

New Trends in Photobiology (Invited Review)

Thionucleobases as intrinsic photoaffinity probes of nucleic acid structure and nucleic acid–protein interactions

Alain Favre^{a,*}, Carole Saintomé^a, Jean-Louis Fourrey^b, Pascale Clivio^b, Philippe Laugâa^a

^a Institut Jacques Monod, CNRS-Université Paris VII, 2, Place Jussieu, 75251 Paris Cedex 05, France

^b Institut de Chimie des Substances Naturelles, CNRS, 91198 Gif-sur-Yvette Cedex, France

Received 3 September 1997; accepted 7 November 1997

Abstract

In the past few years thionucleobases have been extensively used as intrinsic photolabels to probe the structure in solution of folded RNA molecules and to identify contacts within nucleic acids and/or between nucleic acids and proteins, in complex nucleoprotein assemblies. These thio residues such as 4-thiouracil found in *E. coli* tRNA and its non-natural congeners 4-thiothymine, 6-thioguanine and 6-mercaptapurine absorb light at wavelengths longer than 320 nm and, thus, can be selectively photoactivated. Synthetic or enzymatic procedures have been established, allowing the random or site-specific incorporation of thionucleotide(s) within a RNA (DNA) chain which, in most cases, retains unaltered structural and biological properties. Owing to the high photoreactivity of their triplet state (intersystem yield close to unity), 4-thiouracil and 4-thiothymine derivatives exhibit a high photocrosslinking ability towards pyrimidines (particularly thymine) but also purines. From the nature of the photoproducts obtained in base or nucleotide mixtures and in dinucleotides, the main photochemical pathway was identified as a (2 + 2) photoaddition of the excited C–S bond onto the 5, 6 double bond of pyrimidines yielding thietane intermediates whose structure could be characterized. Depending on the mutual orientation of these bonds in the thietanes, their subsequent dark rearrangement yielded, respectively, either the 5-4 or 6-4 bipyrimidine photoadduct. A similar mechanism appears to be involved in the formation of the unique photoadduct formed between 4-thiothymidine and adenosine. The higher reactivity of thymine derived acceptors can be explained by an additional pathway which involves hydrogen abstraction from the thymine methyl group, followed by radical recombination, leading to methylene linked bipyrimidines. The high photocrosslinking potential of thionucleosides inserted in nucleic acid chains has been used to probe RNA–RNA contacts within the ribosome permitting, in particular, the elucidation of the path of mRNA throughout the small ribosomal subunit. Functional interactions between the mRNA spliced sites and U RNAs could be detected within the spliceosome. Analysis of the photocrosslinks obtained within small endonucleolytic ribozymes in solution led to a tertiary folded pseudo-knot structure for the HDV ribozyme and allowed the construction of a Y form of a hammerhead ribozyme, which revealed to be in close agreement with the structure observed in crystals. Thionucleosides incorporated in nucleic acids crosslink efficiently amino-acid residues of proteins in contact with them. Despite the fact that little is known about the nature of the photoadducts formed, this approach has been extensively used to identify protein components interacting at a defined nucleic acid site and applied to various systems (replisome, spliceosome, transcription complexes and ribosomes). © 1998 Elsevier Science S.A.

Keywords: Thionucleobases; 4-Thiouridine; Photoaffinity labelling; Nucleic acid; Protein; Structure; Interaction

1. Introduction

Treatment of genetic information within the cell, requires multiple interactions between nucleic acids as well as between nucleic acids and proteins. Most of these fundamental processes are achieved in nucleoprotein assemblies (replisome, spliceosome, transcription complexes, ribosomes, ...). Understanding the functioning of any of these complexes

requires a detailed knowledge of both the structure of their individual components and the way they mutually contact each other at any step of their functional cycle. In the past few years a number of crosslinking methodologies has been designed to tackle this challenging problem, being able to give physical evidences for several types of biologically important interactions. In this respect, photocrosslinking methods offer a number of decisive advantages due to a better control in the formation of covalent bonds between contacting moieties, provided monofunctional photolabels (preferen-

* Corresponding author. E-mail: corne@ccr.jussieu.fr

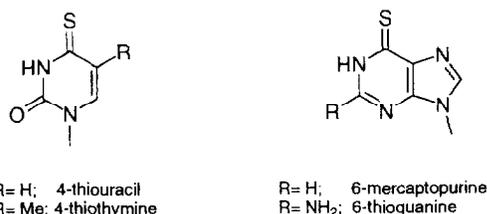


Fig. 1. Structure of 4-thiouridine and some thio-substituted base analogs.

tially nucleoside analogs) can be incorporated in representative biological systems [1–3]. Among these photolabels, 4-thiouridine, thiouridine and its non-natural congeners including 4-thiothymine, 6-mercaptapurine and 6-thioguanine deserve special attention (Fig. 1).

4-Thiouridine (s^4 Urd) was discovered in 1965 as a rare nucleoside present only in position 8 (and exceptionally position 9) of *E. coli* tRNA [4]. The presence of a sulfur atom in place of a keto oxygen at position 4 of the uracil ring results in exceptional spectroscopic properties. The absorption spectrum of s^4 Urd is shifted to the red ($\lambda_{\max} = 330$ nm) while all major RNA components only display end-absorption at wavelengths longer than 300 nm. It can thus be selectively photoactivated leading to the quantitative formation of a photocrosslink between residues 8 and 13 in all native tRNAs containing a s^4 Urd in position 8 and a cytidine in position 13 (50% of molecules from bulk *E. coli* tRNA). This photo-reaction has unraveled for the first time a feature of tRNA tertiary structure common to a large number of tRNAs [5,6]. It also plays a key role in *E. coli* photobiology. Near-UV irradiation of *E. coli* cells triggers a number of photobiological effects such as growth delay, uncoupling of cell growth and division and photoprotection for competent cells [1]. Interestingly, a mutant cell line (*nuvA*) containing Urd instead of s^4 Urd at position 8 of its tRNA grows as well as the wild type cell in the dark, but lacks the photobiological effects mentioned above. Moreover, formation of the 8-13 link occurs in isolated tRNA and in the cells with similar rates and yields. However, some crosslinked tRNA species appeared to be poor substrates of their acylation enzymes, resulting in a marked reduction of protein synthesis and a slower growth rate. Indeed, in cells submitted to solar light, arrest of protein synthesis prevents expression of the inducible mutagenic SOS response due to the lesions triggered in DNA by the UVC and UVB components of light. Thus s^4 Urd acts as an antiphotomutagenic device [1].

In the field of photolabeling experiments the attributes of s^4 Urd as a photoaffinity probe are numerous: (i) it is stable in the dark; (ii) it can be selectively photoactivated by wavelengths between 330 and 370 nm thus keeping undesirable side reactions at a minimum; (iii) it is highly photoreactive towards both nucleic acid bases and amino acid residues (see below); (iv) its base-pairing properties are similar to those of the parent nucleoside uridine allowing polynucleotide chains, at low substitution levels, to retain their standard physico-chemical properties [1,7]. In addition, s^4 Urd (on the triphosphate form) is readily incorporated in RNA chains

both in cells and in a T7 RNA polymerase system so that the term of intrinsic photolabel has been coined to illustrate this essential feature. Since s^4 dThd and other thionucleosides are expected to retain closely similar properties, s^4 Urd obviously constitutes the prototype of an interesting class of nucleic acid photolabels. The photochemistry of s^4 Urd, focused on tRNA work, has been reviewed in 1990 [11].

The purpose of the present paper is to summarize and update our knowledge of thionucleobase photochemistry and to illustrate some of the remarkable achievements obtained by the intrinsic photolabeling technique since 1990. First, the procedures allowing the synthesis of oligonucleotides containing a thionucleobase at a single preselected position will be summarized. Then, the photophysics and photochemistry, so far essentially limited to 4-thiopyrimidines will be described, going from mixtures of bases or nucleotides to dinucleotides and polynucleotides. Applications of this phototechnology to elucidate RNA structure, to unravel nucleic acid–nucleic acid contacts within complex assemblies and to determine proteins contacting a defined nucleic acid portion in nucleoproteins will be presented and discussed.

2. Synthesis of (oligo)polynucleotides site specifically substituted with thionucleotides

Incorporation of s^4 UTP (as well as s^4 TTP, s^6 GTP or s^6 TTP) into polynucleotides is easily achieved using appropriate polymerase systems. The exonuclease free fragment of DNA polymerase I, but not T4 polymerase, incorporates s^4 dTTP into DNA albeit with some unintended primer extension [8]. Synthesis of s^4 U containing polyribonucleotide is in general performed with the T7 RNA polymerase system. At low values of the s^4 UTP/UTP ratio, the analog is incorporated at a rate 4 to 5 times slower than UTP and is found distributed randomly among available positions [9] Enzymatic incorporation of s^6 GTP in polynucleotide chains revealed to be much less efficient than with s^4 UTP ([10] and unpublished data from our laboratory). This is in agreement with the finding that s^6 G incorporation in cellular RNA is far more damaging to the cell metabolism than the corresponding s^4 U incorporation [11].

Generally, in photolabeling experiments, analysis of the crosslinks is facilitated if a single thionucleotide is incorporated at a strategic preselected position. This can be achieved by the above procedure if a single position in the synthesized polynucleotide is available for incorporation of the thionucleotide assuming that the polynucleotide can function despite an artificial sequence context [12]. However, a more general procedure can be used to achieve this goal; it involves chemical synthesis of thionucleotide containing oligonucleotides combined with enzymatic ligation. The first step can be realized using either the phosphoramidite or the H-phosphonate method. Indirect routes involving post synthetic modifications have been considered [13,14] but the use of sulfur-protected nucleosides appears more convenient. The

latter procedures differ by the choice of the protection recommended at the sulfur position to avoid undesirable side reactions during the phosphorylation and oxidation steps of oligonucleotide synthesis. We have proposed the S-(pivaloyloxy) methyl group which can be easily introduced and, in contrast to other suggested groups, can be removed without modification of the protocol used in machine synthesis [15–18]. Recently it was shown that no sulfur protection is required if one uses *tert*-butyl hydroperoxide instead of standard iodine/water for the oxidation step in solid phase synthesis [7]. Such protocols allow synthesis of site-specific substituted RNA up to about 30 nt [19] while for DNA it can exceed 100 nt. If a longer polynucleotide chain is required, then two DNA fragments can be linked together using a bridging oligonucleotide and a DNA ligase. In the case of RNA advantage was taken that a dinucleotide NpG, in occurrence $s^4\text{UpG}$, can serve in place of GTP to initiate transcription by T7 polymerase. Ligation of two RNA chains by means of T4 DNA ligase and a bridging oligodeoxynucleotide has allowed introduction of $s^4\text{U}$ residues at $s^4\text{UpG}$ sites in a RNA chain [20,21]. Combining chemical synthesis of thio base containing oligoribonucleotides and enzymatic ligation should allow the incorporation of the thioresidue at any desired position. Recently a new protocol has been proposed to insert $s^4\text{U}$ in long RNA chains. First, $ps^4\text{Up}$ is ligated to the 3'-end of the 5'-half RNA [22]. Then, the elongated 5'-half and appropriate 3'-half are ligated using the T4 DNA ligase procedure [23].

3. Photophysics

The important photobiological role of $s^4\text{Urd}$ has stimulated extensive investigation of its ground state and excited state properties. This was further extended to parent compounds such as 4-thiouracil $s^4\text{U}$, its tautomeric blocked forms 1-Me $s^4\text{U}$, 3-Me $s^4\text{U}$, 1,3-diMe $s^4\text{U}$ as well as 4-thiothymine $s^4\text{T}$. All $s^4\text{U}$ derivatives exhibit closely similar behavior, previously reviewed in details [1]. Only the main features of this research will be summarized here.

In neutral aqueous medium, 4-thiouracil is essentially in the 2-keto-4-thione form which strongly absorbs light in the near-UV range ($\lambda \approx 330$ nm). Irradiation at this wavelength directly populates the S_2 ($\pi\pi^*$) singlet state which rapidly decays (estimated lifetime 5 ps) and efficiently converts, $\Phi_{isc} \approx 0.9 \pm 0.1$ to the corresponding lowest ($\pi\pi^*$) triplet state T_1 . This state, which has been characterized by a variety of techniques, gives rise in solution to a room-temperature phosphorescence ($\lambda_{max} \approx 550$ nm, $\Phi \approx 3 \times 10^{-4}$) which decays with a lifetime of 200 ns in aerated solution at 298 K. At 77 K, in a glass the phosphorescence yield increases up to 15% and two distinct emissions ($\lambda_{max} = 475$ nm, $\tau \approx 1$ ms and $\lambda_{max} = 550$ nm, $\tau \approx 0.4$ ms) can be detected. The behavior of 4-thiouracil analogs, including 4-thiothymine is similar. Obviously, for these compounds, the high S–T interconversion yield and the short singlet lifetime point to the triplet

state T_1 as the photoreactive state. Indeed, the T_1 state is efficiently quenched in solution (k_q up to $6 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$) by a variety of compounds including halides ions such as Br^- , Cl^- , aminoacids and nucleosides or nucleotides [1]. Of peculiar interest is ground state dioxygen quenching in aerated solution which proceeds with a rate constant of $4 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ and generates singlet oxygen $^1\text{O}_2$ with a yield $\Phi_{\Delta} \approx 0.2$ [24].

4. Photochemistry

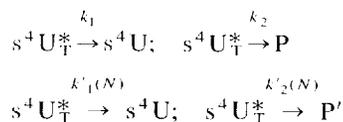
Much attention has been paid to the photocoupling ability of thionucleosides when mixed in solution with the current nucleosides or in dinucleotide models. However, it should be called to mind that these compounds undergo photolysis even when irradiated at low concentrations ($C < 5 \times 10^{-5} \text{ M}$), keeping *trans*-coupling reactions at their minimum.

4.1. Photolysis of $s^4\text{U}$ derivatives in diluted aqueous solution

Dilute solutions of $s^4\text{Urd}$ or analogous compounds are photolyzed with a quantum yield of 10^{-3} in the presence of oxygen and with a ten-fold reduced efficiency in its absence. In aerobic conditions the main photoproduct is Urd. However it is not yet clear whether the $s^4\text{Urd}$ to Urd conversion is due to direct photoreaction of the $s^4\text{Urd}$ triplet state with oxygen and/or whether it predominantly results from oxidation of ground state $s^4\text{Urd}$ by singlet oxygen $^1\text{O}_2$ which is produced in high yield. Formation of $s^4\text{Urd}$ hydrate, which has never been unambiguously characterized, is at most a residual pathway in contrast to the behavior of the parent compound Urd.

4.2. Photoreactivity and photoadducts obtained with mixtures of bases or nucleosides

Regarding the photochemistry of nucleic acids [25] the first attempts to identify mixed photoproducts were limited to pyrimidine adducts with U and C residues [1,26,27]. The finding that in a folded RNA fragment $s^4\text{Urd}$ yields a substantial amount of intramolecular crosslinks with A and G [9] led to a comprehensive reevaluation of the $s^4\text{Urd}$ photocrosslinking potential [28]. All current nucleosides (N) in millimolar range concentrations revealed to be able to quench the $s^4\text{Urd}$ triplet ($s^4\text{U}_T^*$) and to stimulate its photolysis in aqueous aerated solution according to the following collisional scheme:



allowing determination of the rate constants $k_q \approx k'_1 + k'_2$ and of the ratios k'_2/k_q and k'_2/k_2 . The formation of stable pho-

Table 1

Rate constants for s^4 Urd quenching and photoadduct formation by *trans*-coupling with current nucleosides

	T	U	A	C	G
$10^9 \cdot k_q$ ($M^{-1} s^{-1}$)	1	0.3	0.8	0.06	~ 0.1
$10^3 \cdot k'_2/k_4$	80	12	4	7	~ 3
k'_2/k_2 (M^{-1})	(a) 16000	700	700	80	~ 100
	(b) 2000	500	750	180	~ 100

The acceptor nucleosides are ordered following photoreactivity. Data obtained with G can only be considered as crude estimates. The ratio k'_2/k_2 was evaluated using s^4 Urd photolysis stimulation (a) or the direct photocoupling assay (b).

toadducts (P') between N and the deoxytrinucleotide $d(s^4UCC)$ was monitored by a gel retardation assay providing another estimate of k'_2/k_2 . One major photoadduct can be detected for A and G while, at least, two adducts are formed in the presence of pyrimidine nucleosides. As shown in Table 1, Thd is a much better acceptor than any other nucleoside [28] in agreement with previous data showing that T residues are the major targets in s^4 Urd photocrosslinking experiments with DNA [29]. In addition mixed photoadducts are formed with ten times higher overall quantum yield with thymine ($\Phi = 0.017$) than with cytosine in s^4 Urd-base mixtures [30]. The discrepancy observed in Table 1 for the k'_2/k_2 values corresponding to Thd can be ascribed to the use of $d(s^4UCC)$

in place of free s^4 Urd but also to the possibility that Thd could promote photolytic decomposition of s^4 Urd.

In order to gain some insight into tRNA structure, intensive efforts were made to identify the nature of the photoproducts whose formation gave rise to the 8-13 photocrosslink. Thus, 334 nm irradiation of submillimolar mixtures of 4-thiouracil and cytosine in deaerated aqueous solution allowed the isolation of two adducts, namely pyrimidin-2-one (4-5) cytosine, Pyo(4-5)Cyt (Fig. 2) and pyrimidin-2-one (4-5) 4-thiouracil, Pyo(4-5) s^4 Ura. The former was obtained in a low yield with an overall quantum yield of 1.5×10^{-3} while the latter was formed in roughly ten times higher chemical and quantum yields. Similarly irradiation of a 4-thiouracil-uracil mixture yielded Pyo(4-5)Ura [31,32]. These adducts could be interconverted between each other by mild chemical treatments and furthermore the Pyo(4-5)Cyt structure was confirmed by unequivocal multistep synthesis [33]. The spectroscopic properties of these photoadducts (absorption and fluorescence spectra) have been summarized in Wang's review [34].

The chemical nature of the 8-13 link was established first by comparison of the absorption and fluorescence characteristics of authentic diriboside photoadducts isolated from irradiated tRNA^{Val} [35] or bulk *E. coli* tRNA [36] with those of Pyo(4-5)Cyt obtained by irradiation of base mixtures and to Pdo(4-5)Cyd produced by 254 nm irradiation of hemipro-

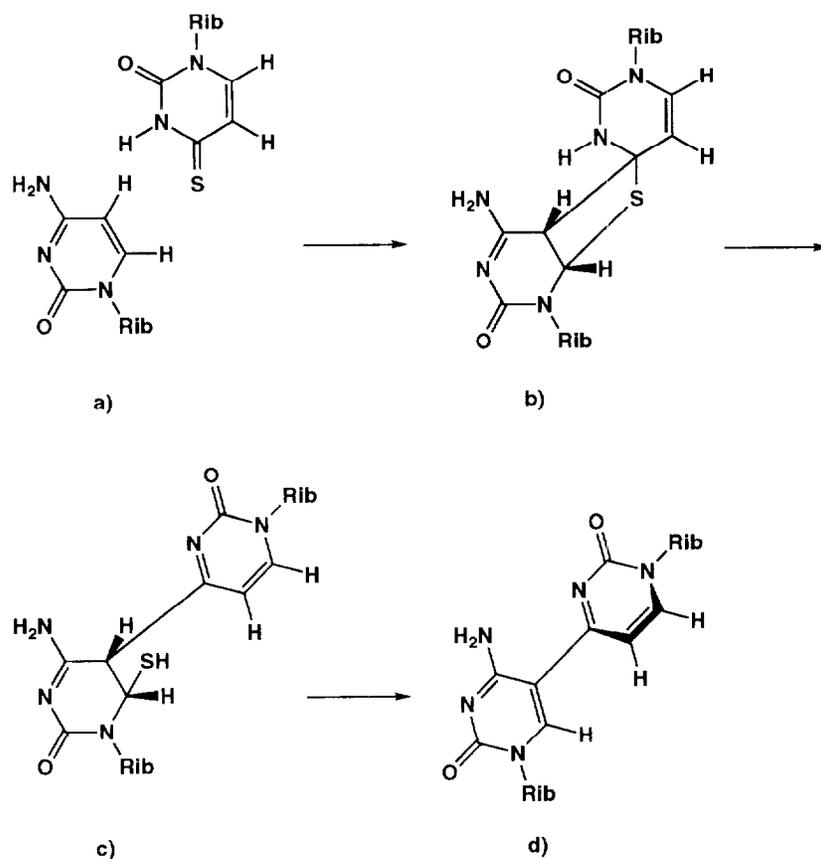


Fig. 2. Reaction scheme leading, in *E. coli* tRNA, to the formation of the 8-13 photoadduct, Pyo(4-5)Cyt (d) via a thietane intermediate (b).

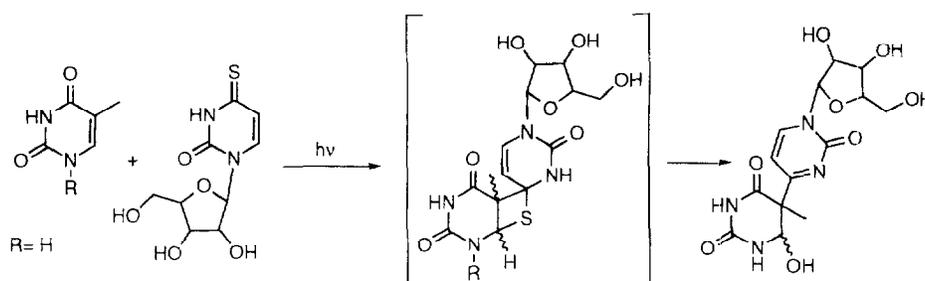


Fig. 3. Structure of O⁵hThy(5-4)Pdo photoadduct obtained by irradiation of a mixture of thymine and s⁴Urd.

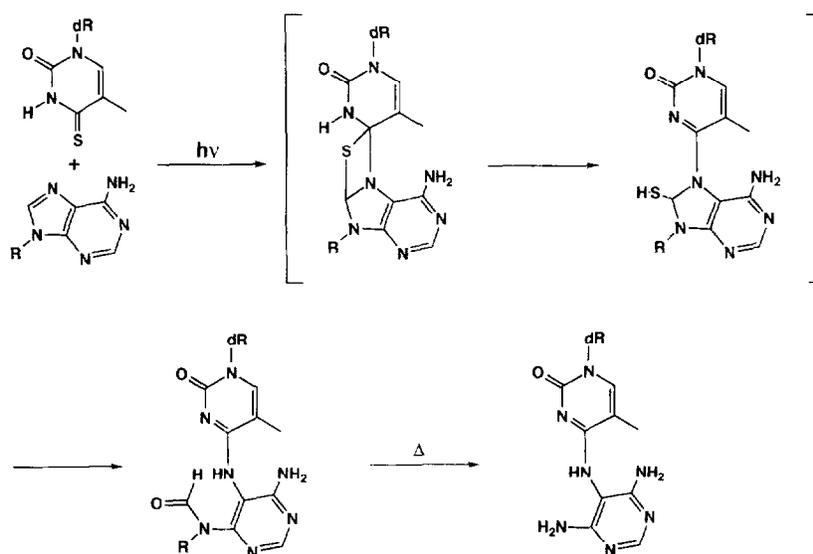


Fig. 4. Postulated reaction scheme yielding (N-6-Fapy-Ade) upon irradiation of N-1 substituted 4-thiothymine in the presence of N-9 substituted adenine.

tonated Cyt puddles [37,38]. It was further confirmed by the finding that NaBH₄ reduction, which converts the 8-13 photoproduct in tRNA into a highly fluorescent compound ($\lambda_{ex} = 390$ nm, $\lambda_{em} = 450$ nm) [39], equally converts the (5-4) photoadducts into their reduced forms. Pyo(4-5)hPyr [31,32,36]. The fluorescent properties of Pyo(4-5)hCyt, when either free in solution or in tRNA, have been thoroughly examined [40]. The formation of the 8-13 link was proposed to occur by a (2+2) cycloaddition involving a thietane intermediate [41,42], thus implying an antiparallel orientation of the s⁴Urd 8 and Cyt 13 glycosidic bonds [32,36], as indeed found in tRNA crystals (Fig. 2).

Owing to the high photoreactivity of the thymine-s⁴Urd mixture Blazek et al. [43] attempted to characterize the corresponding adducts. From the irradiated mixture they were able to separate by HPLC at least 13 products, one of which was characterized as 6-hydroxy-5,6-dihydro-thymine(5-4)-pyrimidine-2-one nucleoside abbreviated O⁵hThy(5-4)Pdo (Fig. 3). It could result from a (2+2) cycloaddition yielding a thietane intermediate with the C5C6 bond of Thy parallel to the C4S4 bond of s⁴Urd followed by ring opening and exchange of the SH group at Thy C6 by an hydroxyl group. The same type of (5-4) photoproducts is also obtained in dinucleotide analogs in which the dThd and s⁴dThd moieties are linked by a flexible polyamide backbone [44].

An unexpectedly clean photoreaction occurred upon irradiation of N-1 substituted 4-thiothymine in the presence of 1 to 2 equivalents of N-9 substituted adenosine to provide a single photoadduct having a N-4 (4-amino-N-6-formyl-pyrimidin-5-yl) 5-methylcytosine structure, irrespective of the nature of the substituents (ribo-, deoxyribo- or carboxymethyl-). Interestingly, in the nucleoside series this (N-6-Fapy-Ade) derivative deformylates in alkaline conditions, while it both deformylates and loses the substituent initially attached at Ade N-9 upon thermal degradation (Fig. 4). Formation of these photoadducts is expected to occur again by a (2+2) cycloaddition to the N7C8 double bond of Ade [45].

4.3. Photoreactions of s⁴U and its analogs in di- and oligonucleotides

Dinucleotides were used to examine the photochemical behavior of sulfur-containing bases as they are relatively easy to handle and furthermore represent the simplest models for a polynucleotide chain. The photoreactivity of deoxydinucleotides s⁴dUpN, Nps⁴dU (N being any current nucleoside) and of the diribonucleotides s⁴UpG, s⁴UpU was determined by spectrophotometry in diluted ($c = 0.5 \times 10^{-4}$ M) aqueous neutral solution. Upon irradiation with 335 or 366 nm light, all dinucleotides with N = A, G or C photoreact at rates com-

parable to the one observed with $s^4\text{Urd}$ in aerated solution ($\Phi_p \approx 10^{-3}$). When $N=U$, the observed photoreactivity is two times higher, irrespective of the dinucleotide sequence. Whatever N , the photochemical behavior of dinucleotides is practically independent upon the presence or absence of oxygen (unpublished data). Interestingly, the dideoxynucleotides $s^4\text{dUpT}$ and Tps^4dU revealed to be far more reactive ($\Phi_p \approx 2 \cdot 10^{-2}$). The nature of the photoadducts formed was thus fully elucidated under anoxic conditions, leading to an unexpected sequence-dependent photochemistry. Irradiation of $s^4\text{dUpT}$ led to two photoadducts not previously described [13]. These products most likely arise by hydrogen abstraction from the methyl group of thymine by excited $s^4\text{U}$ followed by radical recombination (Fig. 5). A precedent for this behavior can be found when $s^4\text{U}$ derivatives were irradiated in alcohols or in the presence of triethylamine [46,47]. A different pathway occurs in the case of Tps^4dU which gave the (6-4) bipyrimidine adduct **5** presumably by a (2+2) cycloaddition. In view of the interest for this type of adduct which corresponds to one of the major photolesions induced by UV in DNA, the photochemistry of Tps^4T was examined. Tps^4T yielded four photoadducts (Fig. 6) including the thietane **7** which was found in 3:1 equilibrium with the (6-4) bipyrimidine **8** which was further photochemically converted into its Dewar valence isomer **9**. An additional new photoproduct **10** formed in a low yield by hydrogen abstraction followed by hydrogen sulfide elimination. In order to stop the photoreaction at the thietane stage the N-3 position of the 4-thiothymidine was blocked by a methyl group. Tpm^3s^4T gave compounds **12**, **13** and **14** in 25, 2 and 24% yields, respectively (Fig. 7). The minor thietane **13** resulted from a

cycloaddition pathway in which the 5'-nucleoside is in the syn conformation while the 3' one is anti. The occurrence of stable thietane derivatives allowed examination of their own photochemistry. Interestingly, 254 nm irradiation of a neutral aqueous solution containing an interconverting mixture of thietane **7** and (6-4) photoadduct **8** quantitatively restored Tps^4T . However, at pH 10 which completely shifts the $7 \leftrightarrow 8$ equilibrium towards **8**, no reversal reaction occurred, suggesting that the photoreversion reaction involves 254 nm photoactivation of thietane **7**. This was confirmed since 254 nm photoreversion of the Tpm^3s^4T thietane **12** could be achieved at both pH 6 and pH 10 [48,49].

Since the photochemistry of $s^4\text{dUpT}$ and Tps^4dU was found absolutely sequence dependent it was of interest to examine whether it could be sensitive to the configuration of the phosphate analog in the (Rp) and (Sp) forms of thymine-3'-methylphosphonate 5'-N-3-methyl-4-thiothymidine (Tpm^3s^4T). While all major photoadducts obtained with Tpm^3s^4T were observed with both the Rp and Sp diastereoisomers, additional products resulting from the hydrolysis of the 3'-end glycosidic bond of **14** were detected only with the latter diastereoisomer. These data together with the observation that the photoreactivity decreases in the order $\text{Tpm}^3s^4\text{T} > (\text{Rp}) > (\text{Sp})$ strengthen the conclusion that the (Rp) diastereoisomer is a better mimic of the natural phosphodiester bond than its (Sp) diastereoisomer [50]. In continuation of this line of research, Clivio et al. [51] synthesized PNA analogs of dinucleotides. PNA are achiral and neutral DNA analogs in which the deoxyribose phosphate backbone has been replaced by a pseudo-peptide chain constituted of N-(2-aminoethyl)-N-methylenecarbonyl glycine units that

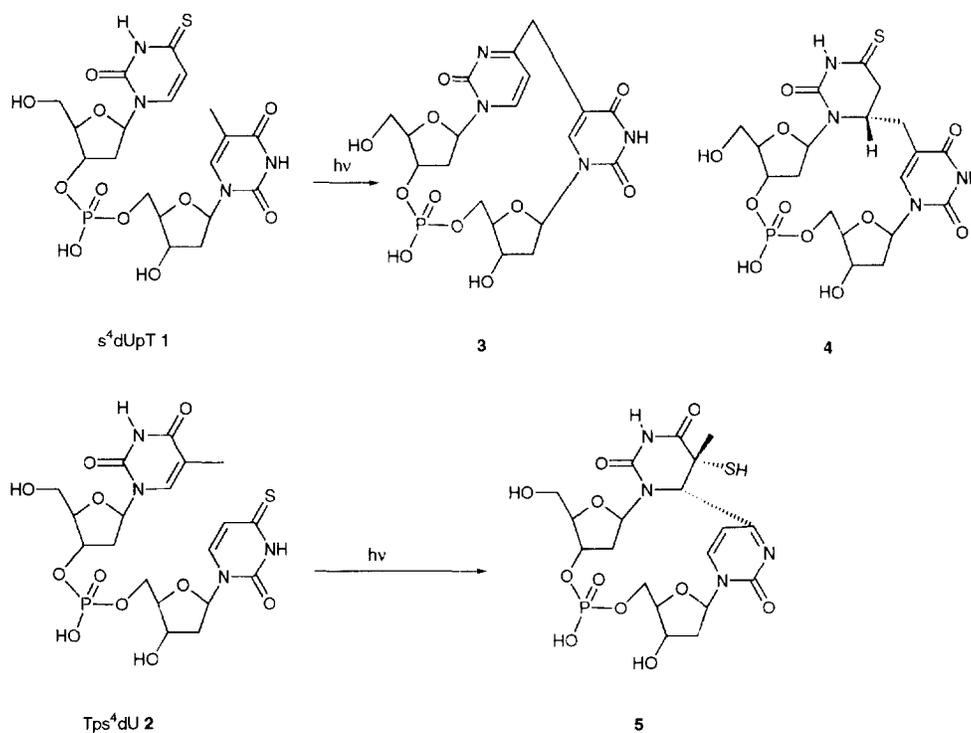
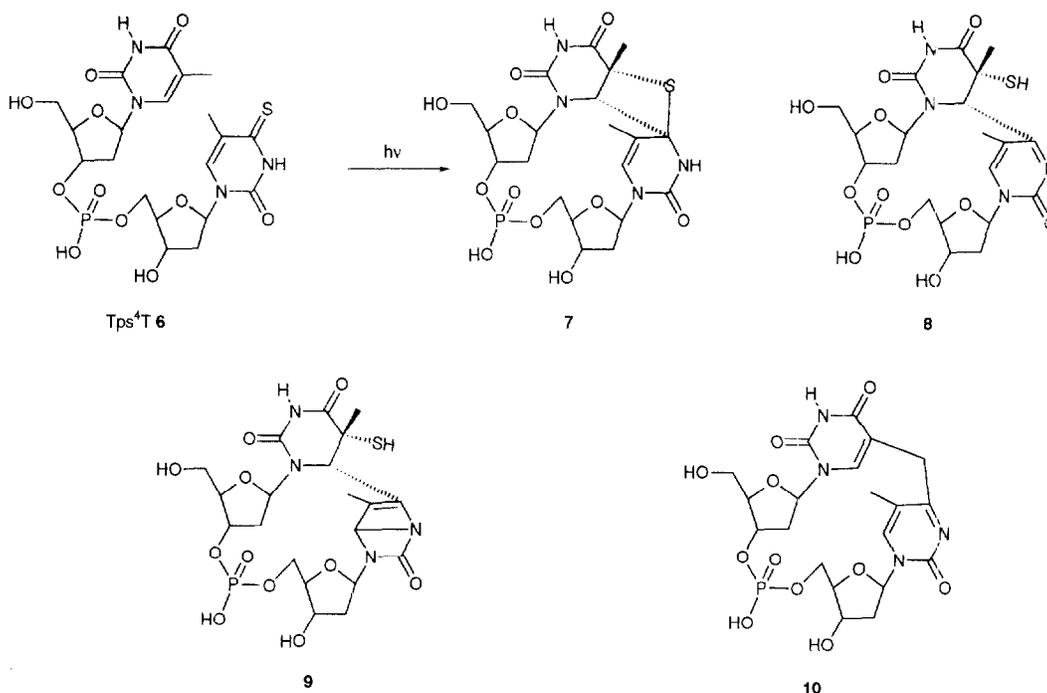
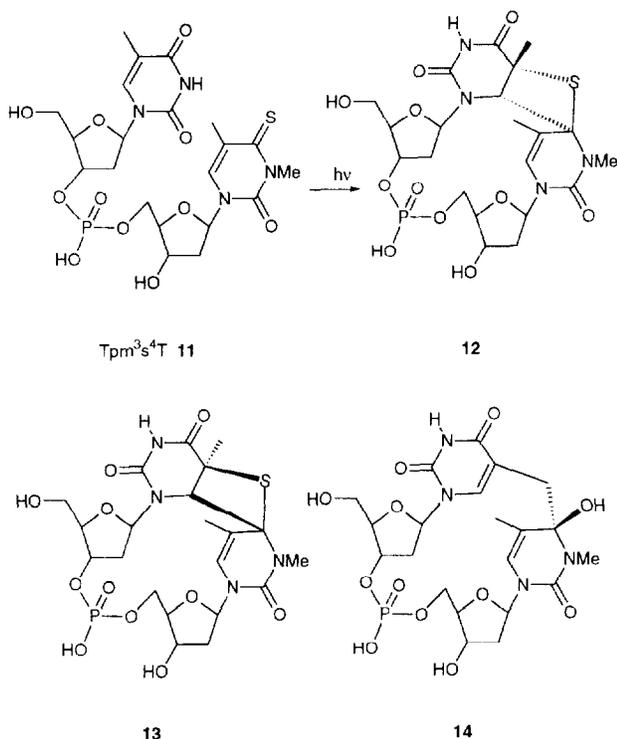


Fig. 5. Structure of the photoproducts obtained upon irradiation of $s^4\text{dUpT}$ (top) or Tps^4dU (bottom).

Fig. 6. Structure of the photoproducts obtained upon irradiation of Tps^4T .Fig. 7. Structure of the photoproducts obtained upon irradiation of $\text{Tpm}^3\text{s}^4\text{T}$.

can bind complementary DNA or RNA strand with high affinity. In comparing the photochemical behavior of two PNA dimers incorporating s^4T either at the N or C terminal end Clivio et al. [51] observed that the orientation of the aminoethyl glycine backbone modulated the nature of the photoproducts. However, none of them is a true mimic of the corresponding dinucleotide. Observations of such reactions

could also be extended to 6-mercaptapurine deoxyriboside (6-thio deoxyinosine, s^6dI) in a dinucleotide context. Thus irradiation of Tps^6dI led in a high chemical yield to a (6-6) adduct whose structure is reminiscent of the (6-4) bipyrimidine **8** [52].

So far, detailed analysis of s^4U , s^4T photochemical behavior in polynucleotides is quite limited. Liu and Taylor [49] showed that the photochemistry observed for the Tps^4T dinucleotide can be reproduced when this motif is inserted in an octanucleotide. As formation of $\text{Py}\delta(4-5)\text{Cyd}$ photoadducts implies a 'head-to-tail' orientation of the reactive pyrimidines, it was of interest to determine the minimal length of s^4U and C containing oligonucleotides for this photoreaction to occur. Using the NaBH_4 reduction test we observed that in the deoxyribooligonucleotide series s^4dUC_N (N , number of Cyt units), formation of the 5-4 adduct requires at least $N=3$ and is optimal when $N \geq 4$.

5. Elucidation of contacts within or between nucleic acid chains

The intrinsic photoaffinity methodology has found extremely remarkable applications for studying interactions within the spliceosome apparatus and the ribosome. In addition, this approach was used for deciphering the folding of small endonucleolytic ribozymes in solution.

5.1. Nucleic acid models

RNAs are known to adopt complex structures in solution. The basic motif in their folding is the hairpin-loop, possibly

containing mismatched base-pairs, unpaired bases and bulges. Tertiary folding is insured by interactions involving both loops and double stranded regions. It was thus important to determine the potential of s^4U to explore such structures. In a seminal set of experiments, Dubreuil et al. [9] investigated the folding in solution of a 218 nt rRNA fragment randomly substituted with s^4U using T7 transcription. Irradiation of the folded monothiolated chains substituted at any of the 41 available positions generated long-range intramolecular crosslinks that were separated and sequenced (17 crosslinks could be mapped). Finally this study demonstrated that crosslinks involving both pyrimidine or purine acceptor residues formed efficiently when the partner residues were located at the border of minihelices or within bulges and loops [9].

Formation of interstrand bridges in DNA, RNA or hybrid minihelices site-specifically substituted with s^4U confirms the above findings. These crosslinks formed efficiently only when the s^4U residue is located at or near the dangling ends (with yields up to 80% in DNA helices) [53,54].

5.2. Spliceosome

A gene can be defined as a DNA segment that is expressed to yield a functional product (RNA or polypeptide). Eucaryotes genes generally have a split structure in which segments of coding sequences (exons) are separated by non-intervening sequences (introns). The entire gene is transcribed to yield a primary transcript, pre-mRNA. The latter biomolecule may be subjected to a number of processing events in the nucleus, including splicing of introns, before being transported as a mature mRNA in the cytoplasm. In the splicing process the pre-mRNA assembles with numerous proteins and five small nuclear RNA (U_1 , U_2 , U_4 , U_5 and U_6 snRNAs). The resulting structure forms a dynamic enzymatic complex (the spliceosome) that catalyses the splicing reaction in the presence of ATP. To probe the interactions of these snRNAs with the active center of mammalian spliceosome, Sontheimer and Steitz [20] synthesized substrate pre-mRNA containing a single s^4U residue adjacent to either splice sites. Crosslinks were induced during the course of the splicing reaction. This demonstrates that an invariant loop of RNA U_5 interacts functionally with the 5'-splice site and aligns the two exons for ligation since the first residue of exon 2 also became crosslinked to U_5 . Several conformational rearrangements of snRNAs within the spliceosome active center could be distinguished including interaction of U_6 conserved sequence with intron (see detailed comments in [55]). Recently, AT-AC introns were found whose removal is catalyzed by a spliceosome that contains only one common (U_5) and four different (U_{11} , U_{12} , U_{4atac} and U_{6atac}) snRNPs compared with the major spliceosome. Site-specific crosslinking allowed to show that the splicing of this rare intron is mechanistically similar to that of the major class of introns. This also implicates that the U_{11} and U_{6atac} snRNPs in the AT-AC spliceosome fulfill analogous roles to U_1 and U_6 ,

respectively, in the major spliceosome [23]. When applied to yeast spliceosome the same approach essentially confirmed the data obtained with mammalian splicing extracts. In addition, this demonstrates that after the first catalytic step the first nucleotide of the 3' exon can also contact U_2 RNA [56]. Interestingly, addition of s^4Urd site specific substituted U_6 RNA to a cell-free yeast system splicing extract depleted of endogenous U_6 snRNA restores functional activity. This also allows the formation of crosslinks between U_6 central domain and the 5' splice site [57]. These photocrosslinking experiments have already largely contributed to a more detailed picture of the dynamic structure of the spliceosome.

5.3. RNA-RNA interactions during translation

An essential feature of protein synthesis is the correct positioning of peptidyl-tRNA and incoming aminoacyl-tRNA in the ribosomal P and A sites respectively, as specified by mRNA through mRNA codon-tRNA anticodon interactions. The ribosomes consist of two well-defined subunits (30S and 50S in *E. coli*). In each subunit, a number of defined r-proteins is organized around rRNA skeletons (16S RNA in 30S particle, 5S and 23S RNA in the 50S particle). While the different steps of protein synthesis (initiation, elongation, termination), the role of soluble factors in these steps and the energy requirements had been well defined in both prokaryotic and eucaryotic systems, the overall structure of the ribosome is known only at low resolution. Several important structural and conformational matters have long remained challenging problems. These include the rRNAs tertiary folding inside the subunits, the nature of tRNA-ribosome contacts in A and P sites (not to speak of the exit site), the mRNA track along the small subunit as well as eventual conformational changes at the different steps of translation.

Sites of contacts of artificial randomly s^4U substituted mRNA (51 nt long) with 16S and 23S RNA in the ribosome were mapped by reverse transcription. Twelve sites on 16S and two on 23S RNA were detected [58] that revealed, to a large extent, independent of the mRNA sequence [59]. The picture was then refined using mRNA analogs containing a Shine-Dalgarno region and a codon for tRNA^{Gly} [60] or tRNA^{Thr} [61], the s^4U residues being introduced at different mRNA positions. Under equilibrium conditions between mRNA and 70S ribosome some mRNA-16S RNA crosslinks were shown to be tRNA dependent. This leads to the suggestion that tRNA binding increases the accessibility of 16S RNA for crosslink formation with mRNA [61].

In a more selective approach Bogdanov and Brimacombe groups used mRNAs which were substituted by s^4U or s^6G residues at a single defined position. The positions of the crosslinks on 16S RNA were mapped by a combination of RNase H and reverse transcriptase analysis [10,12,62–67]. Thus the spacer region between Shine-Dalgarno sequence and the first AUG was shown to contact 16S RNA either around position 1530 as well as 1360 for 7 to 8 nt long spacer or 665 for 4 nt spacers. Downstream to the AUG codon the

following crosslinks were identified: +4 (the position on the 3' side of AUG) to 1402, +6 to 1052, +7 to 1395, +8 and +9 to 1196, +11 and +12 to 532 and +13 to 530. All these crosslinks were found to be entirely tRNA dependent and observed both in initiation and elongation complexes [10,12,62–67]. Interestingly, mRNAs carrying the substituted stop codon s^4 UGA crosslinked also to the 1395–1420 region of 16S RNA. As a future challenge it will be interesting to determine whether a 70S ribosome with mRNA crosslinked to 16S RNA retains biological activity. Application of the same methodology to 80S human ribosomes showed that the few 18S RNA sites (680 and 1111–1112) susceptible to form crosslinks correspond to their equivalent in 16S RNA [68].

In order to analyze the structure of the subunits, s^4 U substitution of rRNA was performed. Thus s^4 U containing 5S RNA was used to reconstitute either 50S or 70S ribosomes. Two crosslinks joining U_{89} of 5S RNA to positions 960 and 2475 of 23S RNA were detected indicating that two distinct domains of this RNA are in close proximity in the 50S subunit [64,69]. Another approach was developed to study the folding of 16S RNA within the 30S particle [22]. The RNA was selectively cut and p^{32} s⁴Up ligated at the 3'-end of one fragment (position 1141). 30S particles were reconstituted and UV irradiated showing that position 1141 is in close contact with residues 1295 and 1272. This allows refinement of rRNA 3D structure in the ribosome.

5.4. Ribozymes

Ribozymes are RNA metalloenzymes able to catalyze specific RNA endonucleolytic cleavage and also RNA ligation reactions. Small endonucleolytic ribozymes are mostly found in plant viroids, virusoids or satellite RNA viruses. Three distinct motifs have been identified, namely the hammerhead [70,71], hairpin [72] and pseudo-knot (or HDV) ribozymes [73,74]. In the presence of divalent metal ions (typically Mg^{2+}), these ribozymes generate 5'-hydroxy and 2'-3' cyclic phosphodiester termini. The small size of these ribozymes as well as the nature of their cleavage products distinguish them from large ribozymes such as group I, II introns and RNase P.

5.4.1. Hammerhead ribozyme

The consensus secondary structure required for activity is predicted to have three duplex stems and a highly conserved core of nucleotides C3-A9 (but U7) and G12-A14 as well as an unpaired nucleotide C17, 5' to the cleavage site (Fig. 8). Within the three duplex stems, only base-pair A15.1:U16.1 in stem III, 5' to C17, is highly conserved. The cleavage occurs specifically at the phosphodiester bond joining nucleotides 17 and 1.1. A large number of mutation and substitution studies has underlined the importance of the conserved nucleotides and nucleotide functional groups present within the central core for the cleavage activity [70,71]. The three-dimensional folding of a transacting hammerhead ribozyme

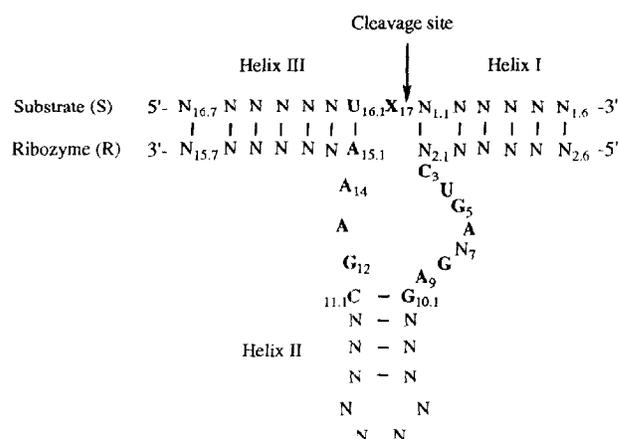


Fig. 8. Hammerhead ribozyme secondary structure with numbering according to Ref. [75]. Bold letters indicate conserved residues of the central core and arrow the cleavage site. X can be any residue but G.

has been studied with non-cleavable substrate analogs substituted at single position by s^4 dU or s^6 dI [52,76]. These positions were chosen in order to bring information related to either the structure of the catalytic site (positions 16.1, 17, 1.1) or the overall structure of the ribozyme (positions 15.7, 1.6). These studies were performed under experimental conditions allowing the cleavage of more than 100 substrates by a single ribozyme ($k_{cat} = 25 \text{ min}^{-1}$ at 37 °C). The use of a full deoxyribose substrate analog likely has little influence on the ribozyme structure since (i) a full deoxysubstrate analog which had a cytidine at the cleavage position was efficiently cleaved and (ii) very similar crosslink patterns were obtained with either a full deoxyribose substrate or its ribo analog, both containing s^4 dU at position 17. A large number of crosslinks was observed with the various substrate analogs. Whereas crosslinks involving either residue 16.7 or 1.6 with, respectively, residues 15.7 and 2.6 are representative of the consensus secondary structure, the crosslink between distant residues 1.6 and L2.4 in stem II loop indicates that, in solution, stems I and II are in a close proximity. Residues 16.1, 17 and 1.1 all yielded multiple crosslinks with the conserved residues of the catalytic core. In particular, the crosslink between residues 16.1 and U7 was formed very efficiently. Most of the tertiary interactions thus determined are likely related to those occurring in a potentially active ribozyme since the crosslinking pattern obtained within the core with a fully inactive mutant is strongly altered. However, it should be noted that the ribozyme still binds properly the substrate as shown by the formation of the 16.7/15.7 crosslink. The multiple crosslinks are indicative of either a large conformational flexibility of the catalytic core or of the existence, in solution, of several closely related folded conformations or both [52,76].

Taking into account these photocrosslinking data, a three-dimensional model could be constructed [77] starting from an A-type stem II extended by a double A9:G12, G8:A13 mismatches supported by thermodynamics and NMR data of small DNA and RNA helices incorporating such double mis-

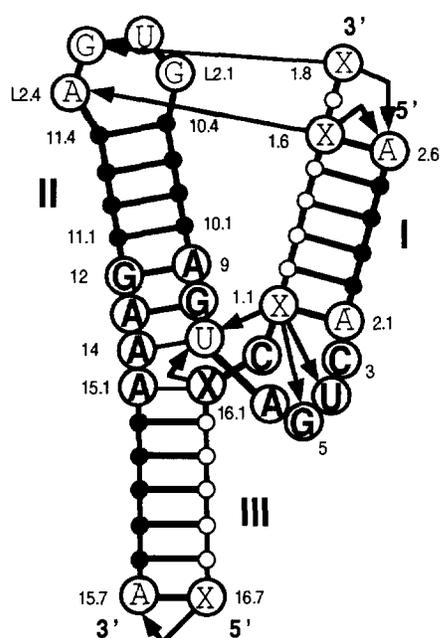


Fig. 9. Planar projection of the 3-dimensional folding of the hammerhead ribozyme. The ribozyme and substrate strands are shown by small black and white circles, respectively. Bold letters mark the conserved residues of the core, and X states for positions that have been substituted by s^4U . Arrows indicate the major photocrosslinks used to build the model [76].

matches and by photocrosslinking experiments at 254 nm. On its 3'-end it was then elongated with A14 and helix III so as to retain an overall A-like type helix. Docking of stem I on this structure was guided by the tertiary contact between residues 1.6 (stem I) and L2.4 (stem II loop) and the necessity to bridge stem I and III with C17. Residues of the catalytic loop (C3-U7) were then added, taking into account the crosslinks in which they were involved (Fig. 9). Overall the structure of the ribozyme looks like a flat Y with the three stems almost coplanar. The central core is stabilized by a network of hydrogen bonds implying a number of conserved residues. In the substrate strand, C17 is located in a pocket made of a C3-A6 turn where it can adopt two alternative positions. One is close to G12 whereas the other is close to G5. In the latter variant the substrate strand is stretched at the cleavable C17-U1.1 phosphodiester bond and minor adjustments of the backbone torsion angles allow an in-line attack of the phosphate by the adjacent ribose 2' hydroxyl. The structure thus obtained is very similar to that provided by X-ray studies [78–81]: overall shape, C3-A6 turn, mismatched G:A base-pairs, C2' *endo* pucker for G8 sugar. However there are significant differences between the two structures. For instance, in the crystal structure, the angle between stems I and II is wider and cannot account for the long-range 1.6-L2.4 crosslink. However, in the crosslink derived structure the C3-A6 turn is much less precisely defined as a consequence of the multiple crosslinks. Another difference lies in the formation, in the crosslink derived structure, of a tertiary interaction involving the conserved residue G5 and the sole conserved base-pair between the ribozyme and the substrate

(U16.1:A15.1). This particular observation, only revealed by the modeled structure could afford an explanation for the conservation of these three residues.

In summary, thio-analog nucleotide-based photolabeling has provided structural information of sufficient quality to allow the construction of a molecular model for the hammerhead ribozyme, very similar to that obtained by crystallography with a few interesting differences. Taken as a whole, 8 of the 11 crosslinks are satisfied by the average model. This shows that the central core is flexible and that, in solution, alternative conformations might exist.

5.4.2. Hairpin ribozyme

The same photolabeling approach was applied to hairpin ribozymes derived from the autolytic (-)sTRSV RNA [82,83]. These ribozymes, which can be reduced to a core of 50 nucleotides, elicit *trans*-cleavage activity against a substrate having an ApG site within a four nucleotide loop. A number of deoxyribose substrate analogs containing either s^4dU or s^6dI in place of, respectively, U or A (G) was prepared. Most of the crosslinks confirm the consensus secondary structure and revealed a high flexibility of the bulge involved in the cleavage reaction. This is also indicative of a preferential folding of this bulge bringing the residue +2 of the substrate in close proximity to the purine at position 6 of the ribozyme. Nevertheless these data were insufficient to build a plausible three-dimensional model for the hairpin ribozyme.

5.4.3. HDV antigenomic ribozyme

A *trans*-acting system has been designed in order to explore the three-dimensional structure of the antigenomic HDV ribozyme. In this system, the substrate is associated by base-pairing to the catalytic RNA, forming helix H1. The ribozyme is able to cleave specifically the RNA substrate as well as a deoxyribose substrate analog containing a single cytidine at the cleavage site (position -1). Using a set of full deoxy or mixed deoxyribose-ribose substrate analogs site-specifically substituted with s^4dU , a number of long range contacts was determined between the substrate and the ribozyme core [84]. In particular crosslinks between substrate positions -1 and -2 with residues C15, G19 and C67, thought to be involved in the ribozyme catalytic site, were detected (Fig. 10). When residue -1 is a deoxyribonucleotide, the probe at position -2 crosslinks with C15, G19 and G31 whereas the sole residue C67 is crosslinked when residue -1 is a ribonucleotide (cleavable substrate). Interestingly the ribozyme/substrate complex becomes inactive upon formation of the -2/C67 crosslink. This is in line with the statement that C67 is a key residue of the antigenomic ribozyme. As such, the photocrosslinking data are too sparse to allow the building of a three-dimensional model. Besides, such a model has been proposed for the genomic ribozyme on the basis of a large set of mutational data [85]. Since both the genomic and antigenomic ribozymes are able to adopt the same secondary pseudo-knot structure, a model for the the antigen-

refinement of the TAR structure in solution and demonstrated that all three positions contact Tat in the complex [94].

6.3. Transcription

Eucaryotic cells have evolved three forms of DNA-dependent RNA polymerases specialized in the transcription of rRNA (pol I), mRNA (pol II), 5s RNA and tRNA (pol III), not including mitochondrial and/or chloroplastic RNA polymerases. All eucaryotic RNA polymerases are complex protein assemblies made of two large and a number (up to 14) of smaller subunits. A variety of affinity labeling techniques has been used to identify the catalytic domain and/or subunits containing the DNA strands or the nascent RNA. As s^4 UTP is a good analog of UTP in transcription reactions it can be easily incorporated in the nascent transcript and it was shown to interact with subunits Ia Ib and a 52 kDa polypeptide of human polymerase I [95]. This approach has been extensively applied to RNA polymerase II from various sources [96–98] in order to map the path of nascent RNA. In the case of RNA pol III, it was shown that the transcripts contact the two largest subunits until the transcription complex reaches the initiation factor TFIIC binding site [96]. In pea chloroplast transcriptional complex, two peptides of 51 and 54 kDa were labeled [99].

Formation of specific transcription complexes permits selective introduction of the s^4 U residue at the 3'-end of the growing RNA chain. This allowed derivatization of the polymerase catalytic domain, in close proximity to the elongation site. The two largest subunits, A_{190} , A_{135} of yeast pol I were thus radiolabeled [100] while only the largest one of wheat germ and yeast pol II could be detected [101]. Interestingly, it was shown that a RNA polymerase II binding factor S_{11} which allows bypass of template arrest sites contacts directly the 3'-end of nascent RNA [102]. Site-specific substitution of T by s^4 T residue in DNA genes (5s or tRNA^{lys} fragments) and reconstitution of RNA pol III transcription complexes demonstrated that 8 out of 15 subunits of this enzyme make direct contacts with DNA. In addition, this suggested that multiple states are involved in a 'precisely positioned' complex [8].

6.4. Translation

Initiation of translation in eucaryotic cells first involves binding of 40S ribosomal subunits at or near the mRNA 5'-end followed by binding of Met-tRNA_i and initiation factors. In a subsequent step, scanning of the mRNA is achieved in a 5' to 3' direction until an appropriate AUG codon is encountered. In a model system, an AUG triplet is recognized as a start signal when embedded in the oligomer ACC AUG G. In order to decipher the mechanism of start site selection, the mRNA analog ACCAs⁴UGG was incubated in a cell-free system. It was shown to crosslink efficiently two *trans*-acting factors of 50 and 100 kDa, one of which revealed to be the La autoantigen [103]. Other studies were devoted to the *E.*

coli system. Thus in initiation complexes formed with the 30S subunits, s^4 U residues inserted upstream to the AUG codon crosslinked specifically to S7 and less specifically to S1, S18 and S21 [12] while in elongation complexes proteins S1, S3, S5 contact mRNA nearby the decoding site [67]. The nature of the ribosomal P site was investigated using randomly substituted tRNA^{Met}. The 50S subunit proteins L1 and L27 attached the tRNA D loop whereas S19 crosslinked to the variable loop [104]. Protein synthesis is achieved, i.e. the free polypeptide chain is released, when a translational stop signal occupies the ribosomal A site, triggering specific binding of appropriate release factors (RF). Highly efficient crosslinking of RF2 to either s^4 UAA or s^4 UGA containing mRNA was observed under these conditions. This is in agreement with models in which the RF has an anticodon-like domain [105,106].

7. Conclusions

The use of thionucleotides as intrinsic photolabels of nucleic acids has yielded a number of remarkable findings (folding of small endonucleolytic ribozymes, the path of mRNA onto the ribosome, nature of RNA-protein interactions within the spliceosome or RNA polymerase complexes...). Not only new contacts between macromolecules could be revealed (and mapped) but in some cases the functionality of the crosslinked complexes was checked (ribozymes and spliceosome), thus shedding some light on their internal dynamics.

By far the photolabels preferentially used were s^4 U (RNA) and s^4 T or s^4 dU (DNA) which can be incorporated within nucleic acids with minimal structural perturbation. Their photophysics is reasonably well understood and the nature of the photoadducts they form with current bases well established, with the notable exception of G derived adducts. Both s^4 U and s^4 T react with U and C pyrimidines to yield either (5-4) or (6-4) photoadducts, depending upon the orientation of the C4-S4 bond with respect to the acceptor 5-6 double bond. As a future development in the field of RNA structure elucidation, identification of the nature, (6-4) versus (5-4), of the photoadducts present within defined crosslinked RNA chains (which requires development of highly sensitive methods) will be most interesting. Indeed, it would allow determination of the mutual orientation of the bases prior to the reaction, which will lead to refinement of the proposed RNA structures.

To date, little is known on the photophysical and photochemical behavior of thiopurine derivatives. Nevertheless, in the dinucleotide Tps⁶dI, s^6 dI yielded a (6-6) photoadduct whose structure is reminiscent of the (6-4) bipyrimidines. These derivatives represent attractive photolabels to substitute A, G and U (T) positions (s^6 I) or G positions (s^6 G) within nucleic acid chains. However, s^6 G and s^4 U display a number of distinctive properties. Indeed, s^6 G exhibits an absorption maximum close to 310 nm instead of 330 nm for

s^4U . Its ex-vivo incorporation is far more damaging to the cell metabolism than s^4U incorporation. Finally, s^6GTP is a much poorer substrate than s^4UTP in transcription reactions catalysed by T7 RNA polymerase. All these data suggest that ground state s^6G could exist in a thiol–thione equilibrium leading to misincorporation during RNA synthesis. Therefore, s^6G or s^6I substitution may lead to significant structural perturbations and the functionality of the substituted nucleic acids should be checked prior to structural analysis.

Although thionucleobases photocrosslink efficiently proteins within nucleoproteins, it should be kept in mind that little is known to date, even in model systems, of their photocoupling ability with amino-acid residues and of the nature of the photoadducts formed. At present, the major problem is the unambiguous identification of the amino-acid residues which had undergone crosslinking within an oligopeptide–oligonucleotide complex. However, this can now be tackled by a modern technique, namely matrix-assisted laser desorption/ionization mass spectrometry [107]. The data obtained can be considered as reliable when the photolabeled RNA (DNA) associates with high affinity to the protein components of the complex. If low affinity complexes are investigated, care should be taken that the crosslinks formed do not result from collisional photoreactions. No doubt that answers will be provided to the pending questions raised above within the next few years, together with the development of new applications of biological interest.

Acknowledgements

We are indebted to S. Dokudovskaya and the referees for a critical reading of the manuscript. This work was supported by grants from ANRS and CNRS (ACC-SV 5).

References

- [1] A. Favre, 4-Thiouridine as an intrinsic photoaffinity probe of nucleic acid structure and interactions, in H. Morrison (Ed.), *Bioorganic Photochemistry: Photochemistry and the Nucleic Acids*, Wiley, New York, 1990, pp. 379–425.
- [2] M.M. Hanna, Photoaffinity cross-linking methods for studying RNA–protein interactions, *Methods Enzymol.*, 180 (1989) 383–409.
- [3] K.M. Meisenheimer and T.H. Koch, Photocross-linking of nucleic acids to associated proteins, *Crit. Rev. Biochem. Mol. Biol.*, 32 (1997) 101–1040.
- [4] M.N. Lipsett, The isolation of 4-thiouridylic acid from the soluble ribonucleic acid of *Escherichia coli*, *J. Biol. Chem.*, 240 (1965) 3975–3978.
- [5] A. Favre, M. Yaniv and A.M. Michelson, The photochemistry of 4-thiouridine in *Escherichia coli* t-RNA Val1, *Biochem. Biophys. Res. Commun.*, 37 (1969) 266–271.
- [6] M. Yaniv, A. Favre and B.G. Barrell, Structure of transfer RNA. Evidence for interaction between two non-adjacent nucleotide residues in tRNA from *Escherichia coli*, *Nature*, 223 (1969) 1331–1333.
- [7] R.K. Kumar and D.R. Davis, Synthesis and studies on the effect of 2-thiouridine and 4-thiouridine on sugar conformation and RNA duplex stability, *Nucleic Acids Res.*, 25 (1997) 1272–1280.
- [8] B. Bartholomew, B.R. Braun, G.A. Kassavetis and E.P. Geiduschek, Probing close DNA contacts of RNA polymerase III transcription complexes with the photoactive nucleoside 4-thio-deoxythymidine, *J. Biol. Chem.*, 269 (1994) 18090–18095.
- [9] Y.L. Dubreuil, B.A. Expert and A. Favre, Conformation and structural fluctuations of a 218 nucleotides long rRNA fragment: 4-thiouridine as an intrinsic photolabelling probe, *Nucleic Acids Res.*, 19 (1991) 3653–3660.
- [10] P.V. Sergiev, I.N. Lavrik, V.A. Wlasoff, S.S. Dokudovskaya, O.A. Dontsova, A.A. Bogdanov and R. Brimacombe, The path of mRNA through the bacterial ribosome: a site-directed crosslinking study using new photoreactive derivatives of guanosine and uridine, *RNA*, 3 (1997) 464–475.
- [11] W.T. Melvin, H.B. Milne, A.A. Slater, H.J. Allen and H.M. Keir, Incorporation of 6-thioguanosine and 4-thiouridine into RNA. Application to isolation of newly synthesised RNA by affinity chromatography, *Eur. J. Biochem.*, 92 (1978) 373–379.
- [12] O. Dontsova, A. Kopylov and R. Brimacombe, The location of mRNA in the ribosomal 30S initiation complex; site-directed crosslinking of mRNA analogues carrying several photo-reactive labels simultaneously on either side of the AUG start codon, *EMBO J.*, 10 (1991) 2613–2620.
- [13] J.-L. Fourrey, J. Gasche, C. Fontaine, E. Guittet and A. Favre, Sequence dependent photochemistry of di(deoxynucleoside) phosphates containing 4-thiouracil, *J. Chem. Soc. Chem. Commun.*, (1989) 1334–1336.
- [14] A. Favre and J.-L. Fourrey, Structural probing of small endonucleolytic ribozymes in solution using thio-substituted nucleobases as intrinsic photolabels, *Acc. Chem. Res.*, 28 (1995) 375–382.
- [15] P. Clivio, J.-L. Fourrey, J. Gasche and A. Favre, Synthesis of dinucleoside phosphates containing sulfur substituted nucleobases: 4-thiouracil, 4-thiothymine and 6-mercaptapurine, *Tetrahedron Lett.*, 33 (1992) 69–72.
- [16] P. Clivio, J.-L. Fourrey and A. Favre, Synthesis of deoxydinucleoside phosphates containing 6-thio-substituted purine nucleobases, *J. Chem. Soc. Perkin Trans.*, 1 (1993) 2585–2590.
- [17] P. Clivio, J.-L. Fourrey, J. Gasche and A. Favre, Synthesis of deoxydinucleoside phosphates containing 4-thiosubstituted pyrimidine nucleobases, *J. Chem. Soc. Perkin Trans.*, 1 (1992) 2383–2388.
- [18] P. Clivio, J.-L. Fourrey, J. Gasche, A. Audic, A. Favre, C. Perrin and A. Woisard, Synthesis and purification of oligonucleotides containing sulfur substituted nucleobases: 4-thiouracil, 4-thiothymine and 6-mercaptapurine, *Tetrahedron Lett.*, 33 (1992) 65–68.
- [19] J.B. Murray, C.J. Adams, J.R. Arnold and P.G. Stockley, The roles of the conserved pyrimidine bases in hammerhead ribozyme catalysis: evidence for a magnesium ion-binding site, *Biochem. J.*, (1995) 487–494.
- [20] E.J. Sontheimer and J.A. Steitz, The U5 and U6 small nuclear RNAs as active site components of the spliceosome, *Science*, 262 (1993) 1989–1996.
- [21] E.J. Sontheimer, Site-specific RNA crosslinking with 4-thiouridine, *Mol. Biol. Rep.*, 20 (1994) 35–44.
- [22] P.V. Baranov, S.S. Dokudovskaya, T.S. Oretskaya, O.A. Dontsova, A.A. Bogdanov and R. Brimacombe, A new technique for the characterization of long-range tertiary contacts in large RNA molecules: insertion of a photolabel at a selected position in 16S rRNA within the *Escherichia coli* ribosome, *Nucleic Acids Res.*, 25 (1997) 2266–2273.
- [23] Y.T. Yu and J.A. Steitz, Site-specific crosslinking of mammalian U11 and u6atac to the 5' splice site of an AT-AC intron, *Proc. Natl. Acad. Sci. USA*, 94 (1997) 6030–6035.
- [24] K. Heihoff, R.W. Redmond, S.E. Braslavsky, M. Rougee, C. Salet, A. Favre and R.V. Bensasson, Quantum yields of triplet and $O_2(^1\Delta_g)$ formation of 4-thiouridine in water and acetonitrile, *Photochem. Photobiol.*, 51 (1990) 635–641.

- [25] J. Cadet and P. Vigny, The photochemistry of nucleic acids, in H. Morrison (Ed.), *Bioorganic Photochemistry: Photochemistry and the Nucleic Acids*, Wiley, New York, 1990, pp. 1–272.
- [26] N.J. Leonard, D.E. Bergstrom and G.L. Tolman, Photoproducts from 4-thiouracil and cytosine and from 4-thiouridine and cytidine: refinement of tertiary tRNA structure, *Biochem. Biophys. Res. Commun.*, 44 (1971) 1524–1530.
- [27] A. Favre and J.-L. Fourrey, Intramolecular cross-linking of single-stranded copolymers of 4-thiouridine and cytidine, *Biochem. Biophys. Res. Commun.*, 58 (1974) 507–515.
- [28] A. Favre, Y. Lemaigre Dubreuil and J.-L. Fourrey, Photocoupling of the intrinsic photoaffinity probe 4-thiouridine to nucleosides and oligonucleotides, *New J. Chem.*, 15 (1991) 593–599.
- [29] A. Ito, F.T. Robb, J.G. Peak and M.J. Peak, Base-specific damage induced by 4-thiouridine photosensitization with 334-nm radiation in M13 phage DNA, *Photochem. Photobiol.*, 47 (1988) 231–240.
- [30] M.J. Peak, W.R. Midden, D.M. Babasick and D.A. Haugen, Photochemistry of 4-thiouridine and thymine, *Photochem. Photobiol.*, 48 (1988) 229–232.
- [31] D.E. Bergstrom and N.J. Leonard, Photoreaction of 4-thiouracil with cytosine. Relation to photoreactions in *Escherichia coli* transfer ribonucleic acids, *Biochemistry*, 11 (1972) 1–9.
- [32] D.E. Bergstrom and N.J. Leonard, Structure of the borohydride reduction product of photolinked 4-thiouracil and cytosine. Fluorescent probe of transfer ribonucleic acid tertiary structure, *J. Am. Chem. Soc.*, 94 (1972) 6178–6184.
- [33] D.E. Bergstrom, I. Inoue and N.J. Leonard, Synthesis of the 335-nm photoproduct of cytosine and 4-thiouracil, *J. Org. Chem.*, 37 (1972) 3902–3907.
- [34] S.Y. Wang, Pyrimidine bimolecular photoproducts, in S.Y. Wang (Ed.), *Photochemistry and Photobiology of Nucleic Acids*, Academic Press, New York, 1976, pp. 296–356.
- [35] A. Favre, A.M. Michelson and M. Yaniv, Photochemistry of 4-thiouridine in *Escherichia coli* transfer RNA¹Val, *J. Mol. Biol.*, 58 (1971) 367–379.
- [36] A. Favre, B. Roques and J.-L. Fourrey, Chemical structures of TU-C and TU-C^{red} products derived from E. Coli tRNA, *FEBS Lett.*, 24 (1972) 209–214.
- [37] D.F. Rhoades and S.Y. Wang, Photochemistry of polycytidylic acid, deoxycytidine, and cytidine, *Biochemistry*, 10 (1971) 4603–4611.
- [38] D.F. Rhoades and S.Y. Wang, A new photoproduct of cytosine. Structure and mechanism studies, *J. Am. Chem. Soc.*, 93 (1971) 3779–3781.
- [39] A. Favre and M. Yaniv, Introduction of an intramolecular fluorescent probe in E. coli tRNA¹Val, *FEBS Lett.*, 17 (1971) 236–240.
- [40] G. Thomas, J.-L. Fourrey and A. Favre, Reduced 8-13 link, a viscosity-dependent fluorescent probe of transfer RNA tertiary structure, *Biochemistry*, 17 (1978) 4500–4508.
- [41] A. Jack, J.E. Ladner and A. Klug, Crystallographic refinement of yeast phenylalanine transfer RNA at 2–5 Å resolution, *J. Mol. Biol.*, 108 (1976) 619–649.
- [42] S.H. Kim, G.J. Quigley, F.L. Suddath, A. McPherson, D. Sneden, J.J. Kim, J. Weinzierl and A. Rich, Three-dimensional structure of yeast phenylalanine transfer RNA: folding of the polynucleotide chain, *Science*, 179 (1973) 285–288.
- [43] E.R. Blazek, J.L. Alderfer, W.A. Tabaczynski, V.C. Stamoudis, M.E. Churchill, J.G. Peak and M.J. Peak, A 5-4 pyrimidine-pyrimidone photoproduct produced from mixtures of thymine and 4-thiouridine irradiated with 334 nm light, *Photochem. Photobiol.*, 57 (1993) 255–265.
- [44] Clivio et al., unpublished data, (1998).
- [45] C. Saintomé, P. Clivio, A. Favre, J.-L. Fourrey and C. Riche, RNA photolabeling mechanistic studies: X-ray crystal structure of the photoproduct formed between 4-thiothymidine and adenosine upon near UV irradiation, *J. Am. Chem. Soc.*, 118 (1996) 8142–8143.
- [46] J.-L. Fourrey and P. Jouin, Thiocarbonyl photochemistry. I. Hydroxyalkylation of 1,3-dimethyl-4-thiouracil, *Tetrahedron Lett.*, (1973) 3225–3227.
- [47] J.-L. Fourrey and J. Moron, Thiocarbonyl photochemistry. VII. Photochemistry of 1,3-dimethyl-4-thiouracil in the presence of triethylamine, *Tetrahedron Lett.*, (1976) 301–304.
- [48] P. Clivio and J.-L. Fourrey, (6-4) Photoproduct DNA photolyase mechanistic studies using s⁵-(6-4) photoproducts, *J. Chem. Soc., Chem. Commun.* (1996) 2203–2204.
- [49] J. Liu and J.-S. Taylor, Remarkable photoreversal of a thio analog of the Dewar valence isomer of the (6-4) photoproduct of DNA to the parent nucleotides, *J. Am. Chem. Soc.*, 118 (1996) 3287–3288.
- [50] P. Clivio, J.-L. Fourrey, T. Szabo and J. Stawinsky, Photochemistry of di(deoxyribonucleoside) methylphosphonates containing N¹-methyl-4-thiothymine, *J. Org. Chem.*, 59 (1994) 7273–7283.
- [51] P. Clivio, D. Guillaume, M.-T. Adeline and J.-L. Fourrey, A photochemical approach to highlight backbone effects in PNA, *J. Am. Chem. Soc.*, 118 (1997) 8142–8143.
- [52] A. Woisard, A. Favre, P. Clivio and J.-L. Fourrey, Hammerhead ribozyme tertiary folding: Intrinsic photolabeling studies, *J. Am. Chem. Soc.*, 114 (1992) 10072–10074.
- [53] C. Saintomé, P. Clivio, A. Favre, J.-L. Fourrey and P. Laugãa, Site specific cross-linking of single stranded oligonucleotides by a complementary sequence equipped with an internal photoactive probe, *J. Chem. Soc., Chem. Commun.* (1997) 167–168.
- [54] M. Tanaka, D. Dos Santos, P. Laugãa, A. Favre and J.-L. Fourrey, The RNA substrate of the tobacco ringspot virus derived hairpin ribozyme and its DNA analog adopt distinct ordered structures in solution: An intrinsic photoaffinity study, *New J. Chem.*, 19 (1995) 469–478.
- [55] J.A. Wise, Guides to the heart of spliceosome, *Science*, 262 (1993) 1978–1979.
- [56] A.J. Newman, S. Teigelkamp and J.D. Beggs, snRNA interactions at 5' and 3' splice sites monitored by photoactivated crosslinking in yeast spliceosome, *RNA*, 1 (1995) 968–980.
- [57] C.H. Kim and J. Abelson, Site-specific crosslinks of yeast U6 snRNA to the pre-mRNA near the 5' splice site, *RNA*, 2 (1996) 995–1010.
- [58] P. Wollenzien, A. Expert-Bezancon and A. Favre, Sites of contact of mRNA with 16S rRNA and 23S rRNA in the *Escherichia coli* ribosome, *Biochemistry*, 30 (1991) 1788–1795.
- [59] R. Bhangu and P. Wollenzien, The mRNA binding track in the *Escherichia coli* ribosome for mRNAs of different sequences, *Biochemistry*, 31 (1992) 5937–5944.
- [60] R. Bhangu, D. Juzumiene and P. Wollenzien, Arrangement of messenger RNA on *Escherichia coli* ribosomes with respect to 10 16S rRNA cross-linking sites, *Biochemistry*, 33 (1994) 3063–3070.
- [61] D.J. Juzumiene, T.G. Shapkina and P. Wollenzien, Distribution of cross-links between mRNA analogues and 16 S rRNA in *Escherichia coli* 70 S ribosomes made under equilibrium conditions and their response to tRNA binding, *J. Biol. Chem.*, 270 (1995) 12794–12800.
- [62] O. Dontsova, S. Dokudovskaya, A. Kopylov, A. Bogdanov, A.J. Rinke, N. Junke and R. Brimacombe, Three widely separated positions in the 16S RNA lie in or close to the ribosomal decoding region: a site-directed cross-linking study with mRNA analogues, *EMBO J.*, 11 (1992) 3105–3116.
- [63] S.S. Dokudovskaya, O.A. Dontsova, S.L. Bogdanova, A.A. Bogdanov and R. Brimacombe, mRNA-ribosome interactions, *Biotechnol. Appl. Biochem.*, (1993) 149–155.
- [64] O. Dontsova, V. Tishkov, S. Dokudovskaya, A. Bogdanov, T. Doring, A.J. Rinke, S. Thamm, B. Greuer and R. Brimacombe, Stem-loop IV of 5S rRNA lies close to the peptidyltransferase center, *Proc. Natl. Acad. Sci. USA*, 91 (1994) 4125–4129.
- [65] A. Bogdanov, O. Dontsova and R. Brimacombe, Messenger RNA path through the prokaryotic ribosome, in: K. Nierhaus, F. Franceschi, A.R. Subramanian, V.A. Erdmann and B. Wittman-Liebold

- (Eds.), The translational apparatus, Plenum Press, New York, 1993, pp. 421–433.
- [66] J. Rinke Appel, N. Junke, R. Brimacombe, S. Dukudovskaya, O. Dontsova and A. Bogdanov, Site-directed cross-linking of mRNA analogues to 16S ribosomal RNA; a complete scan of cross-links from all positions between +1' and +16' on the mRNA, downstream from the decoding site, *Nucleic Acids Res.*, 21 (1993) 2853–1859.
- [67] J. Rinke Appel, N. Junke, R. Brimacombe, I. Lavrik, S. Dokudovskaya, O. Dontsova and A. Bogdanov, Contacts between 16S ribosomal RNA and mRNA, within the spacer region separating the AUG initiator codon and the Shine-Dalgarno sequence: a site-directed cross-linking study, *Nucleic Acids Res.*, 22 (1994) 3018–3025.
- [68] D.M. Graifer, D.I. Juzumiene, G.G. Karpova and P. Wollenzien, mRNA binding track in the human 80S ribosome for mRNA analogues randomly substituted with 4-thiouridine residues, *Biochemistry*, 33 (1994) 6201–6206.
- [69] S. Dokudovskaya, O. Dontsova, O. Shpanchenko, A. Bogdanov and R. Brimacombe, Loop IV of 5S ribosomal RNA has contacts both to domain II and to domain V of the 23S RNA, *RNA*, 2 (1996) 146–152.
- [70] K.R. Birikh, P.A. Heaton and F. Eckstein, The structure, function and application of the hammerhead ribozyme, *Eur. J. Biochem.*, 245 (1997) 1–16.
- [71] F. Eckstein, The hammerhead ribozyme, *Biochem. Soc. Trans.*, 24 (1996) 601–604.
- [72] J.M. Burke, Hairpin ribozyme: current status and future prospects, *Biochem. Soc. Trans.*, 24 (1996) 608–615.
- [73] J. Bartolome, A. Madejon and V. Carreno, Ribozymes: structure, characteristics and use as potential antiviral agents, *J. Hepatol.*, (1995) 57–64.
- [74] M.D. Been, Cis- and trans-acting ribozymes from a human pathogen, hepatitis delta virus, *Trends Biochem. Sci.*, 19 (1994) 251–256.
- [75] K.J. Hertel, A. Pardi, O.C. Uhlenbeck, M. Koizumi, E. Ohtsuka, S. Uesugi, R. Cedergren, F. Eckstein, W.L. Gerlach, R. Hodgson and R.M. Symons, *Nucl. Acids Res.*, 20 (1992) 3252.
- [76] A. Woisard, J.-L. Fourrey and A. Favre, Multiple folded conformations of a hammerhead ribozyme domain under cleavage conditions, *J. Mol. Biol.*, 239 (1994) 366–370.
- [77] P. Laugaa, A. Woisard, J.-L. Fourrey and A. Favre, Hammerhead ribozyme: a three dimensional model based on photo-crosslinking data, *C.R. Acad. Sci. Paris, Life Sci.*, 318 (1995) 307–313.
- [78] H.W. Pley, K.M. Flaherty and D.B. McKay, Three-dimensional structure of a hammerhead ribozyme, *Nature*, 372 (1994) 68–74.
- [79] W.G. Scott, J.T. Finch and A. Klug, The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage, *Cell*, 81 (1995) 991–1002.
- [80] W.G. Scott and A. Klug, Ribozymes: structure and mechanism in RNA catalysis, *Trends Biochem. Sci.*, 21 (1996) 220–224.
- [81] W.G. Scott, J.B. Murray, J. Arnold, B.L. Stoddard and A. Klug, Capturing the structure of a catalytic RNA intermediate: the hammerhead ribozyme, *Science*, 274 (1996) 2065–2069.
- [82] V.D. Dos Santos, J.-L. Fourrey and A. Favre, Flexibility of the bulge formed between a hairpin ribozyme and deoxy- substrate analogues, *Biochem. Biophys. Res. Commun.*, 190 (1993) 377–385.
- [83] V.D. Dos Santos, A.L. Vianna, J.-L. Fourrey and A. Favre, Folding of DNA substrate-hairpin ribozyme domains: use of deoxy 4-thiouridine as an intrinsic photolabel, *Nucleic Acids Res.*, 21 (1993) 201–207.
- [84] C. Bravo, F. Lescure, P. Laugaa, J.-L. Fourrey and A. Favre, Folding of the HDV antigenomic ribozyme pseudoknot structure deduced from long-range photocrosslinks, *Nucleic Acids Res.*, 24 (1996) 1351–1359.
- [85] N.K. Tanner, S. Schaff, G. Thill, K.E. Petit, D.A. Crain and E. Westhof, A three-dimensional model of hepatitis delta virus ribozyme based on biochemical and mutational analyses, *Curr. Biol.*, 4 (1994) 488–498.
- [86] S. Ito, I. Saito, A. Nakata and T. Matsuura, Photochemical addition of L-lysine to 1,3-dimethyl-4-thiouracil, *Photochem. Photobiol.*, 32 (1980) 683–685.
- [87] R. Bensasson, C. Salet and A. Favre, unpublished data. (1998).
- [88] T.T. Nikiforov and B.A. Connolly, Oligodeoxynucleotides containing 4-thiothymidine and 6-thiothiouracil as affinity labels for the Eco RV restriction endonuclease and modification methylase, *Nucleic Acids Res.*, 20 (1992) 1209–1214.
- [89] C.D. Thomas, T.T. Nikiforov, B.A. Connolly and W.V. Shaw, Determination of sequence specificity between a plasmid replication initiator protein and the origin of replication, *J. Mol. Biol.*, 254 (1995) 381–391.
- [90] N. Sheng and D. Dennis, Active site labeling of HIV-1 reverse transcriptase, *Biochemistry*, 32 (1993) 4938–4942.
- [91] Y. Mishima and J.A. Steitz, Site-specific crosslinking of 4-thiouridine-modified human tRNA(3Lys) to reverse transcriptase from human immunodeficiency virus type I, *EMBO J.*, 14 (1995) 2679–2687.
- [92] S.B. Renwick, A.D. Critchley, C.J. Adams, S.M. Kelly, N.C. Price and P.G. Stockley, Probing the details of the HIV-1 Rev-Rev-responsive element interaction: effects of modified nucleotides on protein affinity and conformational changes during complex formation, *Biochem. J.*, (1995) 447–453.
- [93] A. McGregor, M.V. Rao, G. Duckworth, P.G. Stockley and B.A. Connolly, Preparation of oligoribonucleotides containing 4-thiouridine using Fpmp chemistry. Photo crosslinking to RNA binding proteins using 350 nm irradiation, *Nucleic Acids Res.*, 24 (1996) 3173–3180.
- [94] Z. Wang and T.M. Rana, RNA conformation in the Tat–TAR complex determined by site-specific photo-cross-linking, *Biochemistry*, 35 (1996) 6491–6499.
- [95] M. Roberge and E.M. Bradbury, RNA contacts the two large polymerase subunits and a 52-kDa polypeptide in nucleolar RNA polymerase I transcribing complexes, *J. Biol. Chem.*, 263 (1988) 18553–18557.
- [96] B. Bartholomew, M.E. Dahmus and C.F. Meares, RNA contacts subunits Ilo and Iic in HeLa RNA polymerase II transcription complexes, *J. Biol. Chem.*, 261 (1986) 14226–14231.
- [97] D.L. Cadena and M.E. Dahmus, Messenger RNA synthesis in mammalian cells is catalyzed by the phosphorylated form of RNA polymerase II, *J. Biol. Chem.*, 262 (1987) 12468–12474.
- [98] A.B. Chapman and N. Agabian, Trypanosoma brucei RNA polymerase II is phosphorylated in the absence of carboxyl-terminal domain heptapeptide repeats, *J. Biol. Chem.*, 269 (1994) 4754–4760.
- [99] N.C. Khanna, S. Lakhani and K.K. Tewari, Photoaffinity labelling of the pea chloroplast transcriptional complex by nascent RNA in vitro, *Nucleic Acids Res.*, 19 (1991) 4849–4855.
- [100] S. Kelly, N. Sheng and D. Dennis, Yeast RNA polymerase I. Derivatization of the 190 and 135 subunits by 4-thiouridine monophosphate positioned uniquely at the 3' terminus of an enzyme-bound ³²P-containing transcript initiated by a triribonucleotide primer on synthetic single-stranded DNA, *J. Biol. Chem.*, 265 (1990) 7787–7792.
- [101] N. Sheng, E.B. Mougey, S. Kelly and D. Dennis, Wheat germ and yeast RNA polymerase II: photoaffinity labeling by 4-thiouracil 5'-monophosphate positioned uniquely at the 3' end of an enzyme-bound [³²P]-containing transcript, *Biochemistry*, 32 (1993) 2248–2253.
- [102] W. Powell, B. Bartholomew and D. Reines, Elongation factor SII contacts the 3'-end of RNA in the RNA polymerase II elongation complex, *J. Biol. Chem.*, 271 (1996) 22301–22304.
- [103] S. McBratney and P. Sarnow, Evidence for involvement of trans-

- acting factors in selection of the AUG start codon during eukaryotic translational initiation, *Mol. Cell. Biol.*, 16 (1996) 3523–3534.
- [104] K.V. Rosen, R.W. Alexander, J. Wower and R.A. Zimmermann, Mapping the central fold of tRNA²(fMet) in the P site of the *Escherichia coli* ribosome, *Biochemistry*, 32 (1993) 12802–12811.
- [105] C.M. Brown and W.P. Tate, Direct recognition of mRNA stop signals by *Escherichia coli* polypeptide chain release factor two, *J. Biol. Chem.*, 269 (1994) 33164–33170.
- [106] W. Tate, B. Greuer and R. Brimacombe, Codon recognition in polypeptide chain termination: site directed crosslinking of termination codon to *Escherichia coli* release factor 2, *Nucleic Acids Res.*, 18 (1990) 6537–6544.
- [107] H. Urlaub, B. Thiede, E.C. Muller, R. Brimacombe and L.B. Wittmann, Identification and sequence analysis of contact sites between ribosomal proteins and rRNA in *Escherichia coli* 30 S subunits by a new approach using matrix-assisted laser desorption/ionization-mass spectrometry combined with N-terminal microsequencing, *J. Biol. Chem.*, 272 (1997) 14547–14555.