A CONFORMATIONAL RATIONALE FOR THE ORIGIN OF THE MECHANISM OF NUCLEICACID-DIRECTED PROTEIN SYNTHESIS OF 'LIVING' ORGANISMS

R. BALASUBRAMANIAN, P. SEETHARAMULU, and G. RAGHUNATHAN

Department of Biophysics and Crystallography, University of Madras, A.C. College Campus, Madras–600 025, India*

Abstract. The physical basis for the natural evolution of a primitive decoding system is presented using the concepts of molecular interactions. Oligoribonucleotides of five residues having U at the 5′-end, a purine at the 3′-end and any combination of three bases in the middle is taken as a primitive tRNA (PIT). From conformational considerations PIT is expected to have U-turn conformation wherein, N_3-H_3 of base U hydrogen-bonds with phosphate, three residues ahead leaving triplet bases called primitive anticodons (PAC) into a helical conformation, and this creates a cleft between U and PAC. An amino acid can be comfortably nestled into the cleft with the amide hydrogens and carboxyl oxygen hydrogen-bonded to the last purine and the first uridine, while the side-chain can interact with the cleft side of PAC. The other side of PAC is free to base-pair with triplet codons on a longer RNA. Also two PACs can 'recognize' consecutive triplet codons, and this leads to a dynamic interaction in which the amino and carboxyl ends are brought into proximity, making the formation of peptide bond feasible.

The cleft formed by different anticodon triplets, broadly speaking, shows preferences for the corresponding amino acids of the presently known codon assignment.

Thus the nucleicacid-directed protein synthesis, which is a unique feature of all 'living' organisms is shown to be a natural consequence of a particular way of favourable interaction between nucleic acids and amino acids, and our model provides the missing link between the chemical evolution of small organic molecules and biological evolution through the process of mutations in nucleicacids and nucleicacid-directed protein synthesis.

1. Prologue

The phenomenon of nucleicacid-directed protein synthesis may be said to be an important molecular process that distinguishes the 'living' from the 'non-living', since this is a unique feature of all living organisms. Unlike all ordinary chemical interactions, here a class of chain molecule, viz., the nucleic acids, possess some kind of 'coded information' for the synthesis of an entirely different class of chain molecules, viz., proteins, which are responsible for most of the functions of living organisms, and such a protein synthesis is achieved by what may be called as a decoding mechanism making use of catalysts (synthetases), adaptors (tRNAs) and 'jigs' (ribosomes). The evolution of such a complex mechanism from rather simpler chemical evolution of small organic molecules (Calvin, 1969; Lemmon, 1970; and Ponnamperuma, 1971) looks like a mysterious jig-saw puzzle, and it has been elegantly described by Crick *et al.* (1976) as a 'notoriously difficult problem' from which even speculations recoil. No wonder it is a widely discussed topic in the literature. Many of the papers are found cited in some of the recent articles, (e.g.,

---

* Contribution No. 507 from this department.

Lacey and Weber, 1977; Nelsestuen, 1978; Hopfield, 1978; see our list of references also). Experimental studies (Weber and Lacey, 1978, and references therein) and theoretical analyses (Jungck, 1978, and references therein) show broad correlations between anticodon nucleotides and amino acids. In this paper we present a mechanism whereby these relationships manifested themselves in the primitive decoding system and the physical basis for the natural evolution of such a system by giving a conformational rationale for its origin, from the basic concepts of molecular interactions.

2. Postulate

In some prebiotic stage of evolution, it is reasonable to assume, that oligonucleotides (of say, about five residues long) were abundant (Ponnamperuma and Mark, 1965; Ponnamperuma, 1971). They are expected to have taken up conformations very similar to the RNA-11 helix (hereinafter called as helical conformation) since such a structure is known to be one of the most favourable conformations (Arnott *et al.*, 1973). If uridine were at the 5'-end of these oligonucleotides, this residue is expected to take up the energetically favourable U-turn conformation (Kim and Sussman, 1976) wherein the residue U makes an 'about turn' and its $N_3 - H_3$ is hydrogen-bonded to the oxygen (O_L) of the phosphate, three residues ahead. (Usually nucleotides are numbered from the 5'-end of the chain molecule; see IUPAC-IUB Commission, 1970.) Figure 1a shows a
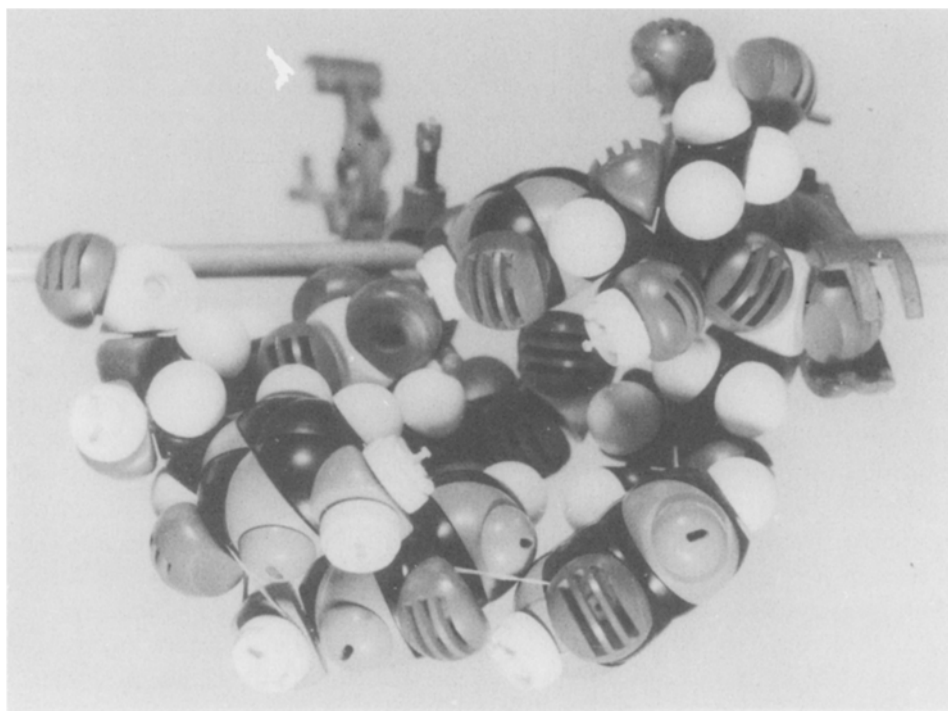

Fig. 1(a).   The photograph of a CPK-model for U-turn conformation. Note the cleft of space near the centre of the figure. Important atoms and bonds may be recognized by comparing it to Figure 1b.
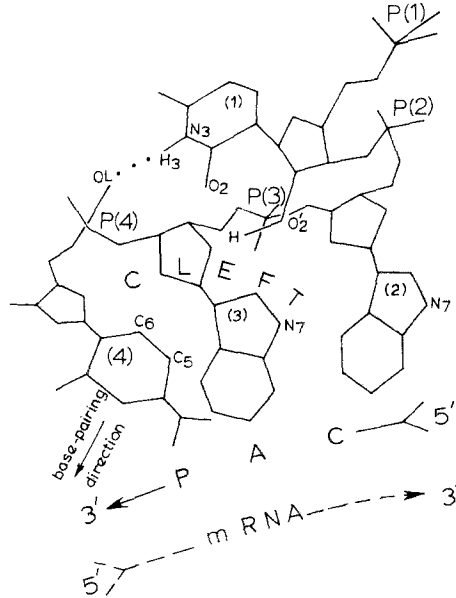
Fig. 1(b). Schematic diagram of important atoms and bonds for U-turn conformation (corresponding to the CPK-model in Figure 1a). Residues are numbered from 5'- to 3 -end. Note that N_3-H_3 of $U(1)$ is hydrogen-bonded (shown with three dots) to phosphate (4) and for this purpose, $U(1)$ is seen to have made an 'about-turn' with respect to the next three bases (2, 3 and 4) such that a cleft is formed and the cleft is 'lined' by groups O_2'-H of sugar (1), phosphate (2), N_7 or C_6 of residues (2, 3 and 4).

photograph of such a U-turn conformation built using CPK-model. This space-filling model shows a cleft in the middle (see the region near the centre of the figure). For convenience of discussions, we have marked important atoms, numbered the residues and shown the bonds in a line drawing (Figure 1b). From these two figures, one can readily see that the cleft is 'lined' by atoms O_2 of the base $U(1)^*$, O_2'-H of the sugar(1), phosphate of residue(2), and N_7 of the purine bases or C_6 of pyrimidine bases (2, 3, and 4). Phosphates of the third and fourth residues are seen at the back of the cleft.

At the same prebiotic stage of evolution, free amino acids too would have been abundantly available (see for example, Ferris *et al.*, 1978). Let us consider a situation wherein an amino acid is nestled in the above mentioned cleft of the oligonucleotides, in the following manner. Let N–H of the amino group be hydrogen bonded to O_2 of $U(1)$ and let C–O of the carboxyl group be hydrogen bonded, to O_2'-H of sugar(1). Now the side chain of the amino acid is in a favourable position to interact intimately with the bases (2, 3 and 4) on their cleft side. An amino acid nestled into the cleft formed by the above mentioned conformation of an oligonucleotide is shown in Figure 2, with the space-filling CPK-model. Actually in this figure, the fifth base is adenine and the atom N_7

---

*Throughout this article the numbers within parentheses refer to residue numbers counted from the first residue $U(1)$.
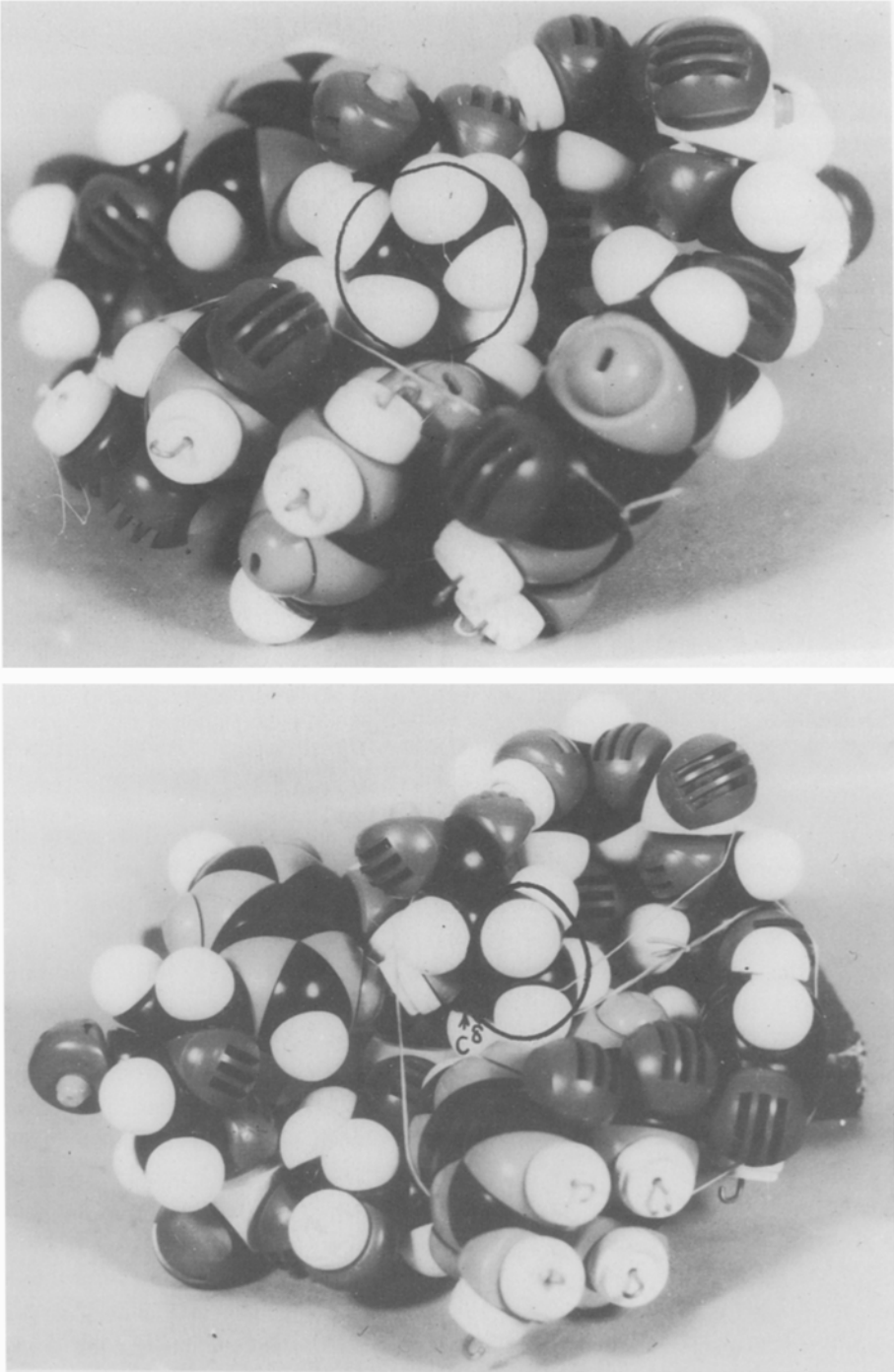
Fig. 2. CPK-model of aminoacid-nucleotide interaction. The primitive tRNA (PIT) of five residues having U at the 5'-end and purine at the 3'-end is seen, nestling an amino acid (Ile in Fig. 2a and Pro in Figure 2b). One of the hydrogens of the amino group is hydrogen-bonded to O_2 of $U(1)$, another to N_7 of $A(5)$ and the third is pointing towards the cleft (for iminoacid, Pro, this hydrogen is replaced by C^δ-atom). One of the oxygen of the carboxyl group is the acceptor to the hydrogen bond where O'_2—H of sugar (1) is the donor. The side chain is pointing towards the cleft-side of the primitive anticodon (PAC) triplet. The sidechains have been shown circled, for recognition.

of this $A(5)$ holds the amino group by $(N-H...N_7)$-hydrogen bond, limiting the possibilities of the side chain of the amino acid for its interaction with the cleft side of the bases (2, 3, and 4).

Let us call the oligonucleotide of five residues having U at its $5'$-end, purine at the $3'$-end and any three bases in the middle, as primitive t-RNA (PIT for short) and name the three bases in the middle as PAC (primitive anticodons). In our conformation, PAC is positioned such that it can simultaneously base-pair with a triplet codon on a longer RNA (which may be termed as primitive mRNA) having approximately an RNA-11 helical conformation, when the PIT can hold an amino acid, on the cleft side of PAC. Thus one side of PAC can interact with a primitive mRNA, while its other side can interact with an aminoacid side-chain and this simple stereochemical entity would act as a primitive decoding system. The conformation of the PIT is such that, the $3'$-side of one PAC can accommodate the $5'$-side of a second PAC so that two PACs can simultaneously interact with two contiguous triplet codons, on a primitive mRNA. This is made feasible because of the fact that the first residue $U(1)$ turns away to hydrogen-bond with the fourth phosphate while the fifth residue also turns away (these two residues also play the crucial role of creating a nest for an amino acid and of holding the amino acid by hydrogen bonds in a proper conformation such that the side chain of the amino acid is in close interaction with the cleft side of PAC), and the triplets of bases at the middle are free from encumbrances (see Figure 2) for a possible extension of their stacked conformation for continuous base-pairing with adjacent codons on primitive mRNA in helical conformation*.

Actually CPK-model-building shows that when a second PAC happens to come to 'recognize' the contiguous triplet (on the $3'$-side of primitive mRNA), the sugar moiety of the fifth residue of PIT_2 has a tendency to push, the first sugar moiety of PIT_1 causing its base $U(1)$ to swing (possibly around the bond $P-O_5$ of second residue in PIT_1) towards PIT_2 bringing the carboxyl group that it holds in proximity to the amino group of the amino acid held by PIT_2 and thus making the formation of peptide bond feasible. This process can repeat itself for a possible primitive protein synthesis. The stereochemical feasibility of this process is shown in Figure 3a where the proximity of the sites that held the carboxyl and amino groups in adjacent PIT's is shown, using the space filling model.

3. Substantiation

The conformation of the PIT in our model is energetically quite favourable. The helical conformation of PAC is a low energy conformation for RNA (Pullman and Saran, 1976) and U-turn is one of the most favourable conformations because of the hydrogen bond N_3-H_3 of $U(1)...O_L$ of $P(4)$ (Kim and Sussman, 1976).

The conformation of the complex between amino acid and primitive tRNA would also

---

*This is also stabilized by a hydrogen bond between $O_2'-H$ of residue (n) ... O_1' of residue $(n + 1)$ (Balasubramanian and Seetharamulu, 1979b).
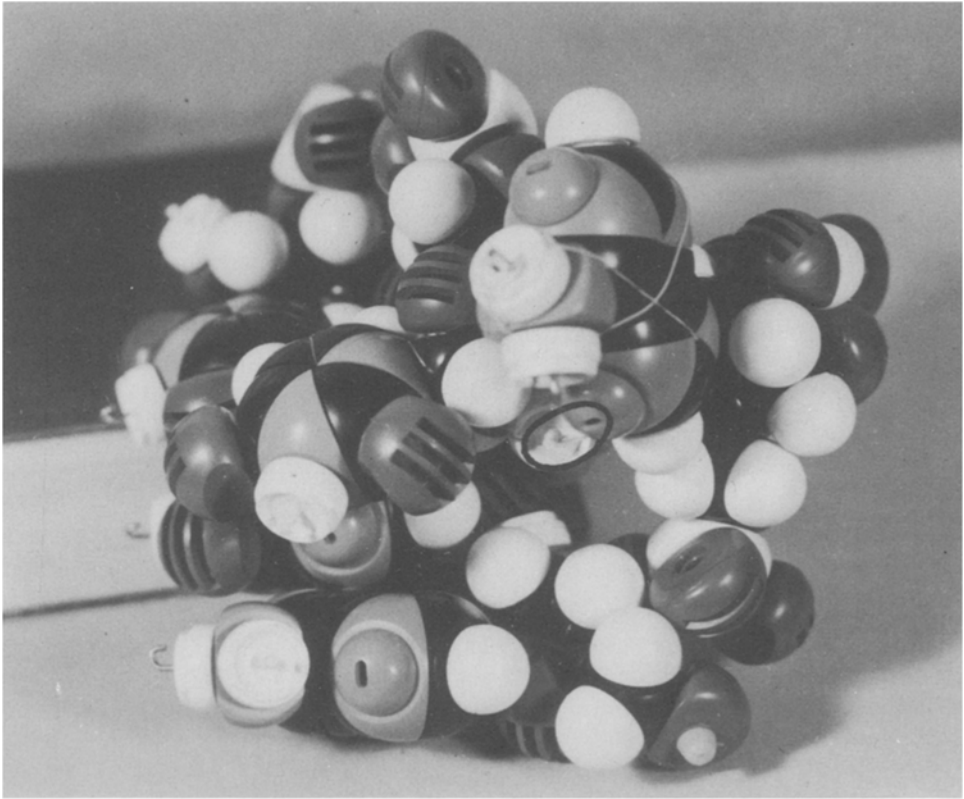
Fig. 3(a). $U(1)$ of the first primitive tRNA swings to a position such that O_2'—H is in proximity with N_7 of the next PIT showing the feasibility of the formation of a peptide bond during a primitive process of protein synthesis. A circle encloses H (of O_2') and N_7 to show their proximity.

be energetically very favourable since the amino acid is nestled into the cleft of the primitive tRNA, and the complex is thus a close packed arrangement.

Triplet coding turns out to be a natural consequence of this interaction. In this conformation where an amino acid is nestled into a cleft of oligonucleotides, U-turn is an important feature and note that in the U-turn conformation the base $U(1)$ is locked up into a hydrogen bond with the fourth phosphate allowing a trinucleotide to remain in a regular helical conformation. Here is thus a 'first-cause' explanation for triplet nucleotides coding for amino acids with a decoding mechanism reading the triplet codons in a non-overlapping comma-free fashion*.

The amino acid is held in the PIT such that the side chain is nearest to the middle base

*Some researchers have considered doublet and even singlet coding (see Dose, 1976; Lacey and Weber, 1976) in a very early stage of evolution. But Crick's (1968) arguments against any transition from singlet or doublet coding, to triplet coding have to be duly considered. Protein evolution has to start all over again from the time of such (hypothetical) transition. Of course the question of any such conceptual difficulties does not arise at all in our model.

Fig. 3(b). Amino acid adenylate ester is shown with carboxyl oxygen of amino acid in proximity with N_7 of the second PIT. The oxygen and N_7 are shown circled. Since one of the carboxyl oxygen is pointing away from the cleft, while the other is hydrogen-bonded to sugar (1), it is possible for the adenylate ester to be held without any steric hindrance in the model and even when $U(1)$ of PIT_1 swings to PIT_2 there is no steric hindrance as is seen in the figure.

of PAC, the third base is a little farther away, while the first is the farthest. This fits very well with the pattern of the genetic code, where the middle base plays the most important role, and the first (corresponding to the third of anticodon) and third (corresponding to the first of anticodon) are playing roles of decreasing importance (Woese, 1967). Only L-amino acids can fit into this model since in D-amino acids the sidechains would be situated in a direction away from the cleft and cannot have close-packed interaction with PAC. Only β–D-ribonucleotides can fit into the scheme as O_2'–H of sugar(1) and O_2 of $U(1)$ hold the amino acid; while N_3–H_3 of $U(1)$ hydrogen bonds with the fourth phosphate. The protein synthesis takes the direction of $5'$- to $3'$-end on mRNA, and amino to carboxyl end in polypeptide, and $3'$-, $5'$-phosphodiester linkage of the nucleotides are essential for our model. Moreover, in the present-day protein synthesis the growing peptide chain is transferred from its tRNA to the amino group, of the freshly arrived aminoacyl tRNA (Ingram, 1972). In our model also, this is the most

probable transfer since $U(1)$ of PIT_n holds the amino acid (when $n = 1$) or a peptide chain (when $n > 1$) by both the amide nitrogen and the carboxyl oxygen, using two hydrogen bonds, and if it swings around the $P-O_5$ bond, as discussed earlier, it transfers the first amino acid or the growing peptide chain to the next amino group of the freshly arrived PIT_{n+1}. The other possible transfer, viz., $A(5)$ swinging back carrying the amino acid with it to PIT_{n-1} is less likely since A's hold on the amino acid residue is less strong than U's hold. Thus the features of the proposed model for primitive tRNA neatly fit into a scheme where the evolution of present-day protein synthesis from its predecessor is but a smooth transition, and all the unique and asymmetric features of the contemporary mechanism are seen to be the consequence of a natural transition from a primitive system which is, incidentally, profoundly asymmetric. Thus our model can be said to give a square explanation to the observed unique features of present day protein synthesis (of course, being a stereochemical model, it cannot offer any specific explanation for the absence of a completely mirrored system in biological organisms).

Amino acid adenylate esters are supposed to have aided the prebiotic peptide synthesis (Lohrmann et al., 1975; Sawai et al., 1975; Weber and Orgel, 1978; see also Hecht, 1977), and, in our model, one of the oxygens of carboxyl group of an amino acid points away from the cleft, and an adenylate ester is stereochemically feasible as seen in the photograph of a CPK model (Figure 3b).

U is an invariant base preceding the anticodon triplets in all contemporary tRNAs, and it is to be remarked that the first base U in our model is a constant feature of PIT, and this U plays two important roles of creating a cleft in the PIT and of holding the amino acid for its side-chain to interact with PAC. The common amino- and carboxyl-groups of amino acids are hydrogen-bonded to the common residue U of the PIT, and the varied side-chain groups interact with the varied triplet bases of PIT, and these are all in keeping with the important general requirements of any reasonable model. Moreover the very occurrence of the so called U-turn conformation in the anticodon loop of contemporary tRNA paves the way for the attractive concept of smooth evolutionary transition from the proposed primitive tRNA to the present day 'adaptors' (Kim et al., 1973; Jack et al., 1976; Stout et al., 1976).

Here we would like to mention that, since the primitive tRNA is small, the growing polypeptide is likely to interact with and possibly wind around the proximate primitive mRNA, interfering with the very process of protein synthesis. This would limit the length of the so synthesised polypeptides to a few amino acid residues. Thus, we envisage that this limitation of the primitive tRNA necessitated the evolution of the present-day large tRNA, wherein the site of addition of amino acids to the growing polypeptide is well separated from the site of recognition of the code.

The residue succeeding the anticodon triplet is always a purine in contemporary tRNAs. Atom N_7 of PIT is common to purines, and in our theory this is hydrogen-bonded to one of the amide hydrogens of the amino acid. This again fits in well with one of the peculiar features of the set of 20 amino acids. Proline is an imino acid, in which, the side chain is covalently bonded to the amino nitrogen. In our postulate one of the

amino hydrogens of an amino acid is hydrogen bonded to O_2 of $U(1)$, another to N_7 of $A(5)$, and the third is constrained to point towards the cleft of PAC, and is free from any hydrogen-bonding. In the case of proline a carbon atom (C^δ) takes its place, and the imino side chain of proline is thus well accommodated into the cleft of the PAC. This is seen in Figure 2b.

That brings us to a study of the differential features of different PAC's with respect to their cleft-side and the possible interactions of PACs (especially the hydrogen bonding interaction) with corresponding amino acids. CPK-models have been built to study these interactions, and the results of our efforts are condensed into Table I. This table can be readily comprehended in conjunction with Figure 4 where the atoms and bonds of appropriate pairs of bases have been projected on the plane of the aromatic rings, bringing out their differential features. In this figure the van der Waals' profile of the cleft-side of the bases have been drawn to show how their space-filling properties differ. For example, note that pyrimidines leave less space in the cleft than purines. The differences in the orientations of CO and NH_2 groups account for the differences in hydrogen bonding possibilities.

A scrutiny of Table I (in conjunction with Figure 4) shows that for many amino acids, the interactions are specific and many other alternate assignment of triplet codons than the ones presently known are not favourable. But some of the interactions between anticodons and amino acids cannot be said to be of the exclusive or all-or-none type (see also Crick, 1968). Experimental studies (Lacey and Pruitt, 1969; Fox *et al.*, 1971; Raszka and Mandel, 1972; Woese, 1973; Saxinger and Ponnamperuma, 1974; Helene, 1977; Weber and Lacey, 1978) and theoretical analyses (Jungck, 1978) show broad correlations between anticodon nucleotides and their corresponding amino acids; but perfect correlations evidencing all-or-none specificity are not exhibited.

We envisage that the primitive system might have come into existence with the not-so-exclusive footing (as indicated by the above mentioned studies and by our model) and would have gradually been evolved into the now-known precision of decoding, where synthetases play a significant role. The imperfect nature of specific interaction was perhaps one of the most important factors (driving force) that has necessitated the evolution of these enzymes, viz., synthetases, which might supplement the cleft of PAC for an all-or-none specificity. The involvement of anticodon in the complexes of tRNAs and synthetases as evidenced by a variety of experimental data (Hayashi and Muira, 1966; Zachau, 1969; Rich and Schimel, 1977) gives credence to this view.

Apart from any possible specific affinity between nucleotide-triplets and the corresponding amino acids there must have been a sound basis for a natural choice (in evolution) of this particular system of chemical synthesis, viz., triplet nucleotides 'coding' for an amino acid and a decoding system with an anticodon translating the information into the sequence of amino acid residues. The above model provides this physical basis, since the coded message and the decoding mechanism are shown to be natural consequence of a particular way of energetically favourable interaction between the two classes of molecules.

TABLE I

Sr. No.	primitive anticodon (PAC) (2) (3) (4)[a]	amino acid	Sidechain[b]	Possible H-bonds between atoms of PAC and sidechain
1	G A A	PHE	CB(H2)-arromatic ring	

[c]Purines allow enough space for the arromatic ring. A(3), A(4) are less hydrophilic than G(3), G(4) and phenylalanine is highly hydrophobic. (See also Sr.No.8)

2	U A A C A A		CB(H2)-CG(H)-CD1(H3) | CD2(H3)	
		LEU		
3	G A G C A G			

Purines in positions (3) and (4) accommodate the branched and a little bulky aliphatic sidechain (compare it with valine)

4	G A U	ILE	CB(H)-CG1(H2)-CD(H3)
	I A U		CG2(H3)

CB is asymmetric carbon atom. Only HB can be cis to amino hydrogen pointing into the cleft. Thus CG2 has to be oriented to have close interaction with the third phosphate and thus fits into the cleft very well (see Fig. 2a). Note that for alloisoleucine CG2(H3) has to be replaced by CG1(H2)-CD(H3) and there is not enough space in the cleft. Probably this is the reason why alloisoleucine has been eliminated from proteins even during early stages of evolution.

5	C A U	MET	CB(H2)-CG(H2)-SD-CE(H3)	N4-H4 of C(2) . . . SD N6-H6 of A(3) . . .SD

In view of the first hydrogen bond C in the wobble position (2) is preferred for methionine. Secondly the bulky sulphur atom demands a purine in the third position and A is preferred because of the possibility of hydrogen bond.

6	G A C I A C U A C	VAL	CB(H)-CG1(H3) | CG2(H3)

The relative choice of the second two nucleotides for valine and alanine is in conformity with their hydrophilicities.

7	I G A			O6 of G(3) . . . H-OG
		SER	CB(H2)-OG(H)	
	U G A			N6-H6 of A(4) . . . OG

The hydrogen bonds are strained. See Sr.No.21. There the bases seem to be better suited for hydrogen bonds. This is in keeping with Weber and Lacey's (1978) comment (see p. 208) on the possibility of these anticodons being later additions.

8	U G G	PRO	CB(H2)-CG(H2)-CD(H2)-NH(of amide group)

Purines in positions (3) and (4) provide enough space for the relatively rigid amino ring. (For conformational details of proline ring see Balasubramanian et al., 1971 and Ramachandran et al., 1970.) A(3), A(4) is less hydrophilic than G(3), G(4) and so much so, phenylalanine is less hydrophilic than proline and this fits well with the relative assignments of anticodons.

TABLE I (*continued*)

Sr. No.	primitive anticodon (PAC) (2) (3) (4)[a]	amino acid	Sidechain[b]	Possible H-bonds between atoms of PAC and sidechain
9	G G U	THR	CB(H)-CG1(H3) | OG2(H)	O4 of U(4) . . . H-OG2

As in Sr.No.4, CB is asymmetric. Only HB can be cis to amino hydrogen pointing into the cleft. Since OG is bonded to U(4), CG(H3) has to point towards G(3). Since this is a unique configuration for hydrogen-bonding interaction allothreonine cannot be fitted to this 'pit' and, probably, that is why it was screened out of protein synthesis even in early stages of evolution. Third base has got to be a purine since CG(H3) has to be accommodated. The choice of G goes well with hydrophilic considerations. Compare this situation with that of serine (Sr.No.21). C(3) for serine gives better packing since CG(H3) of threonine is absent in serine.

Sr. No.	primitive anticodon (PAC) (2) (3) (4)[a]	amino acid	Sidechain[b]	Possible H-bonds between atoms of PAC and sidechain
10	I G C U G C	ALA	CB(H3)	

See comments for valine.

Sr. No.	primitive anticodon (PAC) (2) (3) (4)[a]	amino acid	Sidechain[b]	Possible H-bonds between atoms of PAC and sidechain
11	G U A	TYR	CB(H2)-arromatic ring-OH(H)	O4 of U(3) . . . H-OH N6-H6 of A(4) . . . OH

G at position (3) is not preferable since its O6 will be out of reach for the hydrogen bonding OH-H . . . O6. Similarly C at position (4) is not preferable since its N4-H4 would not be in the preferred hydrogen bonding direction.

Sr. No.	primitive anticodon (PAC) (2) (3) (4)[a]	amino acid	Sidechain[b]	Possible H-bonds between atoms of PAC and sidechain
12	G U G	HIS	CB(H2)-CG--------------ND1(H) | | CD2----NE2----CE1(H)	O6 of G(4) . . . H-ND1

Sr. No.	primitive anticodon (PAC) (2) (3) (4)[a]	amino acid	Sidechain[b]	Possible H-bonds between atoms of PAC and sidechain
13	U U G	GLN	CB(H2)-CG(H2)-CD-NE1(H2) | OE2	O4 of U(3) . . . H1-NE1 O6 of G(4) . . . H2-NE1

For glutamine, when compared to asparagine, there is one additional CH2 in the sidechain. Its fourth base in the anticodon is G instead of U. Thus the addition of CH2 group lengthens the sidechain so as to form better hydrogen bond with G than with U in the fourth position. (In Fig. 4 note that O6 of G is farther away from the glycosidic bond than O4 of U.)

Sr. No.	primitive anticodon (PAC) (2) (3) (4)[a]	amino acid	Sidechain[b]	Possible H-bonds between atoms of PAC and sidechain
14	G U U	ASN	CB(H2)-CG-ND1(H2) | OD2	O4 of U(3) . . . H1-ND1 O4 of U(4) . . . H2-ND1

If either the third or fourth base, or both were G instead of U, only one hydrogen bond is feasible. (Since O6 in G is farther away from the middle of the cleft.)

Sr. No.	primitive anticodon (PAC) (2) (3) (4)[a]	amino acid	Sidechain[b]	Possible H-bonds between atoms of PAC and sidechain
15	C U U or U U U	LYS	CB(H2)-CG(H2)-CD(H2)-CE(H2)-NZ(H2)	N4-H4 of C(2) . . . NZ or O4 of U(2) . . . H2-NZ O4 of U(3) . . . N1-NZ

The longer sidechain for lysine allows it to reach the second (Wobble) base. Note that there could have been rather a better assignments of bases for asparagine (such as GCU) and lysine (such as GGU) for slightly better hydrogen bonding. But then serine and threonine are to be dislocated in the codon table. See text for further discussion.

TABLE I (continued)

Sr. No.	primitive anticodon (PAC) (2) (3) (4)[a]	amino acid	Sidechain[b]	Possible H-bonds between atoms of PAC and sidechain
16	G U C	ASP	CB(H2)-CG-OD1(H) | OD2	O4 of U(3) . . . H-OD1 N4-H4 of C(4) . . . OD2

Compare asparagine and aspartic acid. The latter has only one donor. On comparing the anticodons, U(4) for asparagine is replaced by C(4) for aspartic acid. N4-H4 of C(4) acts as donor for OD2 of aspartic acid. The relative choice of anticodons seems to be in keeping with the optimisation of hydrogen-bonding possibilities.

| 17 | U U C | GLU | CB(H2)-CG(H2)-CD-OE1(H)
 |
 OE2 | O4 of U(3) . . . H-OE1

 N4-H4 of C(4) . . . OE2 |

A in position (4) is not suitable since its N6-H6 is out of reach for OE2 and G in position (3) is also not suitable since its O6 is out of reach for OE1. When compared to aspartic acid G(2) is replaced here by U(2). O4 of U(2) can have a bifurcated hydrogen bond with OE1 and this is feasible because glutamic acid has a longer sidechain than aspartic acid.

| 18 | G C A | CYS | CB(H2)-SG(H) | N4-H4 of C(3) . . . SG
 (a little strained)
 N6-H6 of A(4) . . . SG
 N7 of A(4) . . . H-SG |

C(3) seems to give close packing.

| 19 | C C A | TRP | CB(H2)-CG——————RROMATI
 | A C
 CD1(H)-NE(H)- R I N G | N7 of A(4) . . . H-NE |

C(2) and C(3) seem to give a close packing to the fused ring.

| 20 | I C G | ARG | CB(H2)-CG(H2)-CD(H2)-NE(H)-CZ-NH1(H2)
 |
 NH2(H) | O6 of I(2) . . . H-NH1
 N4-H4 of C(3) . . . NE
 O6 of G(4) . . . H-NH2 |

Note that for long and/or bulky aminoacids base (2) becomes relevant. See also Sr.No.22.

| 21 | G C U | SER | CB(H2)-OG(H) | N4-H4 of C(3) . . . OG
 O4 of U(4) . . . H-OG |

Since hydrogen bonding groups in pyrimidines are nearer to the centre of the cleft, C(3), U(4) as anticodon bases are more suitable than G(3), A(4) as OG-H of serine can reach them better (see comments on Sr.No.7).

| 22 | U C U | ARG | CB(H2)-CG(H2)-CD(H2)-NE(H)-CZ-NH1(H2)
 |
 NH2(H) | O4 of U(2) . . . H-NE
 N4-H4 of C(3) . . . NH2
 O4 of U(4) . . . H-NH1 |

Compare the hydrogen bonds in Sr.No.20 and 22. N4-H4 of C(3) can be donor to either NE or NH2 for different conformations of the long sidechain of arginine. Note that if NH2 were to be acceptor (Sr.No.22), positions (2) and (3) cannot be a purine, for with the present conformation of the sidechain of arginine, H-NE and H-NH1 would not be able to reach the acceptors O6's of purines.

TABLE I (*continued*)

Sr. No.	primitive anticodon (PAC) (2) (3) (4)[a]	amino acid	Sidechain[b]	Possible H-bonds between atoms of PAC and sidechain
23	G C C U C C C C C	GLY	(H)	

C(3), C(4) seems a relevant anticodon for glycine which has the smallest sidechain. Note from Fig 4 that C has the least space on the cleft side of PAC.

[a]Positions of bases in anticodons are given within parentheses. See text and Fig 1. Position (2) refers to wobble base. The presently known anticodon nucleotides (Dayhoff, 1976) have been considered for possible primitive anticodon nucleotides. See Fig. 4. For stereochemical formulae and other details of nucleotides. Note that I is similar to G as far as its cleft-side is concerned.

[b]Formulae for sidechains of aminoacids have been given for ready reference. See IUPAC-IUB Commission (1971) for nomenclature and symbols. We have used roman equivalents of Greek letters (like β = B etc.) as recommended by the commission. Hydrogens attached to C, N, O and S are given within parentheses. Bonding schemes for C, N, O and S are given by -'s.

[c]Brief comments have been interleaved in the table itself for the convenience of ready reference to the interactions between PAC and aminoacids. References to hydrophilicities given in the comments are made with respect to Weber and Lacey (1978).

4. Epilogue

In conclusion we would like to remark: (i) with this favourable possibility of dynamic association between oligonucleotides and amino acids, self-organization that is so essential for the origin of life becomes a natural consequence of chemical evolution of nucleic acids and amino acids and this would provide the missing link between the chemical evolution (evolution of biological small molecules) and biological evolution (evolution of biological organisms through the process of mutations in DNA and nucleicacid-directed protein synthesis); (ii) the model and hence the hypothesis is testable in the sense, oligonucleotides of five residues can be obtained and allowed to interact with amino acids and the conformation of the complex moiety may be studied by various physico-chemical methods; (iii) it suggests a prediction of a possible mechanism for the contemporary phenomenon of charging of tRNA by a synthetase in the following way: some portion of the synthetase perhaps supplements for the all-or-none specificity of interaction between the amino acid and the cleft side of the anti-codons in the contemporary tRNA, and the amino acid is checked against this *composite template* before the tRNA is charged with the amino acid at the CCA-end.
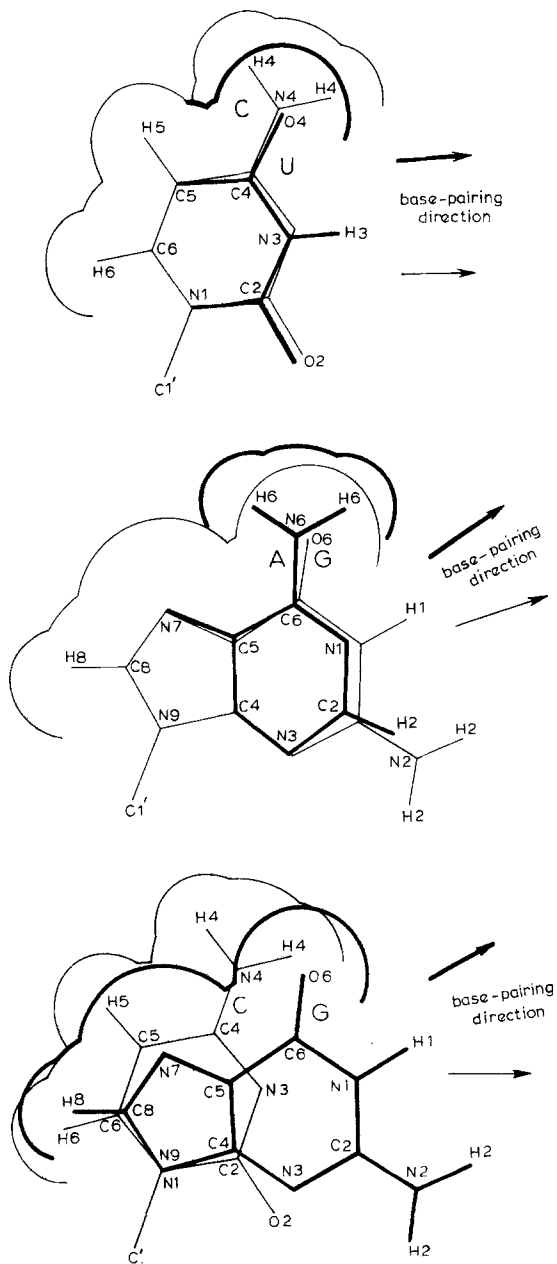
Acknowledgement

Fig. 4. Comparative geometry of bases, taken two at a time. The atoms and bonds have been projected in the plane of the rings (parameters taken from Seeman *et al.*, 1976 and Rosenberg *et al.*, 1976). On the cleft-side of the bases the van der Waals' profiles have been drawn (van der Waals' radii have been taken from Harte (1976). (a) Two pyrimidines have been compared. Note that *C* occupies more space near the edge of the cleft than *U*. (b) Two purines have been compared. Again *A* occupies more space near the edge of the cleft than *G*. (c) A pyrimidine (Py) and a purine (Pu) have been compared. Note that Py occupies more space towards the 'inside' of the cleft. Also note that hydrogen-bonding donors (acceptors) have different positions and orientations with respect to the cleft.

References

Arnott, S., Hukins, D. W. L., Dover, S. D., Fuller, W., and Hodgson, A. R.: 1973, *J. Mol. Biol.* 81, 107–122.

Balasubramanian, R.: 1979, *Ind. J. Biochem. Biophys.* (in press).

Balasubramanian, R., Lakshminarayanan, A. V., Sabesan, M. N., Tegoni, G., Venkatesan, K., and Ramachandran, G. N.: 1971, *Int. J. Pep. Protein Res.* 3, 25–33.

Balasubramanian, R. and Seetharamulu, P.: 1979, *Ind. J. Biochem. Biophys.* (in press).

Balasubramanian, R. and Seetharamulu, P.: 1979a, to be published.

Balasubramanian, R. and Seetharamulu, P.: 1979b, in preparation.

Calvin, M.: 1969, *Chemical Evolution*, Oxford Univ. Press, Oxford.

Carter, C. W., Jr. and Kraut, J.: 1974, *Proc. Natl. Acad. Sci. U.S.A.* 71, 283–287.

Caskey, C. T.: 1970, *Quart. Rev. Biophys.* 3, 295.

Cold Spring Harbor Symp. on Quant. Biol., Vol. 31: 1966, The Genetic Code.

Crick, F. H. C.: 1967, *Nature* 213, 798.

Crick, F. H. C.: 1968, *J. Mol. Biol.* 38, 367–379.

Crick, F. H. C. and Orgel, L. E.: 1973, *Icarus*, 19, 341–346.

Crick, F. H. C., Brenner, S., Klug, A., and Pieczenik, G.: 1976, *Origins of Life* 7, 389–397.

Dose, K.: 1976, in *Protein Structure and Evolution*, L. Fox *et al.* (Eds.), Marcel Dekker, New York, pp. 149–184.

Dunnil, P.: 1966, *Nature* 210, 1267–1268.

Eigen, M.: 1977, *Naturwiss.* 58, 465–523.

Epstein, C. J.: 1966, *Nature* 210, 25–28.

Ferris, J. P., Joshi, P. C., Edelson, E. H., and Lawless, J. G.: 1978, *J. Mol. Evol.* 11, 293–311.

Fox, S. W., Lacey, J. C. Jr., and Nakashima, T.: 1971, in *Nucleic Acid – Protein Interactions*, D. W. Ribbon and J. F. Woessner (Eds.), North-Holland Publ. Co., Amsterdam, p. 113.

Hartf, R. H.: 1976, *Molecules in Three Dimensions*, Ealing Beck Ltd., Watford.

Hartman, H.: 1975, *Origins of Life* 6, 423–429.

Haselkorn, R. and Rothman-Denes, L. B.: 1973, *Ann. Rev. Bioch.* 42, 397–438.

Hayashi, H. and Miura, K. I.: 1966, *Nature* 209, 376–378.

Hecht, S. M.: 1977, *Tetrahedron* 33, 1671–1696.

Helene, C.: 1977, *J. Biol. Chem.* 251, 249–255.

Hoffman, G. W.: 1974, *J. Mol. Biol.* 86, 349–362.

Hoffman, G. W.: 1975, *Ann. Rev. Phys. Chem.* 26, 123–144.

Hopfield, J. J.: 1978, *Proc. Natl. Acad. Sci. U.S.A.* 75, 4334–4338.

Ingram, V. M.: 1972, *Biosynthesis*, W. A. Benjamin Inc., California.

Ishigami, M. and Nagano, K.: 1975, *Origins of Life* 6, 551–561.

IUPAC-IUB Commission: 1971, *Biochem. J.* 121, 577–585.

Jack, A., Ladner, J. E., and Klug, A.: 1976, *J. Mol. Biol.* 108, 619–649.

Jukes, T. H.: 1966, *Molecules and Evolution*, Columbia Univ. Press, New York.

Jukes, T. H.: 1973, *Nature* 246, 22–26.

Jukes, T. H.: 1977, *Comprehensive Bioch.* 24, 235–293.

Jukes, T. H.: 1977, *Science* 198, 319–320.

Jungck, J. R.: 1978, *J. Mol. Biol.* 11, 211–224.

Kim, S. H. and Sussman, J. L.: 1976, *Nature* 260, 645–646.

Kim, S. H., Quingley, G. J., Suddath, F. L., Mcpherson, A., Kim, J. J., Weinzierl, J., and Rich, A.: 1973, *Science* 179, 285–288.

Kuhn, H.: 1972, *Angew. Chem. Inter. Ed.* 11, 798–820.

Lacey, J. C., Jr. and Fox, S. W.: 1972, *Theory and Experiments in Exobiology*, Wolters-Noordhoff, Groningen, pp. 35–63.

Lacey, J. C., Jr. and Pruitt, K. M.: 1969, *Nature* 223, 799–804.

Lacey, J. C., Jr. and Weber, A. L.: 1976, in L. Fox *et al.*, *Protein Structure and Evolution*, Marcel Dekker, New York, pp. 213–222.

Lacey, J. C., Jr. and Weber, A. L.: 1977, *Precamb. Res.* 5, 1–22.

Lacey, J. C., Jr., Weber, A. L., and Pruitt, K. M.: 1977, in R. Duncan and M. Weston-Smith (Eds.) *The Encyclopedia of Ignorance*, Pergamon Press, Oxford, pp. 219–225.

Lacey, J. C., Jr., Weber, A. L., and White, W. E.: 1975, *Origins of Life* 6, 273–283.

Lemmon, R. M.: 1970, *Chem. Rev.* 70, 95–109.

Lohrmann, R., Ranganathan, R., Sawai, H., and Orgel, L. E.: 1975, *J. Mol. Evol.* 5, 57–73.

Miller, S. L. and Orgel, L. E.: 1976, in W. D. Elroy *et al.* (Eds.), *The Origins of Life on the Earth*, Prentice-Hall Inc., New Jersey.

Mizutani, H. and Ponnamperuma, C.: 1977, *Origins of Life* 8, 183–219.

Nagavary, J. and Fendler, J. H.: 1974, *Origins of Life* 5, 357–362.

Nelsestuen, G. L.: 1978, *J. Mol. Evol.* 11, 109–120.

Nirenberg, M., Leder, P., Benfield, M., Brimacombe, R., Trupin, J., Rottman, F., and O'Neal, C.: 1965, *Proc. Natl. Acad. Sci. U.S.A.,* 53, 1161–1168.

Orgel, L. E.: 1968, *J. Mol. Biol.* 38, 381–393.

Paecht-Horowitz, M.: 1974, *Origins of Life,* 5, 173, 187.

Pelc, S. R.: 1965, *Nature,* 207, 597–599.

Pelc, S. R. and Welton, M. G E.: 1966, *Nature* 209, 868–870.

Ponnamperuma, C.: 1971, *Quart. Rev. Biophys.* 4, 77–106.

Ponnamperuma, C. and Mark, R.: 1965, *Science,* 148, 1221–1233.

Pullman, B. and Saran, A.: 1976, *Prog. Nuc. Acid. Res. Mol. Biol.* 18, 215–322.

Ralph, R. K.: 1968, *Bioch. Bioph. Res. Comm.* 33, 213–218.

Ramachandran, G. N., Lakshminarayanan, A. V., Balasubramanian, R., and Tegoni, G.: 1979, *Biochem. Biophys. Acta.* 221, 165–181.

Raszka, M. and Mandel, M.: 1972, *J. Mol. Evol.* 2, 38–43.

Reanney, D. C. and Ralph, R. K.: 1967, *J. Theoret. Biol.* 15, 41–52.

Rendell, M. S., Harlos, J. P., and Rein, R.: 1971, *Biopolymers.* 10, 2083–2094.

Rich, A. and Schimmel, P. R.: 1977, *Nucl. Acid. Res.* 4, 1649–1665.

Rosenberg, J. M., Seeman, N. C., Day, R. O., and Rich, A.: 1976, *J. Mol. Biol.* 104, 145–167.

Sawai, H., Lohrmann, R., and Orgel, L. E.: 1975, *J. Mol. Evol.* 6, 165–184.

Saxinger, C. and Ponnamperuma, C.: 1974, *Origins of Life.* 1, 5–10.

Saxinger, C. and Ponnamperuma, C.: 1974, *Origins of Life.* 5, 189.

Saxinger, C., Ponnamperuma, C., and Woese, C. R.: 1971, *Nature. New Biol.* 234, 172–174.

Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Kim, J. J. P., and Rich, A.: 1976, *J. Mol. Biol.* 104, 109–144.

Sonneborn, T. M.: 1965, in V. Bryson and H. J. Vogel (Eds), *Evolving Genes and Proteins*, Academic Press, New York.

Stout, C., Mizuno, H., Rubin, J., Brennan, T., Rao, S. T., and Sundaralingam, M.: 1976, *Nucl. Acid. Res.* 3, 1111–1123.

Thomas, B. R.: 1970, *Bioch. Bioph. Res. Comm.* 40, 1281–1296.

T'So, P. O. P.: 1974, in P. O. P. T'So (Ed.) *Basic Principles in Nucleic Acid Chemistry* Vol. 1, Academic Press, New York, pp. 1–92.

Von Hippel, P. H. and McGhee, J. D.: 1972, *An.. Rev. Bioch.* 41, 231–300.

Walker, G. W. R.: 1977, *Biosystems* 9, 139–150.

Weber, A. L. and Fox, S. W.: 1973, *Bioch. Bioph. Acta* 319, 174–187.

Weber, A. L. and Lacey, J. C., Jr.: 1978, *J. Mol. Evol.* 11, 199–210.

Weber, A. L. and Orgel, L. E.: 1968, *J. Mol. Evol.* 11, 189–198.

Welton, M. G. E. and Pelc, S. R.: 1966, *Nature,* 209, 870–872.

Woese, C. R.: 1967, *The Genetic Code*, Harper Row Inc., London.

Woese, C. R.: 1968. *Proc. Natl. Acad. Sci. U.S.A.* 59, 110–117.

Woese, C. R.: 1969, *J. Mol. Biol.* 43, 235–240.

Woese, C. R.: 1970, *Bioscience* 20, 471–485.

Woese, C. R.: 1970 *Nature* 226, 817–820.

Woese, C. R.: 1973, *Naturwiss.* 60, 447–459.

Woese, C. R., Dugre, D. H., Dugre, S. A., Kondo, M., and Saxinger, W. C.: 1967, *Cold Spring Harbor Symp. Quant. Biol.* 31, 723–736.

Woese, C. R., Dugre, D. H., Saxinger, W. C., and Dugre, S. A.: 1966, *Proc. Natl. Acad. Sci. U.S.A.* 55, 966–974.

Wong, J. T. F.: 1975, *Proc. Natl. Acad. Sci. U.S.A.* 72, 1909–1912.

Wong, J. T. F.: 1976, *Proc. Natl. Acad. Sci. U.S.A.* 73, 2336–2340.

Yarus, M.: 1969, *Ann. Rev. Bioch.* 38, 841–880.

Zachau, H. G.: 1969, *Angew. Chem. Int. Ed.* 8, 713–727.