Formaldehyde-induced Cross-Linkages in the $\alpha$ Subunit of the Escherichia coli Tryptophan Synthetase

(Received for publication, October 26, 1970)

JOYCE S. MYERS$ and JOHN K. HARDMAN

From the Department of Biology and the McCollum-Pratt Institute, The Johns Hopkins University, Baltimore, Maryland 21218

SUMMARY

Two intramolecular cross-linkages (methylene bridges) have been introduced by formaldehyde in the $\alpha$ subunit of the Escherichia coli tryptophan synthetase without any detectable loss in enzyme activity. These have been assigned to side chains in the sequences between residues 152 to 158, 212 to 218, and 227 to 233. Based on the conditions of reaction, the stability of the methylene bridges, and the availability of reactive residues in these sequences, the specific amino acid pairs linked together have been tentatively designated as asparagine-156-serine-214 and glutamine-218-serine-232.

Structure-function relationships utilizing both enzyme components (the $\alpha$ and $\beta_2$ subunits) of the Escherichia coli tryptophan synthetase have been the subject of investigation in several laboratories (1-6). Attention here has been focused on the $\alpha$ subunit of this enzyme complex. Chemical modification studies, primarily with sulfhydryl group reagents (1-3), and some limited observations of mutational altered proteins (4) have suggested discrete conformations in certain regions of the active enzyme. More recently, a further examination of the three-dimensional structure of this enzyme has been initiated with bifunctional reagents, the results obtained with bis(maleimido)methyl ether (7) and 1,5-difluoro-2,4-dinitrobenzene, reagents which react with the critical cysteinyl residues, have confirmed several suggestions which arose from the initial chemical studies.

In an effort to determine any functionally significant conformation in areas other than the cysteinyl region, formaldehyde, a bifunctional reagent which is less specific for these residues in these sequences, the specific amino acid pairs linked together have been tentatively designated as asparagine-156-serine-214 and glutamine-218-serine-232.

* This investigation was supported by United States Public Health Service Research Grant AI 06766 from the National Institutes of Health. Contribution 627 from the McCollum-Pratt Institute.

† Part of this work is taken from a Ph.D. thesis to be presented to the Department of Biology, The Johns Hopkins University (by J. S. M.) and was supported in part by National Institutes of Health Training Grant GM 57.

1 D. F. Hardman and J. K. Hardman, manuscript in preparation.

Elsall in 1945 (8) and subsequent studies by Fraenkel-Conrat et al. (9-13) have extended the investigation of formaldehyde as a cross-linking reagent both in model systems and with proteins. The primary reaction with many proteins appears to be the formation of an aminomethylol derivative with the $\epsilon$-amino group of lysyl residues. The cross-linking reaction involves the condensation of this group with an active hydrogen on primary amines, guanylyl, phenolic, and imidoxyl groups to form a methylene bridge. In addition, these and other side chain groups (e.g. aliphatic alcohols and thioalcohols) are potentially capable of initiating the cross-linking reaction (8). The conditions for methylene bridge formation vary for the different side chains, although most can react at neutral to alkaline pH and at moderate temperatures. Thus, formaldehyde should be readily applicable for enzyme modification studies, and this report presents the results obtained from an examination of the reaction of this reagent with the $\alpha$ subunit.

MATERIALS AND METHODS

Substrates and Reagents—The preparation of indigolglycerol phosphate and $^{14}$C-indigolglycerol phosphate was described previously (1). $^{14}$C-Formaldehyde and uniformly labeled $^{14}$C-L-lysine were obtained from New England Nuclear. Trypsin-1,1-tosylamido-2-phenylethyl chloromethyl ketone, chymotrypsin, and pepsin were purchased from Worthington.

Enzyme Preparations—The $\alpha$ subunit was purified according to the method of Malkinson and Hardman (3). A crude extract of E. coli strain A2, which contains a defective $\alpha$ subunit, was assayed by the method of Malkinson and Hardman (3). Protein was assayed by the method of Lowry et al. (14). Molecular weight estimations were based on the sodium dodecyl sulfate-polyacrylamide disc electrophoresis method of Weber and Osborn (15). Polyacrylamide disc electrophoresis methods have been described before (3). Gel tracings were made with the linear transport attachment to the Gilford recording spectrophotometer. Radioactivity measurements of aqueous samples were made in a solution of toluene-ethanol-spectrafluor (American-Searle Corporation, Arlington Heights, Illinois; a concentrated liquid scintillator containing 100 g of 2,5-diphenyloxazole and 1.25 g of 1,4-bis[2-(5-phenyloxazolyl)]benzene per liter of toluene) in a volume ratio, 69:29:4.2. A Nuclear-Chicago
model 3000 ambient temperature liquid scintillation counter was used.

**Treatment with Formaldehyde**—The reaction mixture consisted of 1% sodium bicarbonate, enzyme (0.5 mg per ml), and $^{14}$C-formaldehyde ($6 \times 10^6$ cpm per pmole) in a 100:1 molar ratio of reagent to protein. The final pH was 8.2. After incubation at 25°, the reaction mixture was dialyzed for 17 hours at 4° against two 400-volume batches of 0.05 M potassium phosphate buffer (pH 7) containing $10^{-4}$ M dithiothreitol and 0.2% 5,5'-dimethyl-1,3-cyclohexanedione (dimedon). Dimedon removes bound but unreacted formaldehyde. No additional radioactivity could be removed after this dialysis. Subsequently, a similar dialysis was performed in the absence of dimedon to remove this reagent. The extent of reaction for each sample was calculated from radioactivity and protein measurements.

These were the conditions used in all of the experiments described. The only variations were in the volume of the reaction mixture, the length of incubation, and the omission of dithiothreitol from the dialysis buffer when sulfhydryl group content was determined. The volume of the reaction mixture varied, depending on the amount of sample needed for analysis. For the experiments described in Figs. 1 to 4, 1.0-ml volumes were used; for peptide analyses, 30- to 60-ml (15 to 30 mg of protein) volumes were required.

**Peptide Analysis of Formaldehyde-treated Protein**—Prior to digestion for peptides, the protein was lyophilized and the cysteinyl residues were carboxymethylated with iodoacetate according to the procedure of Hardman and Yanofsky (1). This procedure alkylates only the cysteinyl residues. Tryptic digestion was performed as described by Horecker (19). After removal of urea and salts from the digest by filtration through a Sephadex LH-20 column at 4°, the radioactive peptides were isolated by preparative paper electrophoresis and chromatography (16). The radioactive bands were detected with a Packard radiochromatogram scanner and eluted with water. All of the peptides which were analyzed were isolated and determined to be pure by these techniques. Amino acid analyses were performed with the Beckman model 120B amino acid analyzers of the Department of Biology service facility. Chymotryptic and peptic digestion of the isolated tryptic peptide(s) were carried out by the methods of Helinski and Yanofsky (16) and Guest and Yanofsky (17), respectively, and the resulting radioactive peptides isolated as above. The dimethylaminonaphthalene sulfonyl chloride end group procedure was performed as described by Gray (18) and the dimethylaminonaphthalene sulfonyl-amino acids were identified according to the procedure of Morse and Horecker (19).

**RESULTS**

**Extent and Nature of Reaction of α Subunit with Formaldehyde**—Although no extensive study was made regarding the effect of pH on the reaction, preliminary experiments indicated very little reaction (through 48 hours of incubation) at pH 7 in phosphate buffer. Conditions of lower pH were avoided since the enzyme becomes progressively inactive below pH 6. At pH 8.2 (1% sodium bicarbonate), however, substantial reaction occurs and continues linearly from 4 through 24 hours (Fig. 1). During this time, approximately 4 moles of formaldehyde are bound and the pattern of labeling appears to be very reproducible. Longer incubations, up to 40 hours, indicate that 15 to 20 moles of formaldehyde could be bound. In view of the difficulties that were anticipated in determining the sites of reaction at this level of labeling, efforts were directed toward the initial labeling process (4- to 7-hour period), represented by the lag period in Fig. 1. It is seen in Figs. 1 and 2 (upper curves) that even at very short incubation times there was substantial binding. The value of 0.6 to 0.8 mole of formaldehyde bound at “0 hours” incubation represents that reaction which occurs during the time involved in setting up the dialysis and the initial period of dialysis (see “Materials and Methods”). An increase in net binding to about 1.2 to 1.3 moles occurs during the first 4 to 7 hours. In all of the subsequent experiments, the reaction was terminated after 48 hours. In larger scale reaction mixtures, the enzyme preparations routinely contained 1.2 to 1.5 moles of formaldehyde per mole of enzyme.

The finding that the level of bound formaldehyde in all preparations was always larger than unity indicates that some fraction of the protein molecules had more than 1 molecule of formal-
dehydrate. The properties of formaldehyde-treated protein on analytical polyacrylamide gel electrophoresis further indicate a heterogeneous mixture of protein molecules. The electrophoretic pattern shows that about 80 to 85% of the protein migrates substantially faster than untreated, control enzyme preparations, suggesting that on this basis, at least, nearly all of the protein has reacted with formaldehyde. Unfortunately, attempts to resolve the different fractions by preparative polyacrylamide electrophoresis were unsuccessful, all of the protein eluting as a single, large diffuse band. For this reason, the characterization (molecular weight and enzymatic and peptide analyses) of the treated protein was performed on total unfractionated preparations. Despite this drawback, however, these analyses provided relatively clear-cut answers. The peptide analysis, in particular, provides supporting evidence that a major portion (>50%) of the reacted protein molecules contain 2 molecules of formaldehyde.

**Molecular Weight Estimations**—Although the protein concentration was kept relatively low in an attempt to minimize any intermolecular cross-linking, it was important to establish this point. Molecular weight estimation by the sodium dodecyl sulfate-polyacrylamide method (15) was used. Fig. 3 presents the tracing of stained gels containing treated α subunit (solid line) and a mixture of proteins (dashed line) including the untreated enzyme. The position of each protein in the mixture was established in separate gels. There appears to be essentially no difference in mobility between the treated and untreated protein. Molecular weight estimations of the formaldehyde-protein based on the mobilities of the control proteins indicate a molecular weight identical with the normal α subunit.

**Activity Measurements**—Fig. 4 shows the results of activity measurements on the enzyme which was treated with formaldehyde for periods of time up to 4½ hours. Fig. 4A shows the activity of α subunit alone in the reaction, indoleglycerol phosphate to indole; Fig. 4B shows the activity of the α subunit in the presence of the β subunit (the indole plus serine to tryptophan reaction). In neither case is there substantial activity loss. Thus, little if any of the formaldehyde, which reacts either monofunctionally or bifunctionally, appears to alter any side chain or conformation which is essential for normal functioning of the enzyme.

**Bifunctionality of Reaction**—In order to determine whether formaldehyde had formed a cross-linkage or was present as the monomethylol derivative, several techniques were used. The monomethylol derivatives of many amino acids are labile to conditions of acid hydrolysis; the bound formaldehyde is liberated by distillation from acid (8, 9). Accordingly, aliquots of the reaction mixture were removed, dialyzed as described, and hydrolyzed for 24 hours in 6 N HCl at 110° in a vacuum. Subsequent distillation of these samples into dimedon indicated that only about 10 to 20% of the radioactivity was lost from the hydrolysate. These data are shown in the lower curve (●) in Fig. 2. Control experiments in which 14C-formaldehyde was added immediately prior to the hydrolysis or distillation steps indicated that essentially all (92 to 97%) of the radioactivity was recovered in the distillate. Thus, any free formaldehyde liberated under these conditions would have been detected. It would appear from these experiments that little of the bound formaldehyde is in the monomethylol form.

Another test for the presence of monofunctional derivatives was to examine the amount of free lysine which could be bound to the formaldehyde-treated enzyme. After the initial treatment with unlabeled formaldehyde, a second 4½-hour incubation, under similar conditions, was performed in the presence of 14C-lysine. Each sample was then dialyzed as before and analyzed for bound lysine. The concentration of lysine represented a 1000-fold excess over the amount of bound formaldehyde in these samples. As can be seen in Fig. 2 (lower curve, X—X), the
amount of bound lysine consists of about 10 to 20% of the total formaldehyde on the protein. An obvious drawback to this type of experiment, of course, is that conditions of concentration, pH, time, or temperature may not have been optimum for lysine to react. However, these data do agree well with the other non-functional assay, and, furthermore, the peptide analyses presented below strongly suggest that at least 65 to 70% of the formaldehyde is bound as a methylene bridge structure.

Localization of Sites of Formaldehyde Cross-Linkages—The acid stability of the bound formaldehyde suggested cross-linkages other than those involving lysine-arginine or lysine-amide side chains (9, 10). In addition, sulphydryl group titrations (Fig. 4C) indicated that none of the cysteine residues had reacted. Since tryptic peptides are readily identifiable and since limited digestion by trypsin due to blocked lysine or arginine residues was not expected, tryptic digestion was attempted initially to localize the cross-linkages. After tryptic digestion, the urea and digestion buffer were removed by Sephadex LH-20 filtration. Four peaks (A, B, C, and D) of radioactivity were found (Fig. 5). Peaks A and B represent 65 to 70% and 20%, respectively, of the total radioactivity. Peak A yielded seven to eight ninhydrin-staining bands on preparative paper electrophoresis. Of these, only one was radioactive (mobility, 0.28 to 0.32, relative to the quinine sulfate marker (16); this band contained approximately 90% of the total radioactivity). This material was eluted from the paper and tested for homogeneity by re-electrophoresis and paper chromatography (16). A single ninhydrin-staining component was found with both techniques (mobility, 0.30 on electrophoresis; $R_f$, 0.15 in the descending chromatography system (16)). Peak B also contained seven to eight ninhydrin-staining bands on preparative paper electrophoresis. Again, only one major radioactive component was found, containing \( \geq 90\% \) of the radioactivity. The amino acid composition of this band readily established its purity. The remaining peaks (C and D) contained a total of about six to seven ninhydrin-staining components. In none of these, however, was the level of radioactivity high enough to be clearly distinguished from background. Because of this and the fact that 85 to 90% of the labeled peptides had been accounted for in Peaks A and B, the latter two fractions (C and D) were not studied further.

The amino acid analyses of the radioactive peptides from Peaks A and B are shown in Table I. Tryptophan A (from Sephadex Peak A) contained both lysine and arginine in a 2:1 ratio, suggesting that it consisted of three normal tryptic peptides. Calculations were based on these molar ratios. A rigorous comparison with all of the known tryptic peptides in various combinations indicated that it contained TP-29 (residues 145 to 163), TP-13 (residues 201 to 220), and TP-6 (residues 221 to 238). There are the only possible tryptic peptides whose total composition is compatible with that found for this labeled peptide. The compositions of TP-29, TP-13, and TP-6 are shown for comparison in Table I and the sequences of these peptides are given in the legend to the table. Tryptophan B (the radioactive band from Sephadex Peak B) was easily recognized as being identical with TP-7 (residues 256 to 262). No further work was done.

| Amino acid     | Peptide A | Peptide B |
|----------------|-----------|-----------|
|                | Found     | Theory†   | Found   | Theory‡   |
| Lysine            | 2.35      | 2         | 1.00    | 1        |
| Histidine         | 0.86      | 1         | 0       | 0        |
| Arginine          | 1.13      | 1         | 0       | 0        |
| Carboxymethyleysteine | Trace     | 1        | 0       | 0        |
| Aspartic acid     | 7.10      | 8         | 0.18    | 0        |
| Threonine         | 0         | 0         | 0       | 0        |
| Serine            | 3.18      | 3         | 0.19    | 0        |
| Glutamic acid     | 2.62      | 3         | 1.04    | 1        |
| Proline           | 4.48      | 6         | 1.15    | 1        |
| Glycine           | 5.75      | 5         | 0.21    | 0        |
| Alanine           | 13.60     | 12        | 0.31    | 0        |
| Valine            | 2.75      | 3         | 1.79    | 2        |
| Methionine        | 0         | 0         | 0.95    | 1        |
| Isoleucine        | 6.50      | 6         | 0       | 0        |
| Leucine           | 3.20      | 3         | 0.32    | 0        |
| Tyrosine          | 0.55      | 1         | 0       | 0        |
| Phenylalanine     | 2.00      | 2         | 0.96    | 1        |

* Total composition of TP-29, TP-13, and TP-6.
† Composition of TP-7.
‡ Less than 0.3 residue per mole of peptide.

Fig. 5. Elution profile from Sephadex LH-20 of tryptic peptides obtained from \(^{14}C\)-formaldehyde-labeled enzyme (1.3 moles of formaldehyde per mole of enzyme). A tryptic digest (16) of a 30-mg batch of carboxymethylated, formaldehyde treated enzyme preparation was applied to a column (1.5 x 70 cm). The peptides were eluted with water at a flow rate of 10 ml per hour, fractions were collected at 20-min intervals and the absorbance at 232 m\(\mu\) and radioactivity monitored in each fraction.
with this peptide since any cross-linkage which may be present
within this short segment would yield limited information about
the conformation of the enzyme.

The fact that the composition of tryptic peptide A included
three normal peptides suggests several possibilities for its struc-
ture. It could represent, in addition to a single labeled peptide,
the other peptides as contaminants. This possibility seems un-
likely since the constituent peptides (TP-29, TP-13, and TP-6)
are readily separable by either the electrophoresis or chromatog-
raphy systems. For example, the mobilities of TP-29, TP-13,
and TP-6 are 0.36, 0.27, and 0.36, respectively. The Rp values
for TP-29, TP-13, and TP-6 are 0.27, 0.10, and 0.00, respectively.
These values were obtained in separate experiments and agree
well with those obtained from published peptide maps (20).
Furthermore, if any of these peptides were not covalently linked
together, the molar ratios of the residues found in tryptic peptide
A (and in all of the chymotryptic and tryptic peptides derived
from it) would require that the nonlinked peptide(s) had eluted
in a molar ratio which is nearly identical with that of the linked
peptide. From these considerations, it would appear that all
three peptides are linked together. This possibility raises the
question of whether one or more cross-linkages are present in this
peptide. Peptides TP-13 and TP-6 are contiguous in the amino
acid sequence (21) and their presence together may simply rep-
resent the fact that the lysine-alanine bond (residues 220 and
221) between these two peptides was poorly cleaved by trypsin.
This would mean that either TP-13 or TP-6 would be cross-linked
to TP-29 by a single methylene bridge. End group analyses by
the dimethylaminonaphthalene sulfonyl chloride technique (18)
indicated the presence of three amino-terminal residues, histidine,
glutamic acid, and alanine. This finding helps substantiate the
identification of the peptides and further indicates that the lysine-
alanine bond between TP-13 and TP-6 was cleaved by trypsin.
In addition, the specific radioactivity was determined for the
preparation of tryptic peptide A whose analysis is presented in
Table I. The moles of 14C-formaldehyde per mole of lysine,
arginine, and histidine were 1.06, 2.30, and 1.90, respectively.
This gives an average value of 2.14 moles of 14C-formaldehyde per
mole of tryptic peptide A. Several other preparations of this
peptide, in which the specific radioactivity was measured, con-
tained 1.76 to 1.94 moles of 14C-formaldehyde per mole of peptide.
These observations strongly suggest that tryptic peptide A con-
sists of three distinct and separate peptides, which are linked
through two methylene bridges. Subsequent digestion of tryptic
peptide A with chymotrypsin and pepsin was performed in order
to verify this point and to narrow further the areas involved in
the cross-linkages.

Chymotryptic digestion produces four to five peptides, only
one of which was radioactive; pepsin produces six to seven pep-
tides, of which three were radioactive—two were in equivalent
although poor yield, and the third in about one-half the yield of
the other two. The purity of these peptides was determined as
described above for tryptic peptide A. The amino acid analyses
of the chymotryptic fragment obtained from two preparations of
tryptic peptide A are given in Table II and are consistent with
the composition presented for those areas of tryptic peptide A
indicated in Footnote a of Table II. Residues 152 to 162 and
212 to 238 represent, respectively, chymotryptic peptides CP-49
(22) and A 169 CP-M (minus that portion which had been cleaved
by trypsin at lysyl residue 238 (21). Specific radioactivity
measurements on one of these chymotryptic preparations showed

| Amino acid composition of chymotryptic peptide obtained from tryptic peptide A |
|---------------------------------|--------|--------|
| Amino acid                  | Found   | Theory* |
| Lysine                       | -; 1.72 | 2      |
| Histidine                    | -; 0    | 0      |
| Arginine                     | -; 0    | 0      |
| Carboxymethylcysteine        | Trace; trace | 1 |
| Aspartic acid                | 5.04; 5.88 | 6 |
| Threonine                    | 0; 0    | 0      |
| Serine                       | 2.69; 2.35 | 3 |
| Glutamic acid                | 2.16; 2.70 | 1 |
| Proline                      | 3.54; 2.51 | 3 |
| Glycine                      | 5.47; 5.40 | 4 |
| Alanine                      | 8.87; 7.50 | 9 |
| Valine                       | 1.75; 1.62 | 2 |
| Methionine                   | 0; 0    | 0      |
| Isotoucine                   | 4.67; 4.02 | 5 |
| Leucine                      | 2.32; 2.24 | 2 |
| Tyrosine                     | 0; 0    | 0      |
| Phenylalanine                | 0; 0    | 0      |

* Composition of the fragments shown in tritiated tryptic peptide A: TP-29, His-Asn-Val-Ala-Pro-Ile-Phe-Ile-Cys-Pro-Pro-Pro-
Asn-Ala-Asp-Asp-Asp-Leu-Leu-Arg; TP-13, Glu-Tyr-Asn-Ala-Asp-Pro-Leu-Glu-Gly-Phe-Gly-Ile-Ser-Ala-Asp-Gly-
Val-Lys; TP-6, Ala-Ile-Asp-Ala-Oly-Ala-Oly-Ala-Ile-
Seu-Gly-Ser-Ala-2e-Val-Lys.

* Less than 0.2 residue per mole of peptide.
Amino acid composition of peptic peptides I, II, and III obtained from tryptic peptide A

| Amino acid | Peptide I |         | Peptide II |         | Peptide III |         |
|------------|-----------|---------|------------|---------|-------------|---------|
| Lysine     | 0.14      | 0       | 0          | 0       | 1.05        | 1       |
| Histidine  | 0.81      | 1       | 0          | 0       | 0           | 0       |
| Arginine   | 1.16      | 1       | 0          | 0       | 0.93        | 1       |
| Carboxymethyl-cysteine | 0.62 | Trace | 1 | 0 | 0 | 0 |
| Aspartic acid | 7.29 | 8 | 3.97 | 4 | 5.22 | 5 |
| Threonine  | 0.37      | 0       | 0          | 0       | 0           | 0       |
| Serine     | 1.55      | 2       | 1.97       | 2       | 1.84        | 2       |
| Glutamic acid | 3.98 | 3 | 2.07 | 2 | 1.42 | 1 |
| Proline    | 3.50      | 6       | 2.23       | 3       | 1.32        | 1       |
| Glycine    | 4.89      | 5       | 4.97       | 5       | 2.90        | 3       |
| Alanine    | 11.65     | 11      | 6.68       | 6       | 4.89        | 5       |
| Valine     | 1.16      | 1       | 0          | 0       | 2.01        | 1       |
| Methionine | 0         | 0       | 0          | 0       | 0           | 0       |
| Isoleucine | 4.00      | 5       | 2.94       | 3       | 2.11        | 2       |
| Leucine    | 3.24      | 3       | 0.98       | 1       | 2.12        | 2       |
| Tyrosine   | 0.63      | 1       | 0          | 0       | 0           | 0       |
| Phenylalanine | 1.30 | 2 | 1.74 | 2 | 0 | 0 |

^a Composition of the fragments shown in italics in tryptic peptide A: TP-29, His-Asn-Val-Ala-Pro-Ile-Phe-Ile-Cys-Pro-Pro-Asn-Ala-Asp-Asp-Asp-Leu-Leu-Arg; TP-13, Glu-Tyr-Asn-Ala-Ala-Pro-Asp-Leu-Gln-Gly-Phe-Glu-Ile-Ser-Ala-Pro-Asp-Gln-Val-Lys; TP-6, Ala-Ala-Ile-Asp-Ala-Gly-Ala-Gly-Ala-Ile-Ser-Gly-Ser-Ala-Val-Lys.

^b Composition of the fragments shown in italics in tryptic peptide A: TP-29, His-Asn-Val-Ala-Pro-Ile-Phe-Ile-Cys-Pro-Pro-Asn-Ala-Asp-Asp-Asp-Leu-Leu-Arg; TP-13, Glu-Tyr-Asn-Ala-Ala-Pro-Asp-Leu-Gln-Gly-Phe-Glu-Ile-Ser-Ala-Pro-Asp-Gln-Val-Lys; TP-6, Ala-Ala-Ile-Asp-Ala-Gly-Ala-Gly-Ala-Ile-Ser-Gly-Ser-Ala-Ile-Val-Lys.

^c Composition of the fragments shown in italics in tryptic peptide A: TP-29, His-Asn-Val-Ala-Pro-Ile-Phe-Ile-Cys-Pro-Pro-Asn-Ala-Asp-Asp-Asp-Leu-Leu-Arg; TP-13, Glu-Tyr-Asn-Ala-Ala-Pro-Asp-Leu-Gln-Gly-Phe-Glu-Ile-Ser-Ala-Pro-Asp-Gln-Val-Lys; TP-6, Ala-Ala-Ile-Asp-Ala-Gly-Ala-Gly-Ala-Ile-Ser-Gly-Ser-Ala-Val-Lys.

^d Less than 0.15 residue per mole of peptide.

**DISCUSSION**

The results presented here indicate that formaldehyde has introduced two major intramolecular cross-linkages in the α subunit. These cross-linkages can be assigned to residues linking tryptic peptides TP-29, TP-13, and TP-6. The composition of the chymotryptic and peptic peptides are consistent with this conclusion; all can be considered as a family of peptides obtained from these three tryptic peptides.

On the basis of the composition of all of these peptides, the sequences involved in the cross-linkages can be narrowed further. A tentative assignment of these regions is to the sequences between residues 152 to 158, 212 to 218, and 227 to 233. The assignment of the amino-terminal residues for these sequences is based on the composition of the chymotryptic fragment for residue 152, the tryptic peptide and peptide III for residue 212, and peptide peptide III for residue 227. The assignment of the carboxy-terminal residues is based on the composition of peptide II for residue 158, peptides I and II for residue 218, and all three peptic peptides for residue 233. The areas common to the sequences of all of the peptides are shown in Fig. 6.

It is possible that the sequence of residues 152 to 156 can also be eliminated because of the composition of peptic peptide III. However, this assumes peptic cleavage on the carboxy side of a prolyl residue. This is unusual, although the same assumption can be made for peptic peptide II; both require a splitting after a prolyl-proline sequence. For this reason, these amino terminal portions of these peptic peptides cannot be considered as conclusive. The absolute yields of all of the peptic peptides were very low and precluded any reliable amino or carboxy-terminal analyses. Of some importance, also, with regard to these assignments is the designation of residue 233 as a carboxy-terminal residue. The inclusion of an additional seryl residue (residue 234) in these sequences would compromise the conclusions discussed later. This possibility does not seem too likely in view of the fact that serine, which is reportedly susceptible to some loss during acid hydrolysis, has always been recovered in 85 to 90% yield here. Furthermore, the analyses of all three peptic peptides is consistent with the exclusion of an additional seryl residue.

The designation of the specific residues linked by the methylene bridges cannot be made unequivocally at the present time since the isolation of only those residues involved was not possible. On the basis of a number of observations, however, nearly all of the possible cross-linkages can be eliminated. The exclusion of lysyl, arginyl, tyrosyl, histidyl, and cysteinyl side chains is certain, primarily on the basis of the peptide analyses and the sulfhydril group titrations. Indolyl groups need not be considered since the α subunit contains no tryptophan. Of the residues contained within the sequences mentioned above, serine, asparagic acid, glutamine, and asparagine are possibilities at least (8). Several of these would appear unlikely, however, when considering (a) the reaction conditions used (pH 8.3, aqueous medium), (b) the stability of the bound formaldehyde to conditions of acid hydrolysis and distillation, and (c) the fact that two cross-linkages must be present in the available sequences. Thus, methylcarboxy ethers, methylene diamides, and acetals (i.e. cross-linkages involving asparagic acid, amide-amide, or serine-serine residues, respectively) are improbable (8, 9). These considerations vir-
tually eliminate all but seryl-amide cross-linkages in these se-
quences. The stability of formaldehyde bound in this type of
methylene bridge is not clearly defined, although both threonine
and serine have been shown to combine formaldehyde with
phenolic and imidazole compounds in an acid-stable manner (13).
If such cross-linkages are present in the peptides isolated, they
would probably be between asparagine-156 and serine-214 and
between glutamine- and serine-232. These assignments are,
of course, based on the exclusion of serine-234. If this residue
were present, a number of other amide-seryl cross-linkages would
be possible. Regardless of the absolute accuracy of the assign-
ment of the specific cross-linkages, however, it is clear that these
regions of the protein are bound together by formaldehyde and
are probably closely aligned in the three-dimensional structure of
the active enzyme.

In addition, activity measurements suggest that these cross-
linkages do not distort the conformation or hinder access of the
substrate, indoleglycerol phosphate, to the α subunit when it is
acting alone. Neither is there any indication of an impairment
in the ability of the treated α subunit to combine functionally
with the β subunit. Both findings (i.e. the sites of the cross-
linkages and the recovery of activity) are consistent with struc-
ture-function studies using other bifunctional reagents (7) which
indicate that this area of the protein is substantially removed
from the site of substrate interaction.

Acknowledgments—The authors would like to thank Ulysses
Smalls and Joanne Chester for the amino acid analyses and
Deanna Hardman for excellent technical assistance.

REFERENCES
1. HARDMAN, J. K., AND YANOFSKY, C., J. Biol. Chem., 240, 725
   (1965).
2. HARDMAN, J. K., AND YANOFSKY, C., Science, 156, 1369 (1967).
3. Malkinson, A., M., AND HARDMAN, J. K., Biochemistry, 8, 2767, 2777
   (1969).
4. YANOFSKY, C., HORN, V., AND THORPE, D., Science, 146, 1593
   (1964).
5. MILES, E. W., HAYAKAWA, M., AND CRAWFORD, F. P., Biochem-
   istry, 7, 2742 (1968).
6. MILES, E. W., Fed. Proc., 29, 665 (1970).
7. FREEDBERG, W. B., AND HARDMAN, J. K., J. Biol. Chem., 246,
   1439 (1971).
8. FRENCH, D., AND EDSEL, J. T., Advan. Protein Chem., 2, 277
   (1945).
9. FRAENKEL-CORAT, H., COOPER, M., AND OLCOTT, H. S., J.
   Amer. Chem. Soc., 87, 390 (1965).
10. FRAENKEL-CORAT, H., AND OLCOTT, H. S., J. Amer. Chem.
    Soc., 86, 34 (1964).
11. FRAENKEL-CORAT, H., BRANDON, R. A., AND OLCOTT, H. S.,
    J. Biol. Chem., 165, 90 (1947).
12. FRAENKEL-CORAT, H., AND OLCOTT, H. S., J. Amer. Chem.
    Soc., 70, 2073 (1948).
13. FRAENKEL-CORAT, H., AND OLCOTT, H. S., J. Biol. Chem.,
    174, 827 (1948).
14. LAWRY, O. H., ROEBROUGH, N. J., FARR, A. L., AND RANDALL,
    R. J., J. Biol. Chem., 193, 265 (1951).
15. WEBER, K., AND OSBORN, M., J. Biol. Chem., 244, 406 (1969).
16. HELLISKE, D. R., AND YANOFSKY, C., Biochim. Biophys. Acta,
    63, 10 (1962).
17. GUEST, J. R., AND YANOFSKY, C., J. Biol. Chem., 241, 1 (1966).
18. GORMAN, W. R., IN C. W. H. HIRS, S. P. COLOWICK, AND N. O.
    KAPLAN (Editors), Methods in enzymology, Vol. XI, Acad-
    emic Press, New York, 1967, p. 130.
19. MORGAN, D., AND HORECKER, B. L., Anal. Biochem., 14, 429
    (1966).
20. GORMAN, J. R., CARLTON, B. C., AND YANOFSKY, C., J. Biol.
    Chem., 242, 5397 (1967).
21. GORMAN, J. R., CARLTON, B. C., AND YANOFSKY, C., J. Biol.
    Chem., 242, 5402 (1967).
22. CARLTON, B. C., GORMAN, J. R., AND YANOFSKY, C., J. Biol.
    Chem., 242, 5422 (1967).
Formaldehyde-induced Cross-Linkages in the α Subunit of the Escherichia coli Tryptophan Synthetase
Joyce S. Myers and John K. Hardman

J. Biol. Chem. 1971, 246:3863-3869.

Access the most updated version of this article at http://www.jbc.org/content/246/12/3863

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/246/12/3863.full.html#ref-list-1