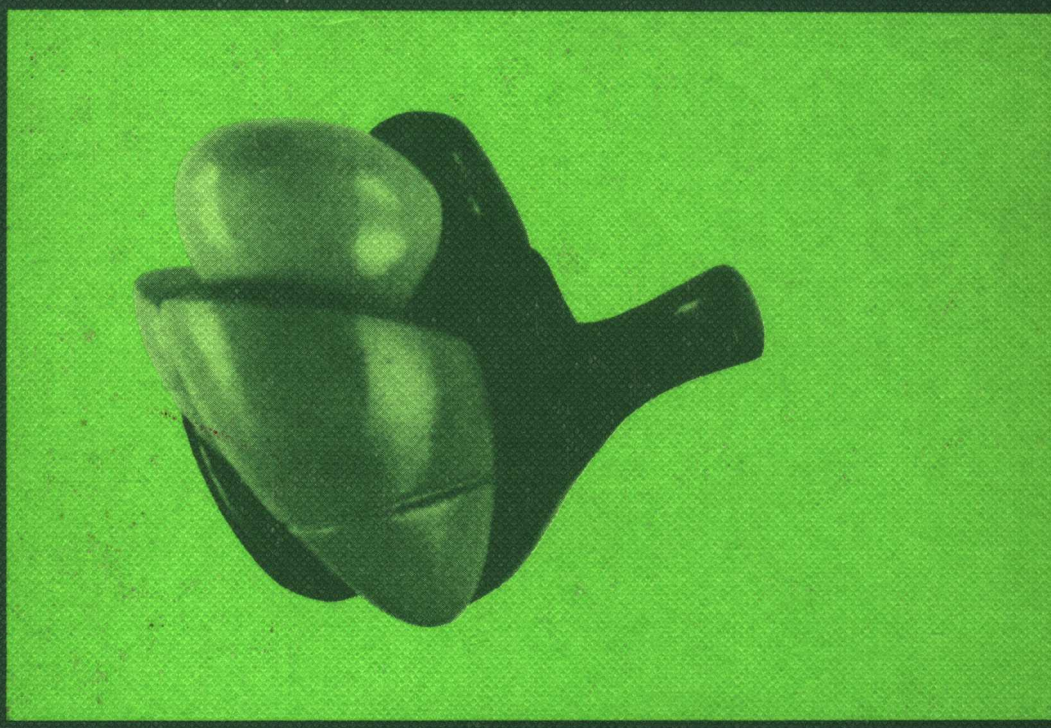
Mechanisms of Protein Synthesis

Structure-Function Relations, Control Mechanisms, and Evolutionary Aspects

Edited by E. Bermek



Springer-Verlag
Berlin Heidelberg New York Tokyo

Mechanisms of Protein Synthesis

Structure-Function Relations, Control Mechanisms, and Evolutionary Aspects

Proceedings of the Symposium
on Molecular Mechanisms in Protein Synthesis
Held at Beyaz Köşk, Emirgan,
Bosphorus, Istanbul

Edited by E. Bermek

With 124 Figures

Springer-Verlag
Berlin Heidelberg New York Tokyo 1985

Prof. Dr. ENGIN BERMEK
Biyofizik Bilim Dali
Istanbul Tip Fakültesi
Istanbul Üniversitesi
Çapa, Istanbul, Türkiye

Cover illustration: The front cover illustrates a photograph of the 70S ribosome model, see p. 95

ISBN 3-540-13653-3 Springer-Verlag Berlin Heidelberg New York Tokyo
ISBN 0-387-13653-3 Springer-Verlag New York Heidelberg Berlin Tokyo

Library of Congress Cataloging in Publication Data. Main entry under title: Mechanisms of protein synthesis. (Proceedings in life sciences). Includes index. 1. Protein biosynthesis–Congresses. I. Bermek, E. (Engin), 1939-. II. Series. QP551.M368 1984 574.19'297 84-13979

Typesetting and printing: Beltz, Offsetdruck, Hemsbach/Bergstr.
Bookbinding: Brühlsche Universitätsdruckerei, Giessen
2131-3130-543210

Preface

This volume contains the papers presented at the international symposium on "Molecular Mechanisms in Protein Synthesis" held on September 26-27, 1983 at the Beyaz Köşk in Emirgan, Bosphorus, Istanbul.

The symposium aimed to create a medium for information exchange and discussions regarding the current developments in the area of protein synthesis. To ensure an informal yet scientifically stimulating and productive atmosphere providing opportunity for relaxed and speculative discussions, the number of presentations was limited to twenty and that of attendants to about sixty. The emphasis in the symposium was laid on structure-function relations in the prokaryotic protein synthesizing systems and on the control mechanisms of eukaryotic protein synthesis, in particular, during chain initiation. Other issues like evolutionary aspects of protein synthesis, translational components genes and proofreading were covered as well.

The manuscripts represent the extended accounts of the oral presentations, and it has been aimed with the concluding remarks at the end of the volume to give a summarizing view of the presentations and the discussions.

The symposium was sponsored by the Istanbul Faculty of Medicine and the Medical and Basic Sciences Research Groups of the Scientific and Technical Research Council of Turkey (TÜBITAK). The support of the Turkish Airlines, Koç Holding A.S., Oesterreichisches Generalkonsulat-Kulturinstitut, Rotary Club, SESA Elektronik San. Tic. A.S., Hoffmann-La Roche Company, Incekara A.Ş., Sandoz-Turkey, Turkish Glassworks Inc., Ünitay Export-Import, Kerman A.S., and Eczacibaşi Holding A.S., in Istanbul is gratefully acknowledged. I would also like to thank all those who have made with their help and contributions the organization of the symposium possible and particularly M. Metinsoy for her ever patient management of the secretarial tasks.

I wish that by propagating the reported findings and the views arisen during the symposium, this volume publication of which has been possible with the willing and competent collaboration of Springer-Verlag, Heidelberg, may provide new stimulus into different directions of research in protein synthesis.

Çapa, October 1984 E. Bermek

Contributors

You will find the adress at the beginning of the respective contribution

Barta, A. 40
Bermek, E. 170, 180
Bilgin-Aktar, N. 180
Cantor, C.R. 2
De Haro, C. 160
De Herreros, A.G. 160
Deng, H.Y. 23
DiSegni, G. 120
Fasano, D. 50
Faulhammer, H. 61
Gökhan, N. 170
Guesnet, J. 50
Hardesty, B. 23
Hopfield, J.J. 106
Hui, C.F. 2
Kaempfer, R. 120
Kan, B. 170
Kang, C. 2
Kanigür, G. 170
Küchler, E. 40
Lai, C.Y. 106
Levin D.H. 144
Lill, R. 76
London, I.M. 144
Matts, R.L. 144
Murphy, R.F. 2
Nurten, R. 180
Ochoa, S. 160
Odabaş, Ü.G. 180
Odom, O.W. 23
Parlato, G. 50
Parmeggiani, A. 50
Paulsen, H. 76
Petryshyn, R. 145
Picone, D. 50
Pizzano, R. 50
Robbins, D. 23
Robertson, J.M. 76
Rosen, H. 120
Rychlik, W. 23
Sayhan, O.Z. 180
Serdyuk, I.N. 92
Spirin, A.S. 92
Sprinzl, M. 61
Steiner, G. 40
Şimşek, M. 194
Thomas, N.S. 144
Tiryaki, D. 170
Vazquez, D. 201
Voorma, H.O. 135
Wintermeyer, W. 76
Wollenzien, P.L. 2
Yamane, T. 106

Contents

Proof-Reading in Translation

Regulation of Eukaryotic Chain Initiation

Eukaryotic Chain Elongation

Evolution and Selective Action of Translational Inhibitors

Ribosome Structure

RNA Structure, Free and on the Ribosome, as Revealed by Chemical and Enzymatic Studies

P.L. WOLLENZIEN, C.F. HUI, C. KANG, R.F. MURPHY, and C.R. CANTOR[1]

1. Introduction

At present, various chemical and enzymatic probes are the experimental techniques capable of providing large amounts of data on the structure and arrangement of large RNA molecules. Direct reactivity measurements yield information on accessibility, while cross-linking provides information on proximity. Taken together with secondary structure estimates derived from analysis of base pairing potential and phylogenetic variations, the results of chemical and enzymatic studies can help evaluate specific models for RNA structures and interactions. It is unlikely that a precise three-dimensional structure could result from this approach. However, it should be possible, through accumulation of a large enough data set, to provide reasonably accurate secondary structures and approximate models of how the secondary structure is arranged in space, alone, or in concert with proteins. A unique advantage of this approach is that it is suited to the detection of conformational changes in RNA molecules. Thus, it will be very useful in establishing functional aspects of altered RNA structures where they exist.

Escherichia coli ribosomal RNAs are the ideal system in which to develop and evaluate particular probes for large RNA structure. They can readily be isolated either as free rRNAs or as ribosomal subunits; the primary structures are completely known (Brosius et al. 1978; Carbon et al. 1979; Brosius et al. 1980). In addition, there are considerable data suggesting that, near physiological conditions, the free RNAs can fold into compact, three-dimensional structures similar to their structures on the ribosome (Vasiliev et al. 1978; Spirin et al. 1979; Allen and Wong 1978; Vasiliev and Zalite, 1980). Detailed secondary structure models have been presented (Noller and Woese 1981; Stiegler et al. 1981a; Zwieb et al. 1981; Noller et al. 1981; Glotz et al. 1981; Branlant et al. 1981) and while these are still evolving it is clear that at the present time we know more about the secondary structures of the 16S rRNA and 23S rRNA than any other RNA of comparable sizes.

The rRNAs have been identified as the structural core for both ribosomal subunits (Stuhrmann et al. 1977; Stuhrmann et al. 1978; Spirin et al. 1979). Since there is considerable experimental evidence for structural changes within the ribosome during protein translation (see Zamir et al. 1974) and for altered topographies of the rRNA under different conditions (Chapman and Noller 1977; Hogan and Noller 1978; Herr and Noller 1979; Brow and Noller 1983), it is reasonable to suspect that changes in the state of the

[1] Department of Human Genetics and Development, College of Physicians and Surgeons, Columbia University, New York, New York 10032, USA

Mechanisms of Protein Synthesis
ed. by Bermek

rRNA are intimately connected to the process of translation. Noller and Woese (1981) have argued, based on evolutionary considerations, that the primordial ribosome may have been RNA alone. In this view the ribosomal proteins are relegated to the secondary role of fine tuning the speed and fidelity of protein translation. In addition, there are several sites in the rRNAs directly implicated in ribosome function. (see Thompson and Hearst 1983b, Küchler et al. 1984). It thus remains of considerable interest to characterize the structure of ribosomal RNAs as fully as possible. It is also of great interest to understand more about the structure of tRNAs and mRNAs when these are bound on the ribosome in various functional states.

The least that direct reactivity measurements could reveal about the structure of an RNA, whether free or in a ribonucleoprotein complex, is the identity of residues on the surface. For a molecule the size of 16S rRNA this could be an appreciable fraction of the whole. The most simpleminded calculations reveal that roughly half of the 1542 residues of the 16S rRNA should be on the surface of the free RNA it if is quite compact; more if it is at all extended.

Most direct chemical and enzymatic studies can also provide information about the secondary structure of the residues involved; however, this pattern is superimposed on the pattern of accessible sites. Examples are enzymes specific for single or double strands and chemical reactions with atoms involved in base pairing in double helices. Unfortunately, we have no way at present of evaluating the effect of RNA tertiary structures on such reaction specificity. Thus, all such inferences about secondary structures and surface accessibility have to be considered with some caution.

Since failure to react with an enzyme or chemical probe can always be the result of an idiosyncratic local environment or conformation, only positive results, i.e., reactivity, are generally interpretable.

To get the maximum value from accessibility studies one needs many probes, preferably with different sizes to try to infer the location of any narrow crevices or channels. It is critical that large amounts of data must be generated easily and even more clear that this data must be as error free as possible. Even a few erroneous positives, i.e., reactive groups that are actually located internally, can badly distort any models built from the data. This is because there is no simple way to cross-check the data. The amount of information needed to reconstruct even crude three-dimensional models is enormous and it is always tempting to try to use all the data available to refine the models. This is fine if all the data is accurate, but it can be disasterous otherwise.

Cross-linking is a direct way of determining the proximity between two residues so long as one can guarantee that the formation of the cross-link has not artifactually brought together two residues normally far apart. Even in this case one knows that the residues can be near each other, but not necessarily in the native state. The advantage of photoreagents for cross-linking is the short time scale of their reactivity. This virtually insures that only residues in normal contact will be cross-linked. However, the current availability of RNA photo cross-linkers is pretty much limited to psoralens which are specific for base paired residues. To probe other types of proximity one must resort to chemical reagents. One of the most useful nucleic acid cross-linkers is the thiol-containing glyoxal derivative $GbzCyn_2Ac$ (Expert-Bezancon and Hayes 1980). This reacts with single stranded guanines analogously with glyoxal. Reduction generates free thiols which can then be cross-linked either by direct oxidation or by bridging with bis thiol

reagents, such as bis maleimides. In this way, one can cross-link nearby single stranded guanine residues, indicating features of the tertiary structure.

In principle, a sufficient battery of cross-linking results could provide a complete three-dimensional structure accurate to about the size of the reagents themselves, typically 10 Å. This would be enormously valuable and might well approach what could be available from X-ray diffraction on such large and potentially floppy structures. There is no way to obtain the correct mirror handedness by cross-linking alone, but this difficulty would be resolved by correlation of the possible structures with protein-RNA cross-linking and positions of proteins determined by immune electron microscopy. However, to achieve the goal of a complete three-dimensional structure at high resolution would require hundreds of cross-links. It would be relatively easy to obtain cross-links for residue pairs accessible to externally added reagents. Indeed, most currently available information on the 16S rRNA has involved the use of rather bulky reagents and, thus, the results obtained are surely a view of the structure heavily biased towards the more accessible regions. For internal proximity measurements one has to resort to more difficult "zero length" cross-links, such as those provided by direct irradiation through a variety of direct and indirect photochemical mechanisms.

Again, as in the case of accessibility measurements, maximally effective cross-linking will involve the collection of large amounts of data, as error free as possible. The availability of extensive secondary structure data will greatly reduce the number of cross-links needed to define the three-dimensional structure and will also provide some useful redundancy checking.

The great potential of cross-linking is offset by the amount of effort that is potentially involved. The approach for RNA is still in its infancy. The 16S rRNA is the most intensively studied species and, thus far, less than 50 distinct cross-links within the 16S rRNA are indicated by currently used reagents (Zwieb and Brimacombe 1980; Wollenzien and Cantor 1982a; Thompson and Hearst 1983a; Expert-Bezancon et al. 1983; Wollenzien et al. 1984). Few of these cross-links have yet been mapped at the level of individual residues. To achieve the potential impact of cross-linking we will have to have many more useful reagents and rapid and accurate ways of analyzing their reaction patterns. In this paper we focus on summarizing the current state of the art in studying psoralen cross-links in 16S rRNA. Many of the obstacles to using psoralens rapidly and effectively are now being overcome. The solutions should offer considerable guidance to those who will use psoralens or other cross-linkers in studies of any large RNA.

2. Electron Microscopic Techniques for Studying Psoralen Cross-linking

Many of the products produced by cross-linkers will always be pairs of residues that lie reasonably close in the nucleotide sequence (Zwieb and Brimacombe 1980; Turner et al. 1982; Expert-Bezancon et al. 1983; Thompson and Hearst 1983a). While this sort of data is useful confirmation of secondary structure models in many cases, it is not often particularly novel or informative. Instead, what one hopes in a cross-linking study, is to see products involving pairs of residues distant in the primary structure. These can provide either tests of long-distance secondary structure contacts, vital in discriminating

between different models, or information about RNA tertiary structure contacts. In either case the long-distance contacts provide information that is invaluable for constructing three-dimensional structure models, and it is information that is almost impossible to obtain by other existing methods.

The major advantage of analysis of cross-link locations by electron microscopy is that one is forced to focus on cross-links between distant residues. The basic idea of the method is very simple and is shown schematially in Fig. 1. A population of RNAs, hopefully homogeneous, in solution or on the ribosome, is treated with a cross-linking agent. The excess agent is removed or inactivated, any proteins present are removed, and then the RNA is subjected to a condition which completely and irreversibly removes all secondary structure. We use treatment with formamide followed by heating in the presence of formaldehyde (Wollenzien and Cantor 1982a). The resulting molecules are then spread on a grid with a cytochrome c technique, and then stained with uranyl acetate and shadowed with platinum and palladium in order to visualize them in the electron microscope.

Molecules containing no cross-links will almost always appear as worm-like rods; loops or circles caused by accidental crossover of two parts of the same molecule are observed less than 3 % of the time for RNAs the size of 16S rRNA (Wollenzien et al. 1984). A typical field of uncross-linked 16S rRNA molecules is shown in Fig. 2 (top). In contrast, with cross-linked molecules, the presence of the extra covalent bond imposed by each cross-link insures that a loop will remain after the RNA structure is denatured. Thus, cross-links between residues distant along the chain will be immediately apparent in the electron microscope by the appearance of looped structures. Figure 2 (bottom) shows a typical field generated by psoralen cross-linking of free 16S rRNA. It is apparent that under the conditions used, roughly a quarter of the RNA molecules show loops. A wide variety of different loop and tail configurations is evident.

Large fields of molecules can be scanned visually to obtain statistics on the frequency of occurrence of cross-links which produce unique and easily discriminated loop patterns. Examples are full circles which indicate cross-links between residues near both ends of the RNA, and "lollipop structures", molecules with a small loop at one end,

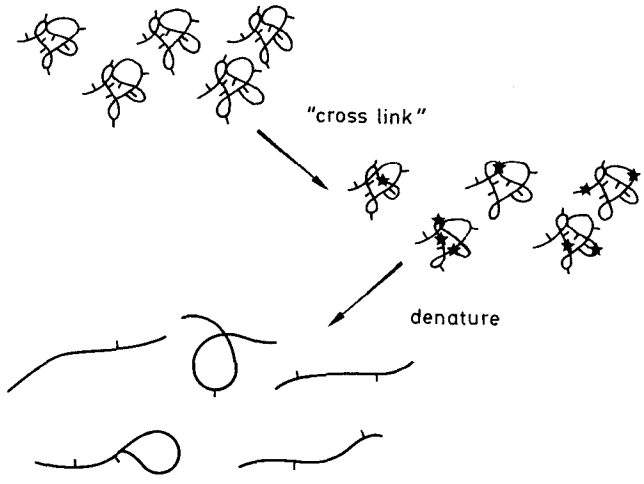


Fig. 1. The analysis of RNA cross-linking by electron microscopy

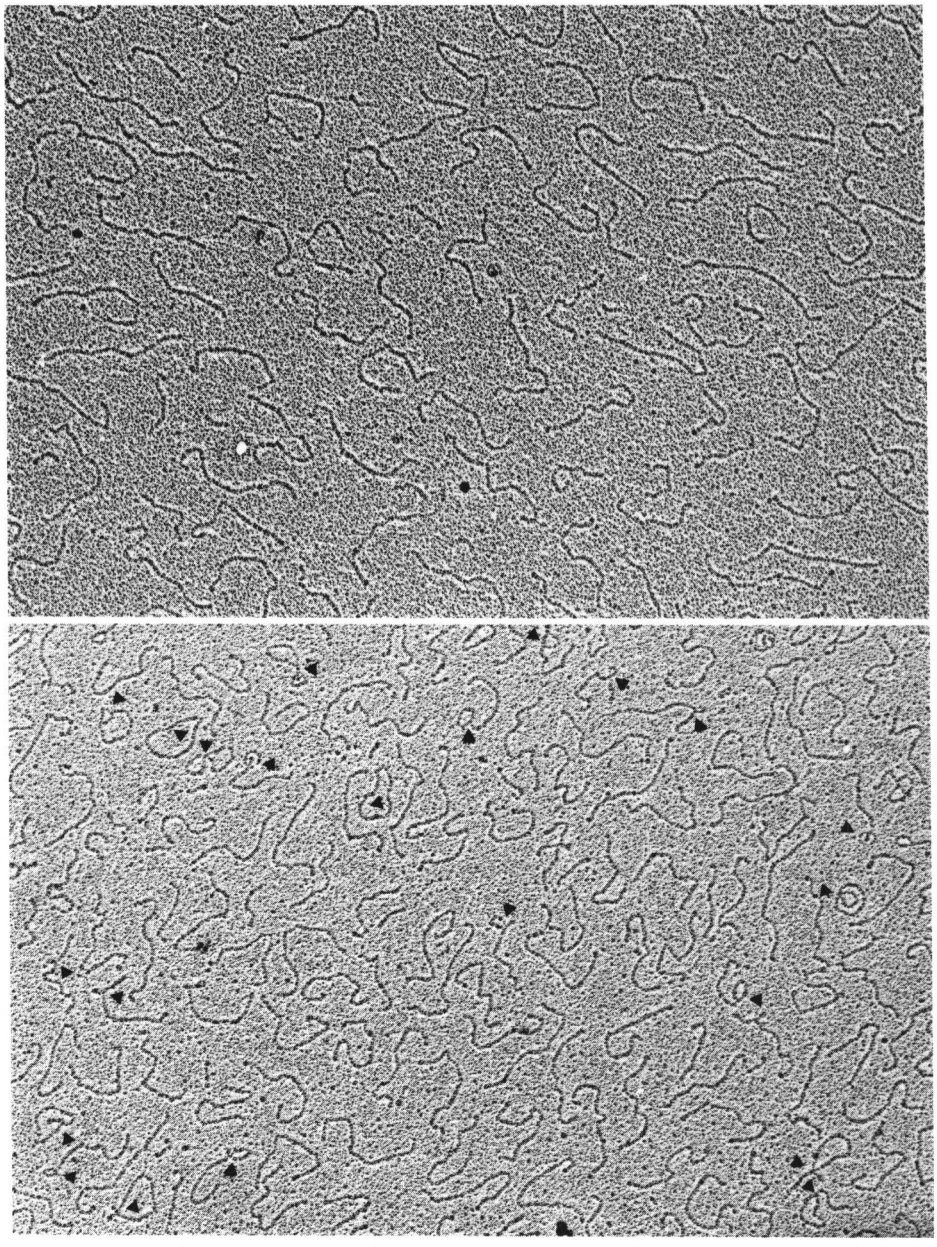


Fig. 2. Typical fields of 16S rRNA molecules as seen in the electron microscope by the spreading techniques in current use. At the *top* is a field of uncross-linked RNAs; at the *bottom* a field of RNAs cross-linked by irradiation in solution with the psoralen AMT. The *arrowheads* show molecules with easily detected cross-linked loops

indicating a cross-link between an end and an internal region somewhere between the middle and that end. For more precise identification of specific cross-links, a much more tedious procedure is required. Individual molecules containing any looped structure are photographed. The lengths of their loops and tails are measured by analyzing the image with a digital planimeter. Because different molecules are stretched to different extents during the preparation of the sample for microscopy, there is significant variation in the quantitative appearance even of molecules with identical cross-links. This is the major limitation in the resolution of specific cross-linked species and the determination of the location of the cross-links. The problem can be partially overcome by observing many examples of each type of cross-linked molecule. Each molecule makes up one member of a large data set that is then analyzed statistically, as will be described later.

A compromise must be made between the yield of molecules with interesting cross-linked features and the ability to analyze these molecules accurately. If too few molecules contain cross-links, the real data will be obscured by noise from the occasional accidental molecular crossover during spreading for microscopy. If too many cross-links are present in a molecule, the pattern of loops that results cannot be traced uniquely and, therefore, all information is lost.

With psoralens, and probably most other cross-linkers, there is another serious problem. Since the desired cross-links between residues distant in the primary structure are not necessarily favored reaction products, by the time these products have appeared at significant yield, the molecules containing them may have one or more additional cross-links between residues nearby in the sequence. If these not visible directly in the microscope, they would alter the apparent length of the RNA chain by a length equal to the distance between the cross-links. Since the distortion can occur, in principle, in any tail or loop, the result would lead to serious deterioration of the sharpness of the lengths characterizing each particular cross-linked loop. A check on this problem can be made by measuring the apparent shortening of contour length of cross-linked molecules compared to control uncross-linked molecules (Wollenzien and Cantor 1982a; Wollenzien et al. 1984).

To overcome the distortion caused by "hidden" cross-links, it is necessary to work at rather low levels of cross-linking. This means that very large numbers of molecules must be examined and one must be scrupulously careful to perform spreading under conditions that minimize accidental crossovers. Techniques are available for fractionating cross-linked species away from uncross-linked molecules and these can be used where necessary to enrich the population of desired products prior to electron microscopy (Wollenzien and Cantor 1982a). There are several consequences of performing cross-linking at low levels that must be kept in mind. As shown schematically in Fig. 1, the curious result of a cross-linking procedure is to take a sample that was once homogeneous and convert it to a heterogeneous mixture of products. Almost any method used for subsequent analysis has to contend with this heterogeneity. This is why microscopy, as a single molecule technique, has an overwhelming potential advantage over any other technique. Representation in the heterogeneous mixture will be determined mainly by the reactivity of the particularly cross-linked residues under specified conditions.

The occurrence of open, untangled structures is a constant feature for all the loops that have been seen by psoralen cross-linking of 16S rRNA and other RNAs. It is for-

tunate since it greatly facilitates the quantitative analysis of the electron microscopic images. It also immediately provides an important piece of structural information: there are no structures with knotted topologies formed by residues within any of the loops (Cantor et al. 1980). RNA helices form one complete turn every 11 base pairs and many possible structures involving multiple turns of helix formed by interrupted stretches of sequence would lead to topologically knotted chain configurations. Once the region was closed into a covalent circle, these knots would be topologically trapped and the resulting molecules would appear to have complex extra looped structures within the circles. The fact that these are never observed forms a powerful structural constraint that will be a critical test of potential three-dimensional folding patterns.

Any inspection of the cross-linked molecules in Fig. 2 will immediately reveal that there is no way to distinguish which end of the molecule is which. The location of each loop has an uncertain polarity. One must have additional information to determine which end of the molecule is the 3' end and which is the 5' end. Earlier psoralen cross-linking studies (Wollenzien et al. 1979a) had to rely on rather indirect ways of doing this (Wollenzien et al. 1979b) and the results led to (what we know now) the incorrect orientation of some of the cross-links made in the free 16S rRNA. What was needed was a way of directly marking one end of the RNA in the electron microscope with a technique that would survive the rather rigorous denaturation conditions required for proper spreading. This demanded a covalent label and the technique that was finally developed took advantage of the unique photochemical properties of psoralens (Fig. 3).

When aminomethyl trioxsalen (AMT) or hydroxymethyl trioxsalen (HMT) are irradiated at 390 nm, near their long wavelength absorption edge, instead of forming cross-links in double stranded DNA or RNA, they form a monoadduct with just one of the strands (Chatterjee and Cantor 1978). At any time later, a second shorter wavelength irradiation, typically at 360 nm, will convert the monoadduct into a cross-link providing that the sequence at that point contains the alternating purine-pyrimidine dinucleotide required for the second photoreaction. This allows the end labeling procedure described schematically in Fig. 4 (Wollenzien and Cantor 1982b).

Appropriate recombinant DNA techniques are used to obtain a plasmid containing part of the 16S rRNA gene in a known position. Digestion of this plasmid DNA with restriction endonucleases and subsequent size fractionation by polyacrylamide gel electrophoresis is used to prepare a restriction fragment containing several hundred base pairs of sequence coding for regions near one end of the 16S rRNA. This fragment is treated

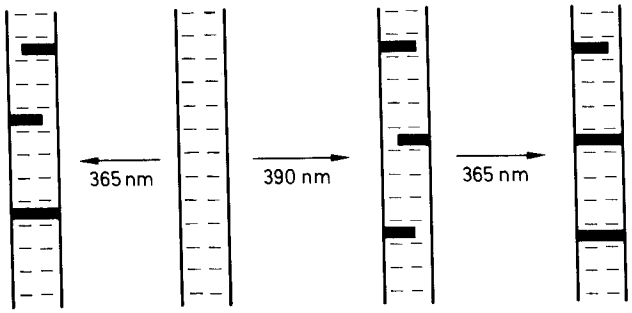


Fig. 3. Psoralen photochemistry. Shown schematically are the production of monoadducts by long wavelength irradiation and their conversion to cross-links by shorter wavelengths. Direct short wavelength irradiation produces a mixture of cross-links and monoadducts.

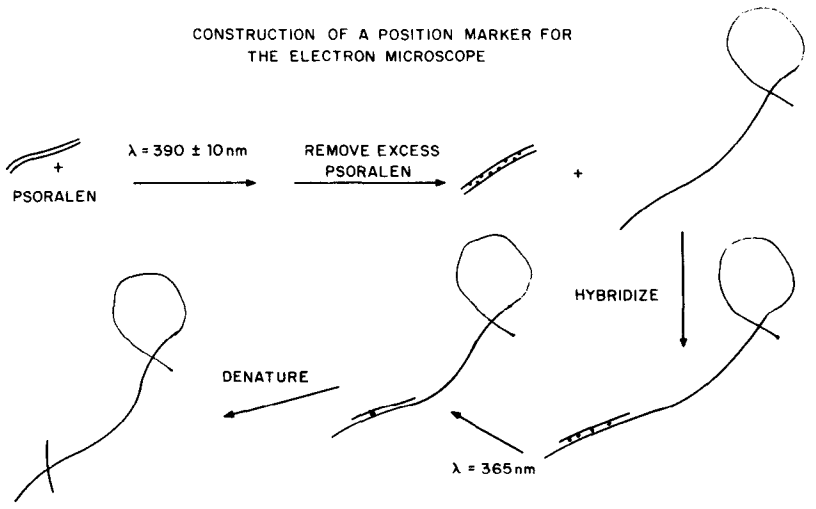
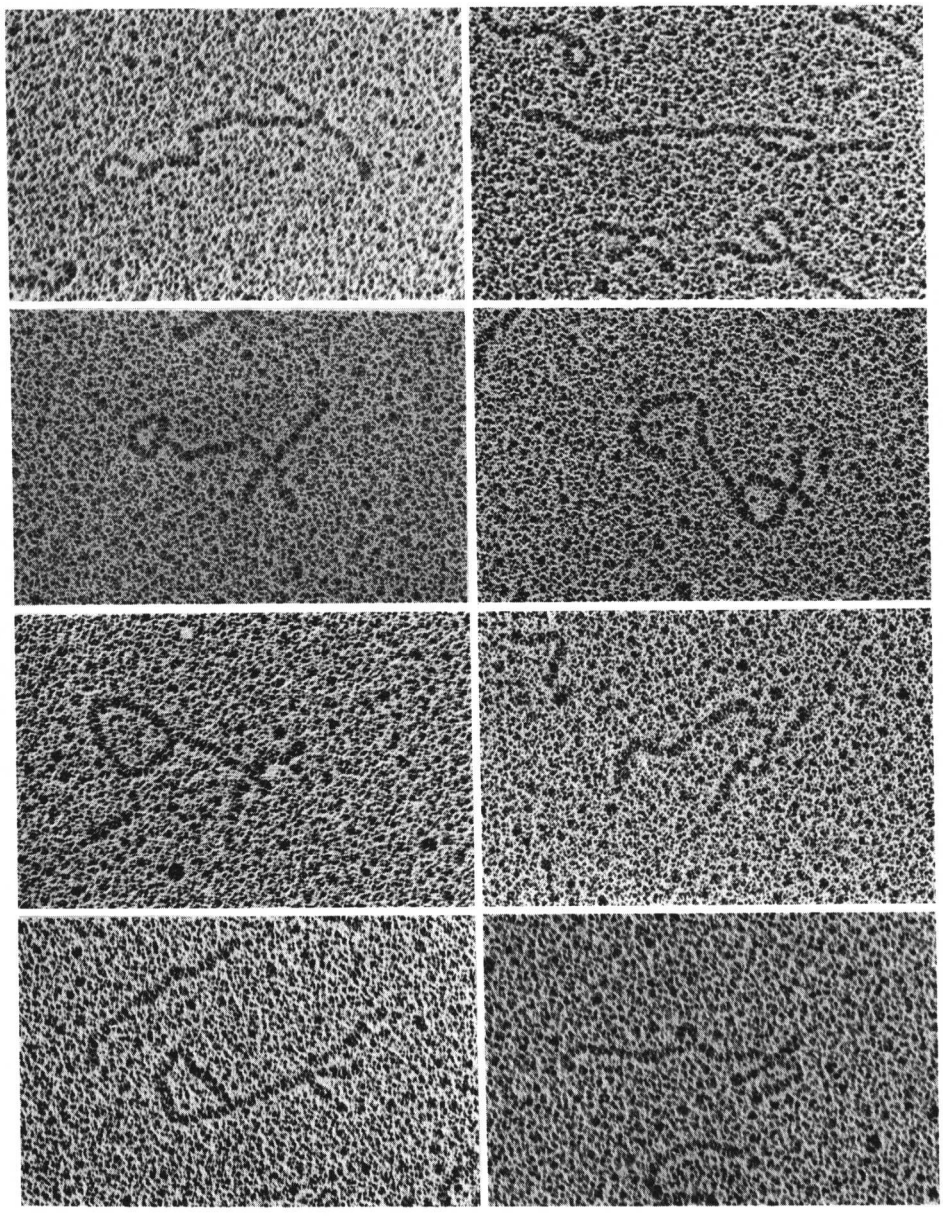


Fig. 4. The technique used to label one end of a cross-linked 16S rRNA in the electron microscope. See the text for further details

with psoralen to produce an average of several monoadducts, but no cross-links. The fragment is denatured by heating and then hybridized to 16S rRNA containing cross-links previously placed there by any agent desired. When the RNA-DNA hybrid is reirradiated at 360 nm, the psoralens form an average of one cross-link, covalently attaching the DNA to the RNA. Now when the sample is denatured and examined in the microscope, the extra DNA fragment appears as a short line that clearly marks the location of the end to which it is attached. The identity of that end is known from the nucleotide sequence of the fragment. Typical molecules labeled in this way are shown in Fig. 5. The technique is simple and unambigous.

In previous studies, the absolute orientation of cross-linked RNA molecules in the electron microscope image was not known. The best that one could do to classify molecules was to measure the position of cross-links from the shortest tail. Using this data each molecule can be represented as a point in a triangular two-dimensional region as shown in Fig. 6a (Wollenzien and Cantor 1982a). Once data on many molecules is accumulated, distinct cross-linked species will manifest themselves as a cluster of many molecules within a small area of the triangular region.

In our recent studies the absolute orientation of the cross-linked molecules is known from the labeling method described above. Thus, each molecule containing a single cross-linked loop can be described by two parameters, an estimate of the distance from the beginning, X, and end, Y, of the loop to the 5, end. From this definition X is always less than Y. Each molecule is then represented as a point in a triangular two-dimensional region as shown in Fig. 6b. Note that this region is twice as large as the corresponding region when unoriented molecules are used. This increase in area results in a markedly enhanced ability to discriminate between many different cross-linked species. Note that the data on unoriented molecules can be generated from data on oriented molecules by folding the triangular region of the latter in half. Thus, once individual cross-linked features are identified with an oriented sample it will be possible



Left

Fig. 5. Examples of end labeled molecules. At *left* are images of the molecules; at *right* are tracings that interpret each of these images: RNA is shown as a *solid line* and DNA as a *dashed line*. The bar indicates 0.5 microns