AMINOACYL-tRNA SYNTHETASE FAMILIES AND THEIR SIGNIFICANCE TO THE ORIGIN OF THE GENETIC CODE

RONALD WETZEL

Dept of Molecular Biophysics and Biochemistry, Yale University, Box 1937 Yale Station, New Haven, Conn. 06520

(Received 30 January, 1978)

Abstract. A correlation of various aspects of the protein structures and substrate and mechanistic specificities of the aminoacyl-tRNA synthetases has led to the identification of at least one family of enzymes probably derived from a common ancestral synthetase. While strong correlations exist only in one part of the array of 64 codons comprising the Genetic Code, this itself may be interpreted as a meaningful pattern, most consistent with a development of the present code from earlier codes containing fewer amino acids and fewer available codons. Specifically, strong correlations in the enzymes whose cognate tRNAs respond to codons containing a central pyrimidine, including the enzyme family of Ile-, Phe-, Val-, Met-, and Leu-tRNA synthetases, suggests that these enzymes evolved last, and that, therefore, an earlier version of the Genetic Code was comprised solely of codons containing a central purine. It is suggested that further study of the historical interrelationships of these enzymes could lead to a fairly detailed picture of how the Genetic Code developed.

1. Introduction

Questions regarding the origin and development of the genetic code have remained a great challenge to experimental and theoretical biologists in the years since the code itself was elucidated. For a time experimental knowledge was limited to the correspondence between amino acids and codons, and this had to serve as the sole basis for speculation (Woese, 1967; Crick, 1968; Orgel, 1968). By the early seventies, much more information was available in the form of nucleotide sequences of many of the transfer RNAs. Phylogenic trees of tRNA were constructed based on relative sequence homologies (Cedergren, 1972; Holmquist *et al.*, 1973; Schwartz *et al.*, 1976), and other analyses of the sequence data led to theories of the code's development based on anticodon patterns (Jukes, 1975; Barricelli, 1977). A model has also been proposed based on known amino acid biosynthetic pathways (Wong, 1975). While much of this information is pertinent, no consistent scheme for the development of the existing genetic code has emerged. It is, perhaps, naive to expect nature to have left us any record of these ancient events fossilized in present biological systems.

On the other hand, not all the available information has been collected. The aminoacyl-tRNA synthetases (Söll and Schimmel, 1974; Kisselev and Favorova, 1974; Ofengand, 1977) are firmly enmeshed in the history of the code, being the intermediaries between the amino acids and the codons of that array. In the past five years much diverse information has accumulated concerning these enzymes and their relationships with the tRNAs. This data, while incomplete, reveals some remarkable patterns which may reflect

Origins of Life 9 (1978) 39–50. *All Rights Reserved*
Copyright © 1978 by D. Reidel Publishing Company, Dordrecht, Holland

aspects of the origin of the genetic code, and, more importantly, indicate promising avenues for further research in the field.

2. The Data

Most of the currently available significant data is summarized in Table I. To facilitate later discussion, the data is presented in groups of enzymes related to the codon groups their tRNAs respond to, as determined by the central nucleotide in the codon.

Macromolecular structural data is shown only for the *E. coli* enzymes. While such information may reflect common ancestry (Koch *et al.*, 1974), much of it is too equivocal to be the sole basis of any interpretation. It is known, for example, that a facile proteolysis can occur during the isolation of some of these enzymes, and this might lead

TABLE I
Some properties of aminoacyl-tRNA synthetases

Codon group	Amino acid	Enzyme quaternary structure[a] and subunit molecular weights (K)		2′/3′ charging specificity		Cognate tRNA discriminator	
XUX:	Leu	α	105	2′	(2′)	A	(A)
	Ile	α	112	2′	(2′)	A	(A)
	Val	α	110	2′	(2′)	A	(A)
	Met	α_2	86	2′	(2′)	A	(A)
	Phe	$\alpha_2\beta_2$	94, 39	2′	(2′)	A	(A)
XCX:	Ser	α_2	48	3′	(3′)	G	(G)
	Pro	α_2	47	3′	(3′)	–	(C, A)
	Thr	α_2	76	3′	(3′)	A	(A)
	Ala	–	–	3′	(3′)	A	(A)
XAX:	Tyr	α_2	48	2′, 3′	(2′, 3′)	A	(A)
	His	α_2	43	3′	(3′)	C	
	Gln	α	69	?	(3′)	G	
	Asn	–	–	2′, 3′	(2′, 3′)	G	
	Lys	α_2	52	3′	(3′)	A	(G, U)
	Asp	–	–	3′	(2′, 3′; 3′)	G	(G)
	Glu	$\alpha\beta$	56, 46	2′		G	(G)
XGX:	Cys	–	–	2′, 3′	(2′, 3′; 3′)	U	(U)
	Trp	α_2	37	2′, 3′	(3′)	G	(A)
	Arg	α	75	2′	(2′)	A	(G, U)
	Ser	α_2	48	3′	(3′)	G	(G)
	Gly	$\alpha_2\beta_2$	80, 33	3′	(3′)	U	(A)

Molecular structural properties of aminoacyl-tRNA synthestases, from *E. coli* only, are listed (Söll and Schimmel, 1974; Kisselev and Favorova, 1974; Ofengand, 1977; Hennecke *et al.*, 1977). The 2′/3′ specificity (Hecht, 1977) and discriminator (Sprinzl and Walker, 1978; Crothers *et al.*, 1972) data for the *E. coli* enzymes are listed first in each case, followed, in parentheses, by available information for yeast or mammalian systems. If the property is the same in both other systems, it is only listed once in parentheses. The value of A for the discriminator of proline tRNA was obtained for tRNA coded on the genome of bacteriophage T4.

[a]Defined as follows: α – single subunit; α_2 – dimer of identical subunits; $\alpha\beta$ – dimer of non-identical subunits; $\alpha_2\beta_2$ – tetramer of non-identical subunits.

to incorrect assignment of molecular weight and quaternary structure. Furthermore, it has been established that some aminoacyl-tRNA synthetases seem to have undergone in their recent past a gene duplication/fusion event (Koch *et al.*, 1974; Waterson and Konigsberg, 1974; Bruton, 1975; Kula, 1973) which would complicate attempts to deduce enzyme lineages from knowledge of structural families. These problems may, however, be limited to certain of the enzymes; in fact, it is interesting that all the large sequence homologies have been reported in enzymes, which, as will be seen, are likely to be related historically.

The 2′/3′ charging specificity data of Table I arise from experiments designed to determine which isomer of aminoacyl-tRNA, the 2′ or 3′ ester, is formed in the enzyme-catalyzed reaction. The method involved substitution of deoxyadenosines for the 3′-terminal adenosine in unfractionated tRNAs to generate two preparations in which tRNAs contained 3′ ends with adenosines lacking either a 2′ or 3′ hydroxyl group. The ability of each set of analogs to be charged by a crude synthetase preparation with each of the 20 amino acids was then measured (Sprinzl and Cramer, 1975; Hecht, 1977). These results were at first surprising in that the preferences for each synthetase were remarkably well-preserved over widely differing species. This is important since it suggests that the patterns in the present system may have been preserved from a much earlier, common ancestor. In the meantime, some doubt has been cast on the mechanistic meaning of the results (i.e., they may not always reflect a preference in the aminoacylation reaction, but rather a preference in the corrective hydrolysis step (von der Haar and Cramer, 1976)), but it is clear that any patterns in the results are likely to derive from some property of the enzymes themselves. This is because a native aminoacyl-tRNA, upon release by the synthetase into solution, can rapidly equilibrate to a mixture of 2′ and 3′ isomers (Wolfenden *et al.*, 1964; Griffin *et al.*, 1966). Thus, there seems to be no possibility of the operation of some unknown organizing principle after aminoacylation which would have selected for 2′ or 3′ specificity in enzymes evolving independently. A remaining question is whether the patterns, once established, were retained by accident or because the site of aminoacylation reflects an undiscovered aspect of aminoacyl-tRNA synthetases which continues to provide selective advantage, as suggested by Hecht (1977).

The special importance of the fourth base from the 3′ end of tRNA, the so-called 'discriminator' position, in aminoacylation was first pointed out by Crothers *et al.* (1972), who demonstrated a correlation between the identity of the discriminator and the chemical structure of the amino acid charged. These ideas derived from the apparent importance of this base in the recognition of the various tRNA species which could be aminoacylated by yeast Phe-tRNA synthetase (Roe *et al.*, 1972). This hypothesis was largely confirmed and somewhat qualified by subsequent experiments. Mutagenesis of phages containing an amber suppressor $tRNA^{Tyr}$ ($\phi 80su_3^+$) gene produced a new suppressor tRNA which could be mischarged with glutamine; the new suppressor proved to be derived from the old by a single A → G conversion at the discriminator site (Shimura *et al.*, 1972; Smith and Celis, 1973). In addition, data from the laboratories of Dudock (Roe *et al.*, 1973), Yarus (Yarus and Mertes, 1973) and Ebel (Ebel *et al.*, 1973;

TABLE II
Most easily mischarged tRNA species in total tRNA

Enzymes:	E. coli Val RS		Yeast Val RS		Yeast Phe RS		E. coli Ile RS		Yeast Asp RS	
Source of tRNA	Yeast[a]	E. coli[b]	E. coli[a]	Yeast[b]	E. coli[a]	Yeast[b]	Yeast	E. coli[a]	E. coli[b]	Yeast[b]
tRNA Cognate for:	Val Ile Met Thr Phe	Val Met Ala	Val Met Ala Pro Ile Thr Leu Phe	Val	Ala Val Phe	Phe Tyr Met Ala Val	–	Ile Met Arg Trp Val Tyr Lys Phe Pro	Asp His Glu Gln Ser Asn	Asp

Experiments were done under normal (a) or perturbed (b, 20% dimethylsulfoxide) conditions. Data from Ebel *et al.* (1973) and Yarus and Mertes (1973).

Giege *et al.*, 1974) showed that misaminoacylation (*via* tRNA mis-recognition) occurs almost exclusively within families defined by the identity of the discriminator base; that is, an enzyme whose cognate tRNA contains A at the fourth position will only misaminoacylate other tRNAs also containing A at that position (see Table II and discussion below). Ebel and colleagues (1973) also made the important point that the discriminator seems to have less to do with any first-order synthetase-tRNA recognition than with proper alignment of the C–C–A end of the already-bound tRNA.

It is useful to view this discriminator data as reflecting a property of the aminoacyl-tRNA synthetases. It has been observed that the data on tRNA mischarging seems to indicate families of enzymes as well as families of tRNAs (Giege *et al.*, 1974; Yarus and Mertes, 1973). This is, the presence of a common recognition element in the tRNAs suggests that their enzymes share a common 'complementary' structural feature required for productive binding, and this feature can only respond properly to the correct nucleotide in the correct orientation. For this reason the discriminator data is presented as an aspect of synthetase structure in Table I. The fidelity with which this property of tRNAs has also generally been maintained across species lines should be noted. Crothers *et al.* (1972) were the first to point out the possible connection between 'discriminator families' and the evolution of the genetic code.

Table II is a list of those tRNAs of total unfractionated tRNA from one source which can be aminoacylated with a particular amino acid and its purified cognate aminoacyl-tRNA synthetase isolated from the same (homologous) or a different (heterologous) source (Ebel *et al.*, 1973; Giege *et al.*, 1974; Yarus and Mertes, 1973). The results were obtained through enzyme assays under standard as well as perturbed (+ 20% dimethylsulfoxide) conditions. It has been observed (Ebel *et al.*, 1973; Yarus and Mertes, 1973) that such data really summarize two aspects of synthetase-tRNA interaction, namely, the ability of a particular tRNA to bind to the enzyme, and the rate of aminoacylation once the tRNA is bound. In many cases the second aspect is what, in fact, determines correct or incorrect aminoacylation, and it is also this rate effect which seems to be associated with the discriminator base, as can be seen by comparing the groups in Table II with discriminator bases from Table I. Nonetheless, it is equally clear that a tRNA cannot be aminoacylated if it is not bound, and binding of tRNAs to synthetases is not universally strong. Thus, a misaminoacylation event is a positive indication of binding as well as correct alignment at the active site, while a lack of misaminoacylation is more equivocal in the absence of supplementary information.

3. Patterns

A brief glance at Table I reveals the existence of at least one family of synthetases. The enzymes specific for leucine, isoleucine, valine and methionine are remarkably similar in subunit molecular weights and, as mentioned earlier, are all known to possess long sequence repeat units. Their mechanistic indicators are also identical, in a charging specificity for the 2′ hydroxyl and a requirement for an adenosine at the discriminator

site of their cognate (and non-cognate, Table II) tRNAs. There is also a high degree of structural similarity in the amino acids these enzymes bind and activate. While more complex in quaternary structure, it is clear that the enzyme for phenylalanine might also easily belong in this family. This is further supported by a recent report of repeated sequences within both the α and β subunits of yeast Phe-tRNA synthetase (Robbe-Saul *et al.*, 1977). With the exception of the $tRNA^{Leu}$ species, the tRNAs for all of these enzymes can be misaminoacylated by the enzymes for valine, isoleucine and phenylalanine (Table II). Their common membership in the group of enzymes whose cognate tRNAs respond to codons containing a central uridine (Figure 1), which may well be

CODON					
1 st letter	2nd letter				3 rd letter
	U	C	A	G	
U	PHE	SER	TYR	CYS	U
	PHE	SER	TYR	CYS	C
	LEU	SER	C.T.	C.T.	A
	LEU	SER	C.T.	TRP	G
C	LEU	PRO	HIS	ARG	U
	LEU	PRO	HIS	ARG	C
	LEU	PRO	GLN	ARG	A
	LEU	PRO	GLN	ARG	G
A	ILE	THR	ASN	SER	U
	ILE	THR	ASN	SER	C
	ILE	THR	LYS	ARG	A
	MET	THR	LYS	ARG	G
G	VAL	ALA	ASP	GLY	U
	VAL	ALA	ASP	GLY	C
	VAL	ALA	GLU	GLY	A
	VAL	ALA	GLU	GLY	G

Fig. 1. The Genetic Code. 'C.T.' stands for 'chain terminator'.

significant, facilitates identification of this family, but it would have stood out in any case.

Other families are more difficult to identify. The enzymes for the small, neutral amino acids in the XCX codon group share a preference for 3′-aminoacylation, but there are at least two discriminators in this group, and the similarity in enzyme quaternary structures cannot be given much weight. The case is more hopeless within the XAX and XGX codon groups, but then, there is no *a priori* reason to feel compelled to include all members of a codon group in a family, or for that matter to look for siblings only within one codon group. Even with the removal of these constraints, however, there is little solid evidence for other extensive families.

It is worth noting some broader patterns in Table I. It is clear, for example, that the enzymes whose cognate tRNAs respond to codons containing central purines (XRX) differ among themselves a great deal more than the enzymes in the central pyrimidine (XYX) groups. The only enzymes which show no charging specificity lie in the XRX subdivision of the genetic code (Figure 1). Every possible discriminator base can be found in the tRNAs of the XRX groups, including the only tRNAs which contain U or C at this position (with the exception of a yeast $tRNA^{Pro}$). The XYX groups, on the other hand, are distinguished by the preponderance of tRNAs with adenosine discriminators.

4. Discussion

The first question to be decided in tackling the origin of the gentic code is the number of amino acids initially coded. If all or most of the amino acids which are utilized today in protein biosynthesis were used in the first organism, the degree of complexity needed in the first protein biosynthesis system would be immense, and rationalization of its *ex nihilo* appearance, correspondingly difficult. If the first code contained all 20 amino acids, then the aminoacyl-tRNA synthetases as we know them must have evolved independently, or from enzymes of lower amino acid specificity. The presence of the XUX family rules out the possibility of completely independent evolution. On the other hand, there is no evidence at present which speaks against a theory that most of the other synthetases have existed in some form since the beginning. Nor is there rigorous proof against a first code consisting of all 20 amino acids, some of which shared the same aminoacyl-tRNA synthetase.

The alternative approach to the development of the code has been to assume that the first working system was much simpler, composed of fewer elements, at least fewer amino acids and cognate enzymes. Through the operation of certain forces more amino acids would later be entered into the code and given new tRNAs and synthetases until the code in its present form was reached, at which point it became inalterable. This approach was first described, and termed the 'Frozen Accident Theory', in a paper by Crick (1968). In his paper Crick laid down the ground rules under which such a development could occur. Few amino acids may have been coded in the first code, but all possible codons must have been read. New amino acids could be introduced into the system by the construction of

new activating enzymes and new tRNAs. Some codons previously coding for an old amino acid would at this point take on new meaning, and the consequent blanket substitution in all the proteins in the cell would have to confer more advantages than disadvantages onto the cellular machinery for the substitution to be retained by natural selection.

As pointed out be Crick (1968), the close structural similarity in the amino acids of the XUX codon group could be explained by this kind of development. If a new amino acid is to be systematically introduced in response to specific 'acquired codons' in all proteins, the effect of this replacement would be least felt if the charge and size of the new amino acid did not differ greatly from the one it replaced. Furthermore, the simplest way for a new activating enzyme to arise is through gene duplication and mutation in a gene for a pre-existing enzyme. Again, the most likely parent of a new enzyme arising in this fashion is one possessing only slightly different amino acid specificity.

The data presented in Table I support this basic model for the development of the code. It shows that, especially in the case of the aliphatic amino acids, there is evidence for the generation of a family of enzymes through divergent evolution from a common ancestral synthetase. One can imagine an early code in which, for example, all the codons with central uridines coded for isoleucine, through an Ile-tRNA synthetase which could bind and aminoacylate all of the tRNAs with anticodons which respond to these codons. At some later time, a duplication of the gene for this enzyme, followed by mutation(s) in one of the copies, would yield an altered synthetase which could better bind one of the other amino acids in this group, as well as a subset of the group's tRNAs. Repetition of this process would eventually lead to a family of enzymes related in mechanistic and structural features.

There are several difficulties with the model as it stands at this point. First, the XUX family is the only group to show clear evidence of common origins, even if one is allowed to range freely over the codon array of Figure 1 in search of sibling enzymes. Does this mean that all the other enzymes, plus the parent of the XUX family, existed in the first coding system? This would leave us little farther than where we started in searching for a realistic initial code. There is an additional problem, in any case, that of the tRNAs. The model as stated requires enough tRNAs in the first organism to code for all 61 sense codons. While the point has been made that it may not have been so difficult for a system randomly generating polynucleotides to produce reasonable tRNA molecules (Orgel, 1968), it is a considerable burden, nonetheless, to require so many at the first level of life. Furthermore, one as yet ignored aspect of the family of leucine, isoleucine, valine, methionine and phenylalanine enzymes is that they do in fact fall into the same codon group. To say that this is because the parent enzyme happened to recognize the corresponding family of tRNAs, and only that family, is to beg the question. How did the relationship between the tRNAs themselves come about? Many synthetases do not require an intact anticodon for recognition of cognate tRNAs. There seems to have been some earlier historical connection between some of the tRNAs.

All of these problems can be dealt with by the idea of the expanding code. The basic

principles of such a process are: The first code, while still a triplet code, is restricted to only certain triplets, say, for example, 16 of the present-day codons. This requires considerably fewer tRNAs, and, at the same time, allows a sufficient number of amino acids to be involved. At a later time some of the restrictions on codon composition are removed, and new codons begin to appear in mRNAs. These 'nonsense' codons can be overcome by 'suppressor' tRNA molecules derived from mutations in the anticodons of duplications of existing tRNAs. Soon a new code, for example of 32 codons, is in operation; at first the newer tRNAs are still charged by the same enzyme as their parent tRNAs. If, however, a new amino acid is made available to the cell, and if a duplicated gene of one of the synthetases can be mutated to create an enzyme specific for the new amino acid, the new enzyme would have available to it some of the tRNAs previously charged by the parent enzyme. In this way the genetic code would eventually expand to the full 64 codons, through a two-step process, in which new tRNAs first are created to respond to new codons, after which new enzymes arise to take advantage of the increased number of tRNAs and the availability of new amino acids.

The idea of an initially restricted code was first mentioned by Crick in his 1968 paper. More recent models (Hartman, 1975; Crick *et al.*, 1976; Barricelli, 1977) for the origin of protein synthesis and the genetic code also lead to proposed early codes which are restricted in size. All these theories demand early codes in which the only allowed codons are those with purines as the first two nucleotides; the first expansion in all but Hartman's model would most likely be to a code of 32 triplets with central purines. It has been observed that the amino acids in this region of the present-day code are those which are likely to have been the earliest available (Crick *et al.*, 1976).

Any scheme in which the genetic code is initially confined to this bottom right-hand corner of Figure 1 is supported by the evidence compiled in Table I. In such a system, all proteins synthesized by a restricted code would obviously be quite primitive in comparison to their modern-day counterparts. The amino acid substitutions inherent in an expansion of an early code of seven amino acids to the final twenty would so alter the properties of original proteins that few subtleties of mechanism or structure would likely survive in their contemporary equivalents. Thus, regardless of whether the aminoacyl-tRNA synthetases of this early code themselves arose independently or from some earlier, common ancestor, they will probably not bear much resemblance to their ancestors or to each other by the time they are exposed to biochemists. As was pointed out in an earlier section, the XRX groups of synthetases are chiefly distinguished by their internal diversity.

If the genetic code developed in a series of stages through expansion from some initially restricted set of codons, the last enzymes to arise in this process would be the ones most likely to retain evidence of common origins. The enzymes in the XUX family clearly fit this description, and this is precisely what one would expect based on the proposed theoretical schemes (Crick, 1968; Hartman, 1975; Crick *et al.*, 1976; Barricelli, 1977).

There exists other evidence which might be brought to bear on this point: phylogenic

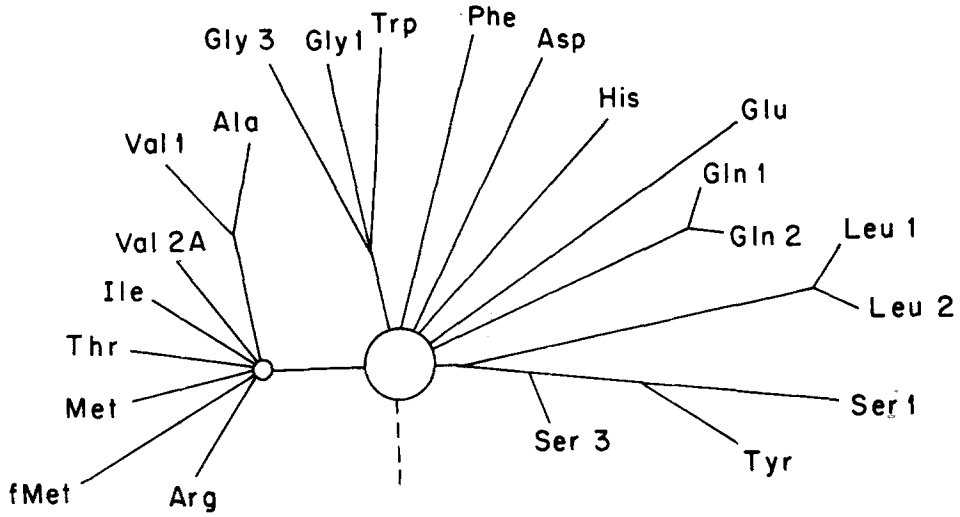

Fig. 2. Evolutionary tree of E. coli transfer RNA. Adapted from Schwartz *et al.*, 1976.

trees of tRNA evolution (Cedergren, 1972; Schwartz *et al.*, 1976) based on sequence homology in these molecules. Nucleotide sequences of tRNA molecules have diverged so much, unfortunately, that it is difficult to construct a trustworthy phylogeny (Holmquist *et al.*, 1973). Nonetheless, the most detailed proposed phylogenic tree to date, shown in Figure 2, is supported by some independent data. Some of the tRNAs (fMet, Met, Ile, Val 1, Val 2) of the XUX codon group are found in a common branch, along with several of the tRNAs (Ala, Thr, Arg) from other codon groups which are found to be easily mischarged by enzymes of the XUX family (Table II). Furthermore, the $tRNA^{Leu}$ species, which surprisingly did not appear in the mischarged groups of Table II, seem to have a separate history from the other XUX tRNAs, more related to the $tRNA^{Ser}$ species. It should be pointed out that the proposed two-step process of the code's expansion allows for some scrambling to occur. For example, present day $tRNA^{Leu}$ species may have entered the code as mutated $tRNA^{Ser}$ molecules, either before or after the appearance of a Leu-tRNA synthetase.

Given a general idea of the way in which the code evolved, knowledge of sequence homologies and phylogenic trees of tRNAs might thus provide some fascinating details of the process. All things considered, however, it seems that information relating to the question of common origins among the synthetases will probably be more useful in elaborating the development of the genetic code.

5. Conclusions

The patterns of Table I, both the similarities within the XUX and XCX codon groups, and the relative variety within the XAX and XGX groups, suggest that the aminoacyl-tRNA synthetases in the XYX half of the present Genetic Code evolved after at least most of

the enzymes in the XRX half. If an earlier code used only codons containing a central purine, later expansions of the code, with XYX enzymes evolving from a few XRX progenitors, would be expected to lead to the kind of distribution of enzyme similarities observed.

While the correlations presented in this work are consistent with this interpretation, the patterns offer no direct proof. Nonetheless, they do provide solid evidence for the existence of at least one enzyme family whose members probably share a common ancestor. This fact in turn suggests routes toward a better understanding of the history of the genetic code. If a number of additional structural and mechanistic criteria can be successfully applied to the aminoacyl-tRNA synthetases, some as yet hidden enzyme inter-relationships might be revealed. Theories of the code's development could be tested against these patterns.

If the code developed through an expansion of available codons as suggested here, it should be possible to learn some further details such as the immediate ancestor of the XUX enzymes, and the extent of inter-relatedness in the XCX group. On the other hand, strong evidence for the earliest amino acids used in a code may not be so readily accessible. Not only will details of mechanism probably have changed in their enzymes as they exist today, but the likelihood of finding sequence homologies, even in mechanistically important peptides, would be quite small, especially since so many changes in codons and codon meanings would have taken place in the interim.

Acknowledgements

I wish to thank Drs D. G. Söll, D. M. Crothers, J. A. Rafalski and J. Kohli for helpful discussions and critical reading of the manuscript. This work was supported by a National Institutes of Health grant GM 22854 to Dr D. G. Söll.

References

Barricelli, N. A.: 1977, *J. Theoret. Biol.* **67**, 85.
Bruton, C. J.: 1975, *Biochem. J.* **147**, 191.
Cedergren, R. J.: 1972, *J. Theoret. Biol.* **37**, 209.
Crick, F. H. C.: 1968, *J. Mol. Biol.* **38**, 367.
Crick, F. H. C., Brenner, S., Klug, A., and Pieczenik, G.: 1976, *Origins of Life* **7**, 389.
Crothers, D. M., Seno, T., and Söll, D. G.: 1972, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3063.
Ebel, J. P., Giege, R., Bonnet, J., Kern, D., Befort, N., Bollack, C., Fasiolo, F., Gangloff, J., and Dirheimer, G.: 1973, *Biochimie* **55**, 547.
Giege, R., Kern, D., Ebel, J. P., Grosjean, H., Henau, S. D., and Chantrenne, H.: 1974, *Eur. J. Biochem.* **45**, 351.
Griffin, B. E., Jarman, M., Reese, C. B., Sulston, J. E., and Trentham, D. R.: 1966, *Biochemistry* **5**, 3638.
von der Haar, F., and Cramer, F.: 1976, *Biochemistry* **15**, 4131.
Hartman, H.: 1975, *Origins of Life* **6**, 423.
Hecht, S. M.: 1977, *Tetrahedron* **33**, 1671.
Hennecke, H., Böck, A., Thomale, J., and Nass, G.: 1977, *J. Bact.* **131**, 943.
Holmquist, R., Jukes, T. H., and Pangburn, S.: 1973, *J. Mol. Biol.* **78**, 91.
Jukes, T. H.: 1975, *Nature* **246**, 22.

Kisselev, L. L., and Favorova, O. O.: 1974, *Adv. Enzymol.* **40**, 141.
Koch, G. L. E., Boulanger, Y., and Hartley, B. S.: 1974, *Nature* **249**, 316.
Kula, M. R.: 1973, *FEBS Lett.* **35**, 299.
Ofengand, J.: 1977, *Molecular Mechanisms of Protein Biosynthesis*, H. Weissbach and S. Pestka (eds.), Academic Press, New York, p. 7.
Orgel, L. E.: 1968, *J. Mol. Biol.* **38**, 381.
Robbe-Saul, S., Fasiolo, F., and Boulanger, Y.: 1977, *FEBS Lett.* **84**, 57.
Roe, B., Sirover, M., and Dudock, B.: 1972, *Fed. Proc.* **31**, 450 Abstr.
Roe, B., Sirover, M., and Dudock, B.: 1973, *Biochemistry* **12**, 4146.
Schwartz, R. M., Gantt, M. J., and Dayhoff, M. O.: 1976, *Atlas of Protein Sequence and Structure*, M. O. Dayhoff (ed.), Natl Biomed. Res. Found., Washington, D.C., Vol. 5, Supplement 2, 271.
Shimura, Y., Aono, H. and Ozeki, H.: 1972, *FEBS Lett.* **22**, 144.
Smith, J. D., and Celis, J. E.: 1973. *Nature New Biol.* **243**, 66.
Söll, D. G. and Schimmel, P. R.: 1974, *The Enzymes*, P. D. Boyer (ed.), 3rd ed., Vol. X, Academic Press, New York, p. 489.
Sprinzl, M. and Cramer, F.: 1975, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3049.
Sprinzl, M. and Walker. R. T.: 1978, *Nucleic Acids Res.* **5**, in press.
Waterson, R. M. and Konigsberg, W. H.: 1974, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 376.
Woese, C. R.: 1967, *The Genetic Code*, Harper and Row, New York.
Wolfenden, R., Rammler, D. H., and Lipmann, F.: 1964, *Biochemistry* **3**, 329.
Wong, J. T.: 1975, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1909.
Yarus, M. and Mertes, M.: 1973, *J. Biol. Chem.* **248**, 6744.