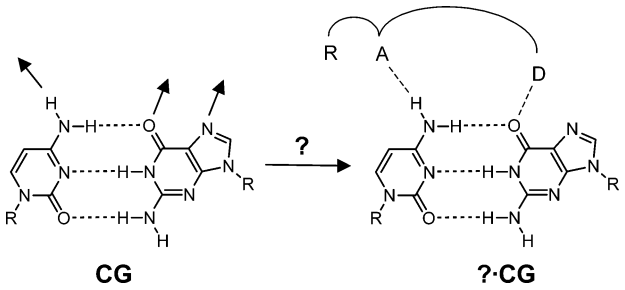
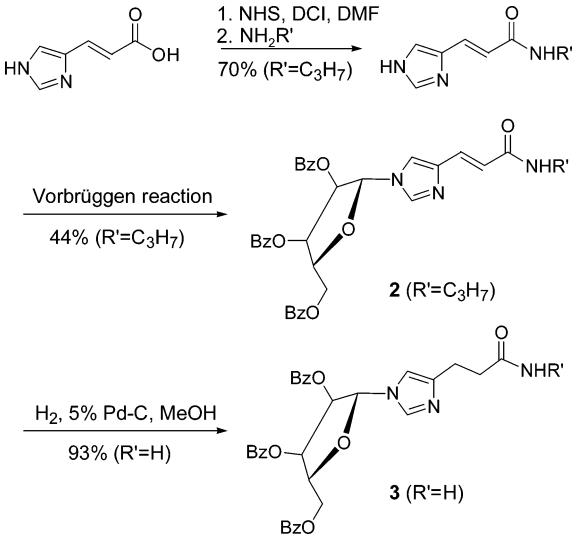


FIGURE 1. Hydrogen bond donor and acceptor sites of a CG base pair and binding of a potential ligand with a complementary acceptor (A)−donor (D) motif; R = *O*-protected ribose sugar.

SCHEME 1. Synthesis of 2 and 3



All the nucleobases are expected to facilitate the formation of two hydrogen bonds to a CG base pair, i.e., to NH_2 of cytosine and the O6 carbonyl oxygen of guanine (see Figure 1). To understand the effect of conformational flexibility on base triple formation, nucleobases **3** and **4** were designed with increased and decreased side chain flexibility, whereas **2** was additionally dedicated to study the effect of an *N*-propyl substituent at the amide function.

As was described for nucleoside **1**[3], **2** was synthesized starting with *trans*-urocanic acid, which after activation using *N*-hydroxysuccinimide (NHS) and *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride (DCI) reacted with propylamine to afford the corresponding *N*-propylurocanamide (Scheme 1). This was glycosylated to the *O*-protected nucleoside with 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-*β*-D-ribofuranose under Vorbrüggen conditions. Compound **1** was smoothly converted to **3** by a Pd-mediated reduction of the double bond which did not affect the sugar or the imidazole moiety. Compound **4** was synthesized based on the Bredereck imidazole synthesis[4] recently also employed by Griffin et al.[5] This procedure affords 2-unsubstituted imidazoles from an α-hydroxy or an α-halogenated ketone (in this case *ω*-bromo-*m*-nitroacetophenone) by using formamide as a

* Corresponding author.

(1) (a) Huang, C.-Y.; Miller, P. S. *J. Am. Chem. Soc.* **1993**, *115*, 10456−10457. (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) Lehmann, T. E.; Greenberg, W. A.; Liberles, D. A.; Wada, C. K.; Dervan, P. B. *Helv. Chim. Acta* **1997**, *80*, 2002−2022. (e) Lecubin, F.; Benhida, R.; Fourrey, J.-L.; Sun, J.-S. *Tetrahedron Lett.* **1999**, *40*, 8085−8088. (f) Lengeler, D.; Weisz, K. *Tetrahedron Lett.* **2001**, *42*, 1479−1481. (g) Guianvarc'h, D.; Benhida, R.; Fourrey, J.-L.; Maurisse, R.; Sun, J.-S. *J. Chem. Soc., Chem. Commun.* **2001**, 1814−1815.

(2) For reviews, see: (a) Doronina, S. O.; Behr, J.-P. *Chem. Soc. Rev.* **1997**, 63−71. (b) Gowers, D. M.; Fox, K. R. *Nucleic Acids Res.* **1999**, *27*, 1569−1577. (c) Purwanto, M. G. M.; Weisz, K. *Curr. Org. Chem.* **2003**, *7*, 427−446.

(3) Purwanto, M. G. M.; Lengeler, D.; Weisz, K. *Tetrahedron Lett.* **2002**, *43*, 61−64.

10.1021/jo035597q CCC: $27.50
Published on Web 12/05/2003

reagent as well as a solvent and a source of ammonia. A Vorbrüggen reaction and a subsequent reduction of the nitro group with Raney nickel and hydrazine as hydrogen source yielded the amino-substituted nucleoside.

By employing concentration- and temperature-dependent ^{1}H NMR measurements, the thermodynamics of binding for each nucleoside analogue toward a CG base pair was examined in methylene chloride and analyzed in terms of enthalpic and entropic contributions. As a first step of our binding studies, we have examined the self-association of the nucleoside analogues through concentration dependent ^{1}H NMR chemical shifts. None of them was found to significantly dimerize ($K_{assoc} \leq 7$ M^{-1} for $T \geq 200$ K) under the given experimental conditions, prerequisite for the reliability of using a simple 1:1 association model in the case of CG binding. Also, the strong binding between C and G in apolar solvents ($K_{assoc} \sim 10^5$ M^{-1}) allowed a 1:1 C:G mixture to be treated as a single species.[1c,f] Correspondingly, NMR measurements were performed in CD_2Cl_2 by titrating the nucleoside analogue (~2 mM) with a 1:1 mixture of di-*O*-triisopropylsilyl-2′-deoxycytidine and di-*O*-triisopropylsilyl-2′-deoxyguanosine. A representative titration experiment is shown for **3** in Figure 2a. Progressive downfield shifts of the hydrogen-bonded NH protons of the analogues were subjected to a nonlinear regression analysis for a 1:1 complex to give an association constant between the nucleoside analogue and the CG base pair (Table 1).[6] These results were further confirmed by reverse titrations observing the non-Watson−Crick bound cytosine amino resonance (not shown).[7]

A van't Hoff analysis of the temperature-dependent association constants as shown in Figure 2b for the **3·CG** asociation yielded enthalpies and entropies of association. As can be seen from Table 1, a similar heat of complex formation $-\Delta H$ was obtained for **1** and **3** complexes, i.e., 14−15 kJ/mol. With the formation of two hydrogen bonds this gives about 7 kJ/mol per hydrogen bond. A smaller value obtained for **4** (4.4 kJ/mol per hydrogen bond) can be attributed to the amino group being a weaker hydrogen bond donor compared to an amide function but also suggests that the rigidity of the base hampers conformational readjustments to optimize the two hydrogen bond contacts.[8]

Overall, the values determined are considerably smaller than the heat of formation of adenine-uracil (AU) association measured in chloroform by Kyogoku et al. (13 kJ/mol per hydrogen bond).[9] This might be attributed to the cooperative nature of the two formed resonance assisted hydrogen bonds and an optimal geometric fit between the monomers in the AU base pair. It is also important to note that the enthalpy determined here not only reflects the formation of hydrogen bonds but also includes the enthalpy of solvation, i.e., the entire enthalpy difference between solvated complex and solvated

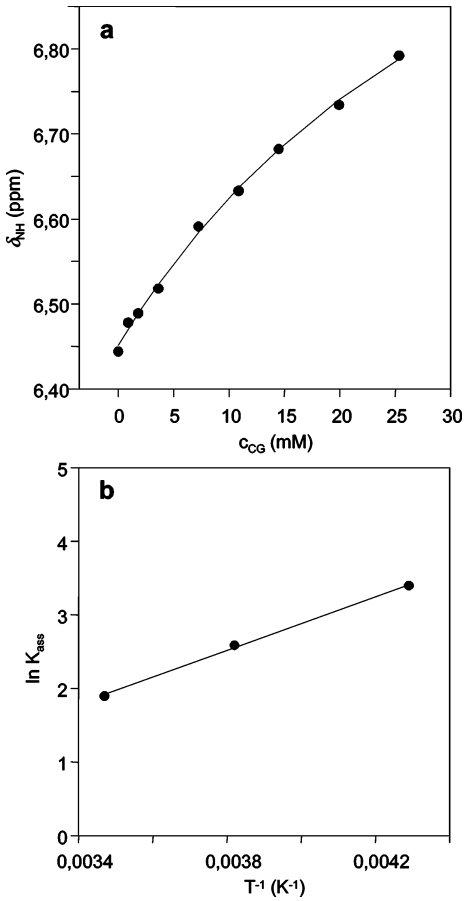


FIGURE 2. (a) NH proton chemical shift of **3** as a function of the CG concentration in CD_2Cl_2 at 233 K; (b) van't Hoff plot of the association constant K_{assoc} for **3**·CG base triple formation in CD_2Cl_2. Lines represent the least-squares fit.

TABLE 1. Summary of K_{assoc}, ΔH, and ΔS for the Association of Nucleoside Analogues 1−4 with a CG Base Pair

	T (K)	K_{assoc} (M^{-1})[a]	ΔH (kJ/mol)	ΔS (J/K·mol)
1	226	212	−14.1 ± 1	−17.9 ± 2
	257	88		
	283	43		
	296	38		
2	207	84	*b*	*b*
	220	62		
	229	66		
	254	31		
3	233	30	−14.7 ± 2	−35.0 ± 4
	262	13		
	288	7		
4	212	28	−8.7 ± 2	−13.5 ± 4
	219	23		
	242	15		

[a] Estimated error about 10%. [b] Not determined due to nonlinearity of van't Hoff plot.

(4) (a) Bredereck, H.; Theilig, G. *Chem. Ber.* **1953**, *86*, 88−96. (b) Theilig, G. *Chem. Ber.* **1953**, *86*, 96−109.

(5) Griffin, L. C.; Kiessling, L. L.; Beal, P. A.; Gillespie, P.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 7976−7982.

(6) For **1** the αH vinyl resonance, which undergoes a downfield shift upon titration with CG in agreement with the proposed binding mode, was used for the determination of K_{assoc}.

(7) Because of severe signal overlap in the spectra, we used specifically ^{15}N amino labeled cytidine for the reverse titration experiments and an ^{1}H−^{15}N filter as implemented in a one-dimensional HMQC experiment for cytosine amino proton observation.

(8) The weak NHO hydrogen bond formed by **4** is also manifested in the small **4** amino downfield shift upon complex formation, i.e., the average limiting chemical shift difference between monomer and complex as obtained by curve fitting amounts to only 0.18 ppm as compared to 0.78 ppm for the **3** amide resonance.

(9) Kyogoku, Y.; Lord, R. C.; Rich, A. *J. Am. Chem. Soc.* **1967**, *89*, 496−504.

monomers. However, assuming a small difference between the enthalpy of solvation for the dimer and monomers, $-\Delta H$ might reasonably well reflect the hydrogen bond energy.

As expected, the entropy upon complex formation $-\Delta S$ closely follows the conformational flexibility of the base analogue, ranging from −13.5 J/K·mol for the most rigid nucleoside analogue **4** to −35.0 J/K·mol for the nucleoside **3** with the most flexible side chain. Consequently, smaller association constants for **4** and **3** compared to **1** arise from less favorable enthalpic and entropic contributions to the free energy of association, respectively.

For the *N*-alkylated analogue **2**, values of K_{assoc} are smaller compared to **1** over the temperature range studied. Moreover, a nonlinearity of the van't Hoff plot precluded the determination of reliable enthalpy and entropy values for the **2**·CG association. Although the *Z*-conformation **2a** is expected to predominate in solution, an additional temperature dependent *Z*-*E* conformational equilibrium through the hindered rotation about the C−N amide bond might attribute to these observations (Figure 3).[10] Clearly, only the **2a** conformation can selectively bind via two hydrogen bonds to the CG base pair in the proposed binding mode.

In summary, less favorable entropic and enthalpic contributions to the hydrogen bond formation for the more flexible **3** and the more rigid **4** imidazole derivative significantly compromise complexation with a CG base pair, respectively. Nucleoside **1** shows the highest affinity toward a CG base pair and seems to be a promising candidate for its incorporation into a triplex-forming oligonucleotide targeting a double-helical DNA with a CG interruption. Corresponding experiments on triplex formation will be reported in due course.

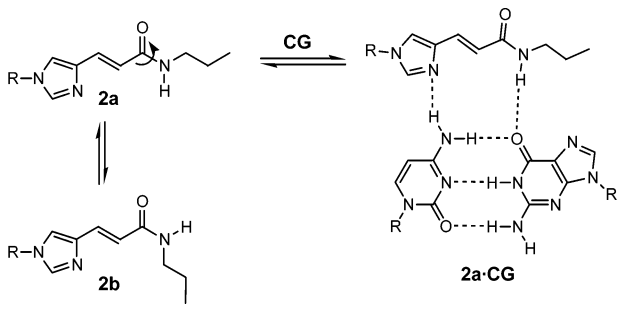


FIGURE 3. Possible equilibria in a **2** + **CG** mixture. R = *O*-protected ribose sugar.

Acknowledgment. We thank the Fonds der Chemischen Industrie for financial support.

Supporting Information Available: Analytical and NMR spectroscopic data including ^{1}H and ^{13}C NMR spectra for compounds **2**−**4**. This material is available free of charge via the Internet at http://pubs.acs.org.

JO035597Q

(10) Similar conformational equilibria have also been discussed by Zimmerman et al. in the case of ligands with *N*-alkylureido substituents. Mertz, E.; Mattei, S.; Zimmerman, S. C. *Org. Lett.* **2000**, *2*, 2931−2934.