6. J. C. Jamieson, in *ibid.*, p. 444.
7. R. H. Christian, *University of California, Lawrence Radiation Laboratory, Livermore, California, UCRL-4900* (1957); B. J. Alder, in *Solids Under Pressure*, W. Paul and D. M. Warschauer, Eds. (McGraw-Hill, New York, 1963), p. 385; *University of California, Lawrence Radiation Laboratory Report UCRL-12473*, p. 61 (1965).
8. J. A. Corll and G. A. Samara, in preparation.
9. Private communications from D. McWhan and I. Borg; see also references 13, 16, and 17 listed in Jamieson's report (*6*).
10. J. C. Jamieson and A. W. Lawson, *J. Appl. Phys.* **33**, 776 (1962).
11. D. B. McWhan and W. L. Bond, *Rev. Sci. Instr.* **35**, 626 (1964).
12. ASTM powder pattern No. 6-0555.
13. V. V. Evdokimova and L. F. Vereshchagin, *Fiz. Tverd. Tela* **2**, 1701 (1960).
14. I thank I. Borg, D. Larson, D. McWhan, G. Smith, D. Stephens, and H. Stromberg for suggestions and comments. This research was done under the auspices of the U.S. Atomic Energy Commission.

5 May 1966

Genetic Code: Aspects of Organization

Abstract. *The pattern of organization of the genetic code decreases to a minimum the phenotypic effects of mutation and of base-pairing errors in protein synthesis. Single base changes, especially transitions, usually cause either no amino acid change or the change to a chemically similar amino acid. The degree of degeneracy of the codons for an amino acid is correlated with their guanine-cytosine content. The code gives greater protection (by both degeneracy and guanine-cytosine content of codons) to those amino acids that appear more frequently in proteins. Increased reliability of the protein-synthesis system afforded by this pattern of organization may have determined the fitness of the present code.*

Schroedinger (*1*) has pointed out that a fundamental characteristic of any mechanism of inheritance must be unusual stability in the face of natural randomizing influences. We now present evidence that the biochemical system for gene expression, the genetic code, is organized to stabilize the phenotype by lessening the effects of mutational processes. Although many elements of the code remain to be elucidated, it has been shown (*2*) that the distribution of codons is nonrandom (*2–5*). We now suggest that the pattern of the present code protects the organism against the consequences of mutation. In addition, the code minimizes the consequences of base-pairing errors occurring in the transcription and the translation of the information in DNA. Ambiguity in translation can be demonstrated in cells and extracts; it is further increased by mutations or drugs that alter the ribosome (*6*). Such noninherited errors were probably even more conspicuous during early stages of the evolution of mechanisms of protein synthesis (*7*). A code capable of buffering the effects of these errors would increase the reliability of the entire system for gene expression and thus be of selective advantage.

Table 1. RNA codon assignments. The RNA codon assignments are those designated principally by Nirenberg and by Khorana, and their co-workers (*2*). A question mark denotes incomplete evidence.

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

The pattern of codon assignment can protect against the phenotypic effects of mutations or reading errors in two ways: (i) degeneracy such that the new triplet still corresponds to the original amino acid, and (ii) codon arrangements such that the new triplet specifies an amino acid whose substitution in the protein does not affect its function. In addition, the base composition of triplets may directly influence the relative rates of errors, and thus the more stable codons may serve a protective function.

Sonneborn (*5*) has reviewed the selective advantages inherent in a highly degenerate code in which different codons for the same amino acid differ in only one nucleotide of the triplet. If each of the 20 amino acids had its own unique codon, and if the remaining 44 base combinations did not specify amino acids, most one-step mutations would lead to nonsense and therefore to unfinished polypeptide chains. In this way, a degenerate code featuring a minimum number of nonsense codons, probably only those essential for punctuation, would serve significantly to decrease the rate of appearance of a mutant phenotype. Furthermore, if the triplets corresponding to the same amino acid share at least two bases in common, then one-step mutations change the identity of the specified amino acids as seldom as possible. Every amino acid appears (Table 1) degenerate at least twice in just this fashion (*8*).

The twice-degenerate amino acids are all of the form *ab purine* or *ab pyrimidine*. This pattern of codon assignment provides the maximum possible protection against mutation in a twice-degenerate system. It affords some protection against transitions (substitution of a purine for a purine, or of a pyrimidine for a pyrimidine) but none against transversions (substitution of a purine for a pyrimidine or a pyrimidine for a purine). Similarly, amino acids having four codons would be most protected if their four degenerate triplets conform to the set *abx*, where *x* is any base (*9*). In such an arrangement, one-third of all transitions and transversions remain within the group. This pattern of degeneracy characterizes all the amino acids that have been shown to be four-times degenerate. Finally, two amino acids that appear to have sixfold degeneracy, Arg and Leu (*10*), have codons of the form *abx* and *a'b purine*. This arrangement is exactly that distribution of six codons which insures that a random base change in the set is least likely to produce a codon not of the set. Serine is also probably six-times degenerate, but of the form *abx, a'b pyrimidine* (*2*). Although highly protective, such an arrangement makes serine the one exception to the generalization that the form of degeneracy minimizes the frequency of base substitutions that lead to different amino acids.

The genetic significance of this pattern of degeneracy is summarized in Table 2, which presents the relative frequencies of new amino acids appearing as a result of single base changes. In these calculations we assume that all one-step mutations are equally likely and that all possible codons for a given amino acid appear in equal amounts. The diagonal represents the frequency with which a given amino acid mutates to itself. A comparison of the diagonals of *A* and *B* of Table 2 shows that the arrangement of the code gives greater insurance

against the effects of transitions than against transversions. This arrangement would be most protective if, as currently believed, transitions are the more frequent mutational events (*11*). In addition, the most frequent base-pairing errors arising in transcription or translation are likely to be phenotypically equivalent to transitions in the DNA.

Replacement of one amino acid by another often does not affect protein function (*12*). Thus the effects of a base change depend both upon the degree of similarity between the new and original amino acid and upon the position in the protein where the replacement occurs. On this basis alone one would anticipate that certain sites in a gene would appear as point mutations far more frequently than others (*13*).

If the code has been selected so that the rate of appearance of a mutant phenotype is as small as possible, amino acids with similar properties should have similar codons. Such an arrangement has in fact been noted (*2–4*) and is evident in Table 1. Woese (*4*) has further shown that amino acids with similar R_F values in a variety of chromatographic solvents frequently have similar codons. We now show quantitatively that the arrangement of the code tends to increase to a maximum the frequency with which amino acids mutate to similar amino acids. The chief difficulty in such a study is in setting criteria of amino acid similarity.

Many amino acids tend to form hydrophobic bonds. The importance of such bonding was emphasized by Kauzmann (*14*), who predicted that the most stable polypeptide conformations have the nonpolar side chains juxtaposed in the interior of the molecule to the exclusion of water. The studies on myoglobin by Kendrew (*15*) and on hemoglobin by Perutz *et al.* (*16*) have supported these views. Thus, exchanges of hydrophobic and hydrophilic amino acids would result most frequently in a protein with highly altered properties. Examination of cytochrome *c* from a variety of species has even led to the suggestion (*17*) that "the entities conserved in evolution are the clusters of hydrophobicity, rather than the individual residues."

One measure of hydrophobic character is the side-chain contribution to the free-energy change when an amino acid is transferred from a nonpolar solvent to water (*18*). On this basis, the hydrophobic amino acids are those for which the free-energy change exceeds that of Gly. In Fig. 1 (bottom) the tendency of each amino acid to mutate to a hydrophobic residue is plotted against its degree of hydrophobic character. The curve shows that the pattern of codons quite effectively prevents mutation of a hydrophobic to a

Table 2. Relative frequencies of new amino acids as a result of a change of one base. *A*. Amino acid changes resulting from all possible single nucleotide substitutions. In *B*, only transitions are considered. The values represent the percentage of all single base changes of the original amino acid that produce codons corresponding to the new amino acid. For example, of the 36 possible base changes of the codons for Ala, 12 (33 percent) do not change the identity of the amino acid and four (11 percent) yield codons of Gly. For each amino acid all codons are used with equal weight in the calculation, and all single base changes are considered equally likely. The amino acids are placed in order of increasing R_F value, averaged from data in six different organic solvents (*4*). This arrangement assures that amino acids contiguous on the chart have similar physical properties. The clustering of frequencies, and particularly those with higher values, about the diagonal is an indication of the tendency of amino acids to mutate to similar amino acids.

Original amino acid	New amino acid																			
	Asp	Glu	Lys	Arg	Asn	Gln	Gly	Cys	Ser	His	Thr	Ala	Pro	Val	Met	Tyr	Ilu	Leu	Try	Phe	Trm
A. Transitions and transversions																					
Asp	11	22			11		11			11		11		11		11					
Glu	22	11	11			11	11					11		11							11
Lys		11	11	11	22	11					11				11						11
Arg			3	33		3	11	3	11	3	3		7		3			7	7		
Asn	11		22		11				11	11	11					11	11				
Gln		11	11	11		11				22			11					11			11
Gly	5	5		16			33	5	5			11		11					5		
Cys				11			11	11	22							11			22	11	
Ser				11	3		3	7	25		11	7	7			3	3	3	3	3	3
His	11			11	11	22				11			11			11		11			
Thr			5	5	5				16		33	11	11		5		5				
Ala	5	5					11		11		11	33	11	11							
Pro				11		5			11	5	11	11	33					11			
Val	5	5					11					11		33	5		5	16		5	
Met			11	11							11			11	11		22	22			
Tyr	11				11			11	11	11						11				11	22
Ilu					11				11		11			11	22		11	11		11	
Leu				7		3			3	3			7	11	7		3	33	3	11	3
Try				22			11	22	11									11	11		11
Phe								11	11					11		11	11	33		11	
Trm		11	11			11			11							22		11	11		11
B. Transitions																					
Asp	33				33		33														
Glu		33	33				33														
Lys		33	33	33																	
Arg			11	33		11	11	11		11									11		
Asn	33				33				33												
Gln				33		33															33
Gly	16	16		16			33		16												
Cys				33				33								33					
Ser					11		11		33				22					11		11	
His				33						33						33					
Thr											33	33			16		16				
Ala											33	33		33							
Pro									33				33					33			
Val												33		33	16		16				
Met											33			33	33						
Tyr								33		33						33					
Ilu											33			33			33				
Leu									11				22					55		11	
Try				33															33		33
Phe									33									33		33	
Trm						33													33		33

hydrophilic amino acid and vice versa. For the hydrophobic residues as a group, the frequency of remaining in the group approaches 90 percent, which far exceeds the fraction of total codons belonging to this group (68 percent). The curve also suggests that, among the hydrophilic amino acids, increasingly hydrophobic character is associated with an increased tendency to mutate to a member of the hydrophobic group.

The correlation is even more striking if only transitions are considered (Fig. 1, top). Here, with the exception of Tyr, hydrophobic amino acids mutate to others of their group with a frequency of 100 percent. In addition, the exclusion of transversions decreases even further the frequency with which a given hydrophilic amino acid mutates to the hydrophobic type. Thus the code seems to be organized so that the exchange of hydrophobic and hydrophilic amino acids as a result of one-step mutations is minimum.

A further grouping of the similar amino acids subdivides the hydrophobic residues into two classes according to the length of their side chains (Table 3*A*). Those nonpolar amino acids with side chains containing at most two carbon atoms (Gly, Ala, Ser, Thr, and Cys) constitute a class intermediate in character between the extreme hydrophobic and the hydrophilic groups. As is evident in Table 3*A*, the frequency with which any amino acid mutates to others in the same group far exceeds that expected of random codon assignments. Furthermore, from Table 2, when the most hydrophobic or the hydrophilic residues do mutate out of their group, they generally go to amino acids of the intermediate class.

A more refined classification would take into account the character of certain amino acids and the special similarities between others. The unique properties of Cys and Pro suggest that these residues are frequently irreplaceable. Further considerations of side-chain structure give the eight groupings shown in Table 3*B*. Again amino acids appear to mutate to codons of the same group far more frequently than expected from random assignment of codons.

We have seen that stability of the phenotype is promoted both by degeneracy and by the relations between the codons for different amino acids. These two contributions can be distinguished by considering the forms of the parts of Table 2. These charts are arranged so that contiguous amino acids are most similar in physical properties, as based on chromatographic criteria (*4*). A degenerate code, of the form discussed, produces in each case a nonzero value for each point on the diagonal. If the pattern of codon assignment does insure that similar amino acids mutate most often to each other, then a graph of this form would show considerable clustering of the higher frequencies about the diagonal. Such clustering is evident in *A* and even more so in *B*, which shows only the effects of transitions. A similar result is obtained if other criteria of amino acid similarity (such as the free energy changes, or simple structural considerations) are employed.

Genetic Significance. The protective organization discussed above may explain certain earlier observations on the nature of mutation. Orgel and Brenner (*19*) and Freese (*20*) have shown that most detected spontaneous mutations in T4 phage arise by nucleotide addition or subtraction (that is, reading frame shifts) and not by single nucleotide substitutions. This difference in frequency may reflect in part the pattern of the code. Undoubtedly most replacements of a single base are not phenotypically recognizable; frame shifts, however, so drastically alter amino acid sequence that essentially every such mutation is expressed.

Similarly, the greater protection which the code affords against the phenotypic expression of transitions means that a larger fraction of transversions will be phenotypically recognizable. Even though transitions occur far more frequently (*11*), examination of naturally occurring changes of amino acids, in the light of our knowledge of the code, shows that a large number of these changes can be accounted for only by transversions. For example, of the 17 human hemoglobinopathies (*21*) that can be explained by single base changes, seven could have arisen only by transversions. However, since these altered hemoglobin molecules were identified electrophoretically, they represent mutations that have produced a change in net protein charge; from Table 2, it is not surprising to find a large number of transversions among mutations that alter amino acid charge.

Base Content of Codons. For the

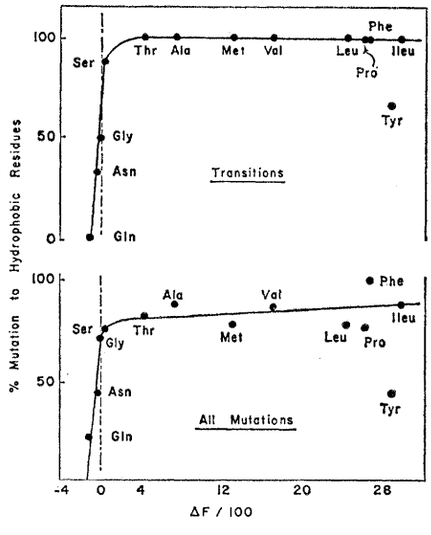

Fig. 1. (Top) The change in free energy of the side chain of a given amino acid, upon transfer from a hydrophobic environment to water (*18*), is plotted against the calculated relative frequency with which the amino acid mutates to the hydrophobic amino acids by single base changes in its codons. It is assumed that all codons occur with equal frequency and that all mutations are equally likely. The amino acids not included are those for which no direct thermodynamic data are available. (Bottom) The change in free energy of the chain of a given amino acid is plotted against the calculated relative frequency with which the amino acid mutates to the hydrophobic amino acids by transition mutations of its codons. The free energy change of Gly is taken as zero. The assumptions are as outlined in Table 2*A*.

Table 3. Mutational pattern of groups of similar amino acids. For each group of similar amino acids, the percentage of all possible mutations which lead to codons of the same group is shown in column 1; column 2 shows only the proportion of all possible transitions. Values in column 3 represent the percentage expected from a random distribution of codons without changing the number of codons in each group (that is, the number of codons in the group, 1/63). The criteria for grouping amino acids are discussed in the text. Calculations are based on values in Table 2, and are based on the assumption that all codons appear with equal frequency.

Type	Mutations (%)	Transitions	Expected from random code (%)
	A. Group		
Hydrophilic (acidic, basic, amidic)	55	70	27
Intermediate (Gly, Ala, Ser, Thr, Cys)	57	53	30
Hydrophobic	59	66	37
	B. Similar amino acids		
Pro	33	33	5
Try, Try, Phe, Ilu, Leu, Met, Val	59	59	30
Cys	11	33	2
Thr, Ser	42	33	14
Gly, Ala	44	33	11
Gln, Asn	11	33	5
Glu, Asp	33	33	5
His, Arg, Lys	33	59	14

Table 4. Relation of protection to frequency of appearance of amino acid in protein. The degree of protection afforded an amino acid and the frequency of the residue in total microbial protein are compared (*25, 26*). The amino acids are divided into "best protected" and "least protected" groups, where protection is defined by GC content and degeneracy. Asn and Try also fall into the least-protected category, but frequency data are not available for these residues. The figure for Trm (not included in the average) represents a maximum estimate, under the assumption that Trm codons appear several times at the end of each peptide chain. Values for Met are probably somewhat higher, since many Met residues, though coded, may not appear in proteins (*30*).

Least protected		Best protected	
Amino acid	Frequency (%)	Amino acid	Frequency (%)
Met	1.7	Ala	11.2
Tyr	2.5	Gly	9.0
Phe	3.5	Leu*	8.7
Lys	5.7	Arg*	5.0
Ileu	5.1	Pro	4.7
Trm	(<3)		
Av.	3.7	Av.	7.7

* Six times degenerate.

different codon sets the base composition and the extent of degeneracy are closely related. By a codon set we mean the group of triplets *abx, ab purine,* or *ab pyrimidine,* which code for the same amino acid. Triplets fall into three classes, according to the nucleotides in the first two positions (*22*): (i) those with only G and C (100 percent GC); (ii) those with one G or C and one A or U (50 percent GC); (iii) those with only A and U (no GC). Examination of Fig. 1 shows that all codon sets with maximum GC content are four-times degenerate; all those lacking GC are twice degenerate; and the sets with an intermediate GC content have an average degree of degeneracy of about three.

The basis of this correlation between GC content and degeneracy is not clear, but a crucial factor is probably the additional hydrogen bond linking GC pairs, as compared with AT pairs. The importance of GC content for the thermal stability of the DNA helix is well known. Possibly the GC content of the triplet set may actually determine the degeneracy, the greater affinity of GC pairs obviating reading of the third nucleotide in those triplets containing only GC in the first two places. Alternatively or in addition, the present correlation may have arisen from selective pressures acting independently on GC content and on degeneracy to achieve maximum protection of certain amino acids; for example, the greater affinity of GC pairs might mean that triplets with a high GC content have a greater pairing specificity and greater resistance to mutation (*23*). A similar suggestion has been advanced for short nucleotide sequences (*24*).

Basis of Protection. The correlations described suggest that protective mechanisms may act at two levels: (i) the nucleotide level, where a high GC content may reduce the rate of mutations and errors in protein synthesis, and (ii) the organizational level where the effects of a base change are minimized by degeneracy and by the connectedness of codons for similar amino acids. The best-protected amino acids would be those with maximum GC content and degeneracy.

If the code has evolved so as to preserve faithfulness in expression and inheritance of the genome, degeneracy and stability should be greater for the codons of those residues that appear more frequently in proteins, for these amino acids will be the ones most often affected by random mutational events and base-pairing errors. However, additional protection should be afforded those amino acids that cannot generally be replaced in protein structures without drastic alteration of function.

In Table 4, the "most-protected" and the "least-protected" classes of amino acids are compared according to their relative frequency in total microbial proteins (*25, 26*). The "most-protected" group of amino acids (those six-times degenerate, or four-times degenerate plus maximum GC) are clearly more common on the average than the "least-protected" ones (twice-degenerate and no GC). Those residues best protected are the preponderant ones in proteins with the exception of Pro and Arg. The unique role of Pro in polypeptide conformations provides an explanation for its added protection. The case of Arg is discussed below.

Included in the "least-protected" group are three of the five least common amino acids and also the chain terminator, which must be among the rarest codons in the genome. Significantly, the two least common residues, Cys and His, both of which appear of special importance in enzyme structure, do not fall in the least-protected class; both have codons containing 50 percent GC. With these two exceptions, the twice-degenerate codons containing 50 percent GC appear more frequently than those lacking GC.

Further evidence of a relationship between the organization of the code and the composition of proteins is seen by grouping similar amino acids and comparing the quantitative importance of these groups in proteins with the relative number of codons allotted to the group. If such a comparison is made for the individual amino acids, no clear relation is seen. Figure 2 shows, however, that the number of codons assigned to a group of similar amino acids (see Table 3*B*) is directly related to their relative frequency in proteins. For example, in the basic group the apparent overprotection of Arg compensates for the underprotection of Lys. An even stronger correlation is obtained on grouping similar amino acids into the three general classes of Table 3*A*.

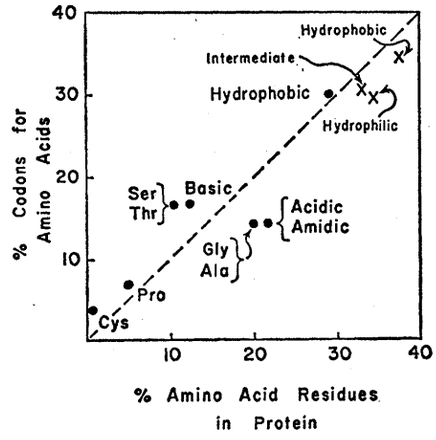

Fig. 2. The correlation between relative proportions of all codons corresponding to a group of amino acids and the relative frequency of appearance of those amino acids in proteins. ● denote groups of amino acids defined in Table 3*B; X* represent the groups of Table 3*A*. Not included in the hydrophobic group is Try, since data on its frequency of appearance are unavailable. The frequency data are taken from the work of Sueoka (*25*) and Fitch (*26*) and represent the average values for seven species of protista. The total number of codons has been taken as 60 (Trm and Try omitted). The dotted line represents the line of identity. Thus the organization of the code and the composition of cellular proteins appears closely related.

Evolution of the Code. The organization of the code thus reflects the composition of cellular proteins in two ways: groups of chemically similar amino acids have similar codons; and the degree of protection afforded a group, as measured by the number of codons allotted to it, is proportional to the frequency of appearance of its amino acids in proteins. In this way the code appears particularly well suited to preserve the pheno-

type in the face of natural randomizing influences. Indeed, among the codes that could be generated by a reassignment of codons, it is doubtful that any would be substantially more protective than the present one. In fact, if the GC content and degree of degeneracy of each amino acid is kept constant, then Table 1 shows that reassignment of codon sets cannot improve the protection of hydrophobic and hydrophilic character.

A different explanation (*4, 27*) for the organization of the code is that it arose through the influence of specific interactions of amino acids and nucleotides before the development of the adaptor RNA. No such chemical affinity has been demonstrated, and such a mechanism cannot adequately explain the fitness of the code to protect amino acid groups in relation to their quantitative importance in proteins.

Sonneborn (*5*) has suggested that protection against lethal mutation may have been the major determinant in the evolution of the code. We suggest, however, that increased accuracy of the systems for gene expression may have been the chief selective influence. The efficiency of cellular processes ultimately depends on the accurate synthesis of enzymes; hence a protective code, apart from its effect on mutation rate, would be advantageous in the same way as a better ribosome or a better sRNA would be. A similar suggestion has been made by Woese (*28*).

Initially, adaptors and the corresponding code were probably nonspecific, recognizing only general aspects of amino acid character. Evolutionary pressures for better enzymes would have selected not only the appropriate amino acid sequences, but also a system for translation that best insured their synthesis. The evolution of the code was thus the development of a larger number of adaptors with greater specificity (*29*) such that related amino acids became associated with closely related adaptors derived from an earlier, less specific one. In this way a genetically protective pattern of coding emerged because of selective pressures for better mechanisms of protein synthesis.

ALFRED L. GOLDBERG
ROBERT E. WITTES
Department of Bacteriology and Immunology, Harvard Medical School, Boston, Massachusetts

References and Notes

1. E. Schroedinger, *What Is Life?* (Cambridge Univ. Press, Cambridge, 1944).
2. R. Brimacombe, J. Trupin, M. Nirenberg, P. Leder, M. Bernfield, T. Jaouni, *Proc. Nat. Acad. Sci. U.S.* **54**, 954 (1965); D. Soll, E. Ohtsuka, D. S. Jones, R. Lohrmann, H. Nayatsu, S. Nishimura, H. G. Khorana, *ibid.*, p. 1378 (1965).
3. S. R. Pelc, *Nature* **207**, 597 (1965).
4. C. R. Woese, *Proc. Nat. Acad. Sci. U.S.* **54**, 71 (1965).
5. T. M. Sonneborn, in *Evolving Genes and Proteins*, V. Bryson and H. J. Vogel, Eds. (Academic Press, New York, 1965), p. 377.
6. S. M. Friedman and I. B. Weinstein, *Proc. Nat. Acad. Sci. U.S.* **52**, 988 (1964); J. Davies, W. Gilbert, L. Gorini, *ibid.* **51**, 883 (1965); W. Szer and S. Ochoa, *J. Mol. Biol.* **8**, 823 (1964).
7. See suggestion of A. Rich, in *Evolving Genes and Proteins*, V. Bryson and H. J. Vogel, Eds. (Academic Press, New York, 1965), p. 453.
8. The observed pattern of degeneracy may reflect two distinct mechanisms: (i) lack of complete specificity of the sRNA in reading the third letter, or (ii) the existence of many different codon-specific sRNA molecules. To what extent the two mechanisms apply remains to be seen. Although the two possibilities have different implications with respect to evolution of the code, both have the same genetic effect in stabilizing the phenotype and do not affect our argument.
9. Another protective arrangement of four codons would be of the form *abc*, *a'bc*, *abc'*, *a'bc'* and where *a* and *a'* and *b* and *b'* are related by transitions. This pattern would give no protection against transversions, but would afford twice the protection against transitions that the present arrangement does. A code based on this pattern, however, does not exhibit certain correlations characteristic of the present code (see discussion of GC).
10. Abbreviations are G, guanine; C, cytosine; A, adenine; U, uracil; T, thymine; Asp, aspartic acid; Glu, glutamic acid; Lys, lysine; Arg, arginine; Asn, asparagine; Gln, glutamine; Gly, glycine; Cys, cystine; Ser, serine; His, histidine; Thr, threonine; Ala, alanine; Pro, proline; Val, valine; Met, methionine; Tyr, tyrosine; Ilu, isoleucine; Leu, leucine; Try, tryptophane; Phe, phenylalanine; Trm, terminating codon (amber and ochre.).
11. D. Krieg, *Progr. Nucleic Acid Res.* **2**, 125 (1963).
12. C. Yanofsky, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 581 (1963).
13. S. Brenner and A. O. W. Stretton, *J. Cell. Comp. Physiol.* **64**, suppl., 43 (1964).
14. W. Kauzmann, in *The Mechanism of Enzyme Action*, W. D. McElroy and B. Glass, Ed. (Johns Hopkins Univ. Press, Baltimore, 1954).
15. J. C. Kendrew, *Brookhaven Symp. Biol.* **15**, 216 (1962).
16. M. F. Perutz, J. C. Kendrew, H. C. Watson, *J. Mol. Biol.* **13**, 669 (1965).
17. E. Margoliash and E. L. Smith, in *Evolving Genes and Proteins*, V. Bryson and H. J. Vogel, Eds. (Academic Press, New York, 1965), p. 221.
18. C. Tanford, *J. Amer. Chem. Soc.* **84**, 4240 (1962).
19. A. Orgel and S. Brenner, *J. Mol. Biol.* **3**, 762 (1961).
20. E. Freese, *Proc. Nat. Acad. Sci. U.S.* **45**, 622 (1959).
21. J. B. Stanbury, J. B. Wyngaarden, D. Fredrickson, Eds., *The Metabolic Basis of Inherited Disease* (McGraw-Hill, New York, ed. 2, 1966).
22. The third position is ignored because each codon set contains at least one G or C and one A or U in that position.
23. Since the completion of this manuscript, Martin and Hoyer have reported [*Fed. Proc. Abstr.* **25**, 779 (1966)] the results of agar hybridization with DNA from two different rodents. They have shown that the regions of greatest affinity are those with a high GC content, while the regions of less affinity are lower in GC. These results are consistent with our suggestion that the GC content of codons may play a protective role by increasing genetic stability.
24. E. K. F. Bautz, in *Evolving Genes and Proteins*, V. Bryson and H. J. Vogel, Eds. (Academic Press, New York, 1965), p. 419.
25. N. Sueoka, *Cold Spring Harbor Symp. Quant. Biol.* **26**, 35 (1961).
26. W. M. Fitch, *Proc. Nat. Acad. Sci. U.S.* **52**, 298 (1964).
27. S. R. Pelc and M. G. E. Welton, *Nature* **209**, 868 (1966); M. G. E. Welton and S. R. Pelc, *ibid.*, p. 870.
28. C. R. Woese, *Proc. Nat. Acad. Sci. U.S.* **54**, 1546 (1965).
29. T. H. Jukes, *Amer. Sci.* **53**, 477 (1965).
30. M. Cappechi and J. Adams, personal communication.
31. We thank Dr. J. Beckwith and Professor B. Davis for advice and encouragement. Discussion with Dr. F. Crick has been rewarding. Supported by NSF GB-1307 to B. D. Davis.

15 April 1966

Ozone and Sulfur Dioxide Synergism: Injury to Tobacco Plants

Abstract. *Tobacco plants displayed ozone-type injury when exposed to mixtures of ozone and sulfur dioxide at subthreshold concentrations. The syndrome suggests synergism between ozone and sulfur dioxide that lowers thresholds to injury; exposure to the individual gases at the mixed-gas concentrations caused no symptoms.*

Certain varieties of tobacco are the most sensitive indicators known of the presence of air-polluting ozone, usually displaying symptoms in the field when concentrations exceed 0.05 ppm (parts per million) (*1, 2*). Development of automated methods for measuring ozone, including those based on oxidation of potassium iodide, has facilitated determination of ozone thresholds. However, we have observed and others (*1*) have reported symptoms resembling ozone fleck after indication of concentrations of ozone, in ambient air, as low as 2 parts per 100 million (pphm). Sulfur dioxide causes 100-percent interference with potassium iodide ozone analysis (*3*), although ozone and sulfur dioxide coexist without antagonism in the gaseous phase at concentrations below 1 ppm (*4*). Interference by sulfur dioxide is prevented by placement of a chromium trioxide scrubber at the beginning of the air-sampling train (*3*). Concentrations of sulfur dioxide tend to be higher in late spring and fall than in summer (*5*). We believe that ozone thresholds determined in the field while sulfur dioxide is present are incorrect because of interference by the sulfur dioxide with the measurement of ozone. Furthermore, the apparently subthreshold concentrations of ozone that caused injury suggested synergistic effect by ozone and sulfur dioxide. The suspicion of synergism is based on the disruption of an indicated balance of oxidized and reduced sulfur in plant