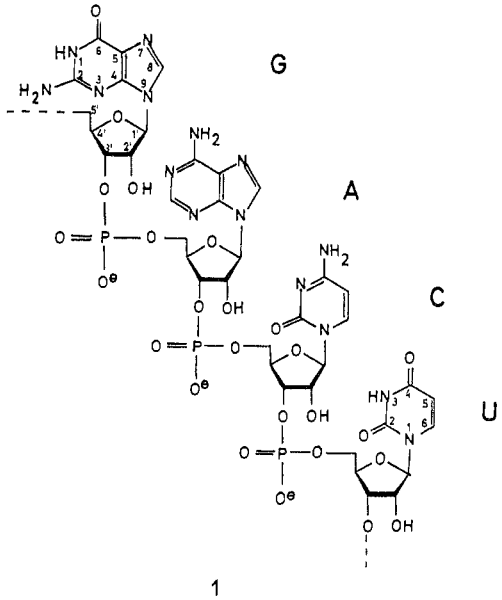
Some Modified Nucleotides as Tools in Nucleic Acid Research[1]

FRIEDRICH CRAMER

Max-Planck-Institut für experimentelle Medizin, Göttingen, Germany

Received April 30, 1969

Naturally occurring nucleotides, the monomeric units of nucleic acids, are relatively uniform in their chemical character in comparison with the amino acids, the basic units of proteins. There are only four major nuclear bases, adenine, guanine, cytosine, and uracil, in RNA (**1**) (or instead of uracil its 5-methyl derivative thymine in DNA) attached to the sugar phosphate backbone. In contrast, 20 different amino acid residues are responsible for the variety of polypeptide chains, not even taking into account the numerous rare amino acids. Only in the tRNAs do some of the nuclear bases



1

occur in modified form, *e.g.*, 5,6-dihydrouracil, N^6-alkyladenine, 4-thiouracil; occasionally some of the ribose residues are methylated.

Chemical reactions leading to modified nucleotides can be important in many respects. (1) Modified nucleotides may be substrates or inhibitors of nucleotide-metabolizing enzymes and thus have potentialities in the elucidation of the mechanisms of these enzymes; *vice versa*, such nucleotides may throw light on the mode of action of naturally occurring nucleotide antimetabolites and antibiotics. (2) When modified nucleotides are incorporated into chemically or biochemically synthesized polymers, they may serve as chromophoric probes in physicochemical studies or in other ways as subjects for various biochemical studies. (3) The chemical modification of nucleotides within a macromolecular structure can influence the physical structure of a nucleic acid and may alter its biological properties. It is known, for instance, that deamination of nuclear bases is mutagenic.[2] The decreased reactivity of certain groups within the macromolecular framework might be a measure of the "degree of exposedness," thus allowing conclusions to be drawn on the folding of the macromolecule.

For each of these aspects examples are given, many from this laboratory. For more comprehensive reviews the reader is referred to some recent publications.[3,4]

Synthesis and Properties of Some Modified Nucleotides

A. 4-Thiopyrimidine Nucleotides. tRNA from *E. coli* contains a small amount of 4-thiouridine,[5] the function of which is, as yet, unclear. As a monomer in $CDCl_3$ the base can still pair with adenosine.[6] Perhaps it is of importance in the formation or stabilization of secondary or tertiary structure of a tRNA since the conformation of the heterocycle with respect to the sugar is different from the other nucleosides, at least in the crystalline state, as shown by X-ray analysis.[7] In the crystal the 4-thiouracil moieties form stacks in which the heterocycle has the *syn* conformation with respect to the sugar moiety.

4-Thiouridine can be synthesized by introducing sulfur into uridine using phosphorus pentasulfide in pyridine.[8] Phosphorylation of the suitably protected nucleoside and subsequent formation of the nucleoside di- and triphosphate can, in principle, be achieved in various ways. The easiest seems to be the use of the phosphoimidazolide method.[9] A nucleoside is allowed

(1) Abbreviations used in this paper: nucleosides: G, guanosine; A, adenosine; C, cytidine; U, uridine; ψ, pseudouridine; nucleotides: GMP, AMP, CMP, and UMP are the corresponding monophosphates; GDP, ADP, CDP, and UDP are the corresponding diphosphates; GTP, ATP, CTP, and UTP are the corresponding triphosphates. Formulas of the deoxy series are indicated by a d: *e.g.*, dT, deoxythymidine. Modified nuclear bases: s^4U, 4-thiouridine, ds^4T, deoxy-4-thiothymidine. Polymers: poly d(A-T) = d-(pApTpApT...) = polymer with alternating sequence of dA and dT; similar abbreviations are used for other polymers, DNA, deoxyribonucleic acid, tRNA, transfer ribonucleic acid, $tRNA^{Phe}$ phenylalanine-specific tRNA.

(2) A. Gierer and K. W. Mundry, *Nature*, **182**, 1457 (1958); *Z. Vererbungslehre*, **89**, 614 (1958); H. Schuster and G. Schramm, *Z. Naturforsch.*, **13b**, 485 (1958).

(3) L. Grossman and K. Moldave, *Methods Enzymol.*, **12**, 2 (1967); **13**, 461 (1968); T. Ueda and J. J. Fox, *Advan. Carbohyd. Res.*, **22**, 307 (1967).

(4) D. R. Kreig, *Progr. Nucleic Acid Res. Mol. Biol.*, **2**, 125 (1963); J. Skoda, *ibid.*, **2**, 197 (1963); C. Heidelberger, *ibid.*, **4**, 2 (1965); R. W. Chambers, *ibid.*, **5**, 350 (1966); E. Goldwasser and R. L. Heinrikson, *ibid.*, **5**, 399 (1966); J. H. Phillips and D. M. Brown, *ibid.*, **7**, 349 (1967); G. B. Brown, *ibid.*, **8**, 209 (1968).

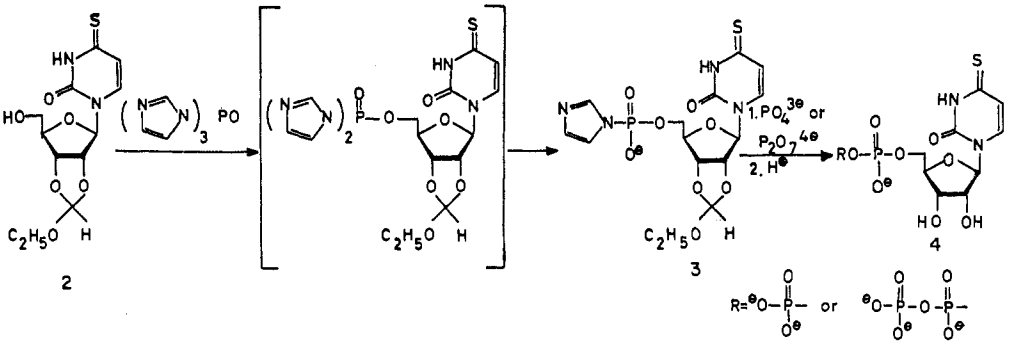
(5) M. N. Lipsett, *J. Biol. Chem.*, **240**, 3975 (1965).

(6) K.-H. Scheit, *Angew. Chem.*, **79**, 190 (1967); *Angew. Chem. Intern. Ed. Engl.*, **6**, 180 (1967).

(7) W. Saegner and K.-H. Scheit, *Angew. Chem.*, **81**, 121 (1969); *Angew. Chem. Intern. Ed., Engl.*, **8**, 139 (1969).

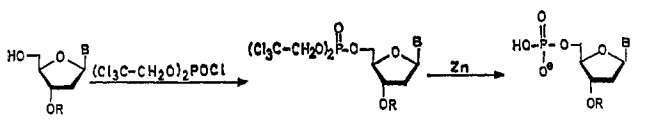
(8) J. J. Fox, D. van Praag, I. Wempen, I. L. Doerr, L. Cheong, J. E. Knoll, M. L. Eidinoff, A. Bendich, and G. B. Brown, *J. Amer. Chem. Soc.*, **81**, 178 (1959).

(9) F. Cramer and H. Schaller, *Chem. Ber.*, **94**, 1634 (1961), F. Cramer and H. Neunhoeffer, *ibid.*, **95**, 1664 (1962).



2 3 4

to react with triimidazolylphosphine oxide. The reaction product, a nucleotide imidazolide,[8] is at the same time a sufficiently activated intermediate for the subsequent reaction with orthophosphoric acid to give the nucleoside diphosphate or with pyrophosphoric acid to give the nucleoside triphosphate (**2**–**4**). Using this sequence 4-thiouridine 5′-diphosphate, 4-thiouridine 5′-triphosphate, and deoxy-4-thiothymidine 5′-triphosphate were obtained.[10] Similar reactions can be carried out with almost any nucleoside. In some cases the route shown below may be preferred.[11]



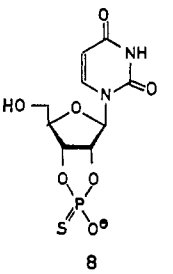
B = nuclear base
R = protecting group

5 6 7

The remarkable feature about the 4-thiopyrimidine nucleotides is their uv spectra,[12] which are very different from those of other nucleotides exhibiting a strong absorption band in the region of 330 mμ. This is a great advantage in studies on enzyme mechanisms, *e.g.*, polymerases and nucleases (see below).

B. Nucleoside Phosphorothioates. A different type of modification may be envisaged by modifying the phosphate group. This, if incorporated, *e.g.*, in an internucleotidic bond, might lead to an oligonucleotide with different stability. Recently nucleoside phosphorothioates[13] as well as diphosphates and triphosphates[14] became available by a modified imidazolide procedure using triimidazolyl-1-phosphine sulfide.

Uridine 2′,3′-O,O-cyclophosphorothioate[15] (**8**) contains an asymmetric phosphorus atom which is bound to the optically active ribose. It is, therefore, a mixture of diastereomers which can be separated.[16] On incubation with pancreatic ribonuclease[16] the higher melting isomer exhibits the same K_m value as the normal substrate, its



8

O analog, which indicates that the sulfur does not play an important role in the binding of this isomer to the enzyme; in contrast the K_m of the lower melting isomer is six times larger than that of the natural substrate. Although a pentacovalent intermediate in ribonuclease-catalyzed hydrolysis cannot be ruled out, pseudorotation does not occur, since ^{35}S-labeled nucleotide was used and no exchange of sulfur was found.

Nucleoside phosphorothioates can be used in chemical oligonucleotide synthesis. Thus dithymidine- (3′–5′)-phosphorothioate can be prepared in 18% yield by condensation of the suitable protected monomeric derivatives in the presence of 2 equiv of triisopropylbenzenesulfonyl chloride.[17]

Incorporation of Modified Nucleotides into Polymers

Although template-dependent polymerases have a surprising degree of fidelity in copying nucleic acid strands both *in vivo* and *in vitro*, it is possible to introduce modified bases, provided that these are similar in geometry and hydrogen-bonding capacity to the natural substrates. Thus 5-bromouridine can be used in place of thymidine; this enabled one to differentiate "light" and "heavy" strands in gradient centrifugation.[18]

A. DNA Polymerase. Deoxy-4-thiothymidine 5′-triphosphate is incorporated with DNA polymerase when poly d(A-T) is used as a template.[19] The product contains the sequence d(pAps^{4}TpAps^{4}T...). The new alternating polymer is a primer for poly d(A-T) synthesis (with dATP and dTTP as substrate) and does not prime the synthesis of new poly d(A-s^{4}T) (Figure 1). Thus the checking system of the polymerase seems to tolerate a

(10) K.-H. Scheit, *Chem. Ber.*, **101**, 1141 (1968).
(11) F. Eckstein and K.-H. Scheit, *Angew. Chem.*, **79**, 317 (1967); *Angew. Chem. Intern. Ed. Engl.*, **6**, 362 (1967); A. Franke, K.-H. Scheit, and F. Eckstein, *Chem. Ber.*, **101**, 2998 (1968); F. Eckstein, *ibid.*, **100**, 2228, 2236 (1967); K.-H. Scheit, *Biochim. Biophys. Acta*, **157**, 632 (1968).
(12) K.-H. Scheit, *ibid.*, **166**, 285 (1968).
(13) F. Eckstein, *J. Amer. Chem. Soc.*, **88**, 4292 (1966).
(14) F. Eckstein and H. Gindl, *Biochim. Biophys. Acta*, **149**, 35 (1967).
(15) F. Eckstein and H. Gindl, *Chem. Ber.*, **101**, 1670 (1968).
(16) F. Eckstein, *FEBS Lett.*, **2**, 85 (1968).
(17) F. Eckstein, *Tetrahedron Lett.*, 1157 (1967).
(18) M. Chamberlin, R. L. Baldwin, and P. Berg, *J. Mol. Biol.*, **7**, 334 (1963).
(19) A. G. Lezius and K.-H. Scheit, *Europ. J. Biochem.*, **3**, 85 (1967).

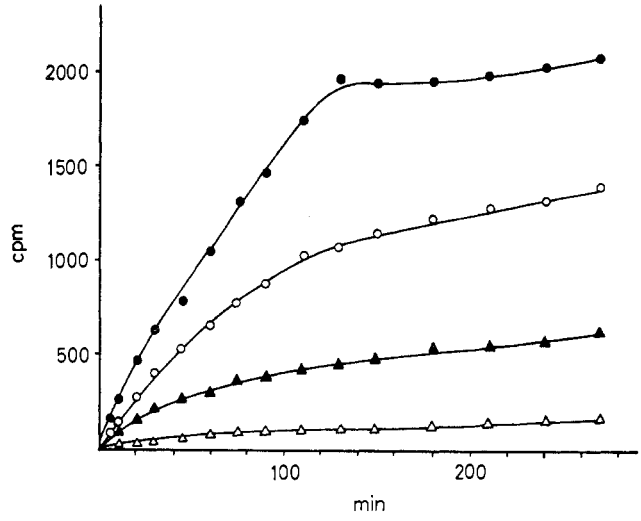


Figure 1. Synthesis of poly d(A-T) and poly d(A-s⁴T) with DNA polymerase and various substrates (●, ▲, dATP + dTTP; ○, △, dATP + ds⁴TTP) and templates (●, ○, poly d(A-T); ▲, △, poly d(A-s⁴T)).[19]

chemical modification on either the substrate or the template site (not equally well), but not on both sites. The fact that the poly d(A-s⁴T) abolishes the function of a template for extensive synthesis can be used for studies on the mechanism of DNA polymerase and its accompanying exonucleases.[20]

B. RNA Polymerase. RNA polymerase shows a comparable behavior with 4-thiouridine 5′-triphosphate[21] (see Table I). Again the checking system seems

Table I
Function of RNA Polymerase with Various Substrates and Templates Containing Modified Pyrimidines[21]

Substrate	Template	Synthesis
s⁴UTP + ATP	Poly d(A-T)	++
UTP + ATP	Poly d(A-s⁴T)	+
s⁴UTP + ATP	Poly d(A-s⁴T)	−
s⁴UTP	Poly dA·poly dT	−
s⁴UTP + ATP + GTP + CTP	Calf thymus DNA	−

to tolerate only one-sided alterations. Moreover, homopolymers or clusters of 4-thiouridine seem to be impossible, since the homopolymer strand poly dA is not copied. Since in natural DNA adjacent adenosine residues occur frequently, 4-thiouridine may bring the synthesis to a halt. Degradation of the 4-thiouridine-containing polymer by ribonuclease seems to proceed normally.

The alternating ribopolymer poly (A-s⁴U) has—as does poly (dA-s⁴T)—interesting spectroscopic properties. The 4-thiouridine chromophore at 330 mμ exhibits an enormous hypochromicity on incorporation into the polymer molecule (Figure 2). Since the absorption of the chromophore is far outside the usual region for nucleotide or protein absorption, enzymatic or physical changes of this polymer can easily be followed, even in

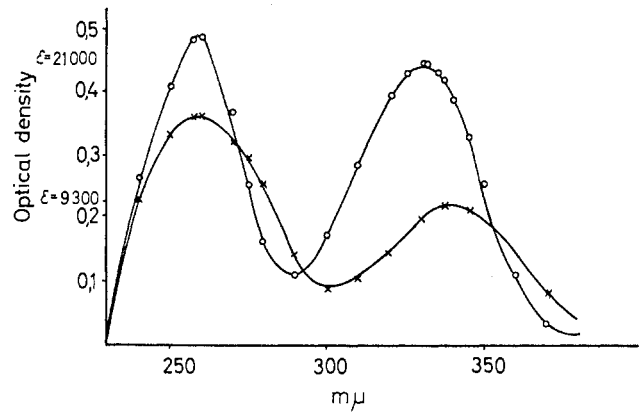
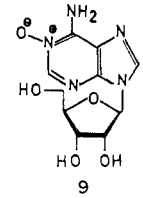


Figure 2. Ultraviolet spectrum of poly (A-s⁴U) before (—×—×—) and after (—○—○—) ribonuclease digestion.[21]

the presence of a large excess of enzyme protein. As an example the spectroscopic assay of poly (A-s⁴U) synthesis is given (Figure 3). The melting profile of this polymer at two wavelengths, 260 and 330 mμ, is shown in Figure 4.

A polynucleotide containing alternating ≻P═O and ≻P═S linkages[22] as well as a polynucleotide containing a thiophosphate backbone[23] was synthesized by RNA polymerase using a poly d(A-T) template and either ATP and uridine 5′-triphosphorothioate or adenosine 5′-triphosphorothioate and uridine 5′-triphosphorothioate as substrates. Both were degraded by catabolic enzymes at considerably slower rates than corresponding unmodified polymers.

C. Polynucleotide Phosphorylase. Polynucleotide phosphorylase, which catalyzes the reaction *n*nucleoside 5′-diphosphate ⇌ polynucleotide + *n*phosphate, is a nontemplate-dependent enzyme: it can therefore tolerate considerable changes in the substrate.[24] In many cases, however, the modified substrate cannot form a homopolymer by itself but requires another nucleotide for copolymerization, thus "diluting out the mistake." This is, for instance, the case with the 5′-diphosphates of adenosine 1-N-oxide[24] (**9**) or 4-thiouridine.[25]



9

In the case of 4-thiouridine 5′-diphosphate polymerization is inhibited by increasing amounts of this substrate and by the sulfur-containing polynucleotide products. When the incubation mixture contains UDP and 4-thiouridine 5′-diphosphate in the ratio of 1:1 the polymer has the ratio of 1.6:1.[25] The assay for poly-

(20) A. G. Lezius and E. Metz, *Europ. J. Biochem.*, **9**, 325 (1969).
(21) F. Cramer, H. Matzura, and K.-H. Scheit, in preparation.
(22) H. Matzura and F. Eckstein, *Europ. J. Biochem.*, **3**, 448 (1968).
(23) F. Eckstein and H. Gindl, *FEBS Lett.*, **2**, 262 (1969).
(24) For a summary see: F. Cramer, K. Randerath, and E. A. Schäfer, *Biochim. Biophys. Acta*, **72**, 150 (1963).
(25) K.-H. Scheit and E. Gaertner, *ibid.*, **182**, 1, 10 (1969).

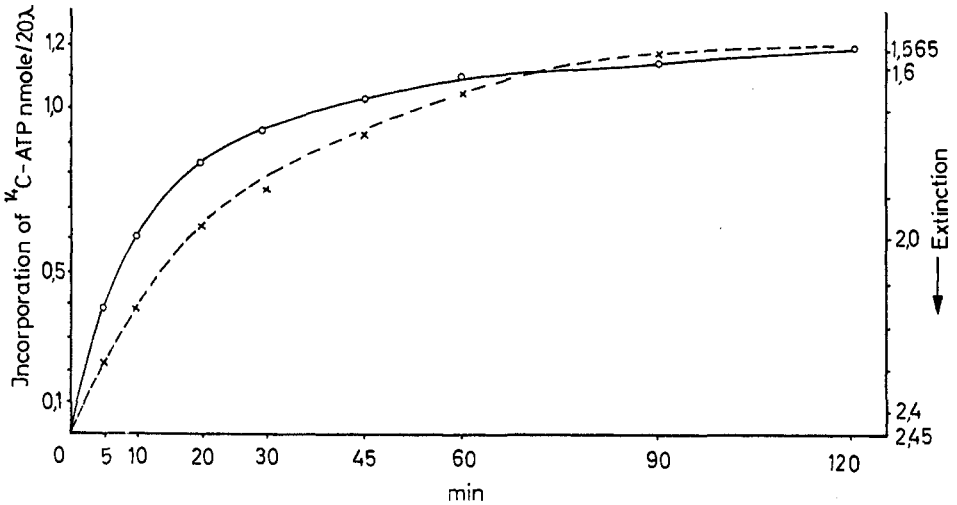


Figure 3. Synthesis of poly (A-s^{4}U) by RNA polymerase assayed by incorporation of ^{14}C-labeled ATP (—) and hypochromicity (- - -).[21]

phenylalanine synthesis[26] with such polymers in an *in vitro* system reveals that the partial exchange of the 4-oxo function by a 4-thio function[25] does not abolish the messenger function. Uridine 5′-diphosphorothioate can serve as a sole substrate for polynucleotide phosphorylase under certain conditions:[23] *n*uridine 5′-diphos-

phorothioate → ...UpUpU.... There is a remarkable reduction in rate of enzymatic degradation of this polymer by either snake venom or spleen phosphodiesterase or pancreatic ribonuclease. The polymer has 45% of the messenger activity of poly U in the above-mentioned assay.

Chemical Modification at the Polymer Level

Chemical modification can also be carried out at the polymer level. Here modifications of tRNA are mainly considered, since the sequence of some of these molecules is known and the function of them is comparatively well understood. Many of these reagents, when applied randomly, eliminate the biological activity of tRNA, namely either the codon recognition by reaction with positions in or adjacent to the anticodon or the amino acid acceptor capacity (charging) by reaction with the CCA end or some other part of the molecule. Some of these modifications can, however, be applied selectively. From these studies there is now ample evidence that an intact anticodon region is not required for the charging with amino acid. When tRNA^{Val} [27] or tRNA^{Phe} [28] is cut in or adjacent to the anticodon by enzymatic or chemical reaction, the modified molecule

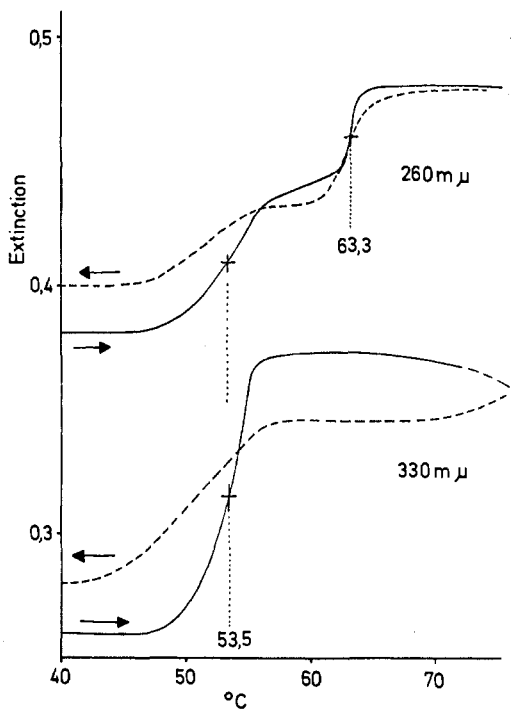


Figure 4. Melting of poly (A-s^{4}U) recorded at two wavelengths.[21]

can still be charged with its amino acid. Also, N-oxidation of the adenosine residues in the anticodon loop of tRNA^{Phe} does not remove chargeability.[29] When tRNA^{Ser} is iodinated at the double bond of N^6-isopentenyladenosine, serine recognition is still preserved.[30] Modification at the CCA end seems to be more controversial. N-Oxidation of the terminal adenosine in tRNA^{Phe} kills charging.[29] One particular photochemical event in the "stem region" inactivates tRNA^{Ala}.[31] On the other hand, the terminal adenosine can be split

(26) A. Parmeggiani, *Biochem. Biophys. Res. Commun.*, **20**, 613 (1968).
(27) A. A. Bayev, T. V. Venkstern, A. D. Mirzabekov, A. I. Krutilina, V. D. Axelrod, L. Li, I. Fodor, L. Ya. Kasarinova, and V. A. Engelhardt, "Structure and Function of Transfer RNA and 5s-RNA," Universitetsforlaget, Oslo, and Academic Press, London and New York, 1968, p 17.
(28) P. Philippsen, R. Thiebe, W. Wintermeyer, and H. G. Zachau, *Biochem. Biophys. Res. Commun.*, **33**, 922 (1968).
(29) F. Cramer, F. v.d. Haar, and E. Schlimme, unpublished results.
(30) F. Fittler and R. H. Hall, *Biochem. Biophys. Res. Commun.*, **25**, 441 (1966).
(31) L. H. Schulman and R. W. Chambers, *Proc. Natl. Acad. Sci. U. S.*, **61**, 308 (1968).

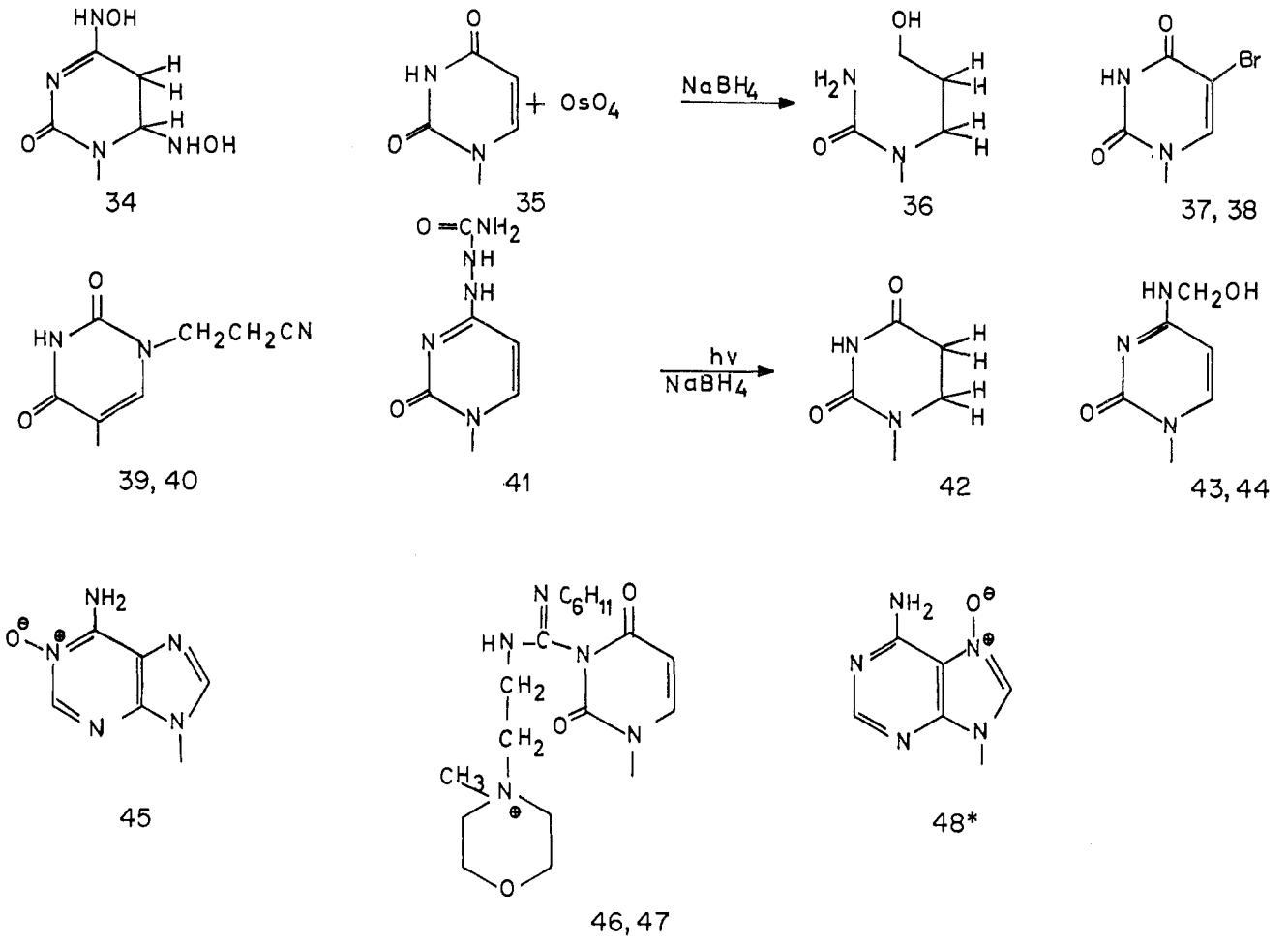


Figure 5. Chemical modifications of nuclear bases in tRNAs (*, in oligonucleotides).

between C-2′ and C-3′[32] without an appreciable loss of acceptor activity. Also introduction of a phosphorothioate into the last internucleotidic bond does not harm the function of the molecule.[33] In Figures 5 and 6[34-50] some of these investigations are depicted.

Some of the above-described modifying reagents show a certain specificity with respect to base pairing or exposedness of the base to undergo reaction. Since the folding (tertiary structure) of tRNA is an important still-unsolved question, these reagents were applied to study the topology of the molecule. $tRNA^{Ala}$ was allowed to react with a water-soluble carbodiimide in the presence of Mg^{2+} at different temperatures. The reaction comes to a plateau at 2° when two residues have reacted and at 38° when six residues have reacted. From these results one can conclude that the Tψ loop is buried in the structure.[47] In a similar study it was shown that bromination of $tRNA^{Ala}$ is consistent with the cloverleaf model. In addition to the less reactive base-paired regions in its stem the TpψpCpG sequence is also protected.[37] The specific cyanoethylation of pseudouridine occurs only at low ionic strength.[40] The dihydrouridine residues which are supposed to be looped out remain reactive in tRNA even in the presence of 0.02 *M* Mg^{2+}.[36] A detailed study has been made using the oxidation of adenosine to adenosine 1-N-oxide with monoperphthalic acid[51] at pH 7 in aqueous medium. This reaction can easily be followed since the uv absorption spectrum is greatly altered in the course of the reaction[45]

(32) F. Cramer, F. v.d. Haar, and E. Schlimme, *FEBS Lett.*, **2**, 136 (1968).
(33) F. Cramer, F. Eckstein, F. v.d. Haar, and E. Schlimme, unpublished; E. Schlimme, Ph.D. Thesis, Technische Universität, Braunschweig, 1969.
(34) D. M. Brown and P. Schell, *J. Chem. Soc.*, 208 (1965).
(35) K. Burton, N. F. Varney, and P. C. Zamecnik, *Biochem. J.*, **99**, 29c (1966).
(36) P. Cerutti and N. Miller, *J. Mol. Biol.*, **26**, 55 (1967).
(37) J. A. Nelson, S. C. Ristow, and R. W. Holley, *Biochim. Biophys. Acta*, **149**, 590 (1967).
(38) J. P. Ebel, J. H. Weil, B. Rether, and J. Heinrich, *Bull. Soc. Chim. Biol.*, **47**, 1599 (1965).
(39) R. W. Chambers, *Biochemistry*, **4**, 219 (1965).
(40) M. Yoshida and T. Ukita, *Biochim. Biophys. Acta*, **123**, 214 (1966); **157**, 455 (1968).
(41) H. Hayatsu, K. Takeishi, and T. Ukita, *ibid.*, **123**, 445 (1966).
(42) G. Balle, P. Cerutti, and B. Witkop, *J. Amer. Chem. Soc.*, **88**, 3946 (1966).
(43) K.-H. Scheit, *Tetrahedron Lett.*, 1031 (1965).
(44) R. Marciello and G. Zubay, *Biochem. Biophys. Res. Commun.*, **14**, 272 (1964).
(45) F. Cramer and H. Seidel, *Biochim. Biophys. Acta*, **91**, 14 (1964).
(46) P. T. Gilham, *J. Amer. Chem. Soc.*, **84**, 687 (1962).
(47) S. W. Brostoff and V. Ingram, *Science*, **158**, 666 (1967).
(48) H.-J. Rhaese, *Biochim. Biophys. Acta*, **166**, 311 (1968).
(49) R. Thiebe and H. G. Zachau, *Europ. J. Biochem.*, **5**, 546 (1968); *Biochem. Biophys. Res. Commun.*, **33**, 260 (1968).
(50) M. L. Gefter and R. L. Russell, *J. Mol. Biol.*, **39**, 145 (1969).
(51) F. Cramer, H. Doepner, F. v.d. Haar, E. Schlimme, and H. Seidel, *Proc. Natl. Acad. Sci. U. S.*, **61**, 1384 (1968).

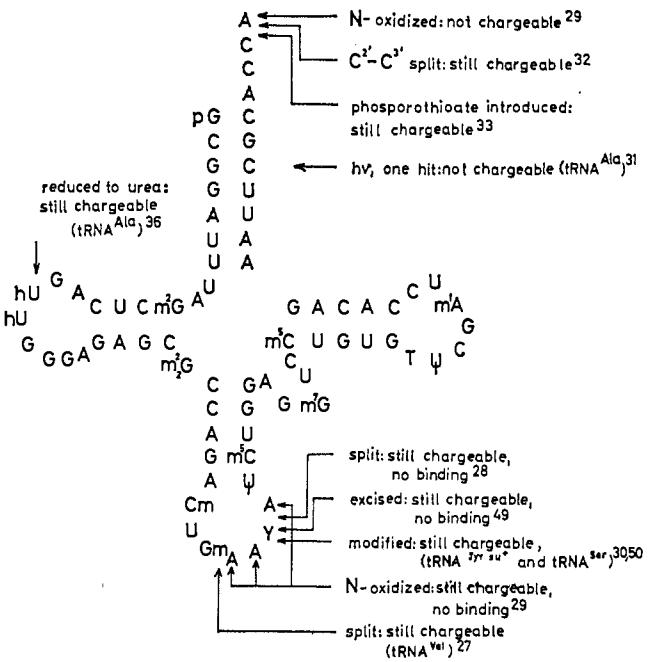


Figure 6. Chemical modification and biological function of $tRNA^{Phe}$ (and others).

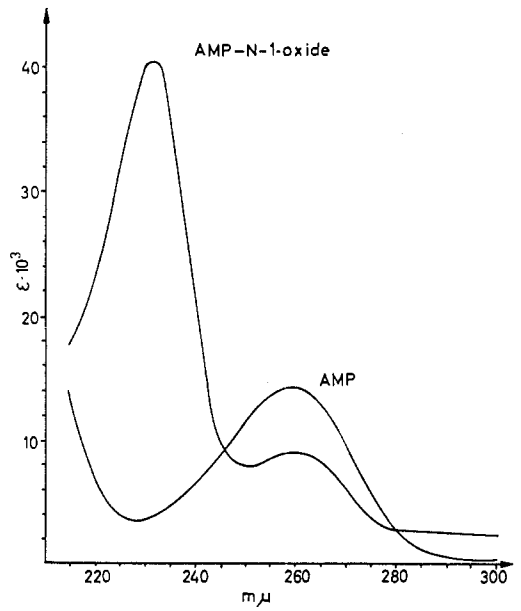


Figure 7. Ultraviolet spectra of AMP and AMP 1-N-oxide.[45,51]

(Figure 7). In poly A over 90% of the adenosine residues are oxidized without chain breaks.[52] On addition of poly U the poly A is protected against oxidation; only 2% of the adenosines remain reactive in the double strand.[52] With poly d(A-T) no oxidation of adenosine can be detected.[53] The results obtained with tRNA are given in Table II.

It can be seen that there is considerably more structure in the molecule at room temperature than one would expect from the cloverleaf. At 40° in $tRNA^{bulk}$ and also in $tRNA^{Ser}$ some adenosine residues become

(52) F. Cramer, F. v.d. Haar, and V. A. Erdmann, unpublished.
(53) S. Hennig, Ph.D. Thesis, Technische Universität, Braunschweig, 1967.

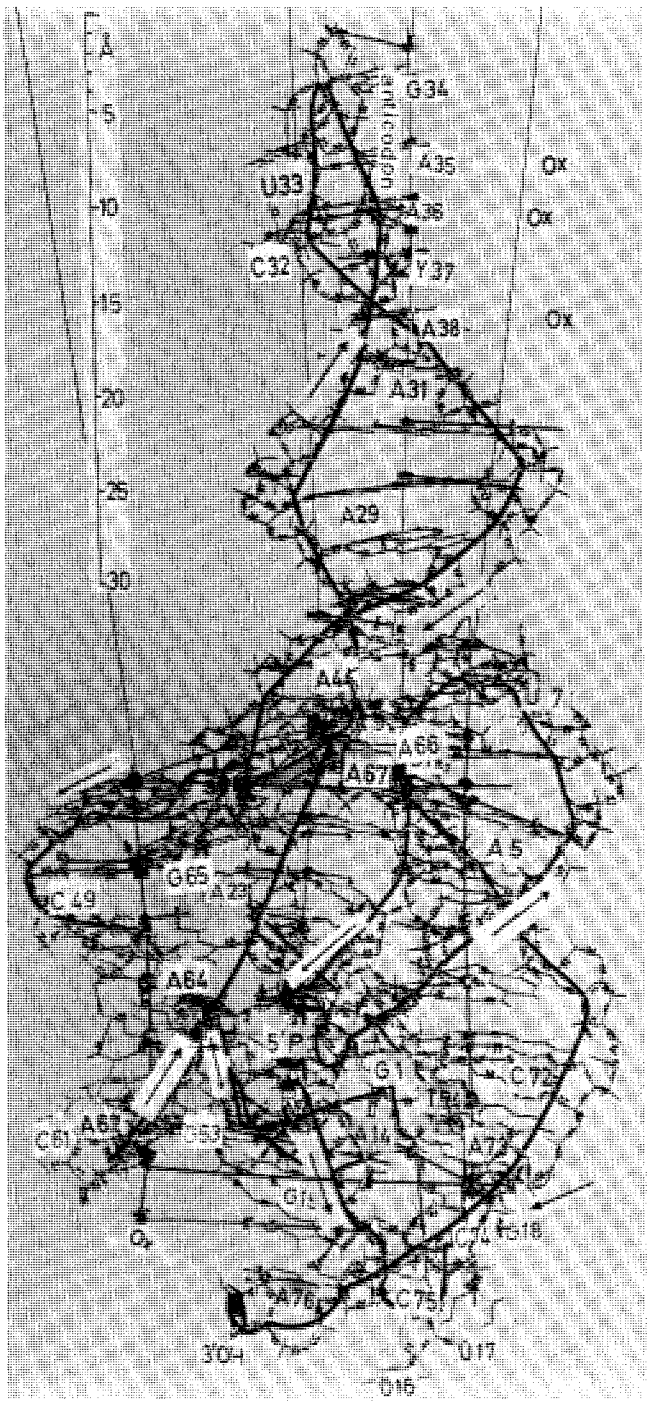


Figure 8. Photograph of a model with the proposed structure of $tRNA^{Phe}$.[51] The cord marks the backbone; the symbol "ox" indicates the positions where adenosine 1-N-oxide is formed; the arrows designate the 5'-3' direction of the backbone (Cambridge Repetition Engineers Ltd., wire models).

Table II

Formation of Adenosine 1-N-Oxide in $tRNA^{51,52}$

tRNA	No. of adenosine 1-N-oxide residues formed at 20°	40°	No. of adenosine residues Exposed in cloverleaf	Total
Bulk	4.1	6.0		16
Phe	4.0	4.0	9.0	17
Ser	6.0	8.0	8.0	16

available, though not in $tRNA^{Phe}$. This is in agreement with the melting behavior.

A more detailed analysis has been done with $tRNA^{Phe}$. From this analysis it was shown that the oxidizable adenosine residues are 35, 36 (anticodon), 38 (anticodon loop), and 76 (3′ terminus). From these data a general three-dimensional model for tRNA has been proposed by folding together stem, dihydrouridine loop, and Tψ loop of the cloverleaf structure. This is depicted for $tRNA^{Phe}$ in Figure 8 and is compatible with all chemical, physicochemical, and biochemical evidence so far available.[51,54]

The author is grateful to Dr. D. Gauss for helping to complete the manuscript. He has also enjoyed close colleagual cooperation with many scientists in the department; their names are gratefully quoted in the references. Most of the investigation described in this paper was supported by grants of the Deutsche Forschungsgemeinschaft.

(54) E. Schlimme, F. v.d. Haar and F. Cramer, *Z. Naturforsch.*, **24b**, 631 (1969).

Water Structure in Organic Hydrates

GEORGE A. JEFFREY

Crystallography Laboratory, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

Received April 1, 1969

Determinations of crystal structures of organic compounds have been concerned mainly with the stereochemistry of the isolated molecule. The results provide numerical information relating to the chemistry of the compound, in the form of bond lengths, valence angles, and conformation angles. The equally detailed information provided by these structures concerning interatomic vectors between molecules has not been so extensively correlated. The organic chemist is not much concerned with it. Most of his experiments are conducted in the liquid or gaseous phases to which the arrangement of the molecules in the solid has little relevance. Consequently, these intermolecular data have attracted the attention of only the comparatively few solid-state chemists or crystallographers who are curious as to how the shape and electron distribution in molecules affects the way they can cohere to form a regular lattice.[1]

In recent years, progress in the use of computers has made it possible to use these intermolecular data more quantitatively. In particular it has been possible to derive numerical parameters for potential energy functions for hydrocarbons by least-squares methods and, conversely, correct crystal structures have been sucessfully predicted by minimizing the lattice repulsive energies calculated from these functions.[2,3]

Since the ubiquitous water molecule occurs in more crystals than any other single molecule, it is tempting to enquire whether the data available from the crystal structure determination of hydrates can be used to obtain a better comprehension of the cohesive forces between water molecules and other molecular or ionic species.

Up to 1963, about 350 inorganic hydrate and 90 organic hydrate crystal structures had been determined,[4,5] and the numbers now probably exceed 1000 and 300, respectively. Attempts to use these data have revealed some general principles in the case of inorganic hydrates,[6,7] but the organic hydrates have scarcely been considered. An important deterrent is the elusive understanding of the electronic nature of the hydrogen bond, which is the principal cohesive force exercised by water molecules in organic hydrates. The absence of a generally accepted expression for the potential energy function of the hydrogen bond is a severe disadvantage, and the pronounced directional character of these forces is an added complication in any attempt to rationalize the intermolecular distances observed in crystalline hydrates. Another difficulty is uncertainty in the location of the hydrogen atoms, which must be known for complete description of a hydrogen bond. Neutron diffraction studies which could provide this information are, for economic reasons, much less numerous than X-ray crystal structure determinations. Hydrogen bonds cannot be recognized with certainty from interoxygen distances alone, in the absence of an observed electron (or proton) density peak for the hydrogen, because oxygen-to-oxygen nonbonding equilibrium distances vary over approximately the same range as do hydrogen bond distances (*i.e.*, 2.5 to 3.0 Å).[8] A hydrogen bond can be described as a directional cohesive force resulting from a redistribution of the electron density which permits the insertion of a proton between two nonbonded oxygens without increasing their

(1) A. I. Kitaigorodskii, "Organic Chemical Crystallography," Consultants Bureau, New York, N. Y., 1961.

(2) D. P. Craig, R. Mason, P. Pauling, and D. P. Santry, *Proc. Roy. Soc.* (London), **A286**, 98 (1965).

(3) D. E. Williams, *J. Chem. Phys.*, **45**, 3770 (1966); *ibid.*, **47**, 4680 (1967); *Science*, **159**, 645 (1968).

(4) R. W. G. Wyckoff, "Crystal Structures," Vol. 3, 2nd ed, Interscience Publishers, New York, N. Y., 1965.

(5) J. R. Clark, *Rev. Pure Appl. Chem.*, **13**, 50 (1963).

(6) R. Chidambaram, *J. Chem. Phys.*, **36**, 2361 (1962).

(7) W. H. Baur, *Acta Cryst.*, **19**, 909 (1965).

(8) Interesting examples where it is difficult to distinguish nonbonded and hydrogen-bonded O···O interactions in the absence of neutron diffraction work are as follows: 5-nitrobarbituric acid trihydrate (B. M. Craven, S. Martinez-Carrera, and G. A. Jeffrey, *ibid.*, **17**, 891 (1964)); 5-hydroxyiminobarbituric acid monohydrate (B. M. Craven and W. J. Takei, *ibid.*, **17**, 415 (1964)); nitranilic acid hexahydrate (E. K. Anderson, *ibid.*, **22**, 204 (1967)).