

# DETERMINATION OF RNA STRUCTURE AND THERMODYNAMICS

*John A. Jaeger, John SantaLucia, Jr.,  
and Ignacio Tinoco, Jr.*

Department of Chemistry and Laboratory of Chemical Biodynamics, University of California, Berkeley, California 94720

**KEY WORDS:** nuclear magnetic resonance, phylogeny, chemical modification, RNA folding, RNA structure prediction

---

## CONTENTS

INTRODUCTION . . . . .	256
PHYLOGENETIC AND MUTATIONAL ANALYSES . . . . .	258
<i>Natural Sequence Variability</i> . . . . .	258
<i>Artificial Phylogeny</i> . . . . .	259
<i>Random Mutations and Specific Selection</i> . . . . .	260
<i>Functional Group Mutations</i> . . . . .	260
OPTICAL SPECTROSCOPY . . . . .	260
<i>Absorbance</i> . . . . .	260
<i>Circular Dichroism</i> . . . . .	263
<i>Fluorescence</i> . . . . .	263
CHEMICAL AND ENZYMATIC REACTIVITY . . . . .	264
<i>Reagents</i> . . . . .	264
<i>Secondary and Tertiary Structure</i> . . . . .	266
<i>Footprinting</i> . . . . .	267
GEL ELECTROPHORESIS . . . . .	268
NUCLEAR MAGNETIC RESONANCE (NMR) . . . . .	268
<i>Proton NMR</i> . . . . .	269
<i>Heteronuclear NMR</i> . . . . .	271
X-RAY DIFFRACTION . . . . .	275
STRUCTURE PREDICTION AND MODELLING . . . . .	277
<i>Secondary Structure Prediction</i> . . . . .	278
<i>Tertiary Structure Prediction and Modelling</i> . . . . .	279
CONCLUSION . . . . .	281
	255

## INTRODUCTION

The central biological importance of RNA has recently become more apparent. In addition to the wide diversity of functions of RNA (1, 2), the list of known catalytic activities has been growing rapidly (3–7). The functional diversity of RNA reflects diversity in its three-dimensional structure. Knowledge of the three-dimensional structures and general rules for RNA folding will be invaluable for deducing more detailed mechanisms of all RNA functions.

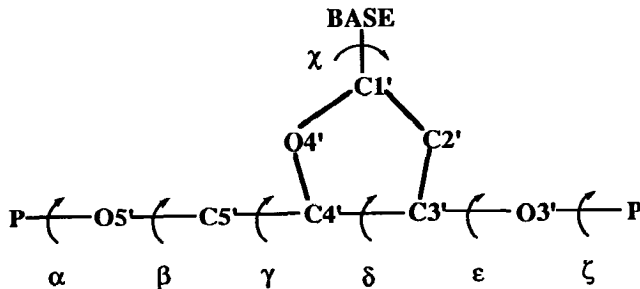
The sequence (primary structure) of an RNA molecule is relatively easy to determine. Methods for determining RNA secondary structure (base pairing) and tertiary structure, however, have not kept pace with the rapid discovery of RNA molecules with interesting functions. Therefore, improved methods for determining and predicting RNA structure are needed. We first review the methods available for characterizing RNA structure and thermodynamics. Methods for structure prediction are also discussed, since they provide valuable information for the design and interpretation of experiments. Accurate prediction of RNA structure requires an understanding of fundamental interactions such as hydrogen bonding, stacking, and hydration in diverse structural contexts.

RNA molecules are polynucleotides containing ribose sugars connected by 3'—5' phosphodiester linkages. The bases are connected to the ribose sugars in the beta position of the anomeric carbon (C1'). Figure 1 shows a diagram of a nucleotide to illustrate the torsion angles (seven per nucleotide) used to specify completely the conformation of an RNA (8).

The sugar pucker (9) is specified by angle  $\delta$ . Constraining the atoms of a five-membered ring to a plane is energetically unfavorable. Therefore, the sugars in nucleic acids usually have either C2' or C3' out of the plane of the other four ring atoms. When the out-of-plane atom is on the same side as the base, the conformation is C2'-endo or C3'-endo, respectively. For isolated nucleotides, the C3'-endo and C2'-endo sugar puckers are nearly equal in energy (10). Thus, the sugars adopt whichever conformation will allow other molecular interactions to be optimized. The phosphate-phosphate distance for C3'-endo pucker (5.9 Å) is shorter than in C2'-endo (7.0 Å) (9).

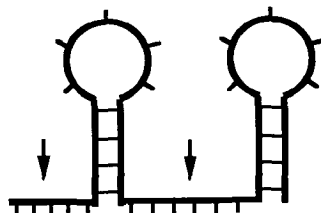
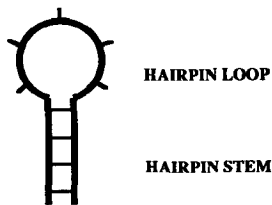
The glycosidic torsion angle,  $\chi$ , specifies the rotation of the base relative to the sugar. In the *anti* conformation the base is positioned away from the sugar ring; in the *syn* conformation the base is rotated by approximately 180 and is positioned above the sugar ring. Typically, bases adopt the *anti* conformation, since it is energetically more favorable than the *syn* conformation.

The other torsion angles (Figure 1) specify the folding of the sugar-phosphate backbone. An RNA chain can fold back upon itself to form hydrogen bonds between bases. Most commonly Watson-Crick base-pairs between A•U



**HAIRPIN**

**SINGLE STRANDED REGIONS**



**BULGE LOOP**

**INTERNAL LOOP**

**MULTI-BRANCHED LOOP or JUNCTION**

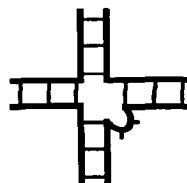


Figure 1 (top) The definition of the torsion angles that define RNA conformation (8). (bottom) Diagrams of RNA secondary structure elements (13).

and G•C are formed, which involve two and three hydrogen bonds, respectively. The consecutive formation of base pairs results in the formation of a double helix; these paired regions define the secondary structure. RNA helices are typically right-handed A-form (9) with 11–12 base pairs per turn and C3'-endo sugars. Left-handed Z-form RNA (11) has been found at high ionic strength (12). DNA helices are typically right-handed B-form with 10 base pairs per turn and C2'-endo sugars. Secondary structural elements formed by RNA include bulges, hairpins, internal loops, and junctions or multi-branched loops; they are shown in Figure 1 (13). Interactions between helices or loop regions are termed tertiary interactions; they include base-triples (14) and pseudoknots (15).

## PHYLOGENETIC AND MUTATIONAL ANALYSES

### *Natural Sequence Variability*

Phylogenetic analysis is the most accurate method for secondary structure prediction. Natural phylogenetic analysis assumes that RNAs of similar function share common structural elements. These elements are proved by compensatory mutations of bases in a double helix, for example a G•C to an A•U pair. Phylogeny cannot predict noncanonical base pairs well, since the rules for these pairs are still vague. This method cannot prove pairing in conserved regions, since there are no compensatory mutations. A recent review describes this method in depth (16). Secondary structures of tRNAs (17), ribosomal RNAs (18–20), group I and group II introns (21, 22), RNAs from snRNPs (23), hammerhead catalytic RNA (7), and RNase P RNAs (24) have been determined by phylogenetic analyses. Further studies have confirmed these structures.

A phylogenetic analysis compares sequences, and looks for similar secondary structures in the same order. Alignments are a convenient tool for studying sequence homology (Figure 2). RNA alignments usually contain gaps to compensate for base deletions and insertions. Computer-generated alignments (see Ref. 25 for a review) are useful, but often fail to align all conserved bases or pairing regions. Interactive alignment programs (26) and multiple sequence editors (27, 28) are useful for manual alignment, provide for insertion and deletion of gaps, as well as for movement of sequence fragments, and provide a consensus sequence of common bases. Secondary structures determined from the alignment must show at least one compensatory base pair mutation per helix.

Tertiary interactions, such as pseudoknots, triple base interactions, and binding sites for tRNA, have been proposed once the secondary structural elements were established. Tertiary canonical base pairings are proved from

	10	20	30	40
Tt	AG <u>UCUCAGGG</u>	GAAACUU <u>UG</u>	AGAUGGCCUU	.GCAA..... <u>.AGGGUA</u>
NcND4L	AGUUA <u>AUAUA</u>	GUA <u>AUAUAU</u>	UAAUGGAA <u>CA</u>	GGUAAUGACU
PaND3	AGCAUUA <u>AUC</u>	GAAAGGU <u>UAA</u>	UUGUGGCUGA	AGUAAU <u>ACU</u>
PaATP6	AGAUAAGACG	GAAACU <u>UCUU</u>	ACCUGGCC <u>UC</u>	AUUA <u>AUGAUU</u>
cons	AGyy..rayr	GAAy.U...	..rUGGcyy.	rguAAu.ayu

```

      A A
      G  A
      G-C
      G-U
      G-U
      A-U      GCA
      C-G      U  A
      U-A      U-A
      C-G      C-G
5' AG U-A UGG C-G UA 3'

```

Figure 2 (top) One possible alignment of four group I intron sequence fragments. Underlined regions can pair; “.” represents gaps. “cons” is the consensus sequence, where capital letters are fully conserved bases, lowercase letters are 75% conserved bases, and r and y refer to purine (G or A) and pyrimidine (C or U) bases, respectively. After Michel & Westhof (29). References: Tt (30); NcND4L (31); PaND3 (32); PaATP6 (31). (bottom) One possible secondary structure of the sequence fragment from Tt (*Tetrahymena thermophila* group I intron).

compensatory changes, just as is secondary structure. For example, phylogenetically proven pseudoknots have been found in group I and group II introns (22), and small-subunit rRNA (33). Isosteric replacement of bases in triples is at least an indication of triple base interactions, and examples have been found in tRNA (34) and the *Tetrahymena thermophila* group I intron (35). Similarly, natural sequence elements from RNAs of different overall function may have common functional subunits. Altman and coworkers, for example, found similarities between the secondary structure of RNase P RNA (16) and the exit site of *Escherichia coli* large-subunit rRNA (36). Altman suggests that this may be a common binding motif for tRNA.

### Artificial Phylogeny

Artificial mutations also confirm structures, and are particularly useful for studies of conserved sequences. Traditionally mutations are introduced in a genome with chemical mutagens, radiation, or UV light. The cells that survive often contain random mutations showing covariation of bases throughout structural RNAs. Hass et al grew *E. coli* cells in the presence of the chemical mutagen hydroxylamine, and found compensatory mutations in the conserved pseudoknot of the RNase P RNA (37). Specific site mutations, deletions, and insertions are quickly and easily made in cloned sequences. These methods

are faster than natural sequence cloning and searching in many cases. Structural studies of ribozymes have extensively used mutations (38); deletion mutants are particularly useful for finding the minimal sequence requirements of catalytic RNAs (e.g. 39). RNA mutations should be carefully checked for alternate structure formation with nucleases or chemical probes.

### *Random Mutations and Specific Selection*

Another way to generate a pool of similar function RNA was introduced independently by Joyce (40), Gold (41), and Szostak (42). SELEX [Systematic Evolution of Ligands by EXponential enrichment, as the method is called by one author (41)] finds sequences with a specific enzymatic activity or ligand-binding affinity. In general, there are four steps per cycle: transcription, selection, reverse transcription, and amplification. The process starts with a random sequence pool of DNA, which is transcribed to RNA. This RNA pool is selected for activity (40) or ligand binding (41, 42), which decreases the number of different mutants in the pool. The selected RNAs are reverse transcribed to DNA, and this DNA is amplified by the polymerase chain reaction. Several SELEX cycles at increasing stringency result in a small population of active RNAs. This powerful technique has been used to generate new RNAs with unique binding properties (42), to probe RNA-protein interactions (41), and to alter the substrate requirements of the *T. thermophila* group I intron (43).

### *Functional Group Mutations*

Functional group mutations probe RNA structure and function at the atomic level. These methods require the synthesis and specific incorporation of nucleotide analogs via chemical RNA synthesis (44, 45). Careful controls for each change must be made, since functional mutations alter the electronic structure of the nucleotide, and thus stacking and hydrogen bonding. Several groups used functional group mutations to probe the stereochemistry of phosphodiester formation and cleavage (46), to identify the guanosine binding site in the *T. thermophila* group I intron (47), and to determine the thermodynamic stability and structure of the  $G^A_{AG}$  internal loop (44) and the UNCG (48) and GNRA (49) hairpin loops.

## OPTICAL SPECTROSCOPY

### *Absorbance*

The temperature-induced transition between nucleic acid native and denatured (single-stranded) states can be conveniently monitored by UV absorbance [reviewed by Puglisi & Tinoco (50)]. The absorbance vs temperature profile

is commonly referred to as a UV absorbance melting curve, and the midpoint of the transition is called the melting temperature ( $T_m$ ). As the temperature increases, the ratio of molecules in the native vs denatured states decreases. Since the native and single-stranded states have different absorptivities, the change in absorbance with temperature effectively monitors the transition. Whereas many methods such as circular dichroism and NMR can monitor thermal denaturation, UV absorbance is the easiest to measure. Approximately 1–10 nanomoles of an oligonucleotide RNA are required for each melting curve.

The UV absorbance melting curve shows a characteristic increase in absorption as the stacked bases in the native state become unstacked in the denatured state; this increase in absorption is called hyperchromicity (51). The decreased absorbance of a native nucleic acid relative to its denatured state is hypochromicity. Hypochromicity is defined as the percent change in absorbance when a nucleic acid is denatured:

$$\text{Hypochromicity (\%)} = 100 (A_{\text{denatured}} - A_{\text{native}})/A_{\text{denatured}} \quad 1.$$

(where  $A_{\text{denatured}}$  is the absorbance of the denatured molecule at high temperature, and  $A_{\text{native}}$  is the absorbance of the native molecule).

If the melting curve depends on only two states (the native and denatured forms), thermodynamic parameters for the transition can be calculated. At each temperature the fraction of molecules in the single-stranded state, and thus the equilibrium constant, is calculated from the absorption. The thermodynamic parameters  $\Delta G_T^\circ$ ,  $\Delta H^\circ$ , and  $\Delta S^\circ$  for the transition (52) are calculated from the temperature dependence for the equilibrium constant:

$$\Delta G^\circ(T) = -RT \ln K(T) = \Delta H^\circ - T\Delta S^\circ. \quad 2.$$

The derivation of thermodynamic parameters from a melting curve depends on several assumptions (see Ref. 50). The most important assumption is that the nucleic acid exists only in two states: native or single-stranded. This two-state assumption can be tested by observing the transition with a different technique (e.g. calorimetry, NMR, or circular dichroism). The observation of isosbestic or isodichroic points in the absorbance or circular dichroism spectra, respectively, is diagnostic of a two-state transition (53). Alternatively, the absorbance profile can be monitored at a different wavelength, so that contributions from different nucleotides are emphasized. If all methods give the same results, the two-state model is validated since the methods have very different sensitivities to different species.

The dependence of the melting temperature,  $T_m$ , on the RNA strand concentration reveals the molecularity of the transition. For example, melting

of an RNA hairpin is a unimolecular process and therefore does not depend on concentration. Melting of a duplex is bimolecular and is concentration dependent for oligonucleotides. Of course, an oligonucleotide can melt as a hairpin (unimolecularly) at low concentrations but become a duplex with an internal loop and melt bimolecularly at high concentration. Higher-order complexes such as triplexes and quadruplexes show stronger dependences on concentration (54). Polynucleotides show negligible concentration dependence of the  $T_m$  because double strand initiation, which is the event dependent on concentration, is only a small fraction of the total free energy involved in the transition.

For bimolecular reactions, thermodynamic parameters can be derived by plotting reciprocal melting temperature vs logarithm of the total strand concentration,  $C_T$  (55). For a duplex formed from a self-complementary oligonucleotide:

$$1/T_m = (R \ln C_T)/\Delta H^\circ + \Delta S^\circ/\Delta H^\circ. \quad 3.$$

The concentration should be varied over a factor of 100 for reliable measurements of  $\Delta H^\circ$  and  $\Delta S^\circ$ . For a non-self-complementary oligonucleotide, the molecularity can be determined from UV absorbance mixing curves (where absorbance is measured while varying the stoichiometry of the reacting strands) (56). Equations 2 and 3 can then be used to test the two-state approximation. For example, if significant concentrations of intermediates are present, then the  $\Delta H^\circ$  from Equation 2 will be smaller than the  $\Delta H^\circ$  from Equation 3 (57). On the other hand, if the transition is known to be two-state, differences in  $\Delta H^\circ$  from Equation 2 and Equation 3 can reveal the molecularity. For example, a triplex would give a  $\Delta H^\circ$  value from Equation 2 that is 50% larger than that from Equation 3, and would require that equations appropriate for triplexes be used (54, 57).

In large molecules, such as tRNA or group I introns, melting is not two state; therefore more complicated models are required to derive thermodynamic parameters. UV melting, however, can reveal information of the stepwise mechanism of thermal denaturation. For example, UV melting of *E. coli* formylmethionine tRNA shows several transitions. First, the tertiary interactions are broken, and then individual helices (secondary structure) are melted (58, 59). The transitions are assigned with the help of other techniques such as circular dichroism, temperature-jump kinetics, NMR, or chemical modification. Since the hypochromicity of A•U and G•C base pairs are maximum at 260 and 280 nm, respectively (60), the ratio of A•U and G•C base pairs broken during each transition can be determined by measuring the hypochromicities at both wavelengths (50). For example, the melting curve of the cyclized form of the *T. thermophila* group I intron involves three major

transitions (61). The first occurs at 36C and is attributed to partial unfolding of the tertiary structure. The second transition occurs at 68C with the 260 nm hypochromicity greater than the 280 nm hypochromicity, suggesting the transition involves an A•U-rich region. The third transition occurs at 72C with the 280 nm hypochromicity greater than the 260 nm hypochromicity, suggesting that the transition involves G•C-rich regions. Chemical modification data support these conclusions (61a). These results are consistent with the general notion that tertiary structure is less stable than secondary structure in RNA. This is one reason for the success of secondary structure predictions for RNA.

### *Circular Dichroism*

Circular dichroism (CD) is the difference in extinction for right and left circularly polarized light. Circular dichroism can be measured in the absorbance bands of an RNA (below about 300 nm) with the same concentration and amount of material required for an ultraviolet absorbance measurement. For nucleic acids the CD spectrum is mainly dependent on the sequence and stacking geometry of the bases. Therefore a CD spectrum can easily distinguish among A-form, B-form, and Z-form double helices (62), and between helices and unstacked single strands (63). Circular dichroism is often used as a sensitive, but qualitative measure of conformation. The details of conformations are usually established by other methods such as NMR. However, CD can sometimes provide quantitative evidence of the amount and type of secondary or tertiary structure in an RNA. Examples include an estimate of the secondary structure and number of A•U, G•C, and G•U base pairs in *E. coli* 5S rRNA (64), and the identification of a pseudoknot in an oligonucleotide (65). The conformations of proteins are easily assayed by CD measurements (66); this can be very helpful in studying RNA-protein binding. Changes in conformation of the protein or the RNA or both can be detected as illustrated by a study of the binding of the HIV-1 Rev protein to the Rev-responsive RNA (67).

### *Fluorescence*

Fluorescence studies of RNA structure and dynamics are limited by the fact that the unmodified bases A, C, G, and U have negligible fluorescence quantum yields ( $\sim 10^{-4}$ ) (68). Therefore, fluorescence studies require a natural modified base such as wyosine in tRNA<sup>Phe</sup>, a synthetic modified base such as ethenoadenosine (69), or synthesis of an RNA modified with a chromophore (e.g. coumarin, pyrene, and fluorescein) connected via a linker (70). Proper controls need to be performed to be sure that inclusion of a fluorophore does not affect the structure of the molecule (71).

Since fluorescence is sensitive to environment, the change in fluorescence

with time can be used to follow kinetic processes such as substrate binding or RNA folding. For example, Bevilacqua et al (72) have investigated the kinetics and thermodynamics for binding a pyrene-modified oligomer substrate to the *T. thermophila* group I intron. For drugs that fluoresce, such as ethidium or proflavin, fluorescence can be used to study the environment near the fluorophore (68). Studies of proteins (73) as well as RNA (74) indicate that decay of fluorescence anisotropy can reveal hydrodynamic properties of a macromolecule or assembly of molecules. This information can be useful for interpreting NMR relaxation data (73).

Fluorescence energy transfer between a donor chromophore and an acceptor chromophore can yield distances from 10 to 80 Å (71, 75). Studies of this type were performed on tRNAs (76), and later X-ray crystal studies confirmed the distances determined by fluorescence (14). However, the fluorescence studies did not reveal the L-shape of the tRNA structure, because not enough distances were available (76). A more recent study used fluorescence energy transfer to deduce the topology of mRNA binding to the 30S ribosomal subunit with respect to the 3'-end of the 16S rRNA (77).

## CHEMICAL AND ENZYMATIC REACTIVITY

The secondary structure of an RNA molecule can be established by using chemical and enzymatic probes that distinguish between double-stranded regions and single-stranded regions on the basis of their reactivities. Accessibility to added reagents can indicate qualitatively the surface of an RNA and thus suggest its folded three-dimensional shape. With <sup>32</sup>P-end-labeled RNA molecules only picomole amounts of material are needed. Each reagent has a distinct specificity, so the larger the number of probes used, the more accurate the view of the folded RNA. Relative rates of reaction of different probes at different sites are most informative. Information about tertiary structure is obtained by use of crosslinking to reveal the proximity of two parts of the RNA widely separated in the sequence. The site of binding of a protein can be investigated by determining its protection of the RNA from chemical reaction. All of these studies can be done on nearly any size RNA; they are described in the following sections.

### Reagents

**NUCLEASES** Most endoribonucleases cleave single-stranded regions of the RNA faster than double-stranded regions. The exception is ribonuclease V<sub>1</sub>, which is specific for double-stranded—or stacked—regions of the RNA. To assess secondary structure, the strategy is to end label the RNA (for example with <sup>32</sup>P), then add an endonuclease so as to produce no more than one cut

**Table 1** Properties of some useful ribo-endonucleases

Enzyme	Specificity	Phosphates	Reference
RNase T <sub>1</sub>	G	2', 3'	79
RNase CL 3	C	2', 3'	79
RNase U <sub>2</sub>	A>G	2', 3'	79
RNase A	C, U	2', 3'	80
RNase Phy M	A, U	2', 3'	81
RNase T <sub>2</sub>	Single strand	2', 3'	79
<i>E. coli</i> RNase I	Single strand	2', 3'	82
Nuclease S <sub>1</sub>	Single strand	5'	83
Mung bean nuclease	Single strand	5'	84
Micrococcal nuclease	Single strand	3'	85
RNase V <sub>1</sub>	Double strand	5'	86

per molecule. The end-labeled fragments are analyzed by gel electrophoresis. To ensure that the fragments arose from single cuts, about 90% of the species seen on the gel should be full-length molecules. Table 1 lists enzymes useful for determination of secondary structure; the position of the phosphate left after hydrolysis is given. All the base-specific enzymes are also single-strand selective, but usually nuclease S<sub>1</sub> and ribonuclease V<sub>1</sub> are used to distinguish between single-strand and double-strand regions.

The use of nucleases in studying secondary and tertiary structure of RNA is reviewed by Knapp (78).

**CHEMICAL PROBES** Two methods are commonly used to determine relative rates of reaction of nucleotides with chemical probes. Cleavage at the site of chemical modification of a <sup>32</sup>P-end-labeled chain provides fragments whose lengths indicate the reaction site. For probes that do not lead to cleavage, the reaction site is identified by chain termination of reverse transcription. A primer is used to start the chain; the end of the chain occurs at the reacted base. Primer extension by reverse transcriptase is most useful for long RNA molecules (79, 87); it is less practical for oligonucleotides of less than about 30 nucleotides. In both methods chain length determination of the fragments is done by gel electrophoresis.

Table 2 lists reagents commonly used for probing secondary and tertiary structure. Double-stranded regions react more slowly with these reagents than do single-stranded regions; this is particularly true for the reagents that react with the Watson-Crick hydrogen-bonding groups. All the probes can be used with primer extension by reverse transcriptase, but only those that cleave the chain are suitable for RNA molecules too short for this method. Good reviews of applications of chemical probes to RNA structure are given by Ehresmann

**Table 2** Chemical probes used for nucleic acid structure determination

Probe	Specificity	Cleavage	Reference
Dimethylsulfate (DMS)	Adenine N1	No	90
Dimethylsulfate (DMS)	Cytosine N3	Yes	90
Dimethylsulfate (DMS)	Guanine N7	Yes	90
Ethylnitrosourea (ENU)	Phosphates	Yes	91
Diethylpyrocarbonate (DEP)	Adenine N7	Yes	92
CMCT <sup>a</sup>	Guanine N1	No	93
CMCT <sup>a</sup>	Uracil or thymine N3	No	93
Kethoxal <sup>b</sup>	Guanine N1–N2	No	93
Bromoacetaldehyde	Adenine N1–N6	No	94
Osmium tetroxide	Pyrimidine C5–C6	Yes	95
Potassium permanganate	Pyrimidine C5–C6	Yes	96
2:1 1, 10-Phenanthroline-Cu <sup>2+</sup>	Single strand	Yes	97
EDTA-Fe <sup>2+</sup>	Solvent accessible regions	Yes	98, 99
Methidiumpropyl-EDTA-Fe <sup>2+</sup>	Double strand	Yes	100
Uranyl acetate	Nonselective (light-activated)	Yes	101
Rh(phen) <sub>2</sub> phi <sup>3+</sup> <sup>c</sup>	Tertiary interaction sites (light-activated)	Yes	102, 103

<sup>a</sup> 1-Cyclohexyl-3-(2-morpholinoethyl) carbodiimide methane-p-toluene sulfonate

<sup>b</sup>  $\beta$ -Ethoxy- $\alpha$ -ketobutyraldehyde

<sup>c</sup> Bis(phenanthroline)(phenanthrenequinonediimine)-rhodium(III)

et al (79), Nielsen (88), and Volume 180 of *Methods in Enzymology*. Sigman & Chen (89) have reviewed the mechanisms and applications of the last five entries in Table 2.

### *Secondary and Tertiary Structure*

The RNA molecules of protein synthesis have been most extensively studied; they include 16S (93), 18S (104), and 5S rRNAs (105, 106), and tRNAs (107, 108). The data have shown that RNA regions that lack Watson-Crick base pairs still may have double-strand properties as seen by ribonuclease V<sub>1</sub> sensitivity and reduced reactivity to chemical probes. An example is loop E of eukaryotic 5S rRNA. The sequence suggests an internal loop of nine nucleotides, but chemical and enzymatic probing indicated four non-Watson-Crick base pairs and either a bulged U (105) or A (109). Determination of the structure of the loop E sequence by NMR (109a) shows four non-Watson-Crick base pairs plus a G bulged into the major groove of the RNA. As chemical reactivity of each base depends not only on base pairing but also on base stacking, solvent accessibility, and the detailed three-dimensional structure, it is probably best to interpret the chemical reactivity qualitatively. Chemical and enzymatic probing methods are ideal for rapid identification of a particular structural element, such as a pseudoknot (110).

Iron-EDTA (EDTA-Fe<sup>2+</sup> where EDTA is ethylenediaminetetraacetic acid)

is a versatile probe of RNA tertiary structure. In solution,  $\text{EDTA-Fe}^{2+}$  discriminates between solvent-accessible and solvent-inaccessible regions (99). Celander & Cech (111, 112) used  $\text{EDTA-Fe}^{2+}$  to study the solution folding of the *T. thermophila* group I intron as a function of different divalent cations.  $\text{EDTA-Fe}^{2+}$  becomes a site-specific probe when it is tethered to a ligand; the hydroxyl radicals attack only surrounding atoms. The guanosine-binding site of the *T. thermophila* group I intron was probed by  $\text{EDTA-Fe}^{2+}$  linked to guanosine monophosphate (113).

**CROSS-LINKING** Reagents with two reactive groups can be used to link covalently nucleotides that are far apart in the sequence; a crosslink thus provides an upper bound to the distance between the nucleotides. The localization of the crosslink can be done precisely by partial hydrolysis of the RNA and identification of the reacted nucleotides. Useful chemical crosslinking reagents include *bis*-(2-chloroethyl)-methylamine or "nitrogen mustard," *p*-azidophenyl acetimidate, and iminothiolane; all have been used to obtain a detailed model of the three-dimensional structure of *E. coli* 16S rRNA in the 30S subunit (114).

Psoralen is one of a class of compounds that photoreacts specifically with nucleic acids; proteins do not react and no nucleic acid-protein crosslinks occur. In the absence of ultraviolet irradiation no reaction occurs. The photo-crosslinking can be done *in vivo* or *in vitro* to probe conformations in a wide range of environments (115). Psoralen photocrosslinks combined with the chemical crosslinking have led to further progress in a three-dimensional structure for 16S rRNA (116).

Ultraviolet light can induce nucleic acid-nucleic acid and nucleic acid-protein photocrosslinks. Further evidence for the structure of the catalytic site of group I introns (117), and a model for the tRNA-binding domain of *E. coli* 23S rRNA (118) have been obtained by this method.

### Footprinting

Chemical and enzymatic reactivity has been widely used in footprinting—assessing the sites of binding of proteins to nucleic acids. Different size probes can give different apparent footprints; the smaller chemical probes provide higher-resolution footprints. Tight-binding enzymes can even perturb the equilibrium and displace the target protein. Some representative examples of chemical footprinting studies include ribosomal protein L11 plus 23S rRNA studied by diethylpyrocarbonate (DEP) and hydrazine (119); *lacZ* messenger RNA plus 70S ribosomes studied by DEP, dimethylsulfate, and copper-phenanthroline (120); pre-mRNA splicing-specific complexes (121) and TFIIIA-5S rRNA complex (122) studied by  $\text{EDTA-Fe}^{2+}$ ; tRNA<sup>Ser</sup> complex with seryl-tRNA synthetase (123) and the HIV Tat-TAR complex (124) studied by ethylnitrosourea. Some novel enzyme footprinting studies include use of

ribonucleases A and T<sub>1</sub> to probe the RNA transcript in a ternary complex of RNA polymerase II, DNA template, and nascent RNA (125); use of nuclease S<sub>1</sub> and complementary deoxyoligonucleotides to find single-stranded regions of 16S rRNA in 70S ribosomes and 30S subunits (126).

## GEL ELECTROPHORESIS

Polyacrylamide gel electrophoresis is a common analytical technique for the purification and sequence analysis of RNA. Denaturing gels use extreme pH, urea, and high temperature (from the gel's resistance to the applied current) to denature secondary and tertiary structure, so that separation is by charge and molecular weight. Native gels can use any buffer, have no denaturant, and use external control of the temperature; native gels separate by structure as well as charge and molecular weight. Temperature (127, 128) or denaturant (128, 129) gradients may be applied perpendicular to the electric field, to quickly optimize separation, or as a probe of stability. Electrophoretic mobility for a given molecule depends on several factors, including the percentage crosslinking and the density of the gel, power level, temperature, and buffer ionic strength. These factors can be adjusted to optimize separation.

Native gel electrophoresis is an important tool for separation by structure. Multimeric species can be separated (130), protein-nucleic acid interactions observed (131), and binding constants of oligomers to catalytic RNAs have been reported (132) but differ from those measured by equilibrium dialysis (132a). Typically, these native gels are run under similar conditions to denaturing gels (percentage crosslinker and total acrylamide concentrations), but at reduced temperature and wattage, and in the absence of denaturant. The reduced temperature helps stabilize weak interactions, while the low wattage prevents gel heating.

Denaturant (127, 128) and temperature (128, 129) gradients are useful to generate melting curves [see the review by Riesner et al (129)]. The gradient is applied perpendicular to the electric field, as is a single RNA sample. Several transitions appear, which are due to tertiary or secondary destabilizations (128, 129); this technique may be more sensitive to tertiary melting transitions than are optical detection methods. The melting behavior of different species can also be detected, assuming their mobilities are different (129). This would be a difficult task using standard optical melting techniques.

## NUCLEAR MAGNETIC RESONANCE (NMR)

NMR techniques measure distances with through-space interactions (nuclear Overhauser effect or NOE) and measure dihedral angles using through-bond interactions (J-coupling) (133). Measurement of enough of these interactions

can define a structure nearly as well as X-ray crystallography (135). Optimally, NMR samples are about 0.5 ml of solution at as high of a concentration as possible (usually around 2 mM), and of high purity. Proton ( $^1\text{H}$ ) NMR is the most common; other nuclei are important at natural abundance ( $^{31}\text{P}$ ,  $^{19}\text{F}$ ) or in enriched samples ( $^2\text{H}$ ,  $^3\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ). These nuclei have characteristic peaks or resonances, which are sensitive to their environment. The frequency dispersion of these peaks, called the chemical shift, is small (measured in parts per million or ppm) relative to the NMR spectrometer's frequency (e.g. 600 MHz). Derome (144) has a good introduction to NMR, Wüthrich (133) provides an introduction to biological NMR, and Varani & Tinoco (137) have a recent review of RNA NMR techniques.

Assignment of resonances is the first step in determining a structure using NMR. RNA A-form helix geometry has characteristic NOEs and J-couplings for proton assignments (134, 136). Protons in nonstandard geometries are grouped by the spin systems (a series of protons with sequential J-coupling, like the protons in a ribose sugar), and then by NOEs, which connect the sugars (137). After the resonances are assigned, the interactions between nuclei are measured, and converted to distance and angle constraints. An experienced NMR spectroscopist can assign peaks and measure all the interactions in less than a year for a small (about 12 nucleotides) RNA; mutants of already assigned RNAs can be analyzed much more quickly. Only partial assignment may be possible for larger RNAs using standard proton NMR techniques (136), and they take much longer to analyze. The highest-resolution RNA NMR structures published, which have most of their proton resonances assigned, are the UUCG (135) and GNRA (138) hairpins. Published medium-resolution NMR structures, in which many protons are not assigned, include Helix I (139) and Loop E (109a, 136) from 5S rRNA, a pseudoknot (140), a hairpin containing an A•C pair in the stem (141), the three-base bulge from the HIV-I TAR stem (142), an RNA G-quartet (143), and a minor groove triple (143a).

### *Proton NMR*

Nucleic acids have two types of protons, nonexchangeable and exchangeable. Nonexchangeable protons are bound to carbon. Exchangeable protons are bound to nitrogen or oxygen, and exchange rapidly with the surrounding water. The rate of exchange is slowed by hydrogen bonding or decreased accessibility by the solvent. This rate is pH dependent and is influenced by the buffer species, which act as catalysts for proton exchange. The exchange rate is an important probe of structure and dynamics (145).

Each nucleotide contains 10 or 11 protons, of which the H1', the aromatic and amino, and the imino protons all have their own distinct  $\sim 1$  ppm (chemical shift) regions; the five remaining sugar protons resonate in a  $\sim 1$  ppm region.

These regions become more crowded for larger molecules. Base composition and the three-dimensional structure of the RNA influence the chemical shift dispersion of the resonances. For example, a GGAC(UUCG) GUCC hairpin structure stacks the C7H5'' above the G8 base, causing an unusual chemical shift of 1.5 ppm above the normal region (135).

Chemical shift dispersion problems are partially alleviated by multidimensional techniques, which spread the interactions into two or more dimensions. For example, the cross peaks in two-dimensional (2D) experiments describe an interaction between two protons. These extra dimensions reduce the resonances' spectral overlap, making assignments and quantitation easier.

Two-dimensional NOESY (NOE Spectroscopy) measures through-space, dipole-dipole interactions. NOESY cross peak volume is related to the time the two proton dipoles interact (the mixing time) (146, 148). The buildup rate  $\sigma$  is the slope of a plot of the NOESY cross peak volume vs mixing time. The initial buildup rate is linear with respect to  $r^{-6}$ , where  $r$  is the distance between two protons. Distances are measured by comparing an unknown buildup rate  $\sigma_u$ , with the buildup rate of two protons at a known, fixed distance—usually a pyrimidine H5-H6 pair (146).

$$\sigma_{5-6}/\sigma_u = r_u^6/r_{5-6}^6 \quad 4.$$

NOESY cross peaks of exchangeable protons are more difficult to quantify, since they include both chemical exchange with the solvent, and NOE. The  $r^{-6}$  dependence limits the measurable NOE to distances less than 5 Å, since the peak intensity falls off rapidly with distance (147). A NOESY cross peak for a 5 Å distance is  $(2.4 \text{ Å}/5.0 \text{ Å})^6 = 0.01$  times the intensity of an H5-H6 cross peak; this intensity is only slightly larger than the noise. The NOE and NOESY are clearly described in Neuhaus & Williamson (148).

At longer mixing times, the buildup rate depends not only on the local pair interaction, but also on magnetization transfer to or from other protons. Relaxation matrix programs use these data by calculating the relaxation of all protons in a molecule simultaneously (149, 150). Recent studies have combined dynamic calculations and relaxation matrix techniques (261).

Through-bond interactions (J-coupling) occur between nuclei separated by one or more chemical bonds. J-couplings are measured using 2D COSY, or COherence Spectroscopy, experiments (144, 147), but it is easier to measure coupling constants in a double-quantum filtered COSY, which removes some of the experimental artifacts of the COSY experiment (151). The size of the coupling constant depends sinusoidally on the dihedral angle between the two nuclei (152). The dihedral angle is calculated from the coupling constant using an empirical equation (called a Karplus relation) fit to model compounds

**Table 3** Dihedral angle measurement via three-bond J-couplings<sup>a,b</sup>

Dihedral angle	three-bond J-coupling				
	H-H	H- <sup>31</sup> P	H- <sup>13</sup> C	<sup>13</sup> C- <sup>13</sup> C	<sup>13</sup> C- <sup>31</sup> P
$\beta$		H5' (")-P <sub>i</sub>			C4'-P <sub>i</sub>
$\gamma$	H4'-H5' (")		C3'-H5' (")		
$\delta$	H3'-H4'		C2'-H4', C5'-H3'		
$\epsilon$		H3'-P <sub>i+1</sub>			C4'-P <sub>i+1</sub> , C2'-P <sub>i+1</sub>
$\chi$			C8-H1', C6-H1'	C8-C2', C6-C2'	
sugar pucker	H1'-H2', H2'-H3'		H1'-C3', H2'-C4', H3'-C1'		

<sup>a</sup>P<sub>i</sub> refers to the phosphate bound to C5'; P<sub>i+1</sub> refers to the phosphate bound to C3'.

<sup>b</sup>H5' (" refers to both H5' and H5'' protons.

(153). Three-bond coupling is the most common in RNA proton NMR; it can define the backbone torsion angle  $\gamma$ , the sugar pucker (the pseudorotation angle or the equivalent torsion angle  $\delta$ ), and sugar pucker amplitude (134, 135, 153) (Table 3). Dihedral angle errors arise from coupling constant measurement errors (primarily due to linewidth of the resonances), and the variation of the Karplus equation with the dihedral angle near the measured coupling constant.

Other coherence transfer experiments are important for assignment of spectra. These often rely on an initial estimate of the coupling constants. TOCSY (TOTally Correlated SpectroscopY) experiments (154) couple all the protons in a sugar. Multiple quantum spectroscopy is useful for assignments of coupled protons with similar chemical shifts (155). Finally, E COSY experiments (156) edit the COSY cross peak, and are useful for measuring small J-coupling constants.

### Heteronuclear NMR

The major impediment to NMR studies of RNA molecules greater than 25 nucleotides is spectral overlap (137). 3D and 4D NMR (157) studies of <sup>13</sup>C and <sup>15</sup>N isotopically enriched proteins (80–150 amino acids) have significantly reduced spectral overlap (158–160), have allowed the measurement of useful J-couplings (161–163), and have allowed characterization of protein dynamics (164–167). Isotopic enrichment is expected to produce similar advances in NMR studies of RNA (168, 169).

Heteronuclear (<sup>13</sup>C and <sup>15</sup>N) NMR is useful for the following reasons:

1. Chemical shifts are dispersed, so that 3D and 4D NMR experiments that utilize at least one  $^{13}\text{C}$  or  $^{15}\text{N}$  dimension significantly reduce the problem of spectral overlap, allowing larger RNA molecules to be studied (168).
2. Spectral editing (170) simplifies a spectrum by selecting the signals of protons that are either bonded to one isotope ( $^{12}\text{C}$ ) or another ( $^{13}\text{C}$ ). Three editing strategies are to selectively enrich only one type of carbon (e.g. label all C1' positions); to label only one type of nucleotide (e.g. adenosine residues); to label one strand at a time in a multistrand complex.
3. Measurement of three-bond  $^1\text{H}$ - $^1\text{H}$ ,  $^1\text{H}$ - $^{31}\text{P}$ ,  $^1\text{H}$ - $^{13}\text{C}$ ,  $^{13}\text{C}$ - $^{13}\text{C}$ , and  $^{31}\text{P}$ - $^{13}\text{C}$  J-couplings allows the determination of backbone and glycosidic torsion angles (171–173).
4.  $^{13}\text{C}$  and  $^{15}\text{N}$  relaxation measurements allow characterization of molecular dynamics (174–176).

$^{13}\text{C}$  enrichment is required because  $^{13}\text{C}$  is only 1.1% naturally abundant, and has a low-sensitivity nucleus (3% as sensitive as protons). Sensitivity is especially important for RNA studies because of limited solubility and aggregation at high concentrations. Sensitivity can be improved with pulse sequences that allow proton detection via heteronuclear multiple quantum coherence (HMQC) (137, 177), or polarization transfer from proton to carbon nuclei (178). The HMQC experiment is useful for making proton and carbon assignments and is sensitive enough to detect  $^1\text{H}$ - $^{13}\text{C}$  correlations in RNA samples at approximately 2 mM concentrations and natural abundance  $^{13}\text{C}$  (179).

The lack of heteronuclear NMR studies of RNA has been due to the lack of sample availability (177). However, methods are now available (180, 181; J. V. Hines, personal communication) for preparing uniformly  $^{13}\text{C}$ - and  $^{15}\text{N}$ -enriched nucleoside triphosphates (NTPs) from nucleoside monophosphates (NMPs) (182, 183) derived from the RNA of cells grown on  $^{13}\text{C}$ - and/or  $^{15}\text{N}$ -enriched media. This method is adequate for the synthesis of uniformly enriched nucleotides, but is inefficient because most of the label is incorporated into other cellular components. The RNA of interest is synthesized from enriched NTPs using transcription by T7 RNA polymerase (182). Alternatively, the 5'-phosphate of NMPs can be removed with acid phosphatase and the base and ribose moieties protected for chemical synthesis (185, 186). A reasonable chemical approach to the synthesis of selectively labeled nucleosides is to synthesize the labeled ribose (187–190) or base (191–193) separately and then couple the base to the ribose ring (173, 194). Selectively  $^{13}\text{C}$ -enriched riboses are commercially available (C1', C2', C3', and C5' from Cambridge Isotope Labs and Omicron Biochemicals). Nucleosides can be

converted to 5' nucleoside monophosphates (195) and NTPs by chemical (196) or enzymatic methods (182, 183).

There are two strategies for  $^{13}\text{C}$  enrichment: uniform and selective. The advantages of uniform enrichment are high sensitivity and the ability to monitor all carbons simultaneously. Disadvantages result from large one-bond  $^{13}\text{C}$ - $^{13}\text{C}$  J-couplings (35–45 Hz) (173). The one-bond  $^{13}\text{C}$ - $^{13}\text{C}$  J-couplings decrease sensitivity and spectral resolution, complicate lineshape thus hindering three-bond J-coupling measurements, and complicate relaxation behavior so that dynamical properties of the molecule are more difficult to extract. Partial  $^{13}\text{C}$ -uniform enrichment reduces the population of sugars with adjacent labeled carbons, but also lowers the sensitivity. Recent studies indicate partial enrichment (30% uniform) allows J-couplings to be accurately measured (J. V. Hines, G. Varani, personal communication). Selective enrichment obviates the problems associated with one-bond J-couplings, but all carbons cannot be simultaneously monitored. An alternative enrichment strategy is to obtain a series of selectively labeled compounds and then mix them together in equal proportions; this is termed "ensemble enrichment." The NMR spectra of the resulting ensemble-enriched RNA would allow all carbons to be simultaneously monitored and would have excellent lineshape, spectral resolution, and relaxation behavior. NMR experiments on selectively or ensemble-enriched RNA will allow relaxation parameters to be measured. Given these parameters, the timescales and amplitudes of molecular motion can be derived for each residue by fitting the results to some motional model (164–167).

One problem in  $^1\text{H}$ - $^1\text{H}$  2D NMR experiments on RNA is the severe overlap of H2', H3', H4', H5', and H5'' sugar resonances between 4 and 5 ppm (136, 137). In large RNA molecules, spectral overlap precludes assignment and measurement of NOESY cross-peak build-up rates. Since  $^{13}\text{C}$  resonances for these positions resonate over a ~50 ppm range (179), heteronuclear 3D NMR NOESY-HMQC (157, 195) (nuclear Overhauser and exchange spectroscopy edited by heteronuclear shift correlation) and HOHAHA-HMQC (Homonuclear Hartman-Hahn spectroscopy edited by heteronuclear shift correlation) resolve overlapping cross peaks and provide more accurate assignments and NOESY cross-peak volumes. For example, interresidue distances involving H2' and H3' to H5'/H5'' protons are sensitive to helix geometry, but are often in crowded regions of 2D experiments (169). Similarly, the interresidue H1' to H5'/H5'' distance is 4.7 Å in an A-form helix, but is 1.8 Å in a B-form helix. Furthermore, RNA molecules with tertiary interactions are likely to bring the sugars of residues separated in the primary sequence close together. Inclusion of these restraints will significantly improve the quality of NMR-derived structures.

2D NMR methods rely heavily on NOESY analysis for making assign-

ments. Since assignments based on NOEs rely on conformational assumptions, assignment errors can occur when the assumptions do not hold (137). A more reliable method is to make assignments based only on through-bond J-couplings (135, 137, 197, 198). Pardi & Nikonowicz have developed methods for unambiguously assigning sugar spin systems based only on one-bond  $^1\text{H}$ - $^{13}\text{C}$  (145–160 Hz) and  $^{13}\text{C}$ - $^{13}\text{C}$  (38–45 Hz) J-couplings (199). Next, each sugar spin system needs to be assigned to the correct residue. One approach is to relay magnetization from one sugar to the next in the chain via the phosphorus (200). A sensitive experiment for transferring magnetization from the protons on one sugar to the protons on the next sugar in the sequence via the connecting phosphorus is the 2D-heteronuclear TOCSY-NOESY (201, 202). In large RNA molecules, spectral resolution would need to be improved by a 3D NMR experiment (e.g. by transferring magnetization to  $^{13}\text{C}$ -labeled carbons). Alternatively, magnetization can be transferred between phosphorus and  $^{13}\text{C}2'$  or  $^{13}\text{C}4'$  and then transferred to protons for sensitive detection (A. Pardi, personal communication). Both of these experiments, however, require a  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{31}\text{P}$  triple resonance probe, which is not commonly available.

Knowledge of the hydrogen bonding between bases is crucial to establishing the structure of an RNA. Overlap of the exchangeable protons in the 2D NMR of large RNAs is extensive (203). The  $^{15}\text{N}$  resonances, however, have moderate spectral dispersion ( $\sim 15$  ppm), allowing proton resonances to be well resolved in 3D  $^{15}\text{N}$  NOESY-HMQC spectra (168). Interactions involving adenine, cytosine, or guanine amino protons are important, but NOEs to amino protons are often difficult to detect because of broadening due to rotation about the C-N bond (137). However, interactions involving amino protons can be indirectly determined by allowing NOESY magnetization transfer from an assigned proton to an amino proton, then transferring the amino proton magnetization to the amino nitrogen, and finally to a base carbon that is detected. Since the one-bond  $^1\text{H}$ - $^{15}\text{N}$  J-coupling is large, the transfer from amino proton to amino nitrogen magnetization is efficient, and occurs faster than the exchange broadening. Preliminary experiments indicate this experiment works well (A. Pardi, personal communication).

By measuring three-bond J-couplings between  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{31}\text{P}$ , several dihedral angles (see Table 3) could be determined using empirical Karplus equations (171–173). Note that the four couplings used to determine the  $\gamma$  dihedral angle allow stereospecific assignment of  $\text{H}5'$  and  $\text{H}5''$  without any conformational assumption. This information will be important since it will allow more precise NOE-derived distance constraints involving  $\text{H}5'$  and  $\text{H}5''$ . This additional information will be especially useful in loop regions of RNA where conventional methods do not characterize the structure well. Measuring the J-couplings to determine  $\chi$  might seem superfluous, since NOEs alone specify the glycosidic torsion angle well, but NOE information can sometimes

be misleading due to spin-diffusion (146), so information redundancy is important.

One potential problem with using J-couplings to determine dihedral angles is that broad linewidths result in considerable cancellation of the antiphase multiplet components, which in turn distorts coupling constant determination and reduces sensitivity (204). In proteins, cancellation of antiphase components can be completely eliminated by using one-bond heteronuclear coupling to resolve heteronuclear three-bond couplings (161–163). In RNA, the same experiment can be performed (J. V. Hines, G. Varani, personal communication) to measure heteronuclear J-couplings and thus obtain information on dihedral angles as listed in Table 3. The properties of this experiment make it especially attractive for studies of large RNA molecules, where slow molecular tumbling results in linewidths on the order of 10 Hz.

Measurement of  $^1\text{H}$ - $^{31}\text{P}$  and  $^{13}\text{C}$ - $^{31}\text{P}$  3-bond J-couplings allow determination of the  $\beta$  and  $\epsilon$  dihedral angles.  $^1\text{H}$ - $^{31}\text{P}$  J-couplings are routinely measured with  $^1\text{H}$ - $^{31}\text{P}$  heteronuclear-COSY (205) for small RNA molecules (135, 139). In principle,  $^{13}\text{C}$ - $^{31}\text{P}$  couplings are directly obtainable from high-digital-resolution proton-decoupled 1D  $^{13}\text{C}$  spectra (assuming resolved resonances and favorable linewidths) (172). Recently, heteronuclear single quantum correlation spectroscopy (HSQC) with high digital resolution in the  $^{13}\text{C}$  dimension has been used to measure accurately  $^1\text{H}$ - $^{31}\text{P}$  and  $^{13}\text{C}$ - $^{31}\text{P}$  J-couplings in DNA (205a). The excellent lineshape and chemical shift dispersion of this experiment should allow  $^1\text{H}$ - $^{31}\text{P}$  and  $^{13}\text{C}$ - $^{31}\text{P}$  J-couplings to be measured in 20–30-nucleotide RNA molecules.

In conclusion, the reduced spectral overlap and additional J-coupling and NOE information provided by NMR studies of  $^{13}\text{C}$ - and  $^{15}\text{N}$ -enriched RNA will yield greatly improved structural information over what is currently obtainable.

## X-RAY DIFFRACTION

X-ray diffraction analysis of single crystals has provided information at atomic resolution for many proteins and DNAs. Until recently, X-ray diffraction by single crystals of RNA has been limited to ApU (206) and GpC (207), and several tRNAs [e.g. tRNA<sup>Phe</sup> (14, 208), tRNA<sup>Asp</sup> (209), tRNA<sup>Gly</sup> (210), tRNA<sup>Fmet</sup> (211, 212), and tRNA<sup>I<sub>met</sub></sup> (213)]. The development of chemical (185, 186) and enzymatic (184) methods for preparing milligram quantities of pure RNA oligonucleotides has allowed the growth of several crystals in high salt (monovalent ions, divalent ions, and spermine) suitable for X-ray crystallographic studies (214, 215). Thus, detailed information on the sequence dependence of RNA conformation is being obtained.

The single-crystal X-ray studies of tRNAs have provided a wealth of

stereochemical information, and have shown that RNA can fold into a complex tertiary structure similar to that of proteins (216). The overall shape of tRNA is an extended L, with the anticodon loop and acceptor-end at opposite ends. The four helices agree largely with average helical parameters deduced from diffraction studies of A-form fibers (217, 218). Secondary structural motifs found include helices with Watson-Crick A•U and G•C base pairs, and a G•U mismatch, hairpin loops, and a four-stem multi-branched loop or junction. Other interactions include two examples of coaxially stacked helices and several unusual hydrogen bonds between bases, 2'-hydroxyls, and phosphates. Tertiary base pairs include a G•A mismatch, a Hoogsteen A•T, parallel strand G•C, and various base triples. The structure of tRNA<sup>Phe</sup> shows that in addition to the many base-base hydrogen bonds, 72 of the 76 residues are involved in base stacking. The locations of four or five (depending on the tRNA) magnesium ions, critical for stabilizing regions of high charge density, were also located. Refinements of tRNAs (219, 220) have provided more structural detail, including the placements of 60–115 ordered water molecules, more precise sugar puckers, and characterization of thermal motions from isotropic (as well as anisotropic) crystallographic temperature factors. The crystal structure of lead-bound tRNA<sup>Phe</sup> provided the first structural characterization of RNA catalysis, i.e. lead-catalyzed cleavage of the tRNA (221). Crystal structures of *E. coli* glutamyl-tRNA synthetase complexed with its cognate tRNA<sup>Gln</sup> and ATP (222), as well as of yeast aspartyl-tRNA synthetase complexed with tRNA<sup>Asp</sup> (223), have revealed modes by which proteins bind to and discriminate among different RNA structures.

Only two high-resolution structures of small RNA molecules have been published: (U(UA)<sub>6</sub>A)<sub>2</sub> (214) and (pGGACUUCGGUCC)<sub>2</sub> (215). Several groups have reported crystals that diffract to less than 3 Å resolution, but these structures are not yet solved [(GGCGAGCC)<sub>2</sub> (S. Holbrook, J. SantaLucia Jr., unpublished results), (GGCGCUUGC GCC)<sub>2</sub> (J. Doudna, personal communication), and (GGCCGCAAGGCC)<sub>2</sub> (S. Holbrook, personal communication)]. The two most applicable methods for determining the structure from the diffraction pattern for macromolecules are isomorphous replacement and molecular replacement. With isomorphous replacement, a heavy atom derivative (e.g. made by replacing C with bromocytosine or replacing magnesium with lead) of a molecule that crystallizes in the same unit cell as the original molecule is used. Molecular replacement is a computational technique where a model of part of the structure is randomly oriented within the unit cell until an optimum match with the diffraction data is obtained (224). The rest of the structure is then solved iteratively. Molecular replacement is particularly appropriate for RNA, since the stems of RNA are usually close to canonical A-form.

The structures of (U(UA)<sub>6</sub>A)<sub>2</sub> and (pGGACUUCGGUCC)<sub>2</sub>—called the

UUCG-duplex—were solved by molecular replacement to 2.25 Å and 2.0 Å resolution, respectively, allowing detailed conformational analysis. Both (U(UA)<sub>6</sub>A)<sub>2</sub> and (pGGACUUCGGUCC)<sub>2</sub> show features typical of A-form helices and superimpose on ideal A-form helices with RMS differences of 1.75 Å and 1.48 Å, respectively. A more detailed look at the structures reveals the structural variability of different steps. The (U(UA)<sub>6</sub>A)<sub>2</sub> helix has two kinks stabilized by intermolecular interactions (due to crystal packing) involving 2'-hydroxyl groups. The UUCG-duplex also forms intermolecular contacts involving 2'-hydroxyl groups. The intermolecular contacts provide clues to how helices might pack in higher-order RNA structures such as ribosomes and ribozymes. Both structures also reveal many water molecules, some of which are integral parts of the structure. For example, in the UUCG-duplex a water molecule bridges a guanine amino to 2'-hydroxyl in a G•U mismatch; this was also seen in a G•U mismatch in tRNA<sup>Phe</sup> (219, 220). A water molecule bridges the H1 and N1 atoms of the U•C mismatches in the UUCG-duplex, thus allowing the pyrimidine-pyrimidine mismatch to have geometry similar to those of Watson-Crick base pairs. Other water molecules are less ordered, but do provide clues to the organization of the first and second hydration layers. This type of information will be important for improving molecular force fields for future molecular dynamics simulations.

The solution structure of pGGAC(UUCG)GUCC (where residues in parentheses are in the loop region) solved by NMR spectroscopy (197), is different from the crystal structure. It is a hairpin in solution, but a duplex with an internal loop in the crystal phase. GGCC(GCAA)GGCC also forms a hairpin in solution (138), and preliminary X-ray crystal results suggest the molecule crystallizes as a duplex (S. Holbrook, personal communication). NMR studies suggest GGCG(CUUG)CGCC can be either a duplex or hairpin in solution, depending on ionic conditions (168), but the crystal is duplex only (J. Doudna, personal communication). These results have been encouraging from the standpoint that the RNA database is increasing, but frustrating from the point of view of understanding the more biologically significant hairpin structures (225).

## STRUCTURE PREDICTION AND MODELLING

RNA secondary and tertiary structures are calculated using free energy minimization techniques. Secondary structure prediction programs minimize the thermodynamic free energies of formation. These free energies are derived from experimental measurements of RNA structural element stability (226). The thermodynamic method is best for predicting stable double helices because of strong hydrogen bonding and stacking interactions. Thermodynamic secondary structure prediction programs are most useful in providing

probable, low free energy structures. They are less useful in choosing the most stable structure, as the data are not good enough to discriminate among different structures with similar free energies. Tertiary structure prediction uses the interactions between atoms or groups to find a structure (227). Standard bond lengths and bond angles are assumed, then van der Waals interactions and electrostatic interactions between nonbonded groups are calculated to obtain a structure. Both methods allow experimental data from phylogeny, mutational analysis, chemical modification, and NMR to guide the calculations to an optimum result.

### *Secondary Structure Prediction*

Secondary-structure prediction programs use thermodynamic parameters derived from oligonucleotides, which form simple base-paired structures. In one well-tested case, the program predicts about 70% of phylogenetically predicted helices in the optimal structure of a wide variety of RNAs (61). Five algorithms for predicting secondary structures are dynamic programming (228), combinatorial (229), Monte Carlo (230, 231), significance (232), and a hybrid phylogenetic-thermodynamic approach (230, 233, 234). These programs all assume that the folded RNA molecules are at equilibrium; there are no kinetic barriers that trap the molecules in thermodynamically unstable states. The programs are particularly useful where only a single sequence is available, since several sequences are needed for phylogenetic analysis. For example, single-base mutations in cloned sequences are easily checked for alternate structure formation, and the mRNA of cloned genes can be scanned for extra-stable structures that may prevent translation. There are three recent reviews of secondary structure prediction programs (226, 235, 236).

These programs use the 10 nearest neighbors involving Watson-Crick G•C and A•U base-pairs (237) and 11 additional nearest neighbors involving G•U mismatches (238); optical melting studies have determined thermodynamic parameters for all 21 of these nearest neighbors. The sequence dependence of thermodynamics of loop regions is an area of active research. So far, results indicate that for bulge loops the surrounding base pairs and the length of the bulge loop are more important than the identity of the bases in the bulge (239). Internal loops and hairpin loops depend strongly on closing base pairs (240, 241), loop sequence (240–243), and loop length (243–245). These effects, except the sequence dependence of internal loops, are included in a current folding algorithm (246).

The dynamic programming or recursive method first calculates the free energies of all sequence fragments with five or more bases (five bases can form the smallest secondary structure considered—a three-base hairpin closed by a single base pair) (228). This list of free energies is used to find the minimum energy structure. Calculating the list is time intensive, but structure

determination is quick. Extensions to the dynamic programming model allow folding of circular sequences and multiple structure generation. Dynamic programming algorithms cannot predict pseudoknots or other tertiary structures. A recent version has folded sequences as large as 4217 nucleotides (M. Zuker, personal communication).

The combinatorial algorithm generates a list of all possible helices, then generates all possible structures from this list. By its design, the combinatorial method can generate multiple structures, and could be used to fold pseudoknots (although currently does not). This algorithm is useful for sequences of about 150 nucleotides or less (229, 235).

The Monte Carlo method generates a table of helices, creates a starting structure, then replaces helices to search for a minimum energy structure (230). As with the combinatorial method, multiple structures and pseudoknots can be predicted (230, 231). The statistical nature of Monte Carlo programs means that there is no guarantee that the lowest energy structure will be found.

Significance methods compare the thermodynamic stability of a given sequence with alternate sequences having the same base composition (232). These methods use a large amount of computer time and space, and require supercomputing facilities.

Recently, three programs have combined phylogenetic analyses and thermodynamic prediction of secondary structure (230, 233, 234). These programs generate multiple structures near the minimum free energy for two or more sequences, and the configurations generated are compared by score (230), pattern analysis (233), or tree search (234). The final structures are those that agree with folds from all sequences. These algorithms show great promise, since the phylogenetic data will compensate for any missing or erroneous thermodynamic data, and vice versa.

Further improvements in these methods will come from more measured thermodynamic parameters from oligonucleotides and better thermodynamic models. For example, there are no thermodynamic measurements of RNA junctions or multi-branched loops, important structural elements where current predictions most often fail (61, 247).

### *Tertiary Structure Prediction and Modelling*

Tertiary structure prediction of RNA is a complex task. The flexibility of the base-sugar and sugar-phosphate linkages is large, and there are seven degrees of freedom per nucleotide (Figure 1). Interactions between bases, phosphates, sugars, and solvent add even more complexity. The goal of these techniques is to produce models consistent with the X-ray, NMR, chemical modification, or nuclease data, with help from structural libraries and chemical information (bond lengths, bond angles, nonbonded interactions). Some recent reviews are gathered in Goodfellow (248).

These prediction methods use interactions between atoms expressed as distance-dependent or angle-dependent potential energies. The potentials allow slight variations of standard bond lengths and bond angles. Energies between nonbonded groups are characterized by van der Waals and electrostatic interactions; hydrogen bonding groups are treated specially. Equilibrium constants (or free energy—see Equation 2) are calculated from the potential energies of different conformations. Most structure calculations omit counterions and solvent, since each added molecule increases the calculation time. The basic calculation methods using these potentials are molecular dynamics (249, 250), where molecular motions are simulated as a function of time, and free energy perturbation methods, which calculate the difference in free energies between two states (250). Although some structures have been generated without any experimental data (251), tertiary-structure modeling clearly relies on experimental data to ensure a particular conformation.

If there are a large number of constraints (for example high-resolution NMR or X-ray crystallography), the global minimum energy structure is often found with simulated annealing, a type of molecular dynamics (252, 253). A model of the structure is minimized using a combined energy function of atomic potentials and artificial forces for the restraints. In general the experimental constraints are weighted much less than the atomic potentials; this ensures a structure that agrees with known bond lengths, angles, and steric constraints (253). A special procedure is used to move the starting structure to the global minimum, since minimization techniques do not alter a structure too much. Elevated temperatures are simulated to push the structures out of local minima, and subsequent cooling moves the structure toward the global minimum. The actual simulated temperatures and time vary from molecule to molecule. A variation on this method, used for certain NMR data, begins with a randomized structure (253); this ensures that the structure is defined by the constraints, not the starting configuration. A variety of programs are available, including X-PLOR (254), Discover (255), and AMBER (256).

The more qualitative information of chemical and enzymatic probes, phylogeny, and functional group substitution, or the limited data of crosslinks or low-resolution NMR, need more help to arrive at a structure. A starting conformation is generated, usually manually, that agrees with known RNA structures and the experimental data, and this structure is energy minimized. The crucial step is the initial structure generation that defines most of the global and local interactions. Many methods use computer-aided structure design, where structural elements from known structures are combined to produce a folded RNA molecule. It is important to note that the model obtained is not necessarily unique—there may be several conformations that agree with the data. Westhof used this method to create structures compatible with chemical modification data (257), and has modelled tRNA (107), 5S rRNA

(258), and the core of group I introns (29). Two systems that position helical elements instead of atoms were described by Malhotra et al (259) and Hubbard & Hearst (116). Recently, a program that automates the structure-building procedure was developed by Major et al (260). This method generates structure using a unique search procedure, and quickly produces a family of structures, all consistent with the available distance and torsion angle data. Using only stacking information and a single distance constraint, the calculated minimum-energy tRNA<sup>Phe</sup> anticodon loop structure agreed the best with the crystal structure. Similarly, the program regenerated a pseudoknot model with limited NMR data in about an hour, as compared to several person-weeks of hand building.

## CONCLUSION

Knowledge of RNA structure can be useful at many levels of resolution. If an RNA sequence is known from a DNA sequence or for a proposed model oligonucleotide, then a secondary structure algorithm is very helpful. The predicted structure can be used to modify the proposed oligonucleotide sequence, or to suggest experiments to test the natural sequence. If many natural sequences for the same RNAs from different organisms are known, phylogenetic methods can reveal common secondary structures. Once an RNA molecule is in hand, chemical and enzymatic probes, crosslinking, and optical and 1D-NMR techniques can test and improve the predicted structures. These methods require only small amounts of material. At this stage of analysis, base pairing, loops, bulges, and junctions are established, and the general arrangement of the parts are evident. The high-resolution techniques of multidimensional NMR and X-ray crystallography are needed to provide a complete three-dimensional structure at atomic resolution. If a crystal is obtained that diffracts X-rays to high resolution, determination of a structure is straightforward. For NMR, the quality of the structure depends greatly on the size of the RNA. For oligonucleotides of less than 20 nucleotides, atomic resolution structures can be obtained if the structure is rigid. Neither crystallography nor NMR can completely define flexible regions. For larger molecules, more powerful 3D-NMR methods using isotope-labeled RNA are being developed.

## ACKNOWLEDGMENTS

We would like to thank Dr. J. V. Hines, Dr. K. J. Luebke, Dr. G. Varani, and Ms. Z. Cai for their critical comments and suggestions. This work was supported in part by DOE (DE-FG03-86ER60406) and NIH (GM 10840) grants to I. T., and NIH postdoctoral grants to J. A. J. (GM 14468-02) and J. S. (GM 14682-01).

## Literature Cited

1. Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J., Weiner, A. M. 1987. *Molecular Biology of the Gene*, pp. 1103–39. Menlo Park, Calif: Cummings. 4th ed.
2. Inouye, M., Delilhas, N. 1988. *Cell* 53:5–7
3. Guerrier-Takada, C., Altman, S. 1984. *Science* 223:285–86
4. Noller, H. F., Hoffarth, V., Zimniak, L. 1992. *Science* 256:1416–19
5. Piccirilli, J. A., McConnell, T. S., Zaig, A. J., Noller, H. F., Cech, T. R. 1992. *Science* 256:1420–24
6. Cech, T. R. 1987. *Science* 236:1532–39
7. Foster, A. C., Symons, R. H. 1987. *Cell* 49:211–20
8. Puglisi, J. D. 1989. *RNA folding: Structure and conformational equilibria of RNA pseudoknots*. PhD thesis. Univ. Calif., Berkeley
9. Saenger, W. 1984. *Principles of Nucleic Acid Structure*, Chapter 10. New York: Springer-Verlag. 2nd corrected printing
10. Olson, W. K., Sussman, J. L. 1982. *J. Am. Chem. Soc.* 104:270–78
11. Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., et al. 1979. *Nature* 282:680–86
12. Hall, K., Cruz, P., Tinoco, I. Jr., Jovin, T. M., van de Sande, J. H. 1984. *Nature* 311:584–86
13. Chastain, M., Tinoco, I. Jr. 1991. *Prog. Nucleic Acids Res. Mol. Biol.* 41:131–77
14. Kim, S.-H., Quigley, G. J., Suddath, F. L., McPherson, A., Sneden, D., et al. 1973. *Science* 179:285–88
15. Pleij, C. W. A. 1990. *Trends Biotechnol.* 15:143–47
16. James, B. D., Olsen, G. J., Pace, N. R. 1989. *Methods Enzymol.* 180:227–39
17. Madison, J. T., Everett, G. A., Kung, H. 1966. *Science* 153:531–34
18. Fox, G., Woese, C. R. 1975. *Nature* 256:505–7
19. Woese, C. R., Magrum, L. J., Gupta, R., Siegel, R. B., Stahl, D. A., et al. 1980. *Nucleic Acids Res.* 8:2273–93
20. Noller, H. F., Kop, J., Wheaton, V., Brosius, J., Gutell, R. R., et al. 1981. *Nucleic Acids Res.* 9:6167–89
21. Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A., Scazzocchio, C. 1982. *Nature* 300:719–24
22. Michel, F., Jacquier, A., Dujon, B. 1982. *Biochimie* 64:867–81
23. Siliciano, P. G., Jones, M. H., Guthrie, C. 1987. *Science* 237:1484–87
24. Pace, N. R., Smith, D. K., Olsen, G. J., James, B. D. 1989. *Gene* 82:65–75
25. Waterman, M. S. 1988. *Methods Enzymol.* 164:765–93
26. Waterman, M. S., Jones, R. 1990. *Methods Enzymol.* 183:221–37
27. Genetics Computer Group, Inc., University Research Park, 575 Science Drive, Madison, Wis. 53711
28. IntelliGenetics, 700 E. El Camino Real, Mountain View, Calif. 94040
29. Michel, F., Westhof, E. 1990. *J. Mol. Biol.* 216:585–610
30. Michel, F., Dujon, B. 1983. *EMBO J.* 2:33–38
31. Cummings, D. J., Michel, F., Domenico, J. M., McNally, K. L. 1990. *J. Mol. Biol.* 212:269–86
32. Cummings, D. J., Domenico, J. M. 1988. *J. Mol. Biol.* 204:815–39
33. Woese, C. R., Gutell, R., Gupta, R., Noller, H. F. 1983. *Microbiol. Rev.* 47:621–69
34. Levitt, M. 1969. *Nature* 224:759–63
35. Michel, F., Ellington, A. D., Couture, S., Szostak, J. W. 1990. *Nature* 347:578–80
36. Guerrier-Takada, C., Lumelsky, N., Altman, S. 1989. *Science* 246:1578–84
37. Haas, E. S., Morse, D. P., Brown, J. W., Schmidt, F. J., Pace, N. R. 1991. *Science* 254:853–56
38. Burke, J. M. 1988. *Gene* 73:273–94
39. Beaudry, A. A., Joyce, G. F. 1990. *Biochemistry* 29:6534–39
40. Joyce, G. F. 1989. *Gene* 82:83–87
41. Tuerk, C., Gold, L. 1990. *Science* 249:505–10
42. Ellington, A. D., Szostak, J. W. 1990. *Nature* 346:818–22
43. Beaudry, A. A., Joyce, G. F. 1992. *Science* 257:635–41
44. SantaLucia, J. Jr., Kierzek, R., Turner, D. H. 1991. *J. Am. Chem. Soc.* 113:4313–22
45. Moore, M. J., Sharp, P. A. 1992. *Science* 256:992–97
46. Ruffner, D. E., Uhlenbeck, O. C. 1990. *Nucleic Acids Res.* 18:6025–29
47. Michel, F., Hanna, M., Green, R., Bartel, D. P., Szostak, J. W. 1989. *Nature* 342:391–95
48. Sakata, T., Hiroaki, H., Oda, Y., Tanaka, T., Ikehara, M., et al. 1990. *Nucleic Acids Res.* 18:3831–39
49. SantaLucia, J. Jr., Kierzek, R., Turner, D. H. 1992. *Science* 256:217–19

50. Puglisi, J. D., Tinoco, I. Jr. 1989. *Methods Enzymol.* 180:304-25
51. Tinoco, I. Jr. 1960. *J. Am. Chem. Soc.* 82:4785-90
52. Petersheim, M., Turner, D. H. 1983. *Biochemistry* 22:256-63
53. Tinoco, I. Jr., Sauer, K., Wang, J. C. 1985. *Physical Chemistry, Principles and Applications in Biological Sciences*, pp. 490-93. Englewood Cliffs, NJ: Prentice-Hall. 2nd ed.
54. Marky, L. A., Breslauer, K. J. 1987. *Biopolymers* 26:1601-20
55. Borer, P. N., Dengler, B., Tinoco, I. Jr., Uhlenbeck, O. C. 1974. *J. Mol. Biol.* 86:843-53
56. Stevens, C., Felsenfeld, G. 1964. *Biopolymers* 2:293-314
57. SantaLucia, J. Jr., Kierzek, R., Turner, D. H. 1990. *Biochemistry* 29:8813-19
58. Cole, P. E., Yang, S. K., Crothers, D. M. 1972. *Biochemistry* 11:4358-68
59. Crothers, D. M., Cole, P. E., Hilbers, C. W., Shulman, R. G. 1974. *J. Mol. Biol.* 87:63-88
60. Felsenfeld, G., Hirschman, S. Z. 1965. *J. Mol. Biol.* 13:407-27
61. Jaeger, J. A., Zuker, M., Turner, D. H. 1990. *Biochemistry* 29:10147-58
- 61a. Banerjee, A. R., Jaeger, J. A., Turner, D. H. 1993. *Biochemistry* 32:153-63
62. Riazance, J. H., Baase, W. A., Johnson, W. C. Jr., Hall, K., Cruz, P., et al. 1985. *Nucleic Acids Res.* 13:4983-89
63. Gray, D. M., Liu, J.-J., Ratliff, R. L., Allen, F. S. 1981. *Biopolymers* 20:1337-82
64. Johnson, K. H., Gray, D. M. 1991. *Biopolymers* 31:385-95
65. Johnson, K. H., Gray, D. M. 1992. *J. Biomol. Struct. Dyn.* 9:733-45
66. Cantor, C. R., Schimmel, P. R. 1980. *Biophysical Chemistry, Part II: Techniques for the Study of Biological Structure and Function*, pp. 425-28. New York, NY: Freeman
67. Daly, T. J., Rusche, J. R., Maione, T. E., Frankel, A. D. 1990. *Biochemistry* 29:9791-95
68. Cantor, C. R., Schimmel, P. R. 1980. See Ref. 66, pp. 433-65
69. Leonard, N. J. 1984. *CRC Crit. Rev. Biochem.* 15:125-99
70. Odom, O. W., Deng, H. Y., Hardesty, B. 1988. *Methods Enzymol.* 164:174-87
71. Fairclough, R., Cantor, C. R. 1977. *Methods Enzymol.* 48:347-79
72. Bevilacqua, P. C., Kierzek, R., Turner, D. H. 1992. *Science* 258:1355-57
73. Weaver, A. J., Kemple, M. D., Prendergast, F. G. 1989. *Biochemistry* 28:8624-39
74. Picking, W. L., Picking, W. D., Ma, C., Hardesty, B. 1991. *Nucleic Acids Res.* 19:5749-54
75. Stryer, L., Haugland, R. P. 1967. *Proc. Natl. Acad. Sci. USA* 58:719-26
76. Cantor, C. R., Schimmel, P. R. 1980. *Biophysical Chemistry, Part III: The Behavior of Biological Macromolecules*, pp. 425-28. New York, NY: Freeman
77. Czworkowski, J., Odom, O. W., Hardesty, B. 1991. *Biochemistry* 30:4821-30
78. Knapp, G. 1989. *Methods Enzymol.* 180:192-212
79. Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.-P., et al. 1987. *Nucleic Acids Res.* 15:9109-28
80. Silberklang, M., Prochiantz, A., Haenni, A.-L., Rajbhandary, U. L. 1977. *Eur. J. Biochem.* 72:465-78
81. Donis-Keller, H. 1980. *Nucleic Acids Res.* 8:3133-42
82. Meador, J., Cannon, B., Cannistraro, V. J., Kennell, D. 1990. *Eur. J. Biochem.* 187:549-53
83. Sambrook, J., Fritsch, E. F., Maniatis, T. 1989. *Molecular Cloning, A Laboratory Manual*, p. 5.78. Cold Spring Harbor, NY: Cold Spring Harbor Lab. Press
84. Laskowski, M. 1980. *Methods Enzymol.* 65:263-69
85. Sulkowski, E., Laskowski, M. 1969. *J. Biol. Chem.* 244:3818-22
86. Auron, P. E., Weber, L. D., Rich, A. 1982. *Biochemistry* 21:4700-47
87. Stern, S., Moazed, D., Noller, H. F. 1988. *Methods Enzymol.* 164:481-89
88. Nielsen, P. T. 1990. *J. Mol. Recognit.* 3:1-25
89. Sigman, D. S., Chen, C.-h. B. 1990. *Annu. Rev. Biochem.* 59:207-36
90. Singer, B., Kusmierek, J. T. 1982. *Annu. Rev. Biochem.* 51:655-93
91. Vlassov, V. V., Giegé, R., Ebel, J.-P. 1981. *Eur. J. Biochem.* 119:51-59
92. Ehrenberg, L., Fedorcsak, I., Solymosy, F. 1976. *Prog. Nucleic Acid Res. Mol. Biol.* 16:189-262
93. Moazed, D., Stern, S., Noller, H. F. 1986. *J. Mol. Biol.* 187:399-416
94. Lilley, D. M. J. 1983. *Nucleic Acids Res.* 11:3097-12
95. Lilley, D. M. J., Palecek, E. 1984. *EMBO J.* 3:1187-92
96. Rubin, C. M., Schmid, C. 1980. *Nucleic Acids Res.* 8:4613-19
97. Murakawa, G. J., Chen, C.-h. B., Kuwabara, M. D., Nierlich, D., Sig-

## 284 RNA STRUCTURE

- man, D. S. 1989. *Nucleic Acids Res.* 17:5361-69
98. Hertzberg, R. P., Dervan, P. B. 1982. *J. Am. Chem. Soc.* 104:313-15
99. Latham, J. A., Cech, T. R. 1989. *Science* 245:276-85
100. Kean, J. M., White, S. A., Draper, D. E. 1985. *Biochemistry* 24:5062-70
101. Kaynor, R., Soultanakis, E., Kuwabara, M., Garcia, J., Sigman, D. S. 1989. *Proc. Natl. Acad. Sci. USA* 86:4858-62
102. Chow, C. S., Behlen, L. S., Uhlenbeck, O. C., Barton, J. K. 1992. *Biochemistry* 31:972-82
103. Chow, C. S., Hartmann, K. M., Rawlings, S. L., Huber, P. W., Barton, J. K. 1992. *Biochemistry* 31:3534-42
104. Mandiyan, V., Boublik, M. 1990. *Nucleic Acids Res.* 18:7055-62
105. Andersen, J., Delihans, N., Hanas, J. S., Wu, C.-W. 1984. *Biochemistry* 23:5752-66
106. Romby, P., Westhof, E., Taukifimpfa, R., Mache, R., Ebel, J.-P., et al. 1988. *Biochemistry* 27:4721-31
107. Dock-Bregeon, A. C., Westhof, E., Giegé, R., Moras, D. 1989. *J. Mol. Biol.* 206:707-22
108. Hall, K. B., Sampson, J. R. 1990. *Nucleic Acids Res.* 18:7041-47
109. Westhof, E., Romby, P., Romaniuk, P. J., Ebel, J.-P., Ehresmann, C., et al. 1989. *J. Mol. Biol.* 207:417-31
- 109a. Wimberly, B., Varani, G., Tinoco, I. Jr. 1993. *Biochemistry* 32:1078-87
110. Wyatt, J. R., Puglisi, J. D., Tinoco, I. Jr. 1990. *J. Mol. Biol.* 214:455-70
111. Celander, D. W., Cech, T. R. 1990. *Biochemistry* 29:1355-61
112. Celander, D. W., Cech, T. R. 1991. *Science* 251:401-7
113. Wang, J.-F., Cech, T. R. 1992. *Science* 256:526-29
114. Brimacombe, R., Atmadja, J., Stiege, W., Schuler, D. 1988. *J. Mol. Biol.* 199:115-36
115. Shi, Y., Lipson, S. E., Chi, D. Y., Spielmann, H. P., Monforte, J. A., et al. 1990. *Bioorganic Photochemistry: Photochemistry and the Nucleic Acids*, ed. H. Morrison, pp. 341-78. New York: Wiley
116. Hubbard, J. M., Hearst, J. E. 1991. *J. Mol. Biol.* 221:889-907
117. Downs, W. D., Cech, T. R. 1990. *Biochemistry* 29:5605-13
118. Mitchell, P., Osswald, M., Schueler, D., Brimacombe, R. 1990. *Nucleic Acids Res.* 18:4325-33
119. Karaoglu, D., Thurlow, D. L. 1991. *Nucleic Acids Res.* 19:5293-300
120. Murakawa, G. J., Nierlich, D. P. 1989. *Biochemistry* 28:8067-72
121. Wang, X., Padgett, R. A. 1989. *Proc. Natl. Acad. Sci. USA* 86:7795-99
122. Darsillo, P., Huber, P. W. 1991. *J. Biol. Chem.* 266:21075-82
123. Dock-Bregeon, A.-C., Garcia, A., Giegé, R., Moras, D. 1990. *Eur. J. Biochem.* 188:283-90
124. Calnan, B. J., Tidor, B., Biancalana, S., Hudson, D., Frankel, A. D. 1991. *Science* 252:1167-71
125. Rice, G. A., Kane, C. M., Chamberlin, M. J. 1991. *Proc. Natl. Acad. Sci. USA* 88:4245-49
126. Rahman, M. A., Schaup, H. W. 1990. *Biochim. Biophys. Acta* 1087:212-18
127. Fischer, S. G., Lerman, L. S. 1979. *Cell* 16:191-200
128. Nishigaki, K., Miura, T., Tsubota, M., Sutoh, A., Amano, N., et al. 1992. *J. Biochem.* 111:151-56
129. Riesner, D., Henco, K., Steger, G. 1991. *Adv. Electrophor.* 4:169-250
130. Henderson, E., Hardin, C. C., Wolk, S. K., Tinoco, I. Jr., Blackburn, E. H. 1987. *Cell* 51:899-908
131. Konarska, M. M. 1989. *Methods Enzymol.* 180:442-53
132. Pyle, A. M., McSwiggen, J. A., Cech, T. R. 1990. *Proc. Natl. Acad. Sci. USA* 87:8187-91
- 132a. Bevilacqua, P. C., Turner, D. H. 1991. *Biochemistry* 30:10632-40
133. Wüthrich, K. 1986. *NMR of Proteins and Nucleic Acids*, pp. 44-116. New York: Wiley
134. Wüthrich, K. 1986. See Ref. 133, pp. 224-32
135. Varani, G., Cheong, C., Tinoco, I. Jr. 1991. *Biochemistry* 30:3280-89
136. Varani, G., Wimberly, B., Tinoco, I. Jr. 1989. *Biochemistry* 28:7760-72
137. Varani, G., Tinoco, I. Jr. 1991. *Q. Rev. Biophys.* 24:479-532
138. Heus, H. A., Pardi, A. 1991. *Science* 253:191-94
139. White, S. A., Nilges, M., Huang, A., Brünger, A. T., Moore, P. B. 1992. *Biochemistry* 31:1610-21
140. Puglisi, J. D., Wyatt, J. R., Tinoco, I. Jr. 1990. *J. Mol. Biol.* 214:437-53
141. Puglisi, J. D., Wyatt, J. R., Tinoco, I. Jr. 1990. *Biochemistry* 29:4215-26
142. Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D., Williamson, J. R. 1992. *Science* 257:76-80
143. Cheong, C., Moore, P. B. 1992. *Biochemistry* 31:8406-14
- 143a. Chastain, M., Tinoco, I., Jr. 1992. *Biochemistry* 31:12733-41
144. Derome, A. E. 1987. *Modern NMR*

- Techniques for Chemistry Research*. New York: Pergamon
145. Leroy, J. L., Bolo, N., Figueroa, N., Plateau, P., Guéron, M. 1985. *J. Biomol. Struct. Dyn.* 2:915-39
146. Wüthrich, K. 1986. See Ref. 133, pp. 93-116
147. Wüthrich, K. 1986. See Ref. 133, pp. 44-92
148. Neuhaus, D., Williamson, M. 1989. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*. New York: VCH
149. Borgias, B. A., James, T. L. 1989. *Methods Enzymol.* 176:169-83
150. Koning, T. M. G., Boelens, R., Kaptain, R. 1990. *J. Magn. Reson.* 90:111-23
151. Müller, N., Ernst, R. R., Wüthrich, K. 1986. *J. Am. Chem. Soc.* 108:6482-92
152. Karplus, M. 1959. *J. Chem. Phys.* 30:11-15
153. Altona, C. 1982. *Recl. Trav. Chim. Pays-Bas* 101:413-33
154. Bax, A. 1989. *Methods Enzymol.* 176:151-68
155. Rance, M., Chazin, W. J., Dalvit, C., Wright, P. E. 1989. *Methods Enzymol.* 176:114-34
156. Griesinger, C., Sorensen, O. W., Ernst, R. R. 1986. *J. Chem. Phys.* 85:6837-52
157. Fesik, S. W., Zuiderweg, E. R. P. 1990. *Q. Rev. Biophys.* 23:97-131
158. Clore, G. M., Gronenborn, A. M. 1991. *Science* 252:1390-99
159. Powers, R., Garrett, D. S., March, C. J., Frieden, E. A., Gronenborn, A. M., et al. 1992. *Science* 256:1673-77
160. McIntosh, L. P., Dahlquist, F. W. 1990. *Q. Rev. Biophys.* 23:1-38
161. Montelione, G. T., Winkler, M. E., Rauenbuehler, P., Wagner, G. 1989. *J. Magn. Reson.* 82:198-204
162. Edison, A. S., Westler, W. M., Markley, J. L. 1991. *J. Magn. Reson.* 92:434-38
163. Wider, G., Neri, D., Otting, G., Wüthrich, K. 1989. *J. Magn. Reson.* 85:426-31
164. Dellwo, M. J., Wand, A. J. 1989. *J. Am. Chem. Soc.* 111:4571-78
165. Kay, L. E., Torchia, D. A., Bax, A. 1989. *Biochemistry* 28:8972-79
166. Clore, G. M., Driscoll, P. C., Wingfield, P. T., Gronenborn, A. M. 1990. *Biochemistry* 29:7387-401
167. Peng, J. W., Thanabal, V., Wagner, G. 1991. *J. Magn. Reson.* 94:82-100
168. Nikonowicz, E. P., Pardi, A. 1992. *Nature* 355:184-86
169. Nikonowicz, E. P., Pardi, A. 1992. *J. Am. Chem. Soc.* 114:1082-83
170. Otting, G., Wüthrich, K. 1990. *Q. Rev. Biophys.* 23:39-96
171. Rinkel, L. J., Altona, C. 1987. *J. Biomol. Struct. Dyn.* 4:621-49
172. Lankhorst, P. P., Haasnoot, C. A. G., Erkelens, C., Altona, C. 1984. *J. Biomol. Struct. Dyn.* 1:1387-405
173. Kline, P. C., Serianni, A. S. 1990. *J. Am. Chem. Soc.* 112:7373-81
174. Williamson, J. R., Boxer, S. G. 1989. *Biochemistry* 28:2819-31
175. Williamson, J. R., Boxer, S. G. 1988. *Nucleic Acids Res.* 16:1529-40
176. Eimer, W., Williamson, J. R., Boxer, S. G., Pecora, R. 1990. *Biochemistry* 29:799-811
177. Griffey, R. H., Redfield, A. G. 1987. *Q. Rev. Biophys.* 19:51-82
178. Sorensen, O. W., Ernst, R. R. 1983. *J. Magn. Res.* 51:477-89
179. Varani, G., Tinoco, I. Jr. 1991. *J. Am. Chem. Soc.* 113:9349-54
180. Batey, R. T., Inada, M., Kujawinski, E., Puglisi, J. D., Williamson, J. R. 1992. *Nucleic Acids Res.* 20:4515-23
181. Nikonowicz, E. P., Sirt, A., Legault, P., Jucker, F. M., Baer, L. M., et al. 1992. *Nucleic Acids Res.* 20:4507-13
182. Simon, E. S., Grabowski, S., Whitesides, G. M. 1989. *J. Am. Chem. Soc.* 111:8920-21
183. Simon, E. S., Grabowski, S., Whitesides, G. M. 1990. *J. Org. Chem.* 55:1834-41
184. Milligan, J. F., Groebe, D. R., Witherell, G. W., Uhlenbeck, O. C. 1987. *Nucleic Acids Res.* 21:8783-98
185. Usman, N., Ogilvie, K. K., Jiang, M.-V., Cedergren, R. J. 1987. *J. Am. Chem. Soc.* 109:7845-54
186. Chou, S.-H., Flynn, P., Reid, B. R. 1989. *Biochemistry* 28:2422-35
187. Serianni, A. S., Nunez, H. A., Barker, R. 1979. *Carbohydr. Res.* 72:71-78
188. Serianni, A. S., Clark, E. L., Barker, R. 1978. *Carbohydr. Res.* 72:79-91
189. Hayes, M. L., Pennings, N. J., Serianni, A. S., Barker, R. 1982. *J. Am. Chem. Soc.* 104:6764-69
190. Serianni, A. S., Cadman, E., Pierce, J., Hayes, M. L., Barker, R. 1982. *Methods Enzymol.* 89:83-92
191. Murray, A. III, Williams, D. C. 1958. *Organic Syntheses with Isotopes*, Part I, pp. 747-67. New York: Interscience
192. Parkanyi, C., Sorm, F. 1963. *Collect. Czech. Chem. Commun.* 28:2491-500
193. Roberts, J. L., Poulter, C. D. 1978. *J. Org. Chem.* 43:1547-50

286 RNA STRUCTURE

194. Vorbruggen, H., Benua, B. 1981. *Chem. Ber.* 114:1279-86
195. Yoshikawa, M., Kato, T., Takenishi, T. 1967. *Tetrahedron Lett.* 50:5065-68
196. Mishra, N. C., Broom, A. D. 1991. *J. Chem. Soc. Chem. Commun.*, pp. 1276-77
197. Cheong, C., Varani, G., Tinoco, I. Jr. 1990. *Nature* 346:680-82
198. Wüthrich, K. 1986. See Ref. 133, pp. 224-32
199. Pardi, A., Nikonowicz, E. P. 1992. *J. Am. Chem. Soc.* 114:1082-3
200. Bax, A., Lerner, L. 1986. *Science* 232:960-67
201. Kellogg, G. W. 1992. *J. Magn. Reson.* 98:176-82
202. Kellogg, G. W., Szewczak, A. A., Moore, P. B. 1992. *J. Am. Chem. Soc.* 114:2727-28
203. Pease, A. C., Wemmer, D. E. 1990. *Biochemistry* 29:9039-46
204. Neuhaus, D., Wagner, G., Vasak, M., Kagi, J. H., Wüthrich, K. 1985. *Eur. J. Biochem.* 151:257-73
205. Sklenar, V., Miyashiro, H., Zon, G., Miles, H. T., Bax, A. 1986. *FEBS Lett.* 208:94-98
- 205a. Schmieder, P., Ippel, J. H., van den Elst, H., van der Mare, G. A., van Boom, J. H., et al. 1992. *Nucleic Acids Res.* 20:4747-51
206. Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Kim, J. J. P., Rich, A. 1976. *J. Mol. Biol.* 104:109-44
207. Rosenberg, J. M., Seeman, N. C., Day, R. O., Rich, A. 1976. *J. Mol. Biol.* 104:145-67
208. Sussman, J. L., Holbrook, S. R., Warrant, R. W., Church, G. M., Kim, S.-H. 1978. *J. Mol. Biol.* 123:607-30
209. Moras, D., Comarmond, M. B., Fischer, J., Weiss, R., Thierry, J. C., et al. 1980. *Nature* 288:669-74
210. Schevitz, R. W., Podjarny, A. D., Krishnamachari, N., Hughes, J. J., Sigler, P. B., et al. 1979. *Nature* 278:188-90
211. Woo, N. H., Roe, B. A., Rich, A. 1980. *Nature* 286:346-51
212. Wright, H. T., Manor, P. C., Beurling, K., Karpel, R. L., Fresco, J. 1979. In *Transfer RNA: Structure, Properties, and Recognition*, ed. P. R. Schimmel, D. Soll, J. N. Abelson, pp. 145-60. Cold Spring Harbor, NY: Cold Spring Harbor Lab. Press
213. Basavappa, R., Sigler, P. B. 1991. *EMBO J.* 10:3105-11
214. Dock-Bregeon, A. C., Dhevriere, B., Podjarny, A., Johnson, J., de Bear, J. S., et al. 1989. *J. Mol. Biol.* 209:459-74
215. Holbrook, S. R., Cheong, C., Tinoco, I. Jr., Kim, S.-H. 1991. *Nature* 353: 579-81
216. Cantor, C. R., Schimmel, P. R. 1980. *Biophysical Chemistry Part I: The Conformation of Biological Macromolecules*, Chapter 3. New York: Freeman
217. Arnott, S., Hukins, D. W. L., Dover, S. D., Fuller, W., Hudgson, A. R. 1973. *J. Mol. Biol.* 81:107-22
218. Arnott, S., Hukins, D. W. L., Dover, S. D. 1972. *Biochim. Biophys. Res. Commun.* 48:1392-99
219. Holbrook, S. R., Sussman, J. L., Warrant, R. W., Kim, S.-H. 1978. *J. Mol. Biol.* 123:631-60
220. Westhof, E., Sundaralingam, M. 1986. *Biochemistry* 25:4868-78
221. Brown, R. S., Hingerty, B. E., Dewan, J. C., Klug, A. 1983. *Nature* 303:543-46
222. Rould, M. A., Perona, J. J., Soll, D., Steitz, T. A. 1989. *Science* 246:1135-42
223. Ruff, M., Krishnaswamy, S., Boeglín, M., Poterszman, A., Mitschler, A., et al. 1991. *Science* 252:1682-89
224. Brünger, A. T. 1990. *Acta Crystallogr. A* 45:46-57
225. Woese, C. R., Winker, S., Gutell, R. R. 1990. *Proc. Natl. Acad. Sci. USA* 87:8467-71
226. Turner, D. H., Sugimoto, N., Freier, S. M. 1988. *Annu. Rev. Biophys. Biophys. Chem.* 17:167-92
227. Brünger, A. T. 1990. See Ref. 248, pp. 137-78
228. Zuker, M. 1989. *Science* 244:48-52
229. Papanicolaou, C., Gouy, M., Ninio, J. 1984. *Nucleic Acids Res.* 12:31-44
230. Martinez, H. M. 1990. *Methods Enzymol.* 183:306-17
231. Abrahams, J. P., van den Berg, M., van Batenburg, E., Pleij, C. 1990. *Nucleic Acids Res.* 18:3035-44
232. Le, S.-Y., Chen, J.-H., Nussinov, R., Maizel, J. V. Jr. 1988. *CABIOS* 4:337-44
233. Konings, D. A. M., Hogeweg, P. 1989. *J. Mol. Biol.* 207:597-614
234. Le, S.-Y., Zuker, M. 1991. *J. Biomol. Struct. Dyn.* 5:1027-44
235. Gouy, M. 1987. In *Nucleic Acid and Protein Sequence Analysis: A Practical Approach*, ed. M. J. Bishop, C. J. Rawlings, pp. 259-84. Oxford: IRL. 417 pp.
236. Zuker, M. 1989. *Methods Enzymol.* 180:262-88
237. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., et al. 1986. *Proc. Natl. Acad. Sci. USA* 83:9373-77

238. He, L., Kierzek, R., SantaLucia, J. Jr., Walter, A. E., Turner, D. H. 1991. *Biochemistry* 30:11124-32
239. Longfellow, C. E., Kierzek, R., Turner, D. H. 1990. *Biochemistry* 29: 278-85
240. Antao, V.P., Lai, S. Y., Tinoco, I. Jr. 1992. *Nucleic Acids Res.* 19: 5901-5
241. SantaLucia, J. Jr., Kierzek, R., Turner, D. H. 1991. *Biochemistry* 30:8242-51
242. Tuerk, C., Gauss, P., Thermes, C., Groebe, D. R., Gayle, M., et al. 1988. *Proc. Natl. Acad. Sci. USA* 85:1364-68
243. Groebe, D. R., Uhlenbeck, O. C. 1988. *Nucleic Acids Res.* 16:11725-35
244. Peritz, A., Kierzek, R., Sujimoto, N., Turner, D. H. 1991. *Biochemistry* 30: 6428-36
245. Gralla, J., Crothers, D. M. 1973. *J. Mol. Biol.* 78:301-19
246. Jaeger, J. A., Turner, D. H., Zuker, M. 1989. *Proc. Natl. Acad. Sci. USA* 86:7706-10
247. Kwakman, J. H., Konings, D. A. M., Hogeweg, P., Pel, H. J., Grivell, L. A. 1990. *J. Biomol. Struct. Dyn.* 8: 413-30
248. Goodfellow, J. M., ed. 1990. *Molecular Dynamics. Applications in Molecular Biology*. Boston, Mass: CRC
249. Williams, M. A., Saqi, M. A. S., Goodfellow, J. M. 1990. See Ref. 248, pp. 179-96
250. Flores, T. P., Moss, D. S. 1990. See Ref. 248, pp. 1-26
251. Mei, H. Y., Kaaret, T. W., Bruice, T. C. 1989. *Proc. Natl. Acad. Sci. USA* 86:9727-31
252. Brünger, A. T. 1990. See Ref. 248, pp. 137-78
253. Clore, G. M., Gronenborn, A. M. 1989. *Crit. Rev. Biochem. Mol. Biol.* 24:479-564
254. Brünger, A. T. 1992. *X-PLOR Manual, Version 3.0*. New Haven, Conn: Yale Univ. 405 pp.
255. Biosym Technologies, 9685 Scranton Rd., San Diego, Calif. 92121
256. Weiner, S. J., Kollman, P. A., Nguyen, D. T., Case, D. A. 1986. *J. Comput. Chem.* 7:230-52
257. Westhof, E., Romby, P., Ehresmann, C., Ehresmann, B. 1990. In *Theoretical Biochemistry and Molecular Biophysics. DNA*, ed. D. L. Beveridge, R. Lavery, 1:399-409. Schenectady, NY: Adenine. 434 pp.
258. Brunel, C., Romby, P., Westhof, E., Ehresmann, C., Ehresmann, B. 1991. *J. Mol. Biol.* 221:293-308
259. Malhotra, A., Tan, R. K.-Z., Harvey, S. C. 1990. *Proc. Natl. Acad. Sci. USA* 87:1950-54
260. Major, F., Turcotte, M., Gautheret, D., Lapalme, G., Fillion, E., et al. 1991. *Science* 253:1255-60
261. Schmitz, U., Kumar, A., James, T. L. 1992. *J. Am. Chem. Soc.* 114: 10654-56



## CONTENTS

FROM BACTERIAL NUTRITION TO ENZYME STRUCTURE: A PERSONAL ODYSSEY, <i>Esmond E. Snell</i>	1
EUKARYOTIC DNA REPLICATION: ANATOMY OF AN ORIGIN, <i>Melvin L. DePamphilis</i>	29
ANALYSIS OF GLYCOPROTEIN-ASSOCIATED OLIGOSACCHARIDES, <i>Raymond A. Dwek, Christopher J. Edge, David J. Harvey, Mark R. Wormald, and Raj B. Parekh</i>	65
PROTEIN TYROSINE PHOSPHATASES, <i>Kevin M. Walton and Jack E. Dixon</i>	101
THE STRUCTURE AND BIOSYNTHESIS OF GLYCOSYL PHOSPHATIDYLINOSITOL PROTEIN ANCHORS, <i>Paul T. Englund</i>	121
STRUCTURAL AND GENETIC ANALYSIS OF PROTEIN STABILITY, <i>Brian W. Matthews</i>	139
GENERAL INITIATION FACTORS FOR RNA POLYMERASE II, <i>Ronald C. Conaway and Joan Weliky Conaway</i>	161
HUMAN GENE THERAPY, <i>Richard A. Morgan and W. French Anderson</i>	191
NUCLEOCYTOPLASMIC TRANSPORT IN THE YEAST <i>Saccharomyces cerevisiae</i> , <i>Mark A. Osborne and Pamela A. Silver</i>	219
DETERMINATION OF RNA STRUCTURE AND THERMODYNAMICS, <i>John A. Jaeger, John SantaLucia, Jr., and Ignacio Tinoco, Jr.</i>	255
hnRNP PROTEINS AND THE BIOGENESIS OF mRNA, <i>Gideon Dreyfuss, Michael J. Matunis, Serafín Piñol-Roma, and Christopher G. Burd</i>	289
MEMBRANE PARTITIONING DURING CELL DIVISION, <i>Graham Warren</i>	323
MOLECULAR CHAPERONE FUNCTIONS OF HEAT-SHOCK PROTEINS, <i>Joseph P. Hendrick and Franz-Ulrich Hartl</i>	349
BIOCHEMISTRY OF MULTIDRUG RESISTANCE MEDIATED BY THE MULTIDRUG TRANSPORTER, <i>Michael M. Gottesman and Ira Pastan</i>	385

CYTOPLASMIC MICROTUBULE-ASSOCIATED MOTORS, <i>R. A. Walker and M. P. Sheetz</i>	429
SIGNALLING BY RECEPTOR TYROSINE KINASES, <i>Wendy J. Fantl, Daniel E. Johnson, and Lewis T. Williams</i>	453
NEW PHOTOLABELING AND CROSSLINKING METHODS, <i>Josef Brunner</i>	483
MEMBRANE-ANCHORED GROWTH FACTORS, <i>Joan Massagué and Atanasio Pandiella</i>	515
STRUCTURE-BASED INHIBITORS OF HIV-1 PROTEASE, <i>Alexander Wlodawer and John W. Erickson</i>	543
INTRONS AS MOBILE GENETIC ELEMENTS, <i>Alan M. Lambowitz and Marlene Belfort</i>	587
THE TUMOR SUPPRESSOR GENES, <i>Arnold J. Levine</i>	623
PATHWAYS OF PROTEIN FOLDING, <i>C. Robert Matthews</i>	653
CONFORMATIONAL COUPLING IN DNA POLYMERASE FIDELITY, <i>Kenneth A. Johnson</i>	685
COGNITION, MECHANISM, AND EVOLUTIONARY RELATIONSHIPS IN AMINOACYL-tRNA SYNTHETASES, <i>Charles W. Carter, Jr.</i>	715
TRANSCRIPTIONAL REGULATION BY cAMP AND ITS RECEPTOR PROTEIN, <i>A. Kolb, S. Busby, H. Buc, S. Garges, and S. Adhya</i>	749
OXIDATION OF FREE AMINO ACIDS AND AMINO ACID RESIDUES IN PROTEINS BY RADIOLYSIS AND BY METAL-CATALYZED REACTIONS, <i>E. R. Stadtman</i>	797
THE RECEPTORS FOR NERVE GROWTH FACTOR AND OTHER NEUROTROPHINS, <i>Simona Raffioni, Ralph A. Bradshaw, and Stephen E. Buxser</i>	823
FUNCTION AND REGULATION OF RAS, <i>Douglas R. Lowy and Berthe M. Willumsen</i>	851
CONTROL OF TRANSCRIPTION TERMINATION BY RNA-BINDING PROTEINS, <i>Asis Das</i>	893
INDEXES	
Author Index	931
Subject Index	991
Cumulative Index of Contributing Authors, Volumes 58–62	1025
Cumulative Index of Chapter Titles, Volumes 58–62	1029