Localisation of the major reactive lysine residue involved in the self-crosslinking of proteinase-activated *Limulus* α_2-macroglobulin

Klavs Dolmer[a], Lise B. Husted[a], Peter B. Armstrong[b,c], Lars Sottrup-Jensen[a,*]

[a]*Department of Molecular and Structural Biology, C.F. Møllers Allé bldg. 130, University of Aarhus, DK-8000 Århus C, Denmark*
[b]*Marine Biological Laboratory, Woods Hole, MA 02543, USA*
[c]*Department of Zoology, University of California, Davis, CA 95616, USA*

Received 26 June 1996

Abstract When α_2-macroglobulin (α_2M) from the American horseshoe crab, *Limulus polyphemus*, reacts with proteinases, its thiol esters, like those of other α-macroglobulins, become activated, leading to the formation of covalently crosslinked species that can be detected as high molecular weight bands in reducing SDS-PAGE. While other α-macroglobulins extensively form crosslinks to the reacting proteinase, *Limulus* α_2M does not. It rather becomes internally crosslinked. It was found from N-terminal sequence analysis of purified [^{14}C]carboxymethylated peptides from *Limulus* α_2M-trypsin complexes that an isopeptide bond formed in approx. 60% yield from the thiol esterified Gln-1002 specifically to Lys-254 in the opposing monomer of the disulphide bridged dimer is the main cause of the internal crosslinking.

Key words: α_2-Macroglobulin superfamily; Proteinase inhibitor; Thiol ester; Crosslink; Limulus

1. Introduction

The α_2-macroglobulin (α_2M) superfamily consists of the high molecular weight mono-, di- or tetrameric proteinase inhibitors, the α-macroglobulins, and the non-inhibitory complement components C3, C4 and C5 [1]. The α-macroglobulins are found in the blood of vertebrates and invertebrates, and in the egg white of birds and reptiles (for reviews, see [2,3]).

Upon proteolytic cleavage in a specific region, designated the 'bait' region, the inhibitor undergoes a conformational change that 'traps' the proteinase within the 'cage' of the di- and tetrameric inhibitors [4]. Along with the entrapment of the proteinase, a domain in the C-terminal part of the polypeptide chain is exposed [5,6] that is recognised by the α_2M-receptor/LDL-receptor related protein (α_2MR/LRP) [7–10], leading to rapid clearance of the complex [11,12]. Coincident with the structural change, an internal β-cysteinyl-γ-glutamyl thiol ester, which is found in the strictly conserved sequence GCGEQNM, is activated, allowing the typical α-macroglobulin to form an ester or amide crosslink through the γ-carbonyl group of the glutamine and hydroxyl or, typically, amino groups on the proteinase [13–15]. The thiol ester can also be cleaved by small nucleophiles such as methylamine with the covalent linkage of methylamine via an amide bond to the γ-carbonyl of the thiol esterified glutamine, which in general leads to the same conformational change as the proteolytic attack [16]. In reducing SDS-PAGE, the covalently linked α-macroglobulin-proteinase complexes are recognised as bands with apparent molecular mass higher than the approx. 180 kDa monomer. However, as shown previously, α_2M from the American horseshoe crab, *Limulus polyphemus* [17–20], does not crosslink biotinylated trypsin [21], although bands of apparent high molecular weight are seen in reducing SDS-PAGE after proteinase treatment, indicative of isopeptide crosslinking.

We here show that *Limulus* α_2M entraps proteinases with high efficiency but forms crosslinks to ^{125}I-labelled trypsin ([^{125}I]trypsin) only to a very limited extent, indicating that the low level of crosslinking of trypsin reported [21] is not the result of lysine derivatisation of the biotinylated trypsin used in that study. We also show that the high molecular weight products mainly are formed through specific crosslinking of the activated thiol ester to Lys-254 (*Limulus* α_2M numbering [22]) in the opposing monomer of the *Limulus* α_2M dimer.

2. Materials and methods

Human α_2M was prepared from outdated plasma [23]. *Limulus* α_2M was purified from freshly prepared plasma as previously described [24]. Tosyl phenylalanyl chloromethyl ketone-treated trypsin was from Cooper Biochemicals, iodo[^{14}C]acetic acid and Na^{125}I were from Amersham, Mono S HR 5/5, benzamidine-Sepharose 6B, and Sephadex G 25 SF (10×100 mm) were from Pharmacia, Nucleosil C-18 100-5 and Vydac C-18 packing materials were from Machery-Nagel and The Separations Group, respectively. Standard chemicals were of analytical grade from Rathburn, Sigma or Merck.

Sequence analysis was performed on an Applied Biosystems 477A equipped with a model 120A online chromatograph. Reagents and solvents were from Applied Biosystems, and the instruments were operated according to the manufacturer's instructions, with some modifications [25].

Bovine β-trypsin was purified by benzamidine-Sepharose chromatography. Trypsin (15 mg) was dissolved in 50 mM Na-acetate, 10 mM $CaCl_2$, pH 5.5, and loaded on a benzamidine-Sepharose 6B column (7 ml) equilibrated with the same buffer. 95% pure β-trypsin was recovered after elution with a pH gradient formed from 50 ml 50 mM Na-acetate, 10 mM $CaCl_2$, pH 5.5, and 50 ml 20 mM citrate-HCl, 10 mM $CaCl_2$, pH 2.0. Iodination of β-trypsin was performed as described in [26].

SDS-PAGE was performed using 10–20% polyacrylamide gels in the standard Tris-glycine buffer system, and non-denaturing PAGE was performed in 5–10% polyacrylamide gels, using the same buffer system without SDS.

Aliquots (20 μg) of human and limulus α_2M were treated with 1.5×10^5 cpm [^{125}I]trypsin for 2 min and then with 1 equivalent of unlabelled trypsin for 1 min at room temperature. The reaction was stopped by adding 3,4-dichloroisocoumarin (DCI) to 1 mM final con-

*Corresponding author. Fax: (45) (89) 42 26 66.

Abbreviations: α_2M, α_2-macroglobulin; α_2MR/LRP, α_2M-receptor/LDL-receptor related protein; CMCys, *S*-carboxymethylcysteine; DCI, 3,4-dichloroisocoumarin; DTE, dithioerythritol; HPLC, high-performance liquid chromatography; LDL, low density lipoprotein; PTH, phenylthiohydantion; RP, reverse phase; TFA, trifluoroacetic acid

PII S0014-5793(96)00852-6

centration. Samples were subjected to SDS-PAGE with and without reduction, and to non-denaturing PAGE. After staining the gels were placed on a PhosphoImager screen, and the radioactivity of the different bands was measured (Fig. 1).

20 mg of *Limulus* α_2M in 50 mM Tris-HCl, pH 8.0 was treated with an equimolar amount of trypsin for 5 min at room temperature before treatment with 1 mM DCI. The thiol groups appearing from thiol ester cleavage were labelled with 70 μCi iodo[^{14}C]acetic acid for 1 h, and any unreacted thiol groups were blocked by treatment with 10 mM unlabelled iodoacetic acid for 15 min. Excess reagents were removed by gel filtration on a Sephadex G-25 SF column in 10% formic acid. After lyophilisation, the protein was denatured and reduced by heating to 60°C in 6 M guanidinium chloride, 30 mM dithioerythritol (DTE), 300 mM Tris-HCl, pH 8.5 for 15 min and the thiol groups were blocked by treatment with 100 mM iodoacetamide. As a control, 2 mg of methylamine-treated *Limulus* α_2M was labelled, reduced and carboxamidomethylated as above. The proteins from both preparations were digested with 1:10 w/w trypsin overnight at 37°C after 10-fold dilution with 50 mM Tris-HCl, 10 mM $CaCl_2$, pH 8.0. After lyophilisation, the peptides were dissolved in 0.1% trifluoroacetic acid (TFA) and loaded on a Nucleosil C-18 column (4×250 mm) equilibrated with the same buffer. The peptides were eluted at a flow rate of 1.0 ml/min with a linear gradient (Fig. 2) formed by 0.1% TFA (A) and 90% acetonitrile containing 0.085% TFA (B) (10–60% B) over 2 h at 50°C. Radioactive fractions were further purified by cation exchange chromatography and RP-HPLC.

For cation exchange a Mono S column equilibrated and eluted at a flow rate of 0.7 ml/min with the solvent system of [27] was used (5 mM H_3PO_4, 25% acetonitrile in water, eluted with a stepwise linear gradient between 10 and 200 mM NaCl over 20 min and between 200 and 350 mM NaCl over 10 min). Finally, the peptides were purified on a Vydac C-18 column (2×220 mm) operated at a flow rate of 0.2 ml/min using the aqueous acetonitrile system described above (0–20% B or 10–30% B over 38 min at 50°C). The identity of the labelled peptides was determined by N-terminal sequence analysis.

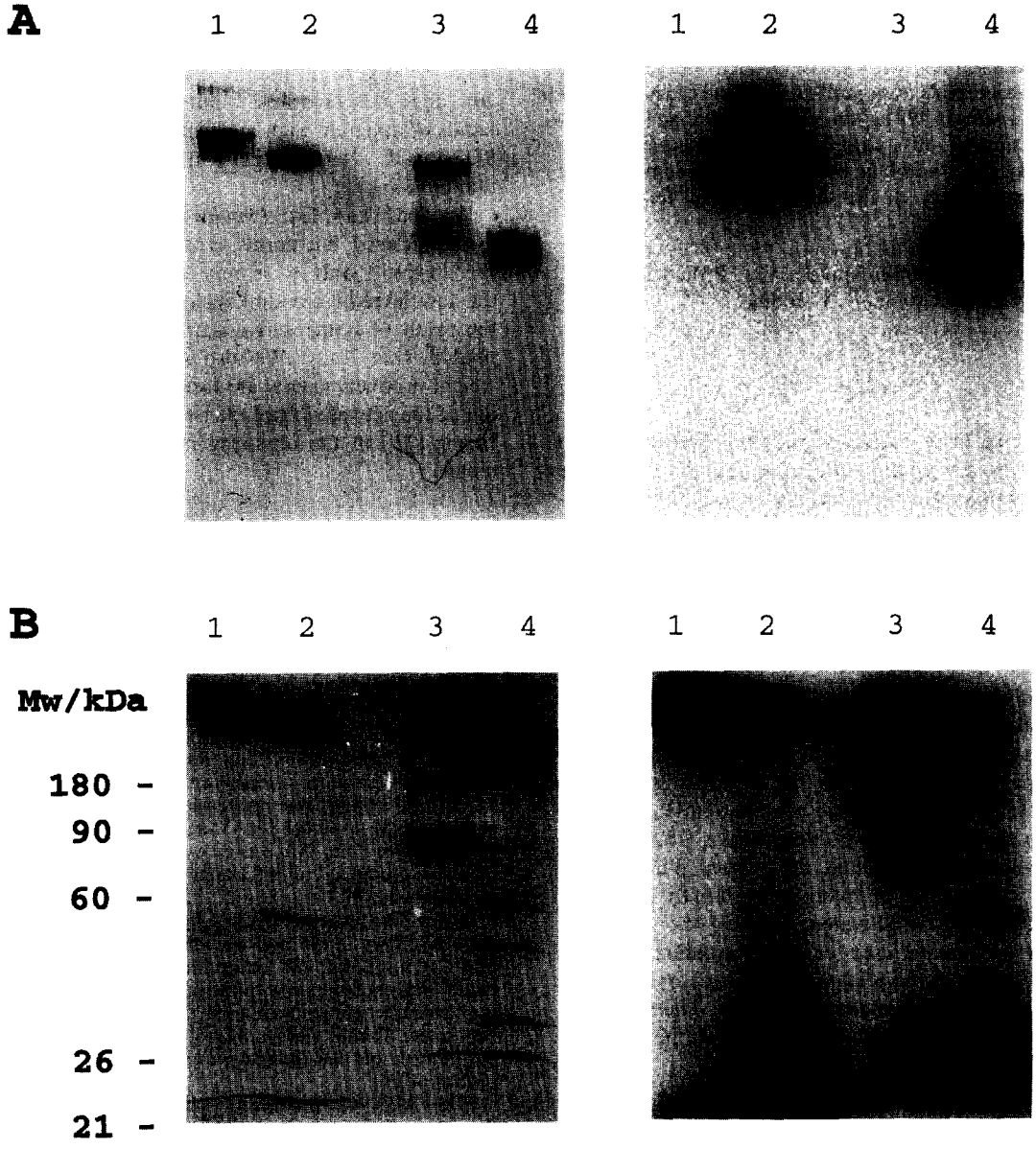

Fig. 1. Nondenaturing PAGE and SDS-PAGE analysis of native and ^{125}I-trypsin-treated human and *Limulus* α_2M. (A) Coomassie brilliant blue-stained nondenaturing gel (left) and the corresponding autoradiogram (right). Lanes: 1,2, native and ^{125}I-trypsin-reacted human α_2M; 3,4, native and ^{125}I-trypsin-reacted *Limulus* α_2M, respectively. 2.5 μg of protein was used per lane. (B) Coomassie brilliant blue-stained SDS-PAGE gel (left) and the corresponding autoradiogram (right) of ^{125}I-trypsin-α_2M complexes. Lanes: 1,2, nonreduced human and *Limulus* α_2M; 3,4, reduced human and *Limulus* α_2M, respectively. 5 μg of protein was used per lane.

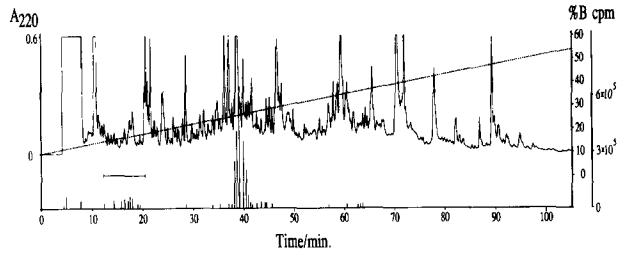

Fig. 2. Elution profile of the tryptic peptides from reduced and [^{14}C]carboxymethylated *Limulus* α_2M-trypsin complex. The peptides were separated on a Nucleosil C-18 column eluted with a linear gradient of acetonitrile in 0.1% TFA (dashed line). The number of cpm in each fraction is indicated by vertical bars beneath the chromatogram. The level of radioactivity in fractions containing less than 1.5×10^4 cpm is not indicated. The radioactive peptides from the methylamine-treated control digest eluted in the region shown by the horizontal bar.

3. Results and discussion

The binding of [^{125}I]trypsin to *Limulus* α_2M and human α_2M as assessed by non-denaturing PAGE is shown in Fig. 1A. Clearly, both proteins bind trypsin efficiently. In addition, both *Limulus* and human α_2M experience the formation of covalent isopeptide crosslinking following reaction with proteinase, seen as protein bands positioned above the unreacted monomer in reducing SDS-PAGE [21]. With human α_2M, these crosslinked bands contain most of the α_2M-bound trypsin, indicative of the efficient covalent crosslinking of α_2M and trypsin [14] (Fig. 1). *Limulus* α_2M in contrast contains very little covalently bound trypsin (Fig. 1B, lane 4), indicating that the majority of crosslinks of proteinase-reacted *Limulus* α_2M must be within the α_2M molecule itself, since this is the only other protein in the system. Approx. 10% of the radioactivity was found in bands corresponding to species larger than 30 kDa in reducing SDS-PAGE of proteinase-reacted *Limulus* α_2M. In contrast, 90% of the radioactivity was found in such species from human α_2M.

Since methylamine-reacted *Limulus* α_2M failed to form high molecular weight products during treatment with trypsin, the activated thiol ester, as in human α_2M, is a participant in the crosslinking [21]. Hence, the peptides from a tryptic digest of *Limulus* α_2M-proteinase complexes containing isopeptide crosslinks will by necessity contain the thiol ester sequence. As the cleavage of the thiol ester by proteolytic attack in the bait region or directly with methylamine leads to the appearance of a thiol group, the strategy for identifying the tryptic peptides containing the crosslinks was to label the freed thiol with iodo[^{14}C]acetic acid to be able to identify the fractions containing the thiol ester peptide during their purification [14]. Sequence analysis of the peptides crosslinked to the thiol ester peptide would reveal the localisation of the crosslink. In order to isolate the peptides of interest, the crude mixture was separated by RP-HPLC on a Nucleosil C-18 column (Fig. 2), and the radioactive fractions recovered. In some cases radioactive fractions were subjected to strong cation exchange on a Mono S column, and finally all fractions were purified by microbore RP-HPLC on a Vydac C-18 column. The identity of the peptide or set of peptides was established by N-terminal sequencing as shown in Fig. 3 and comparison to the complete sequence of *Limulus* α_2M [22].

As seen from Fig. 2, the radiolabelled peptides from the *Limulus* α_2M-trypsin complexes eluted in two major groups. The first group of peptides eluted at approx. 15–20% solvent B, similar to the radiolabelled peptides generated from methylamine-treated *Limulus* α_2M (Fig. 2, horizontal bar). The second group of peptides eluted at 26–29% solvent B.

In the first group the labelled peptides contained only the thiol ester sequence L995-PTGCGEQNMIK, with Gln-1002 [22] being recovered as Glu. As reported earlier [14], this is presumably due to hydrolysis of a fraction of the activated thiol esters. In the peptides isolated from methylamine-treated *Limulus* α_2M Gln-1002 was recovered as Glu-methylamine due to thiol ester cleavage by methylamine.

The multiple peptides eluting in the second group all contained the thiol ester sequence L995-PTGCGEQNMIK and the sequence Y250-NWEKEGVPVIHK in equimolar yields. Upon sequence analysis of these peptides, no PTH-Lys was present in cycle 5, and in cycle 8 where the thiol esterified Gln-1002 is located, the characteristic bis-PTH Glu-Lys [14] was seen in high yield. Hence, all peptides in the second group contained an ε-Lys-γ-Glu crosslink formed from Gln-1002 and Lys-254.

For some peptides sequenced, including those that represented hydrolysis products and those that represented crosslinked products, the level of PTH-Met in cycle 10 was much lower than expected, and in some cases no PTH-Met was seen. Further, in the sequence Y250NWEKEGVPVIHK the yield of PTH-Trp in cycle 3 was generally very low. Hence, the recovery of the peptides 995–1006 and 995–1006 crosslinked to 250–262 in multiple fractions is most likely due to partial oxidation of Met-1004 and Trp-252.

In the peptide separation obtained from *Limulus* α_2M-trypsin complexes (Fig. 2) approx. 10% of the radioactivity was present in the peptides containing the hydrolysed thiol ester, approx. 60% was present in the fraction containing the crosslinked peptides, with the remainder approx. 30% distributed in numerous minor fractions. At least some of these fractions may contain crosslinked peptides involving crosslinks of the thiol esterified glutamine with other lysine residues of α_2M and/or with lysine residues of trypsin (Fig. 1B). However, due to the complexity of the digest and the relatively low amount of radioactivity in these fractions, no attempt was made at their characterisation.

Unlike human α_2M [14], the association between *Limulus* α_2M and bound proteinase is predominantly non-covalent. The thiol esterified Gln-1002 reacts with high efficiency to form covalent crosslinks within the *Limulus* α_2M dimer,

Peptides, eluting at approx. 15-20 %B:

Leu-Pro-Thr-Gly-CMCys-Gly-Glu-Glu[a]-Asn-Met-Ile-Lys (995-1006)

Peptides (crosslinked), eluting at approx. 26-29 %B:

Leu-Pro-Thr-Gly-CMCys-Gly-Glu-Xaa[b]-Asn-Met-Ile-Lys (995-1006)

Tyr-Asn-Trp-Glu-Xaa[c]-Glu-Gly-Val-Pro-Val-Ile-His-Lys (250-262)

Fig. 3. Sequences in *Limulus* α_2M-trypsin complexes identified from ^{14}C-labelled fractions (Fig. 2) containing the major reaction products of the thiol esters. [a]PTH-Glu-NH_2CH_3 or PTH-Glu was seen in this step in the peptides from methylamine-treated *Limulus* α_2M and *Limulus* α_2M-trypsin complexes, respectively; [b]bis-PTH-Glu-Lys was seen in this step, consistent with crosslinking of Gln-1002 and Lys-254; [c]no signal was seen in this step, consistent with the proposed crosslinking.

rather than crosslinking the bound proteinase. The results of the present investigation clearly identify a major role for Lys-254 in the self-crosslinking of *Limulus* α_2M occurring as a result of complex formation with trypsin. However, it cannot be excluded that other residues play a minor role. It has been determined earlier that *Limulus* α_2M-proteinase complexes are dimeric [21]. Hence, the crosslinking reaction identified here must engage one thiol ester group and one Lys-254 residue located on opposing monomers of the disulphide bridged dimer. In order for the crosslinking reaction to proceed efficiently there must be a close spatial relationship between the structural elements bearing the involved residues.

The data presented here reveal a previously undescribed role for the thiol ester. Unlike other thiol ester proteins, that form intermolecular bonds through the γ-carbonyl group of the thiol ester, *Limulus* α_2M evidently forms crosslinks to itself. By comparing the stretch containing Lys-254 of *Limulus* α_2M [22] with the corresponding stretches of other α-macroglobulins, it is found that the sequences are generally highly variable, but intriguingly, none of the other α-macroglobulins contain the equivalent of Lys-254. Hence, the formation of specific intramolecular crosslinks in *Limulus* α_2M when reacted with a proteinase is likely to represent an alternative mechanism for stabilising the complex, fulfilling the same functional role as the intermolecular crosslinking of α-macroglobulins in other species [28].

This is supported by recent results showing that injected proteinase or *Limulus* α_2M-proteinase complexes are cleared in 10–30 min in vivo [29] which is considerably faster than the time it takes to run a non-denaturing gel (about 3 h), indicating that the complex is sufficiently stable to be cleared from circulation without leaking the non-covalently trapped proteinase.

Acknowledgements: Lene Kristensen is thanked for excellent technical assistance. This work was supported by the Danish Biomembrane Centre, the NOVO Nordic foundation, and grant no. MCB9218460 from the National Science Foundation. K.D. received a pre-doctoral training grant from the Danish Science Research Council.

References

[1] Sottrup-Jensen, L., Stepanik, T.M., Kristensen, T., Lønblad, P.B., Jones, C.M., Wierzbicki, D.M., Magnusson, S., Domdey, H., Wetsel, R.A., Lundwall, Å.B., Tack, B.F. and Fey, G.H. (1985) Proc. Natl. Acad. Sci. USA 82, 9–13.
[2] Sottrup-Jensen, L. (1989) J. Biol. Chem. 264, 11539–11542.
[3] Armstrong, P.B. and Quigley, J.P. (1996) Prog. Mol. Subcell. Biol. 15, 101–130.
[4] Barrett, A.J. and Starkey, P.M. (1973) Biochem. J. 133, 709–724.
[5] Van Leuven, F., Marijnen, P., Sottrup-Jensen, L., Cassiman, J.-J. and Van den Berghe, H. (1986) J. Biol. Chem. 261, 11369–11373.
[6] Andersen, G.R., Koch, T.J., Dolmer, K., Sottrup-Jensen, L. and Nyborg, J. (1995) J. Biol. Chem. 270, 25133–25141.
[7] Herz, J., Hamann, U, Rogne, S., Myklebost, O., Gausepohl, H. and Stanley, K.K. (1988) EMBO J. 7, 4119–4127.
[8] Kristensen, T., Moestrup, S.K., Gliemann, J., Bendtsen, L., Sand, O. and Sottrup-Jensen, L. (1990) FEBS Lett. 276, 151–155.
[9] Strickland, D.K., Ashcom, J.D., Williams, S., Burgess, W.H., Migliorini, M. and Argraves, W.S. (1990) J. Biol. Chem. 265, 17401–17404.
[10] Holtet, T.L., Nielsen, K.L., Etzerodt, M., Moestrup, S.K., Gliemann, J., Sottrup-Jensen, L. and Thøgersen, H.C. (1994) FEBS Lett. 344, 242–246.
[11] Imber, M.J. and Pizzo, S.V. (1981) J. Biol. Chem. 256, 8134–8139.
[12] Van Leuven, F., Cassiman, J.-J. and Van den Berghe, H. (1981) J. Biol. Chem. 256, 9016–9022.
[13] Sottrup-Jensen, L., Petersen, T.E. and Magnusson, S. (1980) FEBS Lett. 121, 275–280.
[14] Sottrup-Jensen, L., Hansen, H.F., Pedersen, H.S. and Kristensen, L. (1990) J. Biol. Chem. 265, 17727–17737.
[15] Jacobsen, L. and Sottrup-Jensen, L. (1993) Biochemistry 32, 120–126.
[16] Van Leuven, F., Cassiman, J.-J. and Van den Berghe, H. (1982) Biochem. J. 201, 119–128.
[17] Quigley, J.P. and Armstrong, P.B. (1983) J. Biol. Chem. 258, 7903–7906.
[18] Quigley, J.P. and Armstrong, P.B. (1985) J. Biol. Chem. 260, 12715–12719.
[19] Enghild, J.J., Thøgersen, I.B., Salvesen, G., Fey, G.H., Figler, N.L., Gonias, S.L. and Pizzo, S.V. (1990) Biochemistry 29, 10070–10080.
[20] Armstrong, P.B., Mangel, W.F., Wall, J.S., Hainfeld, J.F., Van Holde, K.E., Ikai, A. and Quigley, J.P. (1991) J. Biol. Chem. 266, 2526–2530.
[21] Quigley, J.P., Ikai, A., Arakawa, H., Osada, T. and Armstrong, P.B. (1991) J. Biol. Chem. 266, 19426–19431.
[22] Iwaki, D., Kawabata, S., Miura, Y., Kato, A., Armstrong, P.B., Quigley, J.P., Sottrup-Jensen, L., Dolmer, K., Nielsen, K.L. and Iwanaga, S., manuscript in preparation.
[23] Sottrup-Jensen, L., Petersen, T.E. and Magnusson, S. (1980) FEBS Lett. 121, 275–279.
[24] Armstrong, P.B., Melchior, R. and Quigley, J.P. (1996) J. Insect. Physiol. 42, 53–64.
[25] Sottrup-Jensen, L. (1995) Anal. Biochem. 225, 187–188.
[26] Jensen, P.H., Christensen, E.I., Ebbesen, P., Gliemann, J. and Andreasen, P.A. (1990) Cell Regul. 1, 1043–1056.
[27] Crimmins, D.L., Thoma, R.S., McCourt, D.W. and Schwartz, B.D. (1989) Anal. Biochem. 176, 255–260.
[28] Sottrup-Jensen, L. (1994) Ann. NY Acad. Sci. 737, 172–187.
[29] Melchior, R., Quigley, J.P. and Armstrong, P.B. (1995) J. Biol. Chem. 270, 13496–13502.