J. molec. Evolution 2, 38—43 (1972)

Is There a Physical Chemical Basis for the Present Genetic Code?

M. Raszka and M. Mandel

Department of Biochemistry and Biophysics, University of Hawaii, Honolulu, Hawaii

Received August 18, 1972
Final version October 9, 1972

Summary. Recent experiments on the interaction of amino acids with homopolyribonucleic acids as measured by proton magnetic resonance spectroscopy and on the interactions of ribonucleoside 5'-monophosphates with immobilized amino acids give relative binding strengths that are not reconcilable with the present genetic code in any simple way.

Key words: Amino Acid — Nucleic Acid Interactions — Genetic Code.

Although the genetic code is now well defined, its origin is still unknown. At present there are essentially two main theories to account for its universality (1) the relationship between triplets and amino acids may be invariant because they were determined in some way by the chemical structures involved leading therefore to a universal code; (2) the structure of the code is a frozen accident. This assumes the common ancestry of all organisms and holds that the code is universal because once established any change in it would be lethal. That the code is evidently non-random can be accounted for in either theory (Crick, 1968; Orgel, 1968; Woese, 1967, 1969).

Without going into speculative details and complexities of these theories which are covered very well in the above references, suffice it to say that the stereochemical theory has two great advantages (1) it makes it easier to see how the system could start and (2) it seems possible to prove experimentally that such specific interactions exist. As far as the specificity is concerned all that is required is a sufficient number of slight preferences (Woese, 1969) and then we could rely on the selection process to bring about our present codon-amino acid catalogue.

If codon-amino acid pairing existed, what is the evidence for such interactions? Zubay and Doty (1958) used equilibrium dialysis to study the binding to calf-thymus DNA of those amino acids that are predominantly found in thymus tissue with essentially negative results. A preference of AT-rich DNA for polylysine has been reported (Leng and Felsenfeld, 1966)

and recent experiments by Fox *et al.* (1970) indicate that polyamino acids rich in arginine tend to react preferentially with polynucleotides rich in adenine and guanine while polyamino acids rich in lysine tend to react preferentially with polynucleotides rich in uridine and cytidine. Woese, Dugre, Saxinger and Dugre (1966) have defined an amino acid polar requirement determined by partition paper chromatography in 2,6-dimethylpyridine: water solvents. The ordering of amino acids by polar requirements bears a striking similarity to the ordering of amino acids by their coding triplets. For example, whenever two amino acids have codons that differ from each other only in the III position $\binom{X\ Y\ Z}{I\ II\ III}$, the two amino acids have similar polar requirements. Woese argues that this ordering of amino acids as defined by their interaction with heterocyclic bases and codon assignments indicates some foundation for the genetic code in amino acid-base interactions.

Another approach has been to build molecular scale models. Such model trinucleotides are flexible enough so that all their bases can be brought into contact with an amino acid (Woese and Brown, private communication). Not enough work has been done with these "cup" models to demonstrate whether they could show specificity.

Lacey and Pruitt (1969) have shown that mononucleotides interact with polyamino-acid chains as observed by the increasing turbidity of a poly L-lysine acid solution, but this interaction is charge dependent and lacks specificity. However, using Corey-Pauling three dimensional atomic models, they have shown that stereospecific interactions are possible between a α helical polypeptide and a helix of associated mononucleotides in the ratio of three nucleotides to one amino acid residue.

In the last year, two groups using completely different techniques have been measuring specific binding interactions between (a) amino acids and homopolyribonucleic acids and (b) between amino acids and ribonucleoside-5′-monophosphates.

Raszka and Mandel (1971, 1972) using proton magnetic resonance spectroscopy have studied the interaction of the 20 amino acids with neutral poly A and of a selected number of amino acids with poly I and poly U. In the case of the aromatic and aliphatic amino acids the chemical shift of the C_2 and C_8 protons of the adenine moiety of poly A is consistent with a destacking of the initially partly stacked polynucleotide chain by the intercalation of the side chain. The strength of this binding depends primarily on the size, or hydrophobic character of the amino acid. Further confirmation of this intercalation model comes from fluorescent studies by C. Héléne (1971) of the interaction of aromatic amino acids with calf thymus DNA. In addition to polynucleotide destacking caused by weak intercalation another mode of binding was deduced for several polar amino acids. In this case the oposite effect, stabilization of the single helix was

Table 1. Poly I C_2—C_8 proton line separation in the absence and presence of aromatic amino acids at 26° and 45° C

Interactants	Line Separation, Hz	
	$T=26°$ C	$T=45°$ C
0.04 M poly I in buffer	very broad overlapping lines	19.0
+0.05 M L-Trp	34.2	27.4
+0.15 M L-Phe	22.6	21.4
+0.25 M L-His	21.5	19.2

indicated presumably by ionic chelation to the phosphate backbone or to individual monomer units. Although binding constants can be calculated from this kind of chemical shift data, a more pragmatic approach was to determine the relative binding strength of the amino acids from measuring the absolute slope of the line obtained from plotting the relative chemical shift of the C_2 proton with respect to the C_8 proton as a function of amino acid concentration. The C_2–C_8 line separation in the absence of amino acid is defined as 1. In Fig. 1 A–C, we present the relative chemical shift for the C_2–C_8 protons of poly A in the presence of aromatic, aliphatic and polar amino acids respectively. From these data the relative binding strength of amino acids to poly A is L-Try > L-Phe > L-His > L-Arg > L-Leu > L-Ile > L-Met > L-Lys > L-Val > L-Pro > Gly.

In the case of poly I the C_2–C_8 line separation was measured in the absence and presence of defined amounts of amino acids as shown in Table 1.

The following amino acids Gly (0.75 M), L-Lys (0.25 M) and L-Pro (0.25 M) were not effective in changing the line separation at 45° or improving the resolution of the poly I spectrum at 26°. The order of binding strength to poly I can be tentatively assigned as follows:

L-Try > L-Phe > L-His > Gly, L-Pro.

Evidence for the interaction of amino acids with poly U was obtained by observing the two overlapping doublet C_1 and C_5 in the absence or presence of various amino acids. Again it was found that the aromatics are the most effective in causing chemical shift. The order of binding to poly U is:

L-Try > L-Phe > L-His > Gly, L-Pro, L-Lys.

Compare these results with those obtained by Saxinger and Ponnamperuma (1971) who studied the binding of ribonucleoside-5′-monophosphates to immobilized L-amino acids. The relative binding strength (the authors use the term "selectivity coefficient") of amino acids to 5′ U*p*, 5′ C*p*, 5′-I*p*, 5′ A*p* and 5′ G*p* is as follows:

L-Tyr > L-Trp > L-Phe > L-His > L-Arg > L-Met > L-Pro > L-Lys > Gly,

independent of the particular ribonucleoside.

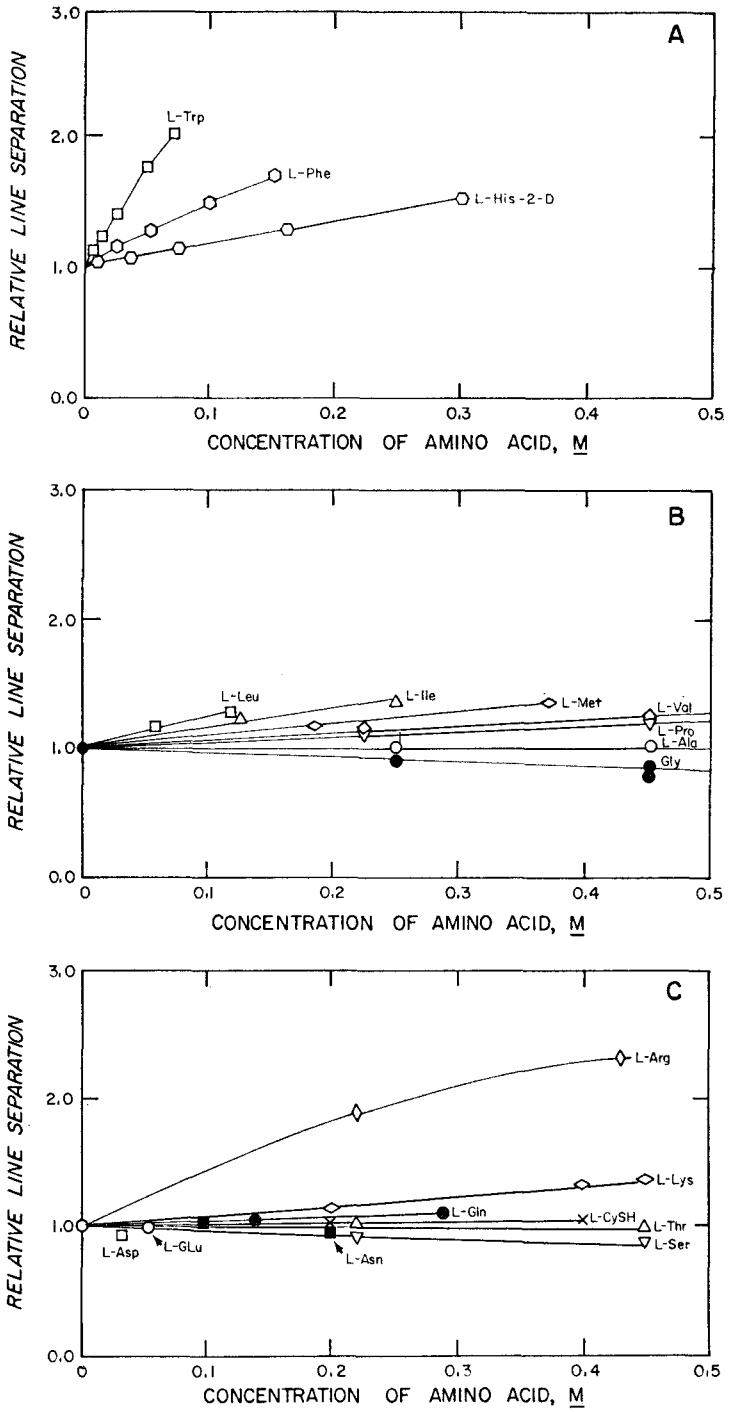

Fig. 1 A–C. Relative C_2–C_8 proton line separation of 0.04 M poly A in the presence of (A) aromatic amino acids, (B) aliphatic amino acids and (C) polar amino acids. All samples in 0.2 M phosphate buffer, pH = 7.0 (pH 7.6 in case of poly A·[2−^{2}H] L-His sample). Line separation of poly A alone = 1.0

In the case of poly A the only difference between the two types of experiment is the position of Lys. The magnetic resonance results with poly I and poly U although not as extensive, confirm the general picture that the 5′-ribonucleoside-P and the corresponding homopolymer interact most strongly with the aromatics amino acids and most weakly with Gly, Pro and Lys.

Although we are to a large extent ignorant of the complete physical nature of these interactions we can ask of what significance are they in the origin of the genetic code. If the present genetic code is representative of a preferred physical interaction, however weak between amino acid and codon, we should most easily recognize this in a study of the four homopolymers (poly A, poly U, poly G, poly C) and the corresponding amino acids that they encode. Depending on the primitive model you chose, phenylalanine for example should interact with the codon poly U or possibly its anticodon poly A. Similarly, lysine should interact most strongly with either poly A or poly U.

The presently available data give strong evidence that both poly A and poly U interact preferentially with phenylalanine over lysine. Even on the momomer level the data from Saxinger and Ponnamperuma confirm that A and U both prefer aromatic amino acids over lysine.

In the case of guanine and cytosine the present codon dictionary would require strong interactions with proline or glycine but we see from the data of Saxinger and Ponnamperuma that again the preferential interaction is with the aromatic amino acids.

The experimental results of the binding of amino acids to homopolymer and of the binding of ribonucleoside 5′-monophosphate to immobilized L-amino acids are not consistent with the present codon dictionary in any simple manner. Thus, if one wants to continue to insist on a physical-chemical basis for the genetic code it will require great speculation on the nature of an early primitive code, the proteins it coded for and how it evolved into the present code. Of course it may well be that the major physical interaction between amino acids and nucleic acid responsible for the NMR line shifts is due to a complex that is of no importance in terms of selection for the genetic code but that a different albeit less prominent form of complex exists on which evolution did operate and which is masked by the major complex studied.

Although the NMR technique has not solved the problem of the origin of the genetic code, extensions of it to the study of nucleic acid-peptide interactions may prove very useful in understanding the nature of protein-nucleic acid interactions in general (for example, repressors, initiators and polymerases).

References

1. Crick, F. H. C.: J. molec. Biol. **38**, 367 (1968).
2. Fox, S., Lacey, J., Yubi, A.: Abstract, Joint Conference Chem. Inst. of Canada and Am. Chem. Soc., Toronto, May 1970.
3. Héléne, C.: Nature (Lond.) New Biol. **234**, 120 (1971).
4. Lacey, J. C., Pruitt, K. M.: Nature (Lond.) **223**, 799 (1969).
5. Leng, M., Felsenfeld, G.: Proc. nat. Acad. Sci. (Wash.) **56**, 1325 (1966).
6. Orgel, L. E.: J. molec. Biol. **38**, 381 (1968).
7. Raszka, M., Mandel, M.: Proc. nat. Acad. Sci. (Wash.) **68**, 1190 (1971).
8. Raszka, M., Mandel, M.: First European Biophysics Congress Baden, Sept. 1971.
9. Saxinger, W. C., Ponnamperuma, C.: J. molec. Evolution **1**, 63 (1971).
10. Woese, C. R.: The genetic code. New York: Harper and Row 1967.
11. Woese, C. R.: J. molec. Biol. **43**, 235 (1969).
12. Woese, C. R., Dugre, D. H., Saxinger, W. C., Dugre, S. A.: Proc. nat. Acad. Sci. (Wash.) **55**, 966 (1966).
13. Zubay, G., Doty, P.: Biochim. biophys. Acta (Amst.) **29**, 47 (1958).

M. Mandel
Department of Biochemistry and
Biophysics, University of Hawaii
Honolulu, Hawaii 96822, USA

M. Raszka
Department of Chemistry
University of California, San Diego
La Jolla, California 92037, USA