Pergamon

Tetrahedron Letters, Vol. 38, No. 43, pp. 7515-7518, 1997
© 1997 Elsevier Science Ltd
All rights reserved. Printed in Great Britain
0040-4039/97 $17.00 + 0.00

PII: S0040-4039(97)10020-X

Design of modified A*U and G*C base pairs with similar stability. *Implication for the DNA sequencing by hybridization*

Hong-Khanh Nguyen, Ulysse Asseline, Daniel Dupret[+] and Nguyen T. Thuong*

Centre de Biophysique Moléculaire, CNRS, Rue Charles Sadron, 45071 Orléans Cedex 02, France
and Appligène-Oncor, Parc d'Innovation, BP 72, 67402 Illkirch, France[+]

Abstract: Modified G*C, or A*U base pairs have been incorporated at the 4th position of a duplex composed with nine base pairs.The influence of these modifications on duplex stabilities has been studied by absorption spectroscopy. It has been found that $G^{4Me}C$ has thermal stability similar to that of the $A^{5Prop}U$ base pair. Thus duplexes involving $G^{4Me}C$ or $A^{5Prop}U$ base pairs exhibit the same Tm. © 1997 Elsevier Science Ltd.

Traditional sequencing techniques require electrophoresis. Recently several authors have proposed a novel DNA analysis method based on the hybridization of a labelled DNA fragment to a complete set of oligonucleotides immobilized as an ordered array on a solid surface[1-3]. However, these approaches have a serious drawback due to the base composition dependence on duplex stabilities. So perfect hybrid built with AT rich sequences would therefore have a similar or even lower stability than do hybrids built with GC rich sequences involving one mismatch[4]. In a recent paper, we described a new system of duplexes built with AT and/or $G^{4Et}C$ base pairs whose thermal stability was independent of their AT and $G^{4Et}C$ ratio content[4]. This new concept consists in the chemical modification of a GC base pair whose stability is very similar to that of an AT base pair. In this paper we describe another system designed to improve the former one by modifying simultaneously AT and GC base pairs in order to obtain duplexes with higher thermal stabilities.

This approach involves the replacement of natural deoxynucleosides by modified nucleosides in both strands of the DNA duplex. Thus for our study it was imperative to use modified nucleosides likely to be incorporated by DNA polymerases for the preparation and amplification of nucleic acid fragments to be analyzed. For an easier chemical synthesis, we chose to modify pyrimidine deoxynucleosides. To increase the stability of AT base pairs we used 5-(1-propynyl)-2'-deoxyuridine ($d^{5Prop}U$) since it is known to form with natural dA a $^{5Prop}UA$ pairing which is more stable than AT pairing[5-6]. To modulate the stability of GC pairing, we used dC analogs in which one hydrogen atom of the amino group at position 4 of dC was substituted by different alkyl, allyl or propargyl groups (R). This strategy would facilitate the selection of a $G^{4R}C$ base pair with a thermal stability similar to that of the $A^{5Prop}U$ base pair. Preliminary studies were carried out with duplexes **1** to **8** containing nine base pairs in which one base pair at position 4 was substituted by one AT, $A^{5Prop}U$, GC or $G^{R}C$ base pair (Table 1).

---

* Fax: (33) 2 38 63 15 17 E-mail: asseline@cnrs-orleans.fr

5’-d (T T T C T C **X** T C G T T) 3’
3’-d (A C A G **Y** A G C) 5’
9 7 6 5 4 3 2 1

Duplex	X	Y	Tm ± 1°C	Δ Tm/1 (°C)
1	A	T	46	
2	A	$^{5Prop}U$	48	+2
3	G	C	51	+ 5
4	G	^{4Me}C	49	+ 3
5	G	^{4Et}C	47	+ 1
6	G	^{4Pr}C	45	- 1
7	G	$^{4Allyl}C$	45.5	- 0.5
8	G	$^{4Propargyl}C$	45.5	- 0.5

Table 1: Melting temperatures at λ = 260 nm of natural duplexes and those involving a modified base pair at position 4. Tms were determined at an oligomer strand concentration of 2 μM in 10^{-2} M sodium cacodylate, pH 7 buffer containing 1M NaCl and 2×10^{-4} M EDTA.

The synthesis of oligonucleotides was performed by phosphoramidite chemistry on solid phase[7] using commercial products for $d^{5Prop}U$. Phosphoramidite derivatives of $d^{4Me}C$, $d^{4Et}C$, $d^{4Pr}C$, $d^{4Allyl}C$ and $d^{4Propargyl}C$ were prepared as described in a previous paper[4]. The absorption studies of the duplexes were carried out with 2 μM of each oligonucleotide in 10^{-2} M sodium cacodylate pH 7, buffer containing 1 M NaCl and 2×10^{-4} M EDTA on a UVIKON 941 cell changer spectrophotometer. Melting temperatures (Tm) were taken as the temperature corresponding to the half-dissociation of the complex. Note that the margin of error was about ± 1°C.

The results obtained were consistent with those expected. Substitution of the methyl group at position 5 of thymine (duplex **1**, Tm_1 = 46°C) by the 1-propynyl group (duplex **2**, Tm_2 = 48°C) stabilizes the modified duplex. The 1-propynyl group is coplanar to uracyl, thus allowing its stacking onto the adjacent 5’-base.

However substitution of dC on the amino function has a destabilizing effect. By changing the nature of the R group, we could easily modulate the stability of the $G^{4R}C$ base pair. Melting data showed that ^{4Me}C and ^{4Et}C would be good candidates (Table 1). Thus duplexes **4** (R = Me, Tm_4 = 49°C) and **5** (R = Et, Tm_5 = 47°C) have thermal stabilities very similar to that of duplex **2** (Tm_2 = 48°C) involving an $A^{5Prop}U$ base pair.

To verify the specific recognition of dG by $d^{4Me}C$ and dA by $d^{5Prop}U$, thermal denaturation experiments were carried out with duplexes involving the mispairing XC, $X^{4Me}C$ (X = T, C or A) and XT, $X^{5prop}U$ (X = T, G or C). Thermal stability data showed that the specificity of the $G^{4Me}C$ and $A^{5Prop}U$ pairing formations was maintained (Table 2). The $X^{4Me}C$ and $X^{5prop}U$ mismatches led to a decrease of about 10 to 20°C in Tm values compared to those obtained with sequences involving $G^{4Me}C$ or $A^{5Prop}U$.

5'-d (T T T C G T C X T C G T T) 3'
3'-d (A G C A G Y A G C) 5'

Y X	T	$^{5\text{Prop}}$U	C	$^{4\text{Me}}$C
A	46	48	27	27
T		32	28	25
G		36.5	51	49
C	28.5	28	21	17

Table 2: Melting temperature values (°C) at λ = 260 nm of perfect duplexes and ones containing a mismatch at position 4. Tms were determined on a sample concentration of 2 μM in 10^{-2} M sodium cacodylate, pH 7 buffer containing 1 M NaCl and 2 × 10^{-4} M EDTA

For further studies, we selected G$^{4\text{Me}}$C rather than G$^{4\text{Et}}$C base pairs for the following reasons:

- G$^{4\text{Me}}$C pairing is slightly more stable than the A$^{5\text{Prop}}$U pairing so that the thermal stability of duplexes built with A$^{5\text{Prop}}$U and/or G$^{4\text{Me}}$C base pairs is higher than those containing A$^{5\text{Prop}}$U and/or G$^{4\text{Et}}$C.
- d$^{4\text{Me}}$C has been found mainly in thermophilic bacteria[8] confirms its hybridization properties. To test the A$^{5\text{Prop}}$U and G$^{4\text{Me}}$C base pair couples selected above, we studied the thermal behavior of two modified duplexes involving nine A$^{5\text{Prop}}$U (duplex **11**) and nine G$^{4\text{Me}}$C base pairs (duplex **12**), respectively (Table 3). Studies were carried out as well on two natural duplexes **9** and **10**, used as references, which were derived from duplexes **11** and **12**, by replacing A$^{5\text{Prop}}$U by AT and G$^{4\text{Me}}$C by GC base pairs, respectively. These sequences representing the maximum of base composition variation allowed us to assess quickly the efficiency of such a system.

		Tm (°C)	Δ Tm/9 (°C)
9	T_2 [T T A] T_2 [A A T]$_3$	18.5	0
10	T_2 [C C G] T_2 [G G C]$_3$	63.5	+ 45
11	T_2 [*U *U A] T_2 [A A *U]$_3$	31	+ 12.5
12	T_2 [G *C G] T_2 [*C G *C]$_3$	33.5	+ 15

Table 3: Melting temperatures of natural and modified duplexes involving C*=$^{4\text{Me}}$C or U*=$^{5\text{Prop}}$U. Tm were determined at an oligomer strand concentration of 2 μM in 10^{-2} M sodium cacodylate, pH 7 buffer containing 1 M NaCl and 2 10^{-4} × M EDTA.

The thermal stability data of fully modified duplexes **11** and **12** were in agreement with the results obtained above with partially modified duplexes **2** and **4**. In fact, duplex **11** containing nine A$^{5\text{Prop}}$U base pairs (Tm_{11} = 31°C) shows thermal stability very close to that of duplex **12** involving nine G$^{4\text{Me}}$C base pairs (Tm_{12} = 33.5°C). The range of variation in transition obtained between both modified duplexes was only about 2.5°C (note that the margin of error on Tm values was about ± 1°C). Concerning double-stranded DNA involving natural nucleosides, their stabilities are highly dependent on the base composition. The Tm difference between natural duplexes **9** (Tm_9 = 18.5°C) and **10** (Tm_{10} = 63.5°C) involving nine AT and nine GC base pairs, respectively, was very great as expected (Δ Tm = 45°C). This new approach involving modified duplexes made with ^{5Prop}U and ^{4Me}C allowed us to obtain a more stable system than the one implying AT and/or G^{4Et}C base pairs described previously[4]. The gain in thermal stability was about 12°C which represents an important advantage since a set of shorter oligonucleotides may be used hence a limited sequence number.

These preliminary results show that duplexes of a given length built with A^{5Prop}U or G^{4Me}C base pairs exhibit a thermal stability independent of their base content in a classical buffer solution. Although these examples represent the maximum variation of the base composition, development of this work implies the synthesis of sequences involving two modified nucleosides d^{5Prop}U and ^{4Me}C to confirm previous results and to study the effect of sequence on hybridization properties. The successful use of such modified oligonucleotides as probes for reverse hybridization analysis requires the incorporation of these modified nucleosides into the DNA target sequences. The ability of d^{5Prop}UTP[9-10] and d^{4Me}CTP[11] to serve as a substrate in DNA polymerase enzyme catalyzed reactions has been demonstrated.

Acknowledgements: This work was supported by a BioMed contract (gene-CT 93-0009).

References

1. Southern, E. M., international patent, **1988**, WO 89/10977.
2. Drmanac, R., Labat, I., Bruckner, I., Crkvenjakov, R., *Genomics*, **1989**, *4*, 114-128.
3. Krapko, K. R., Lysov, Y. P., Khorlin, A. A., Ivanov, I. B., Yershov, A. D., Vasilenko, S. K.,Florentiev, V. L., Mirzabekov, A. D., *J. DNA Sequencing Mapp.*, **1991**, *1*, 375-388.
4. Nguyen, H. K., Auffray, P., Asseline, U., Dupret, D., Thuong, N. T., *Nucleic Acids Res.*, **1997**, *25*, 3059-3065.
5. Froehler, B. C., Wadwani, S., Terhorst, T. J., Gerrard, S. R., *Tetrahedron Lett.* **1992**, *33*, 5307-5310.
6. Sági, J., Scemzö, A., Ébinger, K., Szabolcs, A., Sági, G., Ruff, E., Ötvös. L., *Tetrahedron Lett.* **1993**, *34*, 2191-2194.
7. Beaucage, S. L., Caruthers, M. H., *Tetrahedron Lett.*, **1981**, *22*, 1859-1862.
8. Ehrlich, M., Norris, K. F., Wang, R. Y.-H., Kuo, K. C., Gehrke, C. W., *Bioscience Report*, **1986**, *6*, 387-393.
9. Ötvös. L., Szecsi, J., Sági, J., Kovacs, T., *Nucleic Acids Symp. Ser.* **1987**, *18*, 125-129.
10. Valko, K., Fellegvari, I., Sági, J., Scemzö, A., *J. Liq. Chromatogr.* **1989**, *12*, 2103-2116.
11. Li, S., Haces, A., Stupar, L., Gebeyehu, G., Pless, R. C., *Nucleic Acids Res.* **1993**, *21*, 2709-2714.

(*Received in France* 17 *July* 1997; *accepted* 30 *August* 1997)