Volume 199, number 2 FEBS 3577 April 1986

Amino acid sequence of proteinase K from the mold *Tritirachium album* Limber

Proteinase K – a subtilisin-related enzyme with disulfide bonds

Klaus-Dieter Jany, Georg Lederer and Barbara Mayer

Institut für Organische Chemie, Biochemie und Isotopenforschung der Universität Stuttgart, Pfaffenwaldring 55, D-7000 Stuttgart 80, FRG

Received 20 February 1986

The amino acid sequence of proteinase K (EC 3.4.21.14) from *Tritirachium album* Limber has been determined by analysis of fragments generated by cleavage with CNBr or BNPS-skatole. The enzyme consists of a single peptide chain containing 277 amino acid residues, corresponding to M_r 28 930. Comparison of the sequence with those of the serine proteinases reveals a high degree of homology (about 35%) to the subtilisin-related enzyme. But in contrast to the subtilisins, proteinase K contains 2 disulfide bonds and a free cysteine residue. This finding may indicate that proteinase K is a member of a new subfamily of the subtilisins.

Proteinase K *Subtilisin* *Serine proteinase* *Cystine-containing subtilisin-like proteinase* *Evolution*

1. INTRODUCTION

Proteinase K is an extracellular endopeptidase which is synthesized by the mold *Tritirachium album* Limber. This fungal proteinase belongs to the class of the serine endopeptidases [1]. This class can be divided into the trypsin- and subtilisin-related enzymes. The two families are unrelated to each other [2] but they have evolved independently the same catalytic mechanism [3]. Preliminary sequence determination [4] of proteinase K revealed that the structure around the active site serine residue is homologous to the corresponding ones of the subtilisins. However, all known subtilisins lack disulfide bridges, whereas proteinase K and all trypsin-related enzymes contain such bonds [4,5]. To obtain more insight into the evolutionary relationship of proteinase K, we have determined its amino acid sequence.

Dedicated to Professor Dr Gerhard Pfleiderer on the occasion of his 65th birthday

In this paper the complete covalent structure of proteinase K is presented and its evolutionary relationship to the subtilisin-related enzymes is discussed. This work has provided not only information for completion of the sequence but also for the determination of the three-dimensional structure [6,7] and isolation of the gene (Gassen, H.G., unpublished).

2. MATERIALS AND METHODS

2.1. *Materials*

Proteinase K was a generous gift from Merck, Darmstadt and was obtained in a crystalline form. Iodo[^{14}C]acetic acid and [^{3}H]diisopropyl fluorophosphate (DFP) were purchased from Amersham. Carbobenzoxy-Ala-[^{14}C]Ala chloromethyl ketone was synthesized according to Shaw [8] with some small modifications (Z-Ala-[^{14}C]Ala-CK, spec. act. 0.12 mCi/mmol). The sources of all other reagents are given in the report on the structure of proteinase K [4].

Published by Elsevier Science Publishers B.V. (Biomedical Division)

2.2. *Generation of the peptides*

Prior to the cleavages, proteinase K was either inactivated by [^{3}H]DFP or Z-Ala-[^{14}C]Ala-CK. The inactive protein species were reduced and *S*-carboxymethylated with iodo[^{14}C]acetic acid as described [4]. The labelled proteinase K was rechromatographed on Fracto-gel TSK-HW 50 equilibrated with 30% acetic acid in order to remove autolysis products. CNBr cleavage of the [^{3}H]DFP-labelled proteinase K and the initial separation of the fragments were carried out as described [4]. After purification of 7 CNBr-peptides, 6 fragments were subdigested with trypsin or chymotrypsin in 50 mM NH_4HCO_3 (pH 7.8) at enzyme to peptide ratios of 1:100 to 1:30 depending on the size and solubility of the fragments.

Cleavage of the Z-Ala-[^{14}C]Ala-CK-treated proteinase K at the Trp-X peptide bonds was performed with BNPS-skatole as described in [9]. The resulting fragments were recovered by gel filtration on Sephadex G-100 equilibrated in 30% acetic acid. The two large fragments (BS-2 and BS-3) were treated with citraconic anhydride [10] and afterwards hydrolyzed at the arginine residues with trypsin (50:1, w/w; pH 8.4; 3 h).

Some additional information and overlaps were obtained from subdigestions of BS-2 and BS-3 with *Staphylococcus aureus* V8 protease (Miles) and thermolysin (Boehringer, Mannheim). Digestion with the V8 protease was carried out in 50 mM NH_4HCO_3 (pH 7.8) at a peptide/enzyme ratio of 30:1 (w/w) for 20 h at room temperature. Digestion with thermolysin (1%, w/w) was performed in 0.1 M *N*-ethylmorpholinium acetate (pH 8.0) for 45 min at 55°C.

2.3. *Peptide purification*

Peptides derived from the various digestions were prefractionated by exclusion chromatography on Sephadex G-100, G-50, G-25 and/or Fracto-gel TKS-HW 40, TKS-HW 50 columns equilibrated and eluted with either 25 mM NH_4HCO_3 (pH 8.5) or 30% acetic acid. Most of the peptides were purified further (i) by ion-exchange chromatography on a Chromobeads-P resin using pyridine acetate gradients [9] and more recently (ii) by HPLC on reverse-phase columns (Whatman ODS 3; Partisil 5 C8, 0.4 × 25 cm) using acetonitrile gradients (0–70%) in either 0.1% trifluoroacetic acid or 20 mM NH_4HCO_3 (pH 7.5).

2.4. *Amino acid composition and sequence analysis*

Protein and peptide samples were hydrolyzed in 6 N HCl or, for tryptophan determination in 4 M methanesulfonic acid [11], at 108°C in evacuated tubes for 22–96 h. The compositions of the hydrolysates were determined on a Biotronik LC 6000 analyzer. Sequencing of the peptides was performed manually using the 4-(dimethylamino)-azobenzene-4-isothiocyanate/phenylisothiocyanate double coupling methods [12] and/or the dansyl-Edman procedure [13]. Some C-terminal sequences were determined by kinetic analysis using the carboxypeptidases A, B or Y [14].

2.5. *Disc electrophoresis*

Gel electrophoresis was carried out on 15% gels at pH 4.5 by the method of Davis [15]. SDS-PAGE in 15% gels was performed according to Laemmli [16], except that the gels contained 6 M urea.

3. RESULTS AND DISCUSSION

The commercially available crystalline proteinase K appears to be homogeneous when examined by PAGE on 15% gels at pH 4.3. Also, when using proteinase K labelled at the active site serine residue with [^{3}H]DFP, the whole radioactivity coincided with the protein band (fig.1B). However, in the SDS-PAGE it is evident that the enzyme preparation contains impurities which are probably autolysis products. Electrophoresis of the [^{3}H]DFP-treated proteinase K revealed that most of the radioactivity is focussed on the high-M_r protein band, but activity is also found in the lower-M_r protein bands (fig.1A). This may indicate that the active proteinase K consists of an unhydrolysed protein chain and autolysed polypeptide chains. Most of these autolysis products can be removed by gel filtration after *S*-carboxymethylation of the protein. The starting material (CM-DFP- or CM-Z-Ala-Ala-CK-modified proteinase K) gave essentially a single protein band in the SDS-PAGE corresponding to an M_r of 28000–29000.

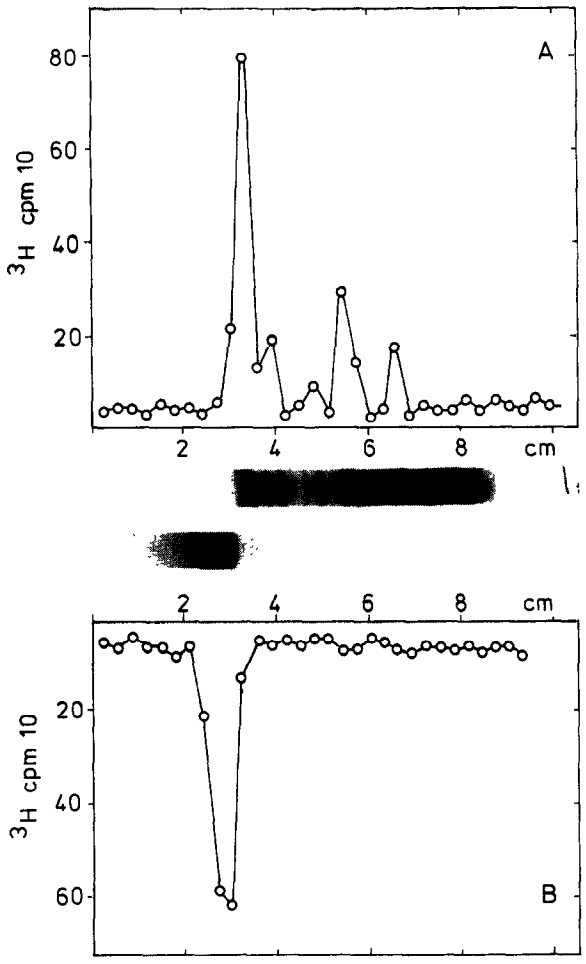

Fig.1. Polyacrylamide gel electrophoresis of [^{3}H]DFP inactivated proteinase K at pH 4.3 (A) and in the presence of SDS (B). 100 μg protein were applied onto the gels. In the middle, the protein pattern is shown and the distribution of the radioactivity along the gels is given in the diagrams.

The strategy employed for determining the sequence is schematically outlined in fig.2. Prior to fragmentation, the enzyme was either inactivated with [^{3}H]DFP or Z-Ala-[^{14}C]Ala-CK which facilitated the identification of the catalytic serine and histidine residue, respectively. Cleavage by CNBr of the DFP inactivated enzyme resulted in 7 fragments including an overlapping peptide derived from an incomplete cleavage of the Met-237–Thr-238 bond. The structure of these fragments was determined separately by a combination of direct analysis and the sequence of

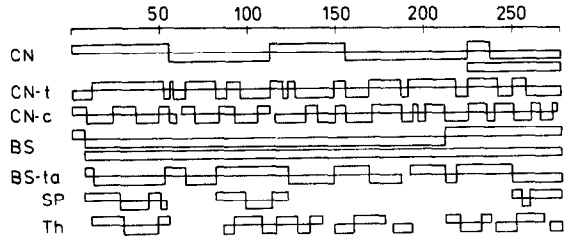

Fig.2. Summary of sequence evidence. Bars represent analysed fragments and the length of each bar is proportional to the size of that peptide. Cyanogen bromide fragments (CN) and their tryptic (CN-t) and chymotryptic (CN-c) subfragments. Peptides generated by treatment with BNPS-skatole (BS) and their arginine peptides (BS-ta). Peptides after digestion of BS-2 and BS-3 with *S. aureus* V8 protease (SP) and thermolysin (Th).

overlapping peptides obtained from tryptic and chymotryptic subdigests. Cleavage of the Z-Ala-[^{14}C]Ala-CK-inhibited proteinase K with BNPS-skatole yielded the 3 expected peptides (BS-1–3) and in addition a fragment derived from the incomplete cleavage at Trp-211. The placement of the CNBr-fragments was achieved by isolation and sequencing of the peptides derived from the citraconylated BS-2 and BS-3 after separate digestion with trypsin. Pairing of the half-cysteine residues was examined from the tryptic fragments after modifications of the single free cysteine residue in the proteinase K with 4-dimethylaminobenzene-4-iodoacetamide [17]. A detailed description will be given elsewhere.

The results accounted for the identification of all residues and the unequivocal placement of all fragments. The covalent structure of proteinase K is shown in fig.3. The molecule contains 277 amino acid residues including 2 disulfide bonds (34–124, 179–248) and a free cysteine (73) in a single polypeptide chain. The M_r of the protein is 28930, which is in excellent agreement with the previous determinations of 28500 based on SDS-PAGE and active site titration [4]. These results demonstrate clearly that the M_r of 18500 estimated previously [1] was erroneous. The amino acid composition calculated from the sequence is in agreement with the results of previous amino acid analyses [4].

As active site residues were identified His-69 and Ser-223. The sequences around both residues are typical to the corresponding ones of the subtilisins.

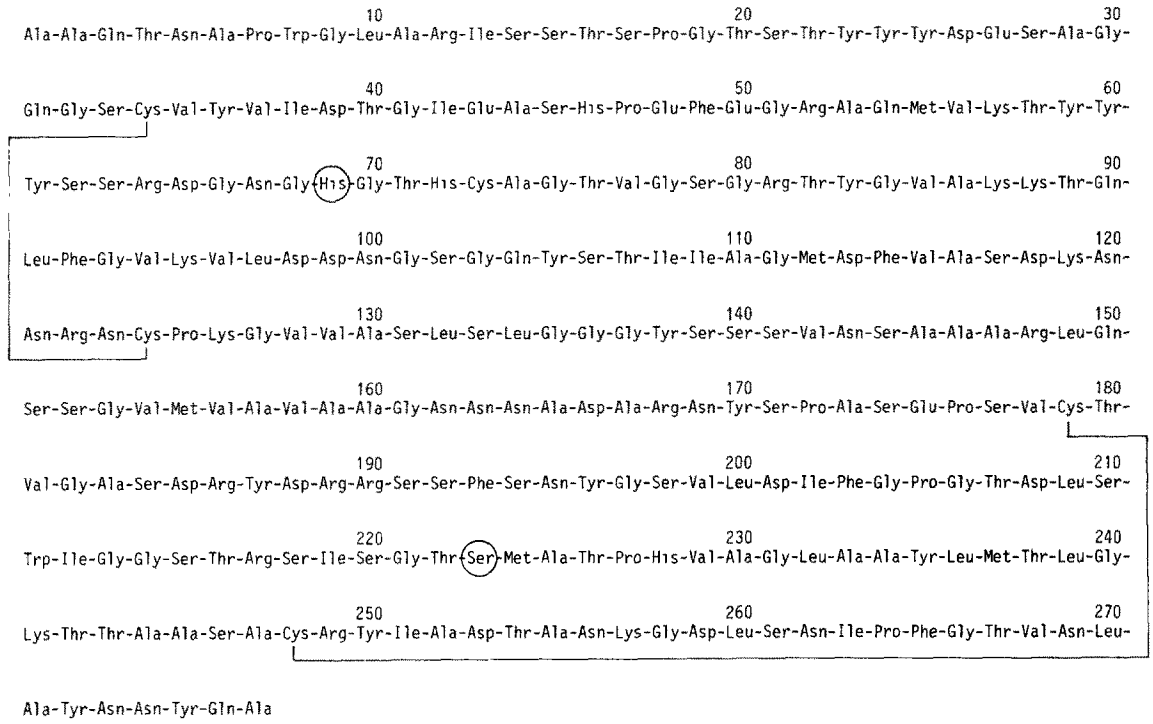

Fig.3. Amino acid sequence of proteinase K from *T. album* Limber. The active site residues are encircled. The disulfide bonds are formed by Cys-34–Cys-124 and Cys-179–Cys-248.

When the sequence of proteinase K is aligned to those of subtilisin novo, carlsberg, DY [18] and thermitase [19] (fig.4) a homologous relationship is evident to the bacterial endopeptidase. The fungal proteinase K has about 35% sequence identity to the subtilisins and 44% to thermitase. The active site residues His-69 and Ser-223 correspond to His-64 and Ser-221 in the subtilisins and in addition Asp-38 (Asp-32) should form the charge relay system in proteinase K, as it is reported for subtilisin novo or chymotrypsin [3]. The sequences around these residues are highly conserved in all subtilisins and proteinase K shares the highest homology in these stretches. However, the homology is not only extended to the catalytic essential residues, but also to regions which are known to be involved in the substrate binding. The sequences Ser-133–Gly-135 and Ala-159–Asn-152 correspond in the subtilisins to Ser-125–Gly-127 and Ala-152–Asn-155. In the subtilisins the sequence Ser-125–Gly-127 forms hydrogen bonds with P2 and P3 residues of the substrate and the same is also found in proteinase K from X-ray studies of the complex with Z-Ala-Ala-CK [7].

The degree of sequence homology and identity in the three-dimensional structure [6,7] indicate a relationship of the fungal proteinase K to the bacterial subtilisins. Proteinase K belongs to the family of the subtilisin-related enzymes. It is likely that these enzymes have evolved from a common ancestral precursor serine proteinase. However, there is a distinct difference between the typical subtilisins and proteinase K. Proteinase K contains 2 disulfide bonds, whereas subtilisins do not possess such bonds. In homologous proteins amino acid replacements occur more frequently than changes in the number of disulfide bonds [2]. The acquirement of new disulfide bonds is a rather rare event as it requires two independent mutations and at best simultaneously. Therefore we assume that two progenitors have diverged from the ancestral proteinase separating the family of the subtilisin-related enzyme into two subclasses. One of these encloses the cysteine-free subtilisins like subtilisin novo, carlsberg or DY. The other subclass is represented by the cysteine-containing subtilisins like proteinase K and thermitase [19].

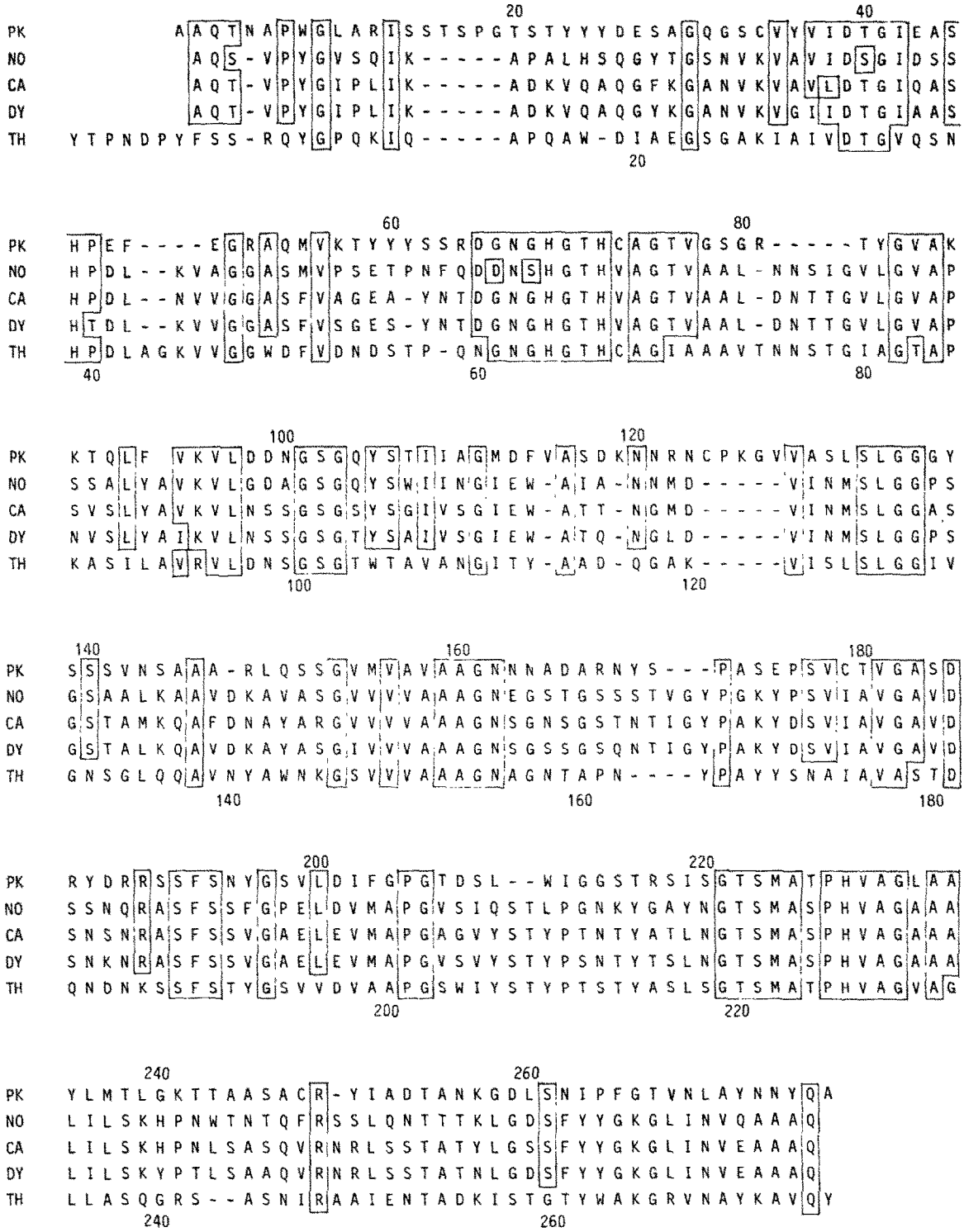

Fig.4. Comparison of the sequence of proteinase K (PK) with those of subtilisin novo (NO), subtilisin carlsberg (CA), subtilisin DY (DY) and thermitase (TH) [19]. The sequences of the subtilisins are taken from [18]. Top: proteinase K numbering; bottom: subtilisin carlsberg numbering.

ACKNOWLEDGEMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (Ja 308/6). We thank E. Nitsche and K. Freimüller for excellent technical assistance.

REFERENCES

[1] Ebeling, W., Hennrich, N., Klockow, M., Metz, H., Orth, H.D. and Lang, H. (1974) Eur. J. Biochem. 47, 91–97.
[2] DeHaën, C., Neurath, H. and Teller, D. (1975) J. Mol. Biol. 92, 225–259.
[3] Robertus, J.D., Alden, R.A., Birktoft, J.J., Kraut, J., Powers, J.C. and Wilcox, P.E. (1972) Biochemistry 11, 2430–2449.
[4] Jany, K.D. and Mayer, B. (1985) Biol. Chem. Hoppe-Seyler 366, 485–492.
[5] Dattagupta, J.K., Fukuwara, T., Grishin, E.V., Lindner, K., Manor, P.C., Pienziazek, J.J., Saenger, W. and Suck, P. (1975) J. Mol. Biol. 97, 267–271.
[6] Pähler, A., Banerjee, A., Dattagupta, J.K., Fujiwara, T., Lindner, K., Pal, G.P., Suck, D., Weber, G. and Saenger, W. (1984) EMBO J. 3, 1311–1313.
[7] Betzel, C., Pal, G.P., Struck, M., Jany, K.D. and Saenger, W. (1986) FEBS Lett. 197, 111–114.
[8] Shaw, E. (1967) Methods Enzymol. 11, 677–686.
[9] Fröschle, M., Ulmer, W. and Jany, K.D. (1984) Eur. J. Biochem. 142, 533–540.
[10] Atassi, N.Z. and Habeeb, A.F.S.A. (1972) Methods Enzymol. 45, 546–553.
[11] Penke, B., Ferezi, R. and Kavacs, K. (1974) Anal. Biochem. 60, 45–50.
[12] Wittmann-Liebold, B. and Lehmann, F. (1980) in: Methods in Peptide and Protein Sequence Analysis (Birr, C. ed.) pp.49–79, Elsevier, Amsterdam, New York.
[13] Chen, R. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 873–886.
[14] Ambler, R.P.A. (1967) Methods Enzymol. 11, 155–166.
[15] Reisfeld, R.A., Lewis, U.I. and Williams, D.E. (1962) Nature 195, 281–283.
[16] Laemmli, U.K. (1970) Nature 227, 680–685.
[17] Chang, J.Y., Knecht, R. and Braun, D.G. (1983) Biochem. J. 211, 163–171.
[18] Nedkov, P., Oberthür, W. and Braunitzer, G. (1985) Biol. Chem. Hoppe-Seyler 366, 421–430.
[19] Meloun, B., Bandys, M., Kostka, V., Hausdorf, G., Frömmel, C. and Höhne, W.E. (1985) FEBS Lett. 183, 195–200.