# A new protein inhibitor of trypsin and activated Hageman factor from pumpkin (*Cucurbita maxima*) seeds

Ramaswamy Krishnamoorthi[1], YuXi Gong[1] and Michael Richardson[2]

[1] *Department of Biochemistry, Kansas State University, Manhattan, KS 66506, USA* and [2] *Department of Biological Sciences, University of Durham, Durham, UK*

Received 3 August 1990

A protein inhibitor (CMTI-V; $M_r$ 7106) of trypsin and activated Hageman factor (Factor $XII_a$), a serine protease involved in blood coagulation, has been isolated for the first time from pumpkin (*Cucurbita maxima*) seeds by means of trypsin-affinity chromatography and reverse phase high performance liquid chromatography (HPLC). The dissociation constants of the inhibitor complexes with trypsin and Factor $XII_a$ have been determined to be $1.6 \times 10^{-8}$ and $4.1 \times 10^{-8}$ M, respectively. The primary structure of CMTI-V is reported. The protein has 68 amino acid residues and one disulfide bridge and shows a high level of sequence homology to the Potato I inhibitor family. Furthermore, its amino terminus consists of an N-acetylates Ser. The reactive site has been established to be the peptide bond between $Lys^{44}$–$Asp^{45}$. The modified inhibitor which has the reactive site peptide bond hydrolyzed inhibits trypsin but not the Hageman factor.

Inhibitor; Serine protease; *Cucurbita maxima*; Pumpkin; Trypsin; Activated Hageman factor; Blood coagulation

## 1. INTRODUCTION

Protein inhibitors of trypsin and activated Hageman factor were isolated from pumpkin (*Cucurbita maxima*) seeds by Polanowski et al. [1] and Hojima et al. [2,3], independently. Elaborate work by the former group led to the isolation of 3 iso-inhibitors ($M_r$ ~3000), termed CMTI-I, CMTI-III, and CMTI-IV, and the determination of their amino acid sequences [4]. These protein inhibitors each have 3 disulfide bridges and are believed to be the smallest known globular trypsin inhibitors. Recently, an X-ray crystallographic structure of the CMTI-I–trypsin complex was reported by Bode et al. [5]. The inhibitor binds to the enzyme in the same way as does a substrate. Our interest in characterizing the solution structures of these proteins by nuclear magnetic resonance (NMR) spectroscopy prompted us to develop a rapid resolution-enhanced method of isolation. As a result, using trypsin-affinity chromatography followed by reversed phase high perfomance liquid chromatography (HPLC), we have now isolated for the first time a higher molecular mass protein ($M_r$ ~7106; CMTI-V) which specifically inhibits both trypsin and activated Hageman factor, but not plasma kallikrein, another serine protease closely related to Hageman factor. We report here the isolation, inhibitory properties, and primary structure of CMTI-V.

*Correspondence address:* R. Krishnamoorthi, Department of Biochemistry, Willard Hall, Kansas State University, Manhattan, KS 66506, USA

*Abbreviations:* CMTI, *Cucurbita maxima* trypsin inhibitor; HPLC, high performance liquid chromatography; NMR, nuclear magnetic resonance spectroscopy; BAPNA, Benzoyl-L-Arg-*p*-nitroanilide; S-2302, D-Pro-Phe-Arg-*p*-nitroanilide; TFA, trifluoroacetic acid; SDS, sodium dodecyl sulfate; DABITC, 4-N,N-dimethylaminoazobenzene-4′-isothiocyanate; PITC, phenylisothiocyanate.

## 2. MATERIALS AND METHODS

Pumpkin seeds (commercially known as Big Max, *Cucurbita maxima*) were purchased from Northrup King Co. Trypsin-affinity column chromatography matrix was purchased from Bio-Rad. Enzymes used were obtained as indicated: bovine trypsin (Sigma), activated Hageman factor (Calbiochem), and plasma kallikrein (Helena Laboratories). The two trypsin substrates used in the determination of dissociation constants of the enzyme–inhibitor complexes were benzoyl-L-Arg-*p*-nitroanilide (BAPNA; Sigma) and D-Pro-Phe-Arg-*p*-nitroanilide (S-2302; Helena Laboratories). All other reagents and supplies were reagent grade or better.

### 2.1. *Isolation of Protein Inhibitors*

The seeds were defatted and the crude proteins were precipitated by salting out with ammonium sulfate. The sediment was dissolved in water, dialyzed, and centrifuged. The clear supernatant liquid was subjected to trypsin-affinity chromatography using an Affi-Prep10 column to which trypsin had been attached. The trypsin inhibitors were liberated by using a glycine buffer (pH 2.0), as described by Lei and Reeck [6]. The fractions were subjected to reverse phase HPLC after dialysis and lyophilization using a Varian system Model 2510 and a Varian C-18 Protein Column: the two solvents employed were 0.1% (v/v) trifluoroacetic acid (TFA) in water (Solvent A) and 0.1% (v/v) TFA in acetonitrile (Solvent B). During elution (flow rate 2.0 ml/min) a gradient was employed such that in 60 min the solvent composition altered from 0% to 40% (v/v) B. It was further changed to 100% B in 15 min and finally changed back to 0% B in 15 min. The eluent was monitored with an ISCO detector at 280 nm. Fractions labeled a, b, c, d, e, and f were collected and re-chromatographed to homogeneity.

*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/90/$3.50 

2.2. *Enzyme Inhibition Assays*

Inhibition of trypsin was followed by monitoring the rate of hydrolysis of BAPNA spectrophotometrically at 405 nm [7]. Similarly, inhibition of activated Hageman factor and of plasma kallikrein was determined by following the hydrolysis of S-2302. The dissociation constants ($K_i$) of the enzyme-inhibitor complexes were determined by the method of Bieth [8]. Stoichiometric concentrations in the $10^{-7}$ M range of the inhibitors and the enzymes were used. the substrate concentration was of the order of $10^{-4}$ M.

2.3. *Polyacrylamide gel electrophoresis*

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed using 18% polyacrylamide and 0.4% N,N-methylene-bis acrylamide as described by Laemmli [9]. Protein bands were visualized by staining with a Coomassie brilliant blue solution. Protein standards (Low range kit; Diversified Biotech) for molecular mass in the range 2–21 kDa were used.

2.4. *Amino acid analysis*

The method described by Bidlingmeyer et al. [10] was used. The inhibitors were hydrolyzed in 6 M hydrochloric acid for 24 h at 105°C and the liberated amino acids were reacted with phenylisothiocyanate (Pierce Chem. Co.) to form the phenylthiocarbamate derivatives wich were then analyzed by reverse phase HPLC on a Pico-Tag column (Waters Associates). Cysteines were determined by the method of Moore [11].

2.5. *Amino acid sequencing*

An applied Biosystems Model 477A Sequencer/Model 120A PTH Analyzer system was employed for the automatic determination of amino acid sequences of the inhibitor molecules using the Edman degradation. The phenylthiohydantoin derivatives were identified by reverse phase HPLC. Cysteines were derivatized by in situ alkylation using 4-vinylpyridine as shown by Andrews and Dixon [12]. CMTI-e and CMTI-f were reduced and S-carboxymethylated as described in [13]. The components present in CMTI-f were subsequently separated by gel filtration on a column (1 × 200 cm) of Biogel P6 in 0.1 M ammonium bicarbonate pH 8.1. Samples of the reduced and carboxymethylated proteins were digested separately with chymotrypsin, pepsin and the protease from *S. aureus* V8 as in [14]. A further sample of CMTI-e was treated with cyanogen bromide (100-fold molar excess) in 70% formic acid for 24 h at 20°C. The peptides resulting from these treatments were purified by reverse phase HPLC on a Varian 5000 HPLC fitted with a Vydac C18 column (25 cm × 4.6 mm; 218TP54, Technicol, Stockport) using variable gradients of 0–70% acetonitrile (HPLC grade, S. Rathburn) in 0.1% trifluoroacetic acid as described in [14]. The peptides obtained were sequenced using the manual 4-N,N-dimethylaminoazobenzene-4′-isothiocyanate (DABITC)/phenylisothiocyanate (PITC) double coupling method [15].

Attempts were made to unblock the N-terminus of CMTI-e by treatment with pyroglutamate aminopeptidase as described in [16], methanolysis with conc. HCl/methanol (1:11 v/v) for 48 h at 20°C as in [17] and by anhydrous TFA as described in [18].

## 3. RESULTS AND DISCUSSION

A typical reverse phase HPLC result obtained after trypsin-affinity chromatography is displayed in Fig. 1A: six peaks, identified a–f, are noted. They are termed CMTI-a through CMTI-f, respectively, following the terminology of Wieczorek et al. [4]. Figs 1B and 1C depict the inhibitory activity towards trypsin and activated Hageman factor, respectively, of each of the eluent fractions collected during the course of time. None of the fractions obtained inhibited plasma kallikrein. Trypsin can cleave the reactive site peptide bond of a virgin inhibitor in the enzyme–inhibitor complex, thus producing its modified form [19]. In order to identify peaks that are produced this way, each of the proteins, CMTI-a, CMTI-b, CMTI-c, and CMTI-e, was reacted with trypsin and the reaction mixture was analyzed by reverse phase HPLC. The results (not shown) reveal that CTMI-a and CMTI-c form one pair of virgin and modified inhibitors, CMTI-b and CMTI-d form a second pair, and, CMTI-e and CMTI-f make the third pair.

Polyacrilamide gel electrophoresis of these inhibitors indicated that CMTI-a through CMTI-d each had a molecule mass of about 3 kDa, whereas CMTI-e and CMTI-f each had a molecular mass of about 7 kDa.

In order to identify the virgin and modified inhibitor in each pair, we resorted to amino acid sequencing. The automated sequencing of amino acid residues for CMTI-a revealed the presence of two amino termini; one had the sequence, RVCPR, and the other, ILMKC, thus unambiguously identifying it as the modified form. Furthermore, the complete sequencing of CMTI-c, the other protein in the pair, established with certainty that it was indeed the same protein as the one termed CMTI-III, by Wieczorek et al. [4]. Similarly complete

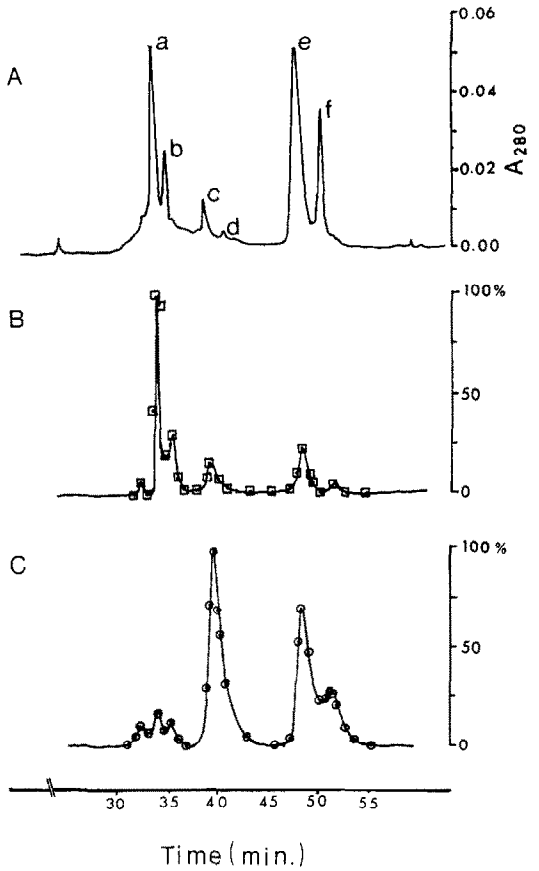


Fig. 1. Isolation and inhibitory activities of pumpkin seed inhibitors. (A) Reverse phase HPLC; see text for elution details. (B) Inhibition of trypsin; (C) Inhibition of activated Hageman factor. Inhibition of the enzyme by each fraction (0.5 ml) is expressed as percentage of maximum inhibition observed.

amino acid sequencing of CMTI-b proved that it had two amino termini and the complete sequence matched that of the protein, identified as CMTI-I by the same workers [4]. Because CMTI-d was produced from CMTI-b by reaction with trypsin, we conclude that CMTI-d is the virgin form of CMTI-b.

Determination of the amino acid composition of CMTI-e (CMTI-V; data not shown) proved that CMTI-e was certainly not a dimer of any of the 3 kDa proteins. It was also found that CMTI-e had only disulfide linkage, which is unusual, and that it had a high proline content of 8.3 residues. Initial attempts to sequence CMTI-e by automated means revealed that it had a blocked N-terminus. However, similar efforts with CMTI-f yielded the sequence: DFRCNRVRIWVN-KRGLVVSTPRI, which was later shown to be the first 23 residues of the C-terminal fragment which resulted from trypsin-catalyzed cleavage of the reactive site peptide bond (P1'; [19]) which occurred during the affinity step of purification. The modified nature of CMTI-f was confirmed when it was shown that after reduction and S-carboxymethylation the protein could be separated into 3 fragments by gel filtration on Biogel P6 or by reverse phase HPLC. The largest fragment (68 residues) was the intact protein, and the two other fragments (44) and 24 residues) were respectively the N- and C-terminal fragments resulting from the cleavage of the reactive site.

CMTI-f could be converted back to CMTI-e by reaction with trypsin followed by HPLC analysis. Thus it is established that CMTI-e is the virgin inhibitor, whereas CMTI-f is the modified form in which the reactive site peptide bond is hydrolyzed. Following the convention of Wieczorek et al. [4], we term the new protein inhibitor CMTI-e as CMTI-V.

Attempts to unblock the N-terminus of the intact CMTI-V by digestion with pyroglutamate aminopeptidase were unsuccessful. However, treatments of the inhibitor with conc. HCl/methanol [17] or with anhydrous TFA [18] subsequently permitted the N-terminal sequence of SSCPGKSS to be determined indicating that the terminal serine was probably N-acetylated. The 400 MHz NMR spectrum of CMTI-V shows a sharp uncoupled peak at ~2.1 ppm (not shown) which could be attributed to the $CH_3$ of the acetyl group. The complete amino acid sequence of inhibitor CMTI-V is shown in Fig. 2 together with the overlapping peptides from which it was deduced.

The structure of CMTI-V shows clear and strong sequence similarities with the potato I iso-inhibitors [20–22], inhibitors from tomato leaves [23], barley [24], bean [25] and also with the leech inhibitor eglin [26]. It should be noted that other workers have previously reported a trypsin inhibitor from seeds of *Momordica charantia* (Cucurbitaceae) which also had a weak homology with the members of the potato I family [27]. Our finding that CMTI-f is a modified form of CMTI-V and the determination of the amino acid sequences of both proteins show that the peptide bond $Lys^{44}$–$Asp^{45}$ is the trypsin reactive (inhibitory) site of the pumpkin inhibitor. When the sequences of the potato I family of inhibitors are aligned (Fig. 3) this bond in the pumpkin CMTI-V protein occurs in exactly the same position as the reactive sites in all of the other inhibitors.

In Table I are collected the dissociation constants of the complexes formed between the various inhibitors and the enzymes, trypsin and activated Hageman factor. The inhibitory activity towards activated Hageman factor is lost in each of the two modified 3kDa proteins (CMTI-a and CMTI-b) and the modified form of the

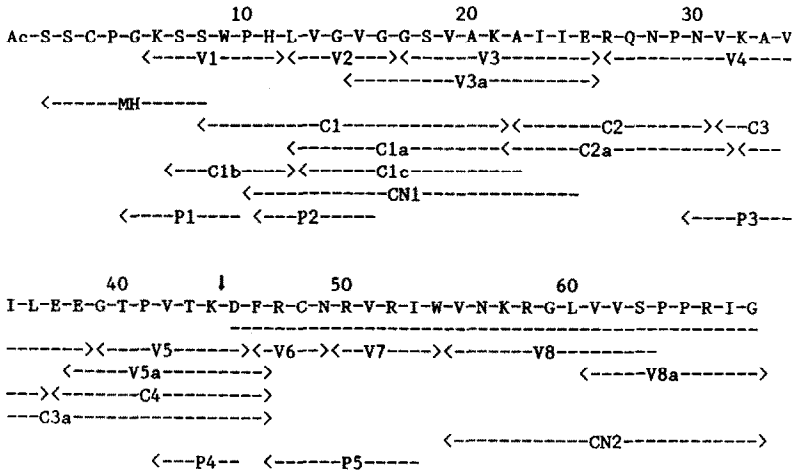


Fig. 2. The amino acid sequence of trypsin inhibitor CMTI-V (CMTI-e) from seeds of pumpkin (*Cucurbita maxima*). V, peptides from digestion with *S. aureus* V8 protease; C, chymotryptic peptides; P, peptic peptides; CN, fragments from cleavage with cyanogen bromide; MH, sequence determined after treatment of CMTI-V with conc. HCl/methanol or anhydrous TFA. (→) indicates residues determined by automated degradation in pulsed phase sequencer; all other residues determined by DABITC/PITC method. (↓) indicates reactive site peptide bond cleaved by trypsin on affinity column which converts CMTI-V to modified form CMTI-f.

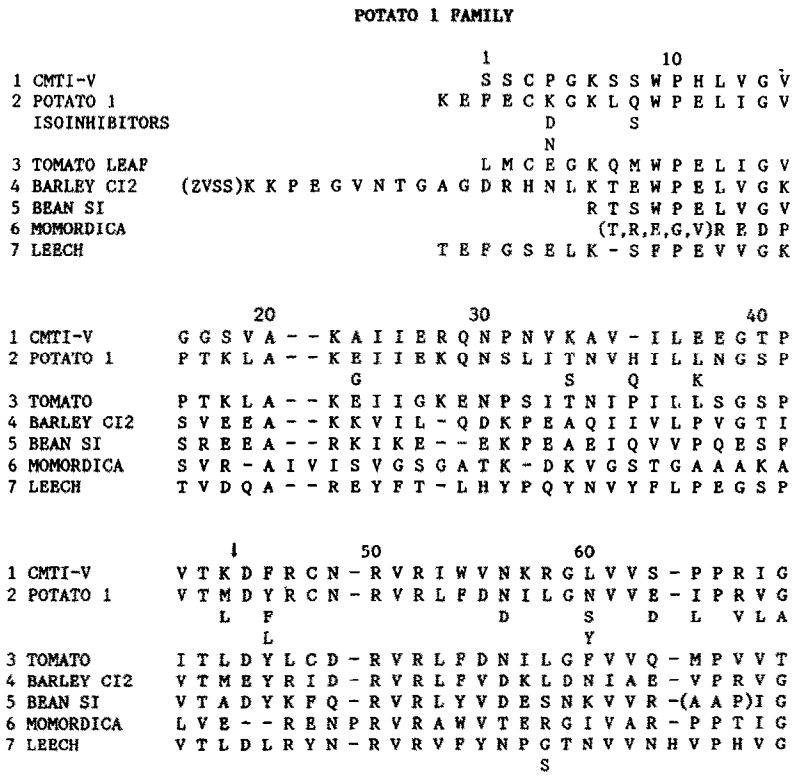


Fig. 3. Alignments of the amino acid sequences of the members of the potato I family of proteinase inhibitors. (1) CMTI-V inhibitor of trypsin and activated Hageman factor from pumpkin (*Cucurbita maxima*) seeds; (2) iso-inhibitors of chymotrypsin from tubers of potato [20–22]; (3) inhibitor I from wounded leaves of tomato [23]; (4) chymotrypsin inhibitor 2 from seeds of barley [24]; (5) subtilisin inhibitor from seeds of broad bean (*Vicia faba*) [25]; (6) trypsin inhibitor from seeds of *Momordica charantia* [27]; (7) inhibitor eglin from leech (*Hirudo medicinaalis*) [26]. (↓) indicates position of reactive site peptide bond (not determined in *Momordica* and leech proteins).

newly isolated 7 kDa protein (CMTI-f). On the other hand, trypsin inhibition is demonstrated by all of them. CMTI-a (which is the same as CMTI-III [4]) and CMTI-d (which is the same as CMTI-I [4]) differ from each other by a single amino acid substitution: Lys-9 in CMTI-III has been replaced by a Glu in CMTI-I ([4] and present work). However, this subsitution at P′4 position has no effect on the stability of the enzyme-inhibitor complexes formed with trypsin and activated Hageman factor (Table I). Furthermore, in spite of their very different primary structures both the 3 kDa proteins (CMTI-I and CMTI-III) and the presently discovered 7 kDa protein (CMTI-V form complexes of similar stability with trypsin and with activated Hageman factor. The results thus indicate not only the stringent structural requirement of an inhibitor to specifically inhibit activated Hageman factor, but also the preservation of similar structural features in the apparently unrelated 3 kDa and 7 kDa trypsin inhibitors from pumpkin seeds. Interestingly, the association equilibrium constants determined by Wyn and Laskowski for CMTI-I and CMTI-III with the Hageman factor reveal that the value for CMTI-III is about 62 times that for CMTI-I [28]. We are currently carrying out NMR studies of CMTI-V to determine its three-dimensional solution structure.

Table I

Comparison of dissociation constants ($K_i$) for the complex of pumpkin seed trypsin inhibitors (CMTI), barley trypsin inhibitor (BTI), and corn Hageman factor inhibitor (CHFI) with trypsin and activated Hageman factor.

| Protein | Trypsin $K_i$ ($M$) | Hageman factor $K_i$ ($M$) | Ref. |
|---|---|---|---|
| CMTI-a | $2.2\times10^{-9}$ | n.d. | This work |
| CMTI-b | $5.1\times10^{-9}$ | n.d. | This work |
| CMTI-c | $2.4\times10^{-8}$ | $2.9\times10^{-8}$ | This work |
| CMTI-d | $1.7\times10^{-8}$ | $2.2\times10^{-8}$ | This work |
| CMTI-V | $1.6\times10^{-8}$ | $4.1\times10^{-8}$ | This work |
| BTI | $2.3\times10^{-8}$ | $1.1\times10^{-7}$ | [7] |
| CHFI | $2.2\times10^{-8}$ | $2.1\times10^{-9}$ | [7] |
| CMTI[a] | $1.3\times10^{-8}$ | $4.1\times10^{-8}$ | [3] |

n.d. not determined because of absence of stoichiometric binding (loose binding). [a] Based on amino acid composition, Hojima et al. [3] suggest this to be the same as CMTI-III of Wieczorek et al. [4].

*Acknowledgements:* The authors thank Prof. G.R. Reeck at the Kansas State University for useful discussions and for permitting the use of his equipment for the determination of amino acid compositions. The financial support of the project by the American Heart Association, Kansas Affiliate, Wesley Foundation, Wichita, Kansas, USA and Kansas Agricultural Experiment Station is gratefully acknowledged. This is publication 91-56-J of the Kansas Agricultural Experiment Station.

## REFERENCES


[1] Polanowski, A., Wilusz, T., Nienartowicz, B., Cieslar, E., Slominska, A. and Nowaka, K. (1980) Acta Biochim Pon. 27, 371–382.

[2] Hojima, Y., Pierce, J.V. and Pisano, J.J. (1980) Thromb. Res. 20, 163–171.

[3] Hojima, Y., Pierce, J.V. and Pisano, J.J. (1982) Biochemistry 21, 3741–3746.

[4] Wieczorek, M., Otlewski, J., Cook, J., Parks, K., Leluk, J., Wilimowska-Pelc, A., Polanowski, A., Wilusz, T. and Laskowski Jr., M. (1985) Biochem Biophys. Res. Commun. 126, 646–652.

[5] Bode, W., Greyling, H.J., Huber, R., Otlewski, J. and Wilusz, T. (1989) FEBS Lett. 242, 285–292.

[6] Lei, M.-G. and Reeck, G.R. (1986) J. Chromatogr. 363, 315–321.

[7] Chong, G.L. and Reeck, G.R. (1987) Thromb. Res. 48, 211–221.

[8] Bieth, J. (1974) in Proteinase Inhibitors (Fritz, H., Tschesche, H., Green, L.J. and Truscheit, E. eds) pp. 463–469, Springer, New York.

[9] Laemmli, U.K. (1970) Nature 227, 680–685.

[10] Bidlingmeyer, B.A., Cohen, S.A. and Tarvin, T.L. (1984) J. Chromatogr. 336, 93–104.

[11] Moore, S. (1963) J. Biol. Chem. 238, 235–237.

[12] Andrews, P.C. and Dixon, J.E. (1987) Anal. Biochem. 161, 524–528.

[13] Crestfield, A.M., Moore, S. and Stein, W.H. (1963) J. Biol. Chem. 238, 622–627.

[14] Richardson, M., Campos, F.D.A.P., Moreira, R.A., Ainouz, I.L., Begbie, R., Watt, W.B. and Pusztai, A. (1984) Eur. J. Biochem. 14, 101–111.

[15] Chang, J.Y., Brauer, D. and Wittman-Liebold, B. (1978) FEBS Lett. 93, 205–214.

[16] Podell, D.N. and Abraham, G.N. (1978) Biochem. Biophys. Res. Commun. 81, 176–185.

[17] Van Beynum, G.M.A., DeGraaf, J.M., Castel, A., Kraal, B. and Bosch, L. (1977) Eur. J. Biochem. 72, 63–78.

[18] Wellner, D., Panneerselam, C. and Horecker, B.L. (1990) Proc. Natl. Acad. Sci. USA 87, 1947–1949.

[19] Laskowski Jr., M. and Kato, I. (1980) Annu. Rev. Biochem. 49, 593–626.

[20] Richardson, M. (1974) Biochem. J. 137, 101–112.

[21] Richardson, M. and Cossins, L. (1975) FEBS Lett. 53, 161.

[22] Richardson, M., Barker, R.D.J., Macmillan, R. and Cossins, L. (1977) Phytochemistry 16, 837–839.

[23] Graham, J.S., Pearce, G., Merryweather, J., Titani, K., Ericksson, L.H. and Ryan, C.A. (1985) J. Biol. Chem. 260, 6555–6560.

[24] Svendsen, I., Boisen, S. and Hejgaard, J. (1982) Carlsberg Res. Commun. 47, 45–53.

[25] Svendsen, I., Hejgaard, J. and Chavan, J.K. (1984) Carlsberg Res. Commun. 49, 493–502.

[26] Seemuller, U.M., Eulitz, M., Fritz, H. and Strobl, A. (1980) Hoppe-Seylers Z. Physiol. Chem. 261, 1841–1846.

[27] Zeng, F.Y., Qian, R.Q. and Wang, Y. (1988) FEBS Lett. 234, 35–38.

[28] Wynn, R. and Laskowski Jr., M. (1990) Biochem. Biophys. Res. Commun. 166, 1406–1410.