Characterization of a Novel, Stable Norcaradiene Adduct Resulting from the Inactivation of Thymine Hydroxylase by 5-Ethynyluracil

Ming-tain Lai, Wei Wu, and JoAnne Stubbe*

Contribution from the Departments of Chemistry and Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received December 5, 1994[⊗]

Abstract: Thymine hydroxylase is an α-ketoglutarate, non-heme iron−dependent dioxygenase that catalyzes the conversion of thymine to its corresponding alcohol, aldehyde, and carboxylic acid in three steps, each accompanied by the conversion of α-ketoglutarate to succinate and CO_2. Studies by Thornburg and Stubbe (*Biochemistry* **1993,** *32*, 14034) showed that incubation of thymine hydroxylase with 5-ethynyluracil resulted in the production of 5-(carboxymethyl)uracil and uracil-5-acetylglycine and inactivation of the protein by covalent modification. Tryptic digestion of the inactivated protein followed by isolation of the modified peptides and their analysis by mass spectrometry revealed sequences (N)SIAFXSNPSLR, in which X was proposed to be a modified tyrosine residue. Recent efforts to clone the gene for thymine hydroxylase fortuitously resulted in isolation of the unmodified peptide. Sequencing of this peptide established that the amino acid residue modified by 5-ethynyluracil is a phenylalanine and not the predicted tyrosine! Two sets of experiments have been carried out to reveal the structure of the 5-ethynyluracil-modified phenylalanine. Incubation of 5-ethynyluracil with thymine hydroxylase in the presence of $^{18}O_2$ revealed, subsequent to tryptic digestion and peptide isolation, that two atoms of oxygen derived from $^{18}O_2$ have been incorporated. A similar experiment using [2-^{14}C]-5-[1′, 2′-$^{13}C_2$]ethynyluracil resulted in the isolation of a sufficient amount of modified peptide for analysis by 1D and 2D NMR spectroscopy. This analysis revealed a novel 7-carboxylated norcaradiene moiety. A mechanism involving partitioning of a carbene intermediate between insertion into the phenylalanine residue of the protein and rearrangement to generate a ketene is proposed to account for the structure of the peptide adduct and the previously characterized small molecule products (5-(carboxymethyl)-uracil and uracil-5-acetylglycine). $^{18}O_2$-Labeling experiments and the presence of a carboxylic acid in the adduct suggest that thymine hydroxylase carries out a second hydroxylation reaction while the oxidized inhibitor is covalently bound in the enzyme's active site.

Introduction

An understanding of the chemistry and enzymology of mononuclear and dinuclear non-heme iron-dependent reactions of biological interest has lagged far behind that of their ostentatious cousins, the heme-requiring systems.[1,2] While the well-characterized heme systems have served as prototypes for the non-heme iron systems,[3−5] it is clear from recent studies on dinuclear non-heme iron requiring proteins such as ribonucleotide reductase[6−8] and methane monooxygenase[9,10] that many surprises are in store.

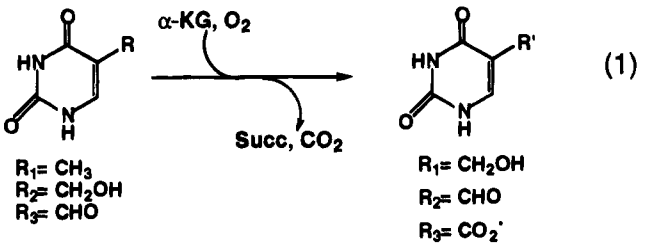
One subclass of the mononuclear non-heme iron-requiring proteins are the α-ketoglutarate (α-KG)-dependent dioxygenases.[11] In mammalian systems, these enzymes play an essential role in posttranslational modification of proline, lysine, aspartic acid, and asparagine residues in polypeptides such as the essential extracellular matrix protein collagen[12] and in the blood clotting cascade.[13] We have focused our efforts toward understanding the mechanism of these reactions on a thymine hydroxylase from *Rhodotorula glutinis*.[14−17] This protein can catalyze three successive, but not processive, hydroxylation reactions (eq 1). Its promiscuity with respect to substrate specificity[18] and the ease of synthesis of the substrate analogs make it an excellent choice for the investigation of mechanistic questions. The ability of thymine hydroxylase to catalyze the epoxidation of vinyluracil led us to successfully design 5-ethy-

* Author to whom correspondence should be addressed.

⊗ Abstract published in *Advance ACS Abstracts*, April 15, 1995.

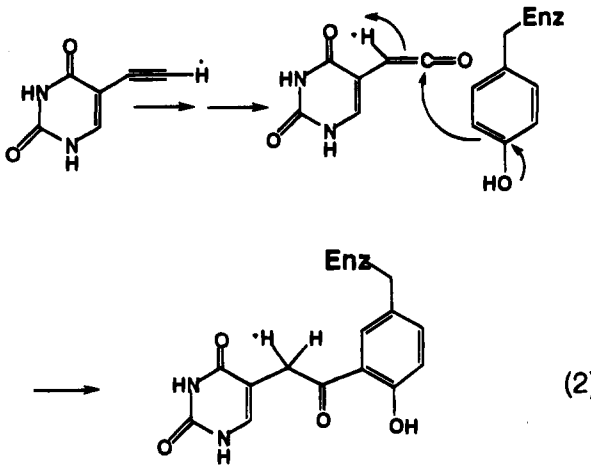
(1) Stubbe, J. *Curr. Opin. Struct. Biol.* **1991**, *1*, 788.
(2) Feig, A.; Lippard, S. J. *Chem. Rev.* **1994**, *94*, 759.
(3) White, R. E.; Coon, M. J. *Annu. Rev. Biochem.* **1980**, *49*, 315.
(4) McMurry, T. J.; Groves, J. T. In *Cytochrome P-450: Structure, Mechanism, and Biochemistry*; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, 1986; p 1.
(5) Ortiz de Montellano, P. R. In *Cytochrome P-450: Structure, Mechanism, and Biochemistry*; Ortiz de Montellano, P. R., Ed.; Plenum press, New York, 1986, p 217.
(6) Bollinger, J. M., Jr.; Tong, W. H.; Ravi, N.; Huynh, B. H.; Edmondson, D. E.; Stubbe, J. *J. Am. Chem. Soc.* **1994**, *116*, 8015.
(7) Bollinger, J. M., Jr.; Tong, W. H.; Ravi, N.; Huynh, B. H.; Edmondson, D. E.; Stubbe, J. *J. Am. Chem. Soc.* **1994**, *116*, 8024.
(8) Ravi, N.; Bollinger, J. M., Jr.; Huynh, B. H.; Edmondson, D. E.; Stubbe, J. *J. Am. Chem. Soc.* **1994**, *116*, 8007.
(9) Liu, K. E.; Wang, D.; Hunyh, B. H.; Edmondson, D. E., Salifoglou, A.; Lippard, S. J. *J. Am. Chem. Soc.* **1994**, *116*, 7465.
(10) Lee, S. K.; Fox, B. G.; Troland, W. A.; Lipscomb, J. D.; Münck, E. J. *J. Am. Chem. Soc.* **1993**, *115*, 6450.
(11) Hanauske-Abel, H. M.; Günzler, V. *J. Theor. Biol.* **1982**, *94*, 421.
(12) Kivirikko, K. I.; Myllyla, R. *Methods Enzymol.* **1987**, *144*, 97.
(13) Jia, S.; McGinnis, K.; Burke, C. J.; VanDusen, W. J.; Kuo, A.; Griffin, P. R.; Elliston, K. O.; Stern, A. M.; Friedman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7227.
(14) Wondrack, L. M.; Hsu, C.-A.; Abbott, M. T. *J. Biol. Chem.* **1978**, *253*, 6511.
(15) Wondrack, L. M.; Warn, B. J.; Saewert, M. D.; Abbott, M. T. *J. Biol. Chem.* **1979**, *254*, 26.
(16) Warn-Cramer, B. J.; Macrander, L. A.; Abbott, M. T. *J. Biol. Chem.* **1983**, *258*, 10551.
(17) Liu, C.-K.; Hsu, C.-A.; Abbott, M. T. *Arch. Biochem. Biophys.* **1973**, *159*, 180.
(18) Thornburg, L. D.; Lai, M.-t.; Wishnok, J. S.; Stubbe, J. *Biochemistry* **1993**, *32*, 14023.

0002-7863/95/1517-5023$09.00/0



(1)

nyluracil as a mechanism-based inhibitor of this protein.[18,19] We have shown that incubation of 5-ethynyluracil with thymine hydroxylase results in a partitioning between production of 5-(carboxymethyl)uracil and uracil-5-acetylglycine (the latter compound is a result of the glycine in the buffer) and inactivation of the protein with stoichiometric labeling. The stability of thymine hydroxylase with the inhibitor allowed isolation and characterization of a modified peptide, (N)-SIAFXSNPSLR, where X was proposed to be a tyrosine residue which was acylated at a carbon rather than at its chemically more reactive oxygen (eq 2).[19]



(2)

Our recent effort to obtain peptide information that would allow us to clone the gene for thymine hydroxylase fortuitously allowed sequencing of the same peptide that was identified in the previous studies with 5-ethynyluracil. The amino acid modified by ethynyluracil is phenylalanine, not tyrosine as originally proposed. These intriguing results provided the impetus for the present studies which describe isolation of sufficient quantities of the major modified peptide (**1**) to establish its structure using 1D and 2D NMR spectroscopic methods. This scale-up procedure allowed identification of a second modified inhibitor adduct (**2**) attached to the same phenylalanine in the same peptide. In addition, the major peptide adduct has been examined by mass spectrometric methods subsequent to tryptic degradation and HPLC purification. A hypothesis is presented involving carbene and ketene intermediates to account for partitioning between the formation of **1** and the small molecule products.

Experimental Section

Materials. [1,2-$^{13}C_2$]-1-(Trimethylsilyl)acetylene (99% ^{13}C enriched) was custom made by Isotec Inc. Thymine, α-KG, sodium ascorbate, protamine sulfate (grade x, from salmon), TPCK-trypsin (trypsin treated with *N*-tosylphenylalanine chloromethyl ketone), and hyamin hydroxide (1 M solution in methanol) were purchased from Sigma. Ferrous sulfate was purchased from Mallinkrodt. Carboxypeptidase Y was supplied by Boehringer Mannheim. Triton X-100 (hydrogenated) was purchased from Calbiochem. Immobilon-P poly(vinyl difluoride) (PVDF) membranes (pore size 0.45 μm) were supplied by Millipore. [1-^{14}C]-α-KG (59 mCi/mmol) and [2-^{14}C]thymine (56 mCi/mmol) were purchased from New England Nuclear and Moravek Biochemicals, respectively. Microbore (1.0 × 250 mm), analytical (4.6 × 250 mm), and semi-preparative (7 × 250 mm) Vydac C_{18} columns were purchased from Vydac Inc. $^{18}O_2$ (99.8%) was obtained from Amersham. *R. glutinis* was supplied by American Type Culture Collection (ATCC no. 2527).

Methods. Thymine hydroxylase from *R. glutinis* (specific activity, 14.5 units/mg) was isolated and assayed according to the procedure of Thornburg et al.[18] [2-^{14}C]-5-[1′,2′-$^{13}C_2$]Ethynyluracil was prepared from [1, 2-$^{13}C_2$]-1-(trimethylsilyl)acetylene as previously described in 30% overall yield.[19] ^{1}H-NMR (DMSO-d_6): δ 3.97 (d, 1H, acetylenic H, $J_{^{13}CCH}$ = 53.5 Hz, $J_{^{13}CH}$ = 251 Hz), 7.82 (s, 1H, C6H), 11.20 (br s, 2H, N1H and N3H). The ^{13}C chemical shifts of the carbons in the acetylenic side chain were observed at 76.2 and 84.7 ppm ($J_{^{13}C-^{13}C}$ = 178 Hz).

The protein concentration was determined according to the method of Lowry et al. using bovine serum albumin (BSA) as a standard.[20] Liquid scintillation counting was performed with a Packard TriCarb 1500 counter using Scint-A (Packard) fluid. HPLC purification was carried out on a Beckman system consisting of a variable wavelength detector (Beckman Model 163). UV spectra were recorded on a Hewlett-Packard 8452A spectrophotometer. All ^{1}H- and ^{13}C-NMR spectra were recorded on either a Varian VXR 500 MHz instrument or a home-built 500 MHz instrument at the Francis Bitter National Magnet Laboratory. Data were then transferred to a Silicon Graphics workstation and processed using Felix software (version 2.3, Biosym Technologies, Inc.). All chemical shifts were referenced to an internal standard, sodium 3-(trimethylsilyl)-1-propanesulfonate (TSP). In the case of the ^{13}C spectrum, the methyl carbon of TSP was used as the reference (0.00 ppm).

Identification of Phenylalanine as the 5-Ethynyluracil-Modified Amino Acid. The following reaction mixture had in a final volume of 220 μL: thymine hydroxylase (48 μg; specific activity, 11.5 units/mg), 0.45 mM α-KG, 2.3 mM ascorbate, 0.11 mM ferrous sulfate, and 45 mM HEPES (pH 7.5). 5-Ethynyluracil (15 nmol in 10 μL of DMSO) was added to start the reaction. In the control reaction mixture, identical to that described above, 10 μL of DMSO was added. The reactions were allowed to proceed for 45 min and were then subjected to SDS–PAGE in a 0.75 mm thick 10% polyacrylamide gel.[21] The protein was transferred to an Immobilon-P PVDF membrane by electroblotting, as described by Matsudaira.[22] Protein was detected by staining in an aqueous solution containing 0.1% Coomassie R-250, 1% acetic acid, and 40% methanol for 10 min, followed by destaining with 50% methanol. The proteins were submitted to the MIT biopolymer laboratory, where they were treated with sequencing grade trypsin and analyzed by microbore Vydac C_{18} reverse phase chromatography using a flow rate of 0.15 mL/min with detection at 220 nm. Solvent A contained 0.1% trifluoroacetic acid (TFA), and solvent B was 0.085% TFA in 80% acetonitrile. Peptides were eluted by a linear gradient from 0 to 37.5% B over 60 min, followed by a linear gradient to 75% B over the next 30 min.

Large Scale Purification of the Major and Minor 5-Ethynyluracil-Modified Peptides (1 and 2). The following reaction mixture had a final volume of 50 mL: 45 mM HEPES (pH 7.5), 10 mM [2-^{14}C]-5-[1′, 2′-$^{13}C_2$]ethynyluracil (specific activity, 5 × 10^6 units/μmol), 0.45 mM α-KG, 0.11 mM $FeSO_4$, 2.3 mM ascorbate, 11 mM glycine, 2.3 mM potassium phosphate and 0.2% DMSO, and 600 mg of thymine hydroxylase (specific activity, 3.7 units/mg). After the reaction mixture was incubated for 45 min at 25 °C, it was loaded onto a Sephadex G-25 column (5.0 × 54 cm) and eluted with ammonium formate (pH 7.5). The fractions containing protein were pooled and lyophilized. The residue was suspended in 100 mL of 100 mM Tris (pH 8.0), 1% Triton X-100, and 10% acetonitrile, 9 mg of TPCK-trypsin was added, and the reaction was allowed to proceed for 24 h at 37 °C. The solution was lyophilized, and the mixture was analyzed on a semipreparative Vydac C_{18} column (flow rate, 3 mL/min) under the same conditions as described above. Radioactivity was detected by scintillation counting

(19) Thornburg, L. D.; Stubbe, J. *Biochemistry* **1993**, *32*, 14034.

(20) Lowry, O. H.; Rosebrough, N. J.; Farr, A. L.; Randall, R. J. *J. Biol. Chem.* **1951**, *193*, 265.

(21) Laemmli, U. K. *Nature (London)* **1970**, *27*, 680.

(22) Matsudaira, P. *J. Biol. Chem.* **1987**, *62*, 10035.

in fractions which eluted at 45 and 47 min. Each fraction was lyophilized and repurified on a Vydac C_{18} column using solvent A, with 20 mM potassium phosphate (pH 7.2), and solvent B, with 100% acetonitrile. Linear gradients were run from 0 to 15% B over 10 min, then from 15 to 30% B over 20 min, followed by isocratic elution at 30% B for 10 min. The major peptide **1** eluted at 27 min, and the minor peptide **2** eluted at 30 min. Desalting of both peptides was carried out by dissolving the peptides in H_2O, injecting them onto the Vydac C_{18} column, and washing the column with H_2O for 15 min, followed by eluting **1** and **2** with 100% CH_3OH for 10 min. The overall recovery of **1** was 30%, and that of **2** was 10% from the original labeled protein. Peptide **1** was analyzed by NMR spectroscopy, and **1** and **2** were examined by FAB-MS.

Degradation of Peptide 1 with Carboxypeptidase Y. Carboxypeptidase Y (25 μg) was added to 50 nmol of peptide **1** in 50 mM sodium citrate (pH 6.0), for a final volume of 1 mL. This reaction mixture was incubated at 37 °C for 3 days, and an additional 25 μg of carboxypeptidase Y was added every 12 h. The digested peptide was analyzed on a Vydac C-18 column using an isocratic elution with 0.1% TFA in 40% methanol (flow rate, 1 mL/min). Radioactive material (35% recovery) eluted at 13.5 min, and the appropriate fractions were pooled, lyophilized, and analyzed by NMR spectroscopy and CID mass spectrometry.

Decomposition of [2-^{14}C]-5-[1′, 2′-$^{13}C_2$]Ethynyluracil Nonenzymatically to 5-Acetyluracil. From the large scale inactivation of thymine hydroxylase with [2-^{14}C]-5-[1′, 2′-$^{13}C_2$]ethynyluracil described above, fractions containing the small molecules from the Sephadex G-25 column were pooled, lyophilized, and analyzed by HPLC on a semipreparative C_{18} Alltech column eluted with H_2O at a flow rate of 3 mL/min. A compound with a retention time of 36.8 min, which was also detected in the control experiment, was collected and characterized. ^{1}H-NMR (DMSO-d_6): δ 2.41 (dd, 3H, $J_{^{13}CH}$ = 125.0, $J_{^{13}CCH}$ = 6.7 Hz, Me), 8.05 (s, 1H, H6). ^{13}C-NMR (DMSO-d_6): δ 30.0 (d, J = 89 Hz, CH_3), 193.8 (d, J = 89 Hz, C=O).

NMR Spectroscopic Analysis of Peptide 1. Peptide **1** was lyophilized twice with 99.96% D_2O and then dissolved in 99.996% D_2O. DQF-COSY and TOCSY[23] (MLEV-17 spin-lock pulse with 35 and 70 ms mixing times) data were recorded at 25 °C. Data sets with 4096 × 512 complex points were acquired with sweep widths of 5500 Hz in both dimensions and 32 scans per t_1 increment. During the relaxation delay period, a 1.5 s presaturation pulse was used for solvent suppression. The t_1 dimension was zero-filled to 4096 data points, and spectra were processed with a combination of exponential and Gaussian weighting functions. Ridges in t_1 were reduced by multiplying the first point in t_1 by 0.5 prior to the Fourier transform.

The one-bond (HMQC, ^{1}H-detected heteronuclear multiple-quantum coherence) and multiple-bond (HMBC, ^{1}H-detected heteronuclear multiple-bond correlation)[24,25] ^{1}H−^{13}C correlation spectra were also recorded at 25 °C. For the HMQC experiment, a data set with 2048 × 256 complex points was acquired with 6000 and 25 000 Hz sweep widths in the proton and the carbon dimensions, respectively, and 128 scans were collected for each t_1 increment. The one-bond heteronuclear coupling (J_{CH}) was set at 160 Hz. The t_1 dimension was zero-filled to 2048 data points. For the HMBC experiments ($\Delta 1$ = 3.4 ms, $\Delta 2$ = 80 ms), a data set with 4096 × 512 complex points was acquired with 6000 and 25 000 Hz sweep widths in the proton and the carbon dimensions, respectively, and 64 scans were collected per t_1 increment. The t_1 dimension was zero-filled to 4096 data points. Both spectra were then processed with exponential weighting functions.

$^{18}O_2$ Labeling Studies. A reaction mixture containing 45 mM HEPES (pH 7.5), 0.45 mM α-KG, 0.11 mM $FeSO_4$, 2.3 mM ascorbate, 23.3 mM 5-ethynyluracil, and 0.2% DMSO in a total volume of 1.5 mL was degassed by eight freeze−evacuate−thaw cycles and equilibrated in an atmosphere of $^{18}O_2$ prior to addition of the enzyme. Thymine hydroxylase (20 μL, 1.5 nmol; specific activity, 11.5 unit/mg) was then added to initiate the reaction, and the resulting solution was incubated at room temperature for 45 min. Modified protein was separated from small molecules by Centricon P-10 ultrafiltration (6000 rpm, 30 min). The protein retentate was then lyophilized, and the resulting residue was digested and purified as described above. Peptide **1** was then submitted for FAB-MS analysis.

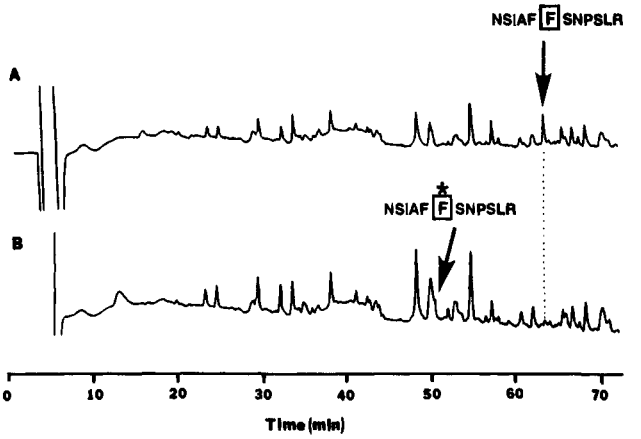


Figure 1. Peptide mapping of the thymine hydroxylase. (A) Trypsin treatment of thymine hydroxylase followed by chromatographic separation of the resulting peptides *via* Vydac C_{18} column chromatography (Methods, Experimental Section). (B). As in (A), except that thymine hydroxylase has been inactivated by 5-ethynyluracil.

Results

Identification of a F Residue in Peptide NSIAF*F*SNPSLR from Thymine Hydroxylase as the Unique Target of Ethynyluracil. A peptide mapping experiment was undertaken in an effort to obtain sequence information on thymine hydroxylase to assist in cloning its gene and in an attempt to identify the amino acid residue in the peptide (NSIAFXSNPSLR) modified by 5-ethynyluracil. Thus, thymine hydroxylase and thymine hydroxylase inactivated with 5-ethynyluracil were purified by SDS−PAGE and electroblotted onto a PVDF membrane. The proteins were then treated with trypsin and analyzed using Vydac C-18 column chromatography. The results are shown in Figure 1. The peak with a retention time of 63.1 min in the native protein is missing in the profile of the thymine hydroxylase inactivated protein. Furthermore, a new peak appears as a shoulder of a peak eluting at 50.1 min. The peak eluting at 63.1 min was collected, sequenced by the automated Edman degradation method, and analyzed by FAB-MS. The amino acid sequence revealed that the amino acid component of X is a phenylalanine. Mass spectrometric analysis of the modified and unmodified peptides (1520.3 vs 1352.4, respectively) revealed a difference of 168 amu, consistent with addition of the inhibitor (136) and two oxygen atoms. Finally, additional peptides were isolated and sequenced, and this information is presently being used in an effort to clone the gene for thymine hydroxylase.

Source of "Putative" Oxygens in the Ethynyluracil-Modified Peptide. Identification of the missing 32 amu as belonging to two oxygen atoms, as well as distinguishing between the two possible sources of these oxygens, O_2 and H_2O, is crucial to defining the chemistry of inhibitor inactivation. Therefore, the inactivation of thymine hydroxylase by 5-ethynyluracil was carried out in the presence of $^{18}O_2$. The modified peptide was isolated and analyzed using FAB-MS (Figure 2). The molecular ion is at 1524.3 amu, 4 units higher than the corresponding peptide from the control experiment with $^{16}O_2$. These results establish that the difference of 32 amu can be attributed to the incorporation of two oxygens and that both oxygens are derived from O_2. Furthermore, a second peak is observed in the mass spectrum at 1522.3 amu, suggesting that one or both of these oxygens can undergo exchange with H_2O

(23) For a review of COSY, TOCSY, HMQC, and HMBC experiments, see: Martin, G. E.; Zektzer, A. S. *Two Dimensional NMR Methods for Establishing Molecular Connectivity*; VCH Publishers: New York, 1988.
(24) Bax, A.; Subramanian, S. *J. Magn. Reson.* **1986**, *67*, 565
(25) Bax, A.; Summers, M. F. *J. Am. Chem. Soc.* **1986**, *108*, 2093

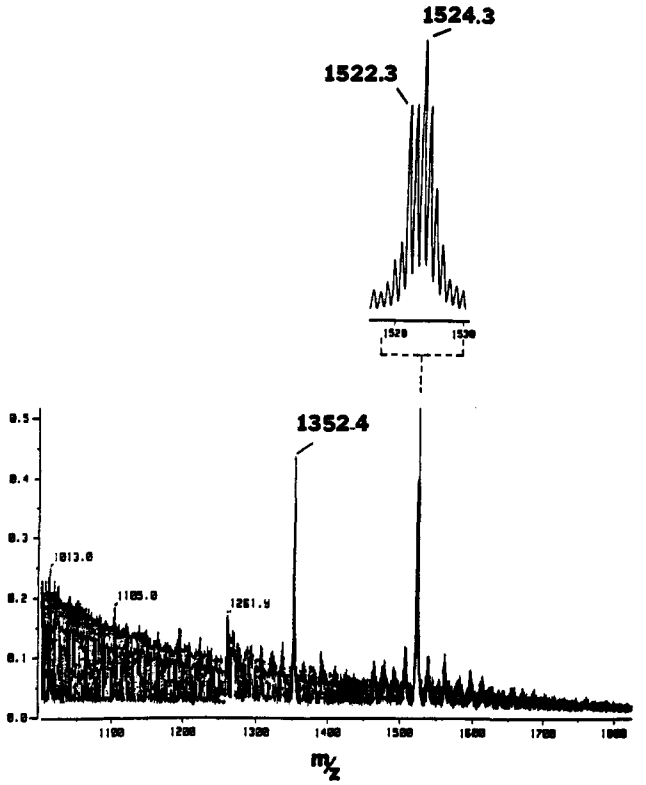


Figure 2. FAB mass spectrum of the major modified peptide **1**. 5-Ethynyluracil was incubated with thymine hydroxylase in the presence of $^{18}O_2$ (Methods, Experimental Section). The inactivated enzyme was isolated, cleaved with trypsin, and analyzed by RPHPLC, as described in Figure 1.

at some stage leading to or subsequent to product formation. A third distinctive feature of the FAB mass spectrum (Figure 2) is the presence of a peak at 1352.4 amu, corresponding to the unmodified peptide. Thus, under the conditions of the mass analysis, the peptide adduct is capable of reverting back to the unmodified peptide. This observation provides insight into the structure of the adduct, indicating that the linkage between the inhibitor and the amino acid is quite labile, as fragmentation is not usually observed under the mild ionization conditions used in FAB-MS.[26]

The $^{18}O_2$-labeling experiments gave surprising results, given that all α-KG dioxygenases catalyze reactions in which one atom of oxygen is found in hydroxylated product and the other oxygen atom is found in succinate.[11] Thus, one interpretation of the presence of two oxygens derived from O_2 is that two successive hydroxylation reactions have been carried out by thymine hydroxylase during its inactivation by 5-ethynyluracil.

Identification of the Structure of the Major Modified Peptide 1 by 1D and 2D NMR Spectroscopic Methods. A large scale inactivation of thymine hydroxylase (600 mg) by [2-^{14}C]-5-[1′,2′-$^{13}C_2$]ethynyluracil was carried out in order to obtain sufficient material to analyze the adduct by NMR spectroscopic methods. Under these conditions, subsequent to tryptic digestion and RP-HPLC chromatography, in addition to the expected peptide **1** (48.5% of the material), a second minor peptide **2** (18% of the material) was identified (Figure 3). Approximately 20% of the radioactive material eluted in the void volume of the column, presumably as labile adducts, and was not further analyzed. Both peptides were further purified by rechromatography, as indicated in Figure 4.

The ^{1}H-NMR spectrum of **1** is shown in Figure 5. This spectrum reveals several remarkable features. There are two resonances, one at 5.87 and a second at 6.20 ppm, which are olefinic in nature, indicating that the aromatic ring of the modified phenylalanine has been converted into a diene. Examination of the region downfield from the olefinic protons reveals a resonance at 7.08 ppm and a group of resonances between 7.27 and 7.40 ppm. The 7.08 ppm peak is assigned to the H6 proton of a uracil ring, while the downfield peaks are assigned to the aromatic protons of a second phenylalanine residue found in peptide **1**. Integration of these groups of protons reveals ratios of 5:1:1:2, from the aromatic region upfield to the olefinic region. These results strongly suggest that the phenylalanine ring modified by the inhibitor is a trisubstituted cyclohexadiene.

A DQF-COSY NMR spectrum of **1** is shown in Figure 6. A cross peak is observed between the resonances at 6.20 and 5.87 ppm, clearly indicating coupling between these two sets of protons. In addition, the proton at 6.20 ppm is coupled to a proton(s) at 2.98 ppm, while a proton at 5.87 ppm is coupled to a proton(s) at 3.00 ppm. The protons with chemical shifts at 3.00 and 2.98 ppm are assigned as arising from the disruption of the aromaticity of the benzene ring of the phenylalanine of **1**, as predicted from the trisubstituted cyclohexadiene structure.[27]

Unfortunately, the two protons at 5.87 ppm exhibit such extensive spectral overlap at 500 MHz that no coupling constants could be obtained. In an effort to obtain this information, peptide **1** was further degraded using carboxypeptidase Y. While this process left much to be desired, a heptameric peptide was obtained that contained the modified amino acid as determined by CID mass spectrometric analysis. A 1D NMR spectrum of this new material revealed that the two protons at 5.87 ppm were now resolved (5.90 and 5.93 ppm), which allowed the coupling constants to be determined. The olefinic proton at 6.20 ppm is composed of a doublet of doublets with $J = 5.5$ and 9.0 Hz, while the proton at 5.93 ppm is a doublet ($J = 3.3$ Hz), as is the proton at 5.90 ppm ($J = 9.0$ Hz). Thus, the proton at 6.20 ppm is coupled to the proton at 5.90 ppm and one of the two protons at ~3 ppm, and the proton at 5.93 ppm is coupled to one of the two protons at ~3 ppm. These coupling constants are also consistent with a trisubstituted cyclohexadiene system.[27]

We originally synthesized 5-ethynyluracil with both of its ^{13}C-acetylenic carbons 100% enriched to assist in the structural assignment of **1**. Using one-bond ^{1}H−^{13}C correlation spectroscopy (HMQC method), we anticipated that we could easily identify protons directly attached to the ^{13}C-labeled carbons in **1** as the cross peaks would have much greater intensity than those attached to non-^{13}C-enriched carbons. To our surprise, no cross peaks with significantly enriched intensities were observed. All of the cross peaks, however, were sufficiently intense that chemical shifts of carbons coupled to previously assigned protons could be made (Table 1, Figures 8 and 9).

The lack of detection of protons attached to ^{13}C-labeled carbons precipitated the acquisition of a 1D ^{13}C-NMR spectrum of **1** (Figure 7). The spectrum revealed two doublets: one at 184.7 ppm and one at 23.5 ppm. The resonance at 184.7 ppm is indicative of a carboxylic acid carbon, while the resonance at 23.5 ppm is indicative of a quaternary carbon with an unusual environment, perhaps part of a cyclopropyl ring. Both of these assignments are consistent with the lack of proton coupling to the ^{13}C-enriched carbons determined by the HMQC method and the proton coupled and decoupled 1D ^{13}C-NMR experiments. The ^{13}C−^{13}C coupling constant of 66.7 Hz for **1** is very similar

(26) Fenselau, C.; Cotter, R. J. *Chem. Rev.* **1987**, *87*, 501.

(27) Cheng, C. C. Ph. D. Thesis, Massachusettes Institute of Technology, Cambridge, MA, 1994.

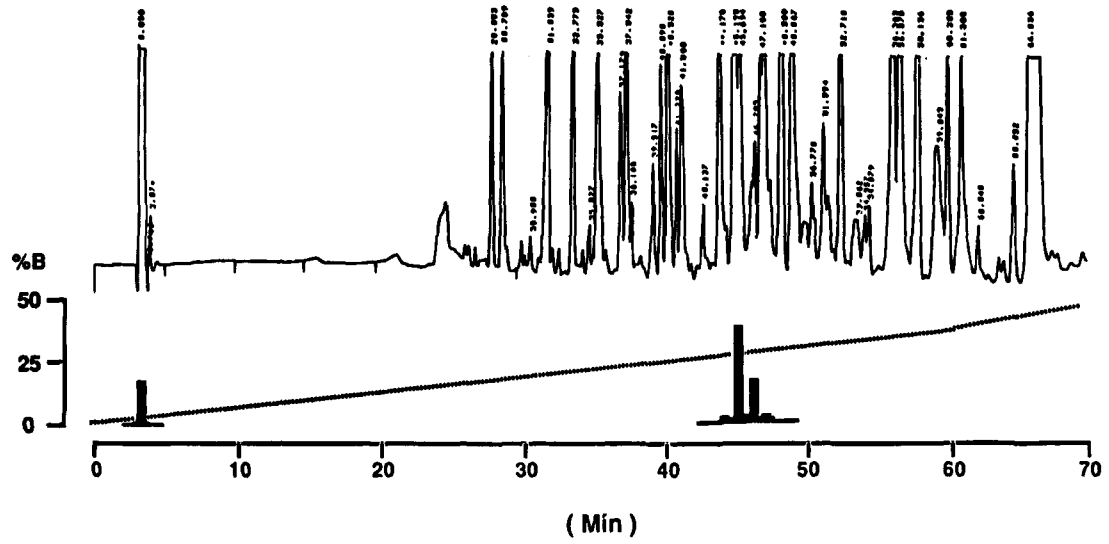


Figure 3. Purification by RP-HPLC using a Vydac C-18 column of the peptide **1** and **2** generated by 5-ethynyluracil inactivation of thymine hydroxylase. (Top) The separation (Methods, Experimental Section) employed a gradient using solvent A (0.1% TFA) and solvent B (0.085% TFA in 80% acetonitrile with monitoring at 220 nm). (Bottom) Radioactivity associated with the various peptides.

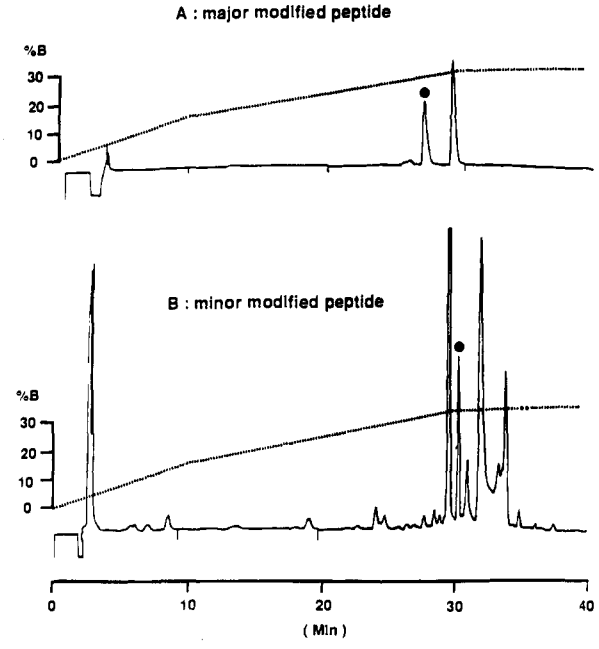


Figure 4. Repurification of peptides **1** (A) and **2** (B) from Figure 3. The separation (Methods, Experimental Section) employed a gradient using solvent A (20 mM potassium phosphate, pH 7.2) and solvent B (acetonitrile). ● indicates radiolabeled peptide.

to the value of 73.7 Hz reported for a one-bond coupling between a cyclopropyl carbon and a carboxylic acid.[28]

The results from both the ^1H- and ^{13}C-NMR experiments are consistent with the norcaradiene (NCD) ring structure shown in Figure 8. Further evidence in support of this structure is derived from multiple-bond $^1H-^{13}C$ correlation spectroscopy (HMBC method) of **1**. On the basis of the structure in Figure 8, we expected to see a number of significant cross peaks. We observed that the proton at 7.08 ppm (C6 H of uracil) is strongly coupled to the carbon at 23.5 ppm (C7 of the norcaradiene). In addition, a moderately strong coupling is also detected between the carboxylic acid carbon at 184.7 ppm and the C1 or C1 and C6 protons of the norcaradiene ring. Very weak couplings are also observed, as indicated in Table 2. These observations are

(28) Marshall, J. L. *Carbon–Carbon and Proton–Proton NMR Coupling*; Verlag Chemie International: Florida, 1983; p 65.

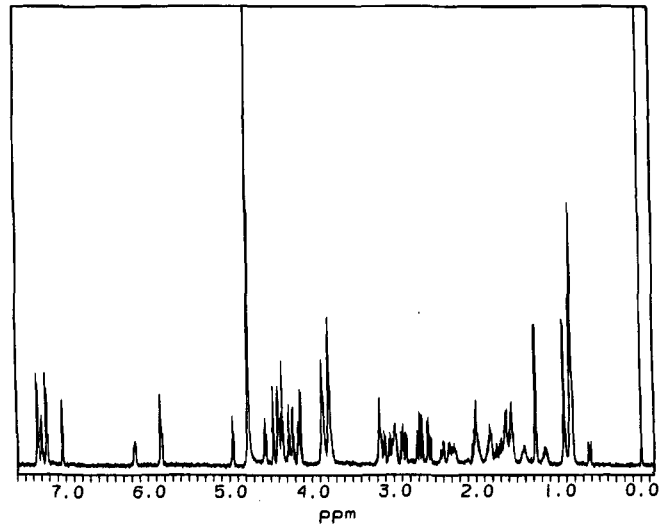


Figure 5. ^1H NMR sepctrum of peptide **1** in D_2O.

all consistent with the structure proposed in Figure 8. In general, NCD ring systems are found to be in equilibrium with their valence tautomers, cycloheptatrienes (CHTs).[29–31] Further evidence supporting our structural assignment and an explanation for the stability of our NCD will be discussed subsequently.

Characterization of the Minor Modified Peptide 2. The minor peptide **2** (Figure 4B) was analyzed by Edman degradation sequencing and by mass spectrometric methods. Sequencing revealed a peptide NSIAFXSNPSLR, identical to peptide **1**. FAB mass spectrometric analysis revealed a molecular weight for **2** of 1490.2, consistent with an unmodified peptide (1352.4), 5-ethynyluracil (136), and two additional protons. Other more esoteric interpretations of these data, however, are also possible. The sequencing data in conjunction with the mass spectrometric data, moreover, indicate that the same phenylalanine is modified as was observed in peptide **1**, but that the modification is unique. Peptide **2** amazingly does not appear to contain any additional oxygens, although the obligate $^{18}O_2$ experiment has not yet been carried out. At present, sufficient information is not available to propose a structure for adduct **2**.

(29) Ciganek, E. *J. Am. Chem. Soc.* **1967**, *89*, 1454.
(30) Hall, G. E.; Roberts, J. D., *J. Am. Chem. Soc.* **1971**, *93*, 2203.
(31) Takeuchi, K.; Senzaki, Y.; Okamoto, K., *J. Chem. Soc., Chem. Commun.* **1984**, 111.

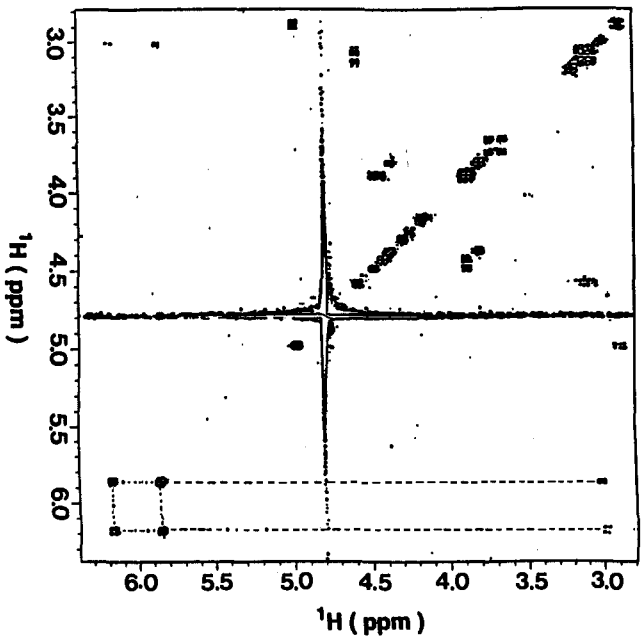


Figure 6. DQF-COSY spectrum of peptide **1** in D_2O. The boxed regions show the cross peaks of interest.

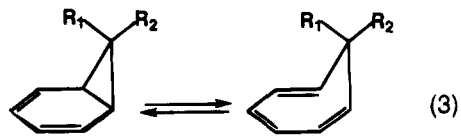
Table 1. Comparison of the Key ^{1}H and ^{13}C Chemical Shifts between Model Compound and Major Adduct

	C6	C1	C7	H6
model compound, Figure 9	39.34	40.20 or 22.94	22.94 or 40.20	2.64
adduct, Figure 8[a]	43.1	43.1	23.5	3.00 or 2.98

[a] 1H Chemical shifts of the olefinic protons of the adduct: δ 6.20 (dd, J = 5.5, 9.0 Hz, H2), 5.93 (d, J = 3.3 Hz, H5), 5.90 (d, J = 9.0 Hz, H3).

Discussion

The NCD structure, Figure 8, that we have proposed for the adduct of peptide **1** is unique, given that almost all previously reported 7-substitiuted NCDs are unstable and in equilibium with their valence tautomers, the CHTs (eq 3).[32] Therefore,



(3)

the basis for our structural assignment and the stability of our adduct require further comment. A search of the literature for molecules similar to that shown in Figure 8 revealed the "unusually stable" γ-lactone ring-fused NCD shown in Figure 9.[32] The structure where X = H was established by an X-ray crystallographic determination. The ^{13}C and ^{1}H chemical shifts for these compound were reported without definitive assignments. Our interpretation of these results in comparison with our own data on adduct 1 is presented in Table 1. Our results suggest that the ^{13}C chemical shift for the model C7 (Figure 9) is 22.94 ppm, in comparison with ours at 23.5 ppm (Figure 8), and that the model C6 shift is at 40.20 ppm, in comparison with our C1 and C6 shifts at 43.1 ppm. It should be pointed out that all of our spectra have been recorded in D_2O, while the spectrum of the model compound was recorded in $CDCl_3$. This difference in solvents is expected to have a greater perturbation on the proton than on the carbon chemical shifts. The chemical shift of the model H6 proton is reported to be 2.64 ppm, while our H1 and H6 protons are at 2.98 and 3.00 ppm. This well-characterized model compound thus supports our structural assignment.

Finally, we would like to come back to the stability of our adduct **1** in comparison with other NCD systems. Studies by many investigators have revealed that the equilibrium between the NCD and the CHT lies far on the side of the CHT.[33] However, a number of factors have been identified to shift the equilibrium toward the NCD side. The most extensively documented is the presence of a π-acceptor substituent at the C7 position. In our adduct, both the carboxyl group and the uracil ring offer the opportunity for efficient conjugation of the occupied cyclopropane Walsh orbital with the unoccupied π-acceptor orbitals of our C7 substituents.[34] This conjugation strengthens the C1−C6 bond. Alkyl substituents at the C2, C3, and or C4 postitions have also been reported to shift the equilibrium toward the NCD side, destabilizing the CHT structure. Thus, the observation that the model compound shown in Figure 9 is stable and that **1** has C7 and C4 substituents favoring the NCD tautomer suggests that our structural assignment is correct and provides at least a tentative explanation for its stability.

Given this structure and the previous identification of 5-(carboxymethyl)uracil and uracil-5-acetylglycine as the products accompanying thymine hydroxylase inactivation,[19] a mechanism can be postulated (Scheme 1) to accommodate the observed results. This model proposes that the active hydroxylating agent in thymine hydroxylase is capable of generating an internal carbene.[35] This carbene can then partition between insertion into one of the double bonds of phenylalanine to generate an aldehydic NCD structure and a 1,2-hydride rearrangement reaction to generate a ketene, which can then be trapped by H_2O, glycine (from the buffer), or some nucleophilic group on the protein.[36,37] The observation that incubation of [^{3}H]ethynyluracil with thymine hydroxylase results in ^{3}H-labeled (carboxymethyl)uracil and uracil-5-acetylglycine is consistent with this proposal.[38]

The mass spectrum of adduct **1** formed during the inactivation reaction, however, indicates, in contrast to expectations based on the normal labeling patterns for dioxygenases, that **1** possesses two atoms of oxygen, both derived from oxygen gas. To account for this observation, one can postulate that thymine hydroxylase during the first turnover generates an aldehydic intermediate (Scheme 1) and then carries out a second hydroxylation reaction to generate the carboxylic acid, while the aldehyde is held covalently within the enzyme's active site. While it has been previously demonstrated that thymine hydroxylase can catalyze a similar oxidation reaction, it should be noted that the position of the carbon undergoing oxidation is not typical of the normal thymine hydroxylase-catalyzed oxidation of 5-formyluracil to its carboxylic acid (eq 1). The fact that the FAB-MS data give molecular mass peaks for mono-

(32) Kohmoto, S.; Kawatsuji, T. Y.; Ogata, K; Yamamoto, M.; Yamada, K. *J. Am. Chem. Soc.* **1991**, *113*, 5476.

(33) Takeuchi, K.; Kitagawa, T.; Veda, A.; Senzaki, Y.; Okamoto, K. *Tetrahedron* **1985**, *41*, 5455−5463.

(34) Lowry, T. H.; Richardson, K. S. *Mechanism and Theory in Organic Chemistry*, 3rd ed.; Harper and Row: New York, 1987; p 941.

(35) There presently is no evidence that an external carbene is generated during the interaction of thymine hydroxylase with ethynyluracil.

(36) Scott, A. P.; Nobes, R. H.; Schaefer, H. F.; Radom, L. *J. Am. Chem. Soc.* **1994**, *116*, 10159−10164.

(37) Toscano, J. P.; Platz, M. S.; Nikolaev, V.; Popic, V. *J. Am. Chem. Soc.* **1994**, *116*, 8146−8151.

(38) Thornburg, L. D.; Stubbe, J. *J. Am. Chem. Soc.* **1989**, *111*, 7632−7633.

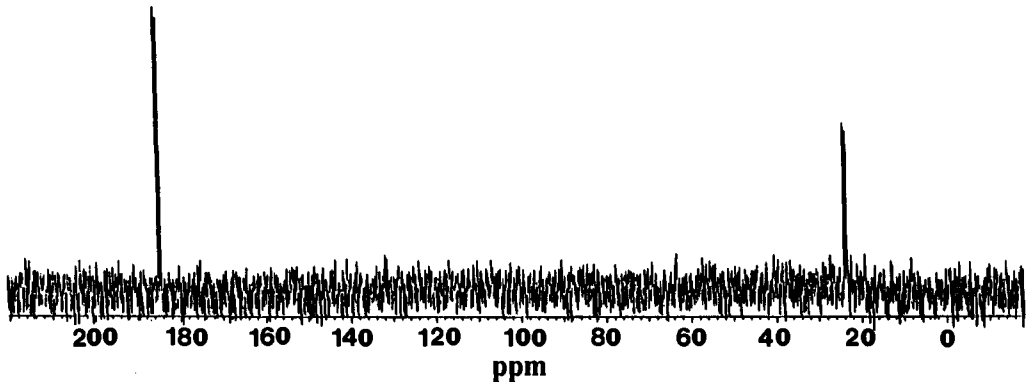


Figure 7. ^{13}C NMR (125 MHz) spectrum of the major modified peptide **1** in D_2O.

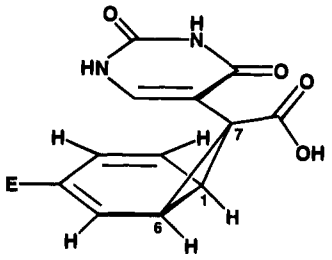


Figure 8. Proposed structure of the peptide **1**.

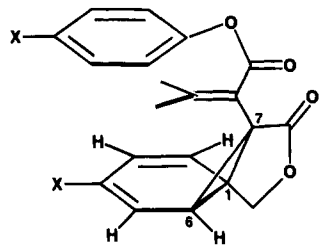


Figure 9. γ-Lactone ring fused norcaradiene, a model for peptide **1** (Figure 8).[32]

Table 2. Cross Peaks Observed in the HMBC Spectrum of **1**

assignments	$^1H-^{13}C$ cross peaks in ppm	size
H6 of uracil to C7 of NCD	7.08–23.5	large
H1 and H6 of NCD to carboxy carbon	2.99[a]–84.7	large
H5 of NCD to C7 of NCD	5.87–23.5	medium
H1 and H6 of NCD to C7 of NCD	2.99[a]–23.5	small
H2 of NCD to C7 of NCD	6.20–23.5	small

[a] Due to overlap, we cannot resolve the H1 and H6 protons of the NCD ring at 2.98 and 3.00 ppm.

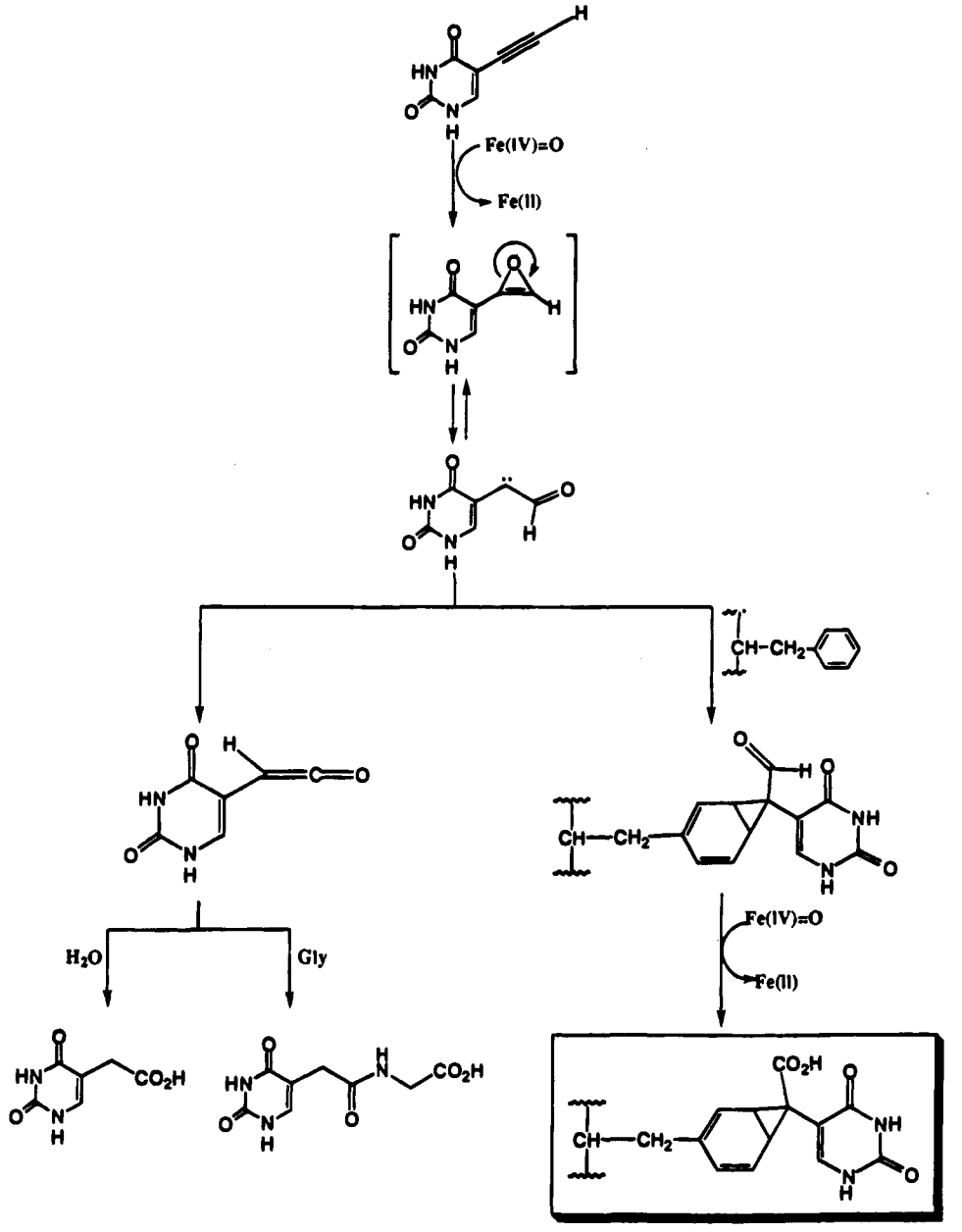
and disubstituted ^{18}O and perhaps an unsubstitued species[44] (Figure 2) suggests that exchange has occurred through an activated iron–oxygenation intermediate or through the putative aldehyde intermediate via a hydration/dehydration mechanism. Exchange has probably not occurred from the carboxylic acid under the workup conditions required to isolate the modified peptide. This interpretation of the results leads to the conclusion that the active site has, in fact, been modified by a reactive intermediate generated from 5-ethynyluracil and that the phenylalanine residue does not play an essential role in this process, as the second hydroxylation can occur despite the fact that its aromaticity has been destroyed.

An alternative interpretation of the observed $^{18}O_2$-labeling pattern is that the incorporation of the second ^{18}O is the result of a nonenzymatic hydroxylation, which could occur during the 45 min incubation reaction between ethynyluracil and thymine hydroxylase to ensure complete inactivation. Iron and ascorbate were present at moderately high concentrations during this period. A control experiment, albeit not the ideal one, was carried out to test this hypothesis. Uracil-5-acetaldehyde was incubated under conditions similar to those used in the inactivation reactions. No oxidation products were detected. Oxidation of our adduct subsequent to Sephadex chromatography of the inactivation mixture could not accommodate the ^{18}O-labeling results, as $^{18}O_2$ has been removed. We thus favor the interpretation that thymine hydroxylase mediates the second hydroxylation.

As outlined in the Results section, scale-up of our inactivation reaction revealed a second minor, as yet incompletely characterized adduct. Sequencing of this peptide by Edman degradation reveals that the same phenylalanine previously identified as the site of attack in the major peptide **1** is also modified in the minor peptide **2**. We have interpreted the FAB-MS data to indicate that no oxygens from O_2 are present in the modified peptide **2** and that **2** has acquired two additional protons in addition to the 5-ethynyluracil moiety. Alternative interpretations involving multiple hydroxylations and loss of formate can also account for the observed results. NMR characterization and an ^{18}O experiment using 5-ethynyluracil are required to distinguish between these possibilities. Identification of the NCD structure associated with the major modified peptide has established that thymine hydroxylase can convert an acetylene to a carbene. These studies contrast with inactivation of cytochrome P450s by terminal acetylenic compounds, such as phenylacetylene.[39–41] In both the heme and non-heme iron cases, ketene intermediates can account for the small molecules (acids) accompanying enzyme inactivation and the ^3H-labeling patterns of these molecules using terminal [^3H]acetylenes. However, the inactivation processes appear to occur by different mechanisms. In the case of the heme systems, inactivation results from alkylation of one of the nitrogens of the pyrrole ligands bound to the iron. No evidence for a carbene intermediate in this process could be obtained. In our case, inactivation

(39) Ortiz de Montellano, P. R.; Komives, E. A. *J. Biol. Chem.* **1985**, *260*, 3330.
(40) Komives, E. A.; Ortiz de Montellano, P. R. *J. Biol. Chem.* **1987**, *262*, 9793.
(41) Komives, E. A.; Tew, D.; Olmstead, M. M.; Ortiz de Montellano, P. R. *Inorg. Chem.* **1988**, *27*, 3112.
(42) Hoffmann, R. *Tetrahedron Lett.* **1990**, *31*, 4657–4660.
(43) Günther, H. *Tetrahedron Lett.* **1970**, *21*, 5173–5176.
(44) Papayannopoulos, I. A.; Biemann, K. Proceedings of the 42nd ASMS Conference on Mass Spectrometry and Allied Topics, Chicago, IL, May 29–June 4, 1994; p 476. In their abstract, these authors report that isotope clusters of protonated peptide molecules often also contain peaks below $(M + H)^+$. The $[(M + H)^+ - 1]$ and $[(M + H)^+ - 2]$ ions are thought to arise from losses of one and two hydrogens, respectively, from the protonated peptide. Thus, in our mass spectrum, the signal at 1520.3 could be due to peptide with no ^{18}O incorporation or to an $[(M + H)^+ - 2]$ ion, where the $(M + H)^+$, at 1522.3 corresponds to the monoisotopic mass of the mono-^{18}O-labeled peptide. In the mass spectrum of the standard peptide NSIAFFSNPSLR, $[(M + H)^+ - 1]$ and $[(M + H)^+ - 2]$ ions are observed.

Scheme 1. Proposed Mechanism for the Formation of the Turnover Products and the Major Adduct



results from labeling of the protein and not an iron ligand. This labeling can most reasonably be accommodated by a carbene type intermediate. Thus, we presently favor a hypothesis that both enzyme inactivation and small molecule formation result from partitioning of a common carbene intermediate (Scheme 1).[36,37] While the active site has been labeled, it is clear that ligands binding Fe^{2+} have not been perturbed, as a second hydroxylation can occur subsequent to phenylalanine modification. A careful examination of all available sequence information on α-KG dioxygenases fails to reveal any matches with the modified peptide. Useful information about groups involved in catalysis and Fe^{2+} coordination thus requires further experimentation. The observed stable norcaradiene adduct, however, has provided new insight into the chemical capabilities of these non-heme iron enzymes and has allowed us to revisit the physical organic world of electrocyclic reactions and their driving forces.[42,43]

Acknowledgment. We thank Dr. L. D. Thornburg for helpful discussions regarding the fermentation of *R. glutinis* and enzyme purification. We thank J. Battiste for assistance in recording NMR spectra. We acknowledge K. M. Downard at the MIT Mass Spectrometry Facility for mass spectra of peptides **1** and **2** and the support of this facility by NIH. RR00317 to K. Biemann. We also thank R. Pascal for a thoughtful review of this paper. This work was supported by NIH Grant GM 32191.

JA943926D