Site specific photo-crosslinking of single stranded oligonucleotides by a complementary sequence equipped with an internal photoactive probe

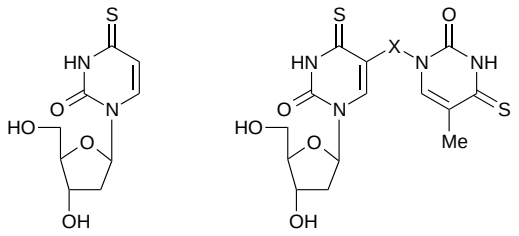
Carole Saintomé,[a,b] Pascale Clivio,[a] Alain Favre,[b] Jean-Louis Fourrey*[a] and Philippe Laugâa[b]

[a] *Institut de Chimie des Substances Naturelles, CNRS 91198, Gif-sur-Yvette Cedex, France*
[b] *Institut Jacques Monod, CNRS-Université Paris VII, 2 Place Jussieu, 75251 Paris Cedex 05, France*

The photo-crosslinking behaviour of oligonucleotide constructs, incorporating photoactive residues 1–5 at a defined position, has been examined in the presence of their DNA (14-mer) and RNA (15-mer) complementary targets using a two-fold excess; it revealed a moderate but promising irreversible binding efficiency, the results of which could be anticipated by molecular mechanics in the case of 4.

There are many instances in which oligodeoxynucleotides (ODN) or nuclease resistant oligonucleotide analogues, conjugated with chemical or photochemical reagents, are utilized as biochemical tools.[1,2] In nucleic acid studies, although simple oligonucleotide binding may not be sufficient to disrupt nucleic acid functions, inactivation will certainly occur if, after specific recognition, a modified oligonucleotide probe has the capacity to undergo irreversible covalent photo-crosslinking to its target.[3] In such probes, a photoactive label is usually introduced at their 5′-terminus, excepted for some examples with psoralens whose applications are specific of AT sites.[4]

Herein, we have examined the photochemical behaviour towards their DNA and RNA targets of strictly complementary constructs having a photoactive nucleoside residue which can be easily placed at a defined position within the sequence. Namely, we have compared the light-induced reactivity of a series of oligonucleotides, containing 4-thiodeoxyuridine (s^4dU) **1** with that of their analogues containing the recently synthesized nucleosides **2**–**5** (Fig. 1).[5] The latter have a chain of variable length at the C-5 position of 2′-deoxyuridine which permits the attachment at its other extremity of N^1-carboxymethyl-4-thiothymine. It is known that, upon light excitation (UVA, $\lambda_{max} \simeq 330$ nm), such sulfur-substituted pyrimidines (4-thiouracil and 4-thiothymine) undergo covalent bond formation with all the current nucleic bases with a higher preference for thymine.[6,7] In this study, nucleoside **1** and the modified analogues **2**–**5** were incorporated at position 1 (for nucleosides **1** and **4**) or position 7 according to a known synthetic and purification procedure[8] by means of their corresponding phosphoramidites,[5] of 14-mer oligonucleotides designated **Oi**-1 and **Oi**-7 (**i** corresponding to nucleosides **1**–**5** (Fig. 1). Oligonucleotides **Oi**-1 and **Oi**-7 were irradiated in two-fold excess in the presence of their RNA or DNA complements, designated **OR** or **OD** respectively, which had been ^{32}P-labelled at their 5′-end (Table 1). Denaturing polyacrylamide gel electrophoresis (PAGE) showed formation of interstrand photoproducts manifested by the appearance of slow moving species on the gels, the yields of which were quantified using a phosphorimager. In contrast, there were no such reactions in the dark or with an oligonucleotide sequence unable to anneal to **Oi**-1 and **Oi**-7.



i = **1** (s^4dU)

i = **2**, **X** = $(CH_2)_2CONH(CH_2)_2NHCOCH_2$
i = **3**, **X** = $(CH_2)_2CONH(CH_2)_4NHCOCH_2$
i = **4**, **X** = $NH(CH_2)_2NHCOCH_2$
i = **5**, **X** = $C{\equiv}CCH_2NHCOCH_2$

Oi-1 5′d(N_i G C C C G T C T G T T G T)3′ with **i** = **1** (s^4dU) or **4**

Oi-7 5′d(T G C C C G N_i C T G T T G T)3′ with **i** = **1** (s^4dU), **2**, **3**, **4** or **5**

OR 3′A C G G G C A G A C A A C A G5′

OD 3′d(A C G G G C A G A C A A C A)5′

Fig. 1 Modified nucleosides **i** (i = **1**–**5**) and oligonucleotides **Oi**-1 and **Oi**-7 incorporating **i** at either position 1 or 7 (N_i corresponds to nucleoside **i**). **OR** and **OD** designate the nucleotide sequences of their RNA (15-mer) and DNA (14-mer) complement, respectively. **OR** was obtained by transcription, and contains an extra G in 5′ which gives efficient initiation.

Oligonucleotide **O1**-1 having **1**, the s^4dU photoreactive label, placed at its 5′-extremity, exhibited a high photo-crosslinking capacity with both its DNA or RNA target, albeit with a lower efficiency in the latter case (55 *vs*. 33%). However, its analogue **O1**-7 having s^4dU at position 7 within the sequence yielded weak (0.5% with DNA) or no detectable (RNA) crosslinks. This indicates that, for a crosslinking reaction to occur, the photoreactive s^4dU should be placed at the 5′-dangling end of the duplex and it suggests that the conformational flexibility at this extremity might be higher for a type B DNA helix than for a type A RNA–DNA helix.

When the extended analogue **4** was placed at the 5′-end of the probe, to give **O4**-1, an efficient cross-linking reaction with the complementary target was observed as in the case of **1**. Again, compared to its RNA analogue, the DNA target proved to be more reactive (21 *vs*. 28%). In the **OR** case, the reaction was very selective and covalent bond formation took place with its complementary residue A15, exclusively (for mapping see below).

Table 1 Crosslinking efficiency, relative yield and position of crosslinks for **Oi**-1 and **Oi**-7

	Total cross-linking efficiency (%)		Relative yields (%) and positions of crosslinks with **OR**				
Probe	**OD**	**OR**	C3	C6	C10	G11	A15
O1-1	53	33					100
O4-1	28	21					100
O1-7	0.5	0					
O2-7	20	8		85	6	7	
O3-7	15	6		88		12	
O4-7	10	3.5		74[a]			
O5-7	22	18	40	60			

[a] Other crosslinks were not characterized due to their instability.

However, in sharp contrast with **1** at position 7 (Fig. 2), when the extended 2′-deoxyuridine residues **2**–**5** were placed within the sequence at this position, covalent crosslinking occurred with both their complementary **OD** and **OR** (Table 1). As for labels **1** and **4** at position 1, the crosslinking yields are better in the DNA series, varying from 10 to 22%. These moderate yields are not unexpected since the complementary sequences **OD** (DNA) and **OR** (RNA), containing the less reactive A, C and G residues compared to T or U, were chosen intentionally to be in the less favourable conditions for crosslinking.[6] On the other hand, protrusion of the label in the major groove of the double helix, which has the possibility to undergo covalent bonding with proximate intra- or inter-strand residues, could also account for these moderate yields. To examine the relative efficiency of the latter process, 5′ ^{32}P-labelled **O5**-7 was first irradiated alone. Denaturating PAGE revealed the formation of intrastrand crosslinked species (60% yield) migrating faster than **O5**-7. When a similar experiment was performed with 5′ ^{32}P-labelled **O5**-7 hybridized to **OR** (two fold excess), it showed a considerable decrease of the fast migrating species (30% yield) together with the appearance of a retarded crosslinked species corresponding to interstrand bridges (22% yield) as revealed by their limited digestion by RNase T1.[9] Interestingly, experiments with **OD** exhibited the same pattern of inter- and intra-strand bridges. Mapping of the exact sites for crosslink formation was carried out using limited alkaline hydrolysis[7c] of the complexes obtained with 5′-end labelled RNA (Fig. 2). The results, given in Table 1, show a preference (74 to 100%) for the residues on the 5′-side of the **OR** target with C6 being preferred (60 to 88%).

Finally, these experimental observations were compared with those predicted by molecular mechanics. A study was undertaken in the case of the **O4**-7–**OR** hybrid in order to determine which residues might participate in inter- and intra-strand photoproduct formation. Model building and computations were performed on a Silicon Graphics Iris 4D/35 equipped with an Evans and Sutherland PS390 graphic device using the SYBYL programme package (Tripos Associates, Saint Louis, Mo). The **O4**-7–**OR** hybrid was built according to the NMR data of Salazar *et al*.,[10] *i.e*. with a C(3′)-*endo* RNA strand and a O(4′)-*endo* DNA strand. The conformational space accessible to the s^{4}T moiety of **4** was explored *via* a systematic search performed along the torsional angles of the linker X (Fig. 1). Residues potentially available for photoaddition were determined through selection of the conformations satisfying both following criteria: (i) a distance of less than 6 Å between the midpoint of the C=S bond of the 4-thiothymine moiety and that of the C(6)=C(5) bond[7g] of the proximal pyrimidines or the midpoint of the C(8)=N(7) bond[11] of the proximal purines, and (ii) an angle between these bonds in the 0 ± 30° (parallel) or 180 ± 30° (antiparallel) ranges. It was found that about 30% of the selected conformations would yield intrastrand crosslinks, while 70% of the selected conformations would yield interstrand crosslinks with residues A5, C6, A7 and G8. When the selection criteria for bond distances and angles were narrowed to 0–4 Å and 0–20° ranges, respectively, the most properly oriented residues appeared to be C6 (50%) and to a lesser extent A7 (21%) and G8 (29%). It is noteworthy that in the case of C6, the plane of the s^{4}T ring is parallel to that of C6, whereas in the case of A7 and G8 it is perpendicular to those of A7 and G8. This might account for the fact that interstrand crosslinking has been detected with a higher preference for C6.

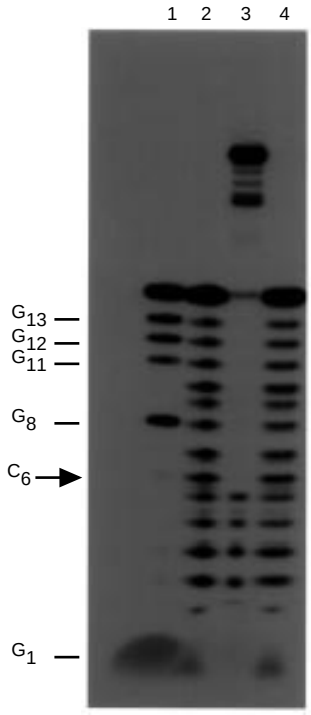


Fig. 2 Identification of RNA residues involved in crosslinking. Sequencing gel of free and crosslinked **OR**. Lane 1: partial digestion of free **OR** with RNase T1; lanes 2 and 4: limited alkaline hydrolysis of **OR**; lane 3: limited alkaline hydrolysis of the major **OR**–**O4**-7 photoadduct. The spots corresponding to G positions are marked. The arrow points to nucleotide C6 involved in the crosslink.

We thank ACC-SV (MST) for financial support.

References

1 U. Englisch and D. H. Gauss, *Angew. Chem., Int. Ed. Engl*., 1991, **30**, 613.
2 C. Hélène, T. Le Doan and N. T. Thuong, in *Photochemical Probes in Biochemistry*, ed. P. E. Nielsen, Kluwer, Dordrecht, 1989, p. 219.
3 D. G. Knorre and V. V. Vlassov, *Genetica*, 1991, **85**, 53.
4 For porphyrin labelling in 5′, see L. Mastruzzo, A. Woisard, M. D. F. Ma, E. Rizzarelli, A. Favre and T. Le Doan, *Photochem. Photobiol*., 1994, **60**, 316; D. Magda, M. Wright, R. A. Miller, J. L. Sessler and P. I. Sansom, *J. Am. Chem. Soc*., 1995, **117**, 3729. For psoralen labelling in 5′ or 3′, see J. M. Kean, A. Murakami, K. R. Blake, C. D. Cushman and P. S. Miller, *Biochemistry*, 1988, **27**, 9113; U. Pieles and U. Englisch, *Nucleic Acids Res*., 1989, **17**, 285; E. H. Chang, P. S. Miller, C. Cushman, K. Devadas, K. F. Pirollo and P. O. Ts'o, *Biochemistry*, 1991, **30**, 8283; J. Woo and P. B. Hopkins, *J. Am. Chem. Soc*., 1991, **113**, 5457; H. E. Johansson, G. J. Belshman, B. S. Sproat and M. W. Hentze, *Nucleic Acids Res*., 1994, **22**, 4591; G. Godard, J.-C. François, I. Duroux, U. Asseline, N. Thuong, C. Hélène, C. and T. Saison-Behmoaras, *Nucleic Acids Res*., 1994, **22**, 4789. For internal labelling with psoralens, see H. B. Gamper, G. D. Cimino and E. Hearst, *J. Mol. Biol*., 1987, **197**, 349; U. Pieles, B. S. Sproat, P. Neuner and F. Cramer, *Nucleic Acids Res*., 1989, **17**, 8967; M. A. Reynolds, T. A. Beck, R. I. Hogrefe, A. McCaffrey, L. J. Arnold and M. M. Vaghefi, *Bioconjugate Chem*., 1992, **3**, 366.
5 C. Saintomé, P. Clivio, J.-L. Fourrey, A. Woisard and A. Favre, *Tetrahedron Lett*., 1994, **35**, 873. Full characterization of these photoactive nucleosides will be described in a forthcoming article.
6 A. Favre, Y. Lemaigre Dubreuil and J.-L. Fourrey, *New J. Chem*., 1991, **15**, 593.
7 (*a*) P. Wollenzien, A. Expert-Bezançon and A. Favre, *Biochemistry*, 1991, **30**, 1785; (*b*) E. J. Sontheimer and J. A. Steitz, *Science*, 1993, **262**, 1989; (*c*) V. D. Dos Santos, A.-L. Vianna, J.-L. Fourrey and A. Favre, *Nucleic Acids Res*., 1993, **21**, 201; (*d*) D. Dos Santos, J.-L. Fourrey and A. Favre, *Biochem. Biophys. Res. Commun*., 1993, **190**, 377; (*e*) A. Woisard, J.-L. Fourrey and A. Favre, *J. Mol. Biol*., 1994, **239**, 366; (*f*) M. Tanaka, J.-L. Fourrey and A. Favre, *New J. Chem*., 1995, **19**, 469; (*g*) A. Favre and J.-L. Fourrey, *Acc. Chem. Res*., 1995, **28**, 375; (*h*) Z. Wang and T. M. Rana, *Biochemistry*, 1996, **35**, 6491.
8 P. Clivio, J.-L. Fourrey, J. Gasche, A. Audic, a. Favre, C. Perrin and A. Woisard, *Tetrahedron Lett*., 1992, **33**, 65.
9 J. M. D'Alessio, in *Gel Electrophoresis of Nucleic Acids*, ed. D. Richmond and B. D. Hames, 1991, p. 173.
10 M. Salazar, O. Y. Fedorov, J. M. Miller, N. S. Ribeiro and B. R. Reid, *Biochemistry*, 1993, **32**, 4207.
11 C. Saintomé, P. Clivio, A. Favre, J.-L. Fourrey and C. Riche, *J. Am. Chem. Soc*., 1996, **118**, 8142.

Received, 13th September 1996; Com. 6/06330F