An Infrared Absorption Method to Titrate Quantitatively the Extent of Self-Association in Peptides[1]

Claudio Toniolo,* Gian Maria Bonora, Fernando Marchiori, and Gianfranco Borin

Contribution from the Biopolymer Research Centre, C.N.R., Institute of Organic Chemistry, University of Padova, 35100 Padova, Italy. *Received July 22, 1983*

Abstract: Self-aggregated species of fully protected, synthetic fragments of (perornithine) thynnin Z 1 in methylene chloride were disrupted by adding increasing amounts of either dimethyl sulfoxide or *N*,*N*-dimethylformamide. This conformational transition was monitored following the disappearance of either the amide A N-H stretching band or the amide I C=O stretching band of strongly intermolecularly hydrogen bonded molecules in the infrared absorption spectra. The effects induced by main-chain length and sequence, type of side-chain protection, and concentration were assessed. The increasing tendency to aggregate shown by these peptides is paralleled by a decrease in their solubility. This study has indicated that infrared absorption is a useful technique to titrate quantitatively the extent of self-aggregation in peptides.

Peptides with main-chain lengths ranging from 5 to 15 amino acids, depending on sequence, have been found to self-aggregate in organic solvents of moderate polarity (methylene chloride, chloroform, tetrahydrofuran, dioxane, ethyl acetate, acetonitrile, etc.).[2-9] The conformational transition from unordered, solvated species to partially or fully ordered, intermolecularly hydrogen bonded species, eventually forming a regular β-structure, is paralleled by a dramatic decrease in solubility. These phenomena strongly interfere with coupling rates during peptide synthesis. It is, therefore, of interest to establish a method to titrate quantitatively the extent of self-aggregation in peptides under the experimental conditions relevant to the coupling step.

For this study we took advantage of an ongoing project in our laboratory on the total synthesis of a fish protamine, thynnin Z 1.[10] The strategy of synthesis of this arginine-rich polypeptide, 34 amino acid residues long, is based on the preparation of the corresponding perornithine analogue, followed by side-chain amidination. Severe solubility problems were encountered during the preparation of the C-terminal fragment 24-34. The fully protected pentapeptide Z-[L-Orn(Boc)]$_2$-(L-Val)$_2$-L-Orn(Boc)-OMe (Z, benzyloxycarbonyl; Boc, *tert*-butyloxycarbonyl; OMe, methoxy), corresponding to the sequence 27-31, prepared as a key intermediate, is sparingly soluble in all solvents commonly used in peptide synthesis. In addition, the homopentapeptide Z-[L-Orn(Boc)]$_5$-OMe, corresponding to the sequence 24-28, an intermediate for an alternative route, shows a solubility that, although higher than that of Z-[L-Orn(Boc)]$_2$-(L-Val)$_2$-L-Orn-(Boc)-OMe, is not completely satisfactory.

We describe here the application of the infrared (IR) absorption technique to investigate the solvent-dependent equilibria between unaggregated and self-aggregated species of the two above-mentioned pentapeptides. In order to study this problem in more detail, we synthesized and examined also (i) the tripeptide Z(L-Val)$_2$-L-Orn(Boc)-OMe and the tetrapeptide Z-L-Orn(Boc)-(L-Val)$_2$-L-Orn(Boc)-OMe, two short sequences of Z-[L-Orn-(Boc)]$_2$(L-Val)$_2$-L-Orn(Boc)-OMe, and (ii) Z-[L-Orn(Boc)]$_3$-L-Orn (Adoc)-L-Orn(Boc)-OMe (Adoc, adamantyloxycarbonyl) and Z-L-Orn (Adoc)-[L-Orn(Boc)]$_2$-L-Orn(Adoc)-L-Orn(Boc)-OMe, two analogues of Z-[L-Orn(Boc)]$_5$-OMe with mixed side-chain protection. The present investigation was performed in CH_2Cl_2, a typical solvent for peptide synthesis, by adding increasing amounts of either dimethyl sulfoxide (Me_2SO) or *N*,*N*-dimethylformamide (DMF), known to form effective (peptide) N-H···O=S (solvent) and (peptide) N-H···O=C (solvent) hydrogen bonds, respectively, thereby disrupting (peptide) N-H···O=C (peptide) hydrogen bonds.[11-14]

Experimental Section

Peptides. The synthesis and characterization of Z-(L-Val)$_2$-L-Orn-(Boc)-OMe and Z-[L-Orn(Boc)]$_5$-OMe have already been reported.[10] The synthesis and characterization of Z-L-Orn(Boc)-(L-Val)$_2$-L-Orn-(Boc)-OMe, Z-[L-Orn(Boc)]$_2$-L-Orn(Boc)-OMe, Z-[L-Orn(Boc)]$_3$-L-Orn(Adoc)-L-Orn(Boc)-OMe, and Z-L-Orn(Adoc)-[L-Orn(Boc)]$_2$-L-Orn(Adoc)-L-Orn(Boc)-OMe will be described elsewhere.[15]

Infrared Absorption. Infrared absorption spectra were recorded by using a Perkin-Elmer Model 580 spectrophotometer and cells with a path length of 0.2, 0.1, and 0.05 mm. The band positions are accurate to ± 1 cm^{-1}. Spectrograde methylene chloride, dimethyl sulfoxide, and *N*,*N*-dimethylformamide were Fluka AG (Büchs, Switerzland) products.

Results

Figure 1 shows the titration curves in CH_2Cl_2/Me_2SO of the fully protected pentapeptides Z-[L-Orn(Boc)]$_2$-(L-Val)$_2$-L-Orn-(Boc)-OMe (and the two related short sequences, the tetrapeptide and the tripetide) and Z-[L-Orn(Boc)]$_5$-OMe at two different concentrations (50 and 10 mg/mL). The higher concentration is in the range typically employed in peptide synthesis. For each compound the relative intensity was calculated from the area of the stretching band of strongly intermolecularly hydrogen bonded C=O groups ($\cong$1630 cm^{-1}),[14,16-18] taking as 1.0 relative intensity the value observed in CH_2Cl_2 (or at saturation, if the compound is not completely soluble in CH_2Cl_2 at the concentration exam-

(1) This paper is part 99 of the "Linear Oligopeptides" series. For part 98 see: Benedetti, E.; Bavoso, A.; Di Blasio, B.; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. M. *Biopolymers* **1983**, *22*, 305-317.

(2) Mutter, M.; Pillai, V. N. R.; Anzinger, H.; Bayer, E.; Toniolo, C. In "Peptides 1980"; Brunfeldt, K., Ed.; Scriptor: Copenhagen, 1981; pp 660-665.

(3) Kent, S. B. H.; Merrifield, R. B. In "Peptides 1980"; Brunfeldt, K., Ed.; Scriptor: Copenhagen, 1981; pp 328-333.

(4) Mororder, L.; Gemeiner, M.; Goehring, W.; Jaeger, E.; Thamm, P.; Wünsch, E. *Biopolymers* **1981**, *20*, 17-37.

(5) Mutter, M.; Maser, F.; Klein, B.; Toniolo, C.; Bonora, G. M. In "Peptides 1982"; Blaha, K.; Malon, P., Eds.; de Gruyter: Berlin, 1983; p 729-734.

(6) Toniolo, C.; Bonora, G. M.; Schaaper, W. M. M. *Int. J. Pept. Protein Res.*, in press.

(7) Toniolo, C.; Bonora, G. M.; Mutter, M., unpublished experiments.

(8) Bonora, G. M.; Palumbo, M.; Toniolo, C.; Mutter, M. *Makromol. Chem.* **1979**, *180*, 1293-1304.

(9) Bonora, G. M.; Toniolo, C.; Pillai, V. N. R.; Mutter, M. *Gazz. Chim. Ital.* **1980**, *110*, 503-510.

(10) Marchiori, F.; Borin, G.; Stivanello, D.; Moretto, V.; Chessa, G. *Hoppe-Seyler's Z. Physiol. Chem.* **1982**, *363*, 1483-1491.

(11) Martin, D.; Hauthal, H. G. "Dimethyl Sulphoxide"; Van Nostrand-Reinhold: Wokingham, England, 1975.

(12) Shaw, Y. H.; Li, N. C. *Can. J. Chem.* **1970**, *48*, 2090-2095.

(13) Fillaux, F.; de Lozē, C. *J. Chim. Phys. Phys.-Chim. Biol.* **1972**, *69*, 36-44.

(14) Baron, M. H.; de Lozē, C.; Toniolo, C.; Fasman, G. D. *Biopolymers* **1978**, *17*, 2225-2239.

(15) Marchiori, F.; Borin, G., unpublished experiments.

(16) Miyazawa, T. In "Poly-α-Amino Acids"; Fasman, G. D., Ed.; Marcel Dekker: New York, 1967, pp 69-103.

(17) Toniolo, C.; Palumbo, M. *Biopolymers* **1977**, *16*, 219-224.

(18) Toniolo, C. In "Bioorganic Chemistry"; van Tamelen, E. E., Ed.; Academic Press: New York, 1977; pp 265-291.

0002-7863/84/1506-1455$01.50/0

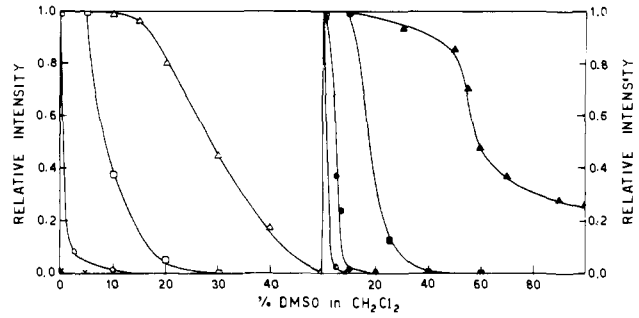


Figure 1. Relative intensity of the amide I C=O stretching band associated with strongly self-aggregated molecules (≅ 1630 cm^{-1}) in the IR absorption spectra of Z-[L-Orn(Boc)]$_5$-OMe [(○) and (●)], Z-[L-Orn-(Boc)]$_2$-(L-Val)$_2$-L-Orn(Boc)-OMe [(△) and (▲)], Z-L-Orn(Boc)-(L-Val)$_2$-L-Orn(Boc)-OMe [(□) and (■)], and Z-(L-Val)$_2$-L-Orn(Boc)-OMe [(×) and (◐)] in CH_2Cl_2/Me_2SO as a function of increasing percentages of Me_2SO at two different concentrations (10 mg/mL, left side; 50 mg/mL, right side).

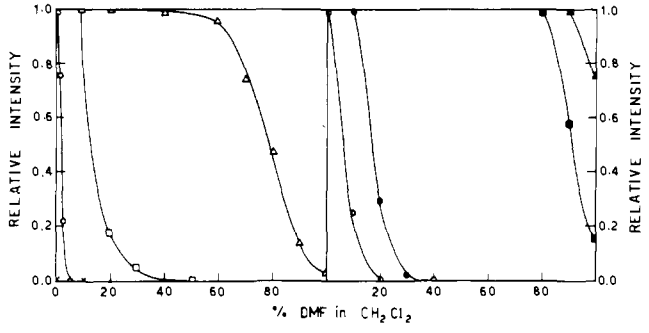


Figure 2. Relative intensity of the amide A N–H stretching band associated with strongly self-aggregated molecules (≅ 3290 cm^{-1}) in the IR absorption spectra of Z-[L-Orn(Boc)]$_5$-OMe [(○) and (●)], Z-[L-Orn-(Boc)]$_2$-(L-Val)$_2$-L-Orn(Boc)-OMe [(△) and (▲)], Z-L-Orn(Boc)-(L-Val)$_2$-L-Orn(Boc)-OMe [(□) and (■)], and Z-(L-Val)$_2$-L-Orn(Boc)-OMe [(×) and (◐)] in CH_2Cl_2/DMF as a function of increasing percentages of DMF at two different concentrations (10 mg/mL, left side; 50 mg/mL, right side).

ined). In Me_2SO-containing mixtures the region of the stretching bands of strongly intermolecularly hydrogen bonded N–H groups (≅3290 cm^{-1})[14,17] is less useful as a probe, owing to overlapping of the peptide–peptide N–H···O=C band to the peptide–solvent N–H···O=S band.[11-14]

Figure 2 illustrates the solvent titration curves of the same four compounds in CH_2Cl_2/DMF mixtures. The two concentrations examined were the same as above. In most cases more than 30% DMF was required to disrupt the self-aggregated structures: as a consequence, we were forced to use the 3290-cm^{-1} band as a probe, since the peptide C=O band near 1630 cm^{-1} is completely hidden by the intense C=O band of DMF in these solvent mixtures.

Figure 3 shows the titration curves in CH_2Cl_2/DMF mixtures of the pentapeptide Z-[L-Orn(Boc)]$_5$-OMe and the two analogues with mixed side-chain protection at the Orn residues (Boc and Adoc) at high concentration (40 mg/mL). At low concentration (10 mg/mL) in DMF and at high concentration (40 mg/mL) in Me_2SO only <10% DMF (or Me_2SO) was enough to destroy completely the self-aggregation of Z-[L-Orn(Boc)]$_5$-OMe, the pentapeptide (among the three pentapeptides reported in Figure 3) with the highest propensity to give intermolecularly hydrogen bonded structures. In these cases the disruption of self-aggregation was successfully monitored at ≅1630 cm^{-1}.

Discussion

Typical IR absorption spectra of these peptides (in CH_2Cl_2) in the N–H stretching (3500–3200 cm^{-1}) and C=O stretching (1800–1600 cm^{-1}) regions show bands near 3440 cm^{-1} (free NHs), 3360 cm^{-1} (weakly hydrogen bonded NHs), 3290 cm^{-1} (strongly intermolecularly hydrogen bonded NHs), 1735 cm^{-1} (ester carbonyl), 1705 cm^{-1} (free and weakly hydrogen bonded urethane carbonyls), 1680–1660 cm^{-1} (strongly hydrogen bonded urethane carbonyls and free and weakly hydrogen bonded peptide carbonyls), and 1630 cm^{-1} (strongly intermolecularly hydrogen bonded peptide carbonsyls).[14,16-19] It should be recalled that the characteristic frequency ranges for the amide A and amide I bands of β-structure forming peptides are found at 3295–3280 and 1635–1625 cm^{-1}, respectively.[14,16-18]

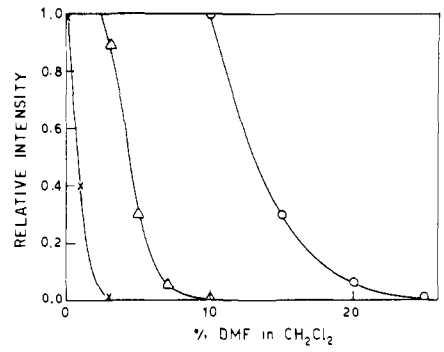


Figure 3. Relative intensity of the amide I C=O stretching band associated with strongly self-associated molecules (≅1630 cm^{-1}) in the IR absorption spectra of Z-[L-Orn(Boc)]$_5$-OMe (○), Z-[L-Orn(Boc)]$_3$-L-Orn(Adoc)-L-Orn(Boc)-OMe (△), and Z-L-Orn(Adoc)-[L-Orn-(Boc)]$_2$-L-Orn(Adoc)-L-Orn(Boc)-OMe (×) in CH_2Cl_2/DMF as a function of increasing percentages of DMF at 40 mg/mL concentration.

Addition of increasing percentage of either Me_2SO or DMF induces the decrease, and eventually the disappearance, of the 3290- and 1630-cm^{-1} bands. Thus, (peptide) N–H···O=C (peptide) hydrogen bonds are replaced by (peptide) N–H···O=S (solvent) and (peptide) N–H···O=C (solvent) hydrogen bonds, respectively.

Figure 1 shows the effects induced by main-chain length, specific amino acid sequence, and concentration on peptide self-aggregation in CH_2Cl_2/Me_2SO mixtures. The tripeptide forms extensive self-aggregation in CH_2Cl_2 solution only at high concentration (50 mg/mL). The other peptides adopt partially developed β-structures in CH_2Cl_2 even at low concentration (10 mg/mL). Upon increasing peptide concentration, all curves shift to considerably higher Me_2SO percentages, indicating that the hydrogen-bonded structures formed are of the intermolecular type. At both concentrations the following hierarchy is observed for the stability of the self-aggregated species: pentapeptide > tetrapeptide > tripeptide. This finding is in keeping with previous observations on main-chain length dependent β-structure formation in peptides.[2,7-9,14]

An additional point is clearly illustrated in Figure 1. The strength of self-aggregation of a peptide with a specific main-chain length is sequence dependent, much more Me_2SO being required to disrupt the β-structure adopted by the pentapeptide Z-[L-Orn(Boc)]$_2$-(L-Val)$_2$-L-Orn(Boc)-OMe than that of the pentapeptide Z-[L-Orn(Boc)]$_5$-OMe. In other words, the "host" pentapeptide [Orn(Boc) is the "host" residue] forms a less stable β-structure than the "host–guest" pentapeptide (Val is the "guest" residue). However, it is worth mentioning that two Val residues in a row, as "guest" residues, are able to disrupt the β-structure formed by Boc-(L-Ala)$_{10}$-NHR in 2,2,2-trifluoroethanol.[5] We conclude that a prediction of the propensity of a peptide to give a stable β-structure, based on the intrinsic potentials of the component amino acids for adopting that ordered secondary structure,[5,20] is not expected to be necessarily successful.

The effects of main-chain length, specific sequence, and concentration on peptide self-aggregation in CH_2Cl_2/DMF mixtures (Figure 2) closely parallel those observed in CH_2Cl_2/Me_2SO mixtures. From Figures 1 and 2 it turns out that Me_2SO is more

(19) Toniolo, C.; Palumbo, M.; Benedetti, E. *Macromolecules* **1976**, *9*, 420–424.
(20) Fasman, G. D. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1978**, *47*, 45–148.

effective than DMF in disrupting the peptide aggregates.

DMF, being less effective, is capable of discriminating the stability of the β-structure of the Boc side chain protected Orn homopentapeptide from those of the same pentapeptide with two different sequences of mixed side-chain protection (Boc and Adoc) at high concentration (40 mg/mL) (Figure 3). Clearly, the intentional interruption of the uniformity of the Boc protecting groups in the Orn side chains decreases the stability of the β-structure of the corresponding peptides. Interestingly, Moroder et al. have recently reported that disruption of the uniformity of the *tert*-butyl alcohol derived protecting groups in a synthetic somatostatin fragment, by blocking the ε-amino function of a Lys residue as the Adoc derivative, exerts a strong solubility effect.[4]

Conclusions

We have shown that main-chain length, specific amino acid sequence, concentration, and uniformity of side-chain protection are all factors playing a role on the stability of the peptide self-aggregated species. It is gratifying to note that the (quantitative) scales of increasing tendency of the peptides to self-aggregate in solution, determined in this work, parallel the (qualitative) scales of their decreasing solubility properties.[10,15] Incidentally, the hierarchies of the propensity of the peptides to assume strongly hydrogen bonded structures in the solid state[21] are in full agreement with those found in solution. We have also demonstrated that IR absorption is a useful technique to titrate quantitatively the extent of self-aggregation in peptides. In general, the titration of the 1630-cm^{-1} band can be carried out more accurately than that of the 3290-cm^{-1} band. However, at high DMF percentages the short-frequency tail of the amide band of the solvent overlaps the 1630-cm^{-1} absorption. Concentrations close to those usually employed in peptide synthesis (40–50 mg/mL) and N^{α}-deblocked peptides in the presence/absence of added tertiary amine[8,9] are also suitable for this study. We believe that this method will be of profit to chemists interested in the synthesis of medium-sized peptides, often poorly reactive owing to strong self-association and low solubility.

Acknowledgment. The expert technical assistance of Vittorio Moretto is gratefully acknowledged.

Registry No. Z-[L-Orn(Boc)]$_5$-OMe, 85571-69-3; Z-(L-Val)$_2$-Orn-(Boc)-OMe, 85571-67-1; Z-L-Orn(Boc)-(L-Val)$_2$-L-Orn(Boc)-OMe, 88376-82-3; Z-[L-Orn(Boc)]$_2$-(L-Val)$_2$-L-Orn(Boc)-OMe, 88376-83-4; Z-[L-Orn(Boc)]$_3$-L-Orn(Adoc)-L-Orn(Boc)-OMe, 88376-84-5; Z-L-Orn(Adoc)-[L-Orn(Boc)]$_2$-L-Orn(Adoc)-L-Orn(Boc)OMe, 88391-94-0.

(21) Toniolo, C.; Bonora, G. M., unpublished observations.

Transition-State Structural Features for Anilide Hydrolysis from β-Deuterium Isotope Effects[1a]

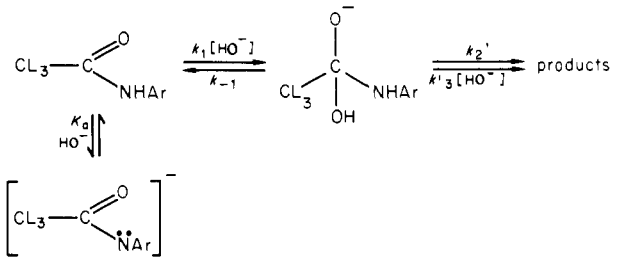
Ross L. Stein,[1b] Hisashi Fujihara, Daniel M. Quinn,[1c] G. Fischer, G. Küllertz, A. Barth, and Richard L. Schowen*

Contribution from the Department of Chemistry, University of Kansas, Lawrence, Kansas 66045, and the Department of Biochemistry, Biosciences Section, Martin Luther University, DDR-4020 Halle, German Democratic Republic. *Received August 8, 1983. Revised Manuscript Received October 14, 1983*

Abstract: The hydrolysis in basic aqueous solution of p-$NO_2C_6H_4NHCOCH_3$ and p-$NO_2C_6H_4NHCOCD_3$ at 30 °C from $[HO^-]$ = 0.002 to 2.31 M generates observed isotope effects k_o^H/k_o^D that begin in the least basic solution at 0.967 ± 0.011, pass through a maximum of around 0.98 at $[HO^-]$ about 0.03 M, and then fall to 0.933 ± 0.020 in the most basic solution. This phenomenon arises from a base-dependent mixture of rate-limiting transition states. At the lowest base concentrations, decomposition of the uninegative tetrahedral adduct, with k_2^H/k_2^D = 0.94 ± 0.02, dominates. In the intermediate range, decomposition of the dinegative adduct, with k_3^H/k_3^D = 1.00 ± 0.02, becomes important. In the most basic solutions, nucleophilic attack of hydroxide on the substrate (k_1^H/k_1^D = 0.94 ± 0.01) assumes the major role. These isotope effects are consistent with quasi-tetrahedral transition states for formation and decomposition of the uninegative adduct and with rate-limiting trapping or diffusional separation of $\{CH_3CO_2^-, ArNH^-\}$ for decomposition of the dinegative adduct. The observed isotope effect at $[HO^-]$ = 0.208 M seems to show an anomalously large dependence on temperature, changing from 0.944 ± 0.016 at 3.8 °C to 0.977 ± 0.019 at 50 °C, which generates on an Arrhenius model A_H/A_D = 1.17 and ΔE_a = 117 cal mol^{-1}. The data are shown, however, also to be fully consistent with a different model: a shift in limitation of the rate away from the k_1 transition state toward the k_3 transition state, as the temperature is raised, and with simple zero-point-energy origins for the individual isotope effects.

Secondary hydrogen isotope effects can be used to deduce the structural features of transition states,[2-5] and the technique is being used for this purpose in the investigation of enzyme catalytic power.[3-5] However, in many complex reaction systems, including both enzymic and nonenzymic examples, a number of different processes may simultaneously contribute to limiting the reaction rate. Then the observed isotope effect may be a complicated average of the effects on the elementary steps.[6,7] Alternatively

Scheme I



(1) (a) Research in Lawrence was supported through Grant GM-20198 by the National Institute of General Medical Sciences. (b) Current address: Pulmonary Pharmacology, ICI Americas Inc., Wilmington, DE 19897. (c) Current address: Department of Chemistry, University of Iowa, Iowa City, IA 52242.

(2) Shiner, V. J., Jr. In "Isotope Effects in Chemical Reactions", Collins, C. J., Bowman, N. S., Eds.; Van Nostrand-Reinhold: Princeton, NJ 1970; Chapter 2.

(3) Kirsch, J. F. In "Isotope Effects on Enzyme-Catalyzed Reactions"; Cleland, W. W., O'Leary, M. H., Northrop, D. B., Eds.; University Park Press: Baltimore, 1977; pp 100–121.

(4) Hogg, J. L. In "Transition States of Biochemical Processes"; Gandour, R. D., Schowen, R. L., Eds.; Plenum Press: New York, 1978; Chapter 5.

(5) Cleland, W. W. *Method. Enzymol.* **1982**, *87C*, 625.

0002-7863/84/1506-1457$01.50/0