ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Biotechnology 123 (2006) 443–452

Journal of BIOTECHNOLOGY

www.elsevier.com/locate/jbiotec

Structural aspects of the *Mucor bacilliformis* proteinase, a new member of the aspartyl-proteinase family

Claudia Machalinski, Marìa L. Pirpignani, Cristina Marino, Adriana Mantegazza, Mirtha Biscoglio de Jiménez Bonino *

Instituto de Química y Fisicoquímica Biológicas (UBA-CONICET), Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Junín 956 (1113), Buenos Aires, Argentina

Received 15 July 2005; received in revised form 2 December 2005; accepted 23 December 2005

Abstract

Bovine chymosin is considered the best milk-clotting enzyme for cheese manufacture; however, the thermophilic *Mucor pusillus* proteinase is also used nowadays. We herein report structural aspects of the aspartyl proteinase from the local mesophilic *Mucor bacilliformis* strain. Sequence data indicate a high similarity degree to those of other family members. The protein is monomeric, not glycosylated, has two disulfide bridges, and mainly includes beta structure. A molecular model was built by using the *Rhizopus chinensis* proteinase structure as the template. Sequence analysis and comparison of our model with bovine chymosin and *M. pusillus* proteinase structures, indicate that the *M. bacilliformis* proteinase is at a similar evolutionary distance on a sequence level; as regards tertiary structure, the *M. bacilliformis* proteinase superimposes on the bovine chymosin structure in a fashion similar to that of the *M. pusillus* proteinase. Overall results suggest that this novel proteinase can be utilized as a good milk-clotting enzyme in the dairy industry.
© 2006 Elsevier B.V. All rights reserved.

Keywords: *Mucor bacilliformis* proteinase; Aspartyl-proteinase family; Mesophilic proteinases; Cheese industry; Primary structure; Milk-clotting enzymes

1. Introduction

Aspartyl proteinases are a protein family widely distributed in nature and involved in multiple important commercial and biomedical processes.

We are interested in basic and applied aspects of the aspartyl proteinase from *Mucor bacilliformis*, a mesophilic mould strain. Fungal thermophilic proteinases – such as those from *Mucor miehei* and *Mucor*

Abbreviations: ACN, acetonitrile; CNBr, cyanogen bromide; CNBr/KI, cyanogen bromide/potassium iodide; PVDF, polyvinylidene difluoride; 4-VP, 4-vinlpyridine; TFA, trifluoracetic acid

* Corresponding author. Tel.: +54 11 49648290; fax: +54 11 49625457.
E-mail address: mbiscoglio@hotmail.com (M.B. de Jiménez Bonino).

0168-1656/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbiotec.2005.12.035

pusillus – are still widely used as substitutes for calf chymosin in the cheese industry; however, these proteinases have a high proteolytic action which impairs cheese organoleptic characteristics. Despite over 50% cheese now being produced by using recombinant chymosin, animal or microbial rennets is the alternative for consumers who refuse to accept bioengineered food.

We isolated, purified (Areces et al., 1992) and characterized structural and kinetic aspects (Venera et al., 1997) of the aspartyl proteinase from *M. bacilliformis*. In this work we have further characterized the protein including the determination of the complete primary structure and built a molecular model by using the crystallographic structure of the *Rhizopus chinensis* proteinase as the template.

Sequence alignment and molecular modeling have enabled us to describe structural similarities and differences between the *M. bacilliformis* proteinase and other members of the protein family. The comparison between the spatial arrangement of this protein and those of bovine chymosin and the *M. pusillus* proteinase is of particular interest. Bovine chymosin has been considered as a prototype in cheese manufacture since it has the best milk-clotting/proteolytic activity ratio; on the other hand, the *M. pusillus* proteinase is currently being used as a substitute milk-clotting enzyme.

2. Materials and methods

The *M. bacilliformis* proteinase was prepared by the method of Venera et al. (1997). Homogeneity of the enzyme preparation was examined by SDS-polyacrylamide gel electrophoresis and N-terminal analysis. Trypsin, chymotrypsin, Staphylococcal V8 proteinase and molecular weight markers were obtained from Sigma–Aldrich, Saint Louis, MO, USA. PVDF and nitrocellulose membranes were from Millipore Corporation, Bedford, MA, USA and the Immun-Blot® kit for glycoprotein detection was from BIO-RAD, Hercules, CA, USA. Mono Q HR 5/5 and PD-10 columns were from Amersham Biosciences, Uppsala, Sweden. Water was Milli Q quality and all other chemicals were AR grade.

2.1. Protein concentration

This was measured either as per Lowry et al. (1951) or by amino acid analysis.

2.2. Amino acid analysis

Samples were hydrolyzed in vacuum-sealed tubes at 110 °C for 20 h in constant boiling HCl containing phenol (1 mg/ml). The hydrolysates were analyzed in duplicate in a Beckman 119 CL amino acid analyzer.

2.3. SDS-polyacrylamide gel electrophoresis

It was performed in slab gels at room temperature under denaturing conditions as described by Laemmli (1970) or according to Schägger and von Jagow (1987). Gels were stained for proteins with Coomassie Blue R-250 or silver staining.

2.4. Cysteine and disulfide bridge contents

Protein reduction and carbamidomethylation was performed according to Crestfield et al. (1963) but using ^{14}C iodoacetamide instead of iodoacetic acid. Reagent excess was separated through a Sephadex G25 column. Fraction radioactivity was measured in a Beckman 7000 counter.

2.5. Peptide sequencing

It was performed at the Facility for Peptide and Protein Sequencing (University of Buenos Aires). An Applied Biosystems pulse liquid–gas phase automated sequencer model 477A was used. Following each cycle, the phenylthiohydantoin amino acid was identified by using an on-line RP-HPLC system. Samples were electroblotted onto a PVDF membrane or applied to a polybrene-coated filter.

Searches for similarities were performed by using the Blast algorithm.

2.6. Protein and peptide electroblotting

After SDS-PAGE, proteins or peptides were blotted onto a PVDF membrane in 10 mM CAPS buffer, 10% methanol, pH 11.0, according to Matsudaira (1987).

2.7. Protein reduction and alkylation

Reduction was achieved with a 10-fold excess of DTT in 6 M guanidinium hydrochloride, 0.25 M Tris–HCl, 1 mM EDTA, pH 8.5, for 1 h at 37 °C under nitrogen (Matsudaira, 1989); the reduced protein was then alkylated with 4-VP (1 μl/100 μg of protein) under nitrogen, at 37 °C during 1 h. Reagent excess was removed through a PD-10 column.

2.8. Proteolytic digestion in solution

Approximately 150 μg of the reduced and 4-VP-treated proteinase was digested with trypsin during 24 h at 37 °C in 100 mM sodium bicarbonate, pH 8.5, at a 1:25 enzyme:protein ratio. *Staphylococcal V8* digestion of the native proteinase was carried out in a 0.1 M phosphate buffer, pH 7.8, at 37 °C for 16 h at a 1:40 enzyme:protein ratio.

2.9. Chemical cleavage

(a) Four hundred micrograms of the reduced and 4-VP-treated *M. bacilliformis* proteinase was digested in 400 μl of 70% formic acid with a two-fold excess of CNBr, in the dark, under nitrogen, at room temperature for 24 h (Matsudaira, 1989).
(b) One hundred micrograms of the proteinase was digested in 240 μl of 70% formic acid containing 200 μg of CNBr and 1 mM KI (Huang and Huang, 1994) under the conditions described in (a). In both cases, after a 10-fold dilution with water the solution was concentrated in a Speed Vac.

2.10. "In-gel" digestion

Protein bands were digested "in-gel" with chymotrypsin at an enzyme:protein ratio of 1:20 (Hellman, 1997). Briefly, after washing with 0.1 M ammonium bicarbonate and ACN, the gel piece was dried up under nitrogen, and a buffer solution containing chymotrypsin was allowed to soak into the gel piece. After overnight incubation at 30 °C, the peptide mixture was recovered by extraction with 60% ACN in 0.1% TFA and analyzed by RP-HPLC.

2.11. HPLC peptide purification

An LKB system from Amersham Biosciences, Uppsala, Sweden, with a diode array detector was utilized. Experimental conditions are detailed for each particular case. When necessary, fractions were rechromatographed.

2.12. Circular dichroism

It was performed in a Jasco J20 spectropolarimeter in the 203–250 nm range at 25 °C with a cell provided with a circular window. The pathlength was 1 mm and the instrument was calibrated with camphorsulfonic acid. Protein concentration – 0.15 mg/ml in 50 mM phosphate buffer with 110 mM NaCl – was determined in quintuplicate. The content of α-helix, β-sheet and disordered structure (Andrade et al., 1993; Merelo et al., 1994) was determined by using the K2d program described by the European Molecular Biology Laboratory. The α-helix content was also estimated by the Zhong and Curtis Johnson method (1992).

2.13. Glycosylation state

An Immuno-Blot Kit for glycoprotein detection was utilized.

2.14. Protein modeling

Sequence of the *M. bacilliformis* proteinase and its template (*R. chinensis* proteinase) were aligned by CLUSTALX (Higgins et al., 1996) followed by manual adjustment. The structure-based alignment of the *M. bacilliformis* model and the aspartyl proteinase HOMSTRAD family (Mizuguchi et al., 1998a) was obtained by using COMPARER (Sali and Blundell, 1990) and the corresponding and output file is given in JOY format (Mizuguchi et al., 1998b). The different structure superimposition and rmsd calculation made during the work were performed by the server for Pairwise Superimposition of Protein 3D-structures (Boutonnet et al., 1995). The model was built using MODELLER (Sali and Blundell, 1993; Fiser et al., 2000). Evaluation of the model was carried out with Verify 3D (Luthy et al., 1992) and Procheck (Laskowski et al., 1993). The realignment cycle, modeling and structure evaluation were repeated until no further improvements on the

structure were observed. Hydrogen bonds of the model and known protein structures were calculated with CSU software (Sobolev et al., 1997) for comparison purposes. The known crystal structure of the *R. chinensis* aspartyl proteinase (pdb code: 2apr.pdb) was chosen as the template for the homology modeling of the *M. bacilliformis* aspartyl proteinase as they have a 67% sequence identity.

3. Results and discussion

3.1. Protein purification

The protein was purified as described by Venera et al. (1997) and met the purity criteria indicated by the authors. Besides, the only protein band present in an SDS-PAGE showing an apparent molecular weight of 32 kDa was electroblotted onto a PVDF membrane and the sequence of the first 33 amino acid residues was then determined (Fig. 1).

3.2. Amino acid sequence

Sequencing of HPLC-purified peptides from the mixtures obtained by trypsin digestion of the 4-VP-treated protein (T1–T21, Fig. 1) and V8 digestion of the native protein (V1–V5, Fig. 1) and their further localization in the sequence by homology with those from other members of the protein family allowed determination of 55.4% of the *M. bacilliformis* protein primary structure, mainly involving the N-terminal region (Fig. 1). Two peptides containing Cys residues were obtained (T17 and T18, Fig. 1).

Distribution of Met residues in homologous proteinases indicated that the CNBr treatment could furnish useful information concerning the missing region. Therefore, the reduced and 4-VP-treated protein was submitted to CNBr cleavage and the peptide mixture resolved by SDS-PAGE (Fig. 2); the two fragments obtained (A and B) were sequenced: A corresponded to an N-terminal fragment and B had Gly 122 as its N-terminal residue. Fragment B was submitted to a chymotryptic “in-gel” digestion and further HPLC of the peptide mixture. Results are shown in Fig. 1. The experiment allowed us to reach 91.6% of the total sequence and important overlapping was achieved. However, other overlappings as well as sequence of segments 47–53, 156–167 and 193–200 had not yet been determined.

On the other hand, Met and Trp distribution in other members of the family led us to utilize the cleavage of the native protein with a CNBr/KI mixture. Fragments were separated (Fig. 3) by SDS-PAGE under the conditions described by Schägger and von Jagow (1987), electroblotted onto a PVDF membrane and submitted to microsequencing. Four out of the nine fragments sequenced (Fig. 1) bore important results: CK2 (42–56) covered the missing sequence and represented an overlapping—though the two cysteine residues could not be identified; CK6 (156–175) covered the second missing sequence and overlapped the peptide following in the structure; CK7 (194–202) and CK8 (195–209) covered the third missing sequence, except W193, and also overlapped the next peptide. The experiment allowed completion of 99.1% of the total sequence. For the purpose of confirming the 42–56 sequence, fragment A (Fig. 2) was digested “in gel” with chymotrypsin; only one peptide showed absorbance at 254 nm and its sequence unambiguously confirmed the presence of cysteine residues and location of one of the two disulfide bridges. All the above results led to the determination of the complete protein primary structure (Fig. 1) although Trp was placed at position 193 on the basis of homology and chemical cleavage specificity. The protein is monomeric, has 323 amino acid residues and – as judged by ^{14}C iodoacetamide incorporation to both the reduced and native proteins – has four cystein residues involved in two disulfide bridges. Members of the aspartyl-proteinase family have one to three disulfide bridges; that located at position 251–286 is conserved in all members of the family. Proteinases containing four cysteine residues have two bridges: 47–50 and 251–286. Apart from localizing them by homology in the *M. bacilliformis* proteinase, we confirmed their positions by protein modeling as described below.

Besides, two protein microheterogeneities were found: peptides containing W or F at position 71 and those containing G or D at position 178 were purified.

On the other hand, the *M. bacilliformis* proteinase is not glycosylated. Although the importance of glycosylation is not well known in the aspartyl-proteinase family, it has been proposed that it stabilizes protein

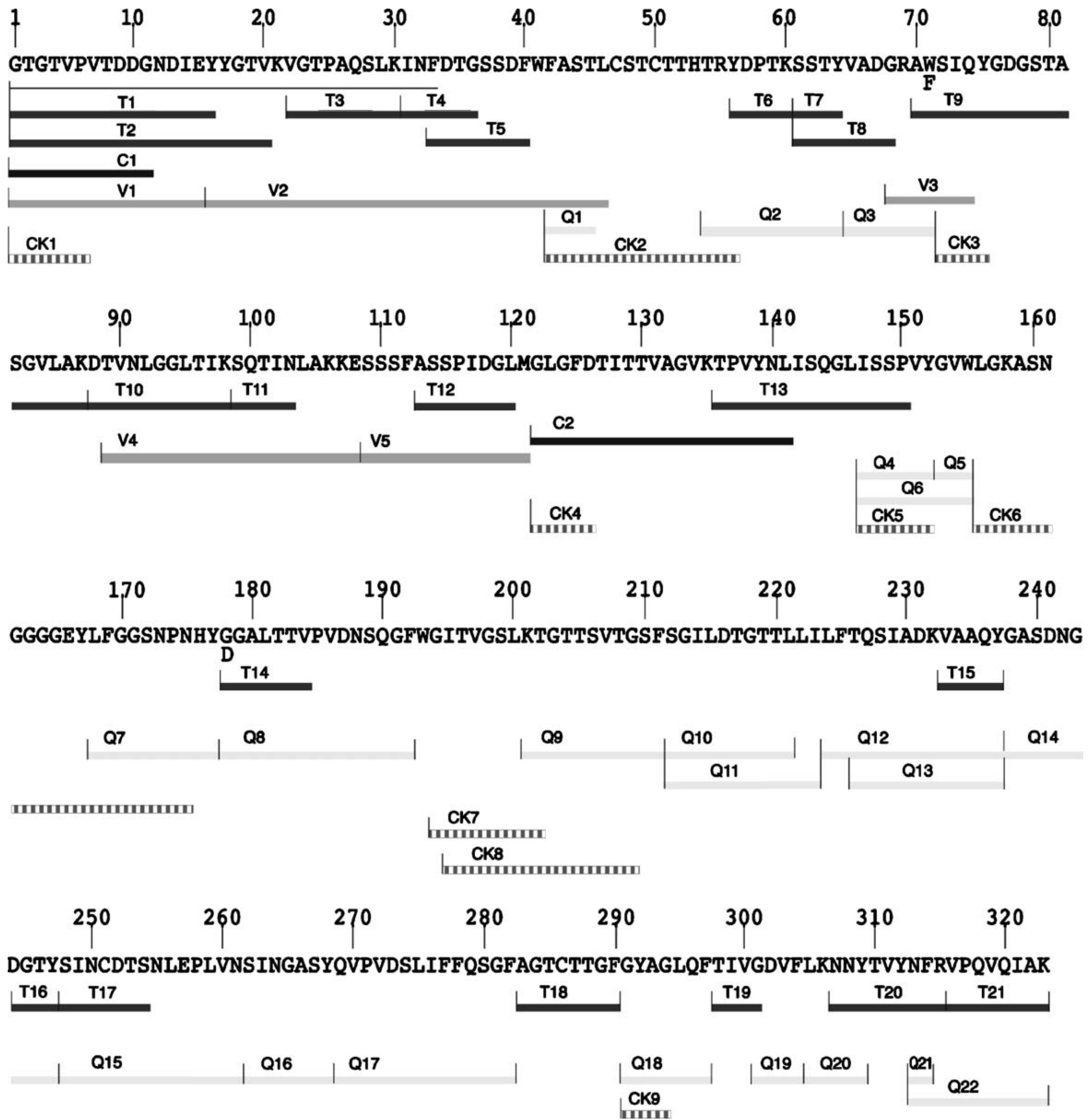

Fig. 1. Primary structure of the *M. bacilliformis* aspartyl proteinase. Thin black line underlines the sequence obtained from the native protein. Peptides obtained from enzymatic and chemical cleavage: trypsin (dark gray full lines); chymotrypsin: (light gray full lines); V8 (medium gray full lines); CNBr (wide black lines); CNBr/KI (dashed gray lines).

conformation (Costa et al., 1997) thus leading to a higher thermostability. In cheese industry a low thermal stability is a particularly useful property in the maturation process. The thermoplilic *M. miehei* proteinase presents high thermostability and has a large carbohydrate content (6%, two glycosylation sites) (Rickert and McBride-Warren, 1974). In contrast, those from mesophilic mold strains such as rhizopuspepsin from *Rhizopus niveus*, penicillopepsin and endothiapepsin – which are thermolabile – are not glycosylated (Gasteiger et al., 2003) as the case is for bovine chymosin.

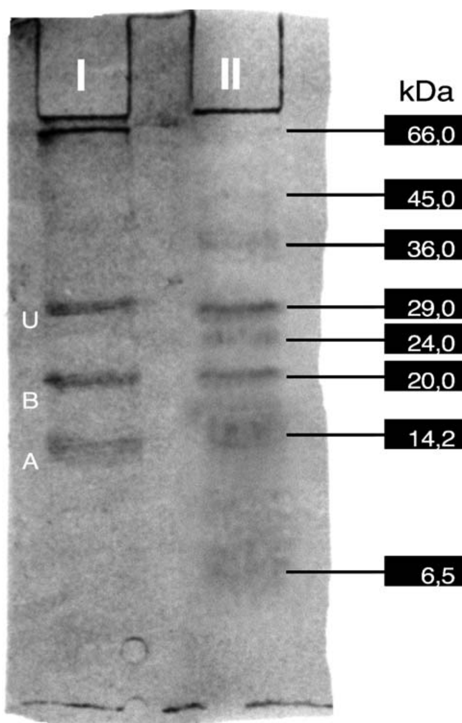

Fig. 2. SDS-PAGE of the reduced and 4-VP-treated *M. bacilliformis* proteinase cleaved by CNBr (Line I). The gel was stained with Coomassie Blue. U: undigested protein. Line II contains molecular markers.

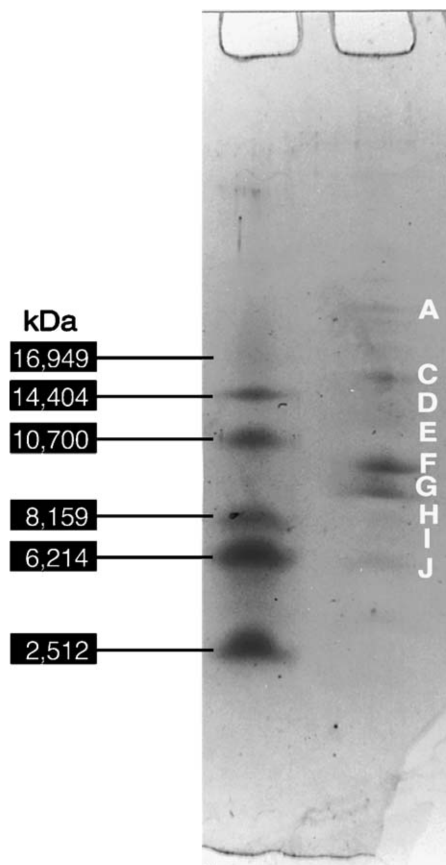

Fig. 3. SDS-PAGE of peptides obtained by CNBr/KI treatment of the *M. bacilliformis* proteinase. Left line contains the molecular mass markers. The gel was stained with Coomassie Blue.

3.3. Sequence alignment and molecular modeling

Fig. 4 shows comparison of the primary and tertiary structure of the *M. bacillormis* proteinase with those from fungal origin including one without a known tertiary structure though important for discussion, and calf chymosin, considered as a prototype in cheese manufacture. There is a high percentage of similarity and identity (McGinnis and Madden, 2004) between the sequence of the *M. bacilliformis* proteinase and that of *R. niveus* I (79% similarity, 65% identity), *R. chinensis* (78% similarity, 67% identity), *Endothia parasitica* (52% similarity, 38% identity), calf chymosin (53% similarity, 33% identity), *M. pusillus* (46% similarity, 32% identity), *M. miehei* (44% similarity, 29% identity). The higher values as well as the lesser gap number are obtained when comparing the *M. bacilliformis* proteinase with that from *Rhizopus* species.

On the other hand, sequence identity of calf chymosin with the *M. bacilliformis* proteinase (33%) is

Fig. 4. Structural comparison of the *M. bacilliformis* proteinase and other family members. *Mucor bacilliformis* proteinase model: muc2; penicillopepsin, *Penicillium janthinellum*: 3app; endothiapepsin, *E. parasitica*: 4ape; rhizopuspepsin, *Rhizopus chinensis*: 2apr; chymosin, *Bos taurus*: 4cms; pepsin, *Mucor pusillus*: 1mpp; aspartic proteinase, *Mucor miehei*: 2asi; aspartic proteinase, *R. niveus*: RHINI_4 (unknown tertiary structure). Alpha helix: red; beta strand: blue; 3_{10} helix: brown; solvent accessible: lower case; solvent inaccessible: upper case; hydrogen bond to main-chain amide: bold; hydrogen bond to main-chain carbonyl: underlined; disulfide bond: cedilla; positive phi torsion angle: italics; amino acids in the hydrogen-bond net involving D34: yellow boxes; amino acids in the hydrogen-bond net involving D216: green boxes, as calculated according to CSU software [23].

```
3app    (  1 )  aasgvatNtPt-anDeeYiTpVtIg--gttLnLnFdTGsADLWVFSteLp-asqgsg
4ape    ( -2 )   stgsatTtpidslDdaYiTpVqIGtpagtLnLdFDTGsSDLWVFSseTt-asevdg
2apr    (  1 )  agvGtVpMtDy-gndieYyGqVtIGtpGkkFnLDFdTgsSDLWIAStlÇt--nÇgsg
4cms    ( -2 )  gevAsVpLtny--ldsqYfgkIyLGtppgeFtVlFdTGsSdFWVPSiyÇk-snAÇkn
1mpp    ( -1 )   gsvdTpGlYD-fdleeYAIpVsIGtpgqdFyLlFdTgsSDTWVphkgÇdnsegÇvg
2asi    (  5 )   gsVdTpGyYD-fdleeYAIpVsIGtpgqdFlLlFdTGsSDTWVphkgÇtksegÇvg
muc2    (  1 )   gtGtVpVtDd-gndieYyGtVkVGtpaqsLkInFdTgsSDFWFAStlÇs---tÇtt
RHINI_4(  1 )   SSGSVPVTDD-GNDIEYYGEVTVGTPGIKLKLDFDTGSSDLWFASTLCT--NCGSS
                                                   D34

3app    ( 54 ) HsvYnPsa--tGkelsgytWsisygdgSsAsgnVftdsVtVggVtAhgQAVQAAqqIs
4ape    ( 54 ) QtiYtPskSttAkllsgAtWsisygdgSsSsgdVyTDtVsVggLtVtgQAVESAkkVs
2apr    ( 55 ) QtkYdpnqSstyqad-grtWsisygdgssAsgiLAkDnVnLggllIkgQtIELAkrEa
4cms    ( 53 ) hqrFdPrkSstfqnl-gkpLsihyg-tgsMqGiLGyDtVtVsnIvDiqQTVGLStqEp
1mpp    ( 55 ) krfFdpssSstfket-dynLnityg-tggAngiYFrDsItVggatVkgQTLAYVdnvs
2asi    ( 60 ) srfFdpsaSstfkat-nynLnityg-t-gAnGlYFeDsIaIgditVtkQiLAyVdnVr
muc2    ( 53 ) htrYdptkSstyvad-graWsiqygdgStAsGvLAkDtVnLggltIksQTINLAkkEs
RHINI_4(54 ) QTKYDPSQSSTYAKD-GRTWSISYGDGSSASGILGKDTVNLGGLKIKNQIIELAKREA

3app    (110) -aqfqqdt-----nnDGlLGLAfssiNtVqp---qsqtTFFdtVks--sLaqpLFAVAL
4ape    (112) -ssfteds-----tiDGlLGLAfstlNtVsp---tqqkTFFdnAka--sLdspVFTAdL
2apr    (112) -asfasg------pnDGLLGLGfdtiTtvr-----gVkTPMdnLisqglIsrpIFGVyL
4cms    (109) gdvFtya------eFdGILGMAYpslAsey------SiPVFdnMmnrhlVaqdlFSVYM
1mpp    (111) g-pTaeQspdselfLDGiFGAAypdnTaMeaeygdtYnTVHvNLykqglIsspvFSVyM
2asi    (115) g-pTaeqspnadifLDGlFGAAypdnTaMeaeygstynTVHvnLykqglIssplFSVyM
muc2    (110) -ssfass------pIDGLMGLGFdtiTtva-----gvkTPvynLisqglIsspvYGVwL
RHINI_4(111) S-SFSSG------PSDGLLGLGFDSITTVSG-----VQTPMDNLISQGLISNPVFGVYL

3app    (158) khg---qpgvYDFGfidsskytgsLtytgVdns---qGfWsFnVdsytAgs--qsg--d
4ape    (160) gyh---apgtYnFGfidttaytgsitytaVstk---qGfWeWtStGyavgsgtfks--t
2apr    (159) GKakngggGeYiFggydstkFkgslttVpIdns---rGwWgItVdrAtvgt-stVa--s
4cms    (156) drng--qeSmLtLGaidpsyytgslhwVpVtv----qqyWqFtVdsVtisg-vvvAÇeg
1mpp    (169) nTnd--ggGqVvFGgvnntLlggdiqyTdVlksrggyffWdAPVtgVkidgsdavsfdg
2asi    (173) nTns--gtGeVVFGgvnntLlggdiayTdVmsryggyyfWdAPVtgItVdgsaavrfsr
muc2    (157) gkasngggGeYlFggsnpnhyggalttVpVdns---qGfWgItvgsLktgt-tsVT--g
RHINI_4(158) GKESNGGGGEYIFGGYDSSKFSGDLTTIAVDNS---NGWYGITIDGASISG-SQVS--D

3app    (207) gfsGIAdTGttlLlLddsVVsqYYsqVs--gAqqdsnAggyVFdçs-t-nL-pdFsVsI
4ape    (211) sIdGIAdTGttlLyLpatVVsaYWaqVs--gAkssssvggyvFpçs-a-tL-psFtFGV
2apr    (212) sfdGILdTGttlLiLpn-niAasVaraY--gAsdng-dgtYtIsçd-tsaf-kpLvFsI
4cms    (208) gçqAILdTgtskLVGps-sdIlnIQqaI--gAtqnq-ygefdIdçdnlsyM-ptVvFeI
1mpp    (226) aqaFtIdTGtnfFiAps-sFAekVVkaALpdAtesq--qGYtVpçskyqdskttFsLvL
2asi    (230) pqaFtIdTGtnfFiMps-saAskIVkaaLpdAtetq--qGwvVpçasyqnskstIsIvM
muc2    (210) sfsGILdTgttlliLftqsIAdkVAaqY--gAsdng-dgtYsIncd-tsnl-epLvnsI
RHINI_4(211) SFSAILDTGTTLLILPS-NVASSVAQAY--NANDNG-DGTYNINCDTSELQ--PLVFTI
                   D 216

3app    (261) s-------gytAtVpgslInygpSgdg-stÇlGGIqsNsgigfSiFGDIFLKsQYVVFd
4ape    (265) g-------sarivIpGdyIdfgpistgsssÇfGGIqsSagiginiFGDVALKAAFVVFn
2apr    (265) n-------gasFqVspdSLvfeef---qgqÇiAGFGyg-nwgfAiIGDTFLKNNYVVFN
4cms    (262) n-------gkmYpLtpsaYTsqd----qgfÇtSgFqse---qkWiLGDvFIreYYSVFD
1mpp    (282) qksgsssdtIdVsVpIskMLlpvdksg-etçmFIVlpd-ggnqfiVGNLFLrfFVNVYD
2asi    (286) qksgsssdtieisVpVskMLlpvdqsn-etçmFIIlpd-GgnqyiVGNLFLrffVNVYD
muc2    (264) n-------gasYqVpvdSLiffqs-gfagtÇtTGFGya-glqfTivGDVFLKNNYTvYn
RHINI_4(264) G-------GSTFEVPTDSLIFEQD---GNTCVAGFGYG-QDDFAIFGDVFLKNNYVVFN

3app    (312) sd-gpqLGFApqa
4ape    (317) gattptLGFAsk
2apr    (313) qg-vpeVqIApVae
4cms    (307) ra-nnlVGLAkAi
1mpp    (339) fg-knrIGFApLasgyeNd
2asi    (343) fg-nnrIGFApLasayeNe
muc2    (314) fr-vpqVqIAk
RHINI 4(312) PQ-VPQVQIAPISN
```

higher than that of proteinases from *M. miehei* (25%) and *E. parasitica* (26%), both of them currently used in cheese manufacture thus indicating that the *M. bacilliformis* proteinase could be a useful substitute in cheese industry. With regard to protein structural architecture, CD determinations indicated an α-helix content of 2.9–4.0% which agrees with values ranging between 1% and 4% for all fungal aspartyl proteinases while pepsin and bovine chymosin reach values of 13% and 24%, respectively. Besides, spectra deconvolution in the 200–240 nm range indicated a 48% of β-sheet thus confirming the large content of said structure in the protein family.

Since the *M. bacilliformis* proteinase shows the greatest identity with and similarity to those from *Rhizopus* origin, we built a molecular model by utilizing the known structure of rhizopuspepsin from *R. chinensis* (Protein Data Bank code, 2apr.pdb) as the template. This structure was solved by X-ray crystallography with a resolution of 1.8 Å (Suguna et al., 1987) and has a high sequence identity (67%) with the target sequence. The structure (muc2) was selected out of 20 models obtained by using MODELLER (Sali and Blundell, 1993; Fiser et al., 2000). The model evaluation with Verify 3D (Luthy et al., 1992) yielded only positive scores and evaluation of its stereo-chemical quality with Procheck (Laskowski et al., 1993) showed that 95.1% of the residues are in the most favored regions, 4.9% of residues are in the additional favored regions, 0% in the generously allowed region and 0% residues in the disallowed region, in the Ramachandran plot. A very good superimposition was found (global rmsd of 0.28 Å in 320 superimposable residues) by using the SSAP Server (Pearl et al., 2005). Only two minor regions have a lower superimposition degree: a loop with an additional residue (Gly 54 in 2apr) and a loop with two additional residues in muc2 (Gly281-Phe282).

Catalytic aspartic residues 34 and 216 of *M. bacilliformis* aspartyl proteinase muc2 are equivalent to aspartic residues 35 and 218 of rhizopuspepsin 2apr and fully superimpose one another. Such residues are part of a hydrogen-bond arrangement very similar to those of rhizopuspepsin 2apr (Suguna et al., 1987), bovine chymosin 4cms (Newman et al., 1991), *M. pusillus* aspartyl proteinase 1mpp (Newman et al., 1993) and porcine pepsin 5pep (Cooper et al., 1990), as calculated with the CSU SOFTWARE (Sobolev et al., 1997). See yellow and green boxes in Fig. 4.

On the other hand, Fig. 5a shows a very good superimposition of bovine chymosin 4cms.pdb (Newman et al., 1991) and muc2 (*M. bacilliformis* proteinase model) global rmsd: 1.77 Å in 311 superimposable residues (96% overlapping). Instead, the

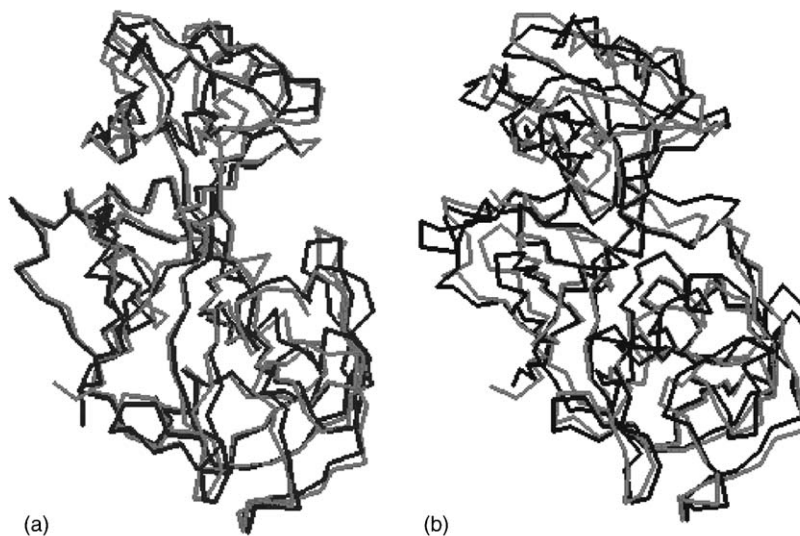

Fig. 5. (a) superimposition of bovine chymosin (4cms.pdb, black) and muc2 (gray); (b) superimposition of the *Mucor pusillus* proteinase (1mpp.pdb, black) and muc2 (gray).

superimposition of *M. pusillus* proteinase 1mpp.pdb (Newman et al., 1993) and muc2 (Fig. 5b) has a global rmsd of 2.24 Å for the 314 aligned residues (87% overlapping). As expected, there is a better superimposition of β-sheets and helices. There are 395 conserved contacts between the model and rhizopuspepsin; instead, there are 246 between the model and bovine chymosin and 223 between the model and the *M. pusillus* proteinase, as calculated according to Boutonnet et al. (1995). Thus, the *M. bacilliformis* proteinase is at a similar evolutionary distance from bovine chymosin and from the *M. pusillus* proteinase on a sequence level (33% and 32% identity, respectively). However, muc2 shows a fuller superimposition on bovine chymosin, on a tertiary structure level (Fig. 5a and b) though it should be pointed out that tertiary structure superimposition is much better on the N-terminal domain (1–178 amino acid residues of *M. bacilliformis* proteinase), something which does not go along with a higher sequence identity.

As bovine chymosin is considered the best milk coagulant in the cheese industry, the above results support the proposal that the mesophilic *M. bacilliformis* proteinase could be used as a better substitute than those from other sources utilized so far.

Acknowledgments

We are greatly indebted to Dr. Osvaldo Cascone for his supplying the *M. bacilliformis* crude extract and to Dr. José M. Delfino for performing CD spectra. We also wish to thank the expertise and advice of †D. Beatti, E. Dacci and S. Linskens as well as to Professor Rex Davis for English supervision. This work was supported in part by grants from the Universidad de Buenos Aires and Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina (CONICET).

References

Andrade, M.A., Chacón, P., Merelo, J.J., 1993. Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. Prot. Eng. 6, 383–390.

Areces, L.B., Bonino, M.B., Parry, M.A., Fraile, E.R., Fernández, H.M., Cascone, O., 1992. Purification and characterization of a milk clotting protease from *Mucor bacilliformis*. Appl. Biochem. Biotechnol. 37, 283–294.

Boutonnet, N.S., Rooman, M.J., Ochagavia, M.E., Richelle, J., Wodak, S.J., 1995. Optimal structure alignments by multiple linkage clustering: application to distantly related proteins. Prot. Eng. 8, 647–662.

Cooper, J.B., Khan, G., Taylor, G., Tickle, I.J., Blundell, T.L., 1990. X-ray analyses of aspartic proteases II. three-dimensional structure of the hexagonal crystal form of porcine Pepsin at 2.3 Angstroms resolution. J. Mol. Biol. 214, 199–222.

Costa, J., Ashford, D.A., Nimtz, M., Bento, I., Frazao, C., Estevez, C.L., Faro, C.J., Kervinen, J., Pires, E., Verissimo, P., Wlodawer, A., Carrondo, M.A., 1997. The glycosylation of the aspartic proteinases from barley (*Hordeum vulgare* L.) and cardoon (*Cynara cardunculus* L.). Eur. J. Biochem. 243, 695–700.

Crestfield, A.M., Moore, S., Stein, W.H., 1963. The preparation and enzymatic hydrolysis of reduced and *S*-carboxymethylated proteins. J. Biol. Chem. 238, 622–627.

Fiser, A., Do, R.K., Sali, A., 2000. Modeling of loops in protein structures. Prot. Sci. 9, 1753–1773.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., Bairoch, A., 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucl. Acids Res. 31, 3784–3788.

Hellman, U., 1997. Preparation, characterization, and microsequencing. In: Kamp, R.M., Choli-Papadopoulou, T., Wittmann-Liebold, B. (Eds.), Protein Structure Analysis. Springer–Verlag, Heidelberg, Germany, pp. 97–104.

Higgins, D.G., Thompsom, J.D., Gibson, T.J., 1996. Using CLUSTAL for multiple sequence alignments. Meth. Enzymol. 266, 383–401.

Huang, S.S., Huang, J.S., 1994. SDS polyacrylamide gel electrophoresis and electroblotting analysis. J. Prot. Chem. 13, 450–451.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T_4. Nature 227, 680–687.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291.

Lowry, O.H., Rosembrough, N.J., Farr, A.L., Randall, R.J., 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. 193, 265–275.

Luthy, R., Bowie, J.U., Eisenberg, D., 1992. Assessment of protein models with three-dimensional profiles. Nature 356, 83–85.

Matsudaira, P.T., 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. 262, 10035–10038.

Matsudaira, P.T., 1989. In: Matsudaira, P.T. (Ed.), A Practical Guide to Peptide and Protein Purification for Microsequencing. Academic Press, Inc., New York, pp. 53–58.

McGinnis, S., Madden, T.L., 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucl. Acids Res. 32, and references there in.

Merelo, J.J., Andrade, M.A., Prieto, A., Morán, F., 1994. Proteinotopic feature maps. Neurocomputing 6, 443–454.

Mizuguchi, K., Deane, C.M., Blundell, T.L., Overington, J.P., 1998a. HOMSTRAD: a database of protein structure alignments for homologous families. Prot. Sci. 7, 2469–2471.

Mizuguchi, K., Deane, C.M., Blundell, T.L., Johnson, M.S., Overington, J.P., 1998b. JOY: protein sequence-structure representation and analysis. Bioinformatics 14, 617–623.

Newman, M., Safro, M., Frazao, C., Khan, G., Zdanov, A., Tickle, I.J., Blundell, T.L., Andreeva, N., 1991. X-ray analyses of aspartic proteinases. IV. Structure and refinement at 2.2 A resolution of bovine chymosin. J. Mol. Biol. 221, 1295–1309.

Newman, M., Watson, F., Roychowdhury, P., Jones, H., Badasso, M., Cleasby, A., Wood, S.P., Tickle, I.J., Blundell, T.L., 1993. X-ray analyses of aspartic proteinases V. Structure and refinement at 2.0 Å resolution of the aspartic proteinase from *Mucor pusillus*. J. Mol. Biol. 230, 260–283.

Pearl, F., Todd, A., Sillitoe, I., Dibley, M., Redfern, O., Lewis, T., Bennett, C., Marsden, R., Grant, A., Lee, D., Akpor, A., Maibaum, M., Harrison, A., Dallman, T., Reeves, G., Diboun, I., Addou, S., Lise, S., Johnston, C., Sillero, A., Thornton, J., Orengo, C., 2005. The CATH domain structure database and related resources gene3D and DHS provide comprehensive domain family information for genome analysis. Nucl. Acids Res. 33, Database Issue D247-D251 and there in.

Rickert, W.S., McBride-Warren, P.A., 1974. Structural and functional determinants of *Mucor miehei* protease. Biochim. Biophys. Acta 336, 437–444.

Sali, A., Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779–815.

Sali, A., Blundell, T.L., 1990. Definition of general topological equivalence in protein structures. A procedure involving comparison of properties and relationships through simulated annealing and dynamic programming. J. Mol. Biol. 212, 403–428.

Schägger, H., von Jagow, G., 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. 166, 368–379.

Sobolev, V., Sorokine, A., Prilusky, J., Abola, E.E., Edelman, M., 1997. Automated analysis of interatomic contacts in proteins. Bioinformatics 15, 327–332.

Suguna, K., Bott, R.R., Padlan, E.A., Subramanian, E., Sheriff, S., Cohen, G.H., Davies, D.R., 1987. Structure and refinement at 1.8 Angstroms resolution of the aspartic proteinase from *Rhizopus chinensis*. J. Mol. Biol. 196, 877–900.

Venera, G.D., Machalinski, C., Zumárraga, H., Biscoglio de Jiménez Bonino, M.J., 1997. Further characterization and kinetic parameter determination of a milk-clotting protease from *Mucor bacilliformis*. Appl. Biochem. Biotechnol. 68, 207–216.

Zhong, L., Curtis Johnson Jr., V., 1992. Environment affects amino acid preference for secondary structure. Proc. Natl. Acad. Sci. (U.S.A.) 89, 4462–4465.