Pergamon

Tetrahedron Letters 41 (2000) 9425–9429

TETRAHEDRON LETTERS

Covalently cross-linked Watson–Crick base pair models. Part 2†

Yao-Ling Qiu, Hong-Yu Li, George Topalov and Yoshito Kishi*

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA

Received 25 August 2000; revised 11 September 2000; accepted 12 September 2000

Abstract

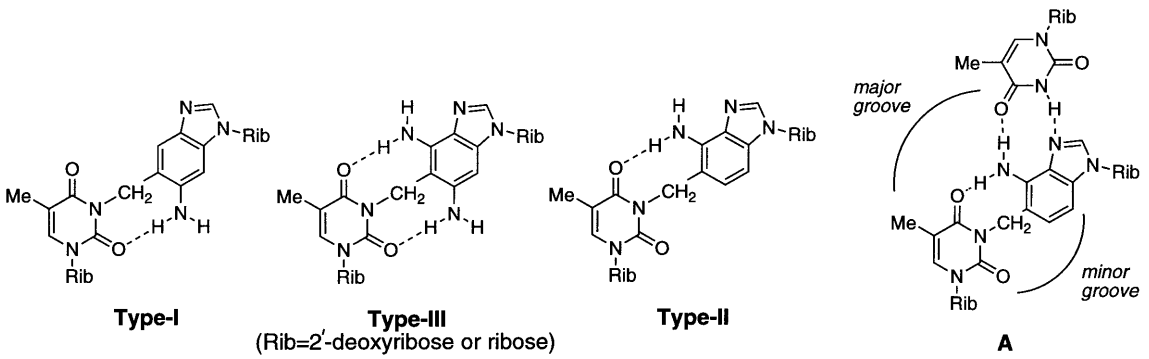
A concise and practical route has been developed for a synthesis of the CH_2-bridged base pairs represented by the type-**II** structure. The structural studies suggest the possibility that these base pairs mimic the molecular architecture of Watson–Crick hydrogen-bonded base pairs.

The concept of covalently linked cross-sections with molecular architecture similar to Watson–Crick hydrogen-bonded base pairs was introduced by Nelson Leonard in the mid-1980s.[1] Since then, several types of covalently linked systems have been developed. However, with the exception of the Leonard system, these systems were generated from preformed double helices as seen in the seminal work by Verdine.[2] Undoubtedly, the Leonard system offers many unique opportunities to address questions on the chemistry of DNA and RNA, but this system has several drawbacks, including difficulty in attempted duplex formation, lack of conformational flexibility between the base pairs, and others.[1] In our view base pairs linked with a CH_2-bridge such as type-**I**–**III** base pairs may be uniquely suited to the chemical exploration of covalently linked nucleosides/nucleotides. In addition to their expected increased chemical stability, these models are expected to adopt only Watson–Crick or reversed Watson–Crick base-pairings but maintain conformational flexibility along the CH_2-bridge. Obviously, there is concern about such models; introduction of a CH_2-bridge does not allow the two bases to be coplanar, and consequently type-**I**–**III** base pairs may be unsuitable as mimics of Watson–Crick base pairs. However, molecular mechanics calculations have suggested that the deformation due to the introduction of the CH_2-bridge to a double-stranded oligomer may be insignificant.[3,4] It should

* Corresponding author. Fax: +1-617-495-5150; e-mail: kishi@chemistry.harvard.edu

† Dedicated to Professor Harry H. Wasserman on the occasion of his 80th birthday.

0040-4039/00/$ - see front matter
PII: S0040-4039(00)01572-0

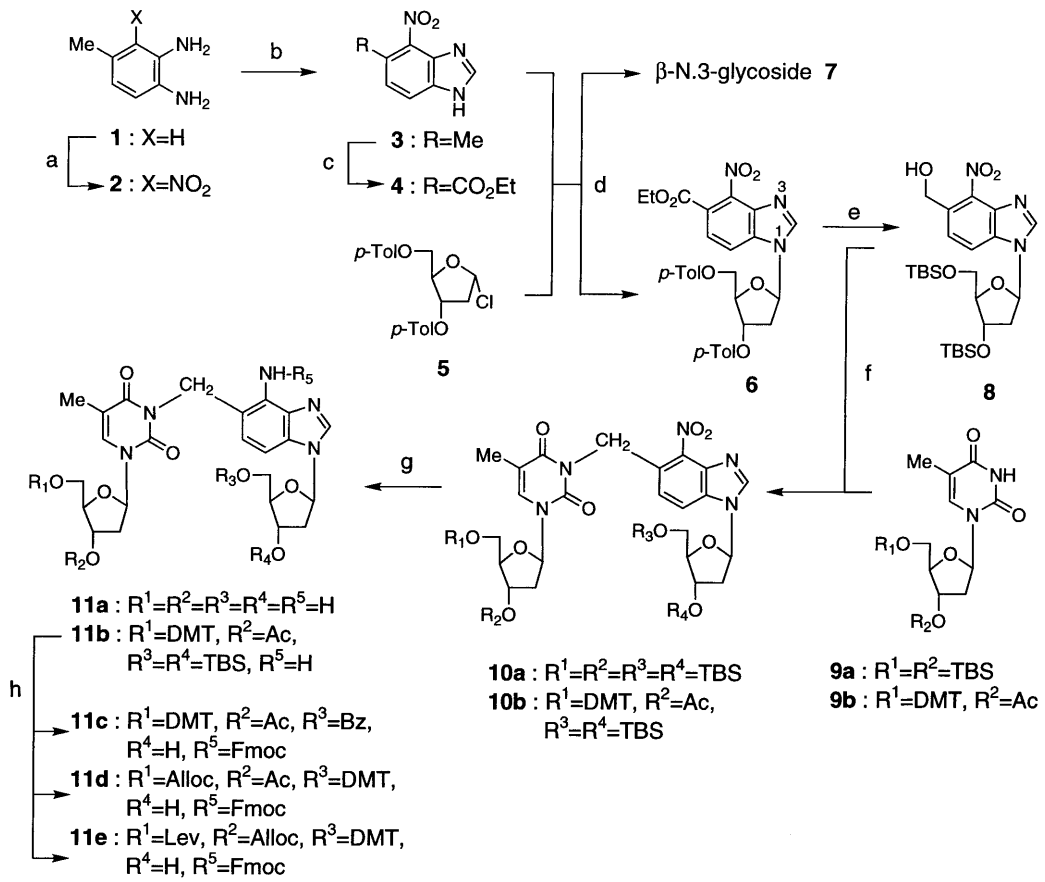


be noted that there is a distinct difference among these three models. Only Watson–Crick or reverse Watson–Crick base-pairings are possible for each model in a primary hydrogen-bonded base-pairing sense, but Hoogsteen triplets such as T–AT and T–GC can be envisioned for type-**II** and **III** models (cf. structure **A**), whereas they are not possible for type-**I** model. We recently reported a concise synthesis and the structural properties of type-**I** base pairs.[3] In this letter, we would like to disclose the synthesis and structural properties of type-**II** base pairs.

Using the procedure developed by Sawicki and Carr,[5] 4-methyl-1,2-phenylenediamine (**1**) was selectively transformed to 3-nitro-4-methyl-1,2-phenylenediamine (**2**), which was then converted to 4-nitro-5-methylbenzimidazole (**3**; mp 220–223°C) (Scheme 1).[6] Permanganate oxidation of **3**, followed by Fischer esterification, gave ethyl 4-nitro-5-benzimidazole carboxylate (**4**; mp >230°C). Following the protocol developed for the type-**I** bases,[3] ethyl 4-nitro-5-benzimidazole carboxylate (**4**) was then transformed to the type-**II** bases. Two observations made in this series are noteworthy. First, the glycosidation of **4** with **5** under phase-transfer conditions[7] gave a 4.6:1 mixture of N1- and N3-glycosides **6** and **7**,[8] compared to a ca. 1:1 mixture in the type-**I** series. Second, as with the previous series, the coupling of **8** with protected forms of thymidine was realized via two different methods (Scheme 1). However, for this series, the coupling efficiency of Method A was significantly superior to Method B. Reduction of the nitro group of **10a**, followed by TBS-deprotection, furnished the CH_2-linked **11a** (mp 128–130°C).

The coupling product **11a** bears one amino and four hydroxyl groups. In order to incorporate a CH_2-bridged base pair into a DNA and/or RNA oligomer, it is necessary to devise a suitable protecting group strategy for these amino and hydroxyl groups. In this context, it is worth noting that, like the type-**I** series, the coupling of **8** is effective with various protected and/or unprotected forms of thymidine, which provides flexibility in the preparation of a monomer with suitable protecting groups. Indeed, the coupling product **10b**, obtained from **8** and **9b** in 87% yield, has been transformed to the CH_2-linked base pairs **11c**–**e** in excellent overall yield.

For the purpose of comparison with the type-**I** series, single crystal X-ray analysis was performed on the CH_2-linked base pair **12**[9] (mp >230°C). The solid phase structure of **12** (Panel B, Fig. 1) compares well with the structure of the corresponding type-**I** base pair **13** (Panel A). Clearly, a hydrogen bond can be seen between the thymidine C4 C=O and the benzimidazole C4 N–H, i.e. the distance between the oxygen and the nitrogen is 2.823 Å, thereby showing that **12** exists in the Watson–Crick conformer (cf. **12a** in Fig. 1) in the solid phase. Thus, like the type-**I** base pair **13**, the type-**II** base pair **12** is nicely superposed on for C–G and A–T base pairs (cf. the superposition depicted in figure 2 of Ref. 3). However, a distinct difference can be seen between the two solid phase structures; on examination of the crystal packing, a Hoogsteen-type



Scheme 1. *Reagents and conditions*. (a) 1. SeO_2, EtOH, reflux, 96%; 2. HNO_3–H_2SO_4, 0°C; 3. HI, HCl, 0°C, 98% over two steps. (b) HCO_2H, reflux, 99%. (c) 1. $KMnO_4$, *t*BuOH–H_2O, 70°C; 2. EtOH, H_2SO_4, 59% over two steps, 10% recovery of **3**. (d) **5**, KOH, 18-crown-6, CH_3CN, rt, 59% (**6**) and 13% (**7**). (e) 1. K_2CO_3, EtOH, rt, 91%; 2. TBSCl, imidazole, DMF, rt, 96%; 3. NaOH, *t*BuOH–H_2O, rt; 4. $ClCO_2Et$, Et_3N, THF, 0°C; 5. $NaBH_4$, EtOH, −78°C, 88% over three steps. (f) Method A: 1. MsCl, Et_3N, CH_2Cl_2, −78°C, 99%; 2. **9a**, K_2CO_3, DMF, 80°C, 80% or **9b**, K_2CO_3, DMF, rt→50°C, 87%; Method B: DEAD, PPh_3, THF, rt, 45% from **9a**, 38% from **9b**. (g) 1. H_2, 10% Pd/C, EtOAc, rt, 99%; 2. TBAF, THF, rt, 80% for **11a**; or H_2, 10% Pd/C, NaOAc, MeOH, rt, 94% for **11b**. (h) 1. FMOCCl, AgOAc, THF, rt, 85%; 2. TBAF, HOAc, THF, rt, 97%; 3. BzCl, Py, 0°C, 92% for **11c**; or 1. FMOCCl, AgOAc, THF, rt, 85%; 2. 3% Cl_3CCO_2H, CH_2Cl_2, rt; 3. AllocCl, Py, CH_2Cl_2, 0°C, 90% over two steps; 4. TBAF, HOAc, THF, rt, 92%; 5. DMTCl, Py, CH_2Cl_2, 93% for **11d**; or 1. 20% NH_3/MeOH, rt; 2. AllocCl, DMAP, CH_2Cl_2, rt, 95% over two steps; 3. FMOCCl, AgOAc, THF, rt, 70%; 4. 3% Cl_3CCO_2H, CH_2Cl_2, rt, 94%; 5. $(Lev)_2O$, Py, rt, 97%; 6. TBAF, HOAc, THF, rt, 94%; 7. DMTCl, Py, CH_2Cl_2, rt, 93% for **11e**

hydrogen bonding network is detected in the type-**II** base pair **12** (Panel C, Fig. 1), whereas it is not seen in the type-**I** base pair **13**.

As in the type-**I** series, the rotational freedom along the CH_2-bridge was studied by temperature-dependent ^1H NMR spectroscopy. However, it should be noted that, due to poor solubility in acetone-d_6, the NMR studies of **12** were conducted in a 9:1 mixture of CD_2Cl_2–CD_3OD. As seen in the Type-**I** base pair series,[3] the proton signals of the CH_2-bridge of **12** were observed as a sharp singlet at room temperature, but were split into an AB quartet ($\Delta\delta = 69.5$ Hz (400 MHz)) at ca. −70°C. The coalescent temperature was found to be around

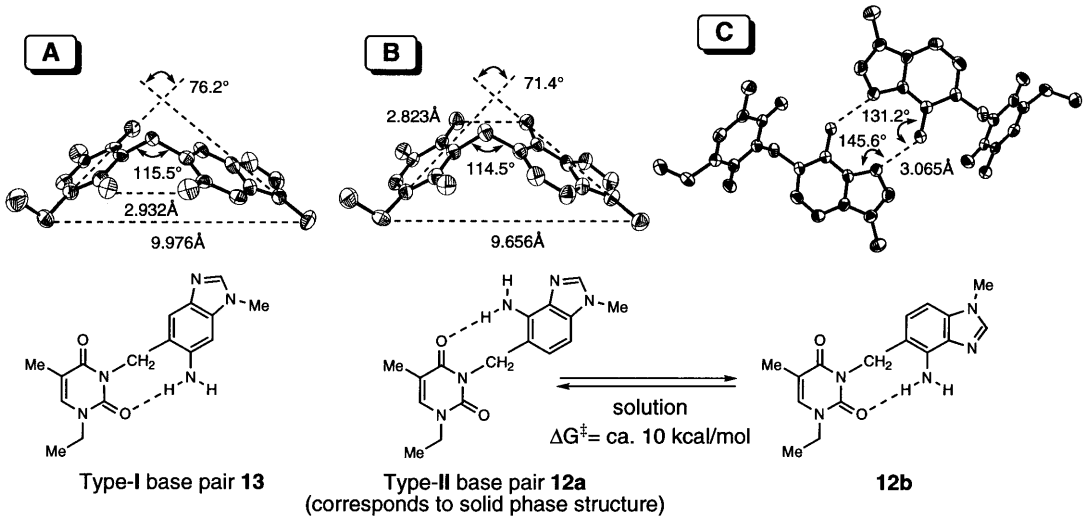


Figure 1.

−60°C, indicating the energy barrier for the rotation to be ca. 10.5 kcal/mol. The same phenomenon was observed for **11a**, with the energy barrier for the rotation being again ca. 11 kcal/mol. Although there is no experimental evidence available at present, it is tempting to suggest that the preferred solution conformer coincides with the solid phase structure which corresponds to the Watson–Crick with base-pairing structure, cf. **12a**, rather than the reverse Watson–Crick base-pairing, cf. **12b**.

In summary, a concise and practical route has been developed for the synthesis of CH_2-bridged base pairs **11a–e** and **12**. Structural studies suggest the possibility that these base pairs mimic the molecular architecture of Watson–Crick hydrogen-bonded base pairs.

Acknowledgements

Financial support from the National Institutes of Health (NS 12108) is gratefully acknowledged.

References

1. Devadas, B.; Leonard, N. J. *J. Am. Chem. Soc.* **1986**, *108*, 5012. For the latest publication on this subject from the Leonard group, see: Bhat, B.; Leonard, N. J.; Robinson, H.; Wang, A. H.-J. *ibid.* **1996**, *118*, 10744.
2. For examples of covalently linked DNA's, see the contributions from Verdine (*J. Am. Chem. Soc.* **1991**, *113*, 4000), Benkovic (*Biochemistry* **1991**, *30*, 788), Mitchell and Kelly (*J. Am. Chem. Soc.* **1991**, *113*, 8994), Lippert (*J. Am. Chem. Soc.* **1992**, *114*, 357), Armstrong (*J. Am. Chem. Soc.* **1992**, *114*, 3144), Hopkins (*J. Am. Chem. Soc.* **1992**, *114*, 4021 and 9240), Seaman and Hurley (*Biochemistry* **1993**, *32*, 12577), and Glick (*J. Am. Chem. Soc.* **1997**, *119*, 1172).
3. For part one, see: Qiao, X.; Kishi, Y. *Angew. Chem., Int. Ed.* **1999**, *38*, 928.
4. We have successfully incorporated both type-**I** and **II** base pairs into several DNA and RNA duplexes. CD and NMR studies have demonstrated that these duplexes mimic the corresponding native duplexes well: Li, H.-Y.; Qiu, Y.-L.; Moyroud, E.; Kishi, Y., unpublished results.

5. Sawicki, E.; Carr, A. *J. Org. Chem.* **1957**, *22*, 503.
6. Direct nitration of 5-methylbenzimidazole is known to yield a ca. 1:1 mixture of **3** and its expected regioisomer: Efros, L. S.; El'tsov, A. V. *Zh. Obshch. Khim. Engl. Transl.* **1958**, *28*, 64.
7. Seela, F.; Bourgeois, W. *Synthesis* **1989**, 912.
8. The structures of **6** and **7** were established by NOESY experiments; the anomeric proton exhibited strong cross-peaks with, in addition to the 4′-proton, the protons at the 2- and 7-positions of benzimidazole ring in **6**, whereas with only the proton at the 2-position in **7**.
9. **12** was prepared via the procedure used for preparation of **11b** (Ref. 3).