*CORRESPONDENCE BETWEEN GENETIC DATA AND THE POSITION OF AMINO ACID ALTERATION IN A PROTEIN**

By Donald R. Helinski† and Charles Yanofsky

DEPARTMENT OF BIOLOGICAL SCIENCES, STANFORD UNIVERSITY

Communicated by Victor C. Twitty, December 22, 1961

The development of methods for the rapid detection of primary structure differences between proteins of similar structure has greatly facilitated studies on the effect of mutations on the primary structure of a protein.[1] Investigations with normal and abnormal forms of human hemoglobin have clearly established single amino acid differences between these proteins.[2] Similarly, some induced mutants of tobacco mosaic virus form proteins containing one to several amino acid changes.[3] Techniques for determining genetic fine structure have also been highly developed, and it is now possible to prepare detailed genetic maps of specific genes.[4] These advances in protein chemistry and genetic analysis permit an examination of the gene-protein relationship at the fine-structure levels of both gene and protein, and, accordingly, several microbial gene-protein systems are being studied at this level.[5–] The present report is concerned with a study of one of these systems, the A gene-A protein system of the tryptophan synthetase of *Escherichia coli*. The experiments described in this paper deal with an examination of the primary structure of mutationally altered A proteins. The results obtained indicate that at least in two groups of mutants there is a correlation between the region of the gene altered by mutation and the location of the amino acid change in the corresponding altered A protein.

Materials and Methods.—The A proteins examined were isolated by the same or slight modifications of the procedure used for the purification of the wild-type A protein.[8] Proteolytic digestion and peptide pattern examination of the mutant proteins were performed as described elsewhere.[9] The trypsin and chymotrypsin preparations used were obtained from Worthington Biochemical Corp. In experiments with the A proteins of mutants A-23 and A-11, all of the protein preparations were heated for 5 min at 100°C prior to proteolytic digestion in order to inactivate trace amounts of other proteolytic enzyme(s) that occasionally were present as contaminants.[9] Amino acid analyses were carried out with a Spinco amino acid analyzer. The peptides were hydrolyzed in twice-distilled 5.7 *N* HCl in sealed, evacuated tubes at 105°C for 48 hr.

Characteristics of Mutants Examined.—The various mutant strains relevant to this study and the sites of their mutational alterations in the A gene are shown in Figure 1. The mutants listed were isolated in the K-12 strain of *E. coli* following ultraviolet irradiation.[10, 11] All of these mutants form altered A proteins that are fully active in the conversion of indole and serine to tryptophan in the presence of the B protein of tryptophan synthetase.[11] However, these altered A proteins are incapable of participating in the physiologically essential reaction, the conversion of indoleglycerol phosphate and serine to tryptophan. The A proteins of many of the mutants listed in Figure 1 have been shown to be distinguishable from each other and from the wild-type A protein on the basis of their physical properties.[11] Mutants A-1, A-23, A-28, A-35, and A-36 form heat-labile A proteins, while

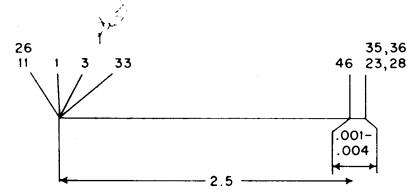

FIG. 1.—Sites of mutational alterations in the A gene in the mutants examined. Sites and distances determined by transduction with phage Plkc.[20] The relative order of A-46 and the A-23 group is not known.

mutants A-3, A-33, and A-46 form A proteins that are more heat-stable than the wild-type A protein. The A proteins of mutants A-11 and A-26 are precipitated at pH 4 while the A proteins of the other mutants and the wild-type A protein do not precipitate at this pH. Mutants A-3 and A-33 have been shown to be distinguishable in other tests;[11] A-3 reverts while A-33 does not, and A-3 is suppressible by a gene that does not affect A-33. Mutants A-11 and A-26 cannot be distinguished in reversion and suppression studies (as well as the tests already mentioned), and the same is true for mutants A-23, A-28, A-35, and A-36.[11, 12]

In deciding whether mutants that appear to map at the same site represent repeat or nonidentical changes at the same mutational site, the sensitivity of the test used to detect genetic separability is obviously of considerable importance. In studies with the A mutants, *cys-tryp* double mutants were used as recipients in transduction tests to increase the sensitivity of the recombination test.[11] It can be calculated that the use of *cys-tryp* double mutants as recipients in transduction recombination experiments should permit the detection of recombination frequencies of 0.0001 per cent before reversion in the donor or the recipient needs to be considered as an interfering factor. Previous studies using this method with many of the mutants listed at identical sites in Figure 1 have shown that there is no recombination at the 0.001 per cent level.[11] However, prototrophic recombinants were detected in these transduction experiments with A-46 and mutants of the A-23 type. The recombination frequencies detected were very low, ranging from 0.0003 to 0.002 per cent in different experiments. More extensive data have now been obtained, and the recombination frequency between mutants A-46 and A-23 has been estimated at 0.001 to 0.004 per cent. This estimate is based on the recovery of only 8 tryptophan-independent recombinants in transductions involving $> 4 \times 10^5$ possible events. The total recombinational length of the A gene is approximately 2.5 per cent. Although mutants A-23 and A-46 map extremely close to each other, A-23 behaves aberrantly in recombination experiments with many other A mutants. For example, A-46 gives 0.21 per cent and 0.45 per cent recombination with mutants A-17 and A-7, respectively, while A-23 gives 1.2 per cent and 1.4 per cent with the same two mutants.

Comparison of Proteolytic Digests of the A Proteins of Mutants Mapping at the A-11 Site.—The A proteins of four mutants that map at the same site (A-3, A-11, A-26, and A-33) have been examined for peptide pattern differences from the wild-type A protein. Trypsin (T), chymotrypsin (C), and trypsin plus chymotrypsin (T + C) peptide patterns of the A protein of mutant A-3 did not show a clear difference from the corresponding wild-type protein peptide patterns, while peptide patterns of the A proteins of mutants A-11, A-26, and A-33 did show differences (Fig. 2).[13] The peptide patterns of the A-11, A-26, and A-33 mutant proteins

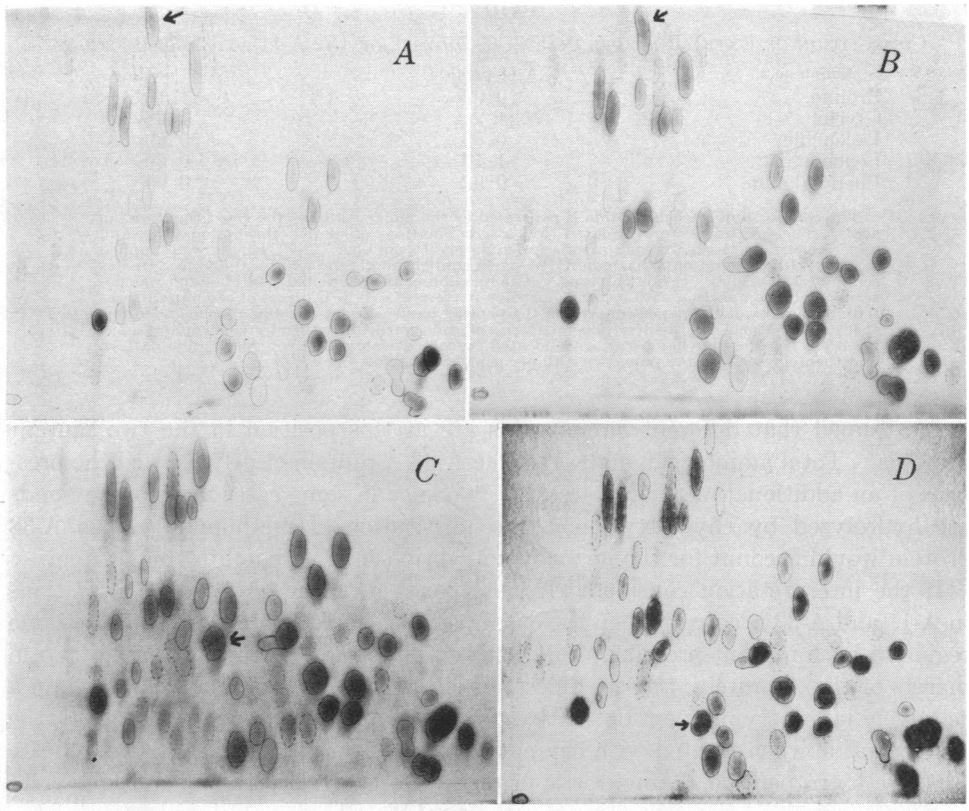

FIG. 2.—Trypsin plus chymotrypsin peptide patterns of the following A proteins: (*a*) A-11, (*b*) A-33, (*c*) A-23, and (*d*) wild-type. The arrows indicate the major peptides involved in the differences between the mutant and wild-type peptide patterns.

contained one peptide in addition to those observed in the corresponding wild-type pattern. This extra peptide was present in the same position in T + C or C peptide patterns of these mutant proteins. Amino acid analyses were carried out on the additional peptides in the T + C digests of the A proteins of mutants A-33 and A-11. The analyses have shown that these peptides have the same amino acid composition (Table 1). As illustrated in Figure 3, the presence of this additional peptide in the

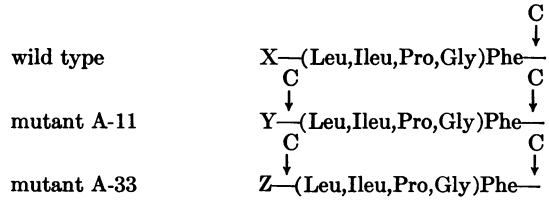

FIG. 3.—Interpretation of peptide alterations in A-33 and A-11. Chymotrypsin-sensitive bonds are designated by the letter C. X, Y, Z represent different amino acids.

A-11 and A-33 peptide patterns could be interpreted as being due to an amino acid change in the same position in the protein, resulting in a new chymotrypsin-sensitive bond. Since the A proteins of A-11 and A-33 have very different properties,

TABLE 1

COMPOSITION OF EXTRA PEPTIDES IN T + C DIGESTS OF THE A-11 AND A-33 PROTEINS

Amino acid	A-11 peptide	A-33 peptide
Proline	0.92	1.10
Glycine	1.21	1.28
Isoleucine	0.85	0.89
Leucine	1.10	0.80
Phenylalanine	0.92	0.94

The composition of each peptide is expressed as the molar ratios of the constituent amino acids. The peptides were isolated by a combination of paper chromatography and paper electrophoresis. In each case, approximately 15 mg of a T + C digest were applied as a 1 × 10 cm band on Whatman 3MM paper and chromatographed by the procedure described for obtaining peptide patterns.[9] The location of the particular peptide band was determined by comparison with a control peptide pattern of a T + C digest of the mutant protein. The area of the paper containing the peptide band was cut out and stitched to a new sheet of Whatman 3MM filter paper and the peptide purified further by electrophoresis at pH 3.7.[9] The peptide band was again located by comparison with a control peptide pattern and then eluted from the paper with 10 per cent acetic acid.

it is assumed that different amino acids are at this position in the two mutant proteins. Total amino acid analysis of the A-33 A protein clearly showed the presence of an additional methionine residue.[14] Since in some cases methionine bonds are hydrolyzed by chymotrypsin,[15] the substitution of methionine in the A-33 protein would account for the new chymotrypsin-sensitive bond in this protein.

If the interpretation considered for the position of the amino acid alterations in A-11 and A-33 is correct, then the extra peptide should be derived from the same region of both mutant proteins. Since the extra peptide is found in chymotrypsin digests of the A protein, trypsin digests of either mutant protein should contain a relatively larger tryptic peptide that contains the amino acid change and yields the extra peptide when treated with chymotrypsin. Several lines of evidence indicate that the extra peptide characteristic of digests of the A proteins of mutants A-11 and A-33 is derived from the same tryptic peptide. When trypsin digests of either of these two mutant proteins were fractionated on a Sephadex column, the tryptic peptide which gives rise to the extra peptide upon addition of chymotrypsin was found in the same fraction (Fig. 4). Furthermore, when these Sephadex fractions were further fractionated on DEAE-Sephadex columns, these tryptic peptides from both mutant digests were again found in approximately the same fractions.

Additional evidence for the similarity of the altered tryptic peptides from the A-11 and A-33 proteins was obtained in studies on the electrophoretic properties of these peptides. When trypsin digests of the A-11 and A-33 proteins were applied as a 1 × 2 cm band to Whatman 3MM filter paper and separated by high-voltage electrophoresis at pH 3.7,[9] the tryptic peptide that gives rise to the extra peptide upon chymotrypsin addition was found at the origin. The failure of this peptide to migrate electrophoretically and, as it was subsequently found, chromatographically, under the conditions used for fingerprinting accounts for the similarity of the T peptide patterns of A-11, A-33, and wild-type A proteins. Experiments are currently being carried out to further purify the relevant tryptic peptides from A-11, A-33, A-26, A-3, and wild-type A proteins.

The results of these studies with mutants A-11, A-26, and A-33 suggest that the amino acid changes in the A proteins of these strains are present at the same position. The fact that mutants A-11 and A-26 are indistinguishable in all tests performed to date further suggests that the A proteins of these particular mutants have the same amino acid change.

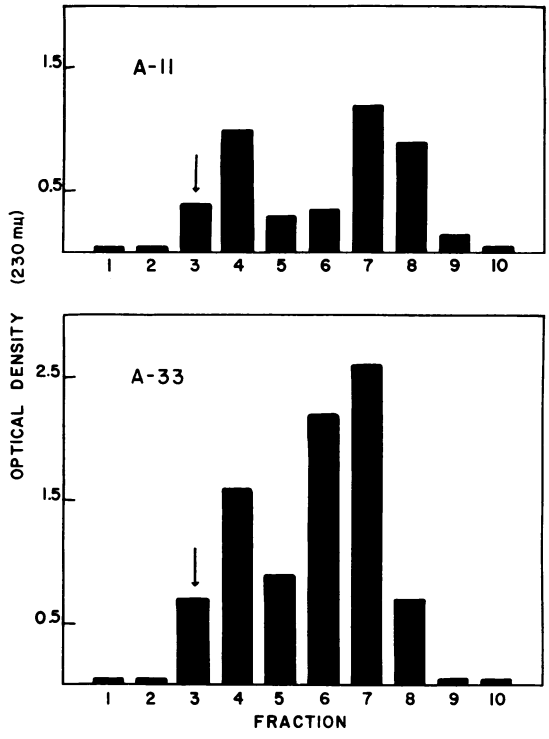

FIG. 4.—Elution patterns of trypsin digests of the A-11 and A-33 proteins. The arrows indicate the fractions containing the tryptic peptides altered in digests of the A-11 and A-33 proteins. The elution patterns were obtained by fractionation of 44 mg of a trypsin digest of A-33 protein and 30 mg of a trypsin digest of A-11 protein on a Sephadex G25 column (1.5 × 100 cm). The peptides were eluted with distilled water and located by measuring absorption at 230 mμ. The flow rate was approximately 10 ml per hr and fractions of 8–12 ml were collected. For the location of the altered tryptic peptides, an aliquot of each fraction was dried *in vacuo* and the peptides dissolved in 0.3 ml of 0.1 *M* ammonium carbonate, pH 8.3. Chymotrypsin was then added and the digestion mixture incubated at 25°C for 90 min. The digests were then dried *in vacuo* and peptide patterns obtained by the procedure described elsewhere.[9]

Studies on the Location of the Amino Acid Change in the A Proteins of Mutants Mapping at the A-23 Site.—The altered A proteins of four mutants, A-23, A-28, A-35, and A-36, all of which map at the same site and form heat-labile A proteins, have also been examined in peptide pattern studies. T + C peptide patterns of each of these mutant proteins showed an identical difference from the corresponding wild-type pattern. A T + C peptide pattern of the A protein of one of these mutants, A-23, is compared with a corresponding wild-type pattern in Figure 2. As was the case with mutants A-11, A-26, and A-33, the peptide pattern studies showed that the same peptide was altered in strains with mutational changes at the same site.

The corresponding mutant and wild-type peptides that differ in the A-23 and wild-type T + C peptide patterns were purified by paper electrophoresis and paper chromatography. The amino acid composition of these peptides is shown in Table 2. It is clear that the mutant peptide contains phenylalanine in addition to the nine amino acids in the wild-type peptide. The corresponding mutant

TABLE 2

COMPOSITION OF A-23 MUTANT PEPTIDE AND CORRESPONDING WILD-TYPE PEPTIDE FROM T + C DIGESTS

Amino Acid	A-23 peptide	Wild-type peptide
Lysine	1.02	1.05
Aspartic (amide)	0.92	0.94
Serine	0.82	0.96
Glutamic (amide)	1.03	1.04
Proline	1.12	1.03
Glycine	1.32	1.06
Alanine	1.02	1.03
Valine	0.93	0.95
Isoleucine	1.04	0.95
Phenylalanine	0.84	—

The composition of each peptide is expressed as the molar ratios of the constituent amino acids. The peptides were isolated by the procedure described in Table 1.

peptide was also isolated from T + C digests of the A-35 protein and was found to have the same composition as the A-23 mutant peptide. The amino acids in the corresponding mutant and normal peptides have been partially sequenced, and the sequences are identical with the exception that the extra phenylalanine is present in the N-terminal position in the mutant peptide.[16]

Further information on the nature of the amino acid change in mutant A-23 was obtained from peptide pattern studies with trypsin digests of the A-23 protein. Comparison of a T peptide pattern of the A-23 protein with the corresponding wild-type pattern clearly showed the absence of one peptide and the presence of two new peptides (Fig. 5). One of the new peptides was located in exactly the

FIG. 5.—A tracing of the major peptides of the peptide pattern of a trypsin digest of the A-23 protein superimposed over a tracing of the corresponding peptide pattern of the wild-type A protein. The major peptides of the A-23 tracing are outlined with a dashed line and the wild-type peptides are outlined with a solid line.

same position as the altered peptide in the T + C peptide pattern of A-23. The other new peptide overlaps a normally occurring peptide and was located by staining A-23 and wild-type peptide patterns for tyrosine-containing peptides.[17]

Since two peptides in the A-23 T peptide pattern replace a single wild-type peptide, an amino acid alteration resulting in the formation of a new trypsin-sensitive bond was considered for the A-23 protein. Such a possibility was also suggested by total amino acid analyses of the A-23 protein, which showed the presence of an additional arginine.[14] The wild-type tryptic peptide which is absent in trypsin digests of the A-23 protein has been isolated,[18] and its composition is shown in Figure 6. This information and the peptide pattern studies suggested the following amino acid sequence for the altered region of the A-23 protein: basic amino acid-phenylalanine-glycine; and X-phenylalanine-glycine for the correspond-

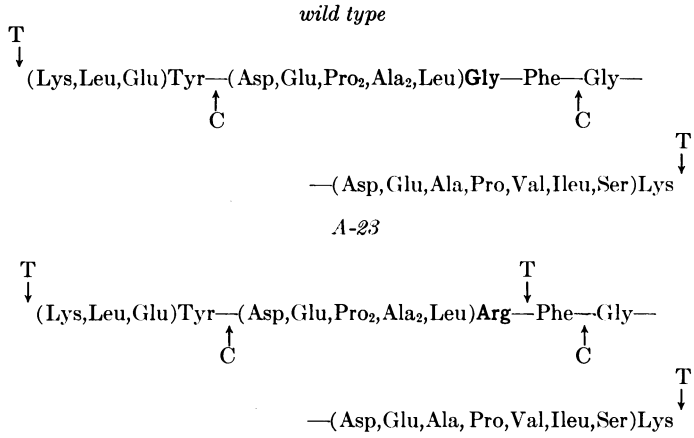

FIG. 6.—Corresponding regions of wild-type and A-23 proteins. The trypsin and chymotrypsin-sensitive bonds are designated by T and C, respectively.

ing wild-type region. The arginine-phenylalanine-glycine sequence in the mutant protein would account for the peptide pattern differences, since the addition of trypsin first, and then chymotrypsin, would result in the hydrolysis of the peptide bond between the arginine and phenylalanine yielding a peptide with an N-terminal phenylalanine. An N-terminal phenylalanine bond has been shown to be resistant to hydrolysis by chymotrypsin.[19] The corresponding sequence in the wild-type protein, X-phenylalanine-glycine, would not be attacked by trypsin but would be hydrolyzed between phenylalanine and glycine by chymotrypsin. The net result of the combined action of trypsin and chymotrypsin on this region of the mutant and wild-type proteins would be one mutant peptide which differs from the corresponding wild-type peptide by the presence of an N-terminal phenylalanine. This peptide would be found in either a T or T + C peptide pattern. In addition, a second mutant peptide would be expected and this peptide should contain a C-terminal arginine in place of an X-phenylalanine C-terminal sequence in the corresponding wild-type peptide.

Such an arginine-containing peptide on the T + C peptide pattern was indicated when a comparison was made of the A-23 and wild-type T + C peptide patterns which had been stained for arginine-containing peptides.[9] The new arginine-containing mutant peptide gave a yellow-brown ninhydrin color reaction and was located slightly below the position of a corresponding wild-type peptide that gives a similar ninhydrin reaction but does not contain arginine.

Since analyses of trypsin peptides of the A-23 protein were complicated by the presence of an additional trypsin-sensitive peptide bond, attempts were made to isolate the arginine-containing mutant peptide from chymotrypsin digests of the A-23 protein. Chymotrypsin peptide patterns of A-23 and wild-type A proteins were stained for arginine-containing peptides, and an additional arginine-containing peptide was found in the A-23 peptide pattern at a slightly lower chromatographic position than the corresponding chymotryptic peptide of the wild-type protein. This arginine-containing peptide was isolated by fractionation of a chymotrypsin digest of A-23 protein on a DEAE-Sephadex column. Amino acid

analyses were performed on this mutant peptide and the compositions of this peptide and the corresponding wild-type chymotryptic peptide are shown in Table 3.

TABLE 3

COMPOSITION OF THE CHYMOTRYPTIC MUTANT PEPTIDE OF A-23 AND THE CORRESPONDING WILD-TYPE PEPTIDE

Amino acid	Wild-type peptide*	A-23 peptide
Arginine	—	0.97
Aspartic (amide)	0.98	1.03
Glutamic (amide)	1.11	1.09
Proline	2.08	2.03
Glycine	1.05	—
Alanine	1.81	1.97
Leucine	1.07	0.99
Phenylalanine	0.92	0.90

The composition of each peptide is expressed as the molar ratios of the constituent amino acids. The A-23 peptide was isolated from a chymotrypsin digest of the A-23 protein by fractionation on a DEAE-Sephadex column. A solution of 38 mg of a chymotrypsin digest of the A-23 protein in 1 ml of pyridine-acetic buffer, pH 7.0 (0.02 *M* with respect to pyridine), was added to a 1.5 × 100 cm column of DEAE-Sephadex, A-50, which had been washed thoroughly with the pyridine-acetic buffer. A linear gradient was then applied to the column with 300 ml of the pyridine-acetic buffer in the mixing flask and 300 ml of 0.2 *M* acetic acid in the inlet flask. The flow rate was approximately 10 ml per hr and 8–12 ml fractions were collected. The fractionation was performed at 4°C. The fractions were subjected to paper electrophoresis at pH 3.7 and the peptide located by staining for arginine-containing peptides.[9] The mutant peptide was further purified by paper electrophoresis at pH 3.7.

* The analysis of this peptide was taken from the accompanying paper by Henning and Yanofsky.[18]

It is clear that the mutant peptide differs from the wild-type peptide by the substitution of an arginine for a glycine. The composition of this chymotryptic peptide and the peptide pattern information indicate that the sequence of amino acids in the altered region of the A-23 protein is as shown in Figure 6.

The A protein of mutant A-46 was also examined in peptide pattern studies. The mutational alteration in this strain maps very close to the alteration in strain A-23. No clear difference was observed between T or T + C peptide patterns of the A-46 and the wild-type A proteins.

Discussion.—Recombinational analyses and peptide pattern studies with a number of A protein mutants indicate that a correlation exists between the region of the genetic map at which a mutation occurs and the position of amino acid alteration in the A protein. Recombination studies with mutants A-11, A-33, and A-3 suggest that the mutational alterations in these strains are at the same site in the A gene. These mutants can be distinguished from one another in reversion[11, 12] and suppressor studies, and, in addition, the A protein of mutant A-11 has different properties from the A proteins of A-3 and A-33. Peptide pattern studies clearly indicate a difference between the A proteins of A-3 and A-11 or A-33. The peptide patterns of the A-11 and A-33 A proteins show the same peptide difference from the corresponding wild-type A protein patterns. The finding that the amino acid changes in the A proteins of A-11 and A-33 are probably in the same tryptic peptide and possibly at the same position in the A protein demonstrates a correlation between corresponding regions of the genetic map and the protein. Studies with the alkaline phosphatase system of *E. coli* have also suggested such a relationship.[5]

Mutant A-26 maps at the same site as A-11 and is indistinguishable from this mutant in reversion[12] and suppressor tests. The proteins of these two mutants have identical properties, and the peptide patterns of the two mutant proteins show the same difference from the corresponding wild-type patterns. These observations suggest that mutants A-11 and A-26 represent identical mutations. The same conclusion was derived from peptide pattern studies and other comparisons of four

mutants, A-23, A-28, A-35, and A-36, that map at one site at the other end of the A gene. Final proof that these mutants represent repeat identical mutations will require amino acid analyses of the mutant peptides of each protein digest.

Mutant A-46 gives tryp$^+$ recombinants with low frequency when crossed with strain A-23, indicating that the genetic alterations in these strains are close to each other. Although the peptide patterns of the A-46 protein are similar to the corresponding patterns of the wild-type protein, evidence is presented in the accompanying paper[18] that the amino acid substitution in the A-46 protein involves the same amino acid in the wild-type protein that is replaced in the A-23 protein. This amino acid, glycine, is replaced by glutamic acid in A-46 and by arginine in A-23.

In the Watson and Crick model of DNA, a nucleotide can be considered to be the smallest possible recombinable unit. Since mutants A-23 and A-46 are recombinable and have mutational changes involving the same amino acid in their A proteins, it can be concluded that at least two nucleotides are required to specify this amino acid in the A protein. Furthermore, the fact that only single amino acids are changed as a result of mutational events rules out, at least in these cases, overlapping nucleotide codes in which each nucleotide is involved in the coding of adjacent amino acids. The same conclusion can be derived from similar observations with other proteins.[2, 3] In view of the fact that the mutational alterations in A-23 and A-46 affect the same amino acid in the A protein, the recombination distance between these sites is of considerable importance. The recombination values obtained suggest that the distance between the A-23 and A-46 mutational alterations is approximately $^1/_{625}$ to $^1/_{2,500}$ of the total recombinational length of the A gene. This estimate can be compared with the length of the A gene in the following way: if three nucleotide pairs code for each of the 280 amino acids in the A protein and if all of the nucleotides are directly involved in specifying the amino acids of the A protein, then the A gene would have a length of about 840 nucleotide pairs. Adjacent nucleotide pairs, therefore, would be expected to give recombination values of $^1/_{840}$ of the recombinational length of the A gene. This value agrees reasonably well with the observed recombination distance between A-23 and A-46.

In the recombination experiments with mutants A-23, A-46, and the other A mutants, it was noted that A-23 behaves aberrantly, giving very high recombination values with mutants with alterations near the same end of the A gene. Since the A-23 mutation involves a single amino acid change and presumably a change in a single nucleotide pair, this aberrant behavior would suggest that recombination frequencies are markedly influenced by the nature of the nucleotide differences in recombining genic regions.

Although the detected alterations in the A-23 and A-46 proteins involve single amino acid substitutions, our present information does not exclude the possibility that other undetected changes have occurred or that some of the other ultraviolet-induced mutants have more complex amino acid changes. Only single peptide alterations have been observed with any of the mutant proteins examined to date, however, indicating that in each case there is only a localized alteration in the primary structure of the A protein. Total amino acid analyses of the mutant proteins are consistent with this conclusion.[14]

The substitution of an arginine or a glutamic acid for glycine as a result of point

mutations suggests a coding relationship between these three amino acids. If the mutational event in each of these two mutants involved a change of a single nucleotide, then it could be concluded that the composition of the coding units for arginine or glutamic acid does not differ from that of the coding unit for glycine by more than one nucleotide. A similar consideration can be applied to the A-11 region, where there are several different mutant types that map at or near the same site and, consequently, may have amino acid alterations in the same position in their A proteins. By examining many mutants at or near the same site, it should be possible to obtain useful information on the coding relationships between amino acids. If the code is degenerate, this approach will be particularly important.

It is apparent that it will be necessary to carry out similar studies on a number of mutants at different sites along the gene in order to conclusively test the concept of collinearity of amino acid sequence and nucleotide sequence. The correlation between the mutational site map and the positions of amino acid changes in the A proteins of the mutants examined, however, is consistent with this concept and establishes the fact that the genetic map does relate to the primary structure of the corresponding protein.

Summary.—Studies of the altered A proteins produced by several mutants of *E. coli* have indicated that at least in two groups of mutants there is a correlation between the region of the gene altered by mutation and the location of the amino acid change in the corresponding mutant protein. In addition, peptide pattern studies have supported genetic evidence indicating that several of the mutant strains in each group represent repeat identical mutations at the same site. The A protein of one mutant, A-23, has been shown to differ from the wild-type protein by the substitution of a glycine by an arginine. This substitution involves the same amino acid that is replaced in the protein of another A-protein mutant, A-46.[18] Recombinational analyses have shown that the mutationally altered site in the A gene of mutant A-46 is very near to but separable from the altered site in mutant A-23.

Note added in proof.—Preliminary electrophoretic studies on cellulose acetate by U. Henning indicate that the A proteins of mutants A3, A11, and A33 are less negatively charged (apparently to the same degree) than the wild type A protein. These findings are consistent with a change of the same amino acid in the A proteins of these mutant strains.

The authors are indebted to Virginia Horn and Janet Lind for their excellent technical assistance.

* This investigation was supported by grants from the National Science Foundation and the U.S. Public Health Service.

† Postdoctoral Fellow of the U.S. Public Health Service. Some of these studies have been reported in a preliminary communication: Helinski, D., and C. Yanofsky, *Federation Proc.*, **20**, 255 (1961).

[1] Ingram, V. M., *Biochim. et Biophys. Acta*, **28**, 539 (1958).

[2] Ingram, V. M., *Hemoglobin and Its Abnormalities* (Springfield, Illinois: Charles C Thomas Co., 1961).

[3] Tsugita, A., and H. Fraenkel-Conrat, *J. Molecular Biology* (in press).

[4] Benzer, S., in *The Chemical Basis of Heredity*, ed. W. D. McElroy and B. Glass (Baltimore: The Johns Hopkins Press, 1957), p. 70.

[5] Rothman, F., see discussion following the paper cited in reference 7.

[6] Streisinger, G., F. Mukai, W. J. Dreyer, B. Miller, and S. Horiuchi, in Cellular Regulatory

Mechanisms, Cold Spring Harbor Symposia on Quantitative Biology, vol. 26 (1961, in press).

[7] Yanofsky, C., D. R. Helinski, and B. Maling, in *Cellular Regulatory Mechanisms*, Cold Spring Harbor Symposia on Quantitative Biology, vol. 26 (1961, in press).

[8] Henning, U., D. R. Helinski, F. C. Chao, and C. Yanofsky, *J. Biol. Chem.* (in press).

[9] Helinski, D. R., and C. Yanofsky, manuscript in preparation.

[10] Yanofsky, C., and I. P. Crawford, these PROCEEDINGS, **45**, 1016 (1959).

[11] Maling, B. D., and C. Yanofsky, these PROCEEDINGS, **47**, 551 (1961).

[12] Allen, M., and C. Yanofsky, unpublished experiments.

[13] Initial studies involving the A-11 protein were performed on the protein isolated from the mutant A-11; however, since the A protein of the double mutant A-11-46 is more stable than the A-11 protein, the double mutant protein was used in later studies on the A-11 amino acid change. Similarly, in studies on the A-23 protein, the initial studies were performed with the protein from the mutant A-23. Later studies were carried out on the more stable A protein obtained from the double mutant A-3-23. In each case, the peptide pattern difference characteristic of the mutant in question was found to be present in the corresponding double mutant peptide pattern.[7]

[14] Helinski, D. R., and C. Yanofsky, unpublished experiments.

[15] Hirs, C. H. W., S. Moore, and W. H. Stein, *J. Biol. Chem.*, **235**, 633 (1960).

[16] Carlton, B. C., and C. Yanofsky, unpublished experiments.

[17] Acher, R., and C. Crocker, *Biochim. et Biophys. Acta*, **9**, 704 (1952).

[18] Henning, U., and C. Yanofsky, these PROCEEDINGS, **48**, 183 (1962).

[19] Neurath, H., and G. W. Schwert, *Chem. Revs.*, **46**, 69 (1950).

[20] Yanofsky, C., and E. S. Lennox, *Virology*, **8**, 425 (1959).

---

*AN ALTERATION IN THE PRIMARY STRUCTURE OF A PROTEIN PREDICTED ON THE BASIS OF GENETIC RECOMBINATION DATA**

BY ULF HENNING† AND CHARLES YANOFSKY

DEPARTMENT OF BIOLOGICAL SCIENCES, STANFORD UNIVERSITY

Communicated by Victor C. Twitty, December 22, 1961

Studies on the fine structure relationships between a gene and the corresponding protein are being carried out with the A protein of the tryptophan synthetase of *Escherichia coli*. The preceding paper[1] describes an investigation with altered A proteins from several mutants and demonstrates that mutational changes at or near the same site in the A gene lead to alterations in the same region of the A protein.

Studies[2, 3] with a large number of A-protein mutants have shown that there are two closely linked sites at one end[4] of the A gene that have mutated frequently. All of the strains with mutational changes at one of these sites form a heat-labile A protein[2] that is distinguishable from the wild-type A protein in peptide pattern studies.[1] Strains with mutational changes at the second site form an altered A protein that is somewhat more heat-resistant than the wild-type A protein[2] but cannot be distinguished from the normal protein in peptide pattern studies.[1] The distance between the two mutational sites is approximately $^1/_{625}$–$^1/_{2500}$ of the total length of the map of the A gene.

Since the A protein appears to be a single polypeptide chain[5] containing approximately 280 amino acids,[6] one would predict that these two mutant types should have amino acid substitutions at or near the same position in the A protein if the number of nucleotide pairs in the A gene is a small multiple of the number of amino