Pergamon

Tetrahedron Letters, Vol. 38, No. 23, pp. 4083-4086, 1997

0040-4039/97 $17.00 + 0.00

PII: S0040-4039(97)00829-0

Studies Towards the Design of a Modified GC Base Pair With Stability Similar to that of the AT Base Pair

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, 67407 Illkirch, France[+]

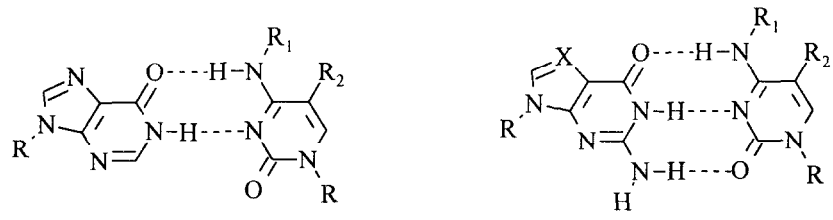
Abstract: Modified G*C, *GC or *G*C base pairs have been incorporated at the 3rd and 8th positions of a self-complementary decadeoxyoligonucleotide. The influence of these modifications on duplex stabilities has been studied by absorption spectroscopy. It has been found that a few of them have thermal stabilities similar to that of the AT base pair. © 1997 Elsevier Science Ltd.

The hybridization of DNA fragments immobilized on solid-support with oligonucleotides allow the detection of mutations[1, 2]. However, the efficiency of this strategy reaches its limit when the number of mutations or defects becomes great as in the case of the β-globine[3] or CFTR[4] genes and more particularly the p53 human gene whose mutations number over 6, 000[5, 6]. Recently a new strategy based on the hybridization of nucleic acids with an array of oligonucleotides of the same length immobilized on a flat area has been developed with the aim of analyzing long pieces of nucleic acids or a great number of nucleic sequences[7-9]. This strategy which implies the specific detection of hybrides without mismatch faces a basic difficulty due to the large difference of duplex stabilities depending on their base composition. Thus GC rich DNA duplexes involving a mismatch can be thermally more stable than perfect AT rich duplexes of the same length. Therefore the discrimination between perfect and mismatched hybrids becomes difficult when a great number of sequences are involved. To solve this problem, many studies were undertaken, however until now none has been successfully completed.

In this paper, we report the results of studies developed in order to obtain nucleic acid duplexes whose stabilities are independent or slightly dependent on their base composition. The method consists in modifying one or both nucleic bases of the GC base pair to obtain modified G*C, *GC or *G*C base pairs with thermal stabilities similar to that of the natural AT base pair. This approach involves the incorporation of the modified base(s) in the probes by a chemical synthesis process and in the nucleic acid fragments to be analyzed enzymatically. The latter constraints involve the choice of modified nucleosides whose 5'-triphosphates can be accepted by the DNA polymerases. For these reasons in the preliminary study we choose deoxycytidine derivatives substituted either at the 5-position by a methyl ($d^{5\text{-}Me}C$) or bromo ($d^{5\text{-}Br}C$) group or at the amino

* Fax: 02 38 63 15 17 E-mail: asseline@cnrs-orleans.fr

group position by an alkyl ($d^{4\text{-}Me}C$, $d^{4\text{-}Et}C$, $d^{4\text{-}Pro}C$) or propargyl ($d^{4\text{-}Propargyl}C$) or allyl ($d^{4\text{-}Allyl}C$) group as well as deoxyinosine (dI) and 7-deazadeoxyguanosine ($d^{7C}G$) as dG analogs. These modified bases are able to form with the natural or modified bases IC, I*C, ^{7C}GC, ^{7C}G*C or G*C base pairs whose thermal stabilities could be modulated by certain R_1 and R_2 substituents (figure).



dI **d*C** **d*G** **d*C**

d*C	R1	R2
dC	H	H
$d^{5\text{-}Me}C$	H	Me
$d^{5\text{-}Br}C$	H	Br
$d^{4\text{-}Me}C$	$-CH_3$	H

d*C	R1	R2
$d^{4\text{-}Et}C$	$-CH_2-CH_3$	H
$d^{4\text{-}Pro}C$	$-CH_2-CH_2-CH_3$	H
$d^{4\text{-}Allyl}C$	$-CH_2-CH=CH_2$	H
$d^{4\text{-}Propargyl}C$	$-CH_2-C\equiv CH$	H

d*G	X
dG	N
$d^{7C}G$	CH

R = 2'-deoxyriboside

Figure : Structures of the modified GC base pairs.

Due to the important work required for the synthesis of duplexes whereby every dC and/or dG has been replaced by their d*C and/or d*G analogs, we chose to study self-complementary duplexes involving ten base pairs whose GC base pair at the 3rd and 8th positions are either modified or replaced by an AT base pair (table). To increase the probability of obtaining a modified GC base pair whose thermal stability is slightly different from that of the AT base-pair, studies were developed with duplexes involving either a base pair formed by a natural and a modified nucleic base (G*C or *GC) or a base pair formed by two modified bases (*G*C). The synthesis of the self-complementary decadeoxynucleotides **1** to **19** was performed by phosphoramidite chemistry on solid phase using commercial products for $d^{5\text{-}Me}C$, $d^{5\text{-}Br}C$, $d^{7C}G$ and dI. The phosphoramidite derivatives of $d^{4\text{-}Me}C$, $d^{4\text{-}Et}C$, $d^{4\text{-}Pro}C$, $d^{4\text{-}Propargyl}C$ and $d^{4\text{-}Allyl}C$ were prepared from modified nucleosides obtained by activation of the C4 position of 2'-deoxyuridine treated with phosphorus oxychloride in the presence of triazole followed by aminolysis to afford the expected alkyl derivatives[10]. The absorption studies of the duplexes were performed with 2 µM of each oligonucleotides in 10^{-2} M sodium cacodylate pH 7 buffer containing 1 M NaCl and 2 10^{-4} EDTA on a UVIKON 941 cell changer spectrophotometer. The melting temperatures (Tm) were taken as the temperature corresponding to the half-dissociation of the complex.

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

Duplex			Tm ± 1 °C	Δ Tm/2	Δ Tm/1
n°	X	Y			
1	A	T	34.0	- 4.5	
2	G	C	38.5		+ 4.5
3	G	$^{5\text{-Me}}$C	42.5	+ 4.0	+ 8.5
4	G	$^{5\text{-Br}}$C	43.0	+ 4.5	+ 9.0
5	G	$^{4\text{-Me}}$C	34.0	- 4.5	0
6	G	$^{4\text{-Et}}$C	32.5	- 6.0	- 1.5
7	G	$^{4\text{-Pro}}$C	26.5	- 12.0	- 7.5
8	G	$^{4\text{-Allyl}}$C	27.0	- 11.5	- 7.0
9	G	$^{4\text{-Propargyl}}$C	27.0	- 11.5	- 7.0
10	G^{7C}	C	36.5	-2.0	+ 2.5
11	G^{7C}	$^{5\text{-Me}}$C	39.0	+ 0.5	+ 5.0
12	G^{7C}	$^{5\text{-Br}}$C	42.0	+ 3.5	+ 8.0
13	G^{7C}	$^{4\text{-Me}}$C	30.5	- 8.0	- 3.5
14	G^{7C}	$^{4\text{-Allyl}}$C	27.0	- 11.5	- 7.0
15	I	C	29.0	- 9.5	- 5.0
16	I	$^{5\text{-Me}}$C	32.5	- 6.0	- 1.5
17	I	$^{5\text{-Br}}$C	35.0	- 3.5	+ 1.0
18	I	$^{4\text{-Me}}$C	25.0	- 13.5	- 9.0
19	I	$^{4\text{-Allyl}}$C	18.0	- 20.5	- 16

Table : Melting temperatures of duplexes (2 μM strand concentration) in 10^{-2} M sodium cacodylate pH 7 buffer containing 0.1 M NaCl and 2 10^{-4} M EDTA and Δ Tm (°C).

The results of the hybridization study show that thermal stabilities of the duplexes are in agreement with the expected values. Thus for duplexes involving the GC (**2**) or d^{7C}GC (**10**), or IC (**15**) base pairs, the replacement of C by $^{5\text{-Me}}$C (**3**, **11**, **16**) or $^{5\text{-Br}}$C (**4**, **12**, **17**) increases the thermal stability of the hybrids, whereas the substitution on the amine function of dC (**5-9**, **13**, **14**, **18**, **19**) has a destabilizing effect. The replacement of G (**2-8**) by d^{7C}G (**10-14**) leads to a slight destabilization while the substitution of G by I (**15-19**) induces a stronger destabilization which could be due to the difference of the hydrogen bond numbers (3 for GC and 2 for IC) as well as to other factors such as the stacking between adjacent nucleic bases and hydration of the complexes. The latter two parameters seem to be important when comparing the stabilities of duplex **1** (Tm_1 = 34°C) and **15** (Tm_{15} = 29°C) having AT and IC base pairs, respectively, which are isomorphous and possess 2 hydrogen bonds. Regarding the thermal stabilitiy of duplexes **3** to **19** owning modified GC base pairs compared to that of duplex **1** involving AT base pairs at the same positions, we can note the following points :

- In the series of duplexes **3** to **9** involving G*C base pairs, duplexes **5** (Tm_5 = 34°C) and **6** (Tm_6 = 32.5°C) which possess the $G^{4\text{-Me}}$C and $G^{4\text{-Et}}$C base pairs, respectively, have thermal stabilities slightly different from that of duplex **1** used as a reference (Tm_1 = 34°C) (it is important to note that the margin of the error on the Tm values is ± 1°C).

- In the series of duplexes **10** to **14** with ^{7C}GC or ^{7C}G*C (with *C = $^{5\text{-Me}}$C, $^{5\text{-Br}}$C, $^{4\text{-Me}}$C and $^{4\text{-Allyl}}$C) base pairs, the Tm values obtained are different from that of duplex **1** involving AT base pairs at the same position.

- In the series of duplexes **15** to **19** involving XY base-pairs equal to IC or I*C, the thermal stabilities of duplexes **16** (Tm_{16} = 32.5°C) and **17** (Tm_{17} = 35°C) built with the base-pairs I^{5Me}C and I^{5Br}C, respectively, are equivalent to that of the duplex **1**.

These preliminary results show that it is possible to find one modified GC base pair whose thermal stability is equivalent to that of the AT base pair. To reach this objective several approaches can be used. The simplest one involves the modification of one of the base pair to reduce the GC base pair stability as in the case of the $G^{4\text{-Me}}$C and $G^{4\text{-Et}}$C. The second more complex approach allowing more possibilities consists in using of two modified bases as in the case of the I^{5Me}C and I^{5Br}C base pairs. The development of this new strategy involves the synthesis of sequences in which every GC base-pair is replaced by the modified base selected to verify that their thermal stability is independent of their base composition. This work is in progress and will be published elsewhere.

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

References

1. Wallace, R. B.; Shaffer, J.; Murphy, R. F.; Bonner, J.; Hirose, T. *Nucleic Acids Res.* **1979**, *6*, 3543-3557.
2. Wallace, R. B.; Jonhson, M. J.; Hirose, T.; Miyake, T.; Kawashima, E. H.; Itakura, K. *Nucleic Acids Res.* **1981**, *9*, 879-894.
3. Orkin, S. H.; Kazazian, Jr. H. H. *Ann. Rev. Genet* **1984**, *18*, 131-171.
4. Tsui, L-C. *Human Mutation*, **1992**, *1*, 197-203.
5. Lane, D. P. *Nature*. **1992**, *358*, 15-16.
6. Hainaut, P.; Soussi, T.; Shomer, B.; Hollstein, M.; Greenblatt, M.; Hovig, E.; Harris, C. C.; Montesano, R. *Nucleic Acids Res.* **1997**, *25*, 151-157.
7. Southern, E. M. international patent WO 89/10977.
8. Drmanac, R.; Labat, I.; Bruckner, I.; Crkvenjakov, R. *Genomics* **1989**, *4*, 114-128.
9. 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.
10. Sung, W. L. *Nucleic Acids Res.* **1981**, *9*, 6139-6151.

(*Received in France* 4 *April* 1997; *accepted* 27 *April* 1997)