Available online at www.sciencedirect.com
SCIENCE DIRECT®
Tetrahedron Letters 45 (2004) 9365–9368
Tetrahedron Letters

Synthesis and hybridization affinity of oligodeoxyribonucleotides incorporating 4-*N*-(*N*-arylcarbamoyl)deoxycytidine derivatives

Kenichi Miyata,[a,c] Ryuji Tamamushi,[a] Akihiro Ohkubo,[a,c] Haruhiko Taguchi,[a,c] Kohji Seio[b,c] and Mitsuo Sekine[a,c,*]

[a]*Department of Life Science, Tokyo Institute of Technology, 4259 Nagatsuta, Midoriku, Yokohama 226-8501, Japan*
[b]*Frontier Collaborative Research Center, Tokyo Institute of Technology, 4259 Nagatsuta, Midoriku, Yokohama 226-8503, Japan*
[c]*CREST, JST (Japan Science and Technology Agency), 4259 Nagatsuta, Midoriku, Yokohama 226-8501, Japan*

Received 24 September 2004; revised 20 October 2004; accepted 21 October 2004
Available online 5 November 2004

Abstract—Modified oligodeoxynucleotides incorporating 4-*N*-(*N*-arylcarbamoyl)-dC derivatives **1a**–**c** were synthesized. The ^{1}H NMR spectra of **1a**–**c** suggest that the carbamoyl group forms an intramolecular hydrogen bond with the cytosine ring nitrogen atom so that formation of a Watson–Crick base pair with the complementary guanine base is inhibited. The hybridization properties of oligodeoxynucleotides containing **1a**–**c** were investigated by use of T_m analysis. The hybridization properties of 4-*N*-(*N*-phenylcarbamoyl)-dC (**1a**) were similar to those of 4-*N*-(*N*-alkylcarbamoyl)-dC derivatives reported previously. In sharp contrast to **1a**, it turned out that 4-*N*-(*N*-napht-1-yl) and (*N*-quionol-5-yl)-dC (**1b**,**c**) have a unique property as a universal base.

A variety of base-modified nucleosides have been extensively synthesized to enhance their original functional roles and to understand their effect on the neighboring circumstance when introduced into the oligonucleotides. Recently, we have reported the synthesis and hybridization properties of oligodeoxynucleotides containing 4-*N*-acyl-dC derivatives,[1,2] 4-*N*-alkoxycarbonyl-dC derivatives,[3] and 4-*N*-alkylcarbamoyl derivatives.[4] Our results showed that such modification of the cytosine 4-*N*-amino group with acyl or alkoxycarbonyl not only allowed formation of the Watson–Crick base pair but also increased their hybridization affinity for the complementary guanine base owing to an intramolecular hydrogen bond between the carbonyl oxygen atom and the 5-H proton of the cytosine ring. On the other hand, it was found that the carbamoyl group of 4-*N*-(carbamoyl)deoxycytidine has a '*distal*' orientation due to an intramolecular hydrogen bond between the carbamoyl group and the cytosine ring nitrogen atom so that formation of a Watson–Crick base pair with the complementary guanine base is inhibited.[4]

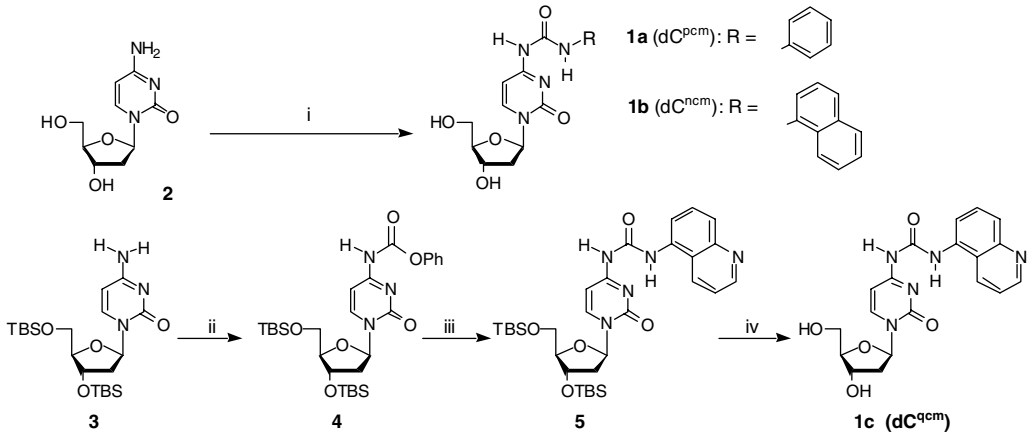
In this letter, we report the synthesis and properties of deoxycytidine derivatives (**1a**–**c**) having various 4-*N*-(*N*-arylcarbamoyl) groups, and also describe their unique duplex stability and base recognition ability when they were incorporated into oligodeoxynucleotides. Tamura et al. reported the reaction of DNA with 2-naphtyl isocyanate gave 4-*N*-(napht-2-yl)carbamoyl-dC as a degradation product.[5] Recently, Sugimoto and his co-workers have reported that 4-*N*-arylcarbonyl-dA derivatives serve as base pair mimics in DNA duplexes.[6]

The synthesis of nucleosides **1a-c** is outlined in Scheme 1. Compounds **1a**,**b** were obtained by selective reaction of dC with aryl isocyanates in high yield. On the other hand, 4-*N*-[*N*-(quinol-5-yl)carbamoyl]-dC (**1c**: dC^{qcm}) was synthesized from 3′,5′-*O*-disilylated deoxycytidine (**3**). The ^{1}H NMR spectra of **1a**–**c** thus obtained are similar to those of 4-*N*-(*N*-alkylcarbamoyl)-dC derivatives. This result suggests that the orientation of the carbamoyl groups is '*distal*'.

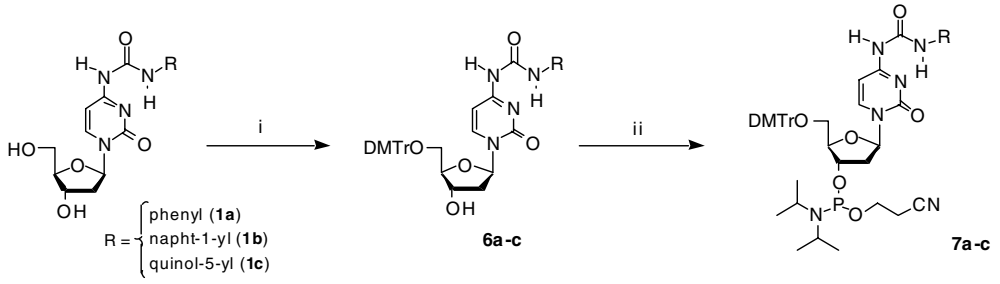
The synthesis of target phosphoramidites **7a**–**c** is outlined in Scheme 2. Treatment of **1a**–**c** with DMTrCl gave the 5′-*O*-dimethoxytritylated products **6a**–**c**. Compounds **6a**–**c** were converted to the phosphoramidite units **7a**–**c** by phosphitylation in the usual way.[7]

Keywords: 4-*N*-carbamoyldeoxycytidine; Oligonucleotide; T_m; Universal base.
* Corresponding author. Tel.: +81 45 924 5706; fax: +81 45 924 5772; e-mail: msekine@bio.titech.ac.jp

0040-4039/$ - see front matter © 2004 Published by Elsevier Ltd.
doi:10.1016/j.tetlet.2004.10.116



Scheme 1. Reagents and conditions: (i) aryl isocyanate (1.0 equiv), triethylamine (1.0 equiv), DMF, rt, 30 min (**1a**: 92%, **1b**: 84%); (ii) phenyl chloroformate (1.5 equiv), CH_2Cl_2–pyridine (1.1, v/v), rt, 1 h (86%); (iii) 5-aminoqunoline (1.5 equiv), pyridine, 70 °C, 3 h (60%); (iv) triethylamine–3HF (5 equiv), rt, 7 h (84%).



Scheme 2. Reagents and conditions: (i) DMTrCl (1.2 equiv), pyridine, rt, 4 h (**6a**: 87%, **6b**: 89%, **6c**: 85%); (ii) chloro(2-cyanoethoxy)(*N*,*N*-diisopropylamino)phosphine (1.1 equiv), diisopropylethyamine (1.3 equiv), CH_2Cl_2, rt, 30 min (**7a**: 84%, **7b**: 78%, **7c**: 75%).

The solid-phase synthesis of oligodeoxynucleotides containing 4-*N*-(*N*-arylcarbamoyl)-dC derivatives was carried out with a DNA/RNA synthesizer by use of the standard phosphoramidite method.[8] The oligomer was released and deprotected from the polymer support by treatment with concd aq NH_3 for 1 h. The product was purified on C_{18}-cartridge by the DMTr-ON purification method, analyzed by anion-exchange HPLC and reversed-phase HPLC, and characterized by MALDI-TOF mass. These results are summarized in Table 1.

From the previous studies of 4-*N*-(*N*-alkylcarbamoyl)-dC derivatives,[4] it was expected that 4-*N*-(*N*-arylcarbamoyl) modified nucleosides would form a Watson–Crick base pair with deoxyguanosine, with their conformational change from '*distal*' to '*proximal*'. Therefore, the thermal stability of DNA duplexes containing a modified nucleoside was investigated in sodium phosphate buffer (pH 7.0) containing 1.0 M NaCl. The T_m values of the duplexes containing modified nucleosides **1a**–**c** are summarized in Table 2. As references, the T_m values of duplexes containing 4-*N*-(*N*-methylcarbamoyl)-dC (dC^{mcm}; entry 2), 4-*N*-(*N*-ethylcarbamoyl)-dC (dC^{ecm}; entry 3), and 4-*N*-(*N*-butylcarbamoyl)-dC (dC^{bcm}; entry 4) are also shown in Table 2.

Table 1. Sequence and isolated yield of oligodeoxynucleotides containing 4-*N*-(*N*-arylcarbamoyl)-dC derivatives

Sequence	Yield (%)	Found[a]	Calcd
d(TTTCTCC^{pcm} TTCTCT)	36	3934.0	3935.7
d(TTTCTCC^{ncm} TTCTCT)	36	3986.5	3985.7
d(TTTCTCC^{qcm} TTCTCT)	66	3987.1	3986.7

[a] MALDI-TOF mass.

It turned out that the T_m values of the duplex containing dC^{pcm}-G (entry 5, 49.3 °C), dC^{ncm}-G (entry 6, 47.5 °C), and dC^{qcm}-G (entry 7, 48.7 °C) base pair were quite lower than that of the unmodified control duplex (entry 1, 53.2 °C). The stability of these modified duplexes was almost the same as that of the duplex substituted with a T–A base pair (entry 7, 49.2 °C). It is inferred that the carbamoyl groups of modified deoxycytidines changed their orientation from '*distal*' to '*proximal*' that does not interrupt the formation of a Watson–Crick base pair with deoxyguanosine, as shown in Figure 1. Compared with the T_m values of the duplexes containing a 4-*N*-alkylcarbamoyl-dC derivative (entries 2–4), these results are in good agreement with those of oligodeoxynucleotides having a 4-*N*-(*N*-alkylcarbamoyl)-dC derivative. The stability of the modified duplexes highly depends on the hydrophilicity of a substituent of the carbamoyl group.

The T_m values of the DNA duplexes having a mismatched base pair (Y = T, C, A) are also summarized in Table 2. The T_m values of the duplexes containing a

Table 2. T_m values[a] for DNA 13 mer duplexes containing 4-*N*-carbamoyldeoxycytidine derivatives
5′-d(TTTCTC**X**TTCTCT)-3′ 10 mM sodium phosphate (pH 7.0), 1.0 M NaCl,
3′-d(AAAGAG**Y**AAGAGA)-5′ 0.1 mM EDTA, 2 μM duplex

Entry	X	Y = G		Y = T		Y = C		Y = A		Selectivity[c]	Range[d]
		T_m	ΔT_m[b]	T_m	ΔT_m[b]	T_m	ΔT_m[b]	T_m	ΔT_m[b]		
1	C	53.2		37.8		30.7		32.8		+14.4	24.5
2[e]	C^{mcm}	51.7	−1.5	39.6	+1.8	32.5	+1.8	38.1	+5.3	+12.1	18.2
3[e]	C^{ecm}	49.7	−3.5	38.8	+1.0	32.4	+1.7	38.8	+6.0	+10.9	17.3
4[e]	C^{bcm}	48.2	−5.0	38.0	+0.2	36.1	+5.4	40.3	+7.5	+7.9	12.1
5	C^{pcm} (**1a**)	49.3	−3.9	43.3	+5.5	44.1	+13.4	42.2	+9.4	+5.2	6.1
6	C^{ncm} (**1b**)	47.5	−5.7	47.7	+9.9	49.4	+18.7	46.4	+13.6	−1.9	3.0
7	C^{qcm} (**1c**)	48.7	−4.5	48.7	+10.9	49.3	+18.6	47.1	+14.3	−0.5	2.2
8	T							49.2			(°C)

[a] The T_m values are accurate within ±0.5 °C.
[b] ΔT_m is the difference in the T_m value between the duplex having a modified base and that having a natural base.
[c] '*Selectivity*' is the difference between the T_m value of the duplex containing a C*–G natch base pair and the highest T_m value of duplexes containing a mismatch base pair.
[d] '*Range*' is the difference between the maximum and minimum T_m value.
[e] Ref. 4; dC^{mcm}, de^{ecm}, and dC^{bcm} represent 4-*N*-(*N*-methylcarbamoyl)-dC, 4-*N*-(*N*-ethylcarbamoyl)1-dC, and 4-*N*-(*N*-butylcarbamoyl)-dC, respectively.

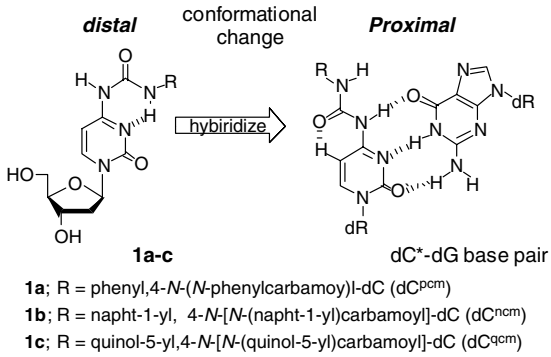


Figure 1. Structures of dC^{pcm} (**1a**), dC^{ncm} (**1b**), and dC^{qcm} (**1c**).

Table 3. T_m values[a] for DNA 13mer duplexes containing 4-*N*-carbamoyldeoxycytidine derivatives and abasic site

Entry	X	Y = abasic	
		T_m	ΔT_m[b](°C)
1	C	31.4	
2	C^{mcm}	33.9	+2.5
3	C^{bcm}	37.6	+6.2
4	C^{pcm}	45.5	+14.1
5	C^{ncm}	48.8	+17.4
6	C^{qcm}	51.1	+19.7

[a] The T_m values are accurate within ±0.5 °C. These sequences of duplexes and used buffer are indicated in Table 2.
[b] ΔT_m is the differences in the T_m value between the duplex having a modified base and that having a natural base.

4-*N*-(*N*-arylcarbamoyl)-dC derivative are higher than those of the duplexes containing a 4-*N*-alkylcarbamoyl-dC derivative. The base recognition abilities of modified dC derivatives are summarized in the column '*selectivity*'. The *selectivity* of the 4-*N*-(*N*-arylcarbamoyl)-dC derivatives was markedly low when compared with those of 4-*N*-(*N*-alkylcarbamoyl)-dC and unmodified dC derivatives. Particularly, the T_m values of the DNA duplexes having a C*–C mismatch base pair (entry 6, dC^{ncm}; 49.4 °C; entry 7, dC^{qcm} ; 49.3 °C) were higher than those of the DNA duplexes having a C*–G match base pair (entry 6, dC^{ncm}, 47.5 °C; entry 7, dC^{qcm}, 48.7 °C). The values in the column '*range*' represent the difference between the highest and the lowest T_m values in each lane. To our surprise, the *range* of 4-*N*-[*N*-(quinol-5-yl)carbamoyl]-dC (entry 7, 2.2 °C) was small enough to be called a universal base. It seemed to us that the orientation of the carbamoyl group was '*distal*' in the mismatched base pair, and the stabilization of the duplex having a mismatch base pair was caused by increase of the stacking effect of the substituent on the carbamoyl groups.

To confirm this assumption, the T_m values of DNA duplexes containing a C*–abasic pair were measured. We used a tetrahydrofuran derivative as an abasic site. These results are summarized in Table 3. It is clear that the T_m value increases with an increase of the surface area of the substituent of the carbamoyl group. The T_m values of the duplexes having a C*–abasic pair were almost close to those of the duplexes having a mismatched base pair. These results suggested that the stabilization of the duplexes containing a mismatch base pair is due to the stacking effect of the substituent of the carbamoyl group. However, the geometry of the mismatch base pair containing a 4-*N*-carbamoyl-dC derivative is still unclear. In addition, the T_m value of the duplex containing dC^{qcm} (entry 6, 51.1 °C) was quite higher than that of the duplex containing dC^{ncm} (entry 5, 48.8 °C). This is due to the nitrogen atom in the quinol-5-yl group.

The fluorescence of 2-aminopurine is known to be sensitive to a delicate change in its surrounding DNA structure. Particularly, it is strongly quenched when hybridized with a base. This inherent property of ap has been used to detect the local structure of bases flipped out by enzymes such as uracil DNA glycosylase[9] and adenine-specific DNA methyltransferase.[10] Therefore, we used this base as a site-specific fluorescent probe in DNA duplexes to see if the mismatched base is flipped out. As the result, the fluorescence of 2-aminopurine in

Table 4. Fluorescence intensity[a] of 2-aminopurine (ap) in DNA 13mer duplexes containing dC^{ncm}

I: 5′-d(AGAGAA**ap**GAGAAA)-3′,
II: 5′-d(TTTCTC**C**TTCTCT)-3′,
III: 5′-d(TTTCTC$\mathbf{C}^{ncm}$TTCTCT)-3′

Entry	Oligo	Fluorescence intensity
1	**Only I**	62
2	**I–II**	15
3	**I–III**	13

[a] Recorded with an exicitation wavelength of 310 nm and an emission wavelength at 370 nm.

DNA duplexes containing a C^{ncm}–ap mismatch base pair was actually quenched as shown in Table 4 (entries 1–3). These results suggest that the mismatched base moiety opposite to 4-*N*-[*N*-(napht-1-yl)carbamoyl]-dC is not flipped out but is involved in stacking of the DNA duplex. However, owing to the steric hindrance of the *N*-(napht-1-yl)carbamoyl group, the modified and complementary base moieties are expected to stack with each other in the DNA duplex.

The conventional universal bases would be classified into two groups. Most artificial universal bases such as 3-nitropyrrole[11] and 5-nitroindole[12] do not stack with the opposite base moiety. On the other hand, intercalator-like nucleobases such as pyrenyl *C*-nucleoside[13,14] act as not only universal bases but also intercalators. Our results mentioned above suggest that 4-*N*-[*N*-(quinol-5-yl)carbamoyl]-dC acts as an intercalator-type universal. On the other hand, it was suggested that, when the guanine base is located on the opposite site of the modified cytosine, the orientation of the *N*-phenylcarbamoyl group changes to '*proximal*' so that a Watson–Crick base pair is formed in the DNA duplex. These results would provide new insight into the design of artificial universal bases that induce change in their geometry. Future work should be done to elucidate the geometry of the C^{qcm}–G base pair in more detail.

Acknowledgements

This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan. This work was also supported by CREST of JST (Japan Science and Technology) and partially supported by the COE21 project.

References and notes

1. Wada, T.; Kobori, A.; Kawahara, S.; Sekine, M. *Tetrahedron Lett.* **1998**, *39*, 6907–6910.
2. Wada, T.; Kobori, A.; Kawahara, S.; Sekine, M. *Eur. J. Org. Chem.* **2001**, 4583–4593.
3. Kobori, A.; Miyata, K.; Ushioda, M.; Seio, K.; Sekine, M. *J. Org. Chem.* **2002**, *67*, 476–485.
4. Miyata, K.; Kobori, A.; Seio, K.; Sekine, M., *J. Org. Chem.*, unpublished work.
5. Tamura, N.; Aoki, K.; Lee, M. S. *Carcinogenesis* **1990**, *11*, 2009–2014.
6. Nakano, S.; Uotani, Y.; Nakashıma, S.; Anno, Y.; Fujii, M.; Sugimoto, N. *J. Am. Chem. Soc.* **2003**, *125*, 8086–8087.
7. *Current Protocols in Nucleic Acid Chemistry*; Beaucage, S. L., Bergstrom, D. E., Glick, G. D., Jones, R. A., Eds.; John Wiley & Sons, Inc.: New York, 2000.
8. Sinha, N. D.; Biernat, J.; Koster, H. *Tetrahedron Lett.* **1983**, *24*, 5843–5846.
9. Stivers, J. T.; Pankiewicz, K. W.; Watanabe, K. A. *Biochemistry* **1999**, *38*, 952–963.
10. Beuck, C.; Singh, I.; Bhattacharya, A.; Hecker, W.; Parmar, V. S.; Seitz, O.; Weinhold, E. *Angew. Chem., Int. Ed.* **2003**, *42*, 3958–3960.
11. Nichols, R.; Andrews, P. C.; Zhang, P.; Bergstrom, D. E. *Nature* **1994**, *369*, 492–493.
12. Loakes, D.; Brown, D. M. *Nucleic Acids Res.* **1994**, *22*, 4039–4043.
13. Matray, T. J.; Kool, E. T. *J. Am. Chem. Soc.* **1998**, *120*, 6191–6192.
14. Raunak; Babu, B. R.; Sorensen, M. D.; Parmar, V. S.; Harrit, N. H.; Wengel, J. *Org. Biomol. Chem.* **2004**, *2*, 80–89.