Available online at www.sciencedirect.com
SCIENCE DIRECT
Tetrahedron Letters
Tetrahedron Letters 47 (2006) 3849–3852

Synthesis and DNA triplex formation of an oligonucleotide containing an urocanamide

Maria G. M. Purwanto[†] and Klaus Weisz[*]

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

Received 27 February 2006; revised 23 March 2006; accepted 29 March 2006
Available online 27 April 2006

Abstract—An urocanamide nucleoside designed and previously tested as its protected ribose derivative in aprotic solvents for binding a cytosine–guanine (CG) Watson–Crick base pair was successfully incorporated into a triplex forming oligonucleotide. Binding affinity and specificity of this nonnatural nucleoside were studied in a triple helix with duplex targets containing all four possible Watson–Crick base pairs opposite the nucleoside analog in the third strand. UV melting experiments indicate the formation of a well-defined triplex with specific binding of the urocanamide analog to a CG inversion of the homopurine–homopyrimidine target. However, binding affinities in the triplex are weak and much lower when compared to the canonical base triads.

As a carrier of genetic information, double-helical DNA and its reliable sequence-specific recognition is central to such basic life processes like replication and transcription. Because the machinery of living cells ultimately derives from genomic DNA, nucleic acids are attractive targets for the artificial regulation of living systems. Thus, manipulations through specific DNA binding may block the expression of particular genes[1,2] or result in site-directed mutagenesis,[3,4] but may also be used for the mapping of genomic DNA.[5,6] Various ligands have been tested in the past for the sequence-specific recognition of double-helical DNA including duplex-invading PNA,[7,8] minor groove binding polyamides,[9,10] and triple helix forming oligonucleotides (TFOs).[11,12] Pyrimidine TFOs can bind in the major groove of a Watson–Crick duplex by T·A or C^{+}·G Hoogsteen base pairing and are particularly attractive with regard to their easy availability and their potential to result in a nearly perfect structural fit upon binding to duplex DNA with its same chemical subunits to form a triple helix. Unfortunately, third strand oligonucleotides containing only natural bases suffer from a limited recognition code and only homopurine·homopyrimidine tracts within the duplex are effectively recognized by the TFO. In order to overcome these limitations, a large number of base analogs have been developed for the binding to all four possible base pairs in the past, however, success has been limited mostly due to our restricted understanding of the interplay between the various interactions within such a triple-helical system.[13]

Some time ago we have initiated studies aimed at getting more detailed information on base–base interactions in DNA triplexes. As a first step, we have designed and synthesized novel nucleoside analogs based on urocanic acid amides for CG Watson–Crick base pair recognition and tested their hydrogen bond interactions with a free CG pair in aprotic organic solvents.[14,15] ^{1}H NMR experiments in methylene chloride clearly indicated formation of a base triple closely isomorphous to canonical T·AT and C^{+}·GC triads within a regular pyrimidine or parallel triplex motif and with the urocanamide specifically bound through two hydrogen bonds to the CG base pair (Fig. 1). We now expand these model studies at the monomeric level to the formation of triple-helical complexes in an aqueous environment by incorporating the N-alkylated urocanamide nucleoside **U2** into the TFO. To shed more light on the binding under aqueous conditions and to compare with the results of our initial model studies, third strand affinities towards all four possible base pairs in the duplex opposite the novel base analog in the TFO are determined through UV melting experiments.

Keywords: Nucleoside analog; Triple helix; UV melting; Urocanamide.
* Corresponding author. Tel.: +49 3834 864426; fax: +49 3834 864427; e-mail: weisz@uni-greifswald.de
† Present address: Faculty of Biotechnology, Universitas Surabaya, Raya Kali Rungkut, 60292 Surabaya, Indonesia.

0040-4039/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2006.03.180

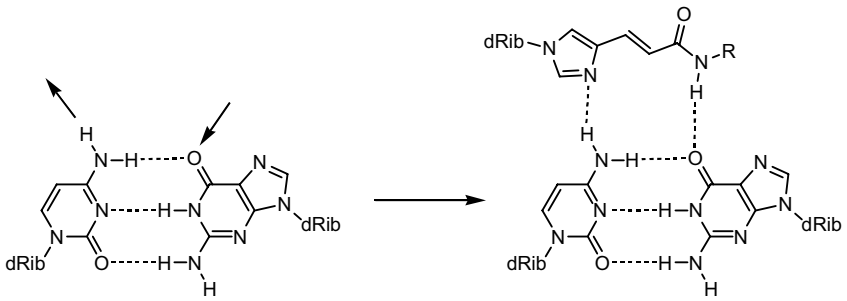
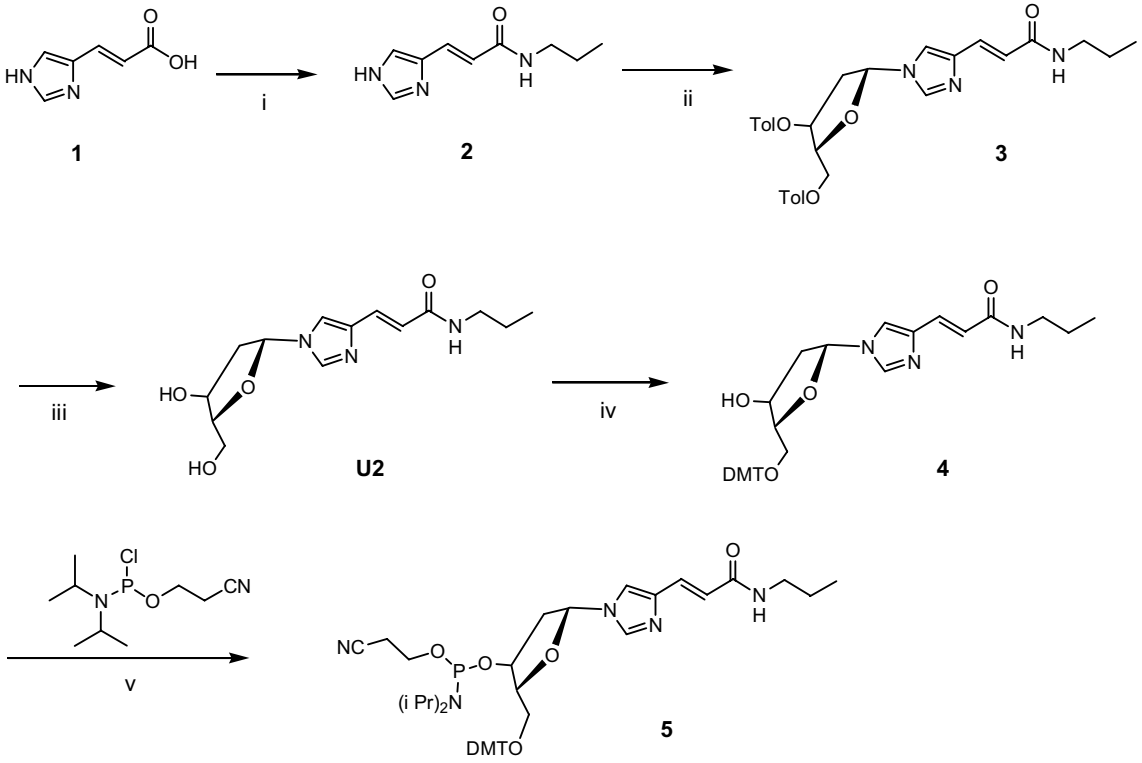


Figure 1. Hydrogen bond mediated recognition of a CG base pair by an urocanamide receptor; arrows denote hydrogen bond acceptor and donor sites, respectively.

The synthetic scheme for the preparation of the 5′-tritylated and 3′-phosphitylated 2′-deoxyribonucleoside **U2**, which was used as synthon for the oligonucleotide synthesis, is outlined in Scheme 1. As described previously,[14,15] N-alkylation of *trans*-urocanic acid **1** with propylamine in the presence of *N*-hydroxysuccinimide (NHS) and *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride (DCI) afforded *N*-propylurocanamide **2**. Glycosylation of **2** with a standard one-pot Vorbrüggen procedure failed but could be achieved by employing the sodium salt of **2**, which was prepared in situ with sodium hydride and reacted with 2-deoxy-3,5-di-*O*-*p*-toluoyl-α-D-ribofuranosyl chloride in acetonitrile.[16] The latter was prepared as previously described[17] with the α-anomer as the major product and subjected to the reaction with the imidazole base without further purification to afford the 2′-deoxynucleoside **3** with an anomeric α:β molar ratio of 1:2. The two anomers were unambiguously assigned based on 2D NOE intensities of sugar protons. After deprotection under alkaline conditions followed by 5′-tritylation, the β-anomer **4** was separated from its α-isomer by HPLC (SiO_2, hexane/ethylacetate 10:1).[18] After conversion to its phosphoramidite, **5** was used as synthon for the synthesis of the oligonucleotide **OL15** using the standard β-cyanoethyl phosphoramidite method. The 15-mer oligonucleotide containing the nucleoside analog was finally characterized by ESI mass spectrometry.

In order to examine the affinity and selectivity toward a duplex target, equimolar amounts of four AT-rich double-helical oligonucleotides were mixed with the **U2** containing TFO **OL15**, which was designed to bind in a parallel motif to the complementary duplex. Sequences



Scheme 1. Reagents and conditions: (i) NHS, DCI, DMF, $CH_3CH_2CH_2NH_2$, 70%; (ii) 2-deoxy-3,5-di-*O*-*p*-toluoyl-α-D-ribofuranosyl chloride, NaH, CH_3CN, 37%; (iii) 1% NaOH in MeOH, 76%; (iv) DMTrCl, Et_3N, pyridine, 40% (β-anomer); (v) *i*-Pr_2EtN, CH_2Cl_2, 90%.

of the 15-mer double-helical targets differ in their Watson–Crick base pair XY located opposite the nonnatural **U2** analog in a triplex and are given below:

OL15: 5'- C T T C **U2** T T T T T C T T T T -3'
5'- G A A G **X** A A A A A G A A A A -3'
3'- C T T C **Y** T T T T T C T T T T -5'

UV melting experiments at pH 6.0 of the 1:1 mixtures of duplex and the TFO are shown in Figure 2. Each mixture displays a sigmoidal high-temperature transition, which corresponds to the duplex to single strand transition. In addition, hyperchromicity effects at lower temperatures indicate the dissociation of the TFO from the duplex. Note, however, that only the triplex with a **U2**·CG base triad exhibits a sharp and cooperative triplex–duplex transition at 21 °C whereas triplexes with **U2**·AT and **U2**·GC triads only show a rather broad and featureless, noncooperative transition at lower temperatures. Compared to the **U2**·CG triplex, a cooperative yet much broader low-temperature transition is observed in case of the **U2**·TA triple-helical nucleic acid. These results point to significant structural heterogeneity of the **U2**·AT and **U2**·GC triplexes but to a well-defined unique binding of the TFO in the **U2**·CG triplex.

However, **U2** contributes only weakly to the stability of a triplex with a CG inversion site. As summarized in Table 1, melting temperatures are lower by $\geqslant 15$ °C when compared to triplexes containing only canonical T·AT and C^+·GC triads and lower by 5 °C compared to a noncanonical but moderately stabilizing G·TA triplex.[19] Such a weak stabilization by the nucleoside analog may be attributed to (i) the lack of a large surface area necessary for optimal π stacking interactions and (ii) the rather hydrophilic nature of the urocanamide. Although we have increased the hydrophobicity of the base by N-alkylation, the amide is still expected to be

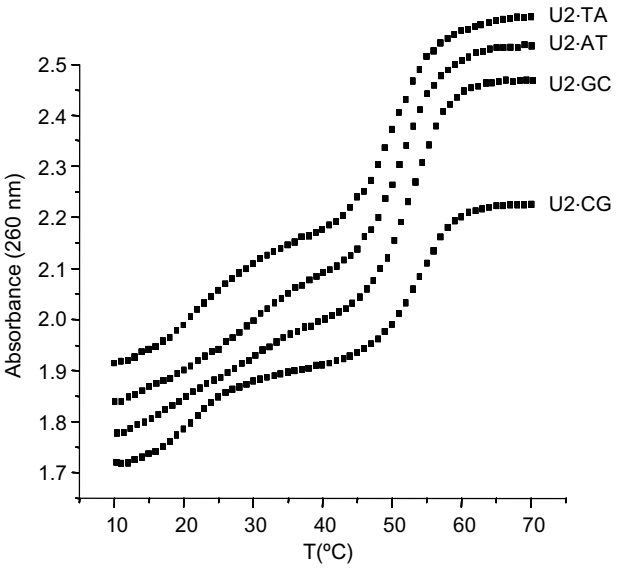


Figure 2. Temperature-dependent absorbance at 260 nm of triplexes with **OL15**. [Triplex] = 5.8 μM, 100 mM NaCl, 20 mM $MgCl_2$, pH = 6.0. Melting curves are offset for the sake of clarity.

Table 1. Summary of triplex–duplex melting temperatures T_m (°C) of triplexes with different Z·XY base triplets

Z·XY base triplet	T_m (°C), pH 6.0	T_m (°C), pH 6.5
C^+·GC	36.2	29.2
T·AT	>39	35.8
G·TA	26.5	21.9
U2·CG	21.0	20.2
U2·TA	21.7	18.4
U2·GC	—[a]	—[a]
U2·AT	—[a]	—[a]

[a] Transition not defined.

significantly hydrated and binding may therefore suffer from a significant initial desolvation. Interestingly, a similar situation with poor triplex stabilization was also encountered for various base analogs carrying strongly hydrophilic ureido substituents.[20,21]

Although designed for the selective hydrogen bond mediated recognition of a CG base pair, the, albeit broad, triplex–duplex transition of the **U2**·TA triplex at a midpoint temperature of 22 °C at pH 6.0 also points to some specific interactions between the **U2** analog and the TA base pair. Such an interaction might possibly involve the formation of a hydrogen bond between a protonated imidazole of **U2** and the 4-carbonyl oxygen of the thymine base as shown in Figure 3. Although the free urocanamide is expected to be largely deprotonated at pH 6.0, the apparent imidazole pK_a within a **U2**·TA base triplet may be significantly elevated. In line with the proposed base triplets, the triplex melting temperature for the **U2**·TA triplex follows the well-known pH dependence of the canonical triplexes with protonated C^+·GC triplets and decreases upon increasing the pH to 6.5 (Table 1). However, an unprotonated urocanamide binding a CG base pair should partially counteract such a pH dependence and accordingly triplex melting for the **U2**·CG triplex hardly changes with increasing pH. Therefore, although still less stable, the **U2**·CG triplex shows a relative stabilization with increasing pH for the sequences tested.

In summary, our results on triplex formation with the **U2** containing TFO in aqueous solution is compatible with the specific binding mode as predicted from our model system of free nucleosides in aprotic solvents and support a hydrogen-bonded **U2**·CG base triplet closely isomorphous to the canonical base triads.

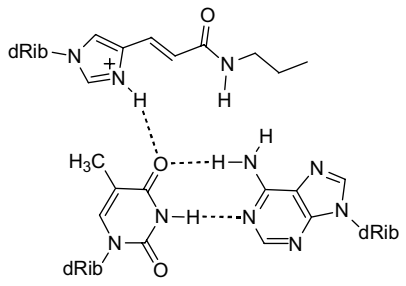


Figure 3. Possible base triplet between analogue **U2** and a TA Watson–Crick base pair.

However, the results also highlight the importance of additional contributions to binding not effective in the aprotic environment, which result in unexpectedly low affinities within a DNA triple helix. Clearly, further optimization with respect to more favorable base–base stacking and hydrophobic interactions while preserving the major structural features and geometry of the base analog are needed for a more efficient CG base pair recognition associated with stronger binding to a corresponding duplex target in aqueous solution.

References and notes

1. Duval-Valentin, G.; Thuong, N. T.; Hélène, C. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 504–508.
2. Kim, H.-G.; Miller, D. M. *Biochemistry* **1998**, *37*, 2666–2672.
3. Havre, P. A.; Gunther, E. J.; Gasparro, F. P.; Glazer, P. M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7879–7883.
4. Jones, W. R.; Stone, M. P. *Nucleic Acids Res.* **1998**, *26*, 1070–1075.
5. Dervan, P. B. *Nature* **1992**, *359*, 87–88.
6. Strobel, S. A.; Doucette-Stamm, L. A.; Riba, L.; Housman, D. E.; Dervan, P. B. *Science* **1991**, *254*, 1639–1642.
7. Cherny, D. Y.; Belotserkovskii, B. P.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.; Berg, R. H.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1667–1670.
8. Nielsen, P. E.; Christensen, L. *J. Am. Chem. Soc.* **1996**, *118*, 2287–2288.
9. Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723–5727.
10. Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 8198–8206.
11. Moser, H.; Dervan, P. B. *Science* **1987**, *238*, 645–650.
12. Thuong, N. T.; Hélène, C. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 666–690.
13. (a) Guianvarc'h, D.; Fourrey, J. L.; Maurisse, R.; Sun, H. S.; Benhida, R. *Bioorg. Med. Chem.* **2003**, *11*, 2751–2759; (b) Rusling, D. A.; Powers, V. E. C.; Ranasinghe, R. T.; Wang, Y.; Osborne, S. D.; Brown, T.; Fox, K. R. *Nucleic Acids Res.* **2005**, *33*, 3025–3032; for reviews, see: (c) Gowers, D. M.; Fox, K. R. *Nucleic Acids Res.* **1999**, *27*, 1569–1577; (d) Doronina, S. O.; Behr, J.-P. *Chem. Soc. Rev.* **1997**, 63–71; (e) Purwanto, M. G. M.; Weisz, K. *Curr. Org. Chem.* **2003**, *7*, 427–446.
14. Purwanto, M. G. M.; Lengeler, D.; Weisz, K. *Tetrahedron Lett.* **2002**, *43*, 61–64.
15. Purwanto, M. G. M.; Weisz, K. *J. Org. Chem.* **2004**, *69*, 195–197.
16. Griffin, L. C.; Kiessling, L. L.; Beal, P. A.; Gillespie, P.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 7976–7982.
17. (a) Hoffer, M. *Chem. Ber.* **1960**, *93*, 2777–2781; (b) Rolland, V.; Kotera, M.; Lhomme, J. *Synth. Commun.* **1997**, *27*, 3505–3511.
18. Spectroscopic data of **4**: ^{1}H NMR (600 MHz, 293 K, CD_2Cl_2): δ (ppm) = 0.91 (t, 3H; CH_3), 1.53 (m, 2H; CH_2), 2.40–2.56 (m, 2H; H2′, H2″), 3.15 (m, 2H; N–CH_2), 3.46–3.50 (m, 2H; H5′, H5″), 3.70 (s, 6H; O–CH_3), 4.10 (m, 1H; H4′), 4.47 (m, 1H; H3′), 5.94 (dd, 1H; H1′), 6.43 (d, 1H; C═C–H), 7.16–7.90 (m, 16H; ArH, C═C–H, Im–H).
19. Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967–971.
20. Guzzo-Pernell, N.; Lawlor, J. M.; Tregar, G. W.; Haralambidis, J. *Aust. J. Chem.* **1998**, *51*, 965–972.
21. Mertz, E.; Mattei, S.; Zimmerman, S. C. *Bioorg. Med. Chem.* **2004**, *12*, 1517–1526.