Brugge, J. S. (1986) *Curr. Top. Microbiol. Immunol. 123*, 1–22.
Catelli, M. G., Binart, N., Jung-Testas, I., Renoir, J. M., Baulieu, E. E., Feramisco, J. R., & Welch, W. J. (1985) *EMBO J. 4*, 3131–3135.
Dalman, F. C., Sanchez, E. R., Lin, A. L., Perini, F., & Pratt, W. B. (1988) *J. Biol. Chem. 263*, 12259–12267.
Danielson, M., Northrop, J. P., & Ringold, G. M. (1986) *EMBO J. 5*, 2513–2522.
Denis, M., Wikstrom, A. C., & Gustafsson, J. A. (1987) *J. Biol. Chem. 262*, 11803–11806.
Denis, M., Gustafsson, J. A., & Wikstrom, A. C. (1988a) *J. Biol. Chem. 263*, 18520–18523.
Denis, M., Poellinger, L., Wikstrom, A. C., & Gustafsson, J. A. (1988b) *Nature 333*, 686–688.
Gametchu, B., & Harrison, R. W. (1984) *Endocrinology 114*, 274–279.
Gehring, U., & Arndt, H. (1985) *FEBS Lett. 179*, 138–142.
Housley, P. R., Sanchez, E. R., Westphal, H. M., Beato, M., & Pratt, W. B. (1985) *J. Biol. Chem. 260*, 13810–13817.
Howe, J. G., & Hershey, J. W. B. (1981) *J. Biol. Chem. 256*, 12836–12839.
Ianotti, A. M., Rabideau, D. A., & Dougherty, J. J. (1988) *Arch. Biochem. Biophys. 264*, 54–60.
Joab, I., Radanyi, C., Renoir, J. M., Buchou, T., Catelli, M. G., Binart, N., Mester, J., & Baulieu, E. E. (1985) *Nature 308*, 850–853.
Laemmli, U. K. (1970) *Nature 227*, 680–685.
Mendel, D. B., & Orti, E. (1988) *J. Biol. Chem. 263*, 6695–6702.
Mendel, D. B., Bodwell, J. E., Gametchu, B., Harrison, R. W., & Munck, A. (1986) *J. Biol. Chem. 261*, 3758–3763.
Moore, S. K., Kozac, C., Robinson, E. A., Ullrich, S. J., & Appella, E. (1987) *Gene 56*, 29–40.
Nakao, K., Myers, J. E., & Faber, L. E. (1985) *Can. J. Biochem. Cell Biol. 63*, 33–40.
O'Farrell, P. H. (1975) *J. Biol. Chem. 250*, 4007–4021.
Okret, S., Wikstrom, A. C., & Gustafsson, J. A. (1985) *Biochemistry 24*, 6581–6586.
Pratt, W. B., Jolly, D. J., Pratt, D. V., Hollenberg, S. M., Giguere, V., Cadepond, F. M., Schweizer-Groyer, G., Catelli, M. G., Evans, R. M., & Baulieu, E. E. (1988) *J. Biol. Chem. 263*, 267–273.
Redmond, T., Sanchez, E. R., Bresnick, E. H., Schlesinger, M. J., Toft, D. O., Pratt, W. B., & Welsh, M. J. (1989) *Eur. J. Cell Biol.* (in press).
Rexin, M., Busch, W., & Gehring, U. (1988a) *Biochemistry 27*, 5593–5601.
Rexin, M., Busch, W., Segnitz, B., & Gehring, U. (1988b) *FEBS Lett. 241*, 234–238.
Riehl, R. M., Sullivan, W. P., Vroman, B. J., Bauer, V. J., Pearson, G. R., & Toft, D. O. (1985) *Biochemistry 24*, 6586–6591.
Sanchez, E. R., Toft, D. O., Schlesinger, M. J., & Pratt, W. B. (1985) *J. Biol. Chem. 260*, 12398–12401.
Sanchez, E. R., Meshinchi, S., Tienrungroj, W., Schlesinger, M. J., Toft, D. O., & Pratt, W. B. (1987) *J. Biol. Chem. 262*, 6986–6991.
Sanchez, E. R., Redmond, T., Scherrer, L. C., Bresnick, E. H., Welsh, M. J., & Pratt, W. B. (1988) *Mol. Endocrinol. 1*, 908–912.
Schuh, S., Yonemoto, W., Brugge, J., Bauer, V. J., Riehl, R. M., Sullivan, W. P., & Toft, D. O. (1985) *J. Biol. Chem. 260*, 14292–14296.
Sullivan, W. P., Vroman, B. T., Bauer, V. J., Puri, R. K., Riehl, R. M., Pearson, G. R., & Toft, D. O. (1985) *Biochemistry 24*, 4214–4222.
Tai, P. K. K., Maeda, Y., Nakao, K., Wakim, N. G., Duhring, J. L., & Faber, L. E. (1986) *Biochemistry 25*, 5269–5275.
Tienrungroj, W., Sanchez, E. R., Housley, P. R., Harrison, R. W., & Pratt, W. B. (1987) *J. Biol. Chem. 262*, 17342–17349.
Welch, W. J., & Feramisco, J. R. (1982) *J. Biol. Chem. 257*, 14949–14959.

Primary Structure of a Zinc Protease from *Bacillus mesentericus* Strain 76†

Stanka Stoeva,‡ Traute Kleinschmidt,*,§ Bohos Mesrob,‡ and Gerhard Braunitzer§,‖
Max-Planck-Institut für Biochemie, Abteilung Proteinchemie, D-8033 Martinsried bei Muenchen, West Germany, and Institute of Chemical Technology, Department of Organic Chemistry, Bu-1156, Sofia, Bulgaria
Received February 28, 1989; Revised Manuscript Received July 13, 1989

ABSTRACT: The amino acid sequence of the neutral zinc protease from *Bacillus mesentericus* strain 76 (MCP 76) has been determined by using peptides derived from digests with trypsin, chymotrypsin, and cyanogen bromide and from cleavage with *o*-iodosobenzoic acid. The peptides were purified by means of gel filtration and reversed-phase high-performance liquid chromatography and analyzed by automatic sequencing. The protein contains 300 amino acid residues. It proved to be identical with the neutral protease deduced from the DNA precursor sequence of *Bacillus subtilis*. The residues for zinc and substrate binding are conserved, whereas the number of calcium binding sites is reduced compared to thermolysin. A classification of the neutral zinc protease is discussed.

The zinc protease isolated from *Bacillus mesentericus* strain 76 is a rennin-like enzyme first described by Emanouilov (1951). It is used in the cheese-making industry as a milk-clotting protease (MCP 76)[1] and belongs to the neutral metalloendopeptidases. These enzymes are inhibited by chealting agents but not by covalent inhibitors of cysteine or serine

† This work was supported in part by a grant from the Alexander von Humboldt Foundation.
* Address correspondence to this author.
§ Max-Planck-Institut, Martinsried.
‡ Institute of Chemichal Technology, Sofia.
‖ Deceased May 27, 1989.

[1] Abbreviations: MCP 76, milk-clotting protease from *Bacillus mesentericus* strain 76 (neutral protease); HPLC, high-performance liquid chromatography.

0006-2960/90/0429-527$02.50/0

proteases. MCP 76 has been shown to have many properties in common with thermolysin, including specificity (Stoeva & Mesrob, 1977; Mesrob & Stoeva, 1978; Mesrob et al., 1981, 1982), metal dependence (Mesrob et al., 1976; Mesrob & Stoeva, 1983), and molecular weight (Mesrob & Vassileva, 1981) as well as the lack of carbohydrates and cysteine (Mesrob & Stoeva, 1983). The amino acid composition of MCP 76 (Mesrob & Stoeva, 1983) is similar to that of other neutral proteases like thermolysin (Titani et al., 1972) and *Bacillus subtilis* NRRLB 3411 (Keay, 1969). In this paper we report the primary structure of MCP 76, which was determined to clarify the relationship between structure and function of the enzyme and to compare it with the sequences of other neutral proteases.

Materials and Methods

TPCK-treated trypsin and α-chymotrypsin were purchased from Worthington; carboxypeptidase A was purchased from Boehringer. Cyanogen bromide and the column used for reversed-phase HPLC (LiChrospher 60 RP-select B, particle size 5 μm, 4 × 250 mm) were obtained from E. Merck. *o*-Iodosobenzoic acid was a product of Pierce Chemical Co.

Isolation of MCP 76. The crude enzyme was obtained after concentration and lyophilization of the bacterial medium of *B. mesentericus* strain 76 according to the method of Emanouilov (1951). It was subjected to three consecutive precipitations with ethanol as described earlier (Stoeva & Mesrob, 1982). The homogeneous neutral enzyme as well as the alkaline protease was obtained by ion-exchange fractionation on CM-Sephadex C-50 and elution with sodium chloride according to the method of Mesrob et al. (1978). To avoid autodigestion of the enzyme, fractions were collected in tubes containing 1 mL of 30% acetic acid. The peak fractions were pooled and lyophilized after dialysis. The purity was checked by polyacrylamide gel electrophoresis in the presence of 8 M urea and Triton X-100 (Alter et al., 1980).

Specific Cleavage of MCP 76. For tryptic digestion 16 mg of MCP 76 (500 nmol) was treated with TPCK-treated trypsin for 4 h at room temperature starting at pH 11 (enzyme/substrate ratio 2:100 w/w). Two core fractions were precipitated at pH 4.8 and 2.5 by titration with 0.1 M HCl and removed by centrifugation.

Cleavage of 16 mg of MCP 76 (500 nmol) with α-chymotrypsin was done in 0.1 M ammonium hydrogen carbonate, pH 9.0, at 37 °C for 7 min. The enzyme/substrate ratio was 5:100 (w/w).

After enzymatic hydrolysis the acidified samples were fractionated on Sephadex G-25 fine (column 2, 6 × 143 cm) in 0.1 M acetic acid. Further separation of the peptides was performed by reversed-phase HPLC using a 342 gradient liquid chromatograph, Controller 421 (Beckman Instruments) and a LiChrospher 60 RP-select B column in 50 mM ammonium acetate (pH 6.0) with a flow rate of 1 mL/min. Most of the peptides were eluted with a linear gradient of 0–40% acetonitrile in 40 min (Kratzin et al., 1980). Large peptides from the first Sephadex peaks were eluted with a gradient from 10 to 40% acetonitrile in 40 min.

The C-terminal amino acid residue was determined by carboxypeptidase A according to the method of Guidotti (1960) and Narita et al. (1975).

Cleavage at tryptophan was based on the *o*-iodosobenzoic acid method described by Mahoney and Hermodson (1979) using 150 nmol of MCP 76.

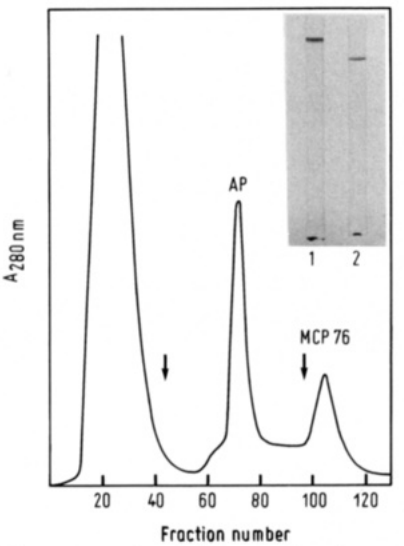

FIGURE 1: Chromatography of 400 mg of crude enzyme from *B. mesentericus* strain 76 on CM-Sephadex C-50. Column: 2.6 × 30 cm. Buffer: 50 mM sodium acetate and 1 mM calcium chloride, pH 6.5 (So). Gradient starting at fraction 42: 300 mL of So + 300 mL of So + 0.6 M NaCl. Final buffer starting at fraction 98: So + 0.6 M NaCl; buffer change is marked by arrows. Flow rate: 10.5 mL in 30 min. AP, alkaline protease; MCP 76, neutral protease. Triton electrophoresis: 1, AP; 2, MCP 76.

The core peptide obtained at pH 2.5 (80 nmol) was specifically cleaved at methionyl residues with cyanogen bromide. The reaction was performed in 70% formic acid with a 200-fold molar excess of cyanogen bromide for 24 h at room temperature in the dark.

Amino Acid Analysis. Samples were hydrolyzed in 5.7 M HCl at 110 °C in vacuo for 20 h and analyzed in an amino acid analyzer LC-5000 (Biotronik). Tryptophan was determined in the presence of 6% thioglycolic acid.

Sequence Determination. All sequences were determined by automatic Edman degradation. For sequence analysis of the N-terminal part of the denaturated MCP 76 chain, a liquid-phase sequencer (Model 890 C, Beckman Instruments) was employed. The program of Edman and Begg (1967) was modified by using 0.25 M Quadrol. For sequencing of the core peptide (pH 2.5) and the mixtures, obtained after *o*-iodosobenzoic acid and cyanogen bromide cleavage, we used 3-(diethylamino)propyne as buffer (Braunitzer et al., 1978) and a Model 890 B sequencer (Beckman Instruments). The sequences of all other peptides were established by the gas-phase method (Hewick et al., 1981) using a noncommercial sequencer (Begg et al., 1986).

Conversion was performed for 14 min at 80 °C in 1.9 M trifluoroacetic acid. The phenylthiohydantoin derivatives of the amino acids were identified by HPLC.

Results

The neutral protease MCP 76 as well as the alkaline one could be isolated from the same batch of *B. mesentericus* (Figure 1). They proved to be homogeneous during electrophoresis under dissociating conditions.

The primary structure of the neutral protease MCP 76, consisting of 300 amino acid residues, is presented in Figure 2. It was mainly determined by analysis of the tryptic peptides (Tp1–Tp25). From the digest, two core fractions could be isolated by precipitation with HCl at pH 4.8 and 2.5. The latter showed to be pure Tp14 (96 amino acid residues), whereas this peptide was contaminated by some other peptides in the pH 4.8 fraction. The soluble peptides were first frac-

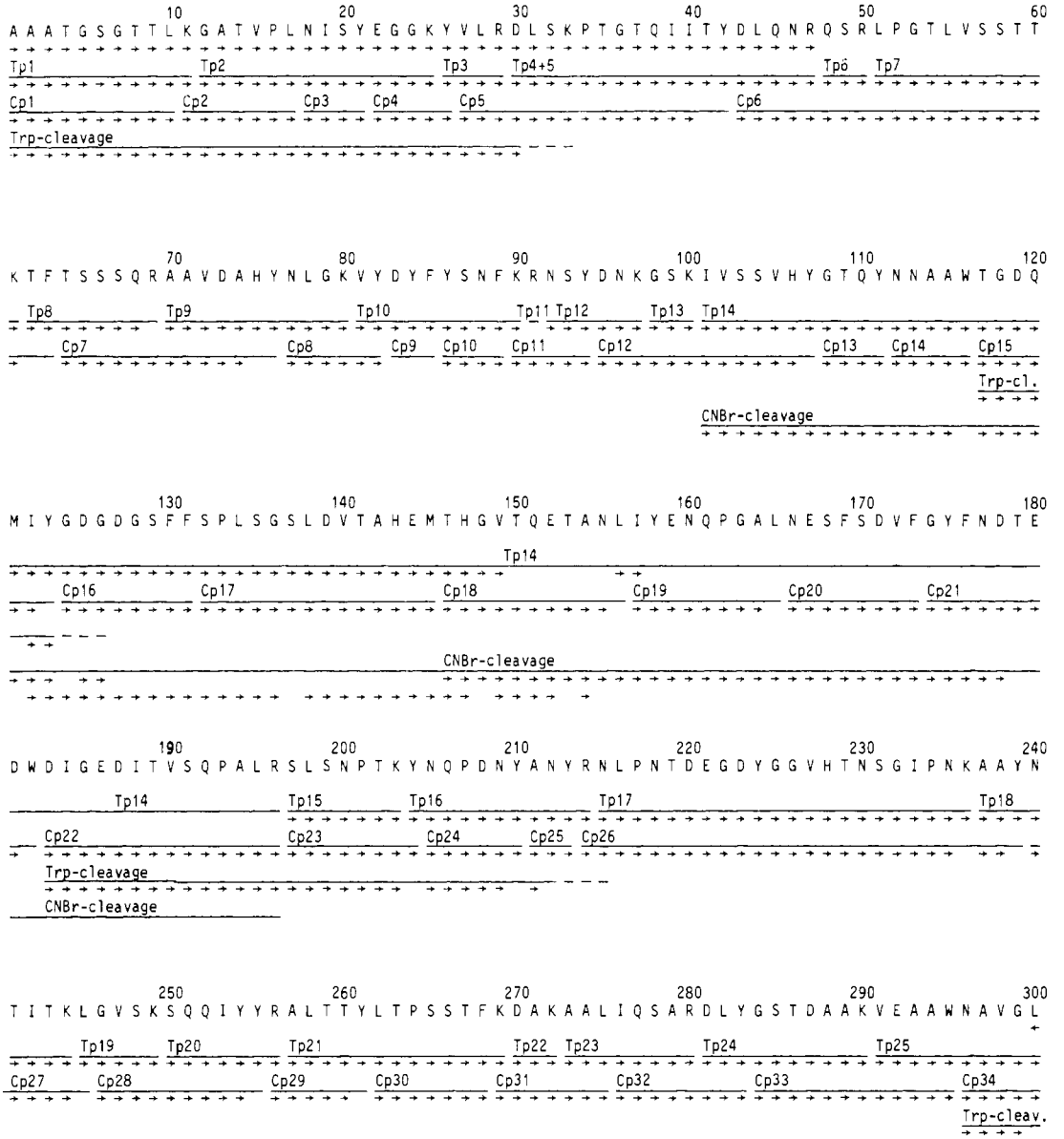

FIGURE 2: Amino acid sequence of the neutral zinc protease from *B. mesentericus* strain 76. The tryptic (Tp1–Tp25) and chymotryptic peptides (Cp1–Cp34) are marked. The amino acids identified after Edman degradation are indicated by →, and the one obtained after digestion with carboxypeptidase A with is indicated by ←. Tryptophan cleavage (Trp) was performed at the polypeptide chain and cyanogen bromide cleavage (CNBr) at the core peptide Tp14.

tionated by gel filtration on Sephadex G-25 in 0.1 M acetic acid (Figure 3A) and then isolated by reversed-phase HPLC of the separated peaks. All peptides could be sequenced to the end except Tp14, where 49 amino acid residues were determined. The alignment of the first tryptic peptides could be achieved with the help of the N-terminal sequence of the polypeptide chain (47 steps). All other peptide were positioned by homology with thermolysin.

Since thermolysin and MCP 76 differ to a relatively high extent in their primary structure and also in their chain length, the alignment was confirmed by analysis of chymotryptic overlapping peptides. The cleavage was started with a suspension of the denaturated MCP 76, which was completely insoluble. The digestion mixture became clear during a cleavage period of 7 min. No precipitation occurred when the reaction was stopped by titration to pH 2.5. The separation of the peptides was achieved as described for the tryptic peptides (Figure 3B). Cp18–Cp22 covered the C-terminal part of the tryptic core peptide Tp14 which could not be sequenced directly. The amino acid compositions of all tryptic (Tp1–Tp25) and chymotryptic (Cp1–Cp34) peptides are given in Tables I and II, respectively.

The core peptide Tp14, containing both methionyl residues of MCP 76, was subjected to a cyanogen bromide cleavage to confirm the alignment of Cp19–Cp21. Since the bond between Met^{145} and Thr^{146} was cleaved incompletely and a separation of the highly hydrophobic peptides in a sufficient yield seemed difficult, the whole mixture was subjected to automatic Edman degradation.

The overlap between Tp14 and Tp15 was achieved by sequencing of the fragment mixture obtained by cleavage at the three tryptophanyl residues of MCP 76 with *o*-iodosobenzoic acid. At the same time the overlap between Cp21 and Cp22 was confirmed. According to its amino acid composition Cp21 contains a tryptophanyl residue. It was the only amino acid residue not detected during sequencing and thus should be placed at its C-terminus. Since the primary structure around the other two tryptophanyl residues is known, the sequence of the third fragment identical with Cp22 should start at position 183 of the chain. The tryptophan cleavage also provided a proof for the C-terminal region.

Tp25 and Cp34 could be sequenced to the end, showing

Table I: Amino Acid Compositions[a] of Tryptic Peptides from MCP 76

peptide	Tp1	Tp2	Tp3	Tp4+5	Tp6	Tp7	Tp8	Tp9	Tp10
position	1–11	12–25	26–29	30–47	48–50	51–61	62–69	70–80	81–90
Asx		1.11 (1)		3.12 (3)				1.90 (2)	2.03 (2)
Thr	2.93 (3)	1.00 (1)		2.95 (3)		2.91 (3)	1.96 (2)		
Ser	1.03 (1)	1.04 (1)		1.05 (1)	1.00 (1)	2.11 (2)	3.04 (3)		1.01 (1)
Glx		0.97 (1)		2.04 (2)	0.90 (1)		1.07 (1)		
Pro		1.08 (1)		1.12 (1)		1.03 (1)			
Gly	2.26 (2)	3.18 (3)		1.15 (1)		1.01 (1)		1.08 (1)	
Ala	2.82 (3)	0.98 (1)						2.72 (3)	
Val		0.92 (1)	0.98 (1)			1.09 (1)		0.93 (1)	1.03 (1)
Met									
Ile		0.88 (1)		1.74 (2)					
Leu	0.95 (1)	0.95 (1)	0.98 (1)	1.78 (2)		1.86 (2)		1.27 (1)	
Tyr		0.91 (1)	1.01 (1)	0.92 (1)				0.91 (1)	2.92 (3)
Phe							0.94 (1)		2.01 (2)
Trp									
Lys	1.01 (1)	0.97 (1)		1.16 (1)		0.98 (1)		1.15 (1)	1.00 (1)
His								1.08 (1)	
Arg			1.02 (1)	0.97 (1)	1.10 (1)		0.99 (1)		
total residues	11	14	4	18	3	11	8	11	10
yield (nmol)	457	430	375	83	75	270	165	377	300

peptide	Tp11	Tp12	Tp13	Tp14	Tp15	Tp16	Tp17	Tp18	Tp19
position	91	92–97	98–100	101–196	197–203	204–214	215–236	237–244	245–249
Asx		3.00 (3)		15.20 (15)	1.01 (1)	4.10 (4)	5.78 (6)	1.03 (1)	
Thr				7.75 (8)	0.97 (1)		2.04 (2)	2.01 (2)	
Ser		1.00 (1)	0.98 (1)	9.14 (9)	1.98 (2)		1.06 (1)		1.01 (1)
Glx				11.01 (11)		1.05 (1)	1.17 (1)		
Pro				3.17 (3)	1.11 (1)	1.22 (1)	1.97 (2)		
Gly			1.05 (1)	10.31 (10)			4.07 (4)		1.05 (1)
Ala				6.13 (6)		1.00 (1)		1.99 (2)	
Val				6.23 (6)			0.92 (1)		0.94 (1)
Met				ND (2)					
Ile				4.34 (5)			0.89 (1)	0.96 (1)	
Leu				5.13 (5)	0.94 (1)		1.17 (1)		0.99 (1)
Tyr		0.96 (1)		4.52 (5)		2.62 (3)	0.90 (1)	0.98 (1)	
Phe				4.87 (5)					
Trp				ND (2)					
Lys		1.03 (1)	0.97 (1)		0.98 (1)		1.07 (1)	1.01 (1)	1.00 (1)
His				3.05 (3)			0.96 (1)		
Arg	1.00 (1)			1.08 (1)		1.01 (1)			
total residues	1	6	3	96	7	11	22	8	5
yield (nmol)	105	373	160	406	415	400	255	343	338

peptide	Tp20	Tp21	Tp22	Tp23	Tp24	Tp25	whole protein	
position	250–256	257–269	270–272	273–280	281–290	291–300	from analysis	from sequence
Asx			1.03 (1)		2.00 (2)	1.05 (1)	42.31	42
Thr		4.02 (4)			0.99 (1)		30.02	30
Ser	0.98 (1)	2.11 (2)		1.00 (1)	0.99 (1)		30.10	30
Glx	1.99 (2)			1.13 (1)		1.03 (1)	23.65	22
Pro		1.09 (1)					11.22	11
Gly					1.12 (1)	1.08 (1)	27.86	26
Ala		1.04 (1)	0.99 (1)	2.81 (3)	2.01 (2)	2.90 (3)	25.68	26
Val						1.89 (2)	13.89	15
Met							1.78	2
Ile	0.98 (1)			0.99 (1)			11.63	12
Leu		1.88 (2)		1.06 (1)	0.97 (1)	1.02 (1)	18.94	21
Tyr	2.00 (2)	0.88 (1)			0.91 (1)		20.05	22
Phe		0.94 (1)					9.23	9
Trp						ND (1)	2.68	3
Lys		1.01 (1)	0.98 (1)		1.02 (1)		15.69	15
His							5.86	5
Arg	1.04 (1)			1.00 (1)			9.43	9
total residues	7	13	3	8	10	10		300
yield (nmol)	372	410	257	170	170	159		

[a] Residues/molecule determined by amino acid analysis or (in parentheses) derived from the sequence (Figure 2). ND = not determined.

leucine as the C-terminal residue. This result was confirmed by enzymatic degradation of MCP 76 with carboxypeptidase A.

The established sequence of MCP 76 is in good agreement with both molecular mass determinations (Mesrob & Vassileva, 1981) and amino acid analysis of the polypeptide chain (Table I).

In addition to the neutral protease MCP 76 the alkaline one was sequenced on the polypeptide chain up to position 50:

```
1                   10                  20                  30
A Q S V P Y G I S Q I K A P A L H S Q G Y T G S N V K V A V
                    40                  50
I D S G I D S S H P D L N V R G G A S F.
```

Table II: Amino Acid Compositions[a] of Chymotryptic Peptides from MCP 76

peptide	Cp1	Cp2	Cp3	Cp4	Cp5	Cp6	Cp7	Cp8	Cp9	Cp10	Cp11	Cp12
position	1–10	11–17	18–21	22–26	27–42	43–63	64–76	77–82	83–85	86–89	90–94	95–107
Asx			1.11 (1)		1.02 (1)	2.11 (2)	1.07 (1)	1.06 (1)	1.06 (1)	1.04 (1)	1.05 (1)	2.09 (2)
Thr	2.96 (3)	1.01 (1)			3.14 (3)	3.85 (4)	0.92 (1)					
Ser	1.01 (1)		0.88 (1)		1.00 (1)	2.83 (3)	2.85 (3)			1.03 (1)	1.08 (1)	2.84 (3)
Glx				1.06 (1)	1.08 (1)	2.15 (2)	1.18 (1)					
Pro		1.02 (1)			1.19 (1)	1.12 (1)						
Gly	2.00 (2)	1.04 (1)		2.07 (2)	1.06 (1)	1.17 (1)		1.10 (1)				1.27 (1)
Ala	3.01 (3)	1.01 (1)					2.92 (3)					
Val		0.93 (1)			1.07 (1)	1.10 (1)	1.02 (1)	1.03 (1)				1.93 (2)
Met												
Ile			1.02 (1)		1.57 (2)							0.90 (1)
Leu	1.01 (1)	0.98 (1)			2.01 (2)	2.80 (3)		1.04 (1)				
Tyr			0.98 (1)	0.85 (1)	0.85 (1)		0.87 (1)	0.77 (1)	0.91 (1)	0.92 (1)	0.76 (1)	0.86 (1)
Phe						0.89 (1)			1.02 (1)	1.00 (1)		
Trp												
Lys		1.02 (1)		1.03 (1)	1.02 (1)	1.03 (1)		1.01 (1)			1.03 (1)	1.98 (2)
His							0.94 (1)					1.13 (1)
Arg					1.00 (1)	1.93 (2)	1.23 (1)				1.08 (1)	
total residues	10	7	4	5	16	21	13	6	3	4	5	13
yield (nmol)	432	420	447	415	135	65	238	430	150	122	275	245

peptide	Cp13	Cp14	Cp15	Cp16	Cp17	Cp18	Cp19	Cp20	Cp21	Cp22	Cp23
position	108–111	112–116	117–123	124–131	132–145	146–156	157–165	166–173	174–182	183–196	197–204
Asx		2.01 (2)	1.06 (1)	1.99 (2)	1.13 (1)	1.06 (1)	1.10 (1)	2.12 (2)	3.05 (3)	1.94 (2)	1.10 (1)
Thr	1.04 (1)		1.06 (1)		1.05 (1)	2.93 (3)			1.00 (1)	0.91 (1)	1.06 (1)
Ser				0.99 (1)	2.89 (3)			1.99 (2)		0.92 (1)	2.03 (2)
Glx	0.96 (1)		1.00 (1)		1.12 (1)	1.85 (2)	2.05 (2)	0.91 (1)	1.06 (1)	1.99 (2)	
Pro					1.05 (1)		1.12 (1)			1.11 (1)	1.03 (1)
Gly	1.26 (1)		1.09 (1)	3.10 (3)	1.08 (1)	1.13 (1)	0.97 (1)		1.16 (1)	1.25 (1)	
Ala		1.99 (2)			1.04 (1)	1.08 (1)	1.08 (1)			1.14 (1)	
Val					0.94 (1)	1.00 (1)		0.91 (1)		1.10 (1)	
Met			ND (1)		ND (1)						
Ile			0.94 (1)				0.84 (1)			1.71 (2)	
Leu					1.73 (2)	1.00 (1)	0.93 (1)			1.02 (1)	0.99 (1)
Tyr	0.74 (1)		0.85 (1)				0.90 (1)		0.90 (1)		0.79 (1)
Phe				1.91 (2)				2.07 (2)	0.94 (1)		
Trp		ND (1)							0.89 (1)		
Lys											1.01 (1)
His					0.97 (1)	0.96 (1)					
Arg										0.90 (1)	
total residues	4	5	7	8	14	11	9	8	9	14	8
yield (nmol)	120	362	165	293	18	200	108	120	53	50	245

peptide	Cp24	Cp25	Cp26	Cp27	Cp28	Cp29	Cp30	Cp31	Cp32	Cp33	Cp34
position	205–210	211–213	214–239	240–245	246–255	256–261	262–268	269–275	276–283	284–295	296–300
Asx	3.04 (3)	1.06 (1)	5.70 (6)	1.03 (1)				1.00 (1)	1.12 (1)	1.07 (1)	1.06 (1)
Thr			2.17 (2)	1.98 (2)		2.05 (2)	1.93 (2)			1.04 (1)	
Ser			1.01 (1)		2.03 (2)		2.04 (2)		1.15 (1)	1.04 (1)	
Glx	1.00 (1)		1.02 (1)		2.02 (2)				1.15 (1)	1.04 (1)	
Pro	1.09 (1)		1.79 (2)				1.03 (1)				
Gly			4.11 (4)		1.10 (1)					0.93 (1)	1.03 (1)
Ala		1.00 (1)	2.08 (2)			1.17 (1)		2.96 (3)	0.94 (1)	4.02 (4)	0.98 (1)
Val			1.08 (1)		1.01 (1)					1.01 (1)	0.95 (1)
Met											
Ile			0.97 (1)	0.97 (1)	1.06 (1)				0.96 (1)		
Leu			1.04 (1)	1.02 (1)		1.04 (1)	1.00 (1)	1.00 (1)	0.90 (1)		0.97 (1)
Tyr	0.87 (1)	0.94 (1)	1.82 (2)		1.81 (2)	0.81 (1)			0.84 (1)		
Phe							1.02 (1)				
Trp										ND (1)	
Lys			0.90 (1)	1.00 (1)	0.98 (1)			2.04 (2)		0.85 (1)	
His			1.16 (1)								
Arg			0.98 (1)			0.92 (1)			0.93 (1)		
total residues	6	3	26	6	10	6	7	7	8	12	5
yield (nmol)	402	60	45	407	225	28	398	287	402	260	239

[a] Residues/molecules determined by amino acid analysis or (in parentheses) derived from the sequence (Figure 2). ND = not determined.

Discussion

The aim of this work was to determine the primary structure of the neutral protease from *B. mesentericus* strain 76 and to compare it with other neutral zinc proteases of microbial origin. In the course of the investigations our data showed not only homology, as expected, but complete identity with the amino acid sequence deduced from the DNA sequence of the precursor of the neutral protease from *B. subtilis* (Yang et al., 1984). Yang et al. compared their sequence with fragments of another neutral protease (Levy et al., 1975) designated a product of *B. subtilis* NRRLB 3411 and found a relatively large number of mismatches. As an explanation they suggest that in the past several commercial strains were denoted *B. subtilis* when they were actually other species of *Bacillus*. For

Table III: Comparison[a] of Neutral Zinc Proteases of Microbial Origin

origin	chain length	*B. amyloliquefaciens*	*B. thermoproteolyticus*	*B. stearothermophilus*	*B. cereus*
B. mesentericus/subtilis	300	11.0	54.9	53.3	52.3
B. amyloliquefaciens	300		52.4	51.7	50.3
B. thermoproteolyticus	316			13.9	27.5
B. stearothermophilus	319				28.5
B. cereus	317				

[a] The number of amino acid substitutions is given in percent. It is calculated on the basis of identical positions after homologous alignment.

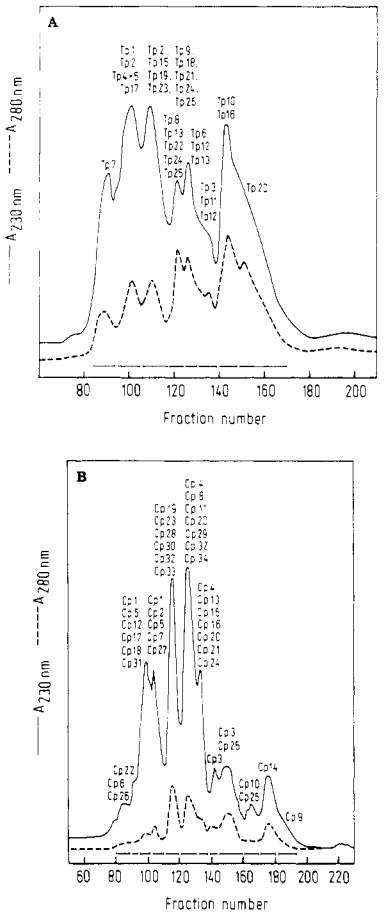

FIGURE 3: Prefractionation of peptide mixtures generated from enzymatic cleavages of 500 nmol of MCP 76. (A) Tryptic cleavage; (B) chymotryptic cleavage. Column: 2.6 × 143 cm Sephadex G-25 fine in 0.1 M acetic acid. Flow rate: 4.5 mL in 20 min. Fractions indicated by horizontal bars were pooled and separated into the peptides by reversed-phase HPLC on LiChrospher RP-select B in 50 mM ammonium acetate.

a better characterization of *B. mesentericus* strain 76 we additionally determined the N-terminal sequence of its alkaline protease up to position 50. It turned out to be identical with the translated major coding sequence of the *B. subtilis* subtilisin (Stahl & Ferrari, 1984). Although the primary structure of the neutral protease and the N-terminal region of the alkaline one from *B. mesentericus* strain 76 and *B. subtilis* are identical, it cannot be concluded that both strains are identical as well. Much more detailed comparison of these bacteria would be needed.

Further studies on the neutral zinc proteases showed that the partial sequence of Levy et al. (1975) is in nearly complete agreement with that of the mature protein of the *Bacillus amyloliquefaciens* neutral protease precursor (Vasantha et al., 1984). The only exception is position 168, where in the protein aspartic acid/asparagine was determined, whereas the DNA sequence coded for serine. It remains unclear whether this one residue differnece is due to a sequence error or whether the neutral protease A of Levy et al. (1975) really differs by one amino acid residue from the one of *B. amyloliquefaciens*.

The five neutral zinc proteases of microbial origin with known primary structures were obtained from *B. mesentericus/subtilis*, *B. amyloliquefaciens* (Levy et al., 1975; Vasantha et al., 1984), *Bacillus thermoproteolyticus* (Titani et al., 1972), *Bacillus cereus* (Sidler et al., 1986), and *Bacillus stearothermophilus* (Takagi et al., 1985). A detailed comparison of the sequences (Table III) shows that the neutral proteases of *B. mesentericus/subtilis* and *B. amyloliquefaciens* have the same chain length (300 amino acid residues) and differ in only 11% of the amino acid positions. On the contrary, the enzymes of *B. thermoproteolyticus*, *B. cereus*, and *B. stearothermophilus* with chain lengths of 316–319 amino acid residues are much less homologous, differing in about 30%. The difference between the latter ones and the first group is about 55%. According to Dayhoff's definition of protein families ("proteins within a family usually differ at fewer than half of their amino acid positions"; Dayhoff et al., 1978), the two groups of neutral zinc proteases should be considered two families. The deviation, however, is not large, and taking into account the similar properties of the proteases, we suggest combining them in one family with different subfamilies.

Among the neutral zinc proteases tertiary structures are only known for the ones of *B. thermoproteolyticus* (Matthews et al., 1974; Holms & Matthews, 1982) and *B. cereus* (Pauptit et al., 1988) with a refinement of 1.6 and 3.0 Å, respectively, but not for the enzymes more closely related to MCP 76. To study the relation between structure and function, the sequences of thermolysin and MCP 76 were aligned (Figure 4). Since the chain length of both proteins differs by 16 amino acid residues, several insertions and deletions had to be assumed for best homology. In MCP 76 143 amino acid residues are conserved compared to thermolysin. Among these are His^{142}, His^{146}, and Glu^{166}, participating in binding the zinc atom essential for catalytic activity. These three amino acid residues are conserved in all five zinc proteases. The functional Asn^{112}, Ala^{113}, Trp^{115}, Glu^{143}, Tyr^{157}, and His^{231} identified by Weaver et al. (1977) and Matthews (1988) as being at the active site of thermolysin are also present at the respective positions of MCP 76 with the exception of Phe^{114}, which is replaced by Ala in MCP 76. This might cause differences in the specificity. From the seven residues defining the hydrophobic pocket for substrate binding in thermolysin (Phe^{130}, Leu^{133}, Val^{139}, Ile^{188}, Gly^{189}, Val^{192}, and Leu^{202}, six are found at corresponding positions, whereas Val^{192} is replaced by the isopolar Ile.

In thermolysin, four calcium ions contribute to the unusual thermal stability and protect surface loops against autolysis (Roche & Voordouw, 1978; Corbett & Roche, 1983; Fassina et al., 1986). The amino acid residues involved in calcium binding in thermolysin (Matthews et al., 1972a,b; Colman et al., 1972) and the residues found at homologous positions in the other enzymes are shown in Table IV. It is obvious that the binding sites for the first and second calcium ions are mainly conserved in all five enzymes. The substitution of glutamic acid with aspartic acid at position 177 might not be

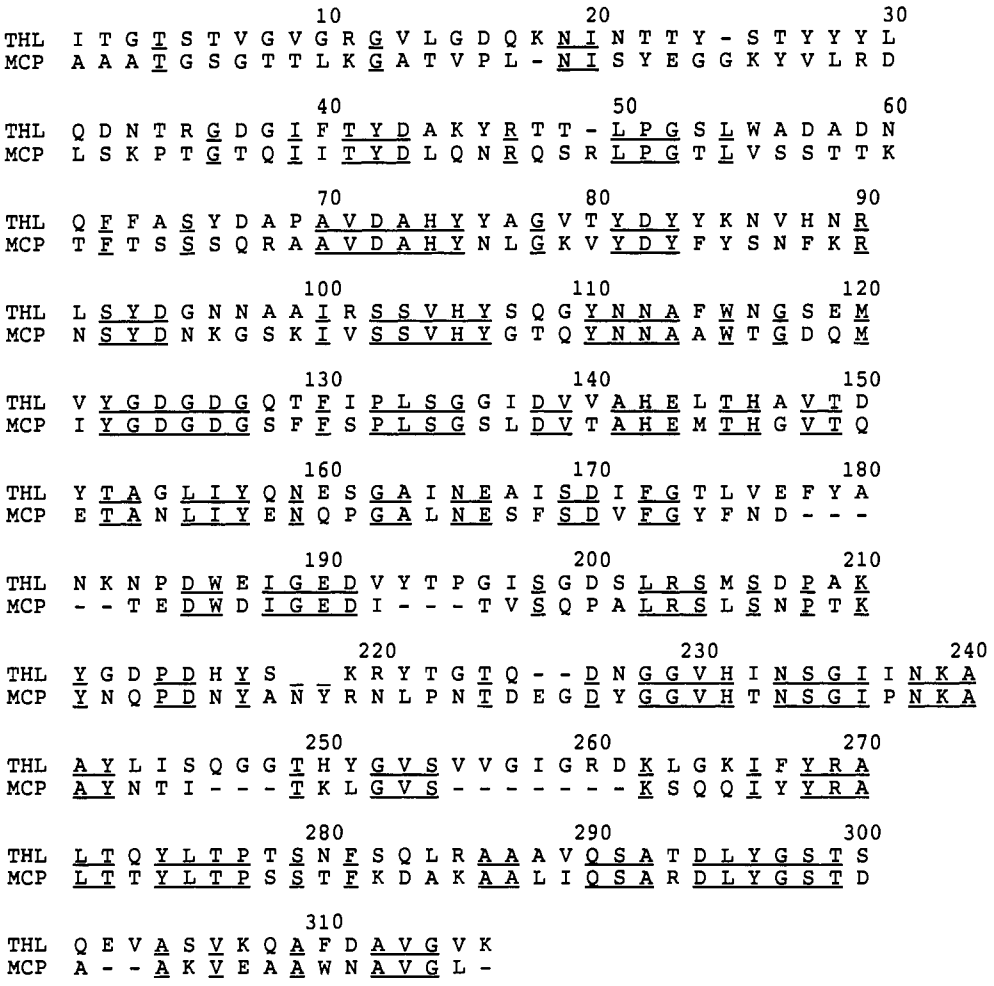

FIGURE 4: Comparison of the sequence of the neutral zinc protease from *B. mesentericus* strain 76 with the sequence of thermolysin (Titani et al., 1972). Identical amino acid residues are underlined. The residue numbering is that of thermolysin.

Table IV: Calcium Binding Sites in Neutral Zinc Proteases[a]

Ca binding site	3	3	3	1, 2	1, 2	1, 2	1, 2	1, 2	1, 2	4	4	4	4
position	57	59	61	138	177	183	185	187	190	193	194	197	200
B. thermoprotelyticus	D	D	Q	D	E	N	D	E	E	Y	T	I	D
B. stearothermophilus	D	D	Q	D	E	N	D	E	E	Y	T	V	D
B. cereus	D	D	V	D	E	N	D	E	E	Y	T	K	D
B. mesentericus/subtilis	S	T	T	D	D	T	D	D	E			V	P
B. amyloliquefaciens	S	T	Q	D	D	T	D	D	E			V	P

[a] The position numbers of the amino acid residues are those of thermolysin.

of importance. Since, according to Pauptit et al. (1988), in the case of Asn[183] the carbonyl moiety is involved in calcium binding, it might well be replaced by threonine without losing its calcium binding properties. These authors also suggest a new binding site for the second calcium ion involving Asp[191], which is present in MCP 76 too. In thermolysin this amino acid residue forms a salt bridge with Lys[182]. In MCP 76 the residue at position 182 is deleted according to our alignment, and thus Asp[191] may be involved in calcium binding too. The amino acid residues of calcium binding sites 3 and 4 are affected by more dramatic differences as can be seen in Table IV. Sidler et al. (1986) suggest that *B. subtilis* neutral protease has one or two calcium ions less than thermolysin and *B. cereus* neutral protease, which might be valid for MCP 76 too. It cannot be excluded that, due to the different chain length, structural rearrangements occur, creating new calcium binding sites, although the lower thermostability of MCP 76 (Mesrob et al., 1976) points to a possible loss of one or two calcium ions. An answer to this question can only be provided by an X-ray crystallogaphic structure analysis.

ACKNOWLEDGMENTS

We are grateful to Josefine Maerz for technical advice and thank Dr. Karl P. Ruecknagel, Barbara Schrank, Renate Gautsch, Reinhardt Mentele, Dr. Karin Rodewald, and Eva Wottawa for sequencer work and amino acid analysis.

Registry No. Metalloproteinase (*Bacillus mesentericus*), 93928-57-5; thermolysin, 9073-78-3; metalloproteinase, 81669-70-7.

REFERENCES

Alter, B. P., Goff, S. C., Efremov, G. D., Gravely, M. E., & Huisman, T. H. J. (1980) *Br. J. Haematol. 44*, 527–543.

Begg, G., Ruecknagel, P., Godovac-Zimmermann, J., Braunitzer, G., & Schuster, W. (1986) *Biol. Chem. Hoppe-Seyler 367*, 81–82.

Braunitzer, G., Schrank, B., Stangl, A., & Scheithauer, U. (1978) *Hoppe-Seyler's Z. Physiol. Chem. 359*, 137–146.

Colman, P. M., Jansonius, J. N., & Matthews, B. W. (1972) *J. Mol. Biol. 70*, 701–724.

Corbett, J. T., & Roche, R. S. (1983) *Biopolymers 22*, 101–105.

Dayhoff, M. O., Barker, W. C., Hunt, L. T., & Schwartz, R. M. (1978) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., Ed.) Vol. 5, Suppl. 3, pp 9–24, National Biomedical Research Foundation, Washington, DC.

Edman, P., & Begg, G. (1967) *Eur. J. Biochem. 1*, 80–91.

Emanouilov, I. (1951) *Izv. Mikrobiol. Inst., Bulg. Akad. Nauk 2*, 1–17.

Fassina, G., Vita, C., Dalzoppo, D., Zamai, M., Zambonin, M., & Fontana, A. (1986) *Eur. J. Biochem. 156*, 221–228.

Guidotti, G. (1960) *Biochim. Biophys. Acta 42*, 177–179.

Hewick, R. M., Hunkapiller, M. W., Hood, L. E., & Dreyer, W. J. (1981) *J. Biol. Chem. 256*, 7990–7997.

Holmes, M. A., & Matthews, B. W. (1982) *J. Mol. Biol. 160*, 623–639.

Keay, L. (1969) *Biochem. Biophys. Res. Commun. 36*, 257–265.

Kratzin, H., Yang, C., Krusche, J. U., & Hilschmann, N. (1980) *Hoppe-Seyler's Z. Physiol. Chem. 361*, 1591–1598.

Levy, P. L., Pangburn, M. K., Burstein, Y., Ericsson, L. H., Neurath, H., & Walsh, K. A. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4341–4345.

Mahoney, W. C., & Hermodson, M. A. (1979) *Biochemistry 18*, 3810–3814.

Matthews, B. W. (1988) *Acc. Chem. Res. 21*, 333–340.

Matthews, B. W., Colman, P. M., Jansonius, J. N., Titani, K. A., & Neurath, H. (1972a) *Nature (London), New Biol. 238*, 41–43.

Matthews, B. W., Jansonius, J. N., Colman, P. M., Schoenborn, B. P., & Dupourque, D. (1972b) *Nature (London), New Biol. 238*, 37–41.

Matthews, B. W., Weaver, L. H., & Kester, W. R. (1974) *J. Biol. Chem. 249*, 8030–8044.

Mesrob, B. K., & Stoeva, S. P. (1978) *Milchwissenschaft 33*, 483–486.

Mesrob, B., & Vassileva, R. (1981) *Int. J. Pept. Protein Res. 17*, 89–92.

Mesrob, B. K., & Stoeva, St. P. (1983) *Int. J. Pept. Protein Res. 21*, 369–373.

Mesrob, B., Ivanov, Ch., & Vassileva, R. (1976) *Dokl. Bolg. Akad. Nauk 29*, 1349–1352.

Mesrob, B. K., Stoeva, St. P., & Vassileva, R. A. (1978) *Int. J. Pept. Protein Res. 11*, 340–344.

Mesrob, B., Stoeva, St., & Simeonova, R. A. (1981) *God. Vissh. Khim.-Tekhnol. Inst., Sofia 28*, 65–77.

Mesrob, B. K., Stoeva, St. P., & Mahdi, R. (1982) *Dokl. Bolg. Akad. Nauk 35*, 653–656.

Narita, K., Matsuo, H., & Nakajima, T. (1975) *Mol. Biol. Biochem. Biophys. 8*, 30–103 (especially pp 90–91).

Pauptit, R. A., Karlsson, R., Picot, D., Jenkins, J. A., Niklaus-Reimer, A. S., & Jansonius, J. N. (1988) *J. Mol. Biol. 199*, 525–537.

Roche, R. S., & Voordouw, G. (1978) *CRC Crit. Rev. Biochem. 5*, 1–23.

Sidler, W., Niederer, E., Suter, F., & Zuber, H. (1986) *Biol. Chem. Hoppe-Seyler 367*, 643–657.

Stahl, M. L., & Ferrari, E. (1984) *J. Bacteriol. 158*, 411–418.

Stoeva, S. P., & Mesrob, B. K. (1977) *FEBS Lett. 80*, 86–88.

Stoeva, S. P., & Mesrob, B. K. (1982) *Milchwissenschaft 37*, 220–224.

Takagi, M., Imanaka, T., & Aiba, S. (1985) *J. Bacteriol. 163*, 824–831.

Titani, K., Hermodson, M. A., Ericsson, L. H., Walsh, K. A., & Neurath, H. (1972) *Nature (London), New Biol. 238*, 35–37.

Vasantha, N., Thompson, L. D., Rhodes, C., Banner, C., Nagle, J., & Filpula, D. (1984) *J. Bacteriol. 159*, 811–819.

Weaver, L. H., Kester, W. R., & Matthews, B. W. (1977) *J. Mol. Biol. 114*, 119–132.

Yang, M. Y., Ferrari, E., & Henner, D. J. (1984) *J. Bacteriol. 160*, 15–21.