Carboxypeptidase Y-Catalyzed Transpeptidation of Esterified Oligo- and Polypeptides and Its Use for the Specific Carboxyl-Terminal Labeling of Proteins[†]

Pierre-François Berne, Sylvain Blanquet, and Jean-Marie Schmitter*

Contribution from the Laboratoire de Biochimie, U.R.A. 240 CNRS, Ecole Polytechnique, 91128 Palaiseau Cedex, France. Received March 19, 1991

Abstract: A systematic study of the specific carboxyl-terminal labeling of oligo- and polypeptides has been undertaken. It relies on chemical esterification of the peptidic substrate followed by carboxypeptidase Y-catalyzed transpeptidation at basic pH with an amino acid derivative as nucleophile. Two nucleophiles, glycine isobutyl ester (Gly-OiBu) and valine amide (Val-NH$_2$), produced transpeptidation with a wide variety of short peptide substrates. Transpeptidation constants, i.e., concentrations of the nucleophile with which transpeptidation occurs at a velocity equal to that of the competing hydrolysis reaction, were in the range 1.2–21 mM for Val-NH$_2$ and 20–240 mM for Gly-OiBu, depending on the esterified peptidic substrate used. Therefore, these compounds were selected as potential nucleophiles for a general method of protein labeling. Using tritiated Val-NH$_2$ as nucleophile, labeling yields obtained with methylated porcine chymotrypsinogen A and methylated equine myoglobin as substrates were 31% and 39%, respectively. Several proteins, including *Escherichia coli* lysyl-tRNA synthetase, were labeled with tritiated Gly-OiBu or Val-NH$_2$ and proteolyzed. In each case, a labeled peptide could be isolated and identified as the carboxyl-terminal peptide.

Introduction

New sequences of proteins are more and more indirectly determined by cloning and sequencing the corresponding genes. In this context, analytical methods which demonstrate the conformity of the encoded protein are necessary, particularly concerning both ends of the polypeptide chain. While the reliable amino-terminal sequencing has long been possible thanks to the Edman chemistry,[1] the characterization of the carboxyl-terminal extremity remains a major source of ambiguity.

Numerous attempts have been made in order to find a solution to this problem. The oldest and most widely used method for the characterization of the carboxyl terminus consists of digestion of the protein with a carboxypeptidase followed by quantitative analysis of the released amino acids.[2,3] However, because of the complex action and the only partly deciphered specificity of carboxypeptidases, this method does not afford an unambiguous determination of the carboxyl-terminal sequence of a polypeptide. For instance, when carboxypeptidase comes to a slowly released residue followed by a quickly released one, both amino acids will appear simultaneously.

A set of methods has been aimed at the isolation of C-terminal peptides of proteins, such as comparison of peptidic maps before and after digestion with a carboxypeptidase,[4] proteolysis in the presence of H$_2$^{18}O followed by mass spectrometric analysis,[5] trypsinolysis followed by affinity chromatography over immobilized anhydrotrypsin[6] or by ion exchange chromatography,[7] and chemical modifications of the carboxyl groups and diagonal electrophoretic analysis of proteolytic fragments.[8,9] Additionally, attempts have been made to develop a carboxyl-terminal sequencing method analogous to the amino-terminal sequencing procedure of Edman, taking advantage of the thiocyanate chemistry.[10] However, in spite of recent improvements,[11] this method still lacks simplicity and general applicability.

Another way to characterize the carboxyl-terminal extremity of a protein may consist of introducing a label at this extremity. This may be achieved with the help of enzyme-catalyzed reactions.

In a previous work,[12] we described the specific carboxyl-terminal labeling of various peptides. The method drew its advantage from the ability of the serine exopeptidase carboxypeptidase Y (CPD-Y) to catalyze transpeptidations. This reaction consists of the replacement of the carboxyl-terminal peptide bond by a new bond, involving a nucleophilic molecule added to the incubation mixture.[13,14] It emerged from our previous work that, under the

conditions employed, the success of the labeling depended on the carboxyl-terminal amino acid sequence of the substrate: the presence of a Pro residue at the penultimate position of the substrate was shown to be a prerequisite for the occurrence of transpeptidation. Transpeptidations have also sometimes been observed with peptides containing a Pro farther from the carboxyl terminus, but such peptides were first shortened by the action of CPD-Y, without transpeptidation, until the Pro was made the penultimate residue.[12,15]

This lack of generality of the transpeptidation conditions can be circumvented by the use of mutagenesis: one may render a protein substrate favorable for labeling with CPD-Y by modifying its carboxyl-terminal sequence. We demonstrated that this strategy allowed the carboxyl-terminal labeling of a protein with a high specific radioactivity. In addition, the enzymatic activity of the protein was left intact in the procedure used. However, as it is based on the prior modification of a known gene, this method is of no help for primary structure determination.

For this reason, we have tried to design a method of carboxyl-terminal labeling that could be applied directly to any protein with any carboxyl-terminal sequence. This method still relies on CPD-Y-catalyzed transpeptidation but now requires a substrate with an esterified carboxyl terminus. The main advantage of using such a substrate is that, under basic pH conditions, CPD-Y ex-

---

(1) Edman, P. *Acta Chem. Scand.* **1950**, *4*, 283–293.
(2) Ambler, R. P. *Methods Enzymol.* **1972**, *25*, 143–154.
(3) Hayashi, R. *Methods Enzymol.* **1977**, *47*, 84–93.
(4) Isobe, T.; Ichimura, T.; Okuyama, T. *Anal. Biochem.* **1986**, *155*, 135–140.
(5) Rose, K.; Savoy, L. A.; Simona, M. G.; Offord, R. E.; Wingfield, P. *Biochem. J.* **1988**, *250*, 253–259.
(6) Kumazaki, T.; Nakako, T.; Arisaka, F.; Ishii, S. *Proteins: Struct. Funct. Genet.* **1986**, *1*, 100–107.
(7) Kawasaki, H.; Imajoh, S.; Suzuki, K. *J. Biochem.* **1987**, *102*, 393–400.
(8) Bietlot, H. P.; Carey, P. R.; Pozsgay, M.; Kaplan, H. *Anal. Biochem.* **1989**, *181*, 212–215.
(9) Dibó, G.; Furka, Á.; Sebestyén, F.; Fehérvári, Á. *Int. J. Peptide Protein Res.* **1986**, *27*, 355–359.
(10) Stark, R. G. *Methods Enzymol.* **1972**, *25*, 369–384.
(11) Bailey, J. M.; Shively, J. E. *Biochemistry* **1990**, *29*, 3145–3156.
(12) Berne, P. F.; Schmitter, J. M.; Blanquet, S. *J. Biol. Chem.* **1990**, *265*, 19551–19559.
(13) Widmer, F.; Johansen, J. T. *Carlsberg Res. Commun.* **1979**, *44*, 37–46.
(14) Breddam, K.; Widmer, F.; Johansen, J. T. *Carlsberg Res. Commun.* **1980**, *45*, 237–247.
(15) Carles, C.; Gueguen, P.; Ribadeau-Dumas, B. *FEBS Lett.* **1987**, *212*, 163–167.

hibits only an esterase activity and no peptidase activity.[16] Thus, transpeptidation would result in the elongation of the chain by one residue, with no chance to observe a preliminary shortening of the polypeptide, as may occur under neutral pH conditions.

Results concerning the transpeptidation of esterified oligopeptides, as model substrates, are presented as the first part of this work. These results led us to select two nucleophiles, valine amide (Val-NH$_2$) and glycine isobutyl ester (Gly-OiBu), which were synthesized as radiolabeled species. Using these compounds, the strategy of labeling was applied to various model proteins in order to establish its generality and efficiency.

Experimental Section

Materials. CPD-Y and *Staphylococcus aureus* protease were obtained from Boehringer (Mannheim, Germany) and trypsin (type XI, TPCK treated) from Sigma (St. Louis, MO). Stock solutions of CPD-Y (80–120 µM) were prepared in 50% glycerol, as previously described.[12] Other proteins were of the best obtainable grade from Serva (Heidelberg, Germany), except lysyl-tRNA synthetase (a gift from Dr. F. Lévêque), which was obtained as described.[17] The peptides YPFPG, YGGFMRGL, and Bz-Y-OEt were obtained from Tebu (France), and all other peptides were from Serva. [^{3}H]Val (24 Ci/mmol) and [^{3}H]Gly (19 Ci/mmol) were purchased from the Commissariat à l'Energie Atomique (Saclay, France).

Preparation of Substrates for Transpeptidation Reactions. Esterification and amidation of amino acids were performed using procedures inspired from those of Fischer[18] and Yang and Rising,[19] respectively. Pyridylethylation of proteins was performed essentially as described.[20] A variant of the procedure of Fraenkel-Conrat et al.[21] was used for the esterification of peptides and proteins. Our detailed procedures are available as supplementary material.

Formalism of the Transpeptidation Reaction. Transpeptidation under initial conditions: The following scheme describes the possible transformation of a peptide ester substrate, P-OX (X is the alkyl group of the ester), and of a nucleophile, Z, in the presence of CPD-Y (at concentration [E]).

$$P\text{-}OX \xrightarrow{v+v''} P \qquad (1)$$
$$v' \downarrow$$
$$P\text{-}Z$$
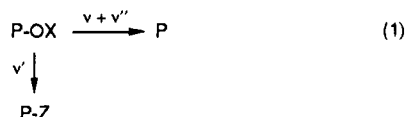
Initial rates v and v' reflect the esterase and transpeptidase activities of CPD-Y, respectively, whereas v'' reflects the spontaneous (nonenzymatic) hydrolysis of the ester bond.

The efficiency of transpeptidation under a given set of conditions is characterized by the initial fraction of transpeptidation, R_0, where $R_0 = v'/(v + v')$. As previously established,[12] the variation of R_0 toward [Z] is expected to be as follows:

$$R_0 = \frac{[Z]}{K + [Z]} \qquad (2)$$

where the transpeptidation constant, K, characterizes the reaction of a given substrate with a given nucleophile under defined conditions of pH and temperature.

Transpeptidation under extended conditions: When transpeptidation proceeds far from initial conditions, the possible hydrolysis of the transpeptidation product P-Z into P has to be taken into account.

$$P\text{-}OX \xrightarrow{\lambda} P \qquad (3)$$
$$\lambda' \downarrow \quad \nearrow \lambda'''$$
$$P\text{-}Z$$
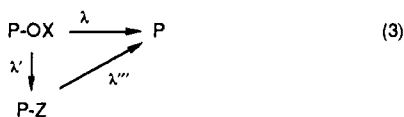
In this scheme, λ, λ', and λ''' are second-order rate constants, which assumes that the concentration of the substrate is far lower than its

(16) Hayashi, R.; Bai, Y.; Hata, T. *J. Biochem.* **1975**, *77*, 69–79.
(17) Lévêque, F.; Plateau, P.; Dessen, P.; Blanquet, S. *Nucleic Acids Res.* **1990**, *18*, 305–312.
(18) Fischer, E. *Chem. Ber.* **1901**, *34*, 433–454.
(19) Yang, P. S.; Rising, M. M. *J. Am. Chem. Soc.* **1931**, *53*, 3183–3184.
(20) Friedman, M.; Krull, L. H.; Calvins, J. F. *J. Biol. Chem.* **1970**, *245*, 3868–3871.
(21) Fraenkel-Conrat, H.; Olcott, H. S. *J. Biol. Chem.* **1945**, *161*, 259–268.

Michaelis constant (K_M) for CPD-Y. The spontaneous hydrolysis reaction is neglected here. Considering the yield of the reaction, r, as the concentration of the transpeptidation product over the initial concentration of the substrate ($r = [\text{P-Z}]/[\text{P-OX}]_0$), the analytical solution of the differential equations relative to eq 3 is the following:

$$r = \frac{\lambda'}{\lambda + \lambda' - \lambda'''}(e^{-\lambda'''t[E]} - e^{-(\lambda+\lambda')t[E]}) \qquad (4)$$

Using the relation between λ and λ', similar to that occurring between the initial rates v and v', as given in eq 2,

$$\frac{\lambda'}{\lambda + \lambda'} = R_0 = \frac{[Z]}{K + [Z]} \qquad (5)$$

it follows that

$$r = \frac{[Z]}{(K + [Z])(1 - \gamma)}(e^{-\gamma(\lambda+\lambda')t[E]} - e^{-(\lambda+\lambda')t[E]}) \qquad (6)$$

where the ratio $\gamma = \lambda'''/(\lambda + \lambda')$ reflects the relative rate of enzyme-catalyzed hydrolysis of the transpeptidation product P-Z to P, as compared to that of the starting substrate P-OX.

Theoretical plots $r = f([E])$ are bell-shaped, and the parameters K, γ, and $\lambda + \lambda'$ can be determined by fitting the experimental data to the formal expression

$$r = A(e^{-k[E]} - e^{-k'[E]}) \qquad (7)$$

Least-squares analyses were performed with a computer-assisted iterative process based on the Newton–Gauss linearization method.

Transpeptidation of Peptide Esters. Prior to incubation, peptide esters were diluted in a neutral buffer (5 mM potassium phosphate, pH 7.0). CPD-Y was diluted to various concentrations in the same buffer and kept on ice. Tubes containing 100 µL of the reaction buffer (0.1 M sodium carbonate buffer, 0.1 mM EDTA, pH 9.5–11) and the nucleophile (0.1–100 mM) were equilibrated at 37 °C. Reactions were then started by the simultaneous addition of 5 µL of the peptide solution (5–50 µM final concentration) and 2 µL of the CPD-Y solution (0.1–100 nM final concentration). Reactions were stopped after 3 min by the addition of 20 µL of 10% aqueous TFA. The resulting pH (close to 1) irreversibly inactivated CPD-Y. The pH of the sample was then raised by the addition of 100 µL of a buffered solution containing 0.2 M sodium acetate, pH 5.5, and the samples were kept at room temperature for 0–12 h until chromatographic analysis. In some cases, the supplementary addition of 20 µL of methanol (10% final concentration) was necessary to prevent the precipitation of hydrophobic peptides during this step.

Analysis of the samples was performed by RPLC. The final mixture (200 µL) was injected onto a Superspher C18 column (250 × 4 mm, 4-µm particles, Merck), and an increasing gradient of acetonitrile in 0.1% TFA was applied at a flow rate of 0.9 mL/min at 42 °C. Detection and quantification of the compounds were based upon the fluorescence of either Tyr ($\lambda_{ex} = 272$ nm, $\lambda_{em} = 303$ nm) or Trp ($\lambda_{ex} = 280$ nm, $\lambda_{em} = 348$ nm).

The initial fraction of transpeptidation, R_0, can be measured as the ratio, $R = [\text{P-Z}]/([\text{P}] + [\text{P-Z}])$, provided that (i) initial rate conditions are satisfied and (ii) the effect of the spontaneous hydrolysis of the esterified substrate is negligible ($v'' \ll v$). Short incubation times and adequate CPD-Y concentrations were not always sufficient to satisfy the latter condition. Therefore, the measured R values were systematically corrected by using the first-order rate constant, k'', of the nonenzymatic ester hydrolysis reaction, which was determined in a separate experiment. Assuming an exponential decrease of the substrate P-OX during the reaction, R values were divided by a factor, f:

$$f = 1 - \frac{k''t}{\ln\left(\dfrac{[\text{P-OX}]_0}{[\text{P-OX}]}\right)} \qquad (8)$$

where the ratio $[\text{P-OX}]/[\text{P-OX}]_0$ represents the fraction of the esterified substrate still present at time t.

Initial fractions, R_0, were measured at various nucleophile concentrations, [Z]. Firstly, a linear regression was performed using the couples $1/[Z]$, $1/R_0$ to ensure that the value of $1/R_0$ at $1/[Z] = 0$ was in the range 0.9–1.1. This control was necessary in order to confirm the validity of eq 2, as the dependence of R_0 toward [Z] was shown to be $R_0 = \rho[Z]/(K + [Z])$, with $\rho < 1$, in some cases.[12] The K constant was then obtained as the mean of the values calculated ($K = [Z](1 - R_0)/R_0$) from couples [Z], R_0, with a weighting factor $R_0(1 - R_0)$ applied to each couple.

Transpeptidation of Esterified Proteins. Reactions were initiated by the addition of 15 µL of 10 mM sodium acetate (pH 5.5) containing the

Carboxyl-Terminal Labeling of Proteins

J. Am. Chem. Soc., Vol. 114, No. 7, 1992 2605

esterified protein (8.5 mg/mL) and various concentrations of the tritiated nucleophile to 15 μL of a buffer containing 0.2 M sodium carbonate (pH 10–11), 0.2 mM EDTA, and 0.8% SDS (37 °C). This was immediately followed by the addition of 2 μL of 5 mM potassium phosphate (pH 7.0) containing CPD-Y (1 nM–7 μM final concentration).

After 3–5 min at 37 °C, the reaction was quenched by the addition of 1 mL of 5% trichloroacetic acid, and the sample was allowed to precipitate on ice. The precipitate was collected on a glass microfiber filter and, after the addition of Picofluor (Packard), the radioactivity was measured by liquid scintillation counting in a Beckman LS1801.

The yield of labeling, r, was calculated as the ratio of the specific radioactivity of the protein to that of the nucleophile. When the experimental plots of r as a function of CPD-Y concentration were bell-shaped, as predicted by the theoretical expression of eq 6, the K and γ values were obtained by fitting the data to eq 7, as indicated above. In cases where the experimental plots were not bell-shaped but regularly increasing, this was assumed to reflect that the parameter γ ≪ 1. In these cases, the value of the transpeptidation constant, K, was estimated using the expression $r_{max} = [Z]/K$, taking for r_{max} the value of r at the highest assayed CPD-Y concentration. This approximation may lead to an overestimation of K.

In some cases, the K constant was more accurately determined by varying the concentration of the nucleophile in the labeling reaction. The parameters of eq 7 were determined by least-squares analysis for transpeptidation experiments at various nucleophile concentrations. Then, the value of $(1 - k/k')/A$, which should represent $1/R_0 = 1 + K/[Z]$, was plotted as a function of $1/[Z]$, and the value of K was obtained from a linear regression.

Preparative Labeling of Proteins. Conditions for analytical transpeptidation experiments were scaled up (typically to 250 μL total volume) for the labeling of proteins in the milligram range. The CPD-Y concentration was chosen after the preliminary transpeptidation experiments were made on smaller amounts of protein substrate (40 μg). The labeling reaction was terminated after 4 min by the addition of 2 μL of DIFP (50 mM) to the incubation mixture (about 260 μL) and by a further 5-min incubation at room temperature. The pH was then raised to 12 by the addition of 2 μL of NaOH (9 M), and the labeled protein was incubated for 3 h at 37 °C in order to hydrolyze any remaining ester groups. Finally, the protein sample was dialyzed first against 1 L of 0.1 M sodium bicarbonate (pH 8.0) and second against the convenient buffer for subsequent proteolysis (2 × 1 L).

Characterization of the Radiolabeled Proteins. Labeled proteins (2 mg/mL), except lysyl-tRNA synthetase, were digested by trypsin (70 μg/mL final concentration) for 3–24 h at 37 °C in a buffer containing 4 M urea, 0.1 M ammonium acetate (pH 8.0), and 20 mM calcium chloride. Lysyl-tRNA synthetase (1 mg/mL) was digested at 25 °C by S. aureus protease (50 μg/mL final concentration) in buffer containing 50 mM potassium phosphate (pH 7.8).

Proteolysis was stopped by the injection of the sample onto a Superspher C18 column (250 × 4 mm, 4-μm particles, Merck), equilibrated at 42 °C in 0.1% TFA. Peptides were eluted at a flow rate of 0.9 mL/min by a linear gradient of acetonitrile (0.8% min⁻¹) in 0.1% TFA. The collected fractions (0.9 mL) of the eluate were analyzed by liquid scintillation counting. After concentration in a vacuum centrifuge, radioactive fractions were further purified using the same chromatographic system with a smoother gradient of acetonitrile (0.16% min⁻¹) at a flow rate of 1.2 mL/min (45 °C).

The amino acid sequences of radioactive peptides were determined using a gas-phase sequencer (Applied Biosystems Model 470A), as previously described,[22] and were verified by plasma desorption mass spectrometry.

Plasma Desorption Mass Spectrometry. A ^{252}Cf fission fragment ionization time of flight mass spectrometer was used for plasma desorption mass spectrometry analyses. This instrument (Depil-X) has been constructed at the Institut de Physique Nucléaire (Orsay, France).[23]

Nitrocellulose backings, produced by electrospraying 25 μL of a 2 mg/mL solution of nitrocellulose in acetone, were used for sample loading. Peptide samples were dissolved at 0.1–1 mM concentrations in 50 mM citric acid solution containing 50% methanol. A 1–3-μL volume was dried on the surface. Spectra were acquired with an acceleration voltage of 10 kV in the positive ion mode.

Results and Interpretation

I. Transpeptidation of Short Peptides. Under basic pH conditions, carboxyl-substituted amino acids have been shown to be

Table I. First-Order Rate Constants, k'', for the Base-Catalyzed Hydrolysis of Peptide YPFPG Esters

peptide	pH	k'' (10^{-3} min⁻¹)	
		25 °C	37 °C
YPFPG-OEt	9.5	1.1	4.5
	10	3.3	13
YPFPG-OMe	9.5	3	11
	10	8	32

much more efficient nucleophiles than unsubstituted amino acids.[13,24] For this reason, we focused on amino acid amides (principally valine amide) and glycine esters.

A. Base-Catalyzed Hydrolysis of Peptide Esters. Each determination of a transpeptidation constant K (cf. Experimental Section) was accompanied by the measurement, under the same conditions, of the rate of spontaneous hydrolysis of the carboxyl-terminal ester of the substrate. Results obtained with the esterified peptide YPFPG (Table I) are representative of others. They indicate that a pH increase of 0.5 unit accelerates the reaction of spontaneous hydrolysis by a factor of 3, which is consistent with a reaction rate proportional to the hydroxyl ion concentration. Similarly, an increase in the temperature from 25 to 37 °C increases the reaction rate 4-fold. In addition, the peptide methyl ester is 2.5 times more labile than the corresponding ethyl ester.

B. Characterization of the Transpeptidation Products. Starting with a peptide ester substrate P-OEt, such as YPFPG-OEt or LWMR-OEt, and using Val-NH₂ and Gly-OX as nucleophiles, the expected transpeptidation products were P-V-NH₂ and P-G-OX, respectively. Upon RPLC analysis, these products appeared at retention times different from that of the hydrolysis product P-G. After RPLC purification, these products were characterized by plasma desorption mass spectrometry (supplementary material).

As mentioned by Breddam et al.,[14] in the presence of amino acid amides, the amidase activity and the repetitive action of CPD-Y may result in an elongation of the product. For instance, P-VV-NH₂, P-VVV-NH₂, etc. might occur in the presence of Val-NH₂. In the presence of high concentrations of Val-NH₂ and CPD-Y, we observed with some peptides a species more hydrophobic than P-V-NH₂ and less hydrophobic than P-OEt, according to its RPLC retention time. However, this species never represented more than 5% of the major product P-V-NH₂. With amino acid esters, polytranspeptidations are also described as possible side reactions of transpeptidations.[25] In the presence of glycine esters, this phenomenon would yield the products P-GG-OX, P-GGG-OX, etc., which may not be separated by chromatography from the product of monotranspeptidation, P-G-OX. Therefore, traces of the elongated products were carefully searched for when mass spectrometry analysis was performed, but only the product of monotranspeptidation was detected.

C. Variation of the Extent of Transpeptidation Reactions. When the reaction time of transpeptidations was varied from 3 to 10 min, only a small variation of the extent of reaction was noticed. This was attributed to the fact that CPD-Y rapidly lost its activity under the experimental conditions, with a half-time of inactivation of about 3 min (data not shown). This instability of CPD-Y was not observed previously when the pH range 4.5–8.5 was used.[12] Therefore, time was not an efficient parameter for control of the reaction, and the extent of reaction was best controlled by varying the CPD-Y concentration employed. A short reaction time (3–4 min) was chosen in order to reduce the extent of nonenzymatic hydrolysis of the ester function of the substrate.

Typical transpeptidation reactions were monitored by measuring the proportion of both the transpeptidation product and the hydrolysis product at a constant reaction time, t, as a function of CPD-Y concentration. As exemplified by the transpeptidation of the peptide LWMR-OEt in the presence of 15 mM Val-NH₂

(22) Beauvallet, C.; Hountondji, C.; Schmitter, J. M. *J. Chromatogr.* **1988**, *438*, 347–357.

(23) Della-Negra, S.; Le Beyec, Y. *Int. J. Mass Spectrom. Ion Proc.* **1984**, *61*, 21–29.

(24) Widmer, F.; Breddam, K.; Johansen, J. T. *Carlsberg Res. Commun.* **1981**, *46*, 97–106.

(25) Widmer, F.; Breddam, K.; Johansen, J. T. *Carlsberg Res. Commun.* **1980**, *45*, 453–463.
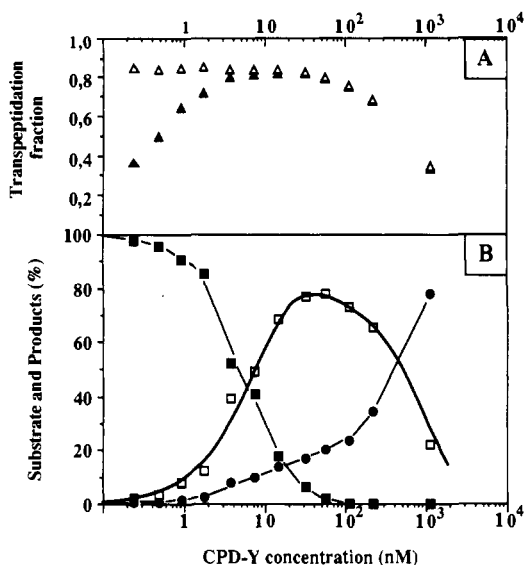
Figure 1. Transpeptidation of the peptide ester LWMR-OEt (10 μM) with Val-NH$_2$ (45 mM) at pH 10 as a function of CPD-Y concentration (3 min at 37 °C). Panel A: The yield of transpeptidation, R = [LWMRV-NH$_2$]/([LWMRV-NH$_2$] + [LWMR]) ($\blacktriangle$) and the corrected yield, R' ($\triangle$), are compared. The correction was made by taking into account a rate constant k'' = 1.12 × 10^{-3} min^{-1} for the spontaneous hydrolysis of the peptide ester substrate. Panel B: The relative percentages of the substrate, LWMR-OEt ($\blacksquare$), and of the two products, LWMR ($\bullet$) and LWMRV-NH$_2$ ($\square$), are plotted. These percentages were corrected in a manner analogous to that of the correction of the yield R. The thicker line is the graph of y (%) = 81.3(exp((−9.49 × 10^{-4})x (nM)) − exp(−0.135x (nM)), resulting from the fit of the experimental values of the percentage of LWMRV-NH$_2$, r, to eq 7 (cf. Experimental Section).

(Figure 1), the crude fraction of transpeptidation, R = [LWMRV-NH$_2$]/[LWMRV-NH$_2$] + [LWMR]), increases from 35 to 80% when the CPD-Y concentration is increased from 0.25 to 4 nM (panel A). This phenomenon results from the fact that a higher CPD-Y concentration, increasing both the [LWMRV-NH$_2$] and [LWMR] terms, reduces the contribution of the spontaneous hydrolysis to [LWMR]. If a correction is made (cf. Experimental Section) in order to compensate for the effect of the nonenzymatic reaction, the calculated value R' = R/f becomes nearly independent of the CPD-Y concentration, as expected (Figure 1). This illustrates well the usefulness of correcting R values in order to interpret the variations of the fraction of transpeptidation as a function of CPD-Y concentration and to accurately determine the initial fraction, R_0.

In the case of LWMR-OEt (Figure 1), R' remains equal to the initial fraction, R_0 (83%), when as much as 95% of the substrate is consumed (30 nM CPD-Y). This behavior indicates that the newly synthesized peptide bond of LWMRV-NH$_2$ is resistant to hydrolysis by CPD-Y under the reaction conditions. By fitting the experimental data to the theoretical relationship (eq 6) as indicated in the Experimental Section, it was demonstrated that the hydrolysis of LWMRV-NH$_2$ by CPD-Y was 150 times slower than that of the starting substrate LWMR-OEt (γ = 0.007).

Similar variations of R' upon increasing the CPD-Y concentration were observed with other peptide ethyl ester substrates, confirming that CPD-Y hydrolyzed the terminal amino acid amide of the transpeptidation product much more slowly than the ethyl ester of the substrate.

D. Influence of the Alkyl Group of the Ester Substrate. Methyl and ethyl groups as the carboxyl-terminal substituent X of the substrate YPFPG-OX were compared in the presence of the Val-NH$_2$ nucleophile. The differences in K values at pH 9.5 or 10 and at 25 or 37 °C are not significant (Table II). This can be explained by the fact that the K value partly reflects the competition between water and the nucleophile, once the transient acyl-enzyme is formed and the methanol or ethanol molecule released. However, the rate of spontaneous hydrolysis of the

Table II. Transpeptidation Constants, K, for Various Peptide Ester/Amino Acid Amide Couples

substrate	nucleophile	pH	K^a (mM)	
			25 °C	37 °C
YPFPG-OEt	Val-NH$_2$	9.5	2.2	1.9
		10	1.8	1.4
YPFPG-OMe	Val-NH$_2$	9.5	2.4	nd^b
		10	1.7	nd^b
Bz-Y-OEt	Val-NH$_2$	9.5	nd^b	1.5
		10	1.5	1.1
	Leu-NH$_2$	9.5	2.1	1.3
		10	1.2	0.7
	Met-NH$_2$	9.5	1.8	1.7
		10	1.4	1.2
	Gly-NH$_2$	10	nd^b	90
	Phe-NH$_2$	10	nd^b	2
	Met	10	nd^b	170
LWMR-OEt	Val-NH$_2$	10	nd^b	11
		11	nd^b	7.2
LWMRFA-OEt	Val-NH$_2$	10	nd^b	16
LLLY-OEt	Val-NH$_2$	10	nd^b	4.5
Cbz-HPW-OEt	Val-NH$_2$	10	nd^b	21
	Leu-NH$_2$	10	nd^b	19

^{a} Relative standard error on K was 10%. ^{b} nd = not determined.

methyl ester is 2.5-fold higher than that of the ethyl ester (Table I). As a result, the corrections performed in order to deduce initial fractions of transpeptidation, R_0 = $\nu'/(\nu + \nu')$, from experimental fractions, R, are greater with the methyl ester substrate than with the ethyl ester one. For this reason, peptide ethyl esters allowed more accurate determinations of K constants and were preferred in subsequent experiments.

E. Test of Amino Acid Amides as Nucleophiles. Amino acid amides were reported to be much more potent nucleophiles in CPD-Y-catalyzed transpeptidations than the corresponding unsubstituted amino acids. All amino acid amides, except proline amide, gave similar fractions of transpeptidation regardless of the nature of the side chain (no transpeptidation occurred with proline amide).[24] Our results are distinguishable from previously published data since, in the present study, the concentrations, K, of nucleophile required to reach an initial fraction R_0 = 50% are considered, while the other authors compared experimental yields obtained in the presence of the highest possible concentration of nucleophile. In spite of this methodological difference, the results that we obtained with the Bz-Tyr-OEt substrate (Table II) confirm those obtained by others.[24] Met is a very poor nucleophile as compared to Met-NH$_2$, the K constants at pH 10 differing by a factor of 150. Met-NH$_2$, Val-NH$_2$, Phe-NH$_2$, and Leu-NH$_2$ give comparable K values, while the value using Gly-NH$_2$ is 80 times higher. The choice of Val-NH$_2$ for further investigations is justified by the fact that, when this residue occupies the carboxyl-terminal position of a peptide, its amide group is resistant to hydrolysis by CPD-Y. Thus, contrary to results observed with other amino acid amides, such as Phe-NH$_2$ or Leu-NH$_2$, transpeptidation with Val-NH$_2$ resulted in little polytranspeptidation.[14] This result is confirmed here by the fact that polytranspeptidation products never accounted for more than 5% with Val-NH$_2$, as mentioned above.

F. Test of Glycine Esters as Nucleophiles. Amino acid esters are also capable of substituting for the carboxyl-terminal ester function of an esterified polypeptide. These compounds are interesting because both the ester alkyl chain and the amino acid side chain can be varied and because they are easily synthesized. When various amino acid esters were compared in a previous study,[25] the authors noticed a particular behavior of glycine derivatives. For all other amino acids, only the methyl ester and, to a lesser extent, the ethyl ester were found to be efficient nucleophiles. In contrast, esters of Gly carrying higher alkyl chains were still able to produce transpeptidation with high yields.

For these reasons, we decided to focus on glycine esters in order to attempt to use one of these compounds, synthesized from [^{3}H]Gly, to label proteins. For the sake of comparison, some of the experiments described below were also made with Val-OEt

Table III. Transpeptidation Constants, K, for Various Peptide Ester/Gly or Val Ester Couples

substrate	nucleophile	pH	K^a (mM)	substrate	nucleophile	pH	K^a (mM)
LWMR-OEt	Gly-OEt	10	100	LWMRFA-OEt	Gly-OiBu	10	92
	Gly-OiPr	10	83		Gly-OtBu	10	55
	Gly-OnPr	10	58	YPFPG-OEt	Gly-OEt	10	58
	Gly-OiBu	9.5	47		Gly-OiBu	10	28
	Gly-OiBu	10	33		Gly-OtBu	10	8
	Gly-OiBu	10.5	20		Val-OMe	10	16
	Gly-OtBu	10	27		Val-OtBu	10	nm
	Gly-OtBu	10.5	9.6	YGGFMRGL-OEt	Gly-OEt	10	nm
	Val-OEt	10	nm^b		Gly-OiBu	10	240
	Val-OtBu	10	nm		Gly-OiBu	10	390
LWMRFA-OEt	Gly-OEt	10	130		Val-OMe	10	nm
	Gly-OiPr	10	135		Val-OtBu	10	nm

^{a}Relative standard error on K was 15%. ^{b}nm = not measurable (transpeptidation was not detected).

and Val-OtBu. However, no transpeptidation could be detected with these two nucleophiles (Table III).

A systematic study of glycine esters was undertaken. The K constant was determined for various combinations of substrates and nucleophiles (Table III). The results confirmed a trend, already observed by others,[25] concerning the alkyl group; an increase in the number of carbon atoms of the alkyl group (methyl → butyl) favored transpeptidation. In addition, when glycine esters of the same number of carbon atoms are compared, it appears that an ester having a more branched alkyl chain behaves as a better nucleophile.

In all cases studied, except with the peptide YGGFMRGL-OEt, the *tert*-butyl ester was the most efficient nucleophile. However, the synthesis of this compound from [³H]Gly is problematic, because tertiary alcohols do not react under the esterification conditions based on the method of Fischer.[18] The synthesis of *tert*-butyl esters of amino acids by another method has been described,[26] but the yield was only about 45% and the given method not suitable for the synthesis of submilligram amounts of a radiolabeled compound. For this reason, protein labeling experiments, which will be presented in part II of this section, were conducted with [³H]Gly-OiBu, obtained from [³H]Gly in nearly theoretical yield.

G. Influence of pH. K constants for various peptide ethyl ester and amino acid amide combinations were compared at two pH values in the range 9.5-11 (Table II) and were found to be systematically lower at the higher pH. The same effect is observed when glycine esters are used as nucleophiles (Table III). As an example, using the LWMR-OEt substrate, the K constant with Gly-OiBu decreases almost 2-fold between pH 10 and 10.5. The effect is stronger in the case of Gly-OtBu, where a 3-fold decrease is observed.

However, the decrease of K upon increasing the pH may be counterbalanced by a higher rate of the spontaneous hydrolysis. For this reason, if one considers crude fractions of transpeptidation rather than calculated K constants, the variation toward pH is expected to exhibit a maximum, as observed in previous studies.[13] The high rate of spontaneous hydrolysis of some substrates at pH values above 10 practically rendered impossible the determination of their transpeptidation parameters.

H. Influence of Temperature. Increasing the temperature has a favorable effect on the transpeptidation constant, K, as measured using Bz-Tyr-OEt as substrate and Val-NH₂ as nucleophile (Figure 2). The same effect is found when K constants are compared at 25 and 37 °C using other substrates (Table II). This behavior is in agreement with observations by others at pH 9.7 using Bz-Ala-OMe as substrate.[13] However, it must be noted that the spontaneous hydrolysis of the carboxylic ester of the substrate can also be accelerated by an increase in temperature (Table I).

I. Influence of the Substrate. The K constant was measured for seven different peptide ethyl esters, at pH 10 and 37 °C, using Val-NH₂ as nucleophile (Table II). K values depend on the sequence of the substrate, with values ranging from 1 to 21 mM. However, these values remain relatively small when compared,
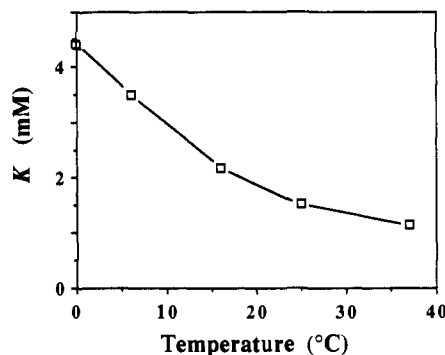

Figure 2. Influence of temperature on the transpeptidation constant, K, for the reaction of Bz-Tyr-OEt (50 μM) with the Val-NH₂ nucleophile at pH 10. Various nucleophile concentrations, in the range 0.5-5 mM, were used for each determination of K. A 3 nM concentration of CPD-Y yielded 4.5% (0 °C), 8.5% (6 °C), or 10% (16-37 °C) transformation of the substrate in 3 min.

for example, to that measured for the transpeptidation of Bz-Tyr-OEt with Met (170 mM). It is, thus, conceivable to use Val-NH₂ as a general nucleophile for the carboxyl-terminal labeling of proteins via transpeptidation.

Similarly, K constants determined with glycine esters depend on the nature of the peptide ester substrate without exhibiting a high disparity. It is noteworthy that favored substrates for transpeptidation with Val-NH₂, such as YPFPG-OEt, are also favored with glycine esters. This behavior argues in favor of the presence in the substrate of some sequence determinants for the transpeptidation reaction which are independent of the nucleophile used. Much more data would be necessary, however, to decipher these determinants.

II. Labeling of Proteins. As a result of the studies presented in part I, tritiated Val-NH₂ and Gly-OiBu were used in order to compare their efficiencies in the labeling of esterified proteins.

A. Esterification of Proteins. The procedure of esterification of proteins in methanolic HCl has already been extensively studied.[21,27] To be complete, the reaction requires a 24-h incubation at 25 °C in the presence of 0.1 N HCl in methanol. The only described side reaction is the methanolysis of the amide function of the Gln and Asn residues.[27] In contrast to a previous study[21] where it was mentioned that this reaction occurred to a low extent under the above conditions, we observed significant degradation of the polypeptides even at a HCl concentration as low as 0.05 N, as revealed by SDS–polyacrylamide gel electrophoresis analysis (performed with the Phast-Gel system from Pharmacia). By performing the reaction for 48 h at a lower temperature (5 °C) in the presence of 0.2 N HCl, degradation was no longer noticed by electrophoretic analyses.

Under these conditions, equilibrium was reached in the reaction of esterification of the terminal carboxyl group, as indirectly indicated by the fact that yields of labeling of proteins (see below)

(26) Roeske, R. W. *Chem. Ind. (London)* **1959**, 1121-1122.

(27) Chibnall, A. C.; Mangan, J. L.; Rees, M. W. *Biochem. J.* **1958**, *68*, 114-118.

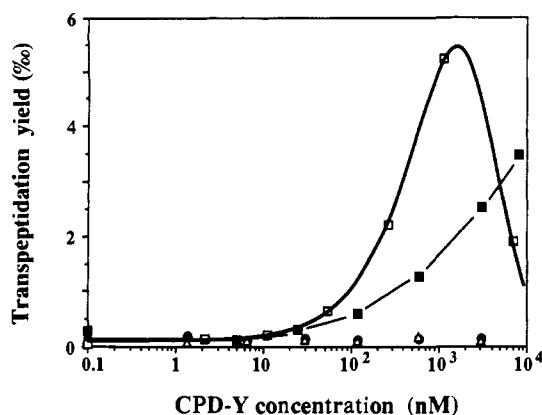2608   *J. Am. Chem. Soc., Vol. 114, No. 7, 1992*

Berne et al.

Figure 3. CPD-Y-catalyzed labeling of methylchymotrypsinogen. The incubation mixture (15 µL) contained 0.1 M sodium carbonate (pH 10), 0.1 mM EDTA, 0.4% SDS, 200 µM esterified protein, catalytic amounts of CPD-Y, and either 10 µM [³H]Val-NH₂ (□) or 30 µM [³H]Gly-OiBu (■). The thicker line is the graph of y (‰) = 9.84(exp((-2.42 × 10⁻⁴)x (nM)) - exp((-1.24 × 10⁻³)x (nM)), resulting from the fit of the experimental data (□) to eq 7. Control experiments were performed under identical conditions with 30 µM [³H]Gly (△) as nucleophile or with 200 µM nonesterified protein as substrate (●). In the latter control experiment, [³H]Gly-OiBu (30 µM) was used as the nucleophile.

were not improved upon prolonging the esterification reaction 3-fold. Since methanol is at least 1000-fold in excess over H₂O, the reaction equilibrium is likely to be fully shifted toward esterification.

Methylated proteins appeared as white powders, soluble at a 10 mg/mL concentration in 10 mM sodium acetate buffer (pH 5.0). As already observed,[21] they precipitated when the pH was raised to 10. However, the suspension produced could be easily solubilized by the addition of SDS (0.3–0.5% final concentration).

B. Transpeptidation of Proteins. The two radioactive nucleophiles, [³H]Gly-OiBu (30 µM) or [³H]Val-NH₂ (10 µM), were incorporated into equine chymotrypsinogen A as a function of CPD-Y concentration (Figure 3). Optimally, 5.2 or 3.5‰ of the protein substrate molecules incorporated a Val-NH₂ or a Gly-OiBu molecule, respectively. Such low yields are explained by the small nucleophile concentrations used, as compared to the corresponding transpeptidation constant values. Much more efficient incorporations (up to 30%) were obtained at higher nucleophile concentrations, as will be shown below. When [³H]Val-NH₂ is used, the incorporation of radioactivity decreases at the highest assayed CPD-Y concentration. This may be attributed to the subsequent hydrolysis by CPD-Y of the carboxyl-terminal Val-NH₂ that has just been incorporated. In contrast, the protein labeled with Gly-OiBu appears more stable, since the yield of labeling regularly increases as a function of CPD-Y concentration (Figure 3).

Seven other protein substrates with various carboxyl-terminal sequences were tested (Table IV). All incorporate labeled nucleophiles in the presence of CPD-Y. In the case of lysozyme, this incorporation was, however, hardly significant, which will be explained later. The plots of incorporated radioactivity versus CPD-Y concentration either were bell-shaped or regularly increased, as illustrated by the two graphs of Figure 3. When the graph is bell-shaped, the ratio, γ, between the rate of hydrolysis of the transpeptidation product and that of the substrate P-OEt can be determined from the experimental data (cf. Experimental Section). In the case of the labeling of chymotrypsinogen and myoglobin with Val-NH₂, γ values of 0.16 ± 0.04 and 0.04 ± 0.01 were obtained, respectively. In other words, the data are consistent with the newly introduced carboxyl-terminal Val-NH₂ residue of chymotrypsinogen and myoglobin being hydrolyzed 6 and 25 times more slowly than the carboxyl-terminal ester, respectively. This stability of the transpeptidation product, as compared to the ester substrate, has already been outlined with synthetic peptides. Non-bell-shaped graphs—the highest yield of labeling was experimentally obtained at the highest assayed CPD-Y

Table IV. Transpeptidation Constants, K, for Various Proteins in the Presence of the Val-NH₂ or Gly-OiBu Nucleophile[a]

protein substrate and its carboxyl-terminal sequence[c]	pH	Gly-OiBu [CPD-Y] (µM)	Gly-OiBu K^b (mM)	Val-NH₂ [CPD-Y] (µM)	Val-NH₂ K^b (mM)
BSA (TQTALA)	10	0.65ᵈ	9 ± 2	0.13ᵈ	24 ± 4
	10.5	3.25ᵉ	9 ± 2		nd
	11	3.25ᵉ	8 ± 2	0.13ᵈ	19 ± 4
	10ᶠ	1.3ᵈ	15 ± 3	0.25ᵈ	44 ± 12
chymotrypsinogen A (QTLAAN)	10	7ᵉ	8 ± 1	1.3ᵈ	1.3 ± 0.4
	10.5	3.25ᵉ	18 ± 3		nd
	10ᶠ	6.3ᵉ	17 ± 3	1.3ᵈ	6.5 ± 1.2
creatine kinase (LMPAQK)	10	3.25ᵉ	26 ± 5	0.65ᵈ	10 ± 2
	10.5		nd	0.65ᵈ	6.6 ± 1.5
	11	3.25ᵉ	19 ± 3	0.65ᵈ	6.1 ± 0.9
ferritin (EEKPKN)	10	3.25ᵈ	49 ± 8	3.25ᵈ	20 ± 3
HSA (QAALGL)	10	3.25ᵉ	25 ± 5	0.65ᵈ	11 ± 2
	10.5	3.25ᵉ	24 ± 5		nd
	11	3.25ᵉ	19 ± 4	3.25ᵉ	15 ± 2
lysozyme (IRGCRL)	10	0.65ᵈ	240 ± 50	0.62ᵈ	82 ± 15
	10ᶠ	0.25ᵈ	59 ± 10	0.25ᵈ	6.0 ± 1.5
myoglobin (ELGFQG)	10	0.13ᵈ	11 ± 1	0.13ᵈ	3.0 ± 0.4
	10.5	0.65ᵈ	9.5 ± 1		nd
trypsin (QTIASN)	10	3.25ᵉ	39 ± 6	3.25ᵉ	6.5 ± 0.8
	11	3.25ᵉ	31 ± 5	3.25ᵉ	7.0 ± 0.8
LysRS (AMRPVK)	10		nd	0.75ᵈ	12 ± 4

[a] nd = not determined. BSA = bovine serum albumin. HSA = human serum albumin. LysRS = lysyl-tRNA synthetase. [b] K values were obtained as described in the Experimental Section from the transpeptidation of the protein substrates in the presence of 6 µM Val-NH₂ or 40 µM Gly-OiBu. [c] Sequences, except that of lysyl-tRNA synthetase (17), were extracted from the NBRF protein data base. [d] CPD-Y concentration producing maximal yield of labeling, r. [e] Highest assayed CPD-Y concentration. The yield of labeling, r, could increase further at higher CPD-Y concentrations. Consequently, the value of K is overestimated. [f] Protein modified with vinylpyridine prior to esterification.

concentration—were more frequently obtained with the Gly-OiBu nucleophile (cf. Figure 3). This may reflect a lower rate of CPD-Y-catalyzed processing of a carboxyl-terminal Gly-OiBu residue as compared to a Val-NH₂ one.

Detailed data concerning the labeling of the seven model proteins are presented in Table IV. Values of the transpeptidation constant, K, were obtained from labeling experiments performed at one given concentration of the nucleophile (typically 6 µM Val-NH₂ or 40 µM Gly-OiBu), as described in the Experimental Section. It is remarkable that, using either nucleophile, K constants with the protein substrates are in the same range (1.2–24 mM for Val-NH₂, 8–59 mM for Gly-OiBu) as with the esterified peptide substrates. As already observed with model peptides, the measured K values remain on the same order of magnitude. This further supports the idea that transpeptidation reactions conducted under constant conditions are not highly dependent on the sequence of the substrate.

The influence of pH on protein labeling with the two nucleophiles was investigated. Transpeptidation does not significantly depend on pH in the range 10–11. This may indicate that the decrease in the K value upon increasing the pH, as observed using short peptides, is compensated by the faster spontaneous hydrolysis of the protein substrate and, in the case of Gly-OiBu, by that of the nucleophile.

Prior reduction of the disulfide bonds with Cys derivatization had an unfavorable effect on the transpeptidation of both chymotrypsinogen and bovine serum albumin. This may be attributed to an incomplete esterification of these pyridylethylated proteins, since they aggregate in methanol during the esterification reaction. In contrast, the reduction of disulfide bonds was necessary to obtain the labeling of lysozyme. This may be explained by the fact that lysozyme contains an antepenultimate Cys residue involved in an

Table V. Parameters for the Preparative Labeling of Esterified Proteins[a]

protein		nucleophile			characterization of the labeling	
	amount involved (nmol)		concn (mM)	specific activity (Ci/mol)	yield (%)[b]	peptides (%)[c]
myoglobin	51	Gly-OiBu	0.7	225	8.5	ELGFQG (36%)[d] ELGFEG (18%)[d] YKELGFQG (17%)[d]
	64	Val-NH$_2$	4.8	5.8	39	ELGFQGV-NH$_2$ (44%)[e] YKELGFQGV-NH$_2$ (33%)[e]
chymo-trypsinogen	35	Gly-OiBu	0.7	225	4.5	VTALVNWVQQTLAAN (51%)[d]
	27	Val-NH$_2$	5.4	6.6	31	VTALVNWVQQTLAANV-NH$_2$ (51%)[e]
bovine serum albumin	13	Gly-OiBu	0.7	225	14	LVVSTQTALA (43%)[d]
lysyl-tRNA synthetase[f]	10	Val-NH$_2$				VILFPAMRFPVKV-NH$_2$ (43%)[d,e]

[a]Proteins (5 mg/mL) were labeled for 4 min with the indicated nucleophile in the presence of an optimal concentration of CPD-Y, as determined in Table IV. [b]An aliquot of the protein was precipitated with 5% trichloroacetic acid and, after filtration, analyzed by liquid scintillation counting. The yield of labeling, r, was the ratio of the specific radioactivity of the labeled protein to that of the nucleophile. [c]Percentages of each peptide species characterized over the total material eluted in the RPLC peptide separation, as determined from liquid scintillation counting. [d]Sequence and/or [e]Mass spectrometric analysis were used for peptide characterization. [f]Labeled as described in the text.

intramolecular bridge, which could prevent the interaction with CPD-Y.

C. Control Experiments. Various control experiments were conducted at pH 10, using the seven assayed proteins, to confirm the requirement of both an esterified protein as substrate and a substituted amino acid as nucleophile.

The first control experiment involved esterified proteins plus the [³H]Gly nucleophile (30 μM). The other controls involved underivatized protein substrates plus [³H]Val-NH$_2$ (10 μM) or [³H]Gly-OiBu (30 μM) as nucleophile. The example of chymotrypsinogen A is shown in Figure 3, where the control experiments are compared to the labeling reactions. In the control experiments, the scattering of the data is consistent with the variance of the experimental background. Similar negative results are obtained with all other proteins, proving that both a substituted amino acid as nucleophile and an esterified substrate are required to observe significant tritium incorporation.

D. Determination of Transpeptidation Constants for Val-NH$_2$ and Proteins. The transpeptidation constants, K, for Val-NH$_2$ and two esterified proteins, chymotrypsinogen and myoglobin, were determined at pH 10 by varying the concentration of the nucleophile in the incubation mixture. The initial fraction of transpeptidation, R_0, which is not directly measurable, is obtained by fitting the data to eq 7, as described in the Experimental Section. In contrast to the case of peptidic substrates when $1/R_0$ was plotted as a function of $1/[Val-NH_2]$ in order to determine the K constant, the straight lines obtained intercepted the vertical axis at a $1/R_0$ value different from 1. In the case of chymotrypsinogen (respectively, myoglobin), a linear regression gave 1 ± 0.1 mM (respectively, 1.8 ± 0.3 mM) for K and 2.5 ± 0.2 (respectively, 1.8 ± 0.2) for the $1/R_0$ intercept. This means that, when the concentration of Val-NH$_2$ indefinitely increases, the fraction of transpeptidation reaches plateau values of 40% and 55% for chymotrypsinogen and myoglobin, respectively. The K constant, in these cases, is the concentration of the nucleophile for which the fraction equals half of this plateau value. Its values are lower than those previously determined (Table IV) by using a single nucleophile concentration ($K = 1.3 \pm 0.4$ mM and 3 ± 0.4 mM for chymotrypsinogen and myoglobin, respectively). This difference results from the fact that the previous estimation assumed a plateau value of 100% for the fraction of transpeptidation at saturating nucleophile concentrations.

It is unlikely that the particular behavior of these protein substrates, for which the labeling yield seems limited to less than 50%, would result from their particular carboxyl-terminal sequence. Indeed, such behavior was not observed when short peptides with various sequences were studied. Several factors may actually interfere with the determination of the labeling yield, such as loss of material during the ether precipitation of the protein after esterification or during precipitation with TCA, or incomplete initial esterification of the protein. Alternatively, since the protein substrate is denatured in the assay by the presence of 0.4% SDS,

only a fraction of the molecules may be available for reaction with CPD-Y.

E. Characterization of the Proteins Labeled with Val-NH$_2$. Preparative-scale labeling experiments were performed at pH 10 using 1 mg of the following proteins: myoglobin, chymotrypsinogen A, and *E. coli* lysyl-tRNA synthetase (the *lysS* gene product[17]). Myoglobin and chymotrypsinogen were labeled in the presence of 5 mM Val-NH$_2$. Since this concentration is 3–5 times higher than the corresponding K values measured above, yields of labeling approaching the maximal yields could be predicted. Indeed, yields of incorporation of [³H] into myoglobin and chymotrypsinogen were 39% and 31%, respectively. These yields are consistent with the data presented in the preceding paragraph. The labeled proteins were submitted to trypsinolysis, and the resulting peptides were purified by RPLC. Radioactive peptides corresponded to the expected carboxyl-terminal peptides elongated by one Val-NH$_2$ residue, as shown by sequence or mass spectrometric analysis (Table V).

To label esterified lysyl-tRNA synthetase, 1 nmol of the protein was incubated in the presence of 40 μM [³H]Val-NH$_2$ (0.25 Ci/mmol), while 7 nmol was incubated with 50 mM cold Val-NH$_2$. After mixing, the sample was further treated as described in the Experimental Section. After digestion with *S. aureus* protease, a radioactive peptide (0.5 nmol) was analyzed by sequencing and mass spectrometry. The mass of 1369.0 ± 0.4, measured for the (M + H⁺) ion, corresponds to the expected carboxyl-terminal peptide labeled with Val-NH$_2$ (VILF-PAMRPVKV-NH$_2$, calculated monoisotopic mass $M = 1367.9$), which confirms the carboxyl-terminal sequence deduced from the DNA sequence of the *lysS* gene.[17]

F. Characterization of the Proteins Labeled with Gly-OiBu. The tritiated Gly-OiBu nucleophile was employed at a concentration of 0.7 mM, a value 12–25 times lower than estimated K constants, depending on the protein substrate (Table IV). Therefore, the yield of labeling could be predicted to be on the order of 10% in the most favorable case. In spite of such a low yield, the monitoring of the radioactivity was expected to lead to the isolation of significant amounts of the carboxyl-terminal peptide, labeled or not, of the proteins. Indeed, an additional Gly residue little modifies the overall hydrophobicity of a peptide[28] and, thus, is not expected to significantly affect its RPLC retention time.

Bovine serum albumin, myoglobin, and chymotrypsinogen A were labeled with measured yields of 14%, 8.5%, and 4.5%, respectively. After trypsinolysis and peptide purification, the radioactively-traced peptides were sequenced (Table V). The sequences coincided with those of the carboxyl-terminal ends of these proteins. The tritiated carboxyl-terminal glycine, present at trace concentrations, was only poorly detected in the ultimate sequencing cycle. As discussed above, this reflects the fact that more than

(28) Guo, D.; Mant, C. T.; Taneja, A. K.; Parker, J. M. R.; Hodges, R. S. *J. Chromatogr.* **1986**, *359*, 499–518.

90% of the isolated product does not carry the additional [3H]Gly residue.

When myoglobin was labeled with Gly-OiBu, a peptide ELG-FEG was observed in addition to the normal carboxyl-terminal peptide ELGFQG. The presence of this peptide could be explained by a partial methanolysis of the penultimate Gln residue to a Glu one during the esterification of the protein. In a previous study,[27] this type of modification was shown to occur to a very slight extent under similar conditions of esterification. Many of the peptides characterized here contained Gln or Asn residues. However, in only one case was an amide group found to be partly converted to a carboxyl group. This may indicate that the methanolysis of amide groups is limited to some particular sites, depending on their amino acid sequence environment.

G. Investigation toward the Possible Formation of Isoaspartyl Linkages. Another chemical modification of polypeptides, the conversion of Asp residues to isoaspartyl residues, could be suspected to take place to some extent upon the base treatment further to transpeptidation.[29] If a quantitative isomerization of an Asp residue lying near the C-terminal extremity occurs, the subsequent Edman degradation process would be blocked by the isoaspartyl linkage. Since none of the proteins assayed in this study had an Asp residue in the vicinity of its C-terminus, we investigated the possible occurrence of isoaspartylation reactions using three short peptides having the sequences WAGGDASGE, MGWMDF, and KKDSGPY.

The peptides were submitted to esterification followed by base treatment under the conditions employed for proteins. Automated Edman degradations were then performed comparatively with the treated and nontreated peptides. MGWMDF was found not to have been altered by the treatment. In contrast, WAGGDASGE and KKDSGPY had undergone 60% and 84% isomerization, respectively, as indicated by the reduction in the yields of sequencing, from the cycle where Asp was released, as compared to the yields for the intact peptides. In all cases, however, a significant part of the peptide molecule remains structurally unaltered, still allowing sequence determination. In the worst case, sequencing of a fully altered carboxyl-terminal peptide could be achieved by the means of mass spectrometry.

Discussion

Choice of Gly-OiBu or Val-NH₂ as Nucleophile. Two radiolabeled nucleophiles, Gly-OiBu and Val-NH₂, were used to develop the labeling method. The most interesting point is the very broad character of the transpeptidation with these compounds: all of the studied proteins can be significantly labeled using either nucleophile. This represents a number of very different carboxyl-terminal sequences.

If one compares the efficiencies of Gly-OiBu and Val-NH₂ as nucleophiles, it appears that transpeptidation works better with Val-NH₂ than with Gly-OiBu. The transpeptidation constants of the latter are higher by a factor of 2–10 for oligopeptidic substrates and by a factor of 2–5 for protein substrates, except in the case of BSA (Table IV). However, some factors argue in favor of the choice of Gly-OiBu in the design of the most convenient labeling procedure: (i) Proteins traced with [3H]Gly-OiBu, even in very low yields, can be utilized for the isolation of their carboxyl-terminal end, as already discussed. Indeed, an additional

Gly residue was shown not to significantly affect the chromatographic retention of peptides. In contrast, an additional Val-NH₂ rendered peptides more hydrophobic, resulting in the fact that the labeled peptide no longer coeluted with the unlabeled one. This imposes the use of higher nucleophile concentrations in order to improve the yield of labeling. (ii) The synthesis of labeled Gly-OiBu from labeled Gly is much easier to perform than that of Val-NH₂ from Val, and the product in the former case needs no purification. The yield of synthesis is close to 100%, while it is only 33% in the case of Val-NH₂.

For these reasons, the Gly-OiBu nucleophile may be useful for a rapid and economical isolation of the carboxyl terminus of a protein in order to check its sequence. In contrast, the choice of Val-NH₂, used at high concentrations, is more suitable for the reliable characterization of proteins of unknown sequence. Indeed, the risk of mistaking the carboxyl-terminal peptide cannot be excluded with the strategy using Gly-OiBu. It will depend on the resolution afforded by the chromatographic separation system.

Prospects. Although not single-step, the procedure proposed here is simple to perform. The optimal CPD-Y concentration could be determined using a few tens of micrograms of protein and transposed to milligram amounts with good reproducibility, regardless of the nucleophile concentration in the range 0–50 mM.

As shown by Carles et al.,[15] one possible purpose of labeling proteins at their carboxyl terminus is to facilitate the isolation of a full set of polypeptides of various lengths suitable for sequencing. These authors demonstrated that a protein labeled at its carboxyl-terminal extremity, with a specific radioactivity of 1.5×10^6 dpm/nmol, could be fully sequenced using a limited number of Edman degradation cycles, after the isolation of fragments by gel electrophoresis. The lack of a general method to label proteins at their carboxyl-terminal extremity was, however, a serious limitation of this procedure. It can be calculated that, with the present method using 1 mCi of [3H]Val-NH₂, 50 nmol of myoglobin would be labeled with comparable specific radioactivity. For less favorable proteins, however, the specific radioactivity resulting from transpeptidation would be smaller, by a factor of 10–100, than that obtained by Carles et al. in the labeling of casein. This may render problematic the detection of protein fragments by fluorography. However, this limitation may be overcome by the use of radiolabeled nucleophiles allowing a more sensitive detection, such as ethyl ester derivatives of [35S]Cys or [125I]iodo-Tyr.

The described method, however, has already demonstrated its usefulness in solving carboxyl-terminal sequences of proteins. Very little radioactivity is consumed (less than 1 mCi of both [3H]Val and [3H]Gly in the overall present study), and the reduction to the nanomole range of the amount of protein required is currently under investigation.

Supplementary Material Available: Experimental procedures for the esterification of amino acids, peptides, and proteins, the amidation of [3H]Val, and the pyridylethylation of proteins and mass spectrometric data of the transpeptidation products (6 pages). Ordering information is given on any current masthead page.

(29) Geiger, T.; Clarke, S. *J. Biol. Chem.* **1987**, *262*, 785–794.