Available online at www.sciencedirect.com
ScienceDirect
Tetrahedron Letters
Tetrahedron Letters 48 (2007) 2143–2145

Synthesis and evaluation of hexitol nucleoside congeners as ambiguous nucleosides

Catia Lambertucci,[a,b] Guy Schepers,[b] Gloria Cristalli,[a] Piet Herdewijn[b] and Arthur Van Aerschot[b,*]

[a]*Dipartimento di Scienze Chimiche, Università di Camerino, I-62032 Camerino (MC), Italy*
[b]*Laboratory of Medicinal Chemistry, Rega Institute for Medical Research, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium*

Received 23 November 2006; revised 16 January 2007; accepted 19 January 2007
Available online 25 January 2007

Abstract—A series of anhydrohexitol nucleoside congeners was synthesized as ambiguous or so-called universal nucleosides and was evaluated for their hybridization potential and discrimination properties. The 1,5-anhydro-2,3-dideoxy-2-(5-nitroindazol-1-yl)-D-*arabino*-hexitol **4e** showed the lower spread in T_m values upon hybridization to the natural bases, with minimal destabilization, and therefore behaved as a true ambiguous nucleoside.

1. Introduction

In the design of oligonucleotide primers or probes, a nucleoside pairing equally well with all four natural bases, a so-called universal or ambiguous nucleobase, could be used to overcome problems due to degeneracy of the genetic code or to incomplete peptide sequence data. Over the last years, different surrogate bases have been evaluated as universal or degenerate nucleosides, which either cannot associate through hydrogen bonding but provide a polarized stackable heterocycle, like 5-nitroindazole and 4-nitroimidazole, or allow a flexible hydrogen bonding pattern mimicking natural bases, like the azole carboxamide derivatives.[1,2] On the other hand, duplex stability of nucleic acids can be increased by the modifications of the carbohydrate moiety, like in constrained hexitol nucleic acids.[3] Therefore, a series of analogs (**4a**–**c**) was prepared bearing a 5-nitroindazole, 4-nitroimidazole or 1,2,4-triazole-3-carboxamide as the base moiety attached to 1,5-anhydro-3-deoxy-D-glucitol as the sugar part. Following incorporation into oligodeoxynucleotides, their base pairing and discriminatory properties have been evaluated.

2. Results and discussion

The synthesis of the nucleoside congeners was accomplished by reacting tosylated precursor **2** with the different heterocyclic bases (**1a**–**c**) in DMF using NaH as basic catalyst (Fig. 1). Heating for 2 h at 110 °C with nitroindazole afforded a mixture of the N^1- and N^2-alkylated congeners (**3a** in 47% and **3b** in 35%, respectively). Prolonged heating (60 h) proved necessary for alkylation with nitroimidazole (yielding **3c**), and a lower temperature of 90 °C was used with the triazole methylcarboxylate affording **3d**. In addition, small amounts of, respectively, the 5-nitroimidazolyl and the triazol-4-yl congeners were obtained and were isolated by column chromatography. Benzylidene deprotection with 80% acetic acid at 70 °C for 1 h afforded **4a**,**c**,**e** in about 85% yield.[4] Treatment of **3d** with methanolic ammonia produced the required amide **3e**. Dimethoxytritylation and phosphitylation afforded the required phosphoramidites **6a**,**c**,**e**.[5] However, phosphitylation of tritylated **5e** was complicated by additional phosphitylation of the carboxamide, reducing the yield to 38%. Phosphitylation of carboxamide moieties of heterocyclic bases was described before as a possible complication.[6] Standard oligonucleotide assembly, deprotection, and purification by anion exchange chromatography, afforded the desired oligonucleotides, which were characterized by HPLC–MS as described before.[7] While phosphitylation of **5e** was troublesome, assembly using the unprotected carboxamide **6e**, afforded flawless coupling upon

Keywords: Universal nucleosides; Anhydrohexitol nucleosides; Duplex stability; Ambiguous base; Degenerate nucleosides.
* Corresponding author. Tel.: +32 16 337388; fax: +32 16 337340; e-mail: arthur.vanaerschot@rega.kuleuven.be

0040-4039/$ - see front matter © 2007 Published by Elsevier Ltd.
doi:10.1016/j.tetlet.2007.01.111

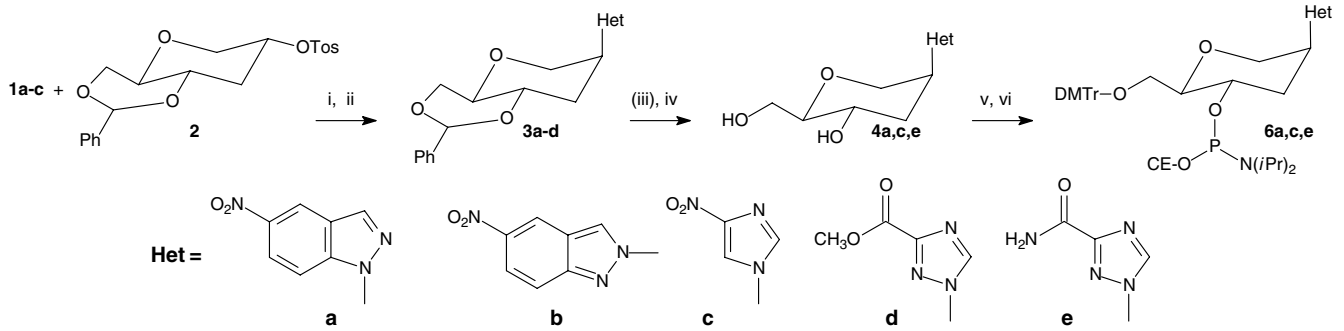

Figure 1. Synthetic scheme for preparation of the new anhydrohexitol nucleoside analogs **4a**,**c**,**e** and their respective amidites **6a**,**c**,**e**. Reaction conditions: (i) NaH / DMF, 60 °C, 1 h (**1a**: 5-nitroindazole; **1b**: 4-nitroimidazole; **1c**: 1,2,4-triazole-3-methylcarboxylate); (ii) add **2**, DMF, T (**3a** 47%; **3b** 35%; **3c** 59%; **3d** 45%); (iii) NH_3/MeOH; (iv) AcOH 80%, 70 °C (**4a** 87%; **4c** 83%; **4e** 86% combined); (v) 4,4′-dimethoxytrityl chloride (DMTrCl, 1.1 equiv), Py, 3 h (**5a** 87%; **5c** 77%; **5e** 80%); (vi) $EtN(iPr)_2$, CH_2Cl_2, $(iPr)_2NP(Cl)OCH_2CH_2CN$ (**6a** 89%; **6c** 77%; **6e** 36%).

monitoring via trityl analysis. Yields for the isolated oligonucleotides (30–45 OD_{260}) proved the same for the different modifications. Smaller amounts of slower eluting products, however, were noticed, especially with multiple incorporation of the triazole carboxamide.

The base pairing properties of the analogs were examined by hybridizing the modified oligomers with natural DNA strands carrying one of the four natural bases in juxtaposition to the modification, using temperature-dependent UV spectroscopy. Hereby, the spread in T_m is given with T_m for the best hybridizing base minus T_m for the least hybridizing base.

Compared to the DNA reference,[7] introduction of one of the new nucleosides (**4a**–**c**) in the sequence 5′-d(CACCG**X**TGCTACC)-3′ led to moderately reduced thermal stability of the resulting duplexes with the respective natural complements (Table 1). Thus, with introduction of 5-nitroindazole nucleoside analog (**4a**), T_m values in the range of 48.6–50.3 °C were observed, which proved to be the least destabilizing congener among the newly tested compounds. In addition, it showed the least spread in T_m values (ΔT_m of 1.7 °C) which is a prime requirement for a true universal nucleoside. With an average destabilization of 7.8 °C versus the A/T matched sequence, this analog perfectly matches the results for the acyclic nitroindazole analog **7** (ΔT_m of 1.7 °C, average destabilization of 7.8 °C, Fig. 2),[6] as well as for the ribosylated 5-nitroindole analog **8** (ΔT_m of 1.0 °C, average destabilization of 8.1 °C).[8]

Table 1. Hybridization studies[a] for the new anhydrohexitol analogs (**X**) within the oligonucleotide duplex 5′-d(CACCG**X**TGCTACC)-3′/3′-d(GTGGCYACGATGG)-5′

X	Y =	T	A	C	G	ΔT_m[b]	Average T_m[c]
T		46.8	57.0	44.4	50.9		
A		57.4	46.6	44.7	53.5		
4a		50.1	49.4	50.3	48.6	1.7	49.6
4c		42.5	43.4	42.7	50.1	7.6	44.7
4e		46.3	47.1	41.8	53.6	11.8	47.2
9		46.2	45.4	45.0	45.0	1.2	45.4

[a] T_m as determined by taking the first derivative of the melting curve by slow heating and cooling (0.2 °C/min) in 0.1 M NaCl, potassium phosphate (0.2 M, pH 7.5) EDTA (0.1 mM).

[b] The spread in T_m (ΔT_m) is calculated as the T_m for the highest melting complementary oligo minus the lowest melting one for each specific modification.

[c] Average T_m versus the matched sequence A/T (57.4 °C).

A shorter acyclic chain carrying the nitroindazole base under the form of the glycol nucleic acid analogue **9**, was evaluated as well and was prepared[9] in analogy with the report of Zhang et al.[10] The higher flexibility of acyclic **7** presumably allows optimal positioning of the base analog, where this seems counterbalanced by the reduced entropic penalty upon duplex formation for the constrained anhydrohexitol **4a**, affording the same overall outcome. In terms of ambiguity **9** performs even slightly better with a ΔT_m of only 1.2 °C. However, this analog gives a larger decrease in hybridization temperature, probably as of the shorter chain length.

Analysis for a second sequence [5′-d(AGTATTG**X**-CCTA)-3′] afforded the same overall picture with a spread in T_m of 2.2 °C for congener **4a**. Finally, the self-complementary sequence [5′-d(CGC**X**AATT**Y**-GCG)-3′] reported before[11] and comprising two points

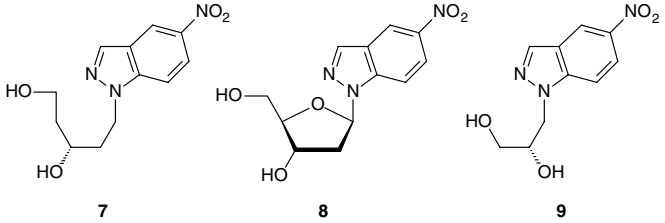

Figure 2. Structures for the acyclic nitroindazole congeners **7** and **9** and the ribosylated 5-nitroindole **8**.

of modification per double stranded complex, was studied with **X** being one of the different new analogs. However, even under high salt conditions (1.0 M NaCl) a mixture of hairpins and duplex structures prevailed, in contrast with data reported for other modified analogs.[11]

3. Conclusions

Three new anhydrohexitol nucleoside analogs were synthesized, characterized, and incorporated into several oligonucleotide sequences for thermal denaturation studies. While all modifications destabilized the double helix upon a single incorporation, 5-nitroindazole congener **4a** was the least destabilizing and showed the least spread in T_m values and therefore is behaving almost like a true ambiguous nucleoside analog comparable to the commercially available 5-nitroindazole deoxyribofuranoside. The performance of **4a** in sequencing and for PCR primers is under study and is clearly important in validating its final usefulness as a universal base.

Acknowledgments

The authors are indebted to Dr. Rozenski for mass determinations, and are grateful to Chantal Biernaux for editorial help. Financial support was obtained from the K. U. Leuven (GOA) and the Research Foundation-Flanders (G.0603.06).

Supplementary data

Supplementary data under the form of experimental procedures and analytical data associated with this article are available. Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.tetlet.2007.01.111.

References and notes

1. Loakes, D. *Nucleic Acids Res.* **2001**, *29*, 2437–2447.
2. Paul, N.; Nashin, V. C.; Hoops, G.; Zhang, P.; Zhou, J.; Bergstrom, D. E.; Davisson, V. J. *Chem. Biol.* **2003**, *10*, 815–825.
3. Hendrix, C.; Rosemeyer, H.; Verheggen, I.; Seela, F.; Van Aerschot, A.; Herdewijn, P. *Chem. Eur. J.* **1997**, *3*, 110–120.
4. *Analytical data for the deprotected nucleoside analogs* **4a**: ^{1}H NMR (DMSO-d_6, 200 MHz) δ 1.99 (m, 1H, H-3′a); 2.51 (m, 1H, H-3′e); 3.24 (m, 1H, H-5′); 3.35 (m, 2H, H-6′); 3.70 (m, 1H, H-4′); 3.79 (dd, 1H, $J = 2.6$ Hz and 12.4 Hz, H-1′a); 4.04 (m, 1H, H-1′e); 4.55 (t, 1H, $J = 5.9$ Hz, 6′-OH); 4.97 (d, 1H, $J = 5.6$ Hz, 4′-OH); 5.11 (br s, 1H, H-2′); 7.91 (d, 1H, $J = 9.2$ Hz, H-7); 8.22 (dd, 1H, $J = 2.2$ Hz and 9.2 Hz, H-6); 8.41 (s, 1H, H-3); 8.83 (d, 1H, $J = 2.2$ Hz, H-4). ^{13}C NMR (DMSO-d_6, 200 MHz) δ 35.1 (C-3′); 54.4 (C-2′); 61.4 (C-6′); 62.6 (C-4′); 68.0 (C-1′); 83.1 (C-5′); 111.0 (C-7); 119.2 (C-4); 120.8 (C-6); 123.1 (C-9); 135.8 (C-3); 140.6 (C-8); 141.9 (C-5). ESMS calcd for $C_{13}H_{16}N_3O_5$ ([M+H]$^+$): 294.1090. Found: 294.1086; Compound **4c**: ^{1}H NMR (DMSO-d_6, 200 MHz) δ 1.88 (m, 1H, H-3′a); 2.28 (m, 1H, H-3′e); 3.16 (m, 1H, H-5′); 3.55 (m, 3H, H-4′ and H-6′); 3.77 (m, 1H, H-1′a); 4.16 (m, 1H, H-1′e); 4.63 (m, 2H, 6′-OH and H-2′); 4.97 (d, 1H, $J = 4.8$ Hz, 4′-OH); 8.00 (s, 1H, H-2); 8.47 (s, 1H, H-5). ^{13}C NMR (DMSO-d_6, 300 MHz) δ 36.7 (C-3′); 54.2 (C-2′); 60.2 (C-6′ and 4′); 67.7 (C-1′); 82.7 (C-5′); 121.1 (C-5); 136.9 (C-2); 146.7 (C-4). ESMS calcd for $C_9H_{14}N_3O_5$ ([M+H]$^+$): 244.0933. Found: 244.0925; Compound **4e**: ^{1}H NMR (DMSO-d_6, 200 MHz) δ 1.89 (m, 1H, H-3′a); 2.50 (m, 1H, H-3′e); 3.16 (m, 1H, H-5′); 3.41 (m, 1H, H-6′); 3.64 (m, 2H, H-4′, H-1′a); 4.30 (m, 1H, H-1′e); 4.60 (t, 1H, $J = 6.0$ Hz, 6′-OH); 4.69 (m, 1H, H-2′); 5.00 (d, 1H, $J = 5.8$ Hz, 4′-OH); 7.61 (s, 1H, NH); 7.83 (s, 1H, NH); 7.85 (s, 1H, H-5). ^{13}C NMR (DMSO-d_6, 300 MHz) δ 35.5 (C-3′); 59.9 (C-2′); 61.1 (C-6′ and C-4′); 67.8 (C-1′); 83.5 (C-5′); 144.6 (C-5); 156.7 (C-3); 161.0 (CO); ESMS calcd for $C_9H_{15}N_4O_4$ ([M+H]$^+$): 243.1093. Found 243.1079.
5. *Experimental data for phosphoramidites*: Compound **6a**: 89% yield; ^{31}P NMR: δ 148.11, 148.26; ESMS calcd. for $C_{43}H_{51}N_5O_8P$ ([M+H]$^+$): 796.3475. Found 796.3460; Compound **6c**: 77% yield; ^{31}P NMR: δ 148.11, 149.15; ESMS calcd for $C_{39}H_{49}N_5O_8P$ ([M+H]$^+$): 746.3318. Found 746.3316; Compound **6e**: 38% yield; ^{31}P NMR: δ 148.49, 149.03; ESMS calcd for $C_{39}H_{50}N_6O_7P$ ([M+H]$^+$): 745.3478. Found 745.3478; double phosphitylated **6e analog**: 35% yield; ^{31}P NMR: δ 148.55, 148.62, 148.92, 149.09; ESMS calcd for $C_{48}H_{67}N_8O_8P_2$ ([M+H]$^+$): 945.4557. Found: 945.4555.
6. Van Aerschot, A.; Rozenski, J.; Loakes, D.; Pillet, N.; Schepers, G.; Herdewijn, P. *Nucleic Acid Res.* **1995**, *23*, 4363–4370.
7. Zhou, D.; Lagoja, I. M.; Rozenski, J.; Busson, R.; Van Aerschot, A.; Herdewijn, P. *ChemBioChem.* **2005**, *6*, 2298–2304.
8. Loakes, D.; Brown, D. M. *Nature* **1994**, *369*, 492–493.
9. Alkylation of nitroindazole was accomplished with toluenesulfonylated (*S*)-isopropylideneglycerol in DMF at 60 °C. 9: ^{1}H NMR (DMSO-d_6) δ 3.39–3.42 (m, 2H, H-3′), 3.96 (m, 1H, H-2′), 4.33–4.44 (dd, 1H, H-1′a, $J =$ 7.6 Hz and 14.0 Hz), 4.51–4.60 (dd, 1H, H-1′b, $J = 3.8$ Hz and 14.0 Hz), 7.79–7.84 (d, 1H, H-6, $J = 9.2$ Hz); 7.986–8.04 (dd, 1H, H-7, $J = 2.2$ Hz and 9.6 Hz), 8.75 (s, 1H, H-3), 8.89–8.90 (d, 1H, H-9, $J = 2.2$ Hz); ^{13}C NMR (DMSO-d_6) δ 57.2 (C-1′), 63.6 (C-3′), 70.7 (C-2′), 118.1 (C-6); 119.6 (C-7); 119.9 (C-4); 120.7 (C-9), 130.5 (C-3), 142.0 (C-8); 149.2 (C-5). ESMS calcd. for $C_{10}H_{12}N_3O_4$ ([M+H]+): 238.0828. Found: 238.0821.
10. Zhang, L.; Peritz, A.; Meggers, E. *J. Am. Chem. Soc.* **2005**, *127*, 4174–4175.
11. Zhang, P.; Johnson, W. T.; Klewer, D.; Paul, N.; Hoops, G.; Davisson, V. J.; Bergstrom, D. E. *Nucleic Acids Res.* **1998**, *26*, 2208.