JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Validation of Recombinant and Bovine Chymosin by Mass Spectrometry

SERGIO LILLA,*,† GIOVANNI MOGNA,‡ AND FRANCESCO ADDEO†

Dipartimento di Scienza degli Alimenti, Università degli Studi di Napoli Federico II, I-80055 Portici, Italy, and Mofin-Alce, Via Pietro Custodi, 12 I-28100 Novara

Mass spectrometry has been used to map chymosin from a fermentative source. The copresence of the two known genetic variants A (Asp^{244}) and B (Gly^{244}) was ascertained in bovine chymosin. By contrast, either the A or the B genetic variant occurred in the three commercial samples of recombinant calf chymosin (RCC). Specific biomarker proteins were searched to identify the enzyme source, in both bovine chymosin and RCC samples. Analyzing the derived tryptic peptides, evidence was provided that RCC and bovine chymosin are mainly formed by (1-323), (3-323), and (40p-323) (suffix "p" denotes residues in the pro-segment region of chymosin), whereas the minor components, (4-323), (5-323), and (6-323), were only detected in bovine chymosin. Additionally, the three commercial RCC samples contained the protein species (1-323), (38p-323), (39p-323), and (40p-323) and the shorter form (3-323). Differentiation of the natural and bioengineered enzyme is based upon the detection of these unique minor components by mass spectrometry.

KEYWORDS: Chymosin; recombinant enzyme; peptide markers; electrospray quadrupole/time-of-flight mass spectrometry

INTRODUCTION

Prochymosin and pepsinogen are inactive zymogens localized in the fourth stomach of ruminants, irreversibly converted into active enzymes at acid pH. "Rennet", the enzyme preparation available in commercial formats extracted from the abomasum, actually consists of a saline solution containing variable proportions of these two enzymes. The ratio of chymosin and pepsin in rennet is mainly determined by the current animal diet and the age of slaughter (*1*). The shortage of calf abomasums has led to the use of DNA recombinant calf chymosin (RCC), at high purity, produced in different microorganisms in large-scale. In 1990, RCC expressed in GRAS microbes was approved by the Food and Drug Administration for cheese making. RCC chymosin is presently used in 70% of U.S. cheese production and for other special cheeses. On the other hand, rennet is the only milk coagulant preparation allowed by the European Union for the manufacture of Protected Denomination of Origin cheeses while RCC is used for common cheeses. The gene encoding for bovine chymosin has been successfully cloned and expressed using a number of plasmid vectors and host organisms including *Escherichia coli* (Chy-Max, Pfizer, Milwaukee, WI), *Kluyveromyces lactis* (Maxiren, Gist-Brocades, Delft, Holland), and *Aspergillus niger* var. *awamori* (Chymogen, Hansen, Denmark). Several forms of the precursor of bovine chymosin have been expressed as a fusion protein and activated by an autocatalysis mechanism into chymosin during product recovery. Cloning the gene coding for preprochymosin A or B has made available to cheese producers two mature A and B chymosin forms differing by the amino acid substitution $Asp^{244}(A) \rightarrow Gly^{244}(B)$. A single chymosin variant containing product was distributed by each of the three chymosin producers.

The RCC preparations have not been characterized in depth as the bovine natural rennet counterpart (*2*). Chymosin expressed by the yeast *K. lactis* was reported to present an amino acid composition and electrophoretic pattern identical to those of calf chymosin, including cheese-making characteristics (*3*).

The lack of studies on mature protein justifies why no analytical methods are available to differentiate rennet and RCC. The only existing procedure is based on recognition by enzyme-linked immunosorbent assay tests of accompanying proteins expressed by genetically engineered microorganisms together with the zymogen (*4*). The reliability of this procedure is seriously limited by the fact that changing the host microorganism could alter the identity of expressed contaminating proteins, making ineffective any recognition of targeted proteins by polyclonal antibodies. Specific methods are required to hinder the illegal practice of spiking natural rennet with lower cost RCC. By using narrow-bore reverse phase liquid chromatography coupled with electrospray mass spectrometry (RPLC/ESI-MS) and quadrupole time-of-flight (Q-TOF) electrospray tandem mass spectrometry, bovine natural rennet preparations and chymosin-based milk-clotting preparations have been previously characterized (*2*). Beside lysozyme and serum albumin, bovine natural rennet was found to be composed of a mixture of chymosins A and B. Each natural bovine chymosin variant was more complex due to the presence of N terminally different forms: one, three residues longer, and the other, two residues shorter.

* To whom correspondence should be addressed. Tel: +39 081 2539361. Fax: +39 081 7762580. E-mail: lilla@unina.it.
† Università degli Studi di Napoli Federico II.
‡ Mofin-Alce.

10.1021/jf0478051 CCC: $30.25
Published on Web 06/07/2005

No comparative study among the three commercially available RCC preparations as well as chemical structure studies is at the moment available. In this paper, we define on a molecular basis the characteristics of DNA recombinant chymosin in comparison with the authentic bovine rennet counterpart. We observe that the analytical procedure was useful to validate any chymosin preparation providing evidence that, during in vivo prochymosin activation, some proteolytic cuts make recognizable the recombinant and natural enzyme.

MATERIALS AND METHODS

Materials. *Enzymes*. The only three RCC preparations available in the market were analyzed. Chy-max was from Pfizer (marketed by Chr. Hansen); Chymogen was from Genencor International (marketed by Chr. Hansen), and Maxiren was from Gist-Brocades. The two reference bovine chymosin samples used were a commercial sample from Sigma (St. Louis, MO) and a crystalline chymosin sample, essentially a laboratory material obtained according to the Foltmann's procedure (*5*). Sequence grade bovine pancreatic trypsin (150 U/mg) was obtained from Boehringer (Mannheim, Germany).

Chemicals. High-performance liquid chromatography (HPLC) grade solvents and reagents were from Carlo Erba (Milan, Italy).

Chromatography. The chymosin samples were desalted using a size exclusion column Econopack 10DG (BioRad, Milan, Italy) equilibrated in 5% acetic acid buffer solution (v/v). The column eluate was collected manually, and the UV absorbance was measured at 280 nm prior to freeze drying.

A Vydac C4 214TP52, 250 mm × 2.1 mm, 5 μm (Hesperia, CA) column was used for narrow-bore LC/ESI-MS, and a Vydac C18 218TP54, 250 mm × 4.6 mm, 5 μm column was used for RP-HPLC fractionation of chymosins N-terminal tryptic peptides. A 15 cm column (75 μm i.d.) packed with C18 PepMap resin and a precolumn cartridge (i.d. 300 μm) also packed with C18 PepMap resin (both from LC-Packings) were used for capillary chromatography.

Enzymatic Hydrolysis. Hydrolysis with trypsin was carried out in 0.4% ammonium bicarbonate, pH 8.5, for 4 h at 37 °C, at an enzyme: substrate ratio of 1:100 (w/w). The reaction was stopped by freeze drying.

RPLC. RP-HPLC on chymosin N-terminal tryptic peptides was done using a Hewlett-Packard HP1100 modular system. The peptides were fractionated on a Vydac C18 218TP54, 250 mm × 4.6 mm, 5 μm column. Solvent A was 0.1% trifluoroacetic acid (TFA) in water, and solvent B was acetonitrile in 0.1% TFA. The samples were injected manually through a Rheodyne valve onto the RP-HPLC column equilibrated with 5% of solvent B. A linear gradient (5−70%) of B was applied at a flow rate of 1 mL/min for 90 min. The elution profile of the proteins was monitored at 220 and 280 nm.

Narrow-Bore LC/ESI-MS. Narrow-Bore LC/ESI-MS was performed using a Hewlett-Packard HP1100 modular system on line connected with a Platform (Micromass, Manchester, United Kingdom) single quadrupole mass spectrometer. Samples were fractionated on a Vydac C4, 214TP52, 5 μm 250 mm × 2.1 mm i.d. column. Solvent A was 0.1% TFA in water, and solvent B was acetonitrile in 0.1% TFA. Samples were dissolved into water and injected onto the RP-HPLC column equilibrated in 30% of solvent A. A linear gradient from 30 to 80% B was applied at a flow rate of 200 μL/min over 60 min for intact protein analysis. The ES-mass spectra were acquired between 400 and 1800 *m*/*z* at a scan time of 5 s/scan. The source temperature was 180 °C, the capillary voltage was 3.5 kV, and the cone voltage was 40 V.

ESI Q-TOF MS. All tryptic digests were analyzed using a Q-TOF hybrid mass spectrometer Q-TOF Ultima from Micromass equipped with a nano Z-spray source operating in positive ion mode. The ionization conditions included a capillary voltage of 2.3 kV, a cone voltage and RF1 lens of 30 and 100 V, respectively, and a collision energy of 10 eV. The source temperature was 70 °C, and the cone gas was nitrogen flowing at 80 L/h; no nebulizing gas was used to obtain spray. Argon was used for collisional cooling and for fragmentation of ions in the collision cell. External calibration with sodium iodide was done over the mass range from 50 to 2500 *m*/*z*. All spectra were acquired with the TOF analyzer in "V-mode" (TOF kV = 9.1).

Capillary LC/ESI Q-TOF. Capillary LC-ESI-QTOF/MS was carried out on all tryptic digests using a Waters CapLC ternary pump system connected through a stream-select valve module to the nano Z-spray source of the mass spectrometer. The sample was injected into the system through the CapLC autosampler using the "microliter pickup" injection method. Solvent A consisted of 95% H_2O and 5% acetonitrile in 0.1% TFA, and solvent B consisted of 95% acetonitrile and 5% H_2O in 0.1% TFA. Solvent C consisted of 0.1% TFA in H_2O. The protein digest was preconcentrated and desalted on a precolumn cartridge (300 μm i.d.) packed with C18 PepMap resin and connected to pump C through the stream-select valve module. The preconcentration/desalting step was done at 30 μL/min over 3 min using pump C. After we switched to pumps A and B, a gradient was applied to the precolumn cartridge and then used to elute the sample from the analytical column (15 cm × 75 μm i.d.) packed with C18 PepMap resin. The column was equilibrated in 5% of solvent B, and a linear gradient from 5 to 70% of B was applied over 60 min at an approximate flow rate of 200 nL/min using a precolumn split with pump delivering at 1 μL/min. The TOF MS analysis was done on-line with capillary chromatographic separation on all of the tryptic digests over the mass range of 400−2000 *m*/*z* at a scan speed of two s/scan. All masses reported are monoisotopic except for molecular masses over 4500 Da.

Peptide Sequence Determination. Tryptic peptides sequenced by tandem MS were fractionated by RP-HPLC, lyophilized, and then suspended in 80% H_2O/20% acetonitrile in 0.1% TFA, prior to injection into the mass spectrometer source at 500 nL/min flow rate. Before fragmentation, an ESI-QTOF/MS mass spectrum (TOF MS mode) was performed over the mass range of 400−2000 *m*/*z*, the precursor ion of interest selected in the first quadrupole and then fragmented in the collision cell (TOF MS/MS mode) with different collision energies, depending on the mass and charge state of the ions. The resulting product ion spectra was acquired in the TOF analyzer and deconvoluted using the MassLynx-MaxEnt 3 algorithm. Singly charged spectra obtained were processed manually using the PepSeq application included in MassLynx.

RESULTS

Comparative Narrow-Bore LC-ESI/MS Analysis of Intact DNA Recombinant and Reference Bovine Chymosin Sample. Desalting chymosin samples by size exclusion chromatography was followed by UV absorption recording at 280 nm of the column effluent. Because of the different protein contents, weight aliquots of RCC and reference chymosin samples were taken according to screen protein composition.

By narrow-bore LC-ESI/MS analysis, the chymosin preparation showed a similar profile regardless of the sample source. As shown in **Figure 1** for the reference chymosin sample, the broad peak at t_R = 40.3 min showed the copresence of two protein chains with masses of 35705.6 and 35647.9 Da matching the bovine chymosin A and B variants, respectively. The other peak, which eluted at t_R = 36.1 min, corresponded to the chymosin fragment 247-323 (*2*). Similar results were obtained by analyzing commercial chymosin samples.

The relative abundance of bovine chymosin genetic variants was measured by integrating the peaks obtained from the intensity plot chromatograms of multicharged ions from 38 to 34 min for each chymosin variant (results not shown). The chymosin A and B proportions in rennet preparations accounted for 35 and 65%, respectively, consistent with the expected values previously measured in rennet preparations by Lilla et al. (*2*).

Repeated measurements by ESI-MS demonstrated the presence of 35648.5 and 35649.7 Da molecular mass components for the RCC samples Chymogen and Maxiren, respectively, each corresponding to chymosin B (expected molecular mass 35647.0 Da). By the same procedure, a molecular mass of 35706.2 Da, corresponding to chymosin A (expected molecular mass, 35705.0 Da), was measured for the RCC preparation Chy-Max. The RCC samples contained the recombinant mature enzyme

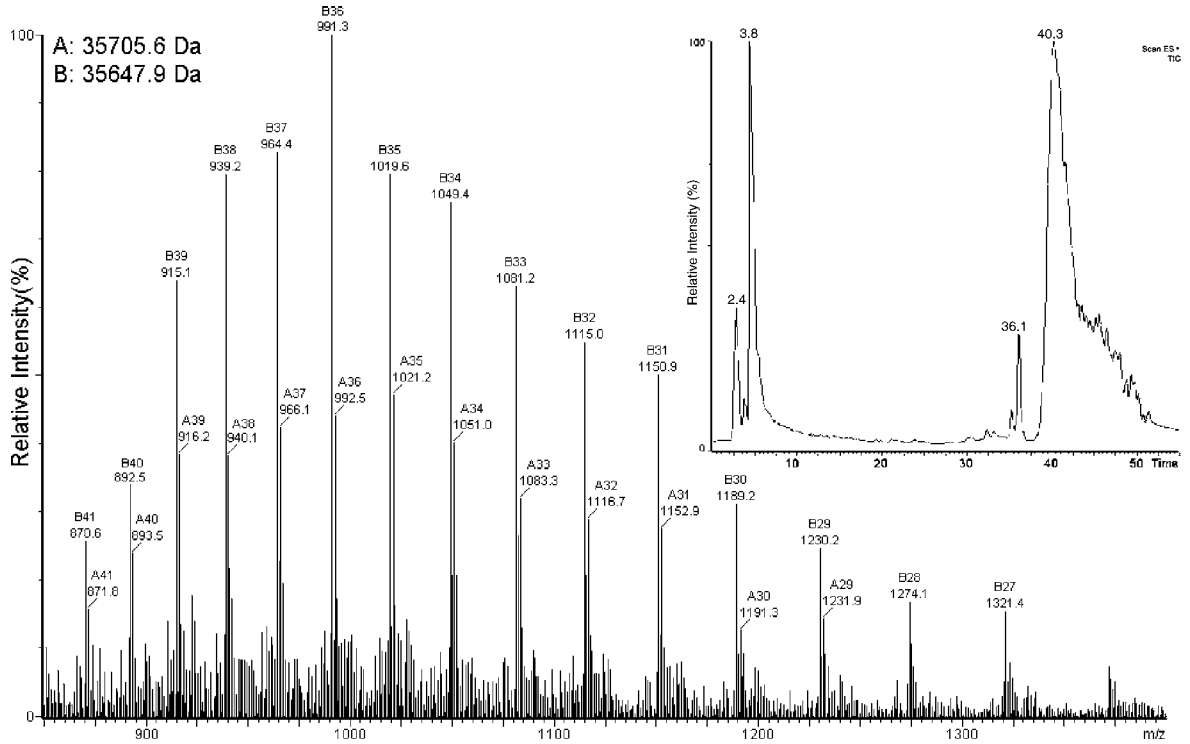

Figure 1. Identification by positive electrospray mass spectrometry of bovine chymosin A and B variants (mass values indicated on spectrum left top) in the peak at t_R = 40.3 min (cf. inset). The chromatographic profile of the bovine chymosin reference sample by narrow-bore LC is shown in the inset.

with the predicted molecular mass identical to that of the mature natural chymosin. Thus, the activated recombinant and natural chymosin should be catalytically indistinguishable.

Comparative Capillary LC-ESI-QTOF/MS Analysis of Reference Bovine and DNA Recombinant Chymosin Samples. The purpose of this experiment was to evaluate the amino acid sequence of chymosin in the RCC samples to determine if this would be identical to that of bovine chymosin. The protein preparations were separately desalted and hydrolyzed with trypsin, and the digests were analyzed by capillary LC-ESI-QTOF/MS. In **Figure 2**, the base peak intensity profile of the trypsin digest of the RCC Maxiren is shown.

The molecular mass of the tryptic peptides obtained from the RCC samples and natural bovine chymosin is shown in **Tables 1** and **2**, respectively. Mass values out of the detection range were not included in the tables.

The protein map of the three RCC samples was similar to that of the bovine counterpart. The differences between natural and RCC were limited to the amino acid substitutions in the disulfide-linked peptide G10 (222-266)-S-S-(267-295) eluting under a single peak. This contained only the tryptic peptide expected for bovine chymosin B (theoretical mass, 8215.99 Da) in Maxiren (measured mass, 8158.38 Da, inset of **Figure 2**) and Chymogen (measured mass, 8158.29 Da) samples. The counterpart tryptic peptide from the Chy-Max sample corresponded to that of chymosin A (measured mass, 8216.21 Da; expected, 8215.99 Da). However, in the tryptic peptide G10 from commercial bovine chymosin sample, shown in **Table 2**, coeluted the two latter peptides arising from bovine chymosin A (measured mass, 8216.18 Da) and B (measured mass, 8158.15 Da), which indicates the copresence of the A and B variants in natural enzyme preparations, as already reported by Lilla et al. (*2*).

A similar result was obtained for the crystalline bovine chymosin reference sample as shown in **Table 2**. The results of this analysis indicated that the RCC samples were chymosin

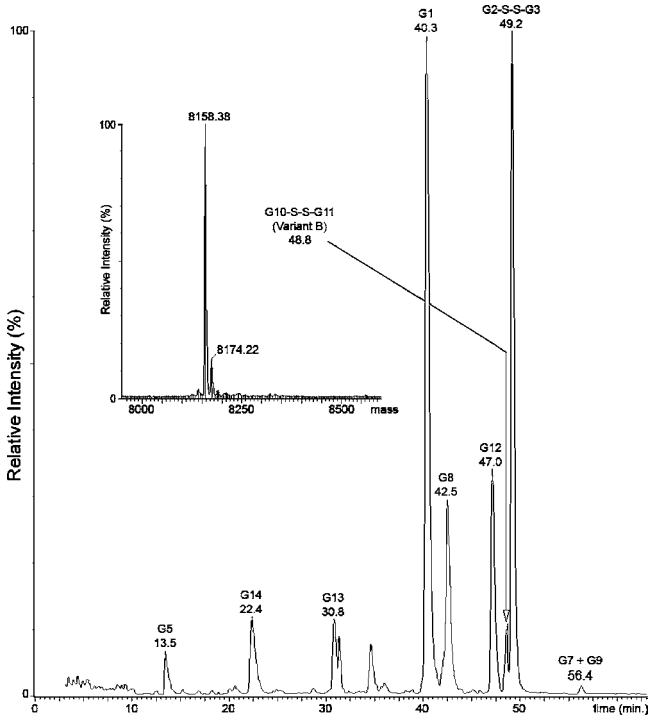

Figure 2. Chromatographic profile of RCC Maxiren tryptic digest by capillary LC-ESI-QTOF/MS; the deconvoluted spectrum in the inset shows the (222-266)-S-S-(267-295) chymosin B tryptic peptide. For peak identification, refer to **Table 1**.

B (Asp^{244}) in Maxiren and Chymogen and chymosin A (Gly^{244}) in Chy-Max. The other two peptides containing disulfide bonds, i.e., (20-48)-S-S-(49-53) and (160-221)-(S-S), occurred in each of the three RCC samples. The latter were found to contain the expected disulfide bond pattern with three bridges linking Cys^{47}-Cys^{52}, Cys^{207}-Cys^{211}, and Cys^{250}-Cys^{283} residues as in the natural chymosin molecule. Moreover, comparing the measured and

Table 1. Identification by Capillary LC/ESI-QTOF/MS of the Derived Tryptic Peptides from Recombinant Chymosin of Three Commercial Samples Assuming Bovine Chymosin A and B as Reference Proteins[a]

RCC							
Maxiren		Chy-max		Chymogen			
peptide label/measured mass (Da)						theoretical mass (Da)	protein sequence
G1	2087.12	P1	2087.14	H1	2087.09	2087.02	(1 G-K 19)
G2,3	3829.01	P2,3	3828.99	H2,3	3828.97	3828.81	(20 I-K 48)-S-S-(49 S-K 53)
G4	553.28	P4	553.29	H4	553.28	553.27	(54 N-R 57)
G5	533.29	P5	533.29	H5	533.28	533.26	(58 F-R 61)
G6		P6		H6		146.11	(62 K)
G7	8988.13	P7	8987.76	H7	8987.99	8988.08	(63 S-R 145)
G8	1692.92	P8	1692.92	H8	1692.92	1692.83	(146 H-R 159)
G9	6610.23	P9	6609.62	H9	6610.27	6610.42	(160 N-K 221) (S-S)
G10,11A		P10,11A	8216.21	H10,11A		8215.99	(222 L-K 266)-S-S-(267 M-K 295) var. A
G10,11B	8158.38	P10,11B		H10,11B	8158.29	8157.96	(222 L-K 266)-S-S-(267 M-K 295) var. B
G12	1117.69	P12	1117.69	H12	1117.68	1117.63	(296 W-R 304)
G13	1077.53	P13	1077.53	H13	1077.53	1077.48	(305 E-R 312)
G14	898.57	P14	898.57	H14	898.56	898.52	(313 A-K 321)
G15		P15		H15		202.13	(322 AI 323)
G1*	2669.35	P1*	2669.39	H1*	2669.23	2669.30	(38p K-K 19)
G2*	2541.16	P2*	2541.12	H2*	2541.25	2541.20	(39p Y-K 19)
G3*	2378.22	P3*	2378.22	H3*	2378.22	2378.14	(40p S-K 19)
G1′	1901.08	P1′	1901.03	H1′	1901.06	1900.96	(3 V-K 19)

[a] All molecular masses are reported as monoisotopic, except that above 4500 Da reported as average. A blank cell means not detected. *Tryptic peptides containing prochymosin residues (the prochymosin segment accounts for 42 amino acid residues). A prime denotes tryptic peptides derived from N-truncated chymosin species.

Table 2. Identification by Capillary LC/ESI-QTOF/MS of the Tryptic Peptides in Two Samples of Bovine Chymosin[a]

bovine chymosin					
commercial		reference			
peptide label/measured mass (Da)				theoretical mass (Da)	protein sequence
S1	2087.09	C1	2087.16	2087.02	(1 G-K 19)
S2,3	3828.97	C2,3	3829.01	3828.81	(20 I-K 48)-S-S-(49 S-K 53)
S4	553.31	C4	553.29	553.27	(54 N-R 57)
S5	533.28	C5	533.32	533.26	(58 F-R 61)
S6		C6		146.11	(62 K)
S7	8988.35	C7	8988.29	8988.08	(63 S-R 145)
S8	1692.90	C8	1692.76	1692.83	(146 H-R 159)
S9	6611.20	C9	6610.28	6610.42	(160 N-K 221) (S-S)
S10,11A	8216.18	C10,11A	8216.21	8215.99	(222 L-K 266)-S-S-(267 M-K 295) var. A
S10,11B	8158.15	C10,11B	8158.31	8157.96	(222 L-K 266)-S-S-(267 M-K 295) var. B
S12	1117.68	C12	1117.56	1117.63	(296 W-R 304)
S13	1077.52	C13	1077.54	1077.48	(305 E-R 312)
S14	898.56	C14	898.56	898.52	(313 A-K 321)
S15		C15		202.13	(322 AI 323)
S1′	1901.06	C1′	1901.15	1900.96	(3 V-K 19)
S2′	1802.01	C2′	1801.97	1801.89	(4 A-K 19)
S3′	1730.96	C3′	1730.78	1730.85	(5 S-K 19)
S4′	1643.81	C4′	1643.81	1643.82	(6 P-K 19)
S3*	2378.20	C3*	2378.18	2378.14	(40p S-K 19)

[a] The molecular mass figures are as in **Table 1**. A blank cell means not detected. *Tryptic peptides containing prochymosin residues (the prochymosin contain 42 amino acid). The prime denotes tryptic peptides derived from N-truncated chymosin species.

expected molecular mass figures, no difference was detected between bovine and RCC chymosin-derived peptides (20-48)-S-S-(49-53) and (160-221)-(S-S) containing the active site Asp^{32} and Asp^{215}, situated in the N- and C-terminal domains, respectively. As part of chymosin characterization, we developed an approach to detect the simultaneous presence of protein forms longer and shorter than mature enzymes (*2*). Different origin chymosin was hydrolyzed with trypsin, and digests were analyzed by LC/ESI-MS. This approach was similar to that previously developed for protein species profiling in chymosin preparations (*2*).

Shorter and Longer Forms of DNA Recombinant and Bovine Chymosin Reference Samples. We needed to investigate whether commercial and reference chymosin samples could contain chymosin species of different length. To that end, we identified (**Table 2**) the main tryptic peptide (1-19) obviously released from the dominant chymosin species and (3-19), (4-19), (5-19), and (6-19) in secondary amounts. In an effort to assign the differential presence of these tryptic peptides to recombinant chymosin preparations, the relative total ion current (TIC) was compared to that of natural chymosin as shown in **Figure 3**. Here, the intensity plot of doubly charged ions of the above indicated peptides extracted from the TIC of two bovine chymosin reference samples and RCC Maxiren sample (panels a−n) is compared with the respective chromatographic profile from 35 to 48 min (panels o−q). While the (3-19), (4-19), (5-19), and (6-19) peptide mass signals simultaneously occurred along the TIC of two bovine chymosin reference sample tryptic digests, only the (3-19) peptide occurred in that from RCC Maxiren sample as well as in the other two RCC brand

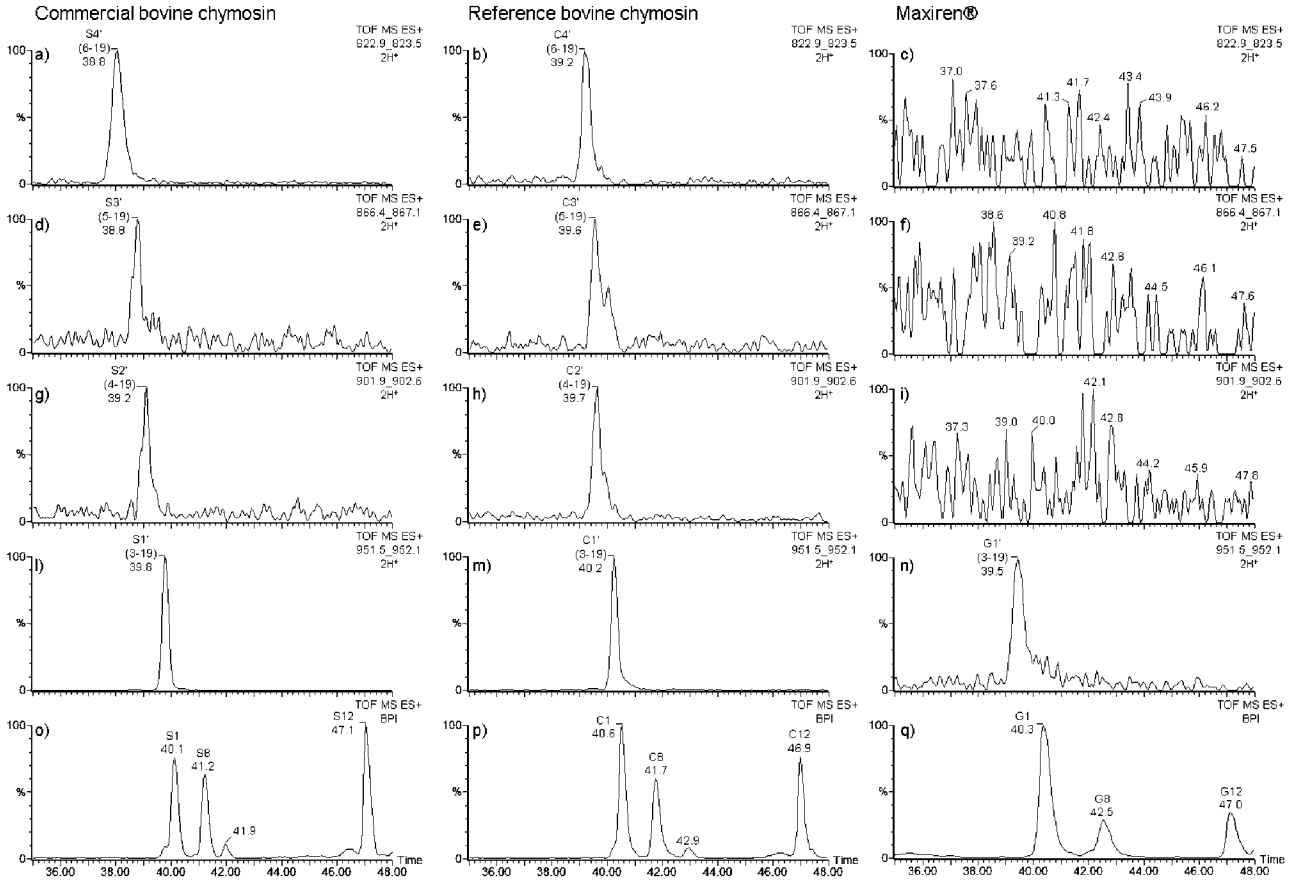

Figure 3. Intensity plot chromatogram of doubly charged tryptic peptide: (3-19) (inset l–n), (4-19) (inset g–i), (5-19) (inset d–f), and (6-19) (inset a–c) of two bovine chymosin preparations and recombinant chymosin Maxiren sample, each compared with the respective chromatographic profile (inset o–q). The signal intensities are scaled to the highest peak (set to 100%) within the respective traces.

preparations (**Table 1**). The presence of a measurable amount of peptides (4-19), (5-19), and (6-19) shorter than (1-19) in the tryptic digest of the two bovine chymosin natural samples could be due to the presence of chymosin species formed during activation of prochymosin. During the autoactivation of chymosin in acid medium, Phe^{42p} mainly directs the cleavage at the Gly^{1} site. The results also indicated that the prochymosin domain was preferentially cleaved at two further sites, after Glu^{2} and Tyr^{39p}, regardless of the proenzyme source. Therefore, the difference between natural and RCC samples was limited to five peptides of which two were specific of recombinant and three of natural chymosin preparations arising from the autocatalytic activation step of prochymosin. The medium composition and environment could conceivably impact the differential cleavage of the protein sites. Therefore, the activation mechanism of two chymosin sources cannot be predicted ex post as the process was not significantly altered in the natural and recombinant prochymosin autoactivation.

This result provided evidence that peptides (4-19), (5-19), and (6-19) are markers for the bovine chymosin. The presence of the species 40p-323 (suffix "p" denotes residues in the prosegment region of chymosin) in natural chymosin signaled by Lilla et al. (*2*) was further confirmed here in the two reference bovine chymosin samples (**Table 2**). Tryptic peptides S3* from commercial chymosin and C3* from reference chymosin sample with 2378.20 and 2378.18 Da molecular mass values, respectively, corresponded each to the (40p-19) peptide starting with the sequence Ser-Gly-Phe followed by the 19 residues of the N-terminal peptide (1-19) (expected molecular mass, 2378.14 Da). This sequence actually corresponds to the C-terminal end of the prosegment whose excision converts prochymosin into the active chymosin 1-323. Identification of this peptide means that a third excision site, in addition to Phe^{27p}-Leu^{28p} of pseudochymosin and Phe^{42p}-Glu^{1} of mature chymosin (*6*), originated from the hydrolysis of the peptide bond Tyr^{39p}-Ser^{40p} of the prosegment for bovine chymosin. A counterpart peptide with a similar molecular mass value was detected in the three RCC samples and labeled as G3*, H3*, and P3* (**Table 1**). The tryptic map of the RCC samples contained other N-terminal tryptic peptides, which originated from protein forms longer than the mature enzyme. The Maxiren sample contained the peptide G1* with a molecular mass of 2669.23 Da (**Table 1**). This peptide was isolated by RP-HPLC, and the amino acid sequence was determined by ESI-TOF MS/MS analysis. The resulting spectrum shown in **Figure 4** inferred the sequence (38p-19), a 24 residue long peptide (expected mass, 2669.30 Da), having as N-terminal residues Lys-Tyr-Ser-Gly-Phe- linked to the 19 residue long N-terminal peptide G1, as shown in the top spectrum.

The peptide (38p-19) should derive from the 38p-323 protein as the N-terminal Lys residue undergoes hydrolysis by trypsin slower than the internal ones. The sequence of the other peptide G2*, characterized in a similar manner, actually corresponds to that of the (39p-19) peptide with the N-terminal sequence Tyr-Ser-Gly-Phe tetramer being the C-terminal end of prochymosin, followed by the 19 residues of N-terminal peptide G1 (data not shown). The measured molecular mass, 2541.16 Da, exactly matches the expected 2541.20 Da value. These peptides also occurred in the other two RCC preparations (**Table 1**).

The intensity plot chromatogram of the (38p-19) triply (**Figure 5** panels a and b) and (39p-19) doubly charged ions (panels d and e) generated from the TICs of RCC samples

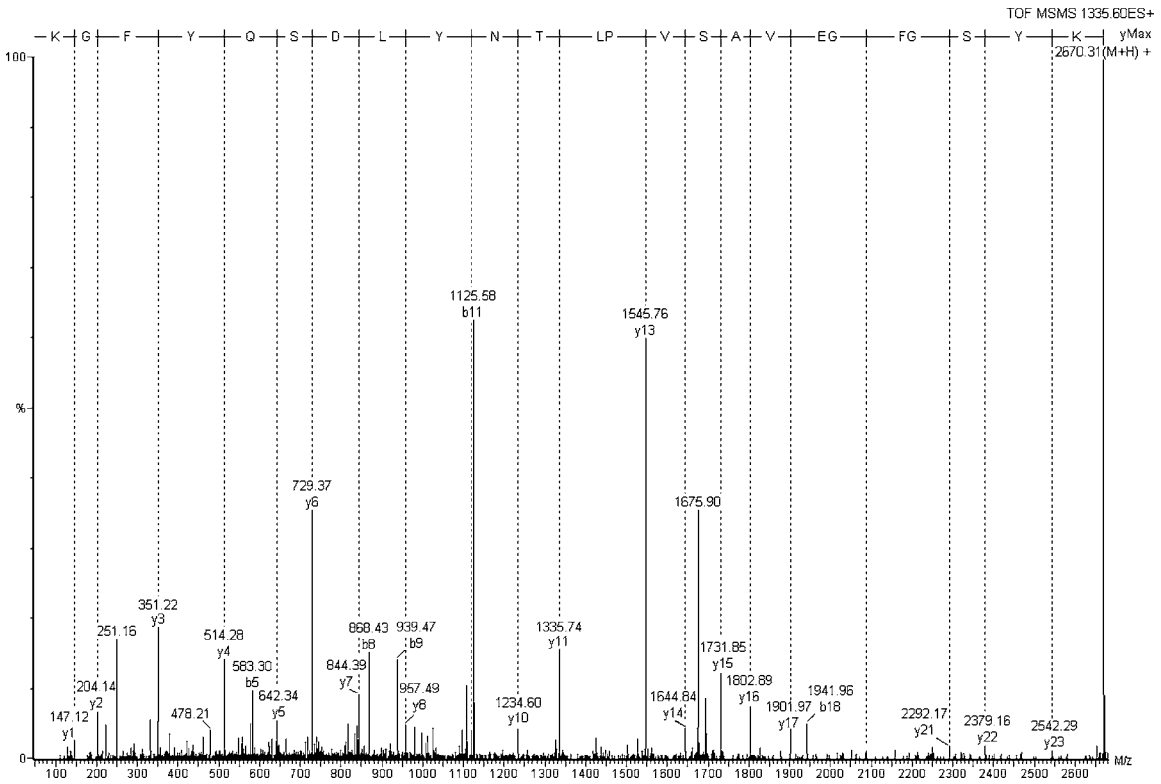

Figure 4. ESI-Q-TOFMS/MS spectrum of the 24 residue long tryptic peptide (38p-19) (G1* in **Table 1**) from recombinant chymosin of Maxiren sample. The amino acid sequence reported on the spectrum indicates the presence of the KYSGF N-terminal sequence of the C-terminal prosegment.

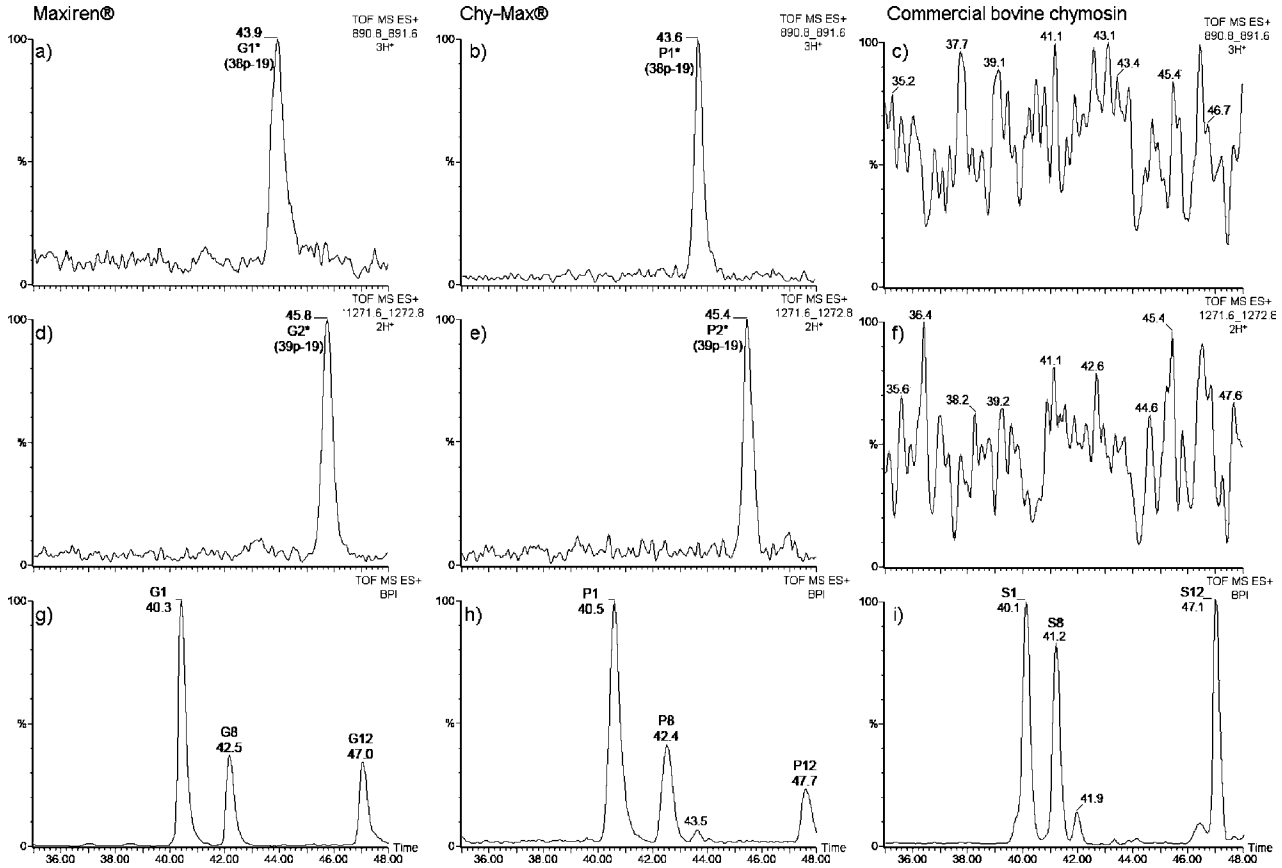

Figure 5. Intensity plot chromatogram profile of triply charged tryptic peptides (38p-19) and doubly charged tryptic peptides (39p-19) relative to the Maxiren (inset a and d), Chy-Max (inset b and e), and commercial bovine chymosin sample (inset d and f) as compared with the respective chromatographic profiles (inset g–i). The signal intensities are scaled to the highest peak (set to 100%) within the respective traces.

Maxiren and Chy-Max was compared with the counterparts of commercial bovine chymosin samples (panels c and f). A further comparison was made with the respective chromatographic profiles within 35 and 48 min (panels g−i). The longer

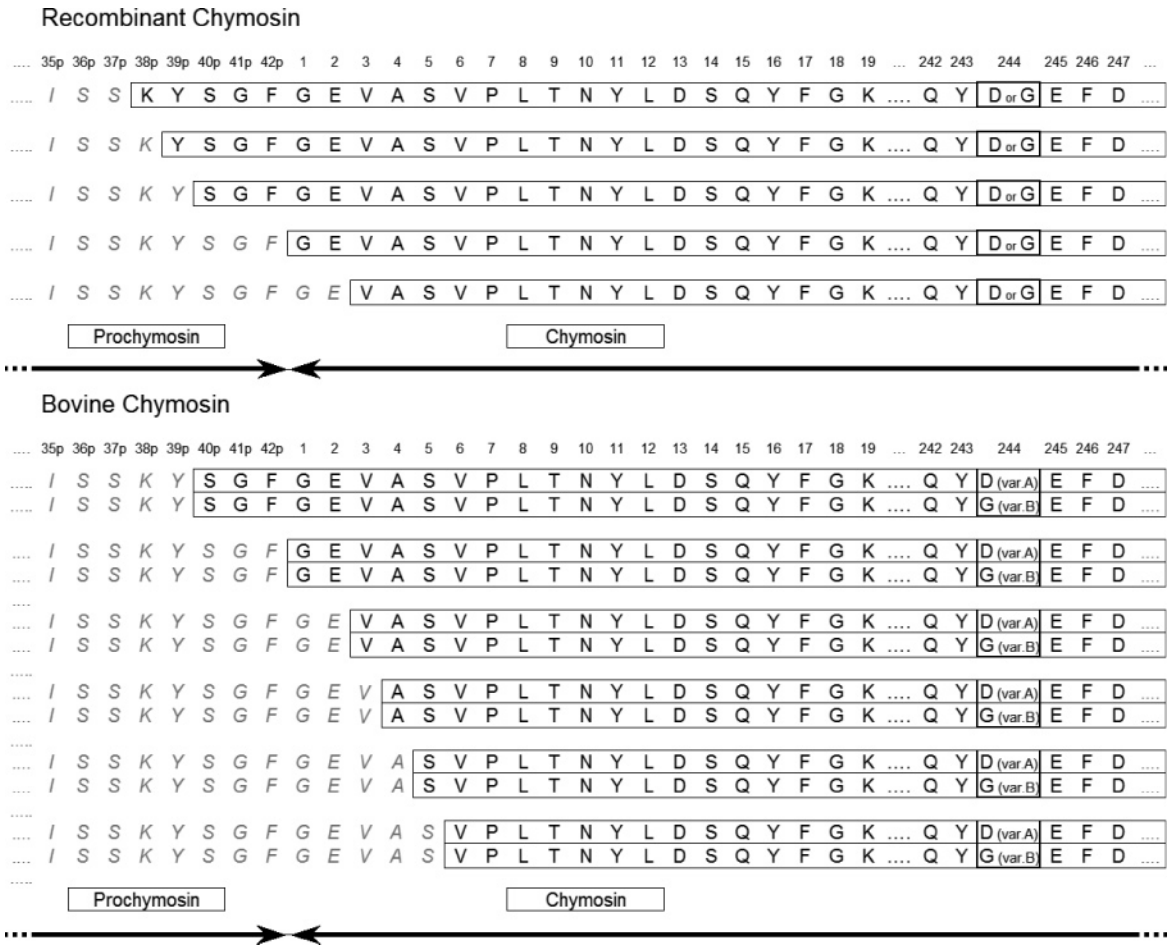

Figure 6. Amino acid sequence of the "strategic" region of bovine and recombinant prochymosin and the chymosin species detected in the present study (in bold). The amino acid substitution at site 244 differentiates A and B variants copresent in bovine chymosin samples.

N-terminal tryptic peptide species behaved similarly, eluting at similar retention times regardless of the RCC brand included the Chymogen sample (refer to **Table 1**).

On the other hand, only a noise signal was detected at the expected retention time for the commercial bovine chymosin sample (the same result was obtained for reference bovine chymosin sample), which indicates that both the (38p-19) and the (39p-19) (panels c and f) peptides were missing in bovine chymosin.

The chymosin samples, either from natural or fermentation sources, contained each peptide (40p-19) whereas (38p-19) and (39p-19) peptides only occurred in the RCC samples. The three parent chymosin species in the RCC samples could probably form during acid prochymosin activation irrespective of the engineered host microorganism. These findings suggest that N-terminal chymosin peptides (38p-19) and (39p-19) could each be used as chemical markers of chymosin from a fermentative source.

Chemistry of Chymosin. In **Figure 6**, the results on the chymosin enzyme from different sources, natural or recombinant, are summarized including those resulting from the present study. The recombinant chymosin commercial samples were either A or B variants at variance with the natural one, which simultaneously contained both of the allelic forms A and B in variable proportions.

The presence of different length protein species in enzyme preparations is not new. Sequencing of pepsin B produced a double sequence, which started at Ala^{1} and Val^{2}. The sequence starting with Val^{2} constituted about 30% of the sequence starting with Ala^{1} (*7*). Pseudo-pepsin B prepared by activation of pepsinogen B also showed a double sequence indicating a cleavage in front of Glu^{17p} and Leu^{22p}. The two sequences occurred at a ratio of 4:1. For a long time, mature chymosin was believed to be produced autocatalytically at acidic pH as a single molecule. However, the present results confirm that natural chymosin is simply heterogeneous. A similar feature has been observed for three recombinant commercial preparations. Little information is available on the vulnerability of the peptide bonds around the active enzyme N terminus during autolytic acid activation of prochymosin. Bovine preprochymosin, a 365 residue long in vivo synthesized protein, is excreted across the cell membrane by losing the 16 residue long signal sequence. The prochymosin (molecular mass, 40477 Da variant B) occurring in calf tissues is converted at around pH 2 into active pseudo-chymosin, a form longer than mature enzyme arising from the Phe^{27p}-Leu^{28p} bond hydrolysis and accounting for 337 amino acid residues. Pseudo-chymosin is converted into chymosin (323 amino acid residues) by further proteolytic cleavage removing its first 15 amino acid residues by rising pH to ~5.5 (*6*). Because prochymosin is only essential for correct folding of the polypeptide chain, recombinant chymosin is usually obtained by transferring the prochymosin gene and recovering the excreted protein counterpart. The autolytic acid activation of the prochymosin would afford, as in the case of the natural counterpart, the formation of pseudochymosin or mature chymosin. The A and B chymosin variants are also indicated in **Figure 6**. Our results additionally show that in bovine chymosin samples, different length enzymes for each

of the two genetic variants are present. The N-terminal tryptic peptides of (1-323), (3-323), and (40p-323), together with minor components, (4-323), (5-323), and (6-323), have been isolated in these samples. In the same way, the protein species (1-323), (38p-323), (39p-323), and (40p-323), together with the shorter form (3-323), occur as a single genetic variant in all of the RCC samples. Summarizing the results shown in **Figure 6** with the fact that two protein species are exclusive of recombinant and three of natural chymosin, one can deduce that both are quite differently excised at the N-terminal end.

DISCUSSION

The purpose of our experiments was to evaluate the amino acid sequence of three commercially available DNA RCC samples to determine if this species was of similar or identical quality with respect to the natural one. Protein mapping by MS revealed that chymosin expressed in the three cases was equivalent although bacteria or yeast cells were the host organisms for the production of heterologous protein. The quality of coexpressed proteins evaluated by MS indicated that the production system and the extraction procedure of chymosin efficiently afforded a high-quality heterologous protein. The amino acid sequence of RCC controlled by MS also excluded the difference with the natural chymosin as various possible posttranslational modifications usually performed by bacteria and yeast, glycosylation being the most common, were missing. Chymosin as well as other enzymes are not glycosylated in nature and therefore must be without any glycosylation pattern in recombinant protein to ensure the correct proteolytic efficacy. The bioengineered chymosin, even if expressed in three different systems, resulted in an authentic protein product. Among the three disulfide bonds of chymosin, Cys^{207}-Cys^{211} was found indispensable for the correct refolding at variance with the dispensable Cys^{47}-Cys^{52} also contributing to substrate stability and specificity of the enzyme (*8*). Further studies have demonstrated that Cys^{206}-Cys^{210} is also dispensable for refolding. It has been demonstrated that the individual prochymosin disulfide bonds have different functions from refolding, but the role of each cysteine residue of a disulfide bond could be also different (*9*). It is worthwhile to mention that although each RCC commercial sample was obtained in different host microorganisms with specific procedures, disulfide bonds were correctly located as in the native natural prochymosin (chymosin). Although rennet is known since ancient times, our understanding of the role of secondary enzyme activation is still insufficient. The only difference acquired within this work was a limited heterogeneity of the polypeptide chain in two chymosin sources. The autoproteolytic activation resulted in a series of differently processed prosegment species as partly shown in **Figure 6**. Envisioning the cleavage specificity, the propeptide should be a reasonably good substrate having the main peptide bond Phe^{42p}-Gly^{1} cleaved. Autoproteolytic activation of prochymosin at low pH could respond to a process carried out by the enzyme itself. On the other hand, the milk clotting activity by pseudo- or mature chymosin starts by splitting the Phe^{105}-Met^{106} of κ-casein, a similarly hydrophobic peptide bond as the Phe^{42p}-Gly^{1} of activated prochymosin. Several lines of evidence exist showing that mutant enzymes are resistant to autoproteolytic processing. This would rule out the possibility of activation cleavage by an accidentally copurified bacterial protease or by chymosin itself. Therefore, it is very likely that the prochymosin is cut in an intramolecular fashion. The activation mechanism leads to 1-323 long protein species directly measured by ESI-MS and confirmed by the isolation of the N-terminal tryptic peptide (1-19) from the tryptic digest of any chymosin preparation. The presence of low amounts of shorter and longer forms of the N-terminal tryptic peptides is indicative of a more complex proteolytic cleavage of peptide bonds within the prosegment and the mature enzyme. In the bovine chymosin, neither (38p-19) nor (39p-19) peptides were formed while (40p-19) was common to RCC samples, demonstrating that activation of the same zymogen proceeded through different intermediate forms, probably dependent on the differently complex medium of the animal tissue extract and the cultured microorganism. The coextracted tissue bovine serum albumin (BSA) and lysozyme might exert different functions among which an antioxidant effect or a protection of chymosin integrity against copresent enzymes.

In proastacin activation, an initial event catalyzed by endogenous trypsin was shown to produce a series of premature astacin species shorter than mature enzyme. Such premature derivatives were found responsible for subsequent cleavages resulting in the mature enzyme (*10*). The activation mechanism of proastacin or other metalloproteinases could exhibit similarities with that of chymosin. An initial cut in the so-called bait region of protein is followed by autoproteolytic cleavages, which release the mature enzyme (*11*).

The observed cleavage within the N-terminal end created a number of truncated forms of chymosin species. Molecular species (4-323), (5-323), and (6-323) were exclusively formed in the natural enzyme preparation, which seems to confirm that enzyme-accompanying BSA and lysozyme could/may only function as antioxidant or antibacterial agents. However, no information about the origin of the peptide bond cleavage according to the possible mechanisms direct, sequential, or concurrent is currently available. An element that might be expected to have an important effect on the enzyme posttranslational processing is the identification of the proteolytic activities in the animal tissues coextracted with zymogen or occurring in the culture medium. Surprisingly, there were only minor changes in the number of the protein species within two different sources possibly formed according to common mechanisms. Therefore, the ability to produce comparable N-truncated enzyme forms could most probably depend on the proximity of the N-terminal end of chymosin to the enzyme active site. In proastacin conversion to astacin, the precise trimming of the amino terminus seems a key element since it enables the correct formation of the buried salt bridge between the ammonium group and the carboxyl group of Glu^{103} important for the structural stability (*10*). In contrast, the autoactivation of procathepsin B was determined as an intermolecular rather than intramolecular process, which would be expected in a concentration-independent process. In presence of 20% glycerol, a ~4-fold deceleration of auto processing was observed, whereas in the case of an intramolecular (unimolecular) process, glycerol should not have affected the processing rate (*12*). Moreover, the addition of small amounts of mature active enzyme accelerated the process ~3-fold, indicating that mature cathepsin B is the catalytic species in the process (*12*). Whatever the mechanisms involved, thanks to MS, the novel enzyme forms that were identified mean that activation processes such as this or maybe other aspartic protease need to be further studied. Also important is the study of the zymogen activation under in vivo conditions to provide further evidence of the similarity and the difference in the processes. In summary, the method using MS is a generally applicable tool to rapidly characterize recombinant proteins to evaluate fine amino acid sequence differences with respect to the natural counterpart.

ACKNOWLEDGMENT

This work was carried out without any funding from public institutions or private companies. We thank Prof. Gilberto de Nucci from the Department of Pharmacology of Institute of Biomedical Sciences of University of São Paulo (Brazil) for the access to the Q-TOF Ultima instrumentation.

LITERATURE CITED

(1) Andren, A.; Bjorck, L. Milk-feeding maintains the prochymosin production in cells of bovine abomasal mucosa. *Acta Physiol. Scand.* **1986**, *126*, 419−427.

(2) Lilla, S.; Caira, S.; Ferranti, P.; Addeo, F. Mass spectrometric characterisation of proteins in rennet and in chymosin-based milk-clotting preparations. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1101−1112.

(3) Swinkels, B. W.; van Ooyen, A. J.; Bonekamp, F. J. The yeast *Kluyveromyces lactis* as an efficient host for heterologous gene expression. *Antonie van Leeuwenhoek Int. J. G.* **1993**, *64*, 187−201.

(4) Collin, J. C.; Moulins, I.; Rolet-Répécaud, O.; Bailly, C.; Lagarde, G. Detection of recombinant chymosins in calf rennet by enzyme-linked immunoassorbent assay. *Lait* **1997**, *77*, 425−431.

(5) Foltmann, B. Prochymosin and chymosin (prorennin and rennin). *Biochem. J.* **1969**, *115*, 3−4.

(6) Pedersen, V. B.; Christensen, K. A.; Foltmann, B. Investigations on the activation of bovine prochymosin. *Eur. J. Biochem.* **1979**, *94*, 573−580.

(7) Nielsen, P. K.; Foltmann, B. Purification and characterization of porcine pepsinogen B and pepsin B. *Arch. Biochem. Biophys.* **1995**, *322*, 417−422.

(8) Wei, C.; Tang, B.; Zhang, Y.; Yang, K. Oxidative refolding of recombinant prochymosin. *Biochem. J.* **1999**, *340*, 345−351.

(9) Chen, H.; Zhang, G.; Zhang, Y.; Dong, Y.; Yang, K. Functional implications of disulfide bond, Cys^{206}-Cys^{210}, in recombinant prochymosin (chymosin). *Biochemistry* **2000**, *39*, 12140−12148.

(10) Yallorous, I.; Kappelhoff, R.; Schilling, O.; Wegmann, F.; Helms, M. W.; Auge, A.; Brachtendorf, G.; Grosse Berkhoff, E.; Beermann, B.; Hinz, H.-J.; Konig, S.; Peter-Katalinic, J.; Stocker, W. Activation mechanism of pro-astacin: role of the propeptide, tryptic and autoproteolytic cleavage and importance of precise amino-terminal processing. *J. Mol. Biol.* **2002**, *324*, 237−246.

(11) Nagase, H. Activation mechanisms of matrix metalloproteinases. *Biol. Chem.* **1997**, *378*, 151−158.

(12) Rozman, J.; Syojan, J.; Kuhelj, R.; Turk, V.; Turk, B. Autocatalytic processing of recombinant human procathepsin B is a bimolecular process. *FEBS Lett.* **1999**, *459*, 358−362.

Received for review December 28, 2004. Revised manuscript received May 6, 2005. Accepted May 9, 2005.

JF0478051