Activation of the Carboxy Terminus of a Peptide for Carboxy-Terminal Sequencing

Victoria L. Boyd,* MeriLisa Bozzini, Piotr J. Guga, Robert J. DeFranco, and Pau-Miau Yuan

Perkin-Elmer, Applied Biosystems Division, 700 Lincoln Centre Drive, Foster City, California 94404

G. Marc Loudon and Duy Nguyen

Department of Medicinal Chemistry, School of Pharmacy, Purdue University, West Lafayette, Indiana 47907

Received November 17, 1994®

Previously, we reported [*Anal. Biochem.* **1992**, *206*, 344–352] a new method of sequencing proteins from the carboxy terminus (C-terminus). The carboxyl group at the C-terminus is activated and derivatized into a thiohydantoin (TH). We reported that, by alkylating the TH formed at the C-terminus, the TH is converted into a readily displaced leaving group. Reaction with $\{NCS\}^-$ under acidic conditions cleaves the alkylated thiohydantoin (ATH) and derivatizes the freshly exposed C-terminus into a new proteinyl-TH. The efficiency of the initial activation of the carboxy group at the C-terminus is critical to the initial yield of the first ATH residue. In order to directly observe the intermediates that form during activation of the C-terminus, a model tripeptide, acetyl-alanine-alanine-alanine-OH (Ac-Ala-Ala-Ala-OH) was subjected to the reagents used to form the peptidyl-TH, Ac-Ala-Ala-Ala-TH. The reaction was monitored by nuclear magnetic resonance spectroscopy. An oxazolone was observed to form immediately at the C-terminus during the reaction with diphenyl chlorophosphate (DPCP), tetraphenyl pyrophosphate (TPPP), or tetramethylchlorouronium chloride (TMU-Cl). The oxazolone was observed to react with an excess of the carboxy group-activating reagents while under basic conditions. Diketopiperazine formation at the C-terminus was also observed. These side reactions prevent or retard the reaction of $\{NCS\}^-$ to form a peptidyl-TH and correlate with a reduced initial yield observed during automated C-terminal protein sequencing. The carboxylic acid-reactive reagents react with the side-chain carboxylic acid groups of aspartic and glutamic acid residues as well as the C-terminus. We found that the side-chain carboxylic acid groups in a protein could be selectively amidated in the presence of the proteinyl-oxazolone at the C-terminus.

Introduction

A chemical degradation method for carboxy-terminal (C-terminal) protein sequencing analogous to the Edman method is a highly desirable and challenging goal. One C-terminal sequencing approach that has been investigated by several researchers involves the derivatization of the C-terminal residue into a thiohydantoin (TH).[1-5] A mixture of acetic anhydride, acetic acid, and ammonium thiocyanate is used to form the C-terminal peptidyl-TH. The derivatized C-terminal amino acid-TH is then cleaved hydrolytically from the parent peptide. The newly exposed carboxyl group of the truncated protein or peptide is reactivated in the subsequent cycle.

Recently, efforts to form a peptidyl- or proteinyl-TH via activation of the C-terminal carboxylic acid have included reagents other than acetic anhydride.[6-10] We previously reported the use of tetramethylchlorouronium chloride (TMU-Cl) and recently have investigated the use of diphenyl chlorophosphate (DPCP) as activating reagents to initiate our automated C-terminal sequencing method.[11] In our sequencing method, which is depicted in Scheme 1, only the initial step requires activation of the carboxyl group. After the initial proteinyl-TH formation, the TH ring is alkylated. The resulting alkylated-TH (ATH) is a readily displaced leaving group. Reaction with $\{NCS\}^-$ under acidic conditions cleaves the ATH and simultaneously generates a new proteinyl-TH from the penultimate C-terminal residue. However, during our sequencing studies, we noted that the efficiency of the initial peptidyl-TH formation using DPCP or TMU-Cl varied. Nuclear magnetic resonance (NMR) spectroscopy was therefore used to study the reaction of DPCP and TMU-Cl with a model tripeptide. Under basic conditions, an oxazolone is the *only* intermediate that we observed by ^{1}H and ^{13}C NMR spectroscopy during the activation of the carboxy group at the C-terminus. Importantly, the oxazolone intermediate has been found to react with the activating reagents in a side reaction that interferes with the initial peptidyl-TH formation. Finally, rearrangement of the oxazolone to a diketopiperazine has been observed under the reaction conditions. This paper reports the details of these observations.

® Abstract published in *Advance ACS Abstracts,* April 1, 1995.

(1) Stark, G. R. *Biochemistry* **1968**, *7*, 1796.

(2) Rangarajan, M.; Darbre, A. *Biochem. J.* **1976**, *157*, 307.

(3) Meuth, J. L.; Harris, D. E.; Dwulet, F. E.; Crowl-Powers, M. L.; Gurd, F. R. N. *Biochemistry* **1982**, *21*, 3750.

(4) Bailey, J. M.; Shively, J. E. *Biochemistry* **1990**, *29*, 3145.

(5) Inglis, A. S.; Moritz, R. L.; Begg, G. S.; Reid G. E.; Simpson, R. J.; Graffunder, H.; Matschull, L.; Wittman-Liebold, B. In *Methods in Protein Sequence Analysis*; Jornvall, H., Hoog, J.-O., Gustavsason, A.-M., Eds.; Birkhauser Verlag: Basel, Switzerland, 1991; pp 23–34.

(6) Boyd, V. L.; Hawke, D. H.; Geiser, T. G. *Tetrahedron Lett.* **1990**, *31*, 3849.

(7) Hawke, D. H.; Boyd, V. L. In *Techniques in Protein Chemistry II*; Villafranca, J. J., Ed.; Academic Press: San Diego, CA, 1991; pp 107–114.

(8) Hawke, D. H.; Boyd, V. L. In *Methods in Protein Sequence Analysis*; Jornvall, H., Hoog, J.-O., Gustavsoon, A.-M., Eds.; Birkhauser Verlag: Basel, Switzerland, 1991; pp 35–45.

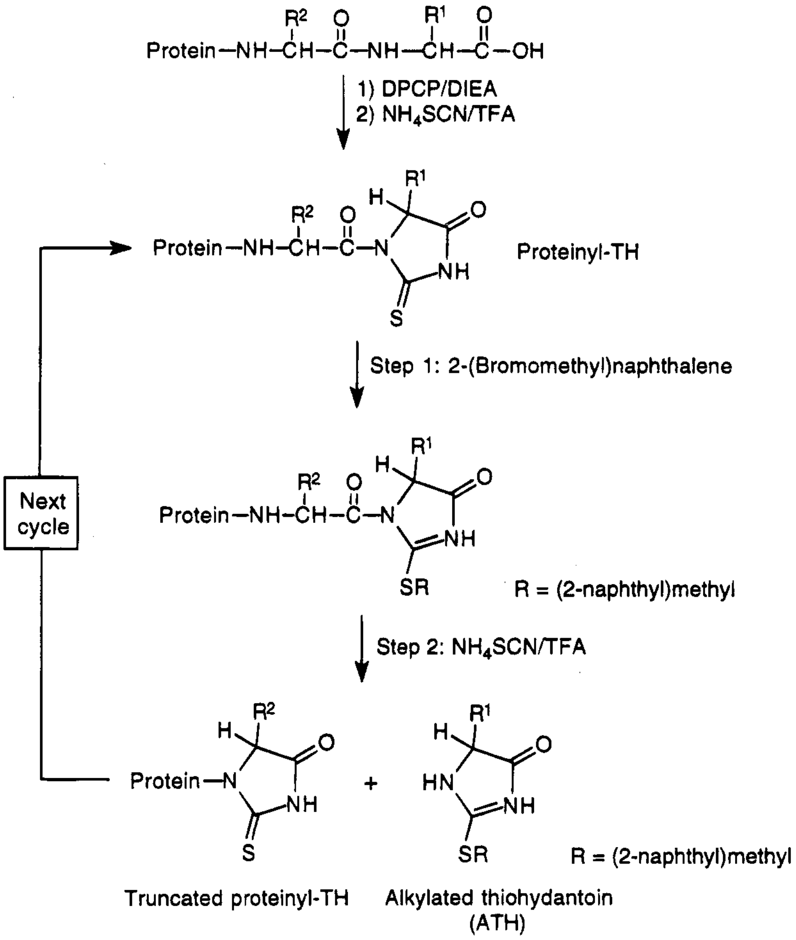
(9) Bailey, J. M.; Nikfarjam, F.; Shenoy, N. R.; Shively, J. E. *Protein Sci.* **1992**, *1*, 1622.

(10) Shenoy, N. R.; Shively, J. E.; Bailey, J. M. *J. Protein Chem.* **1993**, *12*, 193.

(11) Boyd, V. L.; Bozzini, M.; Zon, G.; Noble, R. L.; Mattaliano, R. J. *Anal. Biochem.* **1992**, *206*, 344.

0022-3263/95/1960-2581$09.00/0 © 1995 American Chemical Society

Scheme 1



The results of the tripeptide model study help to interpret the probable byproducts and intermediates, which cannot be observed directly, that occur during automated sequencing of proteins. Having identified the side reactions, it is possible to select reagents and conditions to minimize their occurrence. As a result of the NMR investigations combined with observations made during automated C-terminal sequencing, we discovered that the activated side-chain carboxylic acid groups of aspartic and glutamic acid residues could be amidated selectively in the presence of the oxazolone formed at the C-terminus. The ability to modify the carboxylic acid side chain of aspartic and glutamic acids, and not compromise the formation of a C-terminal-TH, is consistent with the formation of an oxazolone at the C-terminus of a protein during automated sequencing.

Experimental Section

The following reagents and solvents were purchased from Aldrich (Milwaukee, WI): ammonium thiocyanate (NH_4SCN), piperidine thiocyanate, diphenyl chlorophosphate (DPCP), diphenyl phosphate (DPP), deuterated acetonitrile (CD_3CN), triethylamine (Et_3N), 2-(bromomethyl)naphthalene, pyridine, and ethoxycarbonyl isothiocyanate (EtOCONCS). Methylene chloride (CH_2Cl_2) was supplied by Burdick and Jackson (Muskegon, MI). Precoated silica gel preparative plates were obtained from Analtech (Newark, DE), and silica gel (60 A, 70–230 mesh) was purchased from Scientific Products (McGaw Park, IL). Methanol (MeOH), trifluoroacetic acid (TFA; extremely corrosive), acetonitrile (CH_3CN), diisopropylethylamine (DIEA), and ProSpin sample preparation cartridges were supplied by Perkin-Elmer, Applied Biosystems Division (Foster City, CA). Tetraphenyl pyrophosphate (TPPP)[12] and 1,1,3,3-tetramethylchlorouronium chloride (TMU-Cl)[13] were synthesized according to literature reports. *N*-α-acetyl-alanine-alanine-alanine (Ac-Ala-Ala-Ala-OH, **1**), was purchased from Bachem Bioscience Inc. (Philadelphia, PA).

Synthesis of Ac-Ala-Ala-Ala-TH (3). The tripeptide **1** (0.2 g, 0.73 mmol) was suspended in 25 mL of CH_3CN. Two equivalents each of pyridine (0.12 mL, 1.4 mmol) and EtOCONCS (0.2 mL, 1.4 mmol) were added, and the reaction was stirred overnight.[11] The clear solution was concentrated under reduced pressure and the oil was purified by column chromatography by eluting with CH_2Cl_2:MeOH (9.2:0.8, $R_f = 0.5$). The purified product **3** was characterized by ^1H NMR spectroscopy: (300 MHz, CD_3CN) major diastereomer, δ 1.28 (d, J = 7.0 Hz, 3H), 1.37 (d, J = 7.0 Hz, 3H), 1.43 (d, J = 7.0 Hz, 3H), 2.0 (s, 3H), 4.32 (quintet, J = 7.1 Hz), 4.61 (quartet, J = 7.1 Hz, 1H), 6.02 (quintet, J = 7.0 Hz, 1H). Minor diastereomer: δ 1.31 (d, J = 7.0 Hz, 3H), 1.41 (d, J = 7.0 Hz, 3H), 1.42 (d, J = 7.0 Hz, 3H), 1.96 (s, 3H), 4.36 (quintet, J = 7.1 Hz, 1H), 4.68 (quartet, J = 7.1 Hz, 1H), 6.31 (quintet, J = 7.0 Hz, 1H). The quintets became quartets when the N–**H**s were exchanged by adding D_2O to the CD_3CN solution.

Synthesis of Oxazolone 2 with DPCP. Oxazolone **2** formed immediately when 1 equiv of **1** (20 mg, 0.073 mmol) was mixed in CD_3CN with 2 molar equiv of DIEA (0.025 mL, 0.14 mmol) and 1 molar equiv of DPCP (0.015 mL, 0.073 mmol). Both ^1H and ^{31}P NMR spectra were obtained. ^1H NMR (300 MHz, CD_3CN) spectrum: δ 1.24 (d, J = 7.3 Hz, 3H), 1.32 (apparent dd, J = 7.5 and 1.8 Hz, 3H), 1.36 (apparent dd, J = 7.0 and 1.0 Hz, 3H), 1.85 (s, 3H), 4.22 (apparent quartet of triplets, J = 7.6 and 1.8 Hz, 1H), 4.3 (quintet, J = 7.3 Hz,

---

[12] Corby, N. S.; Kenner, G. W.; Todd, A. R. *J. Chem. Soc.* **1952**, 1234.

[13] Knorr, R.; Trzeciak, A.; Bannworth, W.; Gillessen, D. *Tetrahedron Lett.* **1989**, *30*, 1927.

1H), 4.67 (doublet of quintets, $J = 7.2$ and 1.8 Hz, 1 H), 7.4 and 7.6 (br doublets, 2H). Resonance signals for CD_3CN at δ 1.93 (reference), DIEA at δ 1.1 (t), 2.7 (quartet), and 3.22 (quartet) and DPP at δ 6.95–7.3 (m) were also present. The chemical shifts in the ^{31}P NMR (121 MHz, CD_3CN) were corrected relative to phosphoric acid in D_2O as an external reference. A single resonance signal at δ 30.2 was observed. Addition of authentic DPP to this sample enhanced the resonance signal at δ 30.2. The chemical shift of DPP varied with the pH of the solution. The ^1H-decoupled ^{13}C NMR spectrum (75 MHz, CD_3CN) for **2** was acquired and had exactly the expected number of resonance signals: δ 16.8, 17.63 and 17.69, 18.7, 23.1, 44.56 and 44.6, 50.1, 61.3, 166, 171, 174, and 180. Signals for DIEA (δ 14.2, 18.9, 41.8, and 52.9), DPP (δ 121, 126, 131, and 151), and CD_3CN [δ 1.3 (quintet, reference) and 118] were also seen.

Synthesis of Oxazolone 2 with TMU-Cl. Oxazolone **2** formed immediately when 1 equiv of **1** (20 mg, 0.073 mmol) was mixed in CD_3CN with 2 molar equiv of DIEA (0.025 mL, 0.14 mmol) and 1 equiv of TMU-Cl (12.4 mg, 0.073 mmol). In addition to the above resonance signals in the ^1H NMR spectrum for **2** (and DIEA), a signal at δ 2.7 corresponding to the chemical shift for tetramethylurea (TMU) was observed.

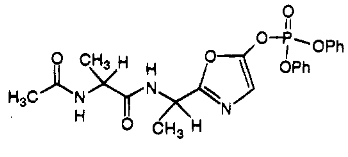
Formation of Adduct 5 from Oxazolone 2 and Excess DPCP. To a CD_3CN solution of oxazolone **2** (0.11 mmol) synthesized with 1 equiv of DPCP were added additional DIEA (0.77 mmol, total) and DPCP (0.44 mmol, total). The resulting ratio of DPCP:**2** was approximately 3:1 (not 4:1) because DPCP also reacts with DPP to form TPPP (see below). Within 5 min, the ^1H NMR spectrum was consistent with the total absence of **2** due to the formation of adduct **5**: (300 MHz, CD_3CN) δ 1.19 (d, $J = 6.9$ Hz), 1.22 (d, $J = 7$ Hz), 1.39 (d, $J = 7$ Hz), 1.84, 1.87, 1.88 (s), 4.3 (quintet, $J = 7.4$ Hz), 4.9 (quintet, $J = 7.0$ Hz), 7.0 (br t, $J = 7.0$ Hz), 7.66 (br d, $J = 7.4$ Hz), 8.16 (br d, $J = 7.4$ Hz). The ^{31}P NMR spectrum (CD_3CN, 121 MHz) had resonance signals at δ 41.01 (DPCP), 33.45 (DPP), 28.95 (oxazolone–DPP adduct), and 21.44 (TPPP).

Formation of Adduct 6 from Oxazolone 2 and Excess TMU-Cl. To the above solution of **2** synthesized with 1 equiv of TMU-Cl were added DIEA (0.025 mL, 0.14 mmol) and TMU-Cl (12.4 mg, 0.073 mmol). The product **6** was formed immediately as evidenced by changes in the ^1H NMR spectrum. Resonance signals for the reaction mixture were seen at (CD_3CN, 300 MHz) δ 1.22–1.32 (several ds), 1.43 (d, $J = 6.9$ Hz), 1.86 and 1.98 (s), 2.67 (s), 3.1 (s), 4.3 (quintet, $J = 7.9$ Hz), 4.85 (quintet, $J = 7.1$ Hz), 8.5 (br d, $J = 7.6$ Hz), 9.0 (br d, $J = 7.6$ Hz). Addition of a third equivalent of TMU-Cl and a total of 6 equiv of DIEA resulted in a resonance signal in the ^1H NMR spectrum at δ 3.26 corresponding to unreacted TMU-Cl. Resonance signals for DIEA were also present as described above.

Formation of Diketopiperazine 4 from Oxazolone 2. A solution of **2** synthesized as described above with 1 equiv (and no more) of either DPCP or TMU-Cl was treated with 1 molar equiv of Et_3N and heated at 60 °C for 5 min. A single new product, **4**, formed from **2** (Scheme 4) and was purified on a silica gel preparative plate by eluting with CH_2Cl_2:MeOH (9.5:0.5). The ^1H NMR spectra of **4** was identical regardless of whether **2** had been generated with DPCP or TMU-Cl. ^1H NMR: (CD_3CN, 300 MHz) δ 1.1–1.2 (four apparent doublets, $J = 7$ Hz), 1.82 and 1.83 (s), 2.16 (apparent dd, $J = 13.5$ and 2.4 Hz), 4.1 (septet, $J = 7.1$ Hz), 4.5 (m), 5.57 (apparent doubled quintet, $J = 2.4$ and 0.7 Hz). ^{13}C NMR: (CD_3CN, 75 MHz) δ 14.12 and 14.27, 15.28, 16.85, 18.4 and 18.45, 49.92 and 49.84, 101.02 and 101.43, 162.96, 163.94, 170.7, 173.12 and 173.29. Signals for residual DPP and CD_3CN were also present at the chemical shifts previously assigned. The chemical ionization mass spectrum was also consistent with the assigned structure **4**: (EI 70 eV) m/z 256 (molecular ion + 1), 114 (base peak, 256 − 142), 143 (diketopiperazine heterocycle MW + 1).

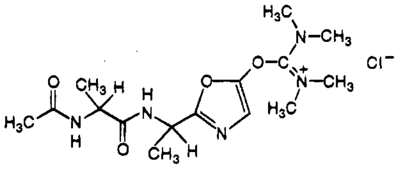
Formation of Peptidyl-TH 3 from Oxazolone 2. A solution of **2**, prepared from 1 equiv (and no more) of either DPCP or TMU-Cl, reacted rapidly with a premixed acetonitrile solution (0.5 mL) containing NH_4SCN (approximately 50 mg) and TFA (approximately 50 μL) to form **3**. The precipitate that formed was identified as NH_4Cl by comparison of its infrared spectrum to that of a library spectrum. The ^1H NMR spectrum of the solution phase was identical to the spectrum of a diastereomeric mixture of **3** previously described but with the diastereoisomers present in equal amounts.

Formation of Peptidyl-TH 3 from Adduct 5. A premixed solution of NH_4SCN (50 mg) and TFA (50 μL) in CD_3CN was added to a solution of the oxazolone–DPP adduct **5** in CD_3CN. NH_4Cl precipitated immediately. After 1 h at 60 °C the ^1H NMR spectrum contained signals identical to those previously described for a diastereomeric mixture of **3** independently synthesized as described above. The ^{31}P NMR spectrum had signals at δ 33.45 (DPP) and 21.44 (TPPP) and no longer had a signal at δ 28.95 for **5**.



5

Formation of Peptidyl-TH 3 from Adduct 6. A premixed solution of NH_4SCN and TFA was added to a solution of **6** in CD_3CN. NH_4Cl precipitated immediately. After 1 h at rt, the ^1H NMR spectrum had signals consistent with the presence of a diastereomeric mixture of **3** as the sole product. The integrated area of the resonance signal at δ 3.1, corresponding to **6**, was greatly reduced. (Resonance signals for DIEA were superimposed on the signal at δ 3.1.)



6

Reaction of Adduct 5 with TFA. A CD_3CN solution of the adduct **5** was acidified with TFA, and a drop of D_2O was added. After 30 min at 60 °C, the ^1H NMR spectrum had signals consistent with the presence of **1** and **5** in an approximately 1:1 ratio. The ^{31}P NMR spectrum showed a reduction in the intensity of the signal at δ 28.95 for **5**.

Reaction of Adduct 6 with TFA. A CD_3CN solution of **6** was acidified with TFA, and a drop of D_2O was added. The integrated intensity for the signal at δ 3.1 assigned to **6** was no longer detected after 3 h at rt. The hydrolysis of **6** resulted in formation of **1**, although signal intensity in the ^1H NMR spectrum for **1** was reduced by deuterium exchange.

Synthesis of Oxazolone 2 with TPPP. TPPP was first formed by premixing DPP (0.22 mmol, 54 mg), DPCP (0.22 mmol, 45 μL), and DIEA (0.22 mmol, 38 μL) in approximately 0.5 mL of CD_3CN. The chemical shift in the ^{31}P NMR, δ 21.44, confirmed the presence of TPPP and the absence of DPCP and DPP. The TPPP solution was added to a solution of **1** (0.22 mmol, 50 mg) and DIEA (0.77 mmol, 95 μL). Oxazolone **2** formed immediately at ambient temperature, and the ^1H NMR spectrum was identical to that of **2** obtained above. The ^{31}P NMR spectrum had a resonance signal at δ 33.5 for DPP.

Reaction of Oxazolone 2 with TPPP. A premixed solution of DPCP (0.22 mmol, 45 μL) DPP (0.22 mmol, 54 mg), and DIEA (0.44 mmol, 76 μL) in CD_3CN (0.5 mL) (to form TPPP) was added to a CD_3CN solution of **2** (0.11 mmol). DPP (0.11 mmol) and DIEA (0.27 mmol) were already present in the solution of **2** generated above. The reaction was monitored by ^1H and ^{31}P NMR spectroscopy. After 30 min at ambient temperature, the integrated intensity for the signals in the ^1H NMR spectrum indicated a 50:50 mix of **2** and **5**.

Automated Sequencing from the C-Terminus. Scheme 1 outlines the automated C-terminal sequencing method previously described.[11] The protein of interest is immobilized on a membrane support and placed in the reaction cartridge of an Applied Biosystems Model 477A protein sequencer. The activation or "begin" cycle initiates the sequencing method by derivatizing the carboxy terminus into an amino acid-TH. The reaction cycle to perform stepwise sequencing is comprised of two steps. First, the newly formed TH is regioselectively alkylated at the sulfur atom. Second, the resultant ATH derivative is cleaved with $\{NCS\}^-$ under acidic conditions. The following reagent changes were made relative to the previous publication:[11] NH_4SCN was used in place trimethylsilyl isothiocyanate, piperidine thiocyanate was used to derivatize the side chains of aspartic and glutamic acids, and TPPP or DPCP was used in place of TMU-Cl.

Results and Discussion

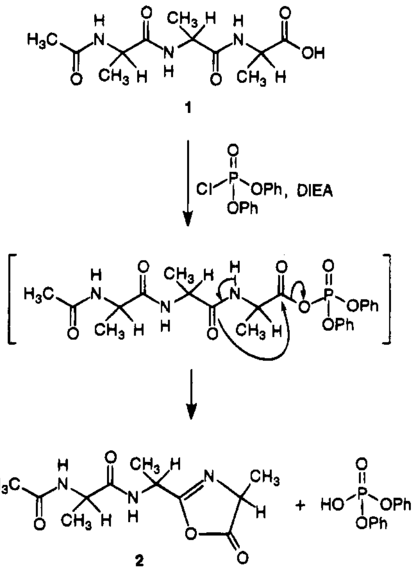
The method of protein sequencing from the C-terminus outlined in Scheme 1 requires an initial activation of the carboxylic acid group, followed by reaction with $\{NCS\}^-$ to form a proteinyl-TH. Because of the difficulty in directly observing the intermediates and byproducts formed during subnanomolar protein sequencing, the corresponding reactions of a model tripeptide were investigated by NMR spectroscopy. The following observations made on the tripeptide model help to interpret the probable intermediates and side reactions that must likewise occur during this automated sequencing method.

A single peptide, Ac-Ala-Ala-Ala-OH, **1**, was used for all the NMR studies. This model peptide was chosen because of the absence of reactive side-chain residues and for the ease of assignments in the ^1H NMR spectra. It should be noted, however, that the rate of formation of the intermediates and products observed in the NMR investigations is likely to vary depending on the peptide or protein sequence.

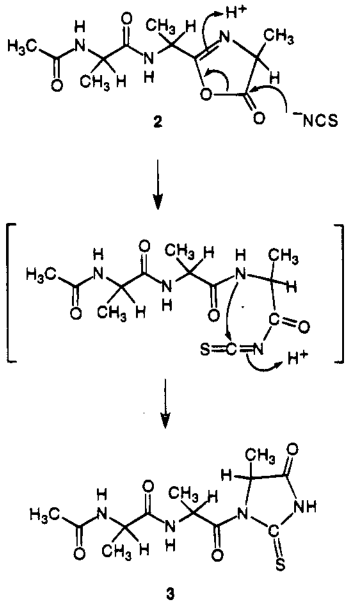
Activation of the Carboxy Group of a Peptide with Diphenyl Chlorophosphate. The reaction of the model peptide **1** with 1 equiv of DPCP and 2 equiv of DIEA was observed by ^1H, ^{31}P, and ^{13}C NMR spectroscopy. Two molar equivalents of base were needed to neutralize the HCl that was generated. Basic conditions also increased the solubility of **1**. Immediately after mixing the reactants in CD_3CN, the ^{31}P NMR spectrum was obtained. Although a resonance signal corresponding to the acyl phosphate–mixed anhydride[14] was anticipated, only a ^{31}P NMR signal for DPP was observed. Since the chemical shift of DPP is sensitive to the pH of the solution, the ^{31}P NMR signal observed was verified by spiking the sample with authentic DPP.

The rapid elimination of the phosphate moiety (DPP) from the presumed acyl phosphate intermediate was consistent with the formation of oxazolone **2** (Scheme 2). In studies involving the activation of a C-terminal carboxy group under acidic conditions with a mixture of acetic acid and acetic anhydride, a peptidyl-oxazolone such as **2** was not proposed as the sole intermediate.[1] Under the basic reaction conditions of the present experiment, **2** was the only activated species observed. A symmetric anhydride, formed from the self-condensation of 2 equiv of **1**, was ruled out because of the observed 1:1 stoichiometry of the reaction of the carboxylic acid with DPCP. The ^1H NMR resonance signals for **2** remained unchanged in solution for at least 1 h at ambient temperature while a ^{13}C NMR spectrum was obtained. (See Experimental Section for chemical shifts.) Oxazolone **2** reacted rapidly and completely with a solution of NH_4SCN acidified with TFA to form the peptidyl-TH, **3**, as an equal mixture of diastereomers. This reaction is shown in Scheme 3. The ^1H NMR signals for **3** were identical to those of **3** synthesized using EtOCONCS, which had been purified by silica gel chromatography (see Experimental Section).

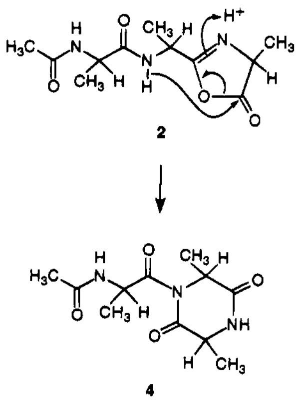
Scheme 2



Scheme 3



Scheme 4



---

[14] Michelson, A. M. *Biochim. Biophys. Acta* **1964**, *91*, 1.

During automated sequencing, reagents are delivered into the reaction cartridge which houses the support-bound protein. The sequencing reagents saturate the membrane and are therefore present in large molar excesses relative to the amount of the protein.

In order to mimic the conditions of proteinyl-oxazolone formation during automated sequencing, oxazolone **2** was mixed with an additional 4 equiv of DPCP and an additional 7 equiv of DIEA. The ensuing reaction was monitored by NMR spectroscopy. The actual ratio of DPCP:**2** was 3:1 since DPCP also reacts with DPP to form TPPP. DPP was already present since it was eliminated during formation of **2** (see above). The solution immediately turned yellow, and the ^{1}H NMR spectrum showed the nearly complete absence of resonance signals for **2**, in addition to new signals. A new resonance signal in the ^{31}P NMR spectrum accompanied the formation of the new product and was assigned to the product of a reaction between DPCP and **2**. Structure **5** depicts the proposed structure for an oxazolone–DPP adduct. Ionization of **2** under basic conditions provides the nucleophilic enolate oxygen which reacts rapidly with DPCP to form **5**.

Adduct **5** formed from oxazolone **2** and excess DPCP was mixed with a solution of NH_4SCN and TFA. Peptidyl-TH, **3**, formed slowly from **5** over 3 h at rt. The slow formation of **3** from **5** contrasts with the rapid formation of **3** from **2** (see above). Thus, *any oxazolone (**2**) that reacts with excess DPCP to form an oxazolone–DPP adduct (**5**) greatly impairs subsequent Ac-Ala-Ala-Ala-TH (**3**) formation.*

A proteinyl-oxazolone, in addition to being subjected to the large excesses of the activating reagent during its formation using automated instrumentation, is also briefly subjected to basic conditions in the absence of the activating reagent and to prolonged acidic conditions (TFA) during the reaction with $\{NCS\}^-$. Therefore, in order to simulate conditions that occur during the stepwise delivery of reagents during automated sequencing, the reaction of oxazolone **2** under acidic and basic conditions was monitored by NMR spectroscopy. A solution of **2**, acidified with TFA in the absence of $\{NCS\}^-$, was rapidly hydrolyzed (presumably by traces of moisture) to the starting peptide **1**. Under basic conditions, **2** underwent an intramolecular rearrangement to a product that was stable to silica gel chromatography. The ^{1}H and ^{13}C NMR spectra, and both the molecular ion and fragmentation pattern in the mass spectrum of the purified material, are consistent with the formation of diketopiperazine **4** (Scheme 4). The rate at which **2** rearranged to **4** was faster in the presence of Et_3N than with DIEA. Diketopiperazine **4** did not react with $\{NCS\}^-$ to form peptidyl-TH, **3**, even after several hours at rt. Therefore, *any diketopiperazine formation at the C-terminus blocks any further attempt to sequence the protein.*

Activation of a Carboxy Group of a Peptide with Tetramethylchlorouronium Chloride. The use of TMU-Cl for the derivatization of a peptide or protein to its corresponding peptidyl-TH as an initial step in the C-terminal sequencing approach presented in Scheme 1 was previously described.[11] The same series of reactions of peptide **1** with TMU-Cl in place of DPCP were monitored by ^{1}H NMR spectroscopy. As with DPCP, oxazolone **2** formed immediately, accompanied by the elimination of tetramethylurea (TMU). The chemical shifts in the ^{1}H NMR spectrum for **2** were identical to those seen previously (using DPCP), and the eliminated TMU corresponded to that of authentic material.

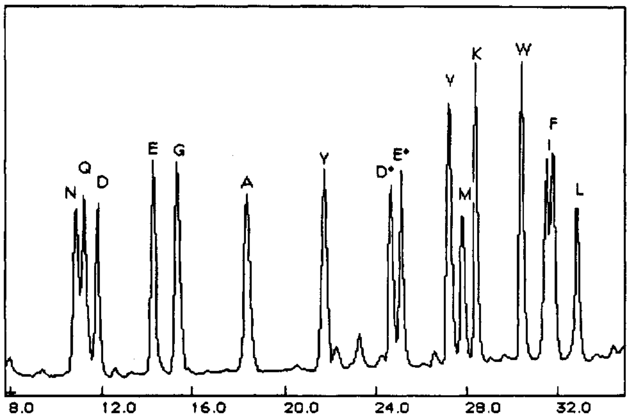


Figure 1. HPLC separation of the ATH reference standards for automated C-terminal sequencing. The ATHs of the amino acids were prepared as previously described[11] using 2-(bromomethyl)naphthalene. The ATH of Gln elutes at 11.5 min and the piperidine amide of Glu (Glu*) is seen at 25 min. The retention times match those of the Gln and Glu* seen in cycle 3 of Figures 2 and 3. The separation was achieved with an Applied Biosystems PTC column using the following solvents and gradient. Solvent A: 35 mM sodium acetate in solvent A3 (3.5% THF). Solvent B: acetonitrile. Column: PTC, C-18. Flow rate: 300 μL/min. HPLC column temperature: 38–42 °C. Gradient (time, % B): (0, 20); (20, 35); (30, 49); (35, 70); (38, 90); (42, 90).

Oxazolone **2** reacted with the excess TMU-Cl under basic conditions much as it does with DPCP to form the adduct **6**. The reaction of **2** with excess TMU-Cl and DIEA was complete within 10 min at rt. Adduct **6** reacted slowly with $\{NCS\}^-$ under acidic conditions, forming peptidyl-TH, **3**, after 1 h at rt. A CD_3CN solution of **6** completely hydrolyzed after 15 min when treated with TFA and a drop of D_2O.

Activation of the Carboxy Group of a Peptide with Tetraphenyl Pyrophosphate. TPPP is generated *in situ* by the reaction of DPCP with DPP. The reaction of peptide **1** with an equivalent of TPPP and DIEA resulted in the immediate formation of oxazolone **2**. The formation of **2** from **1** with TPPP parallels that of the formation of **2** from **1** and DPCP, as shown in Scheme 2. Oxazolone **2** was then mixed with a 2-fold excess of the activating reagent TPPP under basic conditions. Both TPPP and DPCP can form the same adduct **5** from **2**. However, after 30 min at rt only about 50% of **2** had reacted with TPPP to form **5**. Under similiar reaction conditions, **2** reacts with DPCP rapidly to form **5**. *Since oxazolone **2** reacts with $\{NCS\}^-$ more rapidly than oxazolone–adduct **5**, a reduction in the formation of **5** corresponds to an increase in proteinyl-TH formation.*

Modification of the Side-Chain Carboxy Groups of Aspartic and Glutamic Acid Residues during Automated C-Terminal Protein Sequencing. The HPLC separation of the ATH reference standards used for automated C-terminal sequencing is shown in Figure 1. During our automated C-terminal sequencing, aspartic and glutamic acid residues are observed to form asparagine and glutamine if NH_4SCN is used as a source of $\{NCS\}^-$. An example of the amidation of a glutamic acid residue in a protein during C-terminal sequencing is provided in Figure 2. In this α-lactalbumin sample, the glutamic acid residue in cycle 3 is detected as the

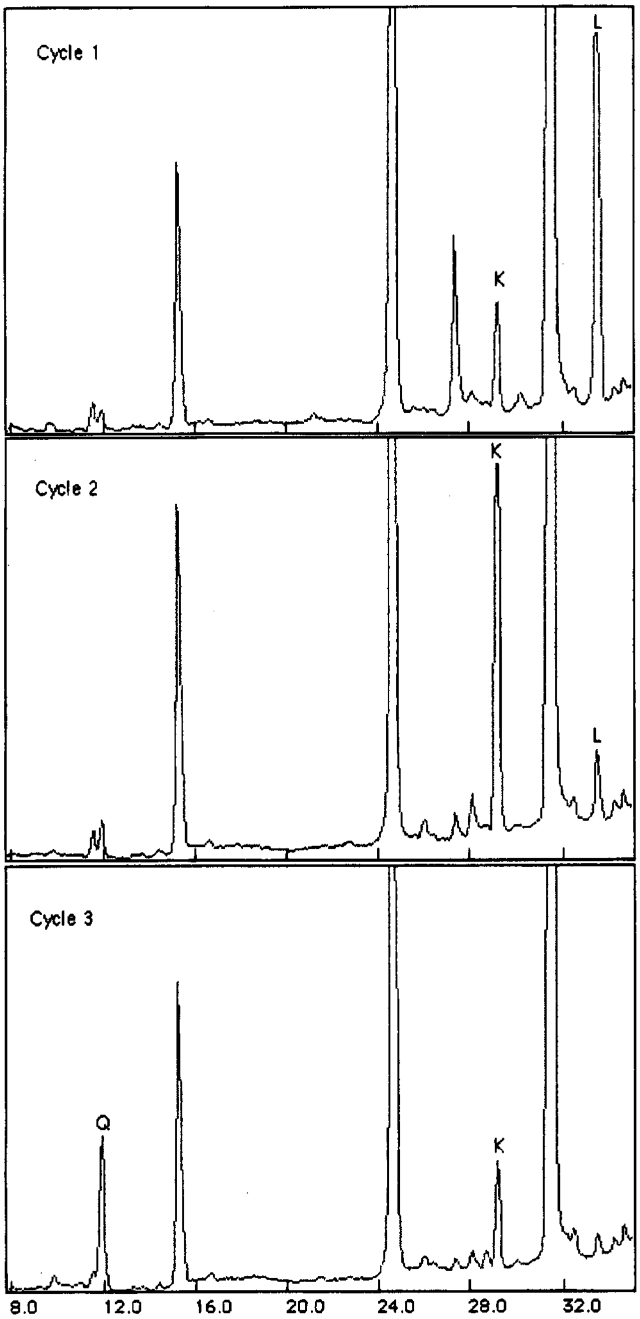


Figure 2. Automated C-terminal sequencing data on 2 nmol of α-lactalbumin applied to PVDF. Deliveries of DIEA, TPPP, and NH_4SCN derivitized the glutamic acid in cycle 3 into glutamine. The ϵ amine group of the lysine ATH residue in cycle 2 was derivatized prior to sequencing with phenyl isocyanate to form the phenylurea.

ATH of glutamine. The initial steps of our automated sequencing protocol require the delivery of base (such as DIEA), the activating reagent (such as TPPP), NH_4SCN, and finally TFA to generate HSCN from the NH_4SCN. Because conditions are initially basic when NH_4SCN is first added to the TPPP-activated protein, NH_4SCN can serve as a source of NH_3. The NH_3 generated from the NH_4SCN has amidated the activated side chain of the glutamic acid residue but not the oxazolone at the C-terminus. The HPLC retention time of the ATH of glutamine formed from glutamic acid during protein sequencing (Figure 2) matches that of the independently synthesized glutamine ATH standard (Figure 1).

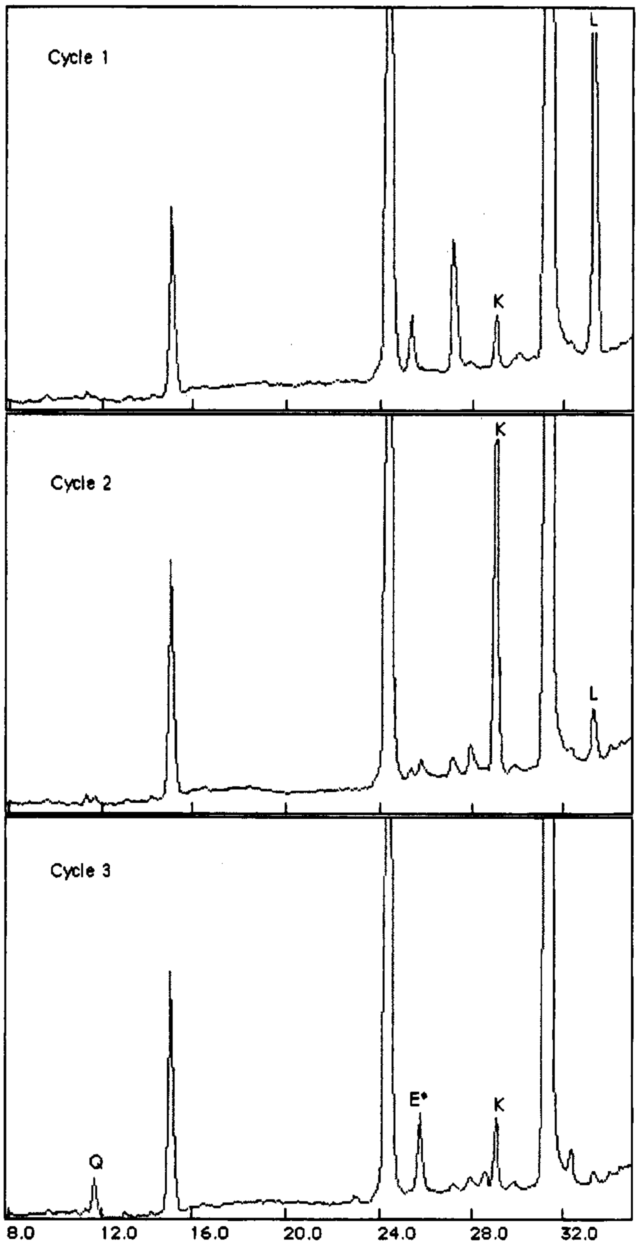


Figure 3. Automated C-terminal sequencing data on 2 nmol of α-lactalbumin including the delivery of piperidine thiocyanate prior to deliveries of NH_4SCN. The sequencing protocol was modified relative to Figure 2 by including a 300 s pause after deliveries of DIEA, TPPP, and piperidine thiocyanate to allow the amidation of the Glu in cycle 3 into the piperidine amide. NH_4SCN was used subsequently for derivitizing the oxazolone into a thiohydantoin, which accounts for the conversion of some of the Glu into Gln in cycle 3.

The reaction of TPPP or DPCP with the side-chain carboxy groups of aspartic acid and glutamic acid residues forms the corresponding acyl phosphates. Studies by NMR spectroscopy could not mimic the reaction of the side-chain carboxy groups of the membrane-bound protein because of the unavoidable formation of the symmetric anhydride from self-condensation that occurs in solution. Anhydride formation between aspartic and glutamic acid residues in a protein can only occur if the residues are in proximity. Under basic conditions, an acyl phosphate formed at a carboxylic acid side chain is uncharged, but an oxazolone formed at the C-terminus is ionized. Addition of a nucleophile such as NH_3 generated from NH_4SCN reacts with the acyl phosphate under basic conditions but does not react with the ionized oxazolone. Therefore the C-terminus is not amidated. Delivery of TFA protonates the oxazolone and generates HSCN which reacts with the oxazolone to form the proteinyl-TH. Substituting piperidine thiocyanate in place of NH_4SCN converts aspartic acid and glutamic acid residues into the corresponding piperidine amides. In comparison with Figure 2, Figure 3 shows an identically prepared sample of α-lacalbumin treated with TPPP, DIEA, and piperidine thiocyanate. In this case, the glutamic acid residue in cycle 3 is detected as the piperidine amide. The HPLC retention time of the piperidine amide corresponded to that of the reference standard (Figure 1). Experiments by NMR spectroscopy confirmed that **2**, while still basic from the excess DIEA present, did not react with NH_4SCN to form the amide Ac-Ala-Ala-Ala-NH_2. *Formation of the oxazolone at the C-terminus therefore allows selective amidation of the side-chain carboxy groups of aspartic acid and glutamic acid residues without also amidating the C-terminus.*

Conclusion

The important conclusions of this work that relate to C-terminal sequencing are the following. 1. Activation of the C-terminal carboxylic acid results in the formation of an oxazolone. If the oxazolone forms an adduct with the carboxy group-activating reagent, proteinyl-TH formation is impaired. 2. Under basic conditions, the oxazolone is converted into diketopiperazine; this reaction completely prevents proteinyl-TH formation. 3. The unique ability to form selectively an oxazolone at the C-terminus, and not at the aspartic acid and glutamic acid side chains, allows selective amidation of the side-chain carboxylic acid groups. The improved C-terminal sequencing performance achieved because of the insights gained from this study will be described in future publications.

JO9419450