The amino acid sequence of the enterotoxin from *Clostridium perfringens* type A

Michael Richardson and Per Einar Granum+

Department of Botany, University of Durham, Durham DH1 3LE, England and +*Norwegian Food Research Institute, N-1432 Ås-NLH, Norway*

Received 3 January 1985

The amino acid sequence of the enterotoxin from *Clostridium perfringens* type A was determined by analysis of peptides derived from the protein by digestion with trypsin chymotrypsin, thermolysin, pepsin, a lysine-specific protease, *S. aureus* V8 protease and a proline-specific protease, and fragments generated by cleavage with cyanogen bromide or by dilute acetic acid in 7 M guanidine HCl. The sequence which is complete except for the definite order of 3 small peptides between residues 88 and 103 consists of 309 amino acids and contains a correction to our preliminary announcement [(1984) FEMS Symp. 24, 329–330]

Clostridium perfringens *Enterotoxin* *Amino acid sequence*

1. INTRODUCTION

The enterotoxin from *Clostridium perfringens* type A is responsible for many cases of a mild type of food poisoning [1]. It is produced in the intestinal tract during sporulation, after ingestion of 10^8–10^9 cells [2].

The enterotoxin is a single polypeptide of M_r ~34000 [3] and contains about 43% hydrophobic amino acids. This toxic protein contains 17 lysine residues, and modification of one, or two at most, of the amino groups results in loss of biological activity without conformational changes [4]. We have previously reported the primary sequence of the first 66 amino acids at the N-terminal end of the molecule and have identified the two peptide bonds hydrolyzed during trypsination of the enterotoxin [5]. The trypsination results in a 3-fold increase in biological activity of the protein [6]. The total amino acid sequence of the enterotoxin was presented in a preliminary report [7], but then a part of the sequence was misplaced.

Here the total sequence of the enterotoxin is reported (except for the definite order of 3 small peptides between amino acids 88 and 103).

2. MATERIALS AND METHODS

2.1. *Protein purification*

Production and purification of the enterotoxin was done as described by Granum and Whitaker [8]. Biological activity was tested by the Vero cell culture technique [9].

2.2. *Reduction and* S-*carboxymethylation*

The native enterotoxin was reduced and *S*-carboxymethylated in 6 M guanidine HCl in 0.1 M Tris, pH 8.6, as described [10].

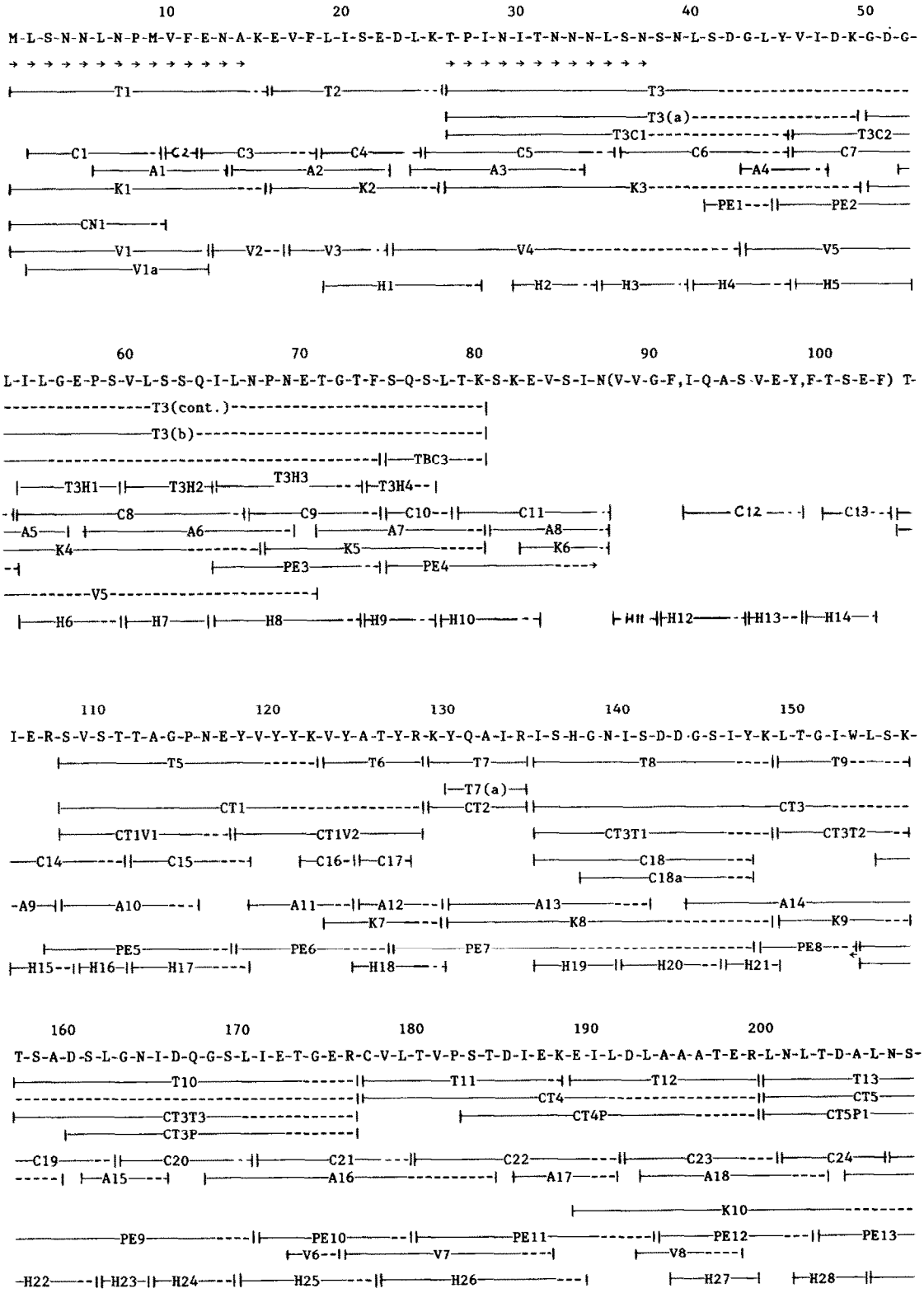
2.3. *Citraconylation*

The reduced and *S*-carboxymethylated enterotoxin (20 mg) in 10 ml of 0.2 M Na borate buffer, pH 8.5, containing 7 M urea was treated with citraconic anhydride (Pierce) as described in [11].

2.4. *Enzyme digestion*

Samples (12–16 mg) of the reduced and *S*-carboxymethylated enterotoxin in 0.2 M *N*-ethylmorpholine HCl buffer, pH 8.5, were digested separately with trypsin (TPCK-treated, Worthington) and chymotrypsin (type A_4, Boeh-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30


10 20 30 40 50
M-L-S-N-N-L-N-P-M-V-F-E-N-A-K-E-V-F-L-I-S-E-D-L-K-T-P-I-N-I-T-N-N-N-L-S-N-S-N-L-S-D-G-L-Y-V-I-D-K-G-D-G-
T1 T2 T3
T3(a)
T3C1 T3C2
C1 C2 C3 C4 C5 C6 C7
A1 A2 A3 A4
K1 K2 K3
PE1 PE2
CN1
V1 V2 V3 V4 V5
V1a
H1 H2 H3 H4 H5
60 70 80 90 100
L-I-L-G-E-P-S-V-L-S-S-Q-I-L-N-P-N-E-T-G-T-F-S-Q-S-L-T-K-S-K-E-V-S-I-N(V-V-G-F,I-Q-A-S-V-E-Y,F-T-S-E-F) T-
T3(cont.)
T3(b)
TBC3
T3H1 T3H2 T3H3 T3H4
C8 C9 C10 C11 C12 C13
A5 A6 A7 A8
K4 K5 K6
PE3 PE4
V5
H6 H7 H8 H9 H10 H11 H12 H13 H14
110 120 130 140 150
I-E-R-S-V-S-T-T-A-G-P-N-E-Y-V-Y-Y-K-V-Y-A-T-Y-R-K-Y-Q-A-I-R-I-S-H-G-N-I-S-D-D-G-S-I-Y-K-L-T-G-I-W-L-S-K-
T5 T6 T7 T8 T9
T7(a)
CT1 CT2 CT3
CT1V1 CT1V2 CT3T1 CT3T2
C14 C15 C16 C17 C18 C18a
A9 A10 A11 A12 A13 A14
K7 K8 K9
PE5 PE6 PE7 PE8
H15 H16 H17 H18 H19 H20 H21
160 170 180 190 200
T-S-A-D-S-L-G-N-I-D-Q-G-S-L-I-E-T-G-E-R-C-V-L-T-V-P-S-T-D-I-E-K-E-I-L-D-L-A-A-A-T-E-R-L-N-L-T-D-A-L-N-S-
T10 T11 T12 T13
CT4 CT5
CT3T3 CT4P CT5P1
CT3P
C19 C20 C21 C22 C23 C24
A15 A16 A17 A18
K10
PE9 PE10 PE11 PE12 PE13
V6 V7 V8
H22 H23 H24 H25 H26 H27 H28

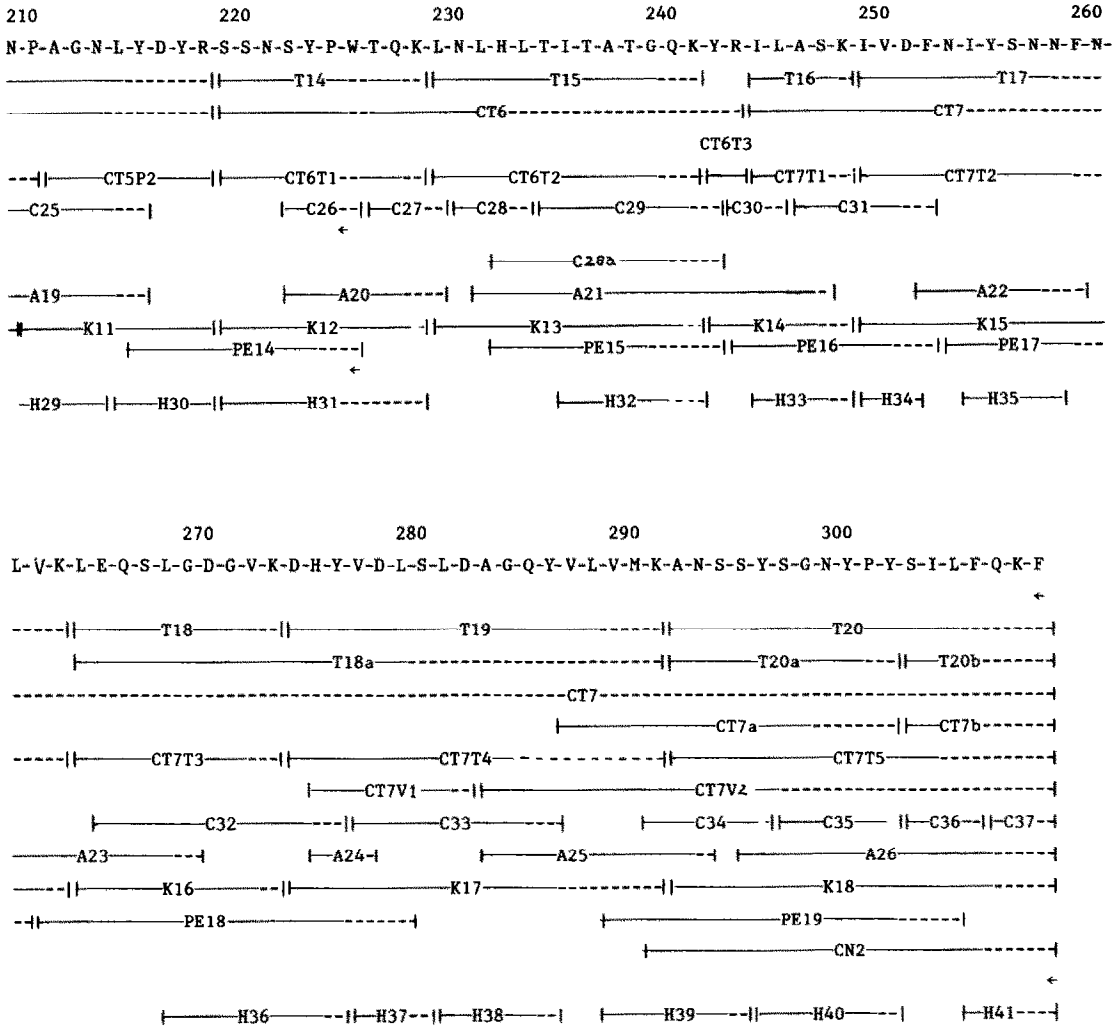


Fig.1. Amino acid sequence of the enterotoxin from *Clostridium perfringens* type A. T, tryptic peptides; CT, peptides obtained from digestion of citraconylated enterotoxin with trypsin; C, chymotryptic peptides; A, peptides obtained from hydrolysis with acetic acid in 7 M guanidine HCl; K, peptides from digest with Lys-specific endoproteinase from *L. enzymogenes*; PE, peptic peptides; V, peptides obtained from digestion with protease from *S. aureus* V8; CN, fragments from cleavage with cyanogen bromide; H, thermolytic peptides; P, peptides obtained from sub-digests of large peptides with Pro-specific protease from *Flavobacterium meningosepticum*. Solid lines indicate regions of peptides sequenced by the DABITC method and/or the dansyl-Edman procedure; dashed lines indicate residues which were not sequenced or yielded unsatisfactory results; →, residues determined by the DABITC method applied to the intact enterotoxin and the native enterotoxin after trypsination [5,6]; ←, residues determined by digestion with carboxypeptidase A.

ringer Mannheim) using a 2% (w/w) enzyme/substrate ratio for 6–10 h at 37°C. The citraconylated enterotoxin (15 mg) in 0.2 M Na borate HCl buffer, pH 8.5, was digested with trypsin for 4 h at 37°C and the digestion terminated by the addition of glacial acetic acid (to pH 2.5) before lyophilization. (The specificities of the proteases employed in this investigation were all as expected except for the anomalous hydrolyses by the Lys-specific enzyme at the peptide bonds

Asn^{67}-Pro^{68}, Asn^{87}-x^{88}, Asn^{209}-Pro^{210} and Arg^{218}-Ser^{219}, by trypsin at Tyr^{286}-Val^{287} and Tyr^{302}-Ser^{303}, and by the proline-specific endoproteinase at Ala^{159}-Asp^{160}.)

Digestion with the protease from *Staphylococcus aureus* V8 (2%, w/w, Miles) was carried out in 0.1 M Na and K phosphate buffer, pH 8.0, for 21 h at 37°C, with the lysine-specific protease from *Lysobacter enzymogenes* (2%, w/w, Boehringer Mannheim) in 0.05 M Tris (pH 7.7) containing 1 mM EDTA for 6 h at 37°C, with pepsin (2%, w/w, Boehringer Mannheim) in 5% formic acid for 2 h at 27°C, and with thermolysin (2%, w/w, Daiwa Kasei KK, Japan; activated before use with 5 mM $CaCl_2$) in 0.1 M NH_4HCO_3 for 3 h at 37°C. Certain large peptides were also redigested with a proline-specific endoproteinase from *Flavobacterium meningosepticum* (2%, w/w, Seikagaku, Japan) in 0.1 M Na phosphate buffer, pH 7.0, for 2 h at 37°C. The C-terminal residues of the protein and certain peptides were determined by the use of carboxypeptidase A (type II, treated with diisopropyl fluorophosphate, Sigma) in 0.2 M NEM acetate buffer, pH 8.5, in combination with the dansyl technique as described [12].

2.5. *Chemical cleavages*

Fragments and peptides were also derived from the reduced and *S*-carboxymethylated enterotoxin by treatment with an excess (60-fold over Met residues) of CNBr in 70% (v/v) formic acid for 24 h at 20°C, and by hydrolysis with 0.25 M acetic acid in 7 M guanidine HCl at 105°C for 40 h.

2.6. *Separation of peptides*

Peptide mixtures resulting from the enzymic digests or chemical cleavage were initially fractionated on columns (1 × 200 cm) of Biogel P-30 or Sephadex G-50 (for larger peptides from tryptic digest of citraconylated protein), Biogel P-6 (200–400 mesh) or Biogel P-4 in 0.05 M ammonium bicarbonate. Fractions of 0.9–1.2 ml tube were collected at a flow rate of 9–12 ml/h, and the absorbance of each measured at 230 nm. The presence of peptides in fractions of low absorbance was also checked for by removing small samples (10–20 μl) for N-terminal analysis by the dansyl/TLC method as in [13]. The individual fractions containing peptides were pooled and lyophilized.

Each of the pooled peaks was subjected to reverse-phase HPLC on a μ-Bondapak C-18 column (0.5 × 25 cm, HPLC Technology) or a Micropac MCH-10 column (0.4 × 30 cm, Varian) in a Varian model 5000 HPC using a linear gradient of 0–70% acetonitrile (HPLC grade S, Rathburn) in 0.1% trifluoroacetic acid [14]. Peptides were detected by measuring the absorbance at 214 nm.

2.7. *Sequence methods*

Peptides derived from the various digests were subjected to micro-sequence analysis using the 4-*N*,*N*-dimethylaminoazobenzene-4′-isothiocyanate (DABITC)/phenylisothiocyanate (PITC) double-coupling method [15]. In the case of certain peptides with both Ile and Leu present in the sequences the identity of these residues was determined using the dansyl-Edman procedure as in [12].

2.8. *Amino acid composition of protein and peptides*

The amino acid composition of the enterotoxin was determined with a Biocal BC 200 automatic amino acid analyzer as described in [6]. The composition of peptide samples was determined in a Varian 5000 HPLC fitted with a Micropak Hydrolysate amino acid column (0.4 × 15 cm, Varian) in the Na form, a post-column *o*-phthalaldehyde reactor system and a Fluorichrom detector as described in [16].

3. RESULTS AND DISCUSSION

Microsequence analysis of the peptides derived from the enterotoxin by digestion with trypsin, chymotrypsin, a Lys-specific protease, pepsin, thermolysin and *S. aureus* V8 protease, together with tryptic peptides from the citraconylated protein, and fragments produced by cleavage with cyanogen bromide or 0.25 M acetic acid in 7 M guanidine HCl at 105°C enabled the sequence to be deduced as shown in fig.1.

This extensive range of fragmentation methods was required to establish the sequence owing to the extremely hydrophobic nature of the protein [3] which prevented the recovery of certain large pep-

tides, notably the expected tryptic peptide covering residues 83–107 and the tryptic fragment from the citraconylated protein covering residues 1–107. As a result of these difficulties we have been unable to establish completely the sequence of amino acids between residues 88 and 103, although chymotryptic and thermolytic peptides for this region were sequenced (fig.1).

The sequence shown in fig.1 is the same as that given in our preliminary report [7] except that the residues covered by fragment CT6 (219–243) were mistakenly inserted earlier in the sequence between residues 128 and 129. The subsequent recovery and analysis of the peptides PE14 and A21 which overlap the ends of this fragment revealed its correct positioning as shown in fig.1, and peptide PE7 clarified the sequence around residues 128 and 129.

The M_r of the enterotoxin calculated from the 309 amino acids found in the sequence is 34262, which is in excellent agreement with the previous estimates of M_r 34000 based on analytical ultracentrifugation and SDS-PAGE electrophoresis [3,17]. Also the amino acid composition calculated from the sequence (table 1) is in very good agreement with the results of previous amino acid analyses [3,18] and the results of titrimetric studies on the content of Lys, Cys [4], Tyr and amides [6].

No microheterogeneity was found in the enterotoxin indicating that the protein is coded for by a single gene. The absence of the N-terminal Met in some polypeptide chains was thought likely to be the result of post-translational processing. Other workers have reported heterogeneous forms of enterotoxin-like proteins from *C. perfringens* type A [19–21]. Comparison of the amino acid compositions of some of these proteins [20] with that of the enterotoxin sequenced here (table 1) suggests that they were derived from the native protein by proteolysis probably in the N-terminal region. Whether this is correct can easily be established by sequencing the first 3–4 amino acids at the N-termini of these proteins.

Apart from the limited homology of residues 26–31 with the N-terminal region of the B subunit of the cholera toxin [22] that we have noted previously [5], the sequence of the enterotoxin from *C. perfringens* type A appears to have no homology with any of the other toxins for which sequence data are currently available.

Eigh amino groups may be titrated by TNBS in the native enterotoxin, and one or at most two amino groups is essential for biological activity [4]. The titratable amino groups are expected to be exposed to the environment, and therefore to be found in hydrophilic parts of the enterotoxin. The 8 titratable amino groups are therefore quite probably the α-amino group on the N-terminal and the ϵ-amino groups on Lys^{15}, Lys^{25}, Lys^{49}, Lys^{80}, Lys^{82}, Lys^{188} and Lys^{273}. The α-amino group is reactive with dansyl chloride and trypsin acts on Lys^{15} and Lys^{25} on the native enterotoxin [5], so these groups are not involved in the active side of the enterotoxin. Since it is difficult to determine

Table 1

The amino acid composition of the enterotoxin from *Clostridium perfringens* Type A

Amino acid	Amino acid residues	
	Analysis[a]	Sequence
Asp (D)	44.7	45
Thr (T)	21.0	21
Ser (S)	33.9	36
Glu (E)	29.1	26
Pro (P)	9.1	9
Gly (G)	21.8	19
Ala (A)	16.2	15
½Cys (C)	1.1	1
Val (V)	17.5	19
Met (M)	2.4	2.5[b]
Ile (I)	22.8	23
Leu (L)	28.8	34
Tyr (Y)	17.8	19
Phe (F)	11.6	10
Try (W)	3.4	2
His (H)	3.3	3
Lys (K)	16.8	17
Arg (R)	6.6	7
Total residues	307.9	308.5[b]
Amides	41.1[c]	28 N, 10 Q } 38
Acids	33.0[c]	17 D, 16 E } 33

[a] Number of residues based on M_r 34000, as determined in [3]

[b] N-terminal Met missing in approx. 50% of molecules

[c] As determined in [6]

whether one or two of the most reactive amino groups (TNBS) is essential for biological activity, our guess for those groups will be the ϵ-amino groups on Lys^{80} and Lys^{82}. This must, however, be established experimentally.

ACKNOWLEDGEMENTS

This work was supported by research grants from the Science and Engineering Research Council and the Smith Kline Foundation. We also thank W. Telle, J. Gilroy and P. Miller for invaluable technical assistance and Professor D. Boulter for providing certain facilities.

REFERENCES

[1] Todd, E.C.D. (1978) J. Food Protect. 41, 559–565.
[2] Hatheway, C.L., Whaley, D.N. and Dowell, V.R. (1980) Food Technol. 34, 77–80.
[3] Granum, P.E. and Skjelkvåle, R. (1977) Acta Pathol. Microbiol. Scand. Sect. B 85, 89–94.
[4] Whitaker, J.R. and Granum, P.E. (1980) J. Food Biochem. 4, 201–217.
[5] Richardson, M. and Granum, P.E. (1983) Infect. Immun. 40, 943–949.
[6] Granum, P.E., Whitaker, J.R. and Skjelkvåle, R. (1981) Biochim. Biophys. Acta 668, 325–332.
[7] Richardson, M. and Granum, P.E. (1984) FEMS Symp. 24, 329–330.
[8] Granum, P.E. and Whitaker, J.R. (1980) Appl. Environ. Microbiol. 39, 1120–1122.
[9] Granum, P.E. (1982) Biochim. Biophys. Acta 708, 6–11.
[10] Crestfield, A.M., Moore, S. and Stein, W.H. (1963) J. Biol. Chem. 238, 622–627.
[11] 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. 144, 101–111.
[12] Richardson, M. (1974) Biochem. J. 137, 101–112.
[13] Woods, K.R. and Wang, K.T. (1967) Biochim. Biophys. Acta 133, 369–370.
[14] Mahoney, W.C. and Hermodson, M.A. (1980) J. Biol. Chem. 255, 11199–11203.
[15] Chang, J.Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[16] Campos, F.A.P. and Richardson, M. (1984) FEBS Lett. 167, 221–225.
[17] Enders, G.L. and Duncan, C.L. (1976) Can. J. Microbiol. 22, 1410–1414.
[18] Hauschild, A.H.W., Hilsheimer, R. and Martin, W.G. (1973) Can. J. Microbiol. 19, 1379–1382.
[19] Skjelkvåle, R. and Duncan, C.L. (1975) Infect. Immun. 11, 1061–1068.
[20] Dasgupta, B.R. and Pariza, M.W. (1982) Infect. Immun. 38, 592–597.
[21] Frieben, W.R. and Duncan, C.L. (1975) Eur. J. Biochem. 55, 455–463.
[22] Lai, C.Y., Mendez, E., Chang, D. and Wang, M. (1977) Biochem. Biophys. Res. Commun. 74, 215–222.