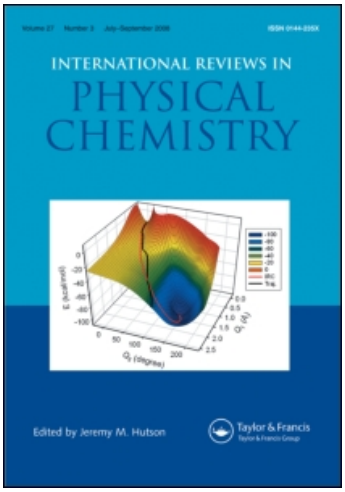


International Reviews in Physical Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713724383

Structures and functions of proteins and nucleic acids in protein biosynthesis

Tatsuo Miyazawa[a]; Shigeyuki Yokoyama[b]
[a] Faculty of Engineering, Yokohama National University, Hodogaya-ku, Yokohama, Japan [b] Faculty of Science, University of Tokyo, Bunkyo-ku, Tokyo, Japan

To cite this Article Miyazawa, Tatsuo and Yokoyama, Shigeyuki(1989) 'Structures and functions of proteins and nucleic acids in protein biosynthesis', International Reviews in Physical Chemistry, 8: 2, 125 — 145
To link to this Article: DOI: 10.1080/01442358909353226
URL: http://dx.doi.org/10.1080/01442358909353226

INTERNATIONAL REVIEWS IN PHYSICAL CHEMISTRY, 1989, VOL. 8, NOS. 2/3, 125–145

Structures and functions of proteins and nucleic acids in protein biosynthesis

by TATSUO MIYAZAWA†‡ and SHIGEYUKI YOKOYAMA§

† Faculty of Engineering, Yokohama National University, Hodogaya-ku, Yokohama 240, Japan

§ Faculty of Science, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Infrared and Raman spectroscopy is useful for studying helical conformations of polypeptides, which are determined by molecular structure parameters. Nuclear magnetic resonance spectroscopy, as well as X-ray analysis, is now established to be important for conformation studies of proteins and nucleic acids in solution. This article is mainly concerned with the conformational aspect and function regulation in protein biosynthesis. The strict recognition of transfer ribonucleic acid (tRNA) by aminoacyl-tRNA synthetase (ARS) is achieved by multi-step mutual adaptation. The conformations of ARS-bound amino acids have been elucidated by transferred nuclear Overhauser effect analysis. Aminoacyl-tRNA takes the 3′-isomeric form in the polypeptide chain elongation cycle. The regulation of codon recognition by post-transcriptional modification is achieved by conversion of the conformational characteristic of the anticodon of tRNA. The cytidine→lysidine modification of the anticodon of minor isoleucine tRNA concurrently converts the amino acid specificity and the codon specificity. As novel protein engineering, a basic strategy has been established for *in vivo* biosynthesis of proteins that are substituted with unnatural amino acids (alloproteins).

1. Introduction

Proteins and nucleic acids are essential cell components for any living organism. The functions of proteins and nucleic acids are recognized to be related to the molecular structures (conformations). In the late 1940s, Professor Mizushima of the University of Tokyo, who was already known for his outstanding contribution in the field of molecular structure and internal rotation (Mizushima 1954), emphasized the importance of studying internal rotation in proteins. Thus, the molecular structure of N-methylacetamide, a simple model of the peptide group in proteins, was studied by a variety of physicochemical methods and the molecule was found to be in the planar *trans* form (Mizushima *et al.* 1950a). This was succeeded by a series of infrared absorption studies on hydrogen bonded structures of model molecules of proteins that have two peptide bonds, including N-acetylamino acid methylamide (Mizushima *et al.* 1950b, 1951, 1952a, b, c, 1954). In solution, a folded structure (B form) with an intramolecular hydrogen bond was found together with an extended form. This B form (now called γ-turn form) was later found in some physiologically active peptides including enkephalin (Higashijima *et al.* 1988) rather than in polypeptides or proteins.

The correlation between infrared spectra and conformations of synthetic polypeptides was found in the early 1950s; the frequencies and dichroism of the amide I band (about 1650 cm^{-1}) and amide II band (around 1550 cm^{-1}) of the α (folded) form

‡ T. M. wishes to dedicate this article to the memory of his professor, Dr San-ichiro Mizushima.

0144–235X/89 $3·00

were different from those of the β (extended) form (Elliott and Ambrose 1950, Elliott 1953). The nature of characteristic amide bands, including amide I and II, was elucidated (Miyazawa *et al.* 1956, 1958, 1960a). This provided a basis for analysing the correlation between amide bands and conformations of polypeptides. The amide I and II bands of polypeptides in helical conformations are split because of intramolecular (and intermolecular) interactions among adjacent peptide groups (Miyazawa 1960b). The frequencies (and dichroism) of amide I, II and V (about 650 cm^{-1}) bands were established (Miyazawa and Blout 1961, Miyazawa *et al.* 1962) for the α-helical form (Pauling *et al.* 1951), the parallel chain β form and the antiparallel chain β form (Pauling and Corey 1951). Characteristic infrared bands and helical conformations of polypeptides and proteins were reviewed in detail (Miyazawa 1962, 1963, Schellman and Schellman 1964).

Helical forms of polymers are determined by the structure of repeating unit that constitutes the polymer chain. Matrix equations for helical parameters (translation d along the axis and rotation θ about the axis per repeating unit) and molecular parameters (bond lengths r, bond angles ϕ, and dihedral angles τ) were obtained (Shimanouchi and Mizushima 1955). Later, new parameters $d\sin(\theta/2)$ and $\cos(\theta/2)$ were explicitly expressed as functions of r, $\phi/2$, and $\tau/2$ (Miyazawa 1961). Numerical equations for a polypeptide main chain $[-NH-CHR-C'O-]_p$ were obtained in terms of dihedral angles τ(N–C) and τ(C–C′) by the use of the standard values of bond lengths and bond angles (Pauling and Corey 1953) and the dihedral angle of $\tau(C'-N)=180°$.

$$\begin{aligned} d\sin(\theta/2) = {} & 2{\cdot}967\cos[\tau(N-C)/2+\tau(C-C')/2] \\ & -0{\cdot}664\cos[\tau(N-C)/2-\tau(C-C')/2], \\ \cos(\theta/2) = {} & -0{\cdot}817\sin[\tau(N-C)/2+\tau(C-C')/2] \\ & +0{\cdot}045\sin[\tau(N-C)/2-\tau(C-C')/2]. \end{aligned}$$

These equations are useful for treating helical conformations of polypeptides. For example, for the α-helical form ($d=1{\cdot}5$ Å and $\theta=100°$), dihedral angles are obtained as $\tau(N-C)=-58°$ and $\tau(C-C')=-47°$. Further a general method was worked out for computation of all the helical parameters from molecular parameters (Sugeta and Miyazawa 1967, 1968).

Later, Raman spectroscopy of polypeptides and proteins became practical by the use of gas laser as excitation source. In 1970, the Raman spectra of lysozyme (Lord and Yu 1970) and ribonuclease (Lord and Yu 1971) were reported, and were followed by a number of reports on Raman spectra of proteins. Raman lines due to the amide III vibration (about 1250 cm^{-1}) (Miyazawa *et al.* 1956, 1958) as well as the amide I vibration are useful for studying helical conformations of polypeptides (Frushour and Koenig 1975).

Further, in contrast with infrared absorption spectra, strong bands due to amino acid side chains are observed in Raman spectra of proteins and polypeptides, which are useful for elucidating hydrogen bond formation and/or conformation of side chain groups. For example, the relative intensities of the Raman bands of tyrosine residue at 850 cm^{-1} and 830 cm^{-1} depend on hydrogen bonding of the tyrosine residue (Yu *et al.* 1972, Siamwiza *et al.* 1975). Raman bands due to S–S stretching vibrations of cystine groups in proteins depend on the conformation of $-CH-CH_2-S-S-CH_2-CH-$; Raman bands are observed at about 510 cm^{-1}, 525 cm^{-1} and 540 cm^{-1} for the *gauche–gauche–gauche*, *gauche–gauche–trans* and *trans–gauche–trans* forms respectively

(Sugeta *et al.* 1972, 1973, Van Wart and Scheraga 1986). This assignment was used for studying the conformations of erabutoxin b (Harada *et al.* 1976), α-lactalbumin (Nakanishi *et al.* 1974) and snake toxins (Tu *et al.* 1976).

Thus vibrational spectroscopy (infrared absorption and Raman scattering) is useful for studying helical conformations of polypeptides (helix content in proteins) and side chain structure (or hydrogen bonding) of some amino acid residues. However, three-dimensional structures of proteins can only be determined by X-ray crystal analyses. The molecular structure of myoglobin was elucidated for the first time by this technique (Kendrew *et al.* 1958, 1960, 1961, Kendrew 1963). Since then, X-ray analyses have been performed of a number of proteins, providing important information on molecular structures and the molecular mechanism of protein functions.

Nevertheless, there are still a number of proteins which have not been crystallized in a form suitable for X-ray analyses. Further, for proteins with high molecular weights, the low resolution allows tracing of the polypeptide main chain only. In particular, for flexible regions in large proteins, interpretation of an electron density map (model building) is difficult. By contrast, a spectroscopic approach provides some detailed information on local structures, however flexible, in proteins.

For proteins and nucleic acids, there is a variety of spectroscopic methods applicable, including ultraviolet absorption, circular dichroism (CD), fluorescence, Raman scattering, infrared absorption, electron spin resonance, and nuclear magnetic resonance (NMR). Among others, NMR spectroscopy has become the most important method for studying conformations and functions of proteins and nucleic acids in solution (Wüthrich 1976, Jardetzky and Roberts 1981).

In particular, two-dimensional NMR spectroscopy (2D-NMR) has been extensively developed (Ernst *et al.* 1987), which allows determination of conformations of small proteins in solution (Wüthrich 1989a, b). Further, proteins may now be labelled with stable magnetic isotopes (^{2}H, ^{13}C and ^{15}N), which allows unambiguous assignment of NMR signals of proteins with higher molecular weight, without referring to X-ray results. In fact, ^{13}C resonances of nearly all the carbonyl groups in *Streptomyces* subtilisin inhibitor were assigned by dual labelling with ^{13}C and ^{15}N (Kainosho and Tsuji 1982, Kainosho *et al.* 1985, 1987), which formed a basis for the complete assignment of proton resonances now in progress. Once (nearly) complete assignment of proton resonances is achieved, the conformation of the protein in solution may be determined by two-dimensional analysis of nuclear Overhauser effects (NOE) (Noggle and Shirmer 1971) in conjunction with computer simulation of the conformation (Wüthrich 1989a, b). Thus NOE analysis of proteins in solution is now recognized to be as important as X-ray crystal analysis. NOE analysis can also provide detailed information on local structures in larger proteins, which is useful for the interpretation of electron density maps of protein crystals.

In recent years, a large number of papers on NMR analyses of proteins and nucleic acids have been published in Japan, which cannot be covered in a short review article. Therefore, in the present article, some recent progress in physicochemical and biochemical studies on the structures and functions of proteins and nucleic acid *in protein biosynthesis* will be reviewed.

2. Protein biosynthesis

In protein biosynthesis (figure 1), the nucleoside sequence of messenger ribonucleic acid (mRNA) is translated into the amino acid sequence of a protein (Watson 1976). A ternary combination of nucleosides uridine (U), cytidine (C), adenosine (A), and

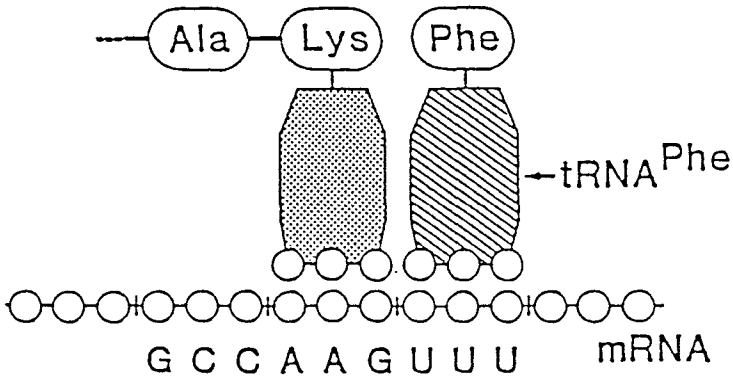


Figure 1. Translation of nucleotide sequence of mRNA into amino acid sequence of protein (schematic).

guanosine (G)) forms a codon of an amino acid (see the table). For example, AAG is a codon for lysine (Lys) and UUU is a codon for phenylalanine (Phe). Codon UUU is recognized by an adaptor, called phenylalanine tRNA. This tRNA^{Phe} is charged with phenylalanine, in advance, by phenylalanyl-tRNA synthetase, one of 20 aminoacyl-tRNA synthetases present in the cell. The tRNA-bound amino acid then accepts a growing polypeptide chain of a protein. Strict recognition of cognate amino acid and tRNA by aminoacyl-tRNA synthetase is essential for correct translation of the nucleotide sequence of messenger RNA into the amino acid sequence of proteins (figure 1).

3. Recognition of amino acid by aminoacyl-tRNA synthetase

The aminoacylation reaction of tRNA by aminoacyl-tRNA synthetase is known to proceed in two steps (Schimmel and Söll 1979). In the first step, an amino acid is

UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC	----	UCC		UAC	----	UGC	----
UUA	Leu	UCA		UAA	term	UGA	term
UUG		UCG		UAG		UGG	Trp
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC		CCC		CAC	----	CGC	
CUA		CCA		CAA	Gln	CGA	
CUG		CCG		CAG		CGG	
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
AUC	----	ACC		AAC	----	AGC	----
AUA	Met	ACA		AAA	Lys	AGA	Arg
AUG		ACG		AAG		AGG	
GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
GUC		GCC		GAC	----	GGC	
GUA		GCA		GAA	Glu	GGA	
GUG		GCG		GAG		GGG	

Genetic codes.

activated to form aminoacyladenosine monophosphate (AA-AMP), and in the second step, the aminoacyl group of AA-AMP is transferred to a cognate tRNA thus forming aminoacyl-tRNA. As for isoleucyl-tRNA synthetase, *L*-valine is also activated to form valyl-AMP, which is hydrolyzed in the second step (proofreading stage) in the presence of isoleucine tRNA (Hopfield *et al.* 1976, Fersht 1977).

Strict recognition of amino acid by aminoacyl-tRNA synthetase has been studied mostly by enzyme kinetic analyses. However, the conformation of an amino acid bound to the active site of the enzyme in solution may now be elucidated by proton NMR analysis, in particular the analysis of transferred nuclear Overhauser effect (TRNOE) (Kohda *et al.* 1987).

The principle of TRNOE is shown in figure 2. Suppose a small ligand is in fast exchange between the enzyme-bound state and the free state in solution. The proton resonances of enzyme-bound ligand are hardly observed, because of slow rotatory motion of the enzyme, while the proton resonances of free ligand are clearly observed. Magnetic saturation (due to irradiation) of a proton (A) in the free state (figure 2 top left) is transferred to the enzyme-bound state, thereby reducing magnetization of another proton (B) in close proximity (figure 2 bottom). The reduced magnetization of proton B is transferred back to the free state (figure 2 top right), thus exhibiting TRNOE for the pair of protons A and B. Systematic analysis of TRNOE for a number of proton pairs of ligand allows determination of the conformation of the enzyme-bound ligand molecule.

As for *E. coli* isoleucyl-tRNA synthetase (IleRS), the conformation of enzyme-bound substrate *L*-isoleucine has been determined by TRNOE analysis (Kohda *et al.* 1987). IleRS-bound *L*-isoleucine is in one form as shown in figure 3 (*a*), although in the free state nine conformers are possible. However, the conformation shown in figure 3 (*a*)

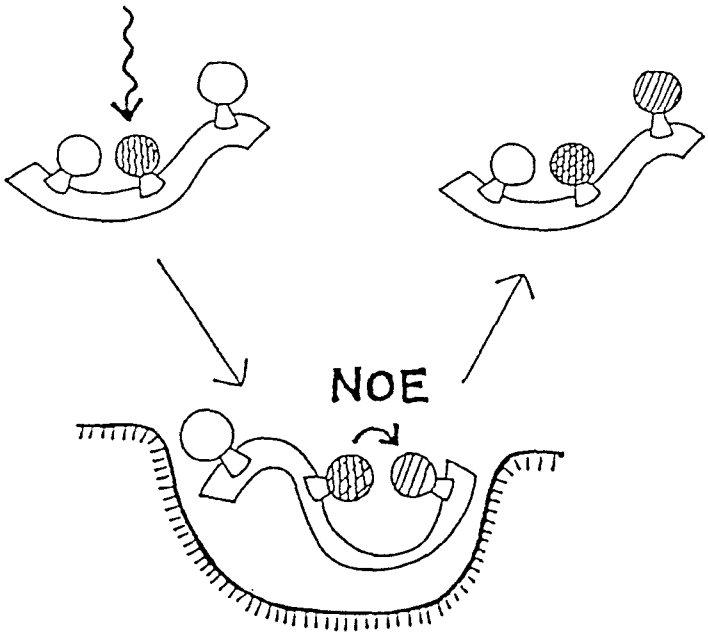
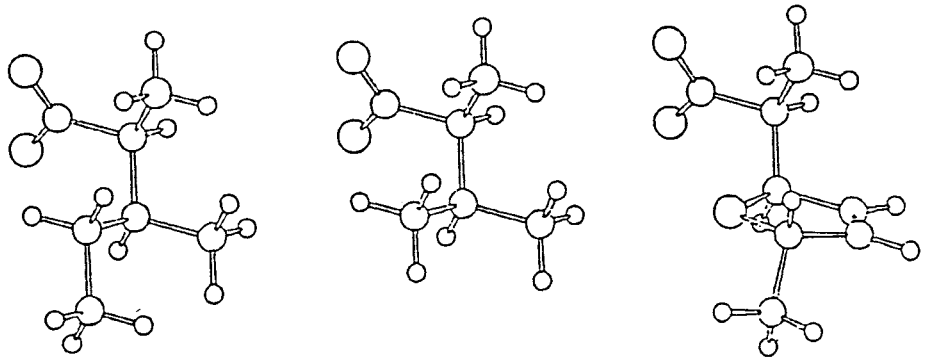


Figure 2. Principle of TRNOE (schematic).



(a) (b) (c)

Figure 3. Conformations of IleRS-bound amino acids, (*a*) *L*-isoleucine, (*b*) *L*-valine, and (*c*) furanomycin.

corresponds to a stable form in the free state. TRNOE analysis has also been made of *L*-valine normally present in the cell. The conformation of IleRS-bound *L*-valine (figure 3(*b*)) is similar to that of *L*-isoleucine, except that the δ-methyl group of *L*-isoleucine is replaced by a hydrogen atom in *L*-valine (Kohda *et al.* 1987). However, binding of *L*-valine with IleRS is much weaker than that of *L*-isoleucine; the δ-methyl group of *L*-isoleucine appears to be important in discrimination of *L*-isoleucine from *L*-valine.

The antibiotic amino acid furanomycin from *Streptomyces* (Nishimura *et al.* 1955) is known to inhibit formation of isoleucyl-tRNA in *E. coli* (Tanaka *et al.* 1969). Unlike *L*-isoleucine, furanomycin has a five-membered heterocyclic ring. However, furanomycin binds tightly with *E. coli* IleRS, in contrast with *L*-valine. The overall conformation of IleRS-bound furanomycin (figure 3(*c*)) is similar to that of *L*-isoleucine (figure 3(*a*)). The pocket of IleRS provided for the δ-methyl group of *L*-isoleucine is probably occupied by the ε-methyl group of furanomycin, significantly stabilizing the complex (Kohno *et al.* 1990). Furthermore, in contrast with *L*-valine, furanomycin is bound to $tRNA^{Ile}$, indicating that furanomycyl-AMP was not hydrolyzed in the proofreading stage in the presence of $tRNA^{Ile}$. Possibly, occupation of a pocket provided for the δ-methyl group of *L*-isoleucine is important for avoiding hydrolysis in the proofreading stage.

4. Recognition of tRNA by aminoacyl-tRNA synthetase

Strict recognition of tRNA species by the cognate aminoacyl-tRNA synthetase is essential for correct translation of the genetic code. The conformations of such tRNA species and aminoacyl-tRNA synthetase have been found to change upon cognate complex formation by the methods of CD (Willick and Kay 1976), *T*-jump (Riesner *et al.* 1976, Krauss *et al.* 1976), fluorescence (Favre *et al.* 1979, Lefevre *et al.* 1980, Ehrlich *et al.* 1980) and nuclease digestion (Yamashiro-Matsumura and Kawata 1981). However, conformation changes are not observed upon formation of complexes of aminoacyl-tRNA synthetases and *noncognate* tRNA species. The mutual adaptation of the cognate tRNA and aminoacyl-tRNA synthetase is probably related to the activation of catalytic site of the enzyme (Renaud *et al.* 1981, Bacha *et al.* 1982).

For a heterologous system of *E. coli* $tRNA^{Glu}$ and *T. thermophilus* GluRS (Hara-Yokoyama *et al.* 1984), the Michaelis constant for *L*-glutamate is increased as KCl concentration is raised. CD and fluorescence analyses indicate that the conformation change of $tRNA^{Glu}$ upon formation of a complex with GluRS is suppressed at higher KCl concentration. *E. coli* $tRNA^{Glu}$ and *T. thermophilus* GluRS form two types of complexes I and II (figure 4). Upon formation of complex I that is stable at high KCl concentration, conformations of GluRS and $tRNA^{Glu}$ are hardly affected. However, at low KCl concentration, complex I is converted to complex II where the conformations of GluRS and $tRNA^{Glu}$ are significantly affected by mutual adaptation and the active site for *L*-glutamate is formed. Selective binding of the substrate *L*-glutamate with GluRS is found only in the presence of $tRNA^{Glu}$ (Hara-Yokoyama *et al.* 1986). Strict recognition of the cognate tRNA by aminoacyl-tRNA synthetase is achieved through multistep formation of the active complex.

5. 2′,3′-Specificity of aminoacyl-tRNA in elongation cycle

Aminoacyl-tRNA (aa-tRNA) is bound with the complex of polypeptide chain elongation factor Tu (EF-Tu) and guanosine triphosphate (GTP) and is involved in the polypeptide chain elongation cycle (figure 5) (Miller and Weissbach 1977, Kaziro 1978). The ternary complex is then brought to the A site of a ribosome, where EF-Tu is released upon hydrolysis of GTP to guanosine diphosphate (GDP). aa-tRNA at the A site is subjected to peptidyl-transfer reaction and forms peptidyl-tRNA (pre-translocational state). This peptidyl-tRNA at the A site is translocated to the P site by the action of the complex of elongation factor G (EF-G) and GTP, thus completing one cycle of polypeptide chain elongation (figure 5). This cycle is as fast as 15–20 amino acid incorporations per second in *E. coli* (Kjeldgaad and Kurland 1963).

In this elongation cycle, aminoacyl-tRNA has been supposed to be subjected to transacylation (conversion between the 2′-isomeric form and the 3′-isomeric form, figure 6), since the transacylation rate has been indirectly estimated to be as fast as $4000\,s^{-1}$ (Griffin *et al.* 1966). Therefore the isomeric form of aminoacyl-tRNA in the elongation cycle has been studied by the use of non-hydrolysable models of aminoacyl-tRNA (Hecht 1979, Sprinzl and Wagner 1979, Wagner *et al.* 1982). In particular,

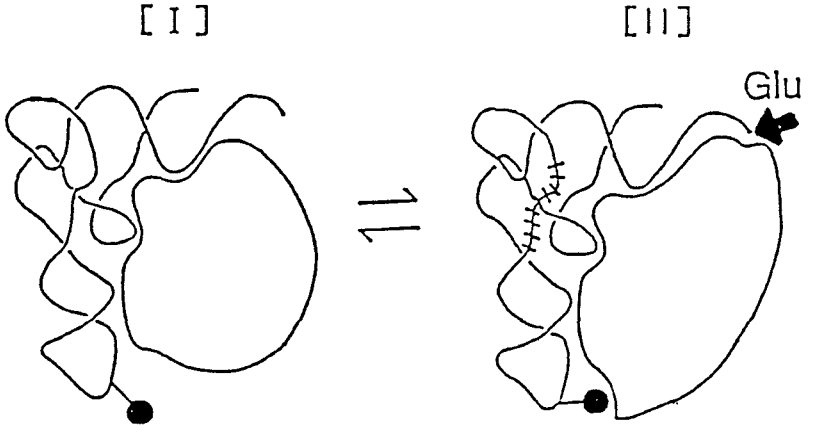


Figure 4. Complex formation of *E. coli* $tRNA^{Glu}$ and *Thermus thermophilus* GluRS. The mnm^5s^2U residue in the anticodon is shown with a closed circle.

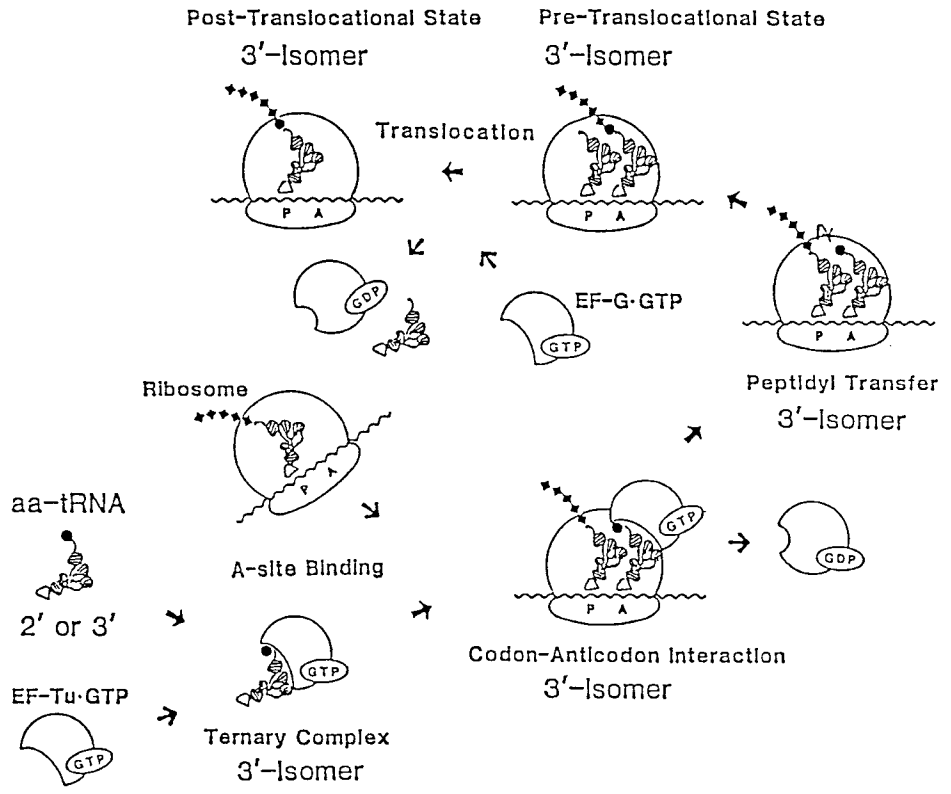


Figure 5. Polypeptide chain elongation cycle (schematic) with 15–20 amino acid incorporations per second.

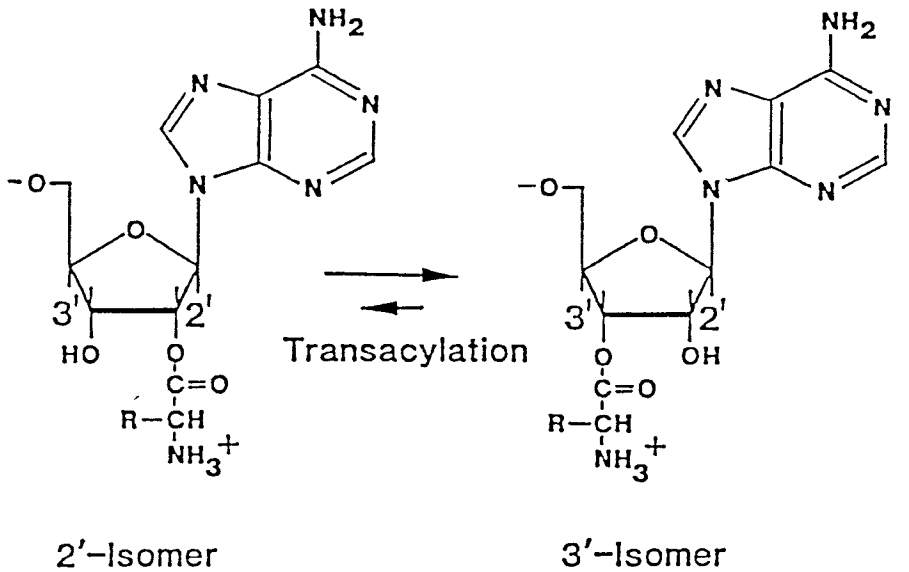


Figure 6. Isomers of aminoacyl-tRNA.

aminoacyl-tRNA has been concluded to be in the 2′-isomeric form in the complex with EF-Tu.GTP (Wagner *et al.* 1982) and then converted into the 3′-isomeric form upon binding to the A site of the ribosome (Sprinzl and Wagner 1979).

Later, however, the abundance ratio of the 2′-isomeric form and the 3′-isomeric form of phenylalanyladenosine has been found to be 0·2 : 0·8 by proton NMR analysis, and further the transacylation rate of phenylalanyladenosine has been directly determined to be slower than $1\ s^{-1}$ from the saturation transfer experiments on proton resonances (Taiji *et al.* 1981, 1983). Note that this transacylation rate is much slower than the rate of polypeptide chain elongation. Actually, however, phenylalanyl-tRNA has been found to be in the 3′-isomeric form in the complex with EF-Tu.GTP (Taiji *et al.* 1985a). Further, the peptidyl transfer reaction of the system involving phenylalanyladenosine (acceptor model) at the A site and N-acetylphenylalanyl-tRNA (donor) at the P site yields as exclusive product the 3′-isomeric form of N-acetylphenylalanylphenylalanyladenosine. Thus aminoacyl-tRNA and peptidyl-tRNA exclusively take the 3′-isomeric form in the polypeptide chain elongation cycle (figure 6), without conversion to the 2′-isomeric form (Taiji *et al.* 1985b). This example clearly demonstrates the importance of NMR analysis in elucidating the dynamic aspect involved in biological function.

6. Codon recognition by modified uridines of anticodon

Codon recognition of tRNA is ascribed to the three base pairs between the codon and the anticodon (figure 7). The third and second letters of anticodon form Watson–Crick base pairs (A.U or G.C) with the first and second letters of the codon

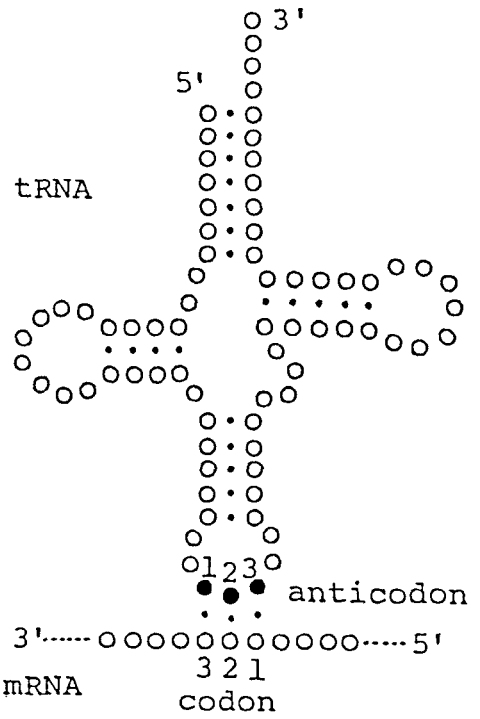


Figure 7. Structure of tRNA (clover leaf model) and recognition of codon of mRNA by anticodon of tRNA (schematic).

respectively. However, certain tRNA species are known to recognize more than one codon. Accordingly, in addition to Watson–Crick base pairs, non-Watson–Crick base pairs are formed between the first letter of the anticodon and the third letter of the codon (Crick 1966). Most of the 20 amino acids are specified by a set of codons that vary in the third letter (see the table).

One of the characteristics of tRNA is that nascent tRNA molecules transcribed from DNA are specifically modified by enzymes. Uridine in the first position of the anticodon (position 34) is modified except for a few cases. As for tRNAs specific to glutamine (Gln), Lys and Glu, U(34) is always modified to 5-methyl-2-thiouridine derivatives (xm^5s^2U, figure 8) (Nishimura 1979). Each of Gln, Lys and Glu is specified by two codons terminating in A or G (table). In the triplet-dependent binding to a ribosome and in the *in vitro* protein synthesis, $xm^5s^2U(34)$ mainly recognizes A as the third letter of the codon and the recognition of G is much less efficient (Sekiya *et al.* 1969, Lustig *et al.* 1981).

On the other hand, in tRNAs specific to valine (Val), serine (Ser), threonine (Thr), proline (Pro) and alanine (Ala), U(34) is modified to a 5-hydroxyuridine derivative (xo^5U, figure 8) (Nishimura 1979). All these amino acids are specified by four codons terminating in U, C, A or G (table). In the triplet-dependent binding to a ribosome (Oda *et al.* 1969, Ishikura *et al.* 1971) and in the *in vitro* protein synthesis (Mitra *et al.* 1979), xo^5U recognizes U in addition to A and G as the third letter of the codon. The molecular mechanism in the regulation of codon recognition by post-transcriptional modification of U(34) has been elucidated on the basis of proton NMR analyses of the conformational characteristics of modified uridines (Yokoyama *et al.* 1985).

7. Conformational characteristics of modified uridines

5-Methylaminomethyl-2-thiouridine 5′-monophosphate (denoted $pmnm^5s^2U$) and 5-methoxyuridine 5′-monophosphate (pmo^5U) are examples of the two types of modified uridine nucleotides pxm^5s^2U and pxo^5U respectively. By analyses of vicinal spin-coupling constants, the fractional populations of the C2′-endo form and the C3′-endo form of the ribose ring are obtained and, from the temperature dependences of fractional populations, the enthalpy differences between the two forms may be determined. The enthalpy difference between the C2′-endo form and the C3′-endo form in unmodified pU is 0·09 kcal mol^{-1}, indicating that the C2′-endo form and C3′-endo form are nearly equally stable without modification of the uracil base (Yokoyama *et al.* 1985).

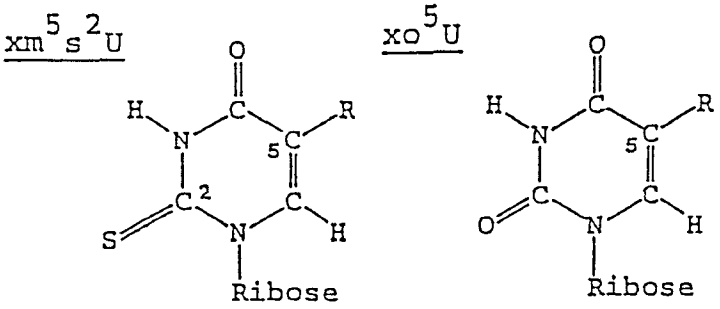


Figure 8. Modified uridines found in the anticodon of tRNA.

On the other hand, for pmnm^{5}s^{2}U, the enthalpy difference between the C2′-endo form and the C3′-endo form is 1·1 kcal mol^{-1}; the C3′-endo form is significantly more stable than the C2′-endo form (Yokoyama *et al.* 1985). This stabilization of the C3′-endo form in 2-thiopyrimidine nucleotides is due to an intramolecular steric effect (figure 9); the C2′-endo form is destabilized by the steric repulsion between the 2-thiocarbonyl group of the uracil base and the 2′-OH group of the ribose ring (Yokoyama *et al.* 1979, Yamamoto *et al.* 1983). The steric effect in the nucleotide unit of 2-thioribothymidine is also responsible for the enhanced thermostability of extreme thermophile tRNA species (Watanabe *et al.* 1979, Horie *et al.* 1985, Yokoyama *et al.* 1987).

By contrast, for pmo^{5}U, the enthalpy difference between the C2′-endo form and the C3′-endo form is obtained as −0·7 kcal mol^{-1}; the C2′-endo form is remarkably more stable than the C3′-endo form. Thus the two types of modifications of uridine found in the first position of an anticodon of tRNA significantly affect the puckering equilibria of the ribose ring moiety; the relative stability of the C3′-endo form in pmnm^{5}s^{2}U is higher by about 2 kcal mol^{-1} as compared to that in pmo^{5}U (Yokoyama *et al.* 1985). The effects of the two types of modifications of uridines in the first position of an anticodon will be discussed later in section 9.

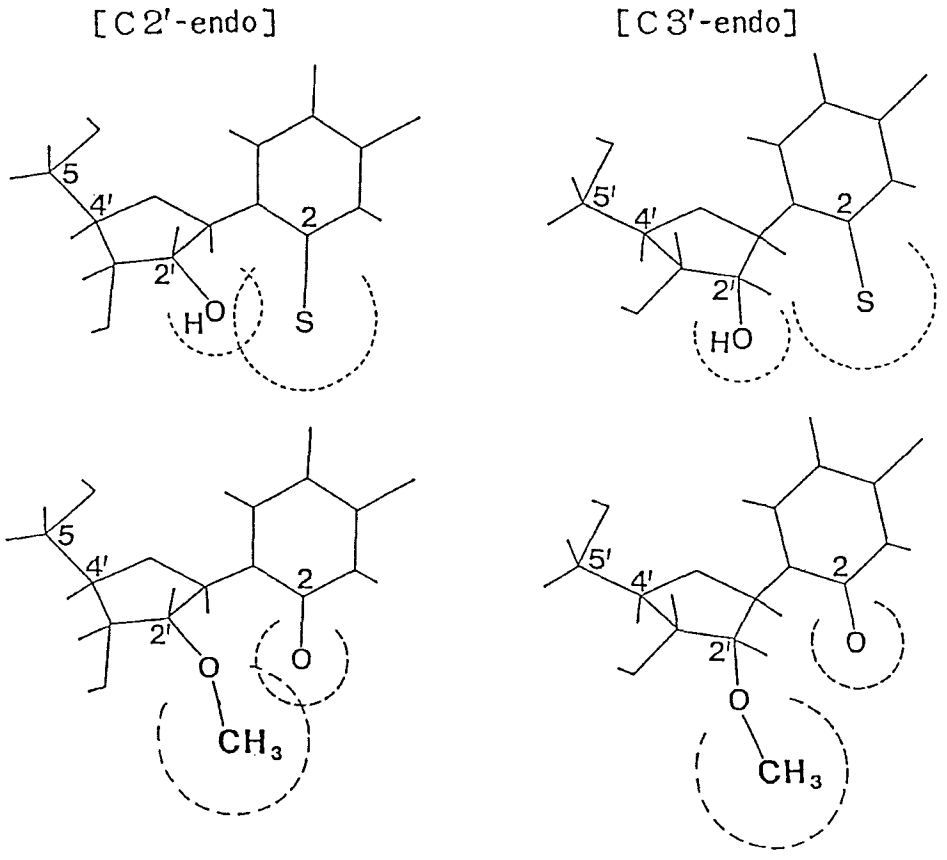


Figure 9. Intramolecular steric repulsion in 2-thiouridine derivatives (top) and in 2′-*o*-methylpyrimidine derivatives (bottom).

8. Lanthanide probe analyses of conformer equilibria in solution

Vicinal spin coupling constants are useful for analysing local conformational equilibria about single bonds. Actually, however, rotamer equilibria about a single bond are interrelated with rotamer equilibria about an adjacent bond. In uridine 3′-monophosphate in aqueous solution (figure 10), rotamer equilibria about the exocyclic C3′–O3′ and C4′–C5′ bonds are closely related with the puckering equilibrium of the ribose ring (Yokoyama *et al.* 1981). For elucidating such interrelations, lanthanide probe analyses in conjunction with spin coupling analyses are important (Inagaki and Miyazawa 1981).

In uridine 3′-phosphate in aqueous solution, lanthanide ion (Pr(III) or Eu(III)) is bound on the bisector of the phosphate group (Yokoyama *et al.* 1981). The lanthanide-induced shift of the resonance frequency of the ith nucleus is proportional to $(3\cos^2\theta_i - 1)/r_i^3$, where r_i and θ_i are spherical polar coordinates of the ith nucleus in the coordinate system of the principal magnetic axes with the lanthanide ion as the origin (Bleaney 1972).

The spin-lattice relaxation rates of ligand nuclei induced by Gd(III) ion are also important for elucidating conformer equilibria. The Gd(III)-induced relaxation rate of the ith nucleus is proportional to r_i^{-6}, provided that the dipolar interactions between the electron spin of Gd(III) ion and the nuclear spins are characterized by a common correlation time. Combined analyses of Gd(III)-induced relaxation rates and lanthanide-induced shifts, together with vicinal coupling constants for all the available resonance nuclei, are useful for elucidating the molecular conformations and conformer equilibria of flexible molecules, such as uridine 5′-monophosphate and uridine 3′-monophosphate (Inagaki *et al.* 1978, Yokoyama *et al.* 1981).

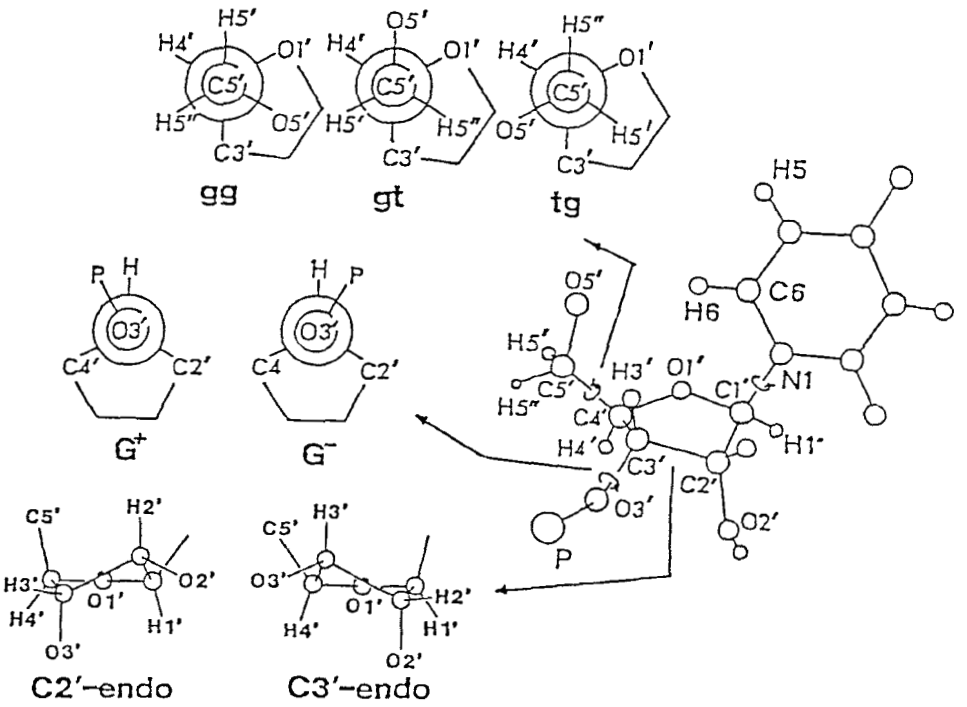


Figure 10. Local conformations of uridine 3′-phosphate.

For uridine 3′-monophosphate, dihedral angles (C4′–C3′–O3′–P) about the exocyclic C3′–O3′ bond (figure 10) have been obtained as $281 \pm 11°$ for the G^+ form and $211 \pm 8°$ for the G^- form. Further, the interrelations between the puckering equilibrium of the ribose ring and the rotamer equilibria about exocyclic bonds have been elucidated. The ribose ring in the C3′-endo form exclusively takes the gg form about the C4′–C5′ bond and the G^- form about the C3′–O3′ bond (figure 10), thus resulting in the rigidity of the ribonucleotide unit. This gg-C3′-endo-G^- form of the ribose phosphate moiety is the basic structure unit in the base-stacked ordered conformation of the RNA A-type helix. By contrast, the ribose ring in the C2′-endo form preferentially takes the gg and gt forms about the C4′–C5′ bond and the G^+ form about the C3′–O3′ bond (figure 10), and thus the ribose 3′-phosphate moiety in the C2′-endo form is more flexible than that in the C3′-endo form. This short range conformational interrelation of the ribose 3′-phosphate moiety is largely due to intranucleotide steric repulsion, affecting the structural feature of the anticodon of tRNA (Yokoyama *et al.* 1981).

9. Regulation of codon recognition by modification of the anticodon

In protein biosynthesis, nucleotide sequences of mRNA are translated to amino acid sequences (figure 1), where tRNA molecules serve as adaptors. The 3′-terminal adenosine of tRNA is charged with the cognate amino acid and the anticodon (position 34–36) recognizes the codon of amino acid, by formation of three base pairs (figure 7). The X-ray analyses of phenylalanine tRNA from yeast have shown that this tRNA is in an L-shaped conformation (Kim *et al.* 1974, Robertus *et al.* 1974); the nucleotide in the first position of the anticodon is in the gg-C3′-endo-G^- form.

The effect of the two types of modification of uridine (position 34) on the conformational aspects of the anticodon will now be discussed. The modification of uridine to xm^5s^2U is to further stabilize the intrinsic conformation of the first letter of the anticodon, namely gg-C3′-endo-G^- (figure 11 top), because of the extreme stability of the C3′-endo form in the xm^5s^2U unit itself. The anticodon is made 'rigid', and $xm^5s^2U(34)$-bearing tRNA species recognize codons terminating in A only. Note that $xm^5s^2U(34)$ is found in tRNAs specific to Gln, Lys and Glu, which have two degenerate codons terminating in A or G (table). The modification to xm^5s^2U does not allow misrecognition of codons terminating in U or C, thus contributing to *correct* translation of codons for Gln, Lys, and Glu (Yokoyama *et al.* 1985).

By contrast, the modification of uridine to $xo^5U(34)$ stabilizes the C2′-endo-G^+ form (figure 11 bottom) as well as the C3′-endo-G^- form, because of the remarkable stability of the C2′-endo form in the xo^5U unit itself. Thus the anticodon is made 'flexible', and xo^5U-bearing tRNA species recognize codons terminating in U and G as well as A. In particular, in the C2′-endo-G^+ form, the modified uracil base is displaced toward the codon and forms a base pair with uridine in the first position of the codon (figure 11 bottom). Note that $xo^5U(34)$ is found in tRNAs specific to Val, Thr, Ser, Pro and Ala, which have four degenerate codons terminating in U, C, A and G. The modification to xo^5U allows recognition of codons terminating in U, A and G, thus contributing to *efficient* translation of codons for these amino acids.

The conformational characteristic of modified uridines in the first position of the anticodon has been confirmed *in situ* by 2D-NMR analyses of tRNA species bearing modified uridines (Kawai 1989). Thus the biological significance of the two types of post-transcriptional modification of uridine is their contribution to the *correct* and *efficient* translation of codons, through the regulation of the conformational

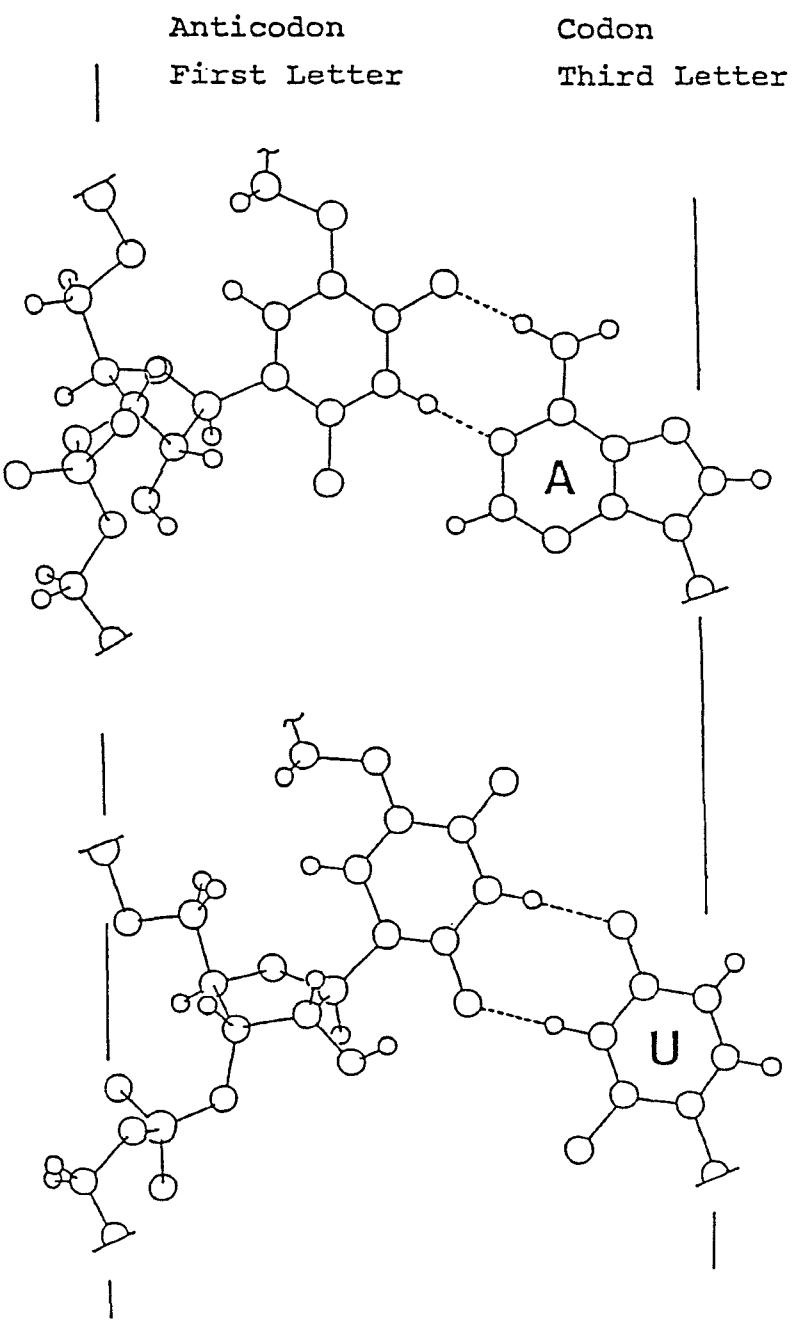
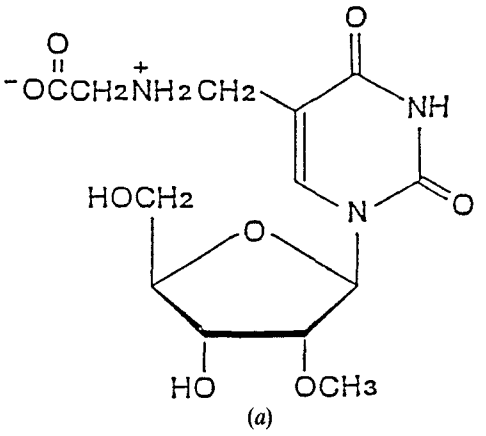


Figure 11. Recognition of adenosine (A) and uridine (U) in the third position of the codon of mRNA by modified uridine in the first position of the anticodon of tRNA (Yokoyama *et al.* 1985).

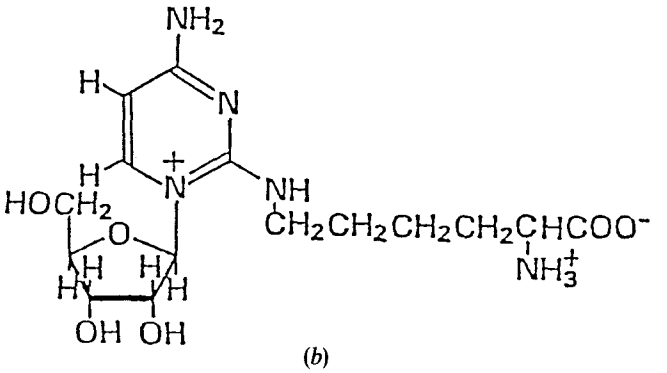
characteristics of the first letter of the anticodon (Yokoyama *et al.* 1985). This is probably the first clear case of an advanced NMR method providing important information for elucidating the conformational aspect (internal rotation) of biomolecules as involved in regulation of biological function.

The molecular mechanism in regulation of codon recognition has also been elucidated in the case of leucine tRNA from *E. coli*. As shown in the table, codon box UUN (where N = U, C, A or G) is divided into two parts; UUA and UUG in the lower half are the codons of leucine (Leu), similar to the cases for Gln, Lys and Glu. These codons are recognized by a minor leucine tRNA having an unknown nucleoside in the first position of the anticodon (Nishimura 1979). For recognition of the two codons in the lower half, rather than in the upper half, of a codon box, the modified nucleoside in this tRNA is expected to take predominantly the C3′-endo form, as in the case of xm^5s^2U.

The chemical structure of this novel nucleoside (figure 12 (*a*)) has been determined as 5-carboxymethylaminomethyl-2′-*o*-methyluridine (cmnm^{5}Um). This modified nucleoside in fact takes preferentially the C3′-endo form as found by proton NMR analyses. The stabilization of the C3′-endo form in this 2′-*o*-methylpyrimidine nucleoside is due to the steric effect; the C2′-endo form is destabilized by the steric



(*a*)



(*b*)

Figure 12. Chemical structures of novel nucleosides, (*a*) 5-carboxymethylaminomethyl-2′-*o*-methyluridine and (*b*) lysidine (N^+).

repulsion between the 2-carbonyl group of the uracil base and the 2′-*o*-methyl group of the ribose ring (figure 7 bottom) (S. Yokoyama, N. Horie, Z. Yamaizumi, Y. Kuchino, S. Nishimura and T. Miyazawa 1990, unpublished). Thus the molecular mechanism in the selective recognition of codons terminating in A and G has been established on the basis of the conformational characteristics of a modified nucleotide in the anticodon of tRNA.

10. Recognition of the minor isoleucine codon

In the genetic code table, the codon box AUN (where N = U, C, A or G) is uniquely divided into the three codons for Ile and one codon for methionine (Met); the regulation of codon recognition here must be different from the cases of Gln, Lys, Glu and Leu. *E. coli* has two $tRNA^{Ile}$ species; the major species recognizes codons AUU and AUC (Yarus and Barrell 1971) while the minor species recognizes codon AUA (Harada and Nishimura 1974). The minor species has an unknown modified nucleoside in the first position of the anticodon (Kuchino *et al.* 1980). The nucleotide sequence of minor $tRNA^{Ile}$ has been determined and the chemical structure of the novel nucleoside in the first position of the anticodon has been found as shown in figure 12(*b*). From the sequence analysis of the gene for minor $tRNA^{Ile}$, this nucleoside has been found to be derived from cytidine and accordingly named as lysidine, since this nucleoside has a lysine chain substituted in position 2 of the pyrimidine ring (Muramatsu *et al.* 1988a).

The anticodon of minor $tRNA^{Ile}$ is coded by CAT (T: thymidine) that is characteristic of $tRNA^{Met}$ (figure 13 top). The nascent anticodon CAU (figure 13 bottom left) is modified after the transcription from DNA and converted to the mature anticodon $N^{+}AU$ (figure 13 bottom right). If an immature tRNA species with the nascent

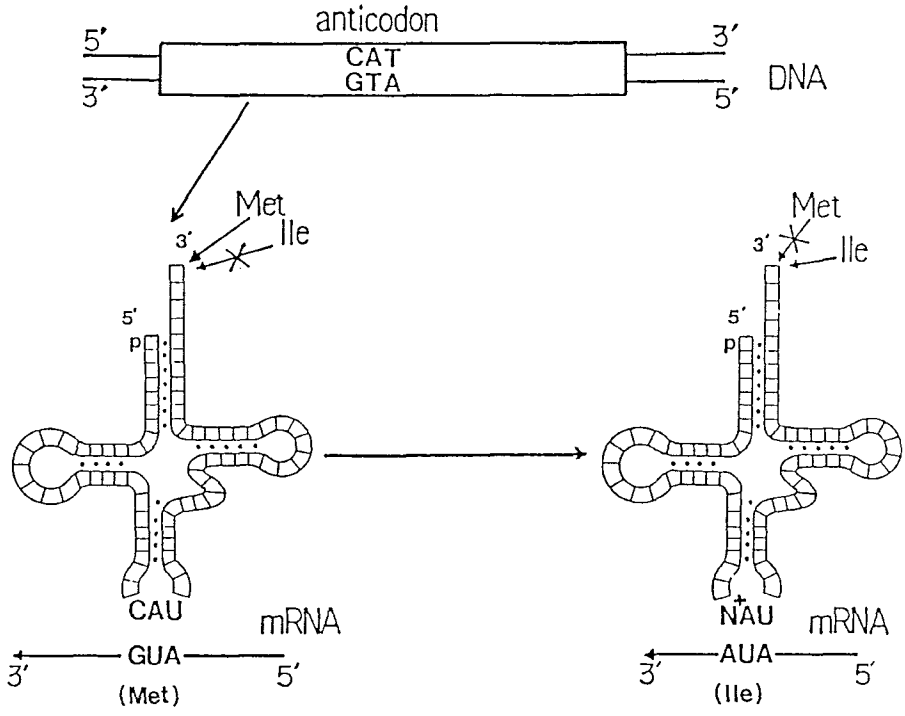


Figure 13. Post-transcriptional modification from $tRNA^{Met}$ to $tRNA^{Ile}$.

anticodon CAU is charged with isoleucine, this should result in the mistranslation of the codon AUG to isoleucine rather than to methionine. To solve this problem, the mature $tRNA^{Ile}$ has been converted back to a precursor tRNA by a series of enzyme treatments and then subjected to aminoacylation experiments. Surprisingly, the precursor tRNA with the anticodon CAU is charged with methionine rather than isoleucine, thus avoiding mistranslation. The post-transcriptional modification of cytidine (C) to lysidine (N^+) is found to convert concurrently the amino acid specificity and the codon specificity of the minor $tRNA^{Ile}$ precursor (figure 14), the first discovery of such modification in molecular biology (Muramatsu *et al.* 1988b, Yokoyama *et al.* 1989).

The conversion of amino acid specificity upon the anticodon modification is actually due to the recognition of tRNA species by aminoacyl-tRNA synthetases. Isoleucyl-tRNA synthetase does not charge with isoleucine tRNA species that have cytidine in the first position of the anticodon; cytidine there is the negative determinant against this enzyme. By contrast, methionyl-tRNA synthetase does not charge with methionine tRNA species that do not have cytidine in the first position of the anticodon; cytidine there is a positive determinant for this enzyme. The strict recognition (discrimation) of tRNA species by aminoacyl-tRNA synthetases is essential for correct translation of nucleotide sequences of mRNA into amino acid sequences of protein, which is a challenging subject for advanced physicochemical studies.

11. Incorporation of unnatural amino acids in protein

An aminoacyl-tRNA synthetase strictly recognizes the cognate amino acid, distinctly from the other 19 amino acids present in the cell. However, substrate recognition of aminoacyl-tRNA synthetases is not necessarily strict in discriminating against nonprotein amino acids. For example, *L*-norleucine, being similar to *L*-methionine in chemical structure, is activated by methionyl-tRNA synthetase and

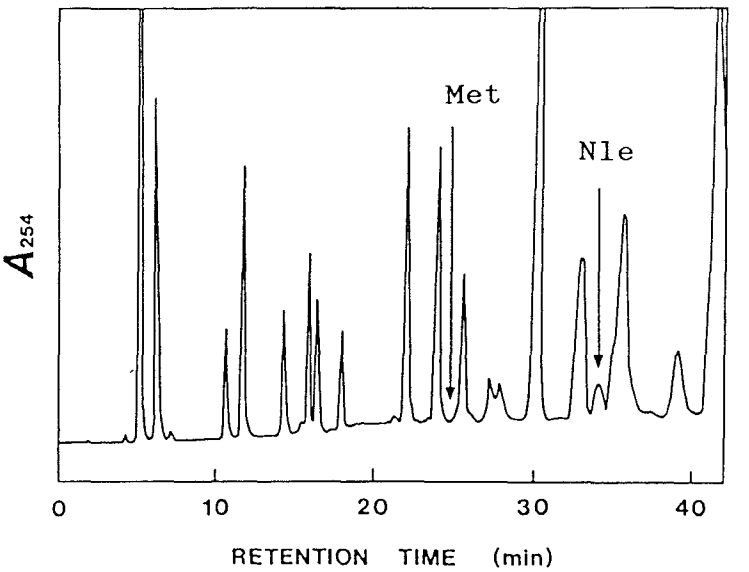


Figure 14. Amino acid analysis of the hydrolyzate of norleucine substituted human EGF (Koide *et al.* 1988); a peak due to norleucine (Nle) is observed but not a peak due to methionine (Met).

incorporated into proteins, although the efficiency of incorporation is not high in the *in vivo* biosynthesis (Anfinsen *et al.* 1969, Naider *et al.* 1972). Canavanine is similar to arginine in chemical structure; this nonprotein amino acid is activated by arginyl-tRNA synthetase and incorporated into proteins (Attias *et al.* 1969). Formation of furanomycyl-tRNA^{Ile} (section 3) suggests the possibility that this nonprotein amino acid could be incorporated in protein. In fact, in *in vitro* protein synthesis with the cell components of *E. coli*, furanomycin is incorporated, in place of *L*-isoleucine, in β-lactamase precursor (Kohno *et al.* 1990). The incorporation of furanomycin indicates that an unnatural amino acid, however different from any protein-constituting amino acid, could be incorporated into proteins provided that the unnatural amino acid is used in the aminoacylation of a tRNA species with an aminoacyl-tRNA synthetase.

Current protein engineering (Ulmer 1983) is useful for studying the structure–function relationship of proteins and for designing novel proteins; an amino acid residue in a protein may be replaced by another member of the family of 20 amino acids. However, the variety of protein-constituting amino acids is limited. If amino acids with side chains bearing 'unnatural' functional groups could be incorporated into proteins, the scope of protein engineering would be drastically expanded. Efficient synthesis of a protein substituted with non-protein amino acid (alloprotein) has much potential application in novel protein engineering.

12. Biosynthesis of alloproteins

Alloproteins may be synthesized in two ways: chemical synthesis and biosynthesis. In principle, any non-protein amino acid can be chemically incorporated into peptides, provided that appropriate methods are designed for protecting functional groups. However, the molecular weight of peptides that can be chemically synthesized is actually limited. By contrast, in protein biosynthesis, molecular weights of synthesized proteins are practically unlimited.

Non-protein amino acids with no cytotoxicity may be readily incorporated into proteins. Tyrosine and tryptophan residues in some proteins may be substituted with *m*-fluorotyrosine and 4-fluorotryptophan respectively, without any effects on protein conformations (Sykes *et al.* 1974, Browne and Otvos 1976). Note that such fluorine substitution does not affect protein function, so that substituted proteins may well be accumulated in the cell. By contrast, cytotoxic amino acids are expected to be useful for effective modification of protein structures and functions.

A toxic amino acid, norleucine, has been successfully incorporated into human epidermal growth factor (hEGF) by an *E. coli* strain harbouring a plasmid that has alkaline phosphatase (*phoA*) promotor and the region encoding *phoA* signal peptide and hEGF. Cells are cultured first in a high-phosphate medium and then, for induction of the hEGF-encoding gene, transferred to a low-phosphate medium containing norleucine (Nle). hEGF and Nle-substituted hEGF are secreted into the periplasm and then recovered. After treatment with hydrogen peroxide, Nle-substituted hEGF is clearly separated from methionine-oxidized hEGF. Substitution of the methionine residue of hEGF with Nle has been confirmed by the amino acid analysis (figure 14). From the successful production of Nle-substituted hEGF, a basic strategy has now been established for preparing proteins substituted with nonprotein amino acid (alloprotein). Induction of the *phoA* promotor and secretion of the product to the periplasm depresses heat shock-like responses and subsequent hydrolysis of the product by cytoplasmic proteases (Koide *et al.* 1988). Incorporation of norleucine into proteins has also been reported by several groups (Gilles *et al.* 1988, Tsai *et al.* 1988, Lu

et al. 1988, Bogosian *et al.* 1989). Physicochemical design of amino acid side groups with specific reactivity or with novel physical properties and development of methods for efficient *in vivo* incorporation of those amino acids in proteins are challenging projects to pursue.

References

ANFINSEN, C. B., and CORLY, L. G., 1969, *J. biol. Chem.*, **244,** 5149.

ATTIAS, J., SCHLESINGER, M. J., and SCHLESINGER, S., 1969, *J. biol. Chem.*, **244,** 3810.

BACHA, H., RENAUD, M., LEFEVRE, J. F., and REMY, P., 1982, *Eur. J. Biochem.*, **127,** 87.

BLEANEY, B., 1972, *J. magn. Reson.*, **8,** 91.

BOGOSIAN, G., VIOLAND, B. N., DORWARD-KING, E. J., WORKMANN, W. E., JUNG, P. E., and KANE, J. F., 1989, *J. biol. Chem.*, **264,** 531.

BROWNE, D. T., and OTVOS, J. D., 1976, *Biochem. Biophys. Res. Commun.*, **68,** 907.

CRICK, F. H. C., 1966, *J. molec. Biol.*, **19,** 548.

EHRLICH, R., LEFEVRE, J. F., and REMY, P., 1980, *Eur. J. Biochem.*, **103,** 145.

ELLIOTT, A., 1953, *Proc. R. Soc., London* A, **221,** 104.

ELLIOTT, A., and AMBROSE, E. J., 1950, *Nature*, **165,** 921.

ERNST, R. R., BODENHAUSEN, G., and WOKAUN, A., 1987, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Oxford: Clarendon).

FAVRE, A., BALLINI, J. P., and HOLLER, E., 1979, *Biochemistry*, **18,** 2887.

FERSHT, A. R., 1977, *Biochemistry*, **16,** 1025.

FRUSHOUR, B. G., and KOENIG, J. L., 1975, *Advances in Infrared and Raman Spectroscopy*, Vol. 1, edited by R. J. H. Clark and R. E. Hester (London: Heyden & Son Ltd.), p. 35.

GILLES, A. M., MARLIERE, P., ROSE, T., SARFATI, R., LONGIN, R., MEIER, A., FERMANDJIAN, S., MONNOT, M., COHEN, G. N., and BARZU, O., 1988, *J. biol. Chem.*, **263,** 8204.

GRIFFIN, B. E., JARMAN, M., REESE, C. B., SULSTON, J. E., and TRENTHAM, D. R., 1966, *Biochemistry*, **5,** 3638.

HARADA, F., and NISHIMURA, S., 1974, *Biochemistry*, **13,** 300.

HARADA, I., TAKAMATSU, T., SHIMANOUCHI, T., MIYAZAWA, T., and TAMIYA, N., 1976, *J. phys. Chem.*, **80,** 1153.

HARA-YOKOYAMA, M., YOKOYAMA, S., and MIYAZAWA, T., 1984, *J. Biochem., Tokyo*, **96,** 1599; 1986, *Biochemistry*, **25,** 7031.

HECHT, S. M., 1979, *Transfer RNA: Structure, Properties, and Recognition*, edited by P. R. Schimmel, D. Söll and J. Abelson (New York: Cold Spring Harbor Laboratory), p. 345.

HIGASHIJIMA, T., WAKAMATSU, K., MILON, A., OKADA, A., and MIYAZAWA, T., 1988, *Peptide Chemistry 1987*, edited by T. Shiba and S. Sakakibara (Osaka: Protein Research Foundation), p. 679.

HOPFIELD, J. J., YAMANE, T., YUE, V., and COUTTS, S. M., 1976, *Proc. natl Acad. Sci. U.S.A.*, **73,** 1164.

HORIE, N., HARA-YOKOYAMA, M., YOKOYAMA, S., WATANABE, K., KUCHINO, Y., NISHIMURA, S., and MIYAZAWA, T., 1985, *Biochemistry*, **24,** 5711.

INAGAKI, F., and MIYAZAWA, T., 1981, *Prog. Nuclear Magn. Reson. Spectrosc.*, **14,** 67.

INAGAKI, F., TASUMI, M., and MIYAZAWA, T., 1978, *Biopolymers*, **17,** 267.

ISHIKURA, H., YAMAO, Y., and NISHIMURA, S., 1971, *Biochim. Biophys. Acta*, **228,** 471.

JARDETZKY, O., and ROBERTS, G. C. K., 1981, *NMR in Molecular Biology* (New York: Academic).

KAINOSHO, M., NAGAO, H., IMAMURA, Y., UCHIDA, K., TOMONAGA, N., NAKAMURA, Y., and TSUJI, T., 1985, *J. molec. Struct.*, **126,** 549.

KAINOSHO, M., NAGAO, H., and TSUJI, T., 1987, *Biochemistry*, **26,** 1068.

KAINOSHO, M., and TSUJI, T., 1982, *Biochemistry*, **24,** 6273.

KAWAI, G., 1989, *Proton Nuclear Magnetic Resonance Studies on Dynamic Structures and Functions of Modified Nucleosides in Transfer RNA*, Dissertation, University of Tokyo.

KAZIRO, Y., 1978, *Biochim. Biophys. Acta*, **505,** 95.

KENDREW, J. C., 1963, *Science*, **139,** 1259.

KENDREW, J. C., BODO, G., DINTZIS, H. M., PARRISH, R. G., WYCKOFF, H. W., and PHILLIPS, D. C., 1958, *Nature*, **181,** 662.

KENDREW, J. C., DICKERSON, R. E., STRANDBERG, B. E., HART, R. G., DAVIES, D. R., PHILLIPS, D. C., and SHORE, V. C., 1960, *Nature*, **185,** 422.

KENDREW, J. C., WATSON, H. C., STRANDBERG, B. E., DICKERSON, R. E., PHILLIPS, D. C., and SHORE, V. C., 1961, *Nature*, **190,** 665.

KIM, S. H., SUDDATH, F. L., QUIGLEY, G. J., MCPHERSON, A., SUSSMAN, J. L., WANG, A. H. J., SEEMAN, N. C., and RICH, A., 1974, *Science*, **185,** 435.

KJELDGAARD, N. O., and KURLAND, C. G., 1963, *J. molec. Biol.*, **6,** 341.

KOHDA, D., KAWAI, G., YOKOYAMA, S., KAWAKAMI, M., MIZUSHIMA, S., and MIYAZAWA, T., 1987, *Biochemistry*, **26,** 6531.

KOHNO, T., KOHDA, D., HARUKI, M., YOKOYAMA, S., and MIYAZAWA, T., 1990, *J. biol. Chem.* (to be published).

KOIDE, H., YOKOYAMA, S., KAWAI, G., HA, J.-M., OKA, T., KAWAI, S., MIYAKE, T., FUWA, T., and MIYAZAWA, T., 1988, *Proc. natl Acad. Sci. U.S.A.*, **85,** 6237.

KRAUSS, G., RIESNER, D., and MAASS, G., 1976, *Eur. J. Biochem.*, **68,** 81.

KUCHINO, Y., WATANABE, S., HARADA, F., and NISHIMURA, S., 1980, *Biochemistry*, **19,** 2085.

LEFEVRE, J. F., EHRLICH, R., and REMY, P., 1980, *Eur. J. Biochem.*, **103,** 155.

LORD, R. C., and YU, N. T., 1970, *J. molec. Biol.*, **50,** 509.

LORD, R. C., and YU, N. T., 1971, *J. molec. Biol.*, **51,** 203.

LU, H. S., TSAI, L. B., KENNEY, W. C., and LAI, P. H., 1988, *Biochem. Biophys. Res. Commun.*, **156,** 807.

LUSTIG, F., ELIAS, P., AXBERG, T., SAMUELSSON, T., TITTAWELLA, I., and LAGERKVIST, U., 1981, *J. biol. Chem.*, **256,** 2635.

MILLER, D. L., and WEISSBACH, H., 1977, *Molecular Mechanism of Protein Biosynthesis*, edited by H. Weissbach and S. Pestka (New York: Academic), p. 323.

MITRA, S. K., LUSTIG, F., AKESSON, B., AXBERG, T., ELIAS, P., and LAGERKVIST, U., 1979, *J. biol. Chem.*, **254,** 6397.

MIYAZAWA, T., 1960a, *J. molec. Spectrosc.*, **4,** 168; 1960b, *J. chem. Phys.*, **32,** 1647; 1961, *J. Polymer Sci.*, **55,** 215; 1962, *Polyamino Acids, Polypeptides and Proteins*, edited by M. Stahmann (University of Wisconsin Press), p. 201; 1963, *Aspects of Protein Structure*, edited by G. N. Ramachandran (London: Academic), p. 257.

MIYAZAWA, T., and BLOUT, E. R., 1961, *J. Am. chem. Soc.*, **83,** 712.

MIYAZAWA, T., MASUDA, Y., and FUKUSHIMA, K., 1962, *J. Polymer Sci.*, **62,** S62.

MIYAZAWA, T., SHIMANOUCHI, T., and MIZUSHIMA, S., 1956, *J. chem. Phys.*, **24,** 408; 1958, *J. chem. Phys.*, **29,** 611.

MIZUSHIMA, S., 1954, *Structure of Molecules and Internal Rotation* (New York: Academic).

MIZUSHIMA, S., SHIMANOUCHI, T., KURATANI, K., TSUBOI, M., SUGITA, T., NAKAGAWA, I., and KUROSAKI, K., 1952c, *Nature*, **169,** 1058.

MIZUSHIMA, S., SHIMANOUCHI, T., NAGAKURA, S., KURATANI, K., TSUBOI, M., BABA, H., and FUJIOKA, O., 1950a, *J. Am. chem. Soc.*, **72,** 3490.

MIZUSHIMA, S., SHIMANOUCHI, T., and TSUBOI, M., 1950b, *Nature*, **166,** 406.

MIZUSHIMA, S., SHIMANOUCHI, T., TSUBOI, M., and SOUDA, R., 1952a, *J. Am. chem. Soc.*, **74,** 270.

MIZUSHIMA, S., SHIMANOUCHI, T., TSUBOI, M., SUGITA, T., KATO, E., and KONDO, E., 1951, *J. Am. chem. Soc.*, **73,** 1330.

MIZUSHIMA, S., SHIMANOUCHI, T., TSUBOI, M., SUGITA, T., KUROSAKI, K., MATAGA, N., and SOUDA, R., 1952b, *J. Am. chem. Soc.*, **74,** 4639.

MIZUSHIMA, S., TSUBOI, M., SHIMANOUCHI, T., SUGITA, T., and YOSHIMOTO, T., 1954, *J. Am. chem. Soc.*, **76,** 2479.

MURAMATSU, T., NISHIKAWA, K., NEMOTO, F., KUCHINO, Y., NISHIMURA, S., MIYAZAWA, T., and YOKOYAMA, S., 1988b, *Nature*, **336,** 179.

MURAMATSU, T., YOKOYAMA, S., HORIE, N., MATSUDA, A., UEDA, T., YAMAIZUMI, Z., KUCHINO, Y., NISHIMURA, S., and MIYAZAWA, T., 1988a, *J. biol. Chem.*, **263,** 9261.

NAIDER, F., BOHAK, Z., and YARIV, J., 1972, *Biochemistry*, **11,** 3202.

NAKANISHI, M., TAKESADA, H., and TSUBOI, M., 1974, *J. molec. Biol.*, **89,** 241.

NISHIMURA, H., NAKAJIMA, K., and SHIMAOKA, N., 1955, *Ann. Rep. Shionogi Res. Lab.*, **5,** 560.

NISHIMURA, S., 1979, *Transfer RNA: Structure, Properties, and Recognition*, edited by R. R. Schimmel, D. Söll and J. N. Abelson (New York: Cold Spring Harbor Laboratory), p. 59.

NOGGLE, J. H., and SHIRMER, R. E., 1971, *The Nuclear Overhauser Effect* (New York: Academic).

ODA, K., KIMURA, F., HARADA, F., and NISHIMURA, S., 1969, *Biochim. Biophys. Acta*, **179,** 97.

PAULING, L., and COREY, R. B., 1951, *Proc. natl Acad. Sci. U.S.A.*, **37,** 729; 1953, *Proc. R. Soc., London* B, **141,** 10.

PAULING, L., COREY, R. B., and BRANSON, H. R., 1951, *Proc. natl Acad. Sci. U.S.A.*, **37,** 205.

RENAUD, M., BACHA, H., REMY, P., and EBEL, J. P., 1981, *Proc. natl Acad. Sci. U.S.A.*, **78,** 1606.

RIESNER, D., PINGOUD, A., BOEHME, D., PETERS, F., and MAASS, G., 1976, *Eur. J. Biochem.*, **68,** 71.

ROBERTUS, J. D., LADNER, J. E., FINCH, J. T., RHODES, D., BROWN, R. S., CLARK, B. F. C., and KLUG, A., 1974, *Nature*, **250,** 546.

SCHELLMAN, J. A., and SCHELLMAN, C., 1964, *The Proteins*, Vol. 2, Second edition, edited by H. Neurath (New York: Academic), p. 1.

SCHIMMEL, P. R., and SÖLL, D., 1979, *Ann. Rev. Biochem.*, **48,** 601.

SEKIYA, T., TAKEISHI, K., and UKITA, T., 1969, *Biochim. Biophys. Acta*, **182,** 411.

SHIMANOUCHI, T., and MIZUSHIMA, S., 1955, *J. chem. Phys.*, **23,** 707.

SIAMWIZA, M. N., LORD, R. C., CHEN, M. C., TAKAMATSU, T., HARADA, I., MATSUURA, H., and SHIMANOUCHI, T., 1975, *Biochemistry*, **14,** 4870.

SPRINZL, M., and WAGNER, T., 1979, *Transfer RNA: Structure, Properties, and Recognition*, edited by P. R. Schimmel, D. Söll and J. Abelson (New York: Cold Spring Harbor Laboratory), p. 473.

SUGETA, H., and MIYAZAWA, T., 1967, *Biopolymers*, **5,** 673; 1968, *Ibid.*, **6,** 1387.

SUGETA, H., GO, A., and MIYAZAWA, T., 1972, *Chem. Lett.*, 83; 1973, *Bull. chem. Soc. Japan*, **46,** 3407.

SYKES, B. D., WEINGARTEN, H. I., and SCHLESINGER, M. J., 1974, *Proc. natl Acad. Sci.*, **71,** 469.

TAIJI, M., YOKOYAMA, S., HIGUCHI, S., and MIYAZAWA, T., 1981, *J. Biochem., Tokyo*, **90,** 885; 1983, *Biochemistry*, **22,** 3220; 1985a, *J. Biochem., Tokyo*, **98,** 1447; 1985b, *Biochemistry*, **24,** 5776.

TANAKA, K., TAMAKI, M., and WATANABE, S., 1969, *Biochim. Biophys. Acta*, **195,** 244.

TSAI, L. B., LU, H. S., KENNEY, W. C., CURLESS, C. C., KLEIN, M. L., LAI, P. H., FENTON, D. M., ALTROCK, B. W., and MANN, M. B., 1988, *Biochem. Biophys. Res. Commun.*, **156,** 807.

TU, A. T., PRESCOTT, B., CHOU, C. H., and THOMAS, G. J. JR, 1976, *Biochem. Biophys. Res. Commun.*, **68,** 1139.

ULMER, K. M., 1983, *Science*, **219,** 666.

VAN WART, H. E., and SCHERAGA, H. A., 1986, *Proc. natl Acad. Sci.*, **83,** 3064.

WAGNER, T., CRAMER, F., and SPRINZL, M., 1982, *Biochemistry*, **21,** 1521.

WATANABE, K., YOKOYAMA, S., HANSSKE, F., KASAI, H., and MIYAZAWA, T., 1979, *Biochem. Biophys. Res. Commun.*, **91,** 671.

WATSON, J. D., 1976, *Molecular Biology of the Gene*, third edition (Menlo Park, California: W. A. Benjamin).

WILLICK, G. E., and KAY, C. M., 1976, *Biochemistry*, **15,** 4347.

WÜTHRICH, K., 1976, *NMR in Biological Research: Peptides and Proteins* (Amsterdam: North-Holland).

WÜTHRICH, K., 1989a, *Science*, **243,** 45; 1989b, *Acc. chem. Res.*, **22,** 36.

YAMAMOTO, Y., YOKOYAMA, S., MIYAZAWA, T., WATANABE, K., and HIGUCHI, S., 1983, *FEBS Lett.*, **157,** 95.

YAMASHIRO-MATSUMURA, S., and KAWATA, M., 1981, *J. biol. Chem.*, **256,** 9308.

YARUS, M., and BARRELL, B. G., 1971, *Biochim. Biophys. Acta*, **43,** 729.

YOKOYAMA, S., MIYAZAWA, T., YAMAIZUMI, Z., and NISHIMURA, S., 1979, *Nucleic Acids Res.*, **6,** 2611.

YOKOYAMA, S., INAGAKI, F., and JIYAZAWA, T., 1981, *Biochemistry*, **20,** 2981.

YOKOYAMA, S., WATANABE, T., MURAO, K., ISHIKURA, H., YAMAIZUMI, Z., NISHIMURA, S., and MIYAZAWA, T., 1985, *Proc. natl Acad. Sci. U.S.A.*, **82,** 4905.

YOKOYAMA, S., WATANABE, K., and MIYAZAWA, T., 1987, *Adv. Biophys.*, **23,** 115.

YOKOYAMA, S., MURAMATSU, T., HORIE, N., NISHIKAWA, K., MATSUDA, A., UEDA, T., YAMAIZUMI, Z., KUCHINO, Y., NISHIMURA, S., and MIYAZAWA, T., 1989, *Pure appl. Chem.*, **61,** 573.

YU, N. T., JO, B. H., and LIU, C. S., 1972, *J. Am. chem. Soc.*, **94,** 7572.