Peptide and Protein Carboxyl-terminal Labeling through Carboxypeptidase Y-catalyzed Transpeptidation*

(Received for publication, May 29, 1990)

Pierre-François Berne, Jean-Marie Schmitter‡, and Sylvain Blanquet
From the Laboratoire de Biochimie, U.R.A. 240 Centre National de la Recherche Scientifique, Ecole Polytechnique, 91128 Palaiseau Cedex, France

A survey of carboxypeptidase Y-catalyzed carboxyl-terminal modification of short peptides in the presence of various amino acids revealed that transpeptidation occurred in significant yield only when peptides containing a proline at the penultimate or antepenultimate position. For these peptides, transpeptidation was shown to occur specifically at the carboxyl side of the proline, thus suggesting a determining role of this residue for transpeptidation. Two model peptides, YPFFGPI and YPFVEPI, were studied in detail. Initial yields of transpeptidation in the presence of various nucleophiles were compared. Among natural amino acids, the highest yield was obtained with methionine, followed by other amino acids bearing hydrophobic side chains.

In order to transpose the method of transpeptidation to a protein, a variant of Escherichia coli methionyl-tRNA synthetase bearing the carboxyl-terminal Glu-Pro-Met sequence was genetically created. Under the conditions optimized for the transpeptidation of YPFFVEPI with methionine, this protein could be labeled specifically at its carboxyl-terminal end. Moreover, the parameters of the labeling reaction were in agreement with those observed in the transpeptidation of the model peptide.

Isotopically labeled polypeptides may be of considerable interest in various fields: protein purification and biochemical characterization, NMR, and physiological studies (2). A gap in the methods of investigation of biochemistry lies, however, in the difficulty of labeling polypeptides specifically at an extremity, as is possible, and very fruitful, with nucleic acids (3).

Chemical reagents are widely used to perform group-specific (4) or site-specific (5) modifications of proteins. However, in order to develop an amino- or carboxyl-terminal labeling process, the use of chemical reagents is problematic because they lack the required specificity, modifying also lysine side chains in one case, aspartic acid and glutamic acid side chains in the other case. This problem has been overcome by Jay (6) and by Jue and Doolittle (7), who designed a procedure for the specific labeling of the amino terminus of polypeptides. This method yielded some valuable information about the primary structure of the labeled protein, but other applications are limited by the fact that the conditions required by the labeling procedure lead to the reduction or abolishment of protein activity.

Therefore, an enzymatic approach, which was expected to preserve the native structure of the protein to be labeled, has been envisaged. It is based on the specificity of carboxypeptidases and on the ability of some of these enzymes to catalyze transpeptidations, i.e. to replace the carboxyl-terminal amino acid of a polypeptide with an exogenous nucleophile. Among this class of enzymes, CPD-Y, a commercially available enzyme isolated from yeast, was especially suited, because its enzymatic properties (8–12), particularly its transpeptidase activity (13–17), have been extensively studied. Available data concerning CPD-Y-catalyzed transpeptidations deal, however, essentially with small unnatural substrates (N-blocked amino acids and dipeptides, mostly with an esterified or amidated carboxyl group) and with extremely high concentrations of the nucleophile.

Recently, CPD-Y was used to perform the labeling of β-casein (18). Surprisingly, transpeptidation did not occur at the carboxyl-terminal position of the native protein, but occurred only after CPD-Y had removed the last 3 residues. This result, as well as preliminary attempts in our laboratory to transpose this method to other proteins, suggested that CPD-Y-catalyzed transpeptidation would be dependent on the carboxyl-terminal sequence of the substrate. One of our aims was therefore to elucidate the sequence determinants favoring transpeptidation. We also endeavored to design optimal conditions, particularly with respect to the nucleophile concentration, to maximize the specific radioactivity incorporated into the product and to minimize the amount of radioactive agent consumed in the labeling reaction. First, optimal conditions were defined using synthetic oligopeptides. Second, the results obtained on short natural substrates were extrapolated to a protein. However, to favor transpeptidation, the carboxyl-terminal extremity of the selected protein, Escherichia coli methionyl-tRNA synthetase, was changed by genetic modification. We show here how this remodelling of the sequence of methionyl-tRNA synthetase allowed its specific labeling at the carboxyl-terminal end.

* The abbreviations used are: CPD-Y, carboxypeptidase Y; Cbz, benzoyloxycarbonyl; RPLC, reversed phase liquid chromatography; SDS, sodium dodecyl sulfate. The residues of a substrate of CPD-Y are numbered according to the Schechter and Berger (1) notation for active site subsites of proteases; starting from the scissile bond, P,, P,, . . . are the residues in the C→N direction and P′,, is the carboxyl-terminal residue, adjacent to the bond. The substrate may thus be represented as: NH₂-P₁-P₂- . . . Pₙ-P₁-P′₁-COOH. The subsites of the active site are abbreviated correspondingly S₁, S₂, . . . Sₙ. The single letter or three letter code is used for amino acids. The NH₂ suffix indicates that an amide group is substituted for the carboxyl group.

†To whom correspondence should be addressed.

This is an Open Access article under the CC BY license.
EXPERIMENTAL PROCEDURES

Materials—Chymotrypsin and the peptides YPFPGPM (i) and YPFPGPI (ii) were obtained from Sigma (St. Louis, MO), RPPGFSPF, insulin (bovine, B chain), and glucagon (bovine) were from Serva (Heidelberg, West Germany), and YPPFVP, YVPFPFV, RPPGFSPF, YGFMGRL, and RYVHPF were from Tebu (France). The concentrations of stock solutions were determined, when possible, from their absorbance at 280 nm using the molar extinction coefficients of free Tyr (εM = 1197) and Trp (εM = 5600), otherwise by weighing 1 mg of the lyophilisate. L-[methy1-3H]Met (40 Ci/mmol) was purchased from the Commissariat a l’Energie Atomique (Saclay, France) and used for transpeptidations on peptides, or from Amersham (United Kingdom) (73 Ci/mmol) for protein labeling experiments. CPD-Y was from Boehringer (Mannheim, West Germany).

Preparation of CPD-Y Solutions—A stock solution (about 20 μM) was prepared in 10 mM K2HPO4 (pH 7.5), 1 mM EDTA, 10% v/v glycerol. Its concentration was determined by absorbance measurement at 280 nm (εn = 1.48 (19)). Enzyme prepared in this manner and stored at −20 °C was found to be fully active after 6 months when assayed using the synthetic substrate Z-Phe-Leu-Val-Asp (20).

Synthesis of Peptides—The peptides YPFPGFM (i) and YPFPGCFR (ii) were synthesized automatically by transpeptidation using the substrate YPFPGFM in buffer containing 150 mM K2HPO4 (pH 6.5) and 1 mM EDTA under the following conditions: (i) 200 μM YPFPGFM, 200 μM Met, 300 mM CPD-Y, 30-min incubation at 37 °C; (ii) 100 μM YPFPGFM, 200 mM Arg, 110 nM CPD-Y, 40-min incubation at 37 °C. Reaction products, containing 80% YPFPGPM (i) or YPFPGPI (ii), were separated by RPLC (see below).

Transpeptidation Reactions—Unless otherwise specified, all reactions were conducted at 37 °C in buffer containing 150 mM K2HPO4 and 1 mM EDTA. CPD-Y from the stock solution was diluted on ice to 1–5 μM in the incubation buffer immediately prior to use. Reactions were started by the addition of CPD-Y (20–300 nM) to the substrate (20 μM) and nucleophile solution. Aliquots were withdrawn at different times, the reaction quenched by the addition of trifluoroacetic acid (5% v/v final concentration), and the sample analyzed by RPLC on a Merck Superspher C18 column, 42 °C, 0.9 ml/min) with a 15-min gradient (18–40%) of acetonitrile in 0.1% trifluoroacetic acid. Peptides were detected either by fluorescence (λex = 272 nm, λem = 303 nm) or by absorbance at 215 nm, and the percentage of each peptide present was quantified by peak area integration. Initial rates of appearance of the products were calculated using at least three points.

A different separation gradient was used to follow the incorporation of [3H]Met into peptides, after the elimination of free Met (15 min at 8% acetonitrile), peptides were eluted sharply by a 1-min rise to 80% acetonitrile. Corresponding fractions were counted in a Beckman LS 1801 after the addition of PicoFlour (Packard).

Amino Acid Analysis—Peptides were submitted to acid hydrolysis using the method of Kaur and Scheffler (21). Amino acids were quantified after precolumn derivatization with phenylisothiocyanate and separation by RPLC (Merck Superspher C18 column, acetonitrile gradient in 50 mM sodium acetate).

19552 Carboxyl-terminal Labeling of Peptide and Protein

The following scheme describes the situation,

\[ E + P - X \xrightarrow{k_1} E - P \xrightarrow{k_2} E + P \]

where both \( \lambda \) and \( \lambda' \) depend on [X]* and exhibit the same relation as \( v \) and \( v' \). Equation 1,

\[ \frac{\lambda}{\lambda + \lambda'} = \frac{[X]^{*}}{K + [X]^{*}} \]

Variation of \( r \) as a function of time yields a bell-shaped curve; the maximal value (\( r_{\text{max}} \)) is given by the following equation.

\[ r_{\text{max}} = \frac{K}{[E] + [X]^{*}} - \frac{K}{K + [X]^{*}} \]

The simplest mechanistic scheme describing this competition is formally equivalent to a scheme including exclusive acceptor sites for water and the nucleophile (24). The initial yield of transpeptidation, \( v'/v + v' \), can be shown to depend on the concentration \([Z]\) of the nucleophile through Equation 1.

\[ \frac{v'}{v + v'} = \frac{[Z]}{K + [Z]} \]

where the constant \( K \) depends on the elementary rate constants but not on the substrate concentration (in the model of Scheme II, \( K = k/k' \)). \( K \) represents the concentration of \( Z \) for which an initial yield of 50% is obtained. Therefore, the determination of this parameter, called the transpeptidation constant, is an accurate way to evaluate the efficiency of transpeptidation in the presence of various nucleophiles and under various conditions. For this purpose, \( v'/v + v' \) was plotted as a function of \([1/Z]\). Experimental diagrams, including at least five points, were well correlated with straight lines. Linear regressions gave an intercept value of 1 and the value of \( K \) as the slope.
Carboxyl-terminal Labeling of Peptide and Protein

If we now suppose that the total rate of processing of the substrate $P_X$, represented by the $\lambda + \lambda'$ constant, is independent of the nucleophile concentration $[X^\star]$, we obtain the following expression for $r$:

$$r = \frac{K}{K + [X^\star]} = e^{-\alpha t}$$  \hspace{5mm} (6)

where $\alpha$ is a time constant directly proportional to $[E]_i$.

The specific radioactivity, $s$, of the substrate at time $t = [P-X^\star]/[P-X]$ and that, $\sigma$, of the nucleophile supposed constant ($[P-X] \ll [X^\star]$), the isotopic dilution with released $X$ is negligible) are related as follows.

$$s = \sigma = \frac{K}{K + [X]} = e^{\alpha t}$$  \hspace{5mm} (7)

Another important feature is the homogeneity, $h$, of the labeled product at time $t$, i.e. the ratio of the concentration of $P-X$ (including $P-X^\star$) to that of the total protein in the assay.

$$h = \frac{K}{K + [X]}$$  \hspace{5mm} (8)

**Modification of the Gene of Methionyl-tRNA Synthetase—Oligonucleotide Preparation and Site-directed Mutagenesis were performed as described previously (25).** The initial vector employed, pBSmetG, resulted from the insertion of a 2316-base pair insert containing the metG gene, which encodes E. coli methionyl-tRNA synthetase (26), between the SacI and KpnI sites of the polylinker of pBluescript KS(+) (Stratagene, San Diego, CA). The PstI-BamHI fragment (439 base pairs) of metG was inserted in the M13mp19 phage and single-stranded DNA was prepared to serve as template for mutagenesis. The synthetic oligonucleotide used, GCGGCCGCTTTTACTT, caused two single point mutations and resulted in a new site at the position corresponding to Ala-553 in the amino acid sequence. The PstI-BamHI fragment of a clone containing the desired mutations was then transferred into pBSmetG. It was now possible to insert, through the unique NotI site of the obtained vector, any sequence followed by a stop codon, generating a protein with the desired carboxy terminus. Variants obtained in this manner would share 553 amino-terminal residues in common with native E. coli methionyl-tRNA synthetase and were expected to be fully active; it has been shown (25) that the genetically truncated species M547, consisting of the 547 amino-terminal residues of native enzyme, retained full activity. The chosen carboxy-terminal sequence (oligonucleotides GGGCGACCGATGTGA and GGGCCGCGCTTTTACTT, leading to the plasmid pBSM556EPM, encoding a 566-residue variant (M556EPM) of methionyl-tRNA synthetase—was then inserted through the unique NotI site of the obtained clone. A second construct, pBSM565EPM, was finally inserted into the NotI site using a synthetic oligonucleotide, CCGGCGCG-CCGCGCGCCGCTTTTACTT, giving rise to a new site at the position corresponding to Ala-553 in the amino acid sequence.

**Purification of the Variant Methionyl-tRNA Synthetase—**The level of expression of M556EPM was measured in crude extracts by assays of aminoacylation of tRNAMet in total tRNA (25). Bacteria cultures of the strain XLI-Blue (Stratagene, San Diego, CA) carrying the plasmid pBSM556EPM over-produced methionyl-tRNA synthetase activity 80-fold compared with the same strain containing pBluescript KS(+). Protein purification, conducted as described (25), produced from 2 liters of culture 20 mg of M556EPM, whose specific activity, as measured by the ATP-PP, isotopic exchange reaction (27), was 60% that of pure M547 protein. An additional purification step involved hydrophobic interaction chromatography (Baker HIC-prolyl, 15-μm particles, 5 × 100-mm column). The column was equilibrated in buffer containing 50 mM K-HPO4 (pH 7), 1 mM 2-mercaptoethanol, and 1.7 mM ammonium sulfate, and the protein (5 mg in 10 ml of the same buffer containing 1.2 mM ammonium sulfate) was injected. A 90-min decreasing gradient of ammonium sulfate down to 0% in the same buffer was then performed at a flow rate of 0.8 ml/min. Methionyl-tRNA synthetase eluted at approximately 0.5 mM ammonium sulfate (19553). Protein purification, conducted as described (25), produced from 2 liters of culture 20 mg of M556EPM, whose specific activity, as measured by the ATP-PP, isotopic exchange reaction (27), was 60% that of pure M547 protein. An additional purification step involved hydrophobic interaction chromatography (Baker HIC-prolyl, 15-μm particles, 5 × 100-mm column). The column was equilibrated in buffer containing 50 mM K-HPO4 (pH 7), 1 mM 2-mercaptoethanol, and 1.7 mM ammonium sulfate, and the protein (5 mg in 10 ml of the same buffer containing 1.2 mM ammonium sulfate) was injected. A 90-min decreasing gradient of ammonium sulfate down to 0% in the same buffer was then performed at a flow rate of 0.8 ml/min. Methionyl-tRNA synthetase eluted at approximately 0.5 mM ammonium sulfate (19553). Protein purification, conducted as described (25), produced from 2 liters of culture 20 mg of M556EPM, whose specific activity, as measured by the ATP-PP, isotopic exchange reaction (27), was 60% that of pure M547 protein. An additional purification step involved hydrophobic interaction chromatography (Baker HIC-prolyl, 15-μm particles, 5 × 100-mm column). The column was equilibrated in buffer containing 50 mM K-HPO4 (pH 7), 1 mM 2-mercaptoethanol, and 1.7 mM ammonium sulfate, and the protein (5 mg in 10 ml of the same buffer containing 1.2 mM ammonium sulfate) was injected. A 90-min decreasing gradient of ammonium sulfate down to 0% in the same buffer was then performed at a flow rate of 0.8 ml/min.
two products varied with the concentration of Z. Amino acid composition confirmed that the transpeptidation product was YPFPGPZ. The peptide YPFPGPI was chosen as model substrate for further studies for two main reasons; (i) it is readily subject to transpeptidation (without prior hydrolysis) using relatively low (1 mM) concentrations of various nucleophiles, and (ii) the substrate and all the expected products carry an amino-terminal Tyr as common chromophore, thus facilitating the quantification of the reactions by absorbance or fluorescence detection. A peptide of close sequence, YPFVEPI, was also selected as a model substrate for similar reasons.

Test of Nucleophiles—The incorporation of the 20 common amino acids (except Ile), three amino acid amides, and norleucine (Nle) by transpeptidation with CPD-Y was tested on the substrate YPFPGPI. The occurrence of transpeptidation was indicated by the appearance of a peptide with a chromatographic retention time different from that of the normal hydrolysis products. In the case of Ala, Val, Val-NH₂, Met, Leu, Nle, Phe, Phe-NH₂, Tyr, Thr, and Arg, the presumed transpeptidation product was submitted to amino acid analysis, verifying unambiguously the composition Pro-Tyr-Phe₁ Gln, Z₁, where Z was the nucleophile involved in the assay. In the presence of Trp, a peptide appeared which contained a Trp residue, as indicated by its fluorimetric signal (λₑₓ = 288 nm, λₑₓₛ = 348 nm), and was thus assigned to the transpeptidation product. In the cases of Cys, Gin, Lys, Asn, Ser, and His, plasma desorption mass spectrometry analysis confirmed that the new peak which appeared was associated to the expected transpeptidation product. No new peptide could be evidenced in the remaining cases (Gly, Gly-NH₂, Asp, Glu, Pro). However, in these cases, the transpeptidation product may have coeluted with one of the products obtained in the absence of added nucleophile. To exclude this possibility, the peaks assigned to YPFPGPI, YPFPGP, and YPFPG were further analyzed by mass spectrometry. No trace of YPFPGPZ could be detected in these cases.

Fig. 1 shows the initial transpeptidation yields observed at 37 °C and pH 6.5 for each studied nucleophile (2 or 20 mM). Higher yields are obtained with hydrophobic amino acids, the optimal side chains being those of Nle and Met. The influence of using a nucleophile whose carboxyl group is substituted for an amide group is not clear; it improves the yield obtained with Phe, but not with Val. Several transpeptidation constants (K) were determined, using either YPFPGPI or YPFVEPI as substrate (Table I). It is noteworthy that transpeptidation occurs more readily with YPFVEPI than with YPFPGP; the transpeptidation constants with all nucleophiles are systematically 2- to 5-fold lower with YPFVEPI.

Effects of pH and Temperature—The transpeptidation constants K for the nucleophile Met, measured with YPFPGPI, at pH values ranging from 4.5 to 8.5, are plotted in Fig. 2. The K value shows little variation in the pH range 6.5–8.5, but is slightly increased under acidic conditions (pH 4.5–5.5). The same effect on K is obtained with Val as nucleophile and the peptide YPFVEPI as substrate (Table II). However, it must be pointed out that, in both cases, the reaction rate decreases considerably with increasing pH, limiting investigation to the pH 4.5–8.5 range for the peptide YPFPGPI and to the pH 4.5–6.5 range for the more anionic substrate YPFVEPI. Contrary to the cases of Met and Val, the transpeptidation constant using Val-NH₂ (YPFVEPI) as substrate shows a sharp decrease when the pH is raised from 6.5 to 8.5. This feature, which is consistent with observations by others (17), will be discussed later.

Finally, the favorable effect of lowering the temperature (pH 6.5, YPFVEPI as substrate) is illustrated by Fig. 3. Using Met as nucleophile, the K constant decreases 2-fold upon lowering the temperature of the assay from 50 to 4 °C.

Influence of the Carboxyl-terminal Residue on Transpeptidation—The influence of the amino acid occupying the P₁ position (at the carboxyl side of the scissile bond, i.e. the carboxyl-terminal position) was investigated by transpeptidation on substrates differing from YPFPGPI only by the carboxyl-terminal residue: YPFPGPM and YPFPGPR. Transpeptidation of YPFPGPR with 1 mM Met was evidenced by following the appearance of the two products YPFPGP and YPFPG (Fig. 4). From this experiment, an initial yield of transpeptidation of 50% can be deduced, indicating a transpeptidation constant K very similar to that obtained with YPFPGPI as substrate (Table I). Using the nucleophile Phe, the transpeptidation constant with YPFPGPM (Table I) is identical to that obtained with YPFPGP. These results were expected, as the competition between water and the nucleophile requires the prior dissociation.
tion of the cleaved amino acid from the acyl-enzyme; therefore, the $K$ value, which reflects this competition, should not depend on the nature of the P residue.

It is noteworthy that, although the $K$ constants for the Met nucleophile using the substrates YPFPGPR and YPFPGPI are almost identical, the reached maximal percentages of the YPFPGPM product are different. The maximal percentages observed are 2.5 and 25% with the substrates YPFPGPR and YPFPGPI, respectively. This discrepancy is due to the fact that the carboxyl-terminal P-R bond in YPFPGPR is hydrolyzed 30-fold more slowly than the P-I bond in YPFPGPI. Consequently, although the YPFPGP product is formed in both cases at the same velocity from the acyl-enzyme, it does not accumulate in the assay with YPFPGPR, because it is further hydrolyzed by CPD-Y much more quickly than is the initial substrate YPFPGP.

**Influence of the Carboxyl-terminal Sequence on Transpeptidation**—Further investigation of the influence of the carboxyl-terminal sequence on transpeptidation was conducted on ten peptides, which were systematically incubated under identical conditions in the presence of 1 mM $[^3H]$Met. Incorporation of $[^3H]$Met into the various substrates was followed as a function of time (Fig. 5A). The kinetics observed with YPFPGPI and YPFVEPI are consistent with the values of the transpeptidation constant given in Table I. Various incorporation optima were observed at various times (Fig. 5B) depending on the substrate. The best optimum was obtained with YPFVEPI; 43% of this peptide incorporated a Met. Comparable yields were obtained with the peptides having a Pro as P residue: RPPGFSPF (33%), RVYIHPF (32%), YPFPGPI (24%), RPPGFSPFR (25%), and YPFVEPI (30%). On the contrary, transpeptidation optimum was very low when glucagon (4.2%), YGGFMRGL (2%), insulin (1.8%), and LWMRFA (1.0%) were used.

**Identification of the Transpeptidation Position in the Glucagon Substrate**—Among the peptides tested for the incorporation of $[^3H]$Met, glucagon was the only one which did not contain a Pro residue and nevertheless showed a significant yield of incorporation (Fig. 5B). In order to determine the position in the sequence where transpeptidation occurred, glucagon was submitted to transpeptidation in the presence of $[^3H]$Met under the conditions described above. The reaction products at various times were analyzed by RPLC (11–35% acetonitrile gradient, detection at 280 nm), and the resulting fractions were counted by liquid scintillation. The radioactivity was shown to be associated with a single absorbance peak which was not observed in a control experiment performed in the absence of Met. The corresponding peptide species (1.5 nmol) was analyzed by plasma desorption mass spectrometry. The measured mass (2592.1 ± 1 for the (M + H)$^+$ ion species) corresponded to a 22-residue polypeptide, consisting in the l-21 fragment of glucagon, to which a carboxyl-terminal Met was added (calculated mass $M = 2592.1$). The resulting carboxyl-terminal sequence of this peptide was thus AQDMEP.

Two unlabelled species isolated by RPLC were characterized by mass spectrometry: (I) [M + H]$^+$ = 2462 and 2609. They corresponded to the 1-21 (-AQD) and 1-22 (-AQDF) fragments of glucagon, respectively. Noteworthy, no intermediate species longer than 22 residues could be detected on the chromatograms, which indicates that the residues 23–28, which are all hydrophobic, were cleaved very quickly by CPD-Y, as compared with the carboxyl-terminal Thr-29 residue.

The 1-22 (-AQDF) fragment was prepared by a 30-min incubation of glucagon in the presence of 230 nM CPD-Y (37 °C, pH 6.5) and was purified by RPLC. It was used as substrate to further determine the parameters of transpeptidation with Met. Contrary to the cases of all other substrates, when $v + v'$/$v'$ was plotted as a function of $1/[\text{Met}]$ in order to determine the constant $K$, the intercept value of the straight line obtained was different from 1. The values given by linear regression were 1.3 ± 0.1 nM for K and 6 ± 0.3 for the intercept. This means that when the concentration of Met is increased, the initial yield of transpeptidation $v'/v'$ reaches a limiting value of 0.17 ± 0.01 instead of 1. The $K$ constant, in this case, is the concentration of the nucleophile for which the yield equals half of this limiting value.

The fact that the limiting value of the initial yield of transpeptidation is lower than 1 reveals that the acyl-enzyme involving the 1-21 fragment of glucagon can be hydrolyzed even though the Met nucleophile is bound to the S$'$ site of CPD-Y. This possibility, which was excluded from the mechanistic scheme (II), should thus be considered in some cases, depending on the substrate.

**Labeling of a Protein**

Having established the determining role of a Pro residue in the carboxyl-terminal sequence, the use of mutagenesis ap-

---

**Table II**

**Effect of pH on the transpeptidation constant $K$ (mM)**

| Substrate   | Nucleophile | pH 4.5 | pH 5.5 | pH 6.5 | pH 7.5 | pH 8.5 |
|-------------|-------------|--------|--------|--------|--------|--------|
| YPFPGPI     | Met         | 4.02 ± 0.08 | 1.21 ± 0.02 | 0.97 ± 0.02 | 1.00 ± 0.03 | 1.2 ± 0.1 |
|             | Val         | ND     | ND     | 3.1 ± 0.1   | 3.16 ± 0.05 | 3.5 ± 0.2   |
|             | Val-NH$_2$ | ND     | ND     | 17.7 ± 0.4   | 2.78 ± 0.07 | 1.46 ± 0.07   |
| YPFVEPI     | Met         | 2.34 ± 0.07 | 0.56 ± 0.02 | 0.31 ± 0.01 | ND     | ND     |
Carboxyl-terminal Labeling of Peptide and Protein

Peared to be the simplest way to allow the labeling of a protein, through the introduction of a favorable, and otherwise lacking, carboxyl-terminal sequence. Methionyl-tRNA synthetase (M547 variant), which has a carboxyl-terminal sequence LVEASK, could not be readily labeled by CPD-Y under the optimal conditions derived from the study of model peptide substrates. A genetic modification of this protein generated the M556EPM variant, whose carboxyl-terminus contained the Glu-Pro-Met sequence. This sequence was chosen because the P₁ and P₂ residues of the modified protein would be identical to those of the peptide YPFVEPI, the best identified substrate for transpeptidation. In addition, as the labeling would be performed with Met, one of the nucleophiles giving the highest yields of transpeptidation, the presence of a carboxyl-terminal Met would increase the carboxyl-terminal homogeneity of the labeled protein.

Incorporation of [³H]Met in the Variant M556EPM Protein—The kinetics of transpeptidation of M547 and M556EPM in the presence of 2 µM [³H]Met (73 Ci/mmol) were followed by trichloroacetic acid precipitation (Fig. 6B). No incorporation of radioactivity was observed with the M547 variant of methionyl-tRNA synthetase. In contrast, with the M556EPM substrate, trichloroacetic acid-precipitable radioactivity increased as a function of time to a maximal value. The time required to reach this rₘₐₓ value can be shown to depend neither on [X] nor on the K constant, provided [X] ≪ K. From this time, it could thus be deduced that the rate of hydrolysis by CPD-Y was about 40-fold lower for the protein substrate M556EPM than for the peptide YPFVEPI (both substrates were compared at a 20 µM concentration). This may be explained by a reduced accessibility to CPD-Y of the carboxyl-terminus of a folded protein as compared with that of a short peptide.

The presence of a denaturing agent (0.05% SDS) was shown to increase the rate of hydrolysis of the protein by a factor of at least 5, probably by partially unfolding the substrate. In the case of the protein substrate M547, it was even shown, by analysis of the released amino acids, that the presence of 0.05% SDS was necessary for any hydrolysis to occur. Incorporation of [³H]Met was thus tested as above, with both protein substrates, but now in the presence of 0.05% SDS. The same optimum level was obtained with the M556EPM variant, whereas again no significant incorporation could be evidenced with M547 (not shown).

In order to determine the K constant for the transpeptidation of M556EPM with Met, concentrations of this nucleophile in the range 0.1–1 mM were employed (Fig. 6A). Initial rates of incorporation v' were calculated and a transpeptidation constant of 0.28 mM was deduced from a linear regression of the double reciprocal plot of 1/v' as a function of 1/[Met].
This value is close to the corresponding value for the peptide YPFVEPI (0.31 mM, Table II). It should, however, be kept in mind that the above determination of $K$ with M556EPM as substrate is based on the hypothesis that $\nu + \nu'$ is independent of Met concentration. This condition was found to be true using YPFVEPI as substrate under similar conditions, at least up to 3 mM Met (not shown). From a mechanistic point of view (Scheme II), this could reflect that either (i) acylation is the limiting step of the reaction or (ii) the protein substrate concentration is far below its $K_m$ value. With the M556EPM substrate (which carries the same $P_1$ and $P_2$ residues as YPFVEPI), criterion (i) is expected to be conserved and the value of $K_m$ on which depends (ii), is likely to increase, because of the reduced accessibility of the carboxyl terminus to CPD-Y.

Preparative Labeling of the M556EPM Protein—A large scale labeling was conducted on 2 mg of M556EPM under the following conditions: M556EPM was 40 $\mu$M, CPD-Y 1.5 $\mu$M, and [3H]Met (73 Ci/mol) 70 $\mu$M. Aliquots (30 pmol of M556EPM) were trichloroacetic acid-precipitated (see “Experimental Procedures”) during the course of the experiment in order to follow the reaction. After 2 h, the maximal level of incorporation (210,000 dpm) was reached and the reaction was stopped by the addition of diisopropylfluorophosphate (0.3 mM final concentration) followed by a 15-min incubation at room temperature. No significant loss of activity of the synthetase occurred during this step, as determined by both ATP-PPi, isotopic exchange (27) and tRNA aminoacylation (29) assays. The labeled protein was then added to the same quantity of unlabeled protein (2 mg) and dialyzed for 4 days at 4 °C against buffer containing 10 mM K$_2$HPO$_4$ (pH 6.5) and 10 mM 2-mercaptoethanol. It was determined by amino acid analysis that the amount of free Met remaining after dialysis was negligible. Direct counting of the protein then allowed the measurement of the final specific radioactivity: 3,700 dpm/pmol. Thus, the reached yield of labeling, $r$ (see “Experimental Procedures”), before isotopic dilution of the protein, was 4.6%. This yield is in reasonable agreement with the 6.7% value predicted from the experimental conditions and the model peptide YPFVEPI. In preliminary experiments (not shown), smaller amounts of protein (<10 $\mu$M in incubation) were labeled with a much better correlation between the predicted and obtained yields. Thus the lower yield observed at high protein concentration may be attributed to the complex formation between methionyl-tRNA synthetase and Met ($K_{diss} = 50 \mu$M (27)), which might lower by 25% the concentration of free Met under the conditions of the preparative labeling experiment.

Reversal of the Labeling—The labeled M556EPM protein (5 $\mu$M) was incubated in buffer containing 10 mM K$_2$HPO$_4$ (pH 6.5) and 1 mM EDTA, either alone or in the presence of CPD-Y (0.35 $\mu$M), with or without SDS (0.05%). The reactions were followed using trichloroacetic acid precipitation as described under “Experimental Procedures.” The results showed that incubation of the protein with CPD-Y resulted in the removal of the radioactive label. After 2 h at 37 °C, only 8% (in the presence of 0.05% SDS) or 19% (no SDS) of the initial counts remained. This confirmed that the labeling of the protein had been caused by a CPD-Y-catalyzed phenomenon. The label of the protein incubated alone was completely stable under the same conditions, even though CPD-Y remained present at a 1/25 molar ratio to the labeled methionyl-tRNA synthetase. This demonstrated that the efficiency of the diisopropylfluorophosphate inactivation of the peptidase at the end of the preparative labeling step.

Characterization of the Radiolabeled Protein—The specificity of CPD-Y for the C-terminal extremity of polypeptide
peptides after a proteolytic digestion of the protein. A chymotrypsin digest of labeled M556EPM was analyzed by RPLC (Fig. 7). Four peaks (I-IV) represented 75% of the injected radioactive material. The remaining 25% was scattered in several small peaks, each accounting for less than 3% of the total. The kinetics of digestion by chymotrypsin was also monitored by injection of aliquots at various times (data not shown). This demonstrated that peak IV was an intermediate species, which upon further digestion gave essentially peak II. Peaks I, II, and III were carboxyl-terminal peptides, as demonstrated by sequencing. They correspond to fragments 549-556, 537-556, and 538-556, respectively.

530 540 550
LYNKIDMVQLVEALVEASKEEVKAAEPM

These results demonstrate the narrow specificity of the labeling of the protein; the label is strictly limited to the Met 556 position.

**DISCUSSION**

**Validity of the Transpeptidation Reaction Scheme—**Experimental validation of the proposed reaction model arose from the study of transpeptidation using the peptides YPPGPGI and YPFVEPI. Noteworthy, with all nucleophiles used, the initial yield tends toward 1 when the concentration of the nucleophile rises. It allows one to describe transpeptidation by the single K value and confirms the mechanistic scheme (II) (see "Experimental Procedures"), presuming that no significant hydrolysis of the acyl-enzyme occurs when the S1 site is occupied by the nucleophile. If this was not so, the maximal value of \( v / (v' + v) \) would strictly remain below 1.

According to our model, the constant of transpeptidation should not depend on the leaving amino acid. This is well established in this study; the nature of the ultimate residue of the substrate can influence only the maximum yield of transpeptidation, through the ratio between the hydrolysis rates of P-X and P-Z. Using a given nucleophile Z, the better the substrate P-X, the higher the maximal yield obtained. According to the known specificity of CPD-Y (12), this means that better yields will be obtained if X is a hydrophobic residue rather than a small or charged residue.

On the contrary, Breddam et al. (14, 17), working on N-blocked dipeptides, such as benzoyl-Ala-X, noticed that the yield of transpeptidation, using Leu-NH2 (0.25 M), Gly (1 M), or Gly-NH2 (1 M) as nucleophiles, basically depended on the leaving amino acid X. When the side chain of X was hydrophobic, experimental yields were very low. This phenomenon was explained by the incomplete dissociation of the leaving amino acid during the deacylation step and by the presence of an independent binding site for a water molecule. Scheme II seems thus to be invalid in this case, probably reflecting the different nature of the used substrate.

**Transpeptidation and Specificity of the S1 Site**—All nucleophiles tested under identical assay conditions on the substrate YPPGPGI are sorted in Fig. 1 according to decreasing efficiency for transpeptidation. It is obvious that transpeptidation works better with hydrophobic nucleophiles, especially with amino acids bearing a long aliphatic side chain and with Phe. Although the transpeptidation constant K may not be simply regarded as the dissociation constant for an acyl-enzyme:nucleophile complex, it partly reflects the interaction of the nucleophile with the S1 site of CPD-Y (the S1 site is the subsite of the active center which interacts with the carboxyl-terminal residue of the substrate). The observed preference for hydrophobic side chains is in agreement with the known specificity of this site (12).

This preference for hydrophobic side chains was not observed by Widmer et al. (16) who worked on the substrate benzoyl-Ala-O-methyl. However, these authors (16), using amino acids and amino acid amides as the nucleophile Z, compared the maximal percentages of benzoyl-Ala-Z formed rather than initial rates of product appearance. For achieving maximal yields of the transpeptidation product, they always used the highest possible concentration of the nucleophile, ranging from 0.15 to 3 M. Moreover, in their case, the substrate was an ester, and the reactions were carried out at pH 0.5.

**Comparison of Amino Acids and Amino Acid Amides as Nucleophiles—**From already mentioned studies, it appears that transpeptidations with amino acid amides differ from those carried out with amino acids; reactions with amino acid amides are generally much more efficient and depend little on the nature of the side chain (15, 16). Transpeptidations with amides, unlike those performed with amino acids, were also shown to crucially depend on pH (17) in the pH 5–8 range. The transpeptidation of Cbz-Ala-Ala with Gly or Gly-NH2 (1 M) was compared at various pH values. With Gly, the yield remains at its maximal value between pH 5 and 8. With Gly-NH2, the yield jumps from nearly 0% at pH 6 to nearly 100% at pH 8.

These differences between amino acids and amino acid amides as nucleophiles in CPD-Y-catalyzed transpeptidation remain unelucidated. It has been proposed (13, 16) that the pK\(_a\) of the α-amino group of the nucleophile, which is generally 1.5 units higher for an amino acid than for the corresponding amide, may play a role, although this property was not sufficient to explain satisfactorily the variations of the yield versus pH (17). This explanation does not, however, take into account the fact that once the nucleophile is bound to CPD-Y, its amino group is surrounded by the extremely powerful charge relay network of the proteolytic enzyme, which may considerably lower the pK\(_a\) value.

To explain the discrepancies, it must thus be imagined that, as suggested by Breddam et al. (11), amino acids and amino acid amides bind differently to the S1 site of CPD-Y. The distinct variations of transpeptidation towards pH may then reflect the ionization of several residues of the enzyme differently involved in the binding.

Our present results concerning transpeptidation of the peptide YPPFGPI with Met, Val, and Val-NH2 are consistent with the previously mentioned results. They further indicate that the variations with pH, specifically observed with amino acid amides, mostly reflect the ionization of a charged residue of the enzyme involved in the fixation of the nucleophile and thus could be expected to be similar for any substrate.

**The Importance of a Penultimate Proline in Optimizing Transpeptidation—**Discrepancies were outlined in the first paragraph of the discussion section between the mechanism of transpeptidations described in the literature and that established in this study using YFPFGPI and YPFVEPI as substrates. In addition, the smallest transpeptidation constant measured in this study (0.3 mM using Met and the peptide YPFVEPI) is 100–1000 times lower than the values extrapolated from the analysis of previous data by others. The only similar situation is that of peptides related to β-casein (18). However, as outlined already, such comparisons should be considered cautiously because of the different experimental approaches used.

More reliable information is obtained from the result of the
incorporation of tritiated Met in various peptide substrates under identical conditions. It clearly appears that these conditions, defined using YPFPGPPI, are not transposable to all other peptides with the same efficiency. Yields of transpeptidation comparable with that of YPFPGPPI are obtained only with the peptides containing a Pro residue at the penultimate or antepenultimate position. A structural explanation may account for this observation; it is known that the presence of a prolyl residue suppresses one degree of freedom of the polypeptide chain. The deformation imposed to the active center of CPD-Y by such a rigid substrate might increase the reactivity of the bound nucleophile.

The poor incorporation yield observed in insulin, which carries the carboxyl-terminal Pro-Lys-Ala sequence, does not refute the hypothesis concerning the role of proline. This case may be similar to that of YPFPGPPI, already outlined; a very slow rate of acylation of the Pro-Lys bond, similar to the Pro-Arg bond in the YPFPGPPI peptide, may lower the maximal yield of transpeptidation, even though the initial yield of transpeptidation could be comparable to that of other peptides containing a penultimate Pro.

A Pro residue is not, however, an absolute requirement for the occurrence of CPD-Y-catalyzed transpeptidation. This is supported by some literature data (1b) and further demonstrated in this study; glucagon, which does not contain a Pro, incorporates a slight but significant amount of tritiated Met under our experimental conditions. In this case, an Asp instead of a Pro residue occupies the P$_i$ position of the substrate when transpeptidation occurs. However, at saturation of the acyl-enzyme by the nucleophile, the transpeptidation/hydrolysis ratio does not exceed 0.2, whereas a complete shift toward transpeptidation is obtained with P$_i$ = Pro substrates.

**Transposition to a Protein**—It is remarkable that the transpeptidation constant for Met and the protein M556EPM is identical to that measured using Met and the model peptide YPFVFPI. Any modified protein with the Glu-Pro-Met carboxyl-terminal sequence should thus be expected to behave as the model peptide regarding transpeptidation. In particular, appropriate conditions for the labeling of such a protein with [H]$^3$Met may be chosen with the help of the theoretical calculations given under “Experimental Procedures.” Depending on the subsequent use of the labeled protein, a compromise has to be found between the consumed radioactivity, C, and the desired parameters s (specific radioactivity) and h (homogeneity) of the labeled product. General pieces of advice can be directly deduced from the equations given under “Experimental Procedures.”

(i) The concentration of the protein in the labeling incubation must be as high as possible. The only possible unfavorable effect of a high protein concentration may come from the isotopic dilution of the nucleophile by the carboxyl-terminal amino acid released during the reaction. This effect will, however, remain negligible if the protein concentration is below $K$ (300 $\mu$m). (ii) For a given amount of protein to label, the ratio s/C, which is the parameter to maximize, is independent of the nucleophile Z concentration, provided that $[Z] < K$. It decreases above $K$.

(iii) The homogeneity $h$ is improved by stopping the reaction before the optimum of $r$ is reached, rather than precisely at the optimum. An alternative method to improve $h$, to the detriment of s, is to label less protein and to further make an isotopic dilution using unlabeled protein.

**Acknowledgments**—We thank Drs. Guy Fayat, Yves Mechulam, and Thierry Meinnel for providing the plasmids and the strains and for the help in the genetic construction and in the purification of the variant methionyl-tRNA synthetase. We also thank Dr. Jean-Pierre Denhez for his assistance in mass spectrometry analysis, Dr. Codjo Hountondji for helpful discussions, and Dr. Mary Lapadat for her careful reading of the manuscript. Finally, we are grateful to Drs. Carles and Ribadeau-Dumas for very stimulating discussions at the beginning of this work.

**REFERENCES**

1. Schechter, I., and Berger, A. (1967) Biochem. Biophys. Res. Commun. 27, 157-162
2. Denegre, R. H. J. (1965) Biochin. Biophys. Acta 789, 151-166
3. Maxam, A. M., and Gilbert, W. (1980) Methods Enzymol. 65, 499-560
4. Glazer, A. N., Delange, R. J., and Sigman, D. S., (1975) in **Laboratory Techniques: Chemical Modification of Proteins** (Work, T. S., and Work, E., eds) North Holland Publishing Co., Amsterdam
5. Jakoby, W. B., and Wiieck, M. (eds) (1977) Methods Enzymol. 46, 1-617
6. Jay, D. G. (1984) J. Biol. Chem. 259, 15572-15578
7. Yue, R. A., and Doolittle, R. F. (1985) Biochemistry 24, 162-170
8. Hayashi, R., and Hata, T. (1972) Biochem. Biophys. Acta 263, 673-679
9. Hayashi, R., Moore, S., and Stein, W. H. (1973) J. Biol. Chem. 248, 2296-2302
10. Hayashi, R., Moore, S., and Stein, W. H. (1973) J. Biol. Chem. 248, 8365-8369
11. Kuhn, R. W., Walsh, K. A., and Neurath, H. (1974) Biochemistry 13, 3871-3877
12. Hayashi, R., Bai, Y., and Hata, T. (1975) J. Biochem. (Tokyo) 77, 69-79
13. Widmer, F., and Johansen, J. T. (1979) Carlsberg Res. Commun. 44, 57-66
14. Breddam, K., Widmer, F., and Johansen, J. T. (1980) Carlsberg Res. Commun. 45, 287-294
15. Breddam, K., Widmer, F., and Johansen, J. T. (1980) Carlsberg Res. Commun. 45, 501-507
16. Widmer, F., Breddam, K., and Johansen, J. T. (1981) Carlsberg Res. Commun. 47, 97-106
17. Rehfeldt, K., Widmer, F., and Johansen, J. T. (1981) Carlsberg Res. Commun. 48, 121-128
18. Carles, C., Gueguen, P., and Ribadeau-Dumas, B. (1987) FEBS Lett. 212, 163-167
19. Johansen, J. T., Breddam, K., and Ottesen, M. (1976) Carlsberg Res. Commun. 41, 1-14
20. Bai, Y., Hayashi, R., and Hata, T. (1975) J. Biochem. (Tokyo) 77, 81-88
21. Tsgita, A., and Scheffler, J. J. (1982) Eur. J. Biochem. 124, 585-588
22. Jonsson, E. B. G., Hedrin, A. B., Hakansson, P. L., Sundqvist, B. U. R., Säve, B. G. S., Nielsen, P. F., Roepstorff, P., Johansson, K. E., Kamensky, I., and Lindberg, M. S. L. (1986) Anal. Chem. 58, 1084-1087
23. Della-Negra, S., and Le Beyec, Y. (1984) Int. J. Mass Spectrom. Ion Proc. 61, 21-29
24. Seydoux, F., and Yon, J. (1967) Eur. J. Biochem. 3, 42-56
25. Mollot, P., Mechulam, Y., Le Corre, D., Blanquet, S., and Fayat, G. (1980) J. Mol. Biol. 138, 429-443
26. Dardel, F., Fayat, G., and Blanquet, S. (1984) J. Bacteriol. 160, 1115-1129
27. Blanquet, S., Fayat, G., and Waller, J.-P. (1974) Eur. J. Biochem. 44, 343-351
28. Beaucamps, C., Hountondji, C., and Schmitter, J. M. (1988) J. Chromatogr. 438, 347-357
29. Lawrence, F., Blanquet, S., Poirot, M., Robert-Gero, M., and Waller, J. P. (1973) Eur. J. Biochem. 50, 234-243