Model studies towards carbohydrate–base pair recognition. Relevance of hydrogen–bonding cooperativity

Manuela López de la Paz,[a] Carlos González[b] and Cristina Vicent*[a]

[a] *Departamento de Química Orgánica Biológica, Instituto de Química Orgánica, CSIC, Juan de la Cierva 3, E-28006 Madrid, Spain. E-mail: iqocv18@iqog.csic.es*

[b] *Instituto de Estructura de la Materia, CSIC, Serrano 119, 28006 Madrid, Spain*

Received (in Cambridge, UK) 16th December 1999, Accepted 3rd February 2000

The importance of intramolecular OH···OH hydrogen-bonds (H-bonds) in the effective molecular recognition of carbohydrates is highlighted; specifically, the 1,3-*cis*-diaxial H-bonded OH groups of 1 are shown to provide an efficient binding motif for bidentate coordination of the amino-carbonyl Hoogsteen site of the CG base-pair through the formation of two cooperative intermolecular H-bonds; this result suggests that intramolecularly H-bonded carbohydrate OH groups may be considered as multidentate units able to H-bond cooperatively.

Deoxygenated oligosaccharides which are present in natural products[1] and aminoglycoside antibiotics[2] are known to directly interact with a number of DNA and RNA sequences, respectively. However there is limited structural information on the molecular basis of such a saccharide–nucleic acid recognition in solution.[3]

The design of many low molecular weight nucleic acid ligands has been based on hydrogen-bonding (H-bonding) recognition of the Hoogsteen sites of the B-DNA grooves.[4] One of the most important characteristics of multiple H-bonded complexes is the non-additivity of the H-bonds therein; this property has given rise to the concept of cooperativity.[5] As part of a general project to study H-bonding cooperativity and its implications in the molecular recognition of carbohydrates, we present here our initial effort to use H-bonding cooperativity to efficiently bind sugars in the grooves of B-DNA.

Carbohydrate 1,2- and 1,3-diol motifs are present in many DNA and RNA binders.[1] We have previously demonstrated that the hydroxy groups of the 1,3-*cis*-diaxial diol **1** are intramolecularly H-bonded (OH-2→OH-4); the presence of this H-bond polarizes the σ O-H bonds and enhances the donor ability of OH-4 and acceptor ability of OH-2 [Fig. 1(a)].[6]

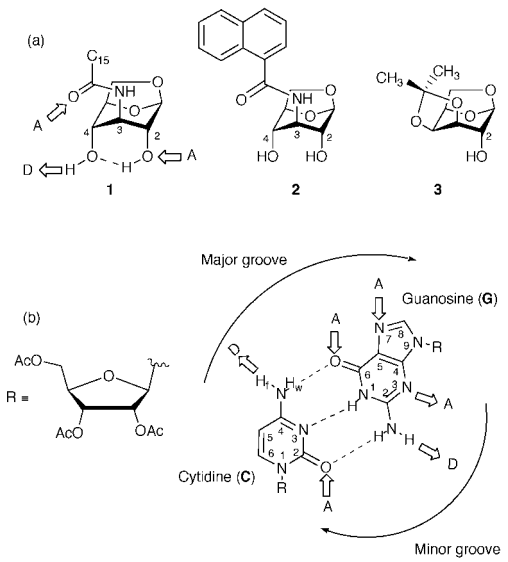

Fig. 1 Carbohydrate-derivatives and **CG** base-pair used in the binding studies.

Molecular modelling studies of the carbohydrate-derivative **1** and the cytidine–guanosine (**CG**) base-pair indicate that the hydroxy groups of **1** are suitably positioned to bridge both amino-carbonyl Hoogsteen binding sites of **CG** in a bidentate fashion [Fig. 1(b)]. A sugar–**CG** complex† could potentially be stabilized by two cooperative intermolecular H-bonds (Fig. 2); this would provide the first example of H-bonding cooperativity in carbohydrate-base pair recognition.

For the reasons outlined above, the binding of **1** to **CG** has been investigated.‡ Binding studies were performed by titrating **1** (0.08 mM) with an equimolar mixture of tri-*O*-acetylguanosine (**G**) and tri-*O*-acetylcytidine (**C**) (2 mM).[7]§ Under these experimental conditions only a 1 : 1 complex of the diol **1** and **CG** could be expected. The titration data were fitted to a 1 : 1 binding model, taking into account the dimerization of **CG** (Table 1).[8]¶ The measured association constant‖ for the **1–CG** complex was 1491 M^{-1}. The binding of the aromatic diol **2** to **CG** was also studied in the same way with the expectation of observing induced chemical shifts of the naphthyl proton resonances and thereby obtaining more data for model fitting. $\Delta G°$ values for the **1–CG** (-4.4 ± 0.1 kcal mol^{-1}) and **2–CG** (-4.2 ± 0.1 kcal mol^{-1}) complexes were greater than expected for a complex stabilized by a single H-bond, or two isolated H-bonds. This result implies the interplay of cooperative H-bonding in **1–CG** and **2–CG** recognition.

To quantify the effect of H-bonding cooperativity on the stabilization of the **1–CG** complex, it is necessary to know the number of intermolecular H-bonds formed between **1** and **CG**. ^{1}H NMR variable temperature experiments, NOESY, and deuterium exchange experiments were carried out with the aim of determining structural information.

The hydroxy proton resonances of **1** were deshielded upon complexation with **CG**, which indicates that the binding process is mediated by H-bonding of both OH groups of the carbohydrate (Table 2). In contrast, the amide proton (NH-3) of diol **1**

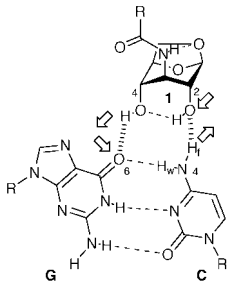

Fig. 2 Schematic representation of the complex formed by the carbohydrate **1** and the **CG** base-pair.

Table 1 Stability parameters of the interaction between the **CG** base-pair and the carbohydrate derivatives **1**–**3** (299 K, $CDCl_3$)

Compound	**1**	**2**	**3**
K_a/M^{-1}	1491	1091	6.7
$\Delta G°$/kcal mol^{-1}	-4.4 ± 0.1	-4.2 ± 0.1	-1.1 ± 0.1

Table 2 $\Delta\delta/\Delta T$ of the exchangeable resonances of **1** in the free and bound state to the **CG** base-pair[a]

Resonance of **1**	$\Delta\delta/\Delta T$ (**1**)[b]/ppb K^{-1}	$\Delta\delta/\Delta T$ (**1–CG**)[c]/ppb K^{-1}
OH-4	−5.2	−13.2
OH-2	−2.8	−6.8
NH-3	−2.9	−4.5

[a] Measured between 293–318 K. [b] [**1**] = 1.1 × 10^{-4} M. [c] [**1**] = 1.1 × 10^{-4} M; [**CG**] = 1.5 × 10^{-3} M (**1**:**CG** = 1:13).

showed minimal displacement on complexation. According to previous reports,[6a,9] comparison of $\Delta\delta/\Delta T$ of the OH resonances of the free ligand **1** and the **1–CG** complex also implies that both OH groups are involved in intermolecular H-bonds (Table 2).

NOESY spectra of mixed samples of **1** and **CG** in different ratios** revealed cross peaks between the OH proton resonances of **1** and the exchangeable protons of **CG**; these could not unambiguously be attributed to chemical exchange or intermolecular NOEs. However, the same experiments showed that the **C**(N^4–H)$_f$ proton resonates at lower field in the presence of a higher concentration of diol **1**, which further suggests its involvement in H-bonding with **1**.

Additional evidence for the preferred complexation site of **CG** was obtained by deuterium exchange experiments. An equal quantity of deuterated diol **1** (**1**–D) was added to separate samples of **CG** (experiment A: **CG** 2 mM) and **1**:**CG** in a 2:1 ratio (experiment B: **1**:**CG** 4 mM:2 mM), to facilitate the formation of the hypothetical 1:2 **CG**:**1** complex. In each case a control experiment was carried out (experiment A: **CG** 2 mM; experiment B: **1**:**CG** 4 mM:2 mM). ^{1}H-NMR spectra were acquired at t = 0 and 17 days. Both samples containing **1**–D showed complete H–D exchange of the **C** amino protons after 17 days, while the amino protons of **G** were only partially deuterated after the same period of time. In the control experiments a small and comparable decrease in the signal intensity of all the exchangeable **CG** protons was observed. From this result we infer that on average the **C**-amino group is in contact with the deuterated hydroxy groups of **1**–D for longer than the amino protons of **G**, and that a 1:1 **1–CG** complex is favoured over a 2:1 **1–CG** complex, which is in agreement with the Job plot determined stoichiometry.¶

Molecular modelling[11] of the **1–CG** complex supported the results of our ^{1}H-NMR experiments and confirmed that the carbonyl group **G**(C^6-O) is the H-bond acceptor best located to form a second H-bond to OH-4 of **1** .†† The calculated **1–CG** structure also showed that the non-exchangeable protons of the **1–CG** complex are very distant from each other, which could explain why only ambiguous intermolecular NOEs were detected.

To quantify the influence of the intramolecular OH-2→OH-4 H-bond on the formation of intermolecular H-bonds between **1** and **CG**, the complexation of monoalcohol **3** with **CG** was studied.‡‡ The K_a for the **3–CG** complex is 7 M^{-1} (Table 1); the formation of one H-bond between the monoalcohol **3** and the **CG** base-pair thus corresponds to a $\Delta G°$ value of −1.1 kcal mol^{-1}, while the $\Delta G°$ for two intermolecular H-bonds in the **1–CG** complex is more than four times greater than this value. This demonstrates that the formation of H-bonds between **1** and **CG** is non-additive (cooperative); furthermore the importance of intramolecular OH···OH H-bonds in the effective molecular recognition of carbohydrates is highlighted.

Our work has shown that in the future we may consider the intramolecularly H-bonded OH groups of carbohydrates as multidentate units capable of H-bonding cooperatively. Specifically, the 1,3-*cis*-diaxial relative configuration of carbohydrate OH groups serves in apolar medium as an efficient binding motif for bidentate coordination of the **C**(N^4–H)$_f$/**G**(C^6–O) site of **CG** through formation of two cooperative intermolecular H-bonds.

Financial support by DGES (Grant PB97–0832) and TMR (FMRX–CT98–0231) are acknowledged. M. L. P. is grateful to the Comunidad Autónoma de Madrid for a predoctoral fellowship. We thank Professor C. A. Hunter (University of Sheffield) for kindly providing the fitting program and Dr Joanne Hawley for critical reading of the manuscript.

Notes and references

† **C** (2',3',5'-tri-*O*-acetylcytidine) and **G** (2',3',5'-tri-*O*-acetylguanosine) were purchased (Sigma) and used without further purification. Carbohydrate-derivatives **1**, **2** and **3** were synthesized.[10] **1**–D was prepared by repeatedly dissolving **1** (1.4 mg, 3.5 mmol) in CD_3OD (5 × 0.5 mL) and evaporating to dryness. The deuterated residue (**1**–D) was dried under high vacuum and heated at 40 °C in the presence of P_2O_5 and dissolved in $CDCl_3$ (2 mL) to give a solution of concentration 1.8 mM.

‡ All binding studies were performed at 299 K using freshly prepared solutions in $CDCl_3$ which were always passed through basic alumina and collected over 4 Å molecular sieves prior to use; the alumina and molecular sieves employed were freshly activated by heating at 600 °C under high vacuum. Each experiment was carried out at least two times and $\Delta G°$ values were reproducible within ±0.1 kcal mol^{-1}.

§ The feasibility of this titration experiment relied on the high stability of the **CG** base-pair[7] in chloroform (K_a = 10^4–10^5 M^{-1}). The imino proton **G**(N^1–H) experienced minimal chemical shift displacement (upfield) upon complexation with **1** ($\Delta\delta$ < 0.1 ppm), which is consistent with the **CG** complex remaining intact during the titrations.[8a]

¶ The 1:1 stoichiometry of the complex **1**:**CG** was determined by a Job plot based on the chemical induced shifts of the hydroxy resonances. The **CG** base-pair dimerizes in chloroform.[8a] We have measured a dimerization constant of 55 M^{-1}.

‖ Reverse titration experiments (**CG** *vs.* **1**) were also performed. Fitting of the induced chemical shifts of the C(C^5–H) proton resonance to a 1:1 complexation model gave a value of K_a (1460 M^{-1}) which was in good agreement with the value determined experimentally from the **1** *vs.* **CG** titrations.

** NOESY spectra (500 ms, 278 K, 600 MHz) were recorded for two samples of different **1**:**CG** molar ratio: (i) **1**:**CG** 1:3, [**1**] = 6.7 × 10^{-4} M; [**CG**] = 2 × 10^{-3} M; (ii) **1**:**CG** 3:1, [**1**] = 6 × 10^{-3} M; [**CG**] = 2 × 10^{-3} M.

†† Molecular mechanics calculations were carried out using MM2[11a] (carbohydrates) and AMBER[11b] (nucleosides, CG base-pair and carbohydrate–CG complexes) with the GB/SA solvent model for chloroform.[11c] Molecular modelling of the complex involing the **C**(C=O) and **G**(N–N) of **CG** (minor groove of the base-pair) indicated that such a complex is not stable. This could be explained on the basis of steric hindrance of the acetylated ribose moieties.

‡‡ The use of **3** allowed us to evaluate the effect of a second OH in a 1,3-*cis*-diaxial relative configuration on the energetics of the recognition process between **1** and **CG**.

1 A. Kirschning, A. F. W. Bechthold and J. Rohr, *Bioorganic Chemistry Models and Applications*, ed. J. Rohr, Springer, Heidelberg, 1997, vol. 184, pp. 1–79.
2 U. von Ahsen and H. F. Noller, *Science*, 1993, **260**, 1501; M. L. Zapp, S. Stern and M. R. Green, *Cell*, 1993, **74**, 969; B. Clouet-d'Orval, T. K. Stage and O. C. Uhlenbeck, *Biochemistry*, 1995, **34**, 11 186.
3 K. C. Nicolaou, B. M. Smith, K. Ajito, H. Komatsu, L. Gómez-Paloma and Y. Tor, *J. Am. Chem. Soc.*, 1996, **118**, 2303; L. Jiang, A. K. Suri, R. Fiala and D. J. Patel, *Chem. Biol.*, 1997, **4**, 35.
4 P. E. Nielsen, *Chem. Eur. J.*, 1997, **3**, 505; S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird and P. B. Dervan, *Nature*, 1998, **391**, 468.
5 H. S. Frank and W. Y. Wen, *Discuss. Faraday Soc.*, 1957, **24**, 133; G. A. Jefrey and W. Sainger, *Hydrogen Bonding in Biological Structures*, Springer-Verlag, Berlin, 1991, p. 569.
6 (*a*) M. López de la Paz, J. Jiménez–Barbero and C. Vicent, *Chem. Commun.*, 1998, 465; (*b*) F. J. Luque, J. M. López, M. López de la Paz, C. Vicent and M. Orozco, *J. Phys. Chem. A*, 1998, **102**, 6690.
7 Y. Kyogoku, R. C. Lord and A. Rich, *Biochim. Biophys. Acta*, 1969, **179**, 10.
8 (*a*) S. C. Zimmerman and P. Schmitt, *J. Am. Chem. Soc.*, 1995, **117**, 10 769; (*b*) N. Branda, G. Kurz and J.-M. Lehn, *Chem. Commun.*, 1996, 2443.
9 E. S. Stevens, N. Sugawara, G. M. Bonora and C. Toniolo, *J. Am. Chem. Soc.*, 1980, **102**, 7048.
10 M. López de la Paz, G. Ellis, S. Penadés and C. Vicent, *Tetrahedron Lett.*, 1997, **38**, 1659; C. W. Holzapfel, J. M. Koekemoer and C. F. Marais, *S. Afr. J. Chem.*, 1984, **37**, 19.
11 (*a*) N. L. Allinger, *J. Am. Chem. Soc.*, 1977, **99**, 8127; (*b*) S. J. Weiner, P. A. Kollman, D. T. Nguyen and D. A. Case, *J. Comput. Chem.*, 1986, **7**, 230; (*c*) W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127.

Communication a909878j