Structural Probing of Small Endonucleolytic Ribozymes in Solution Using Thio-Substituted Nucleobases as Intrinsic Photolabels

ALAIN FAVRE[†] AND JEAN-LOUIS FOURREY*,[‡]

Institut Jacques Monod, CNRS-Université Paris 7, Tour 43, 2 Place Jussieu, 75251 Paris Cedex 05, France, and Institut de Chimie des Substances Naturelles, CNRS, 91198 Gif sur Yvette Cedex, France

Received October 19, 1994 (Revised Manuscript Received March 21, 1995)

Nowadays the growing interest for RNA three-dimensional structural studies can be ascribed to the recent observations on the capacity of these molecules to exercise an increasing number of functions.[1] In the last decade, the discovery of RNA sequences endowed of catalytic activity was received as a revolution in molecular biology. In prokaryotic cells RNase P, which participates in the maturation of tRNA precursors, is a ribonucleoprotein complex, the catalytic activity of which is carried out by the RNA moiety.[2] Class I and class II introns are removed, with exons rejoining, in the absence of any proteins.[3] The multimeric forms of some plant satellite RNA viruses are converted into active monomers by an autolytic process.[4] This observation has led to the design of small ribozymes (termed hammerhead or hairpin) which, in the presence of Mg^{2+}, manifest a trans cleavage activity generating 2′,3′-cyclic phosphate and 5′-OH termini.[5] It is common sense that the formation of the active site of the latter requires a precise three-dimensional folding.

Tertiary structure determination of tRNA proved that X-ray crystallography and to some extent NMR are potentially able to reveal tertiary interactions at atomic resolution within RNA.[6] But their application to natural RNA is so far limited. Within single-stranded nucleic acids chemical probing and thermodynamical analysis can provide direct information regarding the secondary structure. Site-directed mutagenesis and atomic mutagenesis are also very helpful in the exploration of the structure–function relationship. None of these approaches, however, allows a direct identification of the tertiary contacts.[7] In this respect, it is worthy of note that, when an abundance of phylogenic information is available, it becomes possible to construct a plausible model of RNA 3D structure. An example of this approach has been proposed in the case of class I introns.[8]

Obviously, the formation of cross-links between spatially proximate residues in the folded structure, but separated by a stretch of nucleotides in the sequence, can be extremely powerful to detect tertiary interactions. This can be achieved in low yield by direct 254 nm irradiation[9] or by using extrinsic chemical or photochemical probes such as psoralens.[10] A conceptually better alternative is the introduction of a photoactivatable nucleotide analog, at a defined position of the sequence, especially if its incorporation causes little steric perturbation.[11] Optimally this probe, upon selective photoexcitation, should give rise to stable photoadducts with all current nucleosides, and its reactivity within double-stranded regions should be kept minimal in order to only unravel contacts due to tertiary folding. While a variety of nucleic base derivatives[11a] have already been used, thio-substituted analogs present a number of remarkable advantages that make them unique intrinsic photoaffinity probes to study RNA structure.[11b] Interestingly, they are less bulky than the known monofunctional cross-linking agents such as azido-substituted or azidophenacyl-substituted nucleosides.[12] Moreover, they are stable in the dark and can be selectively photoactivated to give a highly reactive excited state able to interact with any nucleic acid bases.[11b,13] A very illustrative example is the case of 4-thiouridine (s^4U) located at position 8 of a number of *Escherichia coli* tRNA. Upon selective irradiation, this residue, which strongly absorbs UVA light (λ_{max} ~ 330 nm), undergoes a cycloaddition reaction with the 5,6-double bond of cytidine at position 13 to form

A. Favre graduated from Ecole Normale Supérieure in Paris and obtained his Doctorat d'Etat in biochemistry under the direction of A. M. Michelson. Structural and photochemical studies of nucleic acids are among his main research interests, which are focused on the structure–activity relationships of small endonucleolytic ribozymes. He is now professor at Université Paris VI and Head of the Département Organisation and Développement du Génome at Jacques Monod Institute in Paris.

J.-L. Fourrey received his engineer degree from Ecole Nationale Supérieure de Chimie de Paris. He obtained his Doctorat d'Etat in chemistry under the guidance of J. Poionsky. His research interests include synthetic methodological developments to be applied to the synthesis of natural products. His main activity is in the domain of nucleoside and oligonucleotide chemistry and photochemistry. He is presently Directeur de Recherche at Institut de Chimie des Substances Naturelles in Gif sur Yvette.

[†] Institut Jacques Monod.

[‡] Institut de Chimie des Substances Naturelles.

(1) Gesteland, R. F.; Atkins, J. F. *The RNA world*; Cold Spring Harbor Press: New York, 1993. Inouye, M.; Delihas, N. *Cell* **1988**, *53*, 5–7.

(2) Guerrier-Takada, C.; Altman, S. *Science* **1984**, *223*, 285–286.

(3) Cech, T. R. *Annu. Rev. Biochem.* **1990**, *59*, 543–568. Cech, T. R. *Science* **1987**, *236*, 1532–1539. Cech, T. R.; Bass, B. L. *Annu. Rev. Biochem.* **1986**, *55*, 543–568.

(4) Symons, R. H. *Annu. Rev. Biochem.* **1992**, *61*, 641–671.

(5) Haseloff, J.; Gerlach, W. L. *Nature* **1988**, *334*, 585–591. Uhlenbeck, O. C. *Nature* **1987**, *328*, 596–600.

(6) Rich, A.; Rajbhandary, U. L. *Annu. Rev. Biochem.* **1976**, *45*, 805–860.

(7) (a) Grasby, J. A.; Gait, M. J. *Biochimie* **1994**, *76*, 1223–1234. (b) Bratty, J.; Chartrand, P.; Ferbeyre, G.; Cedergren, R. *Biochim. Biophys. Acta* **1993**, *1216*, 345–349.

(8) Michel, F.; Westhof, E. *J. Mol. Biol.* **1990**, *216*, 581–606.

(9) Cadet, J.; Vigny, P. *Bioorganic Photochemistry: Photochemistry and the Nucleic Acids*; Morrison, H., Ed.; J. Wiley & Sons: New York, 1990; Vol. 1, pp 1–272.

(10) Shi, Y.; Lipson, S. E.; Chi, D. Y.; Spielman, H. P.; Monforte, J. A.; Hearst, J. E. *Ibid.*; pp 341–378.

(11) (a) Willis, M. C.; Hicke, B. J.; Uhlenbeck, O. C.; Cech, T. R.; Koch, T. H. *Science* **1993**, *262*, 1255–1257. (b) Favre, A. *Bioorganic Photochemistry: Photochemistry and the Nucleic Acids*; Morrison, H., Ed.; J. Wiley & Sons: New York, 1990; Vol. 1, pp 317–378.

(12) Hanna, M. M. *Methods Enzymol.* **1989**, *180*, 383–409.

(13) Favre, A.; Lemaigre Dubreuil, Y.; Fourrey, J.-L. *New J. Chem.* **1991**, *15*, 593–599.

0001-4842/95/0128-0375$09.00/0

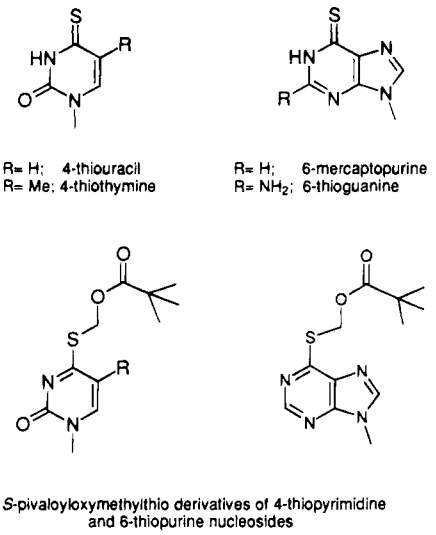

Figure 1.

the stable photoadduct Pdo(4−5)Cyd in native tRNA.[14] Because of this reaction, the tertiary interaction between residues 8 and 13 was demonstrated before it could be confirmed by X-ray analysis of the tridimensional structure of tRNA.[15]

However, the full development of a photoaffinity labeling technique based on thio-substituted nucleobase photoreactivity such as 4-thiouracil or its nonnatural congeners 4-thiothymine, 6-mercaptopurine, and 6-thioguanine (Figure 1) requires (i) efficient methods for the introduction of such modified residues in oligonucleotides by chemical and/or enzymatic synthesis and (ii) a good knowledge of their photoreactions, including photoproduct identification, when incorporated into oligomers. Indeed, the formation of Pdo(4−5)Cyd presumably through a thietane intermediate allowed definition of the respective orientation of residues 8 and 13 in tRNA prior to the photoreaction.[14]

Synthesis of Oligonucleotides Containing Thio-Substituted Nucleobases

Chemical Synthesis. Many efforts have been undertaken to synthesize, by either the phosphoramidite[16] or the H-phosphonate[17] method, modified oligonucleotides having nucleobases in which oxygen is replaced by sulfur.[18] Indirect routes involving postsynthetic modifications have been considered; one route used 4-triazolopyrimidinone chemistry to give 4-thiopyrimidinone.[18n] More convenient procedures, starting from already sulfur-modified nucleosides, have been proposed. They differ by the recommended choice of protection at the sulfur position. Indeed, because of its high nucleophilicity, this function might interfere during the phosphitylation step; moreover, it might also be very sensitive to oxidation by iodine, which is necessary for the conversion of P(III) into P(V). We have proposed the *S*-(pivaloyloxy)methyl group, which can be easily introduced and, in contrast to the other suggested groups, can be removed without modification of the protocol which serves in machine synthesis of oligonucleotides.[18a-d]

Enzymatic Synthesis. Incorporation of s^4UTP (or its 4-thiothymidine analog, s^4TTP) into polynucleotides is easily achieved using appropriate polymerase systems. For example, the exonuclease free fragment of DNA polymerase, but not T4 polymerase, incorporates s^4TTP into DNA, albeit with some unintended primer extension.[19] Although *in vitro* synthesis of s^4U containing RNA can be performed with yeast RNA polymerase II,[20] it is in general achieved using T7 RNA polymerase. Thus, short transcripts are obtained with single-strand DNA templates that are double-stranded at the promoter site.[21] At low values of the s^4UTP/UTP ratio, the analog is incorporated at a rate 4−5 times slower than UTP and is found randomly distributed among all the available positions.[22] However, s^4U can be introduced specifically at the 5′-end of any transcript since the dinucleotide s^4UpG can be used in place of GTP in T7 polymerase transcription initiation. This procedure, together with the possibility of ligating two RNA strands by means of T4 DNA ligase and a bridging oligonucleotide,[23] allows the specific introduction of s^4U residues at any s^4UpG site within an RNA chain.[24] Chemical synthesis of modified oligonucleotides combined with enzymatic ligation should allow the incorporation of thionucleotides at any selected position of an RNA or DNA sequence.

Photochemistry of Oligonucleotides Containing Thio-Substituted Nucleobases

The full scope of thionucleoside photo-cross-linking ability was determined in oligonucleotide model systems. Hence, in aqueous solution the trinucleotide d(s^4UCC) gave stable photoadducts with all the current nucleosides at concentrations $>10^{-3}$ M in the decreasing order of reactivity T ≫ U ~ A > C ~ G.[13] The structures of a number of these adducts were elucidated in the case of pyrimidine deoxydinucleotides incorporating 4-thiodeoxyuridine (s^4dU) or 4-thio-

(14) Favre, A.; Roques, B. P.; Fourrey, J.-L. *FEBS Lett.* **1972**, *24*, 209−214. Bergstrom, D. E.; Leonard, N. J. *Biochemistry* **1972**, *11*, 1−9. Favre, A.; Michelson, A. M.; Yaniv, M. *J. Mol. Biol.* **1971**, *58*, 367−379. Favre, A.; Michelson, A. M.; Yaniv, M. *Biochem. Biophys. Res. Commun.* **1969**, *37*, 266−271.

(15) Kim, S. H.; Quigley, G. J.; Suddath, F. L.; McPherson, A.; Sneden, D.; Kim, J. J.; Weinzierl, J.; Rich, A. *Science* **1973**, *179*, 364−367.

(16) Beaucage, S.; Iyer, R. P. *Tetrahedron* **1993**, *49*, 6123−6194. Beaucage, S.; Iyer, R. P. *Ibid.* **1992**, *48*, 2223−2311.

(17) Froelher, B.; Ng, P.; Matteuci, M. *Nucleic Acids Res.* **1986**, *14*, 5359−5407. Garegg, P. J.; Redberg, T.; Stawinski, J.; Stromberg, R. *Chem. Scr.* **1985**, *25*, 280−282.

(18) (a) Clivio, P.; Fourrey, J.-L.; Favre, A. *J. Chem. Soc., Perkin Trans. 1* **1993**, 2585−2590. (b) Clivio, P.; Fourrey, J.-L.; Gasche, J.; Favre, A. *Tetrahedron Lett.* **1992**, *33*, 69−72. (c) Clivio, P.; Fourrey, J.-L.; Gasche, J.; Audic, A.; Favre, A.; Perrin, C.; Woisard, A. *Ibid.* **1992**, *33*, 65−68. (d) Clivio, P.; Fourrey, J.-L.; Gasche, J.; Favre, A. *J. Chem. Soc., Perkin Trans. 1* **1992**, 2383−2388. (e) Nikiforov, T. T.; Connolly, B. A. *Nucleic Acids Res.* **1992**, *20*, 1209−1214. (f) Nikiforov, T. T.; Connolly, B. A. *Tetrahedron Lett.* **1992**, *33*, 2379−2382. (g) Xu, Y.-Z.; Zheng, G.; Swann, P. F. *J. Org. Chem.* **1992**, *57*, 3839−3845. (h) Xu, Y.-Z.; Zheng, G.; Swann, P. F. *Tetrahedron* **1992**, *48*, 1729−1740. (i) Sudhakar, T.; Jayaraman, K.; Durland, R. H.; Revankar, G. R. *Tetrahedron Lett.* **1992**, *33*, 7651−7654. (j) Coleman, R. S.; Siedlecki, J. M. *J. Am. Chem. Soc.* **1992**, *114*, 9229−9230. (k) Nikiforov, T. T.; Connolly, B. A. *Tetrahedron Lett.* **1991**, *32*, 3851−3854. (l) Ishikawa, T.; Yoneda, K.; Tanaka, K.; Fuji, K. *Bioorg. Med. Chem. Lett.* **1991**, *1*, 523−526. (m) Connolly, B. A.; Newman, P. C. *Nucleic Acids Res.* **1989**, *17*, 1209−1214. (n) Fourrey, J.-L.; Gasche, J.; Fontaine, C.; Guittet, E.; Favre, A. *J. Chem. Soc., Chem. Commun.* **1989**, 1334−1336.

(19) Bartholomew, B.; Braun, B. R.; Kassavetis, G. A.; Geiduschek, E. P. *J. Biol. Chem.* **1994**, *269*, 18090−18095.

(20) Sheng, N.; Mougey, E. B.; Kelly, S.; Dennis, D. *Biochemistry* **1993**, *32*, 2248−2253.

(21) Milligan, J. F.; Uhlenbeck, O. C. *Methods Enzymol.* **1989**, *180*, 42−50.

(22) Lemaigre-Dubreuil, Y.; Expert-Bezançon, A.; Favre, A.; *Nucleic Acids Res.* **1991**, *19*, 3653−3660.

(23) Moore, M. J.; Sharp, P. A. *Science* **1992**, *256*, 992−997.

(24) Sontheimer, E. J.; Steitz, J. A. *Science* **1993**, *262*, 1989−1996.

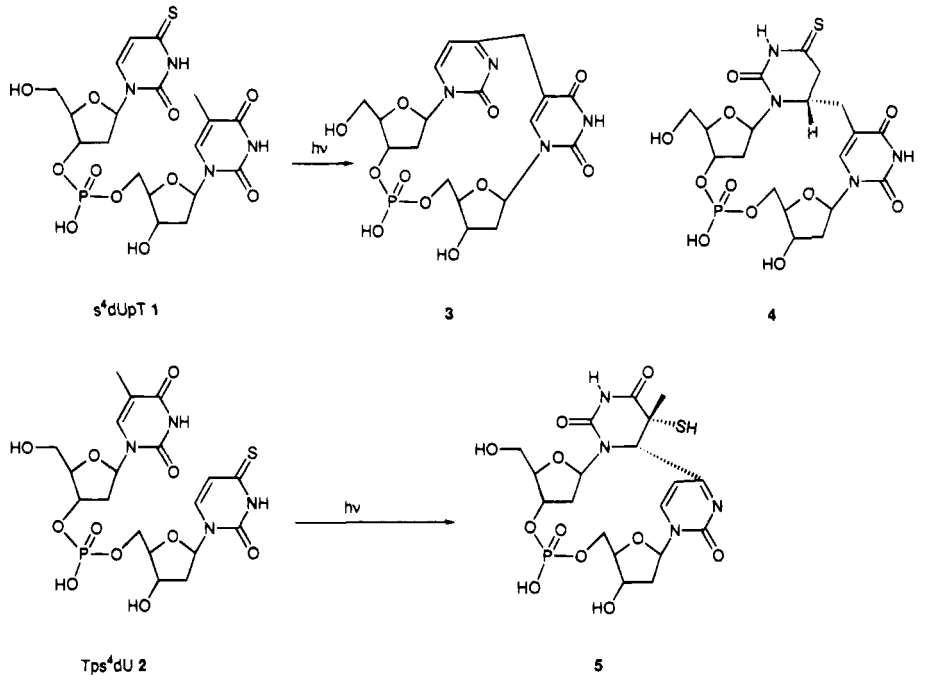

Figure 2.

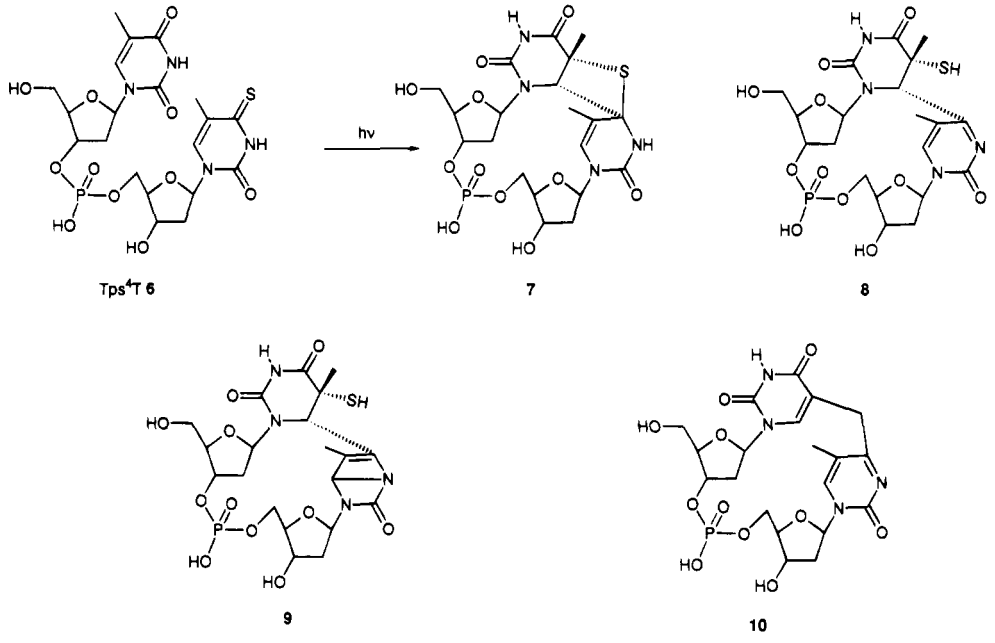

Figure 3.

thymidine (s^4T). Starting from s^4dUpT (**1**) and Tps^4dU (**2**) an unexpected sequence-dependent photochemistry was discovered (Figure 2). While s^4dUpT (**1**) yielded two new photoproducts, **3** and **4**, by a radical mechanism, Tps^4dU (**2**) gave the (6–4) bipyrimidine **5**.[18n,25]

In this reaction the plausible intermediate is a thietane which was indeed observed as the main photoproduct of Tps^4T (**6**) irradiation (Figure 3).[26] In this case thietane **7** was characterized and found to be in equilibrium with the (6–4) photoadduct **8**. The **7** + **8** mixture was further photochemically converted into its Dewar valence isomer **9**. An additional new photoproduct **10**, which was formed by a radical pathway, was also isolated. In order to stop the reaction at the thietane stage and avoid the formation of secondary products such as the Dewar compound, the photochemistry of thymidylyl (3′–5′) N^3-methyl-4-thiothymidine (Tpm^3s^4T, **11**) was examined.[27] Irradiation of **11** gave three compounds, **12**, **13**, and **14**, in 25, 2, and 24% yields, respectively (Figure 4). Interestingly, the minor thietane **13** resulted from a cycloaddition pathway in which the 5′-nucleoside is in the syn conformation while the 3′-nucleoside is in the anti.[27]

The results of this study of the photochemical behavior of the Tps^4T system are important (i) in providing definitive experimental evidence indicating that (6–4) bipyrimidine photoproducts in DNA could

(25) Clivio, P.; Favre, A.; Fontaine, C.; Fourrey, J.-L.; Gasche, J.; Guittet, E.; Laugâa, P. *Tetrahedron* **1992**, *48*, 1605–1616.
(26) Clivio, P.; Fourrey, J.-L.; Gasche, J.; Favre, A. *J. Am. Chem. Soc.* **1991**, *113*, 5481–5483.
(27) Clivio, P.; Fourrey, J.-L.; Gasche, J.; Favre, A. *Tetrahedron Lett.* **1992**, *33*, 1615–1618.

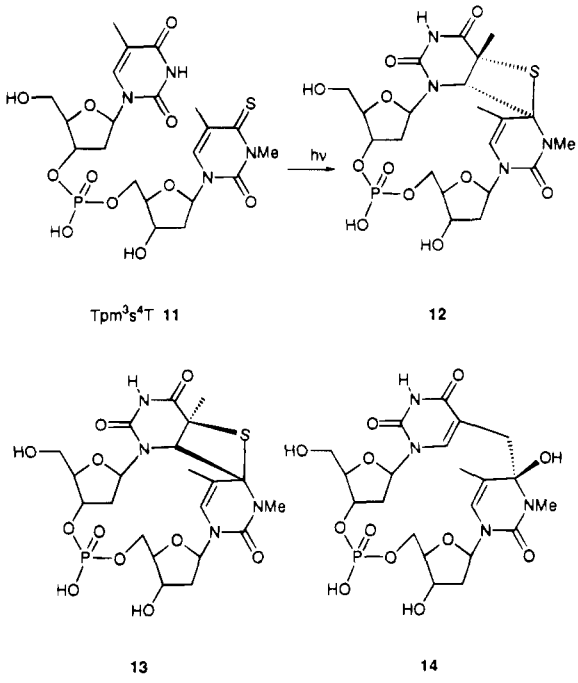

Figure 4.

arise *via* a four-membered-ring intermediate as already postulated in the case of TpT and TpdC[9] and (ii) in establishing the stereochemical course of this cycloaddition reaction.

In continuation of this work, the photoreactivity of (R_p) and (S_p) thymidinyl-3′-methylphosphonate 5′-N^3-methyl-4-thiothymidine (T*p*m^{3}s^{4}T) (**15a,b**) was compared with that of **11**.[28] The rate of photolysis for each dimer was found to decrease in the order **11** > **15a** > **15b**. However, four photoproducts, **16a**−**19a**, could be isolated from irradiated R_p methylphosphonate isomer **15a**, whereas eight photoproducts, **16b**−**19b** and **20**−**23**, have been isolated from the irradiation mixture of the S_p methylphosphonate **15b** (Figure 5). While the (R_p) derived photoadduct **18a** was found to be hydrolytically stable, the corresponding S_p isomer **18b** yielded compounds **20** and **21** resulting from the hydrolysis of the 3′-end glycosidic bond. These observations indicate that s^{4}T photochemistry is sensitive enough to probe the conformation of the methylphosphonate analogs. It can be concluded that (R_p) **15a** is a better mimic of the natural phosphodiester **11** than its diastereoisomer (S_p) **15b**. This is in agreement with the view that oligonucleotide methylphosphonates with predominantly the R_p configuration afford more stable hybrids with complementary DNA than their S_p analogs.[29] Moreover, the study led to the discovery of a new type of photoproduct corresponding to **19**.

One thus observes that, when incorporated in model oligonucleotides, 4-thiouracil and 4-thiothymine give rise after irradiation to highly reactive excited intermediates. These species undergo both cycloaddition and radical reactions leading to stable photoadducts by covalent bond formation. This is in agreement with earlier reports on the photochemical behavior of 4-mercaptopyrimidine and 6-mercaptopurine derivatives.[30,31] Consequently, it is beyond question that thio-substituted residues should be perfectly suitable for photochemical labeling.

Intrinsic Photoaffinity Labeling in the Ribozyme Series

As an illustration, the results obtained in the small endonucleolytic hairpin and hammerhead ribozyme series will be summarized.

Hammerhead Ribozymes. The consensus hammerhead secondary structure found in some small autolytic RNA has allowed the design of ribozymes that are catalytically active in trans. They are made up of three base-paired regions connected by conserved sequences.[5] A number of mutation and substitution studies have underlined the importance for the cleavage activity of the conserved nucleotides and nucleotide functional groups which are present within the central core.[7b]

In order to explore the 3D folding of the hammerhead ribozyme system shown in Figure 6, we first synthesized a noncleavable full deoxy substrate analog dSI containing 6-thiodeoxyinosine (s^{6}dI) at position 16.1, *i.e.*, next to the cleavage site.[32] Irradiation of the radiolabeled ribozyme (R) in the presence of an excess of dSI induced an interstrand cross-link in 25% yield.[33] In this experiment, residue 16.1 was shown to be covalently linked to ribozyme residue U_7. Interestingly the formation of a (6−6) purine−pyrimidine photoproduct was observed in the case of the dinucleotide Tps^{6}dI, suggesting that a similar photoproduct might have been obtained in the ribozyme system. These studies were performed under conditions allowing the cleavage of more than 100 RNA substrates by a single ribozyme molecule (k_{cat} ~ 25 min^{-1} at 37 °C). Furthermore, in a control experiment an appropriate substrate containing s^{6}dI at position 16.1 and a single ribocytidine at position 17 was efficiently cleaved by the ribozyme. These photo-cross-link data suggested that, in the ribozyme domain, positions 16.1 and 7 are proximal.

Subsequently, a systematic investigation was undertaken using s^{4}dU, which more closely mimics the U residues present in the substrate.[34] Thus, full deoxy 14-mer analogs of the substrate (S), dS$_i$, were synthesized (Table 1). In this series each dS$_i$ contained a single s^{4}dU probe at position *i* occupied by U in the parent substrate. When combined with 5′-^{32}P labeled R, each dS$_i$ generated upon irradiation from one to four cross-linked species which could be separated by denaturating gel electrophoresis and analyzed (Table 1). Cross-link formation was absolutely s^{4}dU and light dependent. Furthermore, the finding that a 14-mer control, containing s^{4}dU but unable to anneal to R, did not yield any detectable cross-links together with the results of competition experiments (with S and dS)

(28) Clivio, P.; Fourrey, J.-L.; Szabo, T.; Stawinski, J. *J. Org. Chem.* **1994**, *59*, 7273−7283.
(29) Samstag, W.; Engels, J. W. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 1386−1388. Lesnikowski, Z. J.; Jaworska, M.; Stec, W. J. *Nucleic Acids Res.* **1990**, *18*, 2109−2115.
(30) Fombert, C.; Fourrey, J.-L.; Jouin, P.; Moron, J. *Tetrahedron Lett.* **1974**, 3007−3008. Fourrey, J.-L.; Jouin, P.; Moron, J. *Ibid.* **1974**, 3005−3006. Fourrey, J.-L.; Jouin, P. *Ibid.* **1973**, 3229−3232.
(31) Fourrey, J.-L.; Henry, G.; Jouin, P. *Tetrahedron Lett.* **1976**, 2375−2378. Fourrey, J.-L. *Ibid.* **1976**, 297−300. Fourrey, J.-L.; Jouin, P. *Ibid.* **1973**, 3225−3227.
(32) Hertel, K. J.; Pardi, A.; Uhlenbeck, O. C.; Koizumi, M.; Ohtsuka, E.; Uesugi, S.; Cedergren, R.; Eckstein, F.; Gerlach, W. L.; Hodgson, R.; Symons, R. H. *Nucleic Acids Res.* **1992**, *20*, 3252−3252.
(33) Woisard, A.; Favre, A.; Clivio, P.; Fourrey, J.-L. *J. Am. Chem. Soc.* **1992**, *114*, 10072−10074.
(34) Woisard, A.; Fourrey, J.-L.; Favre, A. *J. Mol. Biol.* **1994**, *239*, 366−370.

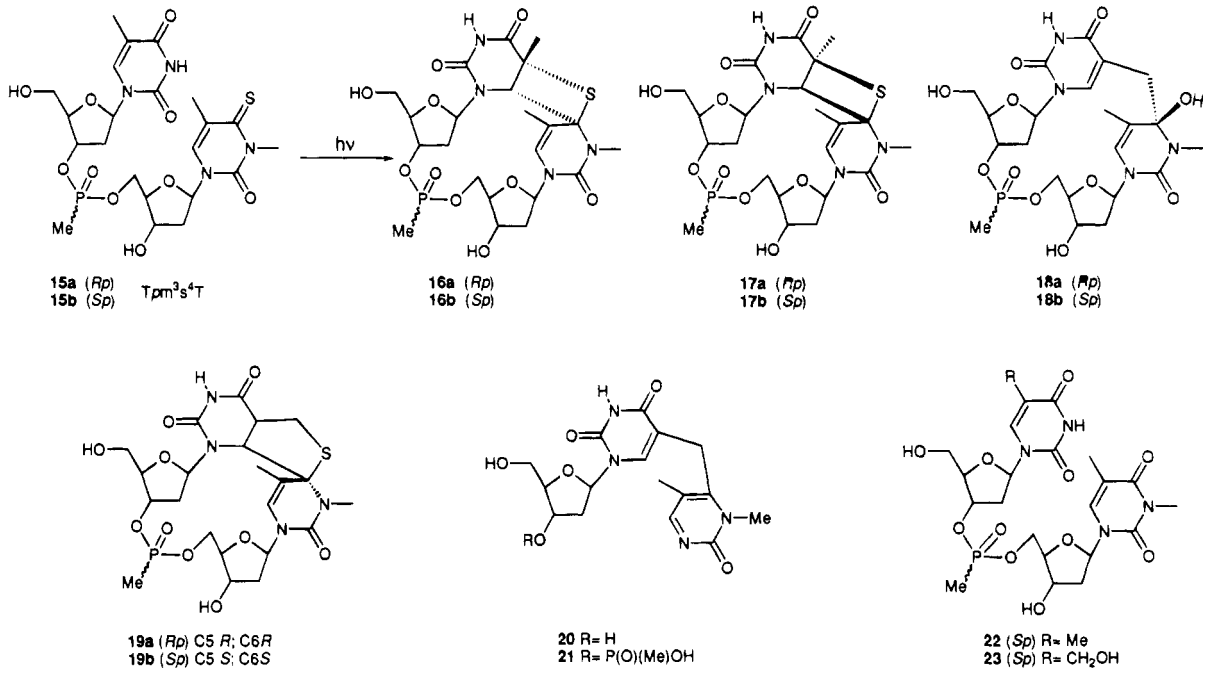

Figure 5.

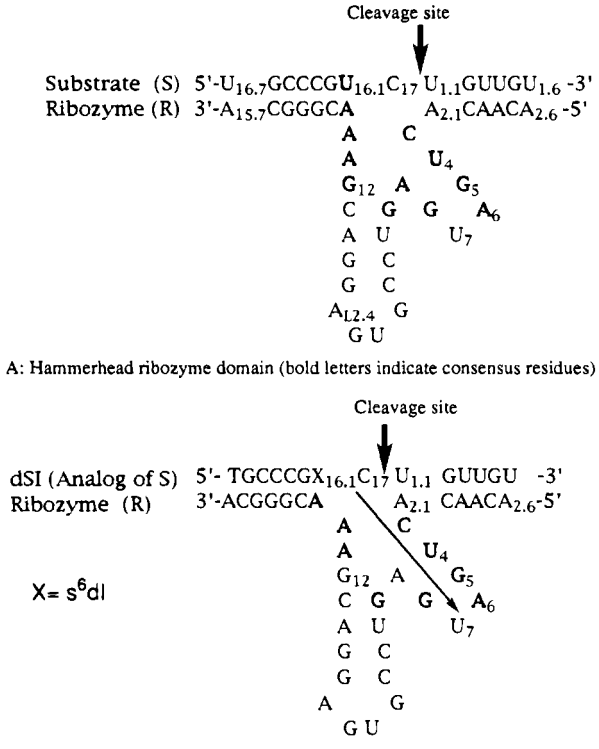

Figure 6.

Table 1. Photo-Cross-Link Formation between Ribozyme (R) and Its Substrate Analogs dS_i and rdS_{17}

	Substrate analogs	total cross-link efficiency (%)	cross-linked residues
$dS_{16.7}$	XGCCCGTCTGTTGT	25	$A_{15.7}$
$dS_{16.1}$	TGCCCGXCTGTTGT	28	U_7 U_4 $A_{2.1}$
$dS_{10.1}$	TGCCCGTCXGTTGT	18	A_6 G_5 U_4 $A_{2.1}$
rdS_{17}	UGCCCGUXUGUGU	12	G_{12} U_7 G_5 U_4
dS_{17}	TGCCCGTXTGTTGT	4.8	G_{12} U_7 G_5 U_4

established that cross-linking occurred within the specific R–dS_i complexes. The obtained data confirm the consensus secondary structure (cross-link between residues 16.7 and 15.7) and provide new insight into the ribozyme domain 3D folding.[34] In particular, positions 16.1 and 1.1 of dS yielded multiple cross-links with the ribozyme 5′-conserved region.

A salient question is whether a full hammerhead ribozyme RNA domain would fold similarly to the corresponding DNA–RNA domain studied herein. Indeed, in other experiments DNA substrates containing a single ribocytidine at the cleavage site were cleaved with reduced rates.[35] In the present case the folding of the DNA–RNA domain appears only slightly affected since (i) a full deoxysubstrate analog which had a ribocytidine at the cleavage position was efficiently cleaved and (ii) very similar cross-link patterns were obtained with either a full deoxy substrate (dS_{17}) or its ribo analog (rdS_{17}) both containing s^4dU at position 17 (Table 1). Having unraveled a number of tertiary interactions, one should ask whether these contacts are related to the interactions occurring during the cleavage step. An experiment was devised in which substitution of A_{14} by U led to a fully inactive mutant ribozyme (R′). The latter was found to cross-link with $dS_{16.7}$, which had s^4dU at its 5′-end, indicating substrate binding capacity with R′. In contrast, interactions within the central core were strongly altered. Thus, $dS_{16.1}$ yielded a single cross-link with U_7 of R′ in low yield (3%) as compared to the multiple cross-links obtained with R (28%). The multiple cross-links generated within the core of the hammerhead domain, as observed with $dS_{16.1}$, dS_{17}, and $dS_{1.1}$, definitely rule out the existence of a single tertiary folded structure in solution (a unique 8–13 cross-link was found in tRNA[14]). On the other hand, the observed cross-links cannot be due to random motion of the conserved residues just held in proximity to the s^4dU label within the secondary structure since cross-links to the A-rich 3′-region were never detected. Therefore, the data indicate that in solution a ri-

(35) Hendry, P.; McCall, M. J.; Santiago, F. S.; Jennings, P. A. *Nucleic Acids Res.* **1992**, *20*, 5737–5741. Yang, J.-H.; Perreault, J. P.; Labuda, D.; Usman, N.; Cedergren, R. *Biochemistry* **1990**, *29*, 11156–11160.

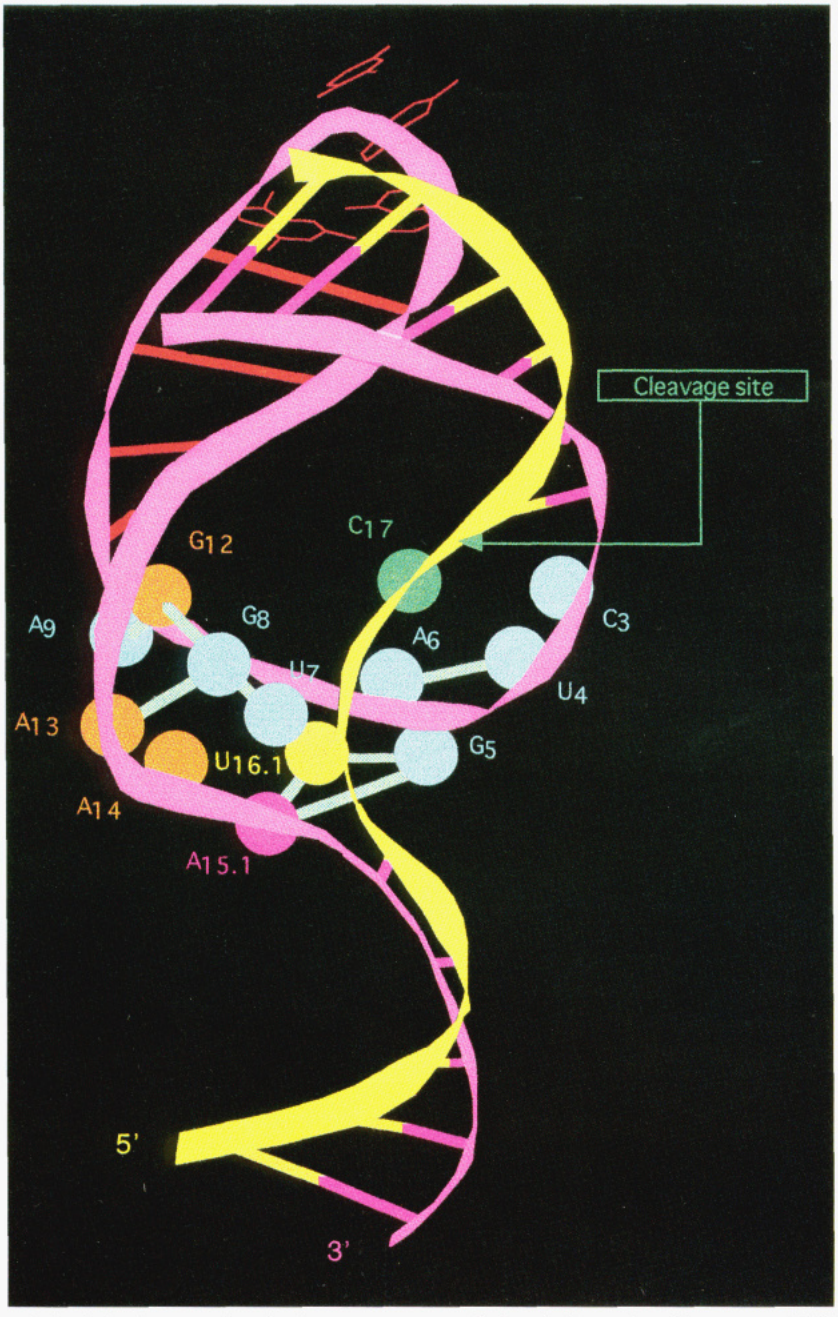

Figure 7. Ribbon 3D model of the hammerhead ribozyme system deduced from the photo-cross-linking data. Yellow ribbon: Substrate. Pink ribbon: Ribozyme. The model indicates the close juxtaposition of helices I and II. The balls represent the conserved residues. Tertiary contacts are represented by bold green lines.

Table 2. Central Core Residues Involved in the Formation of the Photo-Cross-Links[a]

	ribozyme				
substrate	U_4	G_5	A_6	U_7	G_{12}
$U_{16.1}$	*		*	+	
C_{17}	+	+		+	+
$U_{1.1}$	+	*	+		

[a] Plus signs (+) denote the cross-links which were used to construct the 3D model. [b]Asterisks (*) denote the cross-links for which the model does not account.

bozyme domain may experience several closely related folded conformations.

Model of the Hammerhead Ribozyme Domain. On the basis of the cross-links obtained within the ribozyme central core (Table 2), a plausible model has been built by molecular modeling.[36] Essential to this construction was the recent finding that a new long-range cross-link was obtained between position 1.6 of a full deoxy substrate analog and the adenosine residue located in position L2.4 of loop II (Figure 6).[37] Because of this covalent bridge which was formed in significant yield (~4%) under cleavage conditions stems II and III must be placed in close proximity with $U_{1.6}$ facing $A_{L2.4}$. Stems I, II, and III were built as A-type helices. Helix II was extended by a double A_9–G_{12}, G_8–A_{13} mismatch supported by substitution data.[38] On its 3′-end it was then elongated with A_{14} and helix III so as to retain an overall A-type helix structure. Docking of stem I on this structure was guided by the tertiary contact between C_{17} and G_{12} and

(36) Laugâa, P.; Woisard, A.; Fourrey, J.-L.; Favre, A. *C. R. Acad. Sci. Paris, Life Sci.* **1995**, *318*, 307–313.
(37) Woisard, A.; *et al.* To be published.
(38) Slim, G.; Gait, M. J. *Biochem. Biophys. Res. Commun.* **1992**, *183*, 605–609.

by the long-range cross-link $U_{1.6}$,$A_{L2.4}$. The structure was then completed by taking into account other cross-links formed in the central core and energy minimized.

Overall the ribozyme domain was shown to adopt a Y shape with the three stems almost coplanar, the angles between the helix axes being ~140° (stems II and III), ~150° (stems I and III), and ~40° (stems I and II). The central core is stabilized by a network of hydrogen bonds implying a number of conserved residues schematically represented in Figure 7. In the substrate strand, residue C_{17} is placed in a pocket made up of the conserved C_3–A_6 turn and it can adopt two alternative locations, one close to G_{12} and the other close to G_5. In the latter variant structure, the substrate strand is stretched at the cleavable C_{17}–U_{11} phosphodiester bond and upon minor adjusment of the backbone torsion angles it can allow an in-line attack of the phosphate by the adjacent ribose 2′-OH.

A similar Y-shaped structure was deduced from fluorescence energy transfer experiments.[39] Hydrodynamic measurements performed with all RNA strands having extended arms predicted, however, a broader angle between helices I and II.[40] The proposed structure is consistent with the one derived from X-ray crystallographic analysis in a number of details (C-2′ endo pucker of the G_8 sugar, mismatched G–A base pairs, weakening of the conserved base pair $A_{15.1}$–$U_{16.1}$).[41] However, the C_3–A_6 turn is less structured as a consequence of photo-cross-link constraints. We have also found additional hydrogen bonds involving G_5 imino and amino groups with the C-2 carbonyl of U_{16} and the 2′-hydroxyl of A_{15}.

Overall the network of hydrogen bonds identified within the core of the ribozyme domain appears to be in agreement with published mutation and substitution studies aimed at probing the structure–function relationship.[7b] However, the proposed model satisfies only seven out of the 10 cross-links which were found in the central core (Table 2). This strongly suggests that in solution the ribozyme domain can adopt at least one additional conformation.

Hairpin Ribozymes. The same photolabeling approach was applied to hairpin ribozymes derived from the autolytic (–) sTRSV RNA.[4] Such ribozymes can be reduced to a core of 50 nucleotides and manifest trans cleavage activity against any properly designed substrate having an ApG site within a four-nucleotide loop (Figure 8). As above, a number of deoxy substrate analogs have been prepared that contain either s^4dU in place of U or s^6dI in place of A or G. The cross-link data supported the secondary structure which was proposed for this system.[42,43] However, some minor cross-links suggested the existence of alternative conformers which do not match the consensus structure.[42] As a bonus the photo-cross-linking method allowed the determination of the apparent dissociation constant governing the equilibrium between the R–dS

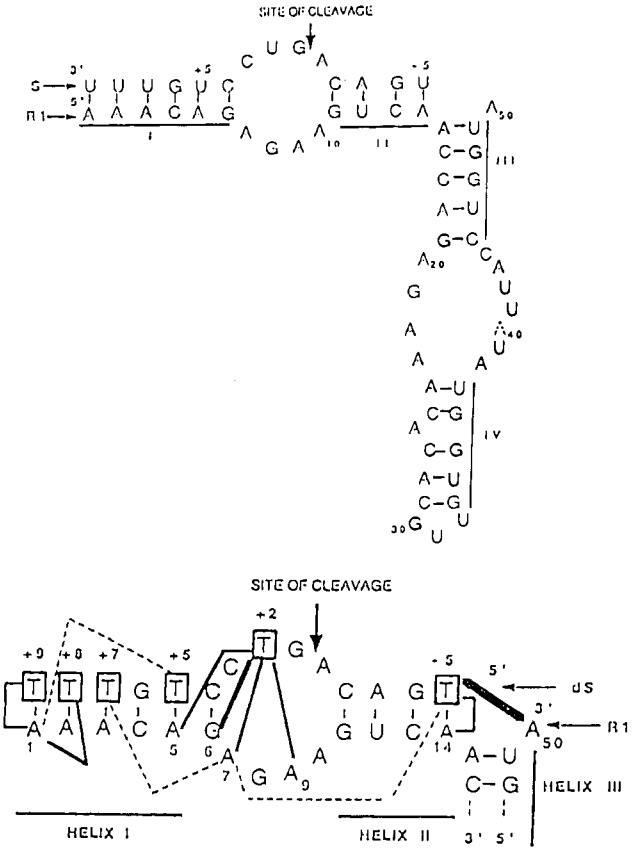

Figure 8. Top: Minimal hairpin ribozyme–substrate complex derived from (–) sTRSV RNA with the current designation of bihelical regions. Bottom: Photo-cross-links obtained with the deoxy substrate analogs having an s^4dU label at defined positions. The width of each line corresponds to the efficiency of crosslinking. The dashed line refers to minor cross-links not expected from the secondary structure.

complex and its components.[43] More important, the data revealed a high flexibility of the bulge involved in the cleavage reaction as well as a preferential folding of this bulge locating the residue +2 of the substrate in close proximity to the purine at position 6 of the ribozyme (Figure 8).

Conclusion

Modified oligonucleotides or polynucleotides containing sulfur-modified nucleobases such as 4-thiouridine manifest a remarkable photochemical behavior which can be exploited to probe their three-dimensional architecture. In the ribozyme series the use of photoactivatable substrate analogs, incorporating these residues, has allowed the identification of a number of tertiary contacts which were used as constraints to build a plausible model of the hammerhead ribozyme domain. This intrinsic photoaffinity methodology has already been used successfully to analyze the folding of an rRNA fragment[22] and to unravel the pathway of mRNA on the ribosome.[44] Introduction of a single s^4U at a strategic position of a pre-mRNA allowed the *in vitro* identification of its interactions with spliceosomal components.[24] Finally, it should be mentioned that this methodology proved to be powerful to analyze RNA–protein[22] or DNA–protein[18e,m] interactions.

(39) Tuschl, T.; Gohlke, C.; Jovin, T. M.; Westhof, E.; Eckstein, F. *Science* **1994**, *266*, 785–789.
(40) Bassi, G. S.; Mollegaard, N. E.; Murchie, A. I. H.; von Kitzing, E.; Lilley, D. M. J. *Struct. Biol.* **1995**, *2*, 45–54. Gast, F. U.; Amiri, K. M. A.; Hagerman, P. J. *Biochemistry* **1994**, *33*, 1788–1796.
(41) Pley, H. W.; Flaherty, K. M.; McKay, D. B. *Nature* **1994**, *372*, 68–74.
(42) Dos Santos, V. D.; Vianna, A.; Fourrey, J.-L.; Favre, A. *Nucleic Acids Res.* **1993**, *21*, 201–207.
(43) Dos Santos, V. D.; Fourrey, J.-L.; Favre, A. *Biochem. Biophys. Res. Commun.* **1993**, *190*, 377–385.
(44) Wollenzien, P.; Expert-Bezançon, A.; Favre, A.; *Biochemistry* **1991**, *30*, 1785–1795.

The herein proposed intrinsic photoaffinity methodology should be particularly useful to explore the folding of single-stranded DNA as well. Indeed, single-stranded DNA plays an important biological role at the replication forks and during transcription. A number of studies indicate that DNA and its homologous RNA exhibit a distinct ability to adopt an ordered conformation such as hairpin folding.[45] Interestingly, a single-stranded DNA can adopt a tRNA-like tertiary structure since a tDNA analog has been found to be functional[46] and a DNA oligomer was shown to manifest a ribozyme-like activity.[47]

Another important field of research which might justify the use of thio-substituted nucleobase photochemistry is DNA photodamage.[9] It is well-known that UVA excitation gives rise to a number of photoproducts. An accurate knowledge of the mechanisms of their formation might be important for elucidating the pathways which are responsible for the formation of DNA photolesions. The reactions which we have described above gave new insights into the mechanism of formation of the (6–4) bipyrimidine photoadducts.[9] The new structures obtained with model oligonucleotides might also be formed in irradiated DNA but could have been overlooked due to isolation problems.

In many systems the extensive development of the thio-modified nucleoside photochemistry is desirable. Thus in view of obtaining cross-link data at longer range a series of new probes have been constructed. They contain a photoactivatable 4-thiothymine residue at the extremity of a more or less flexible chain of variable length attached to the C-5 position of deoxyuridine.[48] Their reactivity is currently being explored in the ribozyme series as well as in other model systems.

We are confident that this new powerful analytical technique is of general applicability. It is complementary to other approaches and will serve in elucidating various structural problems involving nucleic acids in their native state by producing information which could not be obtained otherwise.

Special thanks are due to Dr. P. Clivio, V. D. Dos Santos, P. Laugâa, and A. Woisard for their vital participation in the work described.

AR9400709

(45) Tanaka, M.; Dos Santos, V.; Laugâa, P.; Favre, A.; Fourrey, J.-L. *New J. Chem.* **1995**, *19*, 469–478.
(46) Paquette, J.; Nicoghosian, K.; Qui, G. A.; Beauchemin, N.; Cedergren, R. *Eur. J. Biochem.* **1990**, *189*, 259–265.
(47) Breaker, R. R.; Joyce, G. F. *Chem. Biol.* **1994**, *1*, 223–229.
(48) Saintomé, C.; Clivio, P.; Fourrey, J.-L.; Woisard, A.; Favre, A. *Tetrahedron Lett.* **1994**, *35*, 873–876.