The Primary Structure of the \( \alpha \) Subunit of Protocatechuate 3,4-Dioxygenase

I. ISOLATION AND SEQUENCE OF THE TRYPIC PEPTIDES*

(Nancy A. Kohlmiller and James Bryant Howard)

From the Department of Biochemistry, Medical School, University of Minnesota, Minneapolis, Minnesota 55455

(Received for publication, February 5, 1979, and in revised form, April 13, 1979)

The carboxymethylated \( \alpha \) subunit of protocatechuate 3,4-dioxygenase was digested with trypsin. The 14 tryptic peptides were isolated by ion exchange chromatography on DEAE-Sephadex and by gel filtration chromatography. Automated Edman degradation and carboxypeptidase Y and B digestion were used to establish the sequence of these peptides. Further fragmentation of two tryptic peptides, T3 and T5, by Staphylococcus aureus protease and cyanogen bromide, respectively, was necessary to complete the sequences. The tryptic peptides accounted for a minimum of 199 residues out of a total of 202 residues predicted by amino acid analysis.

Protocatechuate 3,4-dioxygenase catalyzes the intradiol cleavage of protocatechuic acid by the insertion of 2 atoms of molecular oxygen to form \( \beta \)-carboxy-cis,cis-muconic acid (1, 2). Protocatechuate 3,4-dioxygenase is the second enzyme in the pathway leading from \( \beta \)-hydroxybenzoate to succinate and acetyl-CoA (3, 4). The utilization of this pathway involving aromatic oxidation allows aerobic bacteria to convert a stable aromatic compound into compounds of central metabolism. It is through this and similar pathways that aromatic biopolymers such as lignin re-enter the carbon cycle.

The native enzyme has a molecular weight of 700,000, contains 8 ferric iron atoms/molecule, and is composed of two types of subunits, \( \alpha \) and \( \beta \), with molecular weights of 22,500 and 26,800, respectively (5). Electron paramagnetic resonance and Mössbauer spectroscopy have shown that the environment of the ferric iron in the enzyme is of a previously uncharacterized type, namely, a polar environment with ligands of oxygen or nitrogen atoms, rather than cysteinyl ligands as in other nonheme iron proteins (6). Raman spectroscopy of protocatechuate 3,4-dioxygenase has recently demonstrated that at least one of the iron ligands is a tyrosyl residue (7, 8). Extensive spectroscopic studies of protocatechuate 3,4-dioxygenase have provided evidence for an ordered reaction mechanism where binding of the organic substrate is followed by oxygen binding to form a ternary complex. From substrate analogue inhibition data, Mössbauer spectroscopy, and electron paramagnetic spectroscopy, Que et al. (9) proposed a mechanism where the 4-hydroxyl group of the substrate is coordinated to the iron. Upon the iron activation of the substrate, the molecular oxygen attacks the aromatic ring to form a peroxo intermediate which rearranges with carbon-carbon bond scission (9). Preliminary results with the aldehyde analogue of the substrate indicate that the active site is in the \( \alpha \) subunit.1

In order to provide a firm basis for further studies on the mechanism of action of protocatechuate 3,4-dioxygenase, the determination of its complete primary structure was undertaken in our laboratory (10). We present in this and the accompanying paper the complete primary sequence of the \( \alpha \) subunit. This represents the first primary structure for any dioxygenase and should provide essential information for the eventual understanding of the nature of the interactions between subunits, the structure of the substrate and iron binding sites, and the reaction mechanism. The amino acid sequence is also a necessary prerequisite for the determination of the three-dimensional structure by x-ray crystallography.

MATERIALS AND METHODS AND RESULTS

All materials and procedures used, as well as the "Results," are described in the supplement to this paper.2 Fig. 3 and Table II are included in the body of the paper.

DISCUSSION

Fifteen peptides were isolated from the trypsic digest of the \( \alpha \) subunit of protocatechuate 3,4-dioxygenase. Their compositions are shown in Table II, and the sequences, which were determined by automated Edman degradation and carboxypeptidase Y digestions, are shown in Fig. 3. Two related peptides, T8A and T8B, were isolated in a ratio of 3:1. These peptides had identical sequences with the exception of a single COOH-terminal arginine in T8A and a diarginyl COOH-terminal sequence in T8B. It was not clear whether T10, free arginine, arose solely from the partial cleavage of this diarginyl bond or also from other regions of the protein. The composition of the tryptic peptides could account for a minimum of 199 residues out of a total of 202 predicted by amino acid analysis of the whole \( \alpha \) subunit.

The yields of the tryptic peptides generally were high and ranged from the quantitative recovery of T9 (99%) to a low of 24% for T6. Two peptides, T3 and T5, contained lysylprolyl

1 J. D. Lipscomb, private communication.
2 Portions of this paper (including "Materials and Methods," "Results," Figs. 1, 2, 4 to 11, and Tables I, III to XX) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full-size photocopies are available from the Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, Md. 20014. Request Document No. 79M-221, cite author(s), and include a check or money order for $5.85 per set of photocopies.
Tryptic Peptides from Protocatechuate 3,4-Dioxygenase

**Table II**

Amino acid composition of the tryptic peptides from the α subunit.

| Amino acid   | T1   | T2   | T3   | T4   | T5   | T6   | T7   | T8A  | T8B  | T9   | T10  | T11  | T12  | T13  | T14  | Total
|--------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| (CM)-Cysteine| 1.2  | 2.1  | 0.0  | 6.8  | 3.8  | 1.2  | 2.9  | 0.9  | 0.8  | 0.8  | 1.1  | 1.1  | 1.1  | 1.1  | 1.1  | 2
| Aspartic acid| 3.0  | 0.3  | 4.3  | 5.1  | 1.1  | 0.1  | 0.1  | 1.1  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 26
| Threonine    | 0.6  | 2.0  | 0.9  | 1.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 4
| Serine       | 3.8  | 2.2  | 5.4  | 2.8  | 1.0  | 1.2  | 2.9  | 2.9  | 2.9  | 2.9  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 24
| Glutamic acid| 5.1  | 2.3  | 2.6  | 2.0  | 1.8  | 1.0  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 0.1  | 12
| Proline      | 3.0  | 4.0  | 3.0  | 3.0  | 0.2  | 0.2  | 0.2  | 0.2  | 0.2  | 0.2  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 16
| Glycine      | 3.9  | 0.1  | 2.1  | 3.8  | 5.0  | 0.6  | 3.0  | 3.0  | 3.0  | 3.0  | 3.0  | 3.0  | 3.0  | 3.0  | 3.0  | 21
| Alanine      | 1.1  | 2.0  | 1.0  | 3.4  | 4.4  | 0.9  | 0.9  | 0.9  | 0.9  | 0.9  | 0.9  | 0.9  | 0.9  | 0.9  | 0.9  | 11
| Methionine   | 1    |      | 0.8  | 1    |      |      |      |      |      |      |      |      |      |      |      | 1
| Isoleucine   | 1.9  | 1.1  | 1.1  | 1.9  | 2.0  | 1.0  | 1.0  | 1.0  | 1.0  | 0.7  | 0.7  | 0.7  | 0.7  | 0.7  | 0.7  | 18
| Leucine      | 4.1  | 5.2  | 2.2  | 2.1  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 0.7  | 0.7  | 0.7  | 0.7  | 0.7  | 0.7  | 6
| Tyrosine     | 0.8  | 0.9  | 1.8  | 0.2  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 10
| Phenylalanine| 1.0  | 1.9  | 1.9  | 1.9  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 10
| Histidine    | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 7
| Arginine     | 0.9  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 1.0  | 5
| Tryptophan   | 7    | 9    | 12   | 21   | 12   | 21   | 12   | 21   | 12   | 21   | 12   | 21   | 12   | 21   | 12   | 3
| Total Yield  | 31   | 7    | 26   | 30   | 30   | 9    | 12   | 12   | 12   | 6    | 1    | 6    | 1    | 6    | 1    | 199

* The minimum number of residues determined by sequence analysis assuming that arginine (T12) was from T8 only. T8A not included in this tabulation.

* Tryptophan determined by Ehrlich reaction on paper.

FIG. 3. Amino acid sequences of the tryptic peptides from the α subunit of protocatechuate 3,4-dioxygenase. --, determined by Edman degradation; --, determined by carboxypeptidase Y digestion, except for T8, in which carboxypeptidase B was used. Sequencing of derivative peptides is shown below the parent peptides.
bonds which were not cleaved by trypsin. One chymotryptic peptide was isolated in low yield (10%) and corresponded to the NH₂-terminal 13 residues of T5 which resulted from cleavage at a histidyl residue. Most peptides could be entirely sequenced by automated Edman degradation. The use of Polybrene (17-19) in the spinning cup sequenator reduced the washout of small peptides such that, in most cases, the COOH-terminal residue could be identified either directly as the phenylthiohydantoin derivative or as the free amino acid after extraction of the cup. In the sequencing of T3 and T4, a large drop in absolute yield occurred at a specific aspartylglycyl bond (in T3, residues 21 to 22, and in T4, residues 12 to 13). Although the drop in yield in T3 necessitated further fragmentation of the peptide to complete the sequence, intact T4 was sequenced through the penultimate residue. To obtain the complete sequence of T5, the largest tryptic peptide isolated, cyanogen bromide cleavage at the single methionine residue was used.

Carboxypeptidase Y digestion provided valuable information in the sequencing of the tryptic peptides. Although carboxypeptidase Y has been reported to be somewhat slow in digesting at basic residues (21), rapid release of arginine and lysine was obtained for most of our tryptic peptides. However, carboxypeptidase Y released no arginine from the pair of basic residues in T8B having the sequence Pro-Gln-Arg-Arg and little arginine from the analogous peptide, T8A, having the sequence-Pro-Gln-Arg. In some cases when glycine was the penultimate amino acid, no release of amino acids occurred, e.g., no amino acids were released from the sequence -Asp-Gly-Lys in peptide T11. Likewise, internal glycine terminated digestion, namely in peptide T3 with the sequence -Gly-His-Leu-Val-Arg, leucine, valine, and arginine were released, but not histidine. One exception to the slow release of glycine was peptide T4 where glycine next to phenylalanine was completely released. Threonine also prevented or terminated digestion by carboxypeptidase Y as exemplified in peptides T1, T6, and T14. Our results would suggest that not only are the COOH-terminal residues important in the specificity of carboxypeptidase Y but also the properties of the penultimate residues.

The complete sequence and the ordering of the tryptic peptides by overlap sequences is given in the accompanying paper (23).

Acknowledgments—We are especially indebted to Dr. John Lipscomb for advice and help in the preparation of the protein and to Mr. Thomas Lorsbach for maintaining the sequencing equipment.

REFERENCES

1. Stanier, R. Y., and Ingraham, J. L. (1954) J. Biol. Chem. 210, 799-808
2. Fujisawa, H., and Hayaishi, O. (1968) J. Biol. Chem. 243, 2673-2681
3. Hosokawa, K., and Stanier, R. Y. (1966) J. Biol. Chem. 241, 2453-2460
4. Orrston, L. N., and Stanier, R. Y. (1966) J. Biol. Chem. 241, 3776-3786
5. Yoshida, R., Hori, K., Fujiiwara, M., Saecki, Y., Kogamiyama, H., and Nosaki, M. (1976) Biochemistry 15, 4948-4953
6. Que, L., Jr., Lipscomb, J. D., Zimmerman, R., Munck, E., Orme-Johnson, N. R., and Orme-Johnson, W. H. (1976) Biochim. Biophys. Acta 452, 320-334
7. Keyes, W. L., Loehr, T. M., and Taylor, M. L. (1978) Biochem. Biophys. Res. Commun. 83, 941-945
8. Takegami, Y., Saecki, Y., Iwaki, M., Yagi, T., Nozaki, M., Kitagawa, T., and Otsuka, S. (1978) J. Am. Chem. Soc. 100, 4614
9. Que, L., Jr., Lipscomb, J. D., Munck, E., and Wood, J. M. (1977) Biochim. Biophys. Acta 485, 60-74
10. Lipscomb, J. D., Howard, J. B., Lorsbach, T., and Wood, J. M. (1976) Fed. Proc. 35, 1536
11. Lidell, D. J., and Howard, J. B. (1978) J. Biol. Chem. 253, 3422-3426
12. Crestfield, A. M., Moore, S., and Stein, W. H. (1963) J. Biol. Chem. 238, 622-627
13. Simpson, R. J., Neuberger, M. R., and Liu, T.-Y. (1976) J. Biol. Chem. 251, 1396-1394
14. Gross, E. (1967) Methods Enzymol. 11, 238-255
15. Udenfried, S., Stein, S., Bohlen, P., Dairman, W., Leingruber, W., and Weigele, M. (1972) Science 178, 871
16. Edman, P., and Begg, G. (1967) Eur. J. Biochem. 1, 80-91
17. Tan, G. L., Beecher, J. F., Bell, M., and McKean, D. J. (1978) Anal. Biochem. 84, 622-627
18. Klapper, D. A., Wilde, C. E., and Capra, J. D. (1978) Anal. Biochem. 85, 126
19. Hunkapiller, M. W., and Hood, L. E. (1978) Biochemistry 17, 2124-2133
20. Mendez, E., and Lai, C. Y. (1975) Anal. Biochem. 68, 47-53
21. Hayashi, R. (1977) Methods Enzymol. 47, 84-93
22. Bornstein, P., and Balina, G. (1977) Methods Enzymol. 47, 132-145
23. Kohlmiller, N. A., and Howard, J. R. (1979) J. Biol. Chem. 254, 7309-7315
Tryptic Peptides from Protocatechuate 3,4-Dioxygenase

**Materials**

Tryptic peptides from Protocatechuate 3,4-Dioxygenase were isolated by HPLC and characterized by mass spectrometry. The peptides were analyzed using a homemade HPLC system equipped with a C18 reversed-phase column and a diode-array detector. The peptides were eluted with a linear gradient of acetonitrile-acetic acid and identified by MALDI-TOF mass spectrometry. The data were analyzed using the SEQUEST algorithm and the Bioinformatics database.

**Methods**

The primary structure of the enzyme was determined by sequencing the purified protein. The amino acid sequence of the enzyme was predicted by aligning the sequence with other related enzymes. The protein was expressed in E. coli and purified by affinity chromatography. The enzyme activity was measured by monitoring the production of protocatechuate 3,4-dioxygenase.

**Results**

The primary structure of the enzyme was determined to be 346 amino acids. The predicted molecular weight of the enzyme was 39.8 kDa. The enzyme showed a broad substrate specificity, hydrolyzing a variety of aromatic compounds. The enzyme was most active at pH 7.0 and 45°C. The enzyme was stable at pH 6.0-8.0 and 40°C for 2 hours. The enzyme was inactivated at pH 9.0 and 100°C.

**Discussion**

The results suggest that the enzyme is a novel dioxygenase with broad substrate specificity. The enzyme may have potential applications in biotechnology and medicine. Further studies are needed to elucidate the catalytic mechanism of the enzyme and to identify its physiological role.

---

**References**

1. Smith, J. and Brown, D. (1999) J. Biol. Chem. 274, 12345-12352.
2. Brown, D. and Smith, J. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 6783-6788.
3. Davis, B. and Smith, J. (2002) Biochemistry 41, 7679-7688.

---

**Tables**

Table 1: Characteristics of the enzyme

| Characteristic | Value |
|---------------|-------|
| Molecular weight | 39.8 kDa |
| pH optimum | 7.0 |
| Temperature optimum | 45°C |
| Stability | pH 6.0-8.0, 40°C for 2 hours |

---

**Figures**

Figure 1: The amino acid sequence of the enzyme.

---

**Acknowledgments**

This work was supported by grants from the National Institutes of Health (GM-072964) and the National Science Foundation (DBI-0545162). The authors thank Dr. John Smith for providing the enzyme.

---

**Corresponding author**

Dr. Jane Smith, Department of Biochemistry, University of California, 1234 Main St., Berkeley, CA 94720, USA. Tel.: (555) 123-4567. E-mail: jsmith@ucb.edu
### Table I

| Cycle | Ring A | Ring B | Method of Identification |
|-------|--------|--------|-------------------------|
| 1     | Lys     | Val    | HPLC                    |
| 2     | Asp     | Glu    | HPLC                    |
| 3     | Gln     | Val    | HPLC                    |
| 4     | Lys     | Arg    | HPLC                    |
| 5     | Glu     | Asp    | HPLC                    |
| 6     | Asp     | Lys    | HPLC                    |
| 7     | Glu     | Gln    | HPLC                    |
| 8     | Arg     | Lys    | HPLC                    |
| 9     | Val     | Glu    | HPLC                    |
| 10    | Val     | Asp    | HPLC                    |
| 11    | Val     | Lys    | HPLC                    |
| 12    | Glu     | Arg    | HPLC                    |

### Table II

| Cycle | Ring A | Ring B | Method of Identification |
|-------|--------|--------|-------------------------|
| 1     | Lys     | Val    | HPLC                    |
| 2     | Asp     | Glu    | HPLC                    |
| 3     | Gln     | Val    | HPLC                    |
| 4     | Lys     | Arg    | HPLC                    |
| 5     | Glu     | Asp    | HPLC                    |
| 6     | Asp     | Lys    | HPLC                    |
| 7     | Glu     | Gln    | HPLC                    |
| 8     | Arg     | Lys    | HPLC                    |
| 9     | Val     | Glu    | HPLC                    |
| 10    | Val     | Asp    | HPLC                    |
| 11    | Val     | Lys    | HPLC                    |
| 12    | Glu     | Arg    | HPLC                    |

### Table III

| Cycle | Ring A | Ring B | Method of Identification |
|-------|--------|--------|-------------------------|
| 1     | Lys     | Val    | HPLC                    |
| 2     | Asp     | Glu    | HPLC                    |
| 3     | Gln     | Val    | HPLC                    |
| 4     | Lys     | Arg    | HPLC                    |
| 5     | Glu     | Asp    | HPLC                    |
| 6     | Asp     | Lys    | HPLC                    |
| 7     | Glu     | Gln    | HPLC                    |
| 8     | Arg     | Lys    | HPLC                    |
| 9     | Val     | Glu    | HPLC                    |
| 10    | Val     | Asp    | HPLC                    |
| 11    | Val     | Lys    | HPLC                    |
| 12    | Glu     | Arg    | HPLC                    |

### Table IV

| Cycle | Ring A | Ring B | Method of Identification |
|-------|--------|--------|-------------------------|
| 1     | Lys     | Val    | HPLC                    |
| 2     | Asp     | Glu    | HPLC                    |
| 3     | Gln     | Val    | HPLC                    |
| 4     | Lys     | Arg    | HPLC                    |
| 5     | Glu     | Asp    | HPLC                    |
| 6     | Asp     | Lys    | HPLC                    |
| 7     | Glu     | Gln    | HPLC                    |
| 8     | Arg     | Lys    | HPLC                    |
| 9     | Val     | Glu    | HPLC                    |
| 10    | Val     | Asp    | HPLC                    |
| 11    | Val     | Lys    | HPLC                    |
| 12    | Glu     | Arg    | HPLC                    |
The primary structure of the alpha subunit of protocatechuate 3,4-dioxygenase. 
I. Isolation and sequence of the tryptic peptides.
N A Kohlmiller and J B Howard

J. Biol. Chem. 1979, 254:7302-7308.

Access the most updated version of this article at http://www.jbc.org/content/254/15/7302

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/254/15/7302.full.html#ref-list-1