Tricorn protease (TRI) interacting factor 1 from *Thermoplasma acidophilum* is a proline iminopeptidase

Tomohiro Tamura, Noriko Tamura, Friedrich Lottspeich, Wolfgang Baumeister*

Max-Planck-Institute for Biochemistry, D-82152 Martinsried, Germany

Received 7 October 1996

Abstract Tricorn protease (TRI), a high molecular mass complex from the archaeon *T. acidophilum*, forms the core of a modular proteolytic system; upon interacting with low molecular mass factors intrinsic activities are enhanced and novel activities are generated. Here we characterize the first factor, F1, which turns out to be homologous with several bacterial proline iminopeptidases (PIPs). Surprisingly, it cleaves not only typical PIP substrates such as H-Pro-AMC, but a wide spectrum of amino acid substrates and several peptide substrates without a proline at the N-terminus. The *pip* gene encodes a 293 amino acid residue protein with a molecular mass of 33 487 Da. By means of site-directed mutagenesis we identified Ser^{105} and His^{271} as the active site nucleophile and proton donor, respectively. Experiments with inactive mutant PIPs indicate that the activities elicited by interacting with TRI are contributed by PIP.

Key words: Proline iminopeptidase; Tricorn protease; *Thermoplasma acidophilum*

1. Introduction

We have recently found a novel proteolytic system in the archaeon *Thermoplasma acidophilum*, unrelated to the 20S proteasome, which consists of a high molecular mass protease named Tricorn protease (TRI) and at least two different low molecular mass components: F1 and F2 [1]. Upon interacting with each other, multicatalytic peptidase activities are generated. TRI in conjunction with F1 yielded Suc-Leu-Leu-Val-Tyr-AMC (Suc-LLVY-AMC) cleaving activity and enhanced its intrinsic H-Ala-Ala-Phe-AMC (H-AAF-AMC) cleaving activity [1]. Since purified TRI has no Suc-LLVY-AMC cleaving activity, F1 must be involved in generating this novel activity. We have purified and cloned the F1 gene which turned out to be homologous to proline iminopeptidases.

Proline iminopeptidase (EC 3.4.11.5) was first identified in *Escherichia coli* based on the ability of proline auxotrophs to utilize poly-L-proline as a source of proline [2,3]. Thereafter, the enzyme has been found in several eubacterial species [4–9]. PIP hydrolyses peptide substrates containing proline at the amino terminus and releases free proline. PIP was also shown to have low levels of cleaving activity against several dipeptides not containing proline. However, no tri- or tetrapeptide endopeptidase activities of PIP have yet been reported.

In this communication, we characterize *Thermoplasma acidophilum* PIP. We have identified its active site residues by site-directed mutagenesis experiments. Furthermore, we sequenced a eukaryotic PIP homologue from *Arabidopsis thaliana* and partially characterized its enzymatic activities.

2. Materials and methods

2.1. Materials

Chemicals and chromatography resins for protein purification were purchased from Sigma, Pharmacia, Merck and Bio-Rad. Enzymes for DNA restriction and modification were obtained from New England Biolabs and Stratagene. Ni-NTA resin for histidine-tagged recombinant PIP purification was obtained from Diagen. Oligonucleotides for PCR reactions were synthesized on an Applied Biosystems 380A DNA synthesizer. The synthetic fluorescent amino acids and peptides were obtained from Bachem.

2.2. Purification of TRI interacting F1 protein

The *T. acidophilum* crude cell extract was prepared as described previously [1]. For purification of TRI interacting factor 1, *T. acidophilum* crude cell extracts were loaded onto a Sepharose 6B column (3×90 cm) pre-equilibrated with buffer A (25 mM Tris-HCl, 1 mM DTT, pH 7.5). The active fractions were pooled and loaded onto a Q-Sepharose column (2×13 cm) pre-equilibrated with buffer A. Bound proteins were eluted using a 0–400 mM NaCl linear gradient in buffer A. The active fractions were pooled and dialyzed against 10 mM potassium phosphate buffer, 1 mM DTT, pH 7.0. The dialyzed sample was applied to a hydroxylapatite column (2×6 cm), pre-equilibrated with 10 mM potassium phosphate buffer. Bound proteins were eluted using a 10–300 mM phosphate linear gradient. The active fractions were pooled and dialyzed against buffer A. The dialyzed sample was loaded onto a Mono-Q column (5 mm×5 cm) pre-equilibrated with buffer A. Bound proteins were eluted using a 0–500 mM NaCl linear gradient in buffer A. The active fractions were pooled, dialyzed against buffer A and stored at 4°C.

2.3. Cloning of the F1 encoding gene from T. acidophilum

Partial protein sequences were obtained by amino acid sequencing as described previously [10]. *T. acidophilum* chromosomal DNA was purified by phenol/chloroform extraction. PCR was carried out by using 4–8-fold degenerated amplification primers which were designed from the peptide sequences, VNGIYIYYK (residues 12–20 aa) and DWDITDK (residues 222–228 aa). The primer pair was GT[G/T] AAT GG[G/T] ATT TAT ATT TAT TA[T/C] AA as a sense primer and TTG TC[G/T] GTG ATG TCC CAG TC[T/C] TT as an antisense primer, respectively. The amplified 651-bp DNA fragment, which covered 74% of open reading frame (ORF) encoding factor 1 protein, was sequenced. To obtain the missing N- and C-terminal ends, two homologous pairs of inverse PCR primers were constructed at 5′- and 3′-edge regions of the analyzed 651-bp DNA fragment. Primers for the 5′-edge region were TTG CTT TCT CCT CAG GGG CT and GCT AAT GAC CAT GCA CGG GG, while primers for 3′-edge region were GGG CCG TTC ATT ATC CTG TA and TGA GTT CAC CAT AAC CGG CA. *T. acidophilum* chromosomal

*Corresponding author. Fax: (49) (89) 8578 2641.
E-mail: tamura@vms.biochem.mpg.de

Abbreviations: SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; DTT, dithiothreitol; AMC, 7-amino-4-methylcoumarin; Suc, succinyl; Z, benzoxycarbonyl; Boc, *t*-butoxycarbonyl; Bz, benzony

The nucleotide sequences of *T. acidophilum* PIP and updated *A. thaliana* PIP will appear in the GenBank/EMBL Data Bank under Accession Numbers U72710 (*T. acidophilum* PIP) and U72711 (*A. thaliana*).

PI S0014-5793(96)01163-5

DNA was digested with *Ava*I or *Sau*96I, circularized by T4 ligase and subjected to inverse PCR [11]. A pair of 5′-edge region inverse primers gave rise to an approximately 1-kb DNA fragment, using *Ava*I-digested and circularized chromosomal DNA as a template. The 3′-edge region encoding primers amplified an approximately 1.1-kb DNA fragment using *Sau*96I digested and circularized chromosomal DNA as a template. These amplified fragments were sequenced.

2.4. Construction and purification of mutant recombinant PIPs

The *T. acidophilum pip* gene was amplified by PCR using an oligonucleotide pair with a $(His)_6$ tag at the N-terminal end and the second at the C-terminal end, both with a flanking *Nco*I site. The amplified DNA fragments were cloned into the *Nco*I site of pRSET6d [12] yielding pRSET-PIP. The mutations were introduced by inverse PCR mutagenesis using pRSET-PIP as a template. The PCR reaction and ligations were performed as described [13]. Incorporation of nucleotide exchanges was confirmed by DNA sequencing. The $(His)_6$-tagged PIP was expressed in *E. coli* BL21(DE3) cell [14], purified on a Ni-NTA resin and dialyzed against 50 mM Tris-HCl (pH 7.5), 20% glycerol.

2.5. Determination of peptidase activity

Throughout protein purification, the Suc-LLVY-AMC cleaving activity was monitored by adding pooled TRI-containing fractions from a 10–40% glycerol density gradient as described in [1]. For PIP characterization, proteins were incubated in 50 mM Tris-HCl (pH 8.0) with 100 nmol of substrate and bovine serum albumin for stabilization of PIP; these conditions optimized PIP peptidase activity to levels 10-fold higher than described previously [1]. A 100 μl reaction was terminated with 100 μl of 10% SDS and 1 ml of 0.1 M Tris-HCl (pH 9.0). Cleavage of the peptide substrate was monitored by measuring the release of AMC fluorometrically.

2.6. Sequence and construction of recombinant Arabidopsis thaliana PIP homologue

The GenBank database was used to search for eukaryotic homologues of PIP. This search produced matches with several *A. thaliana* expressed sequence tagged clones (R64970, R65321, R90623, N96120, N96979, T43193). We sequenced the cDNA clone (GenBank No. R64970), kindly provided by Dr. T. Newman, Michigan State University, via Arabidopsis Biological Resource Center (Ohio State University). For recombinant PIP expression, the ORF was amplified by PCR using an oligonucleotide pair with a $(His)_6$ tag at the N-terminal end with a flanking *Nde*I site and the second at the 3′ non-coding region with a flanking *Nco*I site. The amplified DNA fragment was digested with *Nde*I and *Nco*I, and cloned into the same restriction enzyme sites of pRSET6a [12] yielding pRSET-ATPIP. The recombinant protein was expressed and purified as described above.

3. Results

3.1. Purification of TRI interacting factor 1

The activity of TRI interacting factor 1 (F1) throughout the purification procedure was monitored as an enhancement of Suc-LLVY-AMC cleaving activity upon mixing it with TRI as described before [1]. In four chromatographic steps, we obtained 30 μg of homogeneous F1 protein from 9 g of *T. acidophilum* cell pellet. In the final purification step with a Mono-Q column, the protein eluted at 90 mM NaCl concentration. The purified protein consisted of a single 34-kDa band in SDS-PAGE (Fig. 1). It migrated in the same molecular mass range on Superose 12 FPLC chromatography (data not shown), indicating that it is a monomeric protein.

3.2. Primary structure of the F1 protein

The purified protein was subjected to amino acid microsequencing, and we obtained five partial sequences including the N-terminal end. In order to clone the gene encoding F1, we constructed oligonucleotides to carry out linear and inverse PCRs using *T. acidophilum* chromosomal DNA as a template.

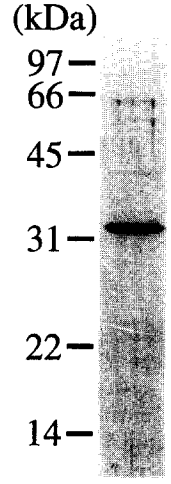

Fig. 1. Electrophoresis of purified F1 protein. Purified F1 protein (1 μg) resolved by 12.5% of SDS-PAGE. Protein was stained with Coomassie brilliant blue.

Three overlapping DNA fragments, which covered the full F1 gene, were amplified and sequenced. The DNA sequence revealed a 882-bp ORF encoding 293 amino acids, which contained all the peptide sequences obtained by microsequencing, with a calculated molecular mass of 33487 Da and an estimated pI of 5.4 (Fig. 2). The primary sequence derived from the cloned gene showed significant similarity to eubacterial proline iminopeptidases (PIP) or prolinases (22.5–34.7% identity). The highest similarity was found with PIP from *Lactobacillus delbrueckii* subsp. *lactis* DMS 7290 (Fig. 2). The N-terminal half is obviously more conserved than the C-terminal half. The PILEUP program in the GCG package showed that PIPs fall into two groups: the first group includes Ta, Bc, Ld and Lh proteins; the second group consists of Xc, Mg and At proteins. The functional significance of these two groups is as yet unknown. Surprisingly, our sequence comparisons revealed that PIPs are members of the α/β-hydrolase superfamily [15]. The highest similarity was found with bromoperoxidase A2 (BPO-A2) from *Streptomyces aureofaciens* (25% identity) (Fig. 2), suggesting that PIPs may have a topology very similar to the α/β hydrolases.

3.3. Characterization of T. acidophilum PIP

We constructed a recombinant PIP (rPIP) expression vector, which carried a $(His)_6$ tag fused to PIP at the N-terminal end, expressed and purified the protein by Ni-NTA affinity column. Purified rPIP was tested for peptidase activity using synthetic fluorogenic substrates (Table 1). As expected, rPIP had strong H-Pro-AMC cleaving activity, but no Suc-Gly-Pro-AMC or Suc-Gly-Pro-Leu-Gly-Pro-AMC hydrolyzing activity. This indicates that rPIP has indeed no prolyl-endopeptidase activity. Surprisingly, however, rPIP showed broad substrate specificity against neutral and hydrophobic amino residues. Particularly H-Phe-AMC and H-Ala-AMC hydrolyzing activities were as high as or even higher than the H-Pro-AMC cleaving activity. Furthermore, rPIP showed cleaving activity against H-AAF-AMC, but not against a substrate that was negatively charged at the N-terminus (Suc-AAF-AMC). Z-Gly-Gly-Leu-AMC is cleaved, though at much lower rates than H-Pro-AMC. PIP has no peptidase activity

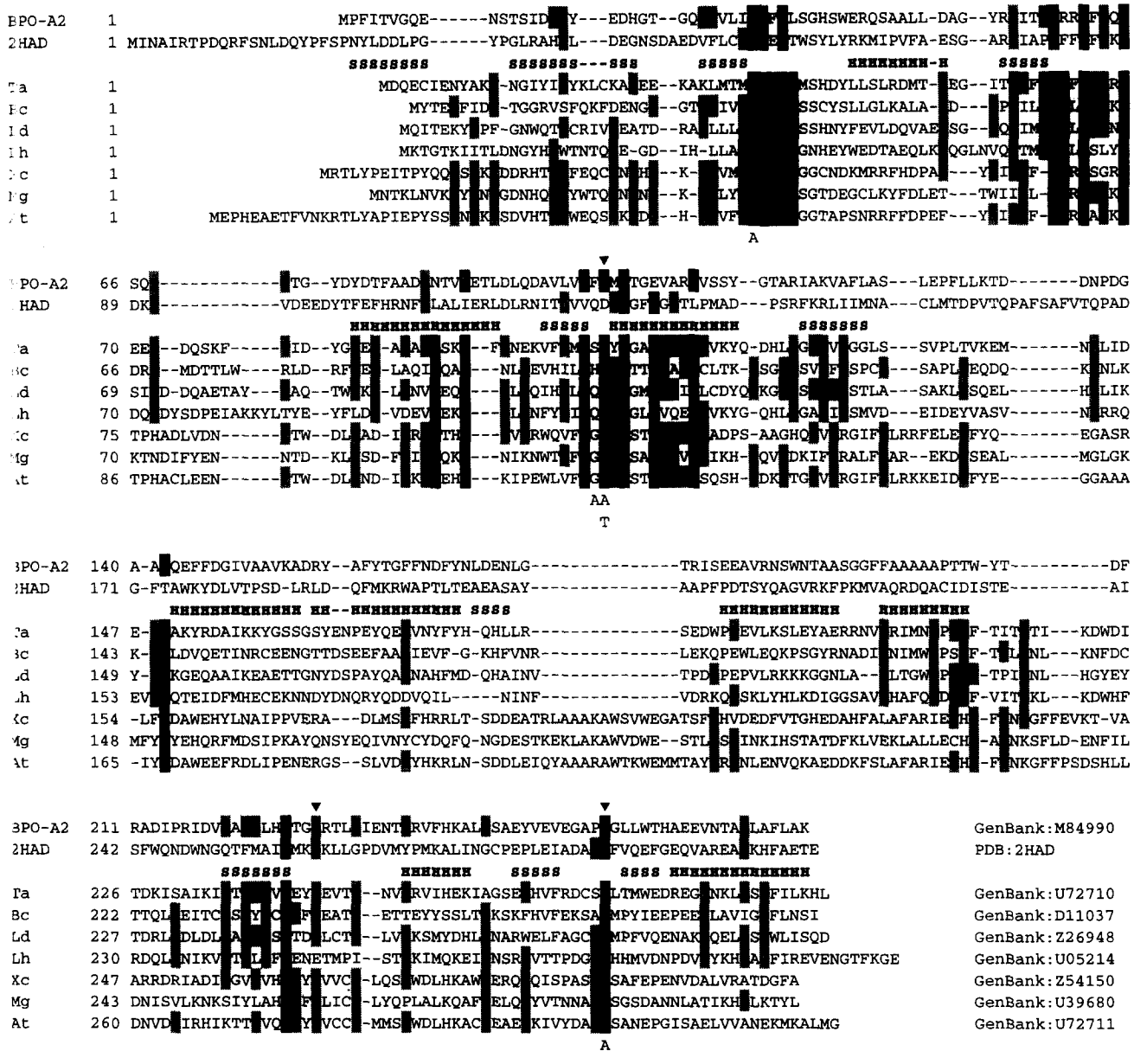

Fig. 2. Alignment of proline iminopeptidases with two α/β hydrolases of known structure. For a structural classification of α/β hydrolases see the SCOP site at http://scop.mrc-lmb.cam.ac.uk/scop/. The seven proline iminopeptidases are from Ta, *Thermoplasma acidophilum*; Bc, *Bacillus coagulans*; Ld, *Lactobacillus delbrueckii* subsp. *lactis* DSM 7290; Lh, *Lactobacillus helveticus* CNRZ32; Xc, *Xanthomonas campestris*; Mg, *Mycoplasma genitalium*; At, *Arabidopsis thaliana*. The two α/β hydrolases of known structure that are most similar to PIPs are bromoperoxidase A1 from *Streptomyces aureofacies* (BPO-A2) and *Xantobacter autotrophicus* haloalkane dehalogenase (2HAD). Their consensus secondary structure is shown beneath their sequence (H, α helix; S, β strand) and the residues of the catalytic triad (Ser/Asp-His-Asp/Glu) are marked by ▼. The shading of residues follows the conservation pattern of PIPs: invariant positions are shown in inverted type and positions with at least 50% identity are shown in shaded type. Mutated residues are indicated underneath the alignment. Numbers to the left refer to the amino acid position of the respective protein sequence. Accession numbers are shown at the end of sequences.

against Suc-LLVY-AMC. This activity was generated only upon mixing F1 with TRI.

An alignment between the various PIPs and the α/β hydrolases revealed that the identified active site residues of the hydrolases are completely conserved in the PIPs. This suggested that S105, D244 and H271 form the catalytic triad in *Thermoplasma* PIP (Fig. 2). In order to confirm this and to verify the PIP tripeptide hydrolyzing activity, mutants were generated; S105, H35 and H271 were mutated to alanine (Table 2). Analysis of the purified mutant rPIPs demonstrated that the S105 mutant rPIP (S105A) had completely lost both the proline iminopeptidase activity and the peptide hydrolyzing activity. The S104A mutant (not a highly conserved residue) rPIP showed drastically reduced levels of all of peptidase activities, but not a complete loss. All peptidase activities were abolished in the mutant rPIP H271A, while the H35A mutant rPIP retained activity. The mutagenesis in conjunction with the sequence comparisons thus identified S105 and H271 as the active site nucleophile and proton donor, respectively.

3.4. Characterization of the interaction between PIP and TRI

As described previously [1], PIP and TRI can, upon interacting with each other, enhance intrinsic peptidase activities or generate new ones. Both PIP and TRI show H-AAF-AMC

Table 1
Hydrolysis of synthetic substrates by recombinant PIP from *Termoplasma acidophilum*

Substrate	μmol/h/mg
H-Pro-AMC	2225.6
H-Ala-AMC	3338.4
H-Phe-AMC	2223.4
H-Leu-AMC	1546.8
H-Val-AMC	298.2
H-Try-AMC	282.7
H-Gly-AMC	353.9
H-Arg-AMC	0
H-Glu-AMC	0
Suc-Gly-Pro-AMC	0
Suc-Gly-Pro-Leu-Gly-Pro AMC	0
H-Ala-Ala-Phe-AMC	52.8
Suc-Ala-Ala-Phe-AMC	0.2
Z-Gly-Gly-Leu-AMC	28.4
Suc-Leu-Tyr-AMC	0
Suc-Leu-Leu-Val-Tyr-AMC	0
Boc-Leu-Arg-Arg-AMC	0
Z-Ala-Arg-Arg-AMC	0
Bz-Val-Gly-Arg-AMC	0
Z-Gly-Gly-Arg-AMC	0

Activity was assayed by incubating recombinant protein with 100 nmol of fluorogenic substrate and 3 μg of BSA for 15–30 min at 60°C. Fluorescence of AMC was measured.

cleaving activity, but no Suc-LLVY-AMC cleaving activity. To find out whether the PIP's enzymatic activities contribute to the cleavage, mutant rPIPs were tested in reconstitution experiments with TRI (Table 3). Wild type rPIP and TRI generated Suc-LLVY-AMC cleaving activity, however, S105A and H271A mutant rPIPs failed to generate this activity. In conclusion, these results demonstrate that the peptidase activities generated by interaction with TRI are contributed by PIP.

3.5. *Characterization of the eukaryotic PIP homologue from A. thaliana*

We searched GenBank for eukaryotic PIP homologues and found several *A. thaliana* expressed sequence tagged clones (see Section 2). The cDNA was sequenced and revealed a 990-bp ORF encoding 329 amino acids with a calculated molecular mass of 37350 Da and a predicted pI of 5.5 (Fig. 2). The primary sequence of *A. thaliana* PIP showed significant similarity to PIPs; the highest similarity was with *Xanthomonas campestris* PIP (57.1% identity). Furthermore, the active

Table 2
Peptidase activities of mutated recombinant PIPs as percent of wild type

	H-P-AMC	H-AAF-AMC	Z-GGL-AMC
Wild type	100	100	100
S104A	8	11	8
S105A	0	0	0
SS104/5AA	0	0	0
S105T	0	0	0
H35A	34	45	21
H271A	0	0	0

Activity was assayed by incubation of wild type (5–100 ng) or a mutant (10–500 ng) rPIP with 100 nmol of fluorogenic substrate and 3 μg of BSA for 15 min at 60°C. Fluorescence of AMC was measured.

Table 3
Effect of mutant PIPs on TRI-dependent peptidase activity

	Suc-LLVY-AMC	H-AAF-AMC
Wild type	6.8	352.0
S104A	0.6	29.1
S105A	0	0
H35A	3.0	156.8
H271A	0	0

The activity values are given as TRI-induced activities in μmol/h/mg. Activity was assayed by incubation of 20–300 ng of rPIPs with 100 nmol of fluorogenic substrate and 3 μg of BSA for 15 min at 60°C. Fluorescence of AMC was measured.

site residues identified in *T. acidophilum* PIP are conserved in the translated *A. thaliana* cDNA clone. We constructed an expression vector encoding the *A. thaliana* PIP ORF with a $(His)_6$ tag at the N-terminal end. Recombinant proteins were expressed in *E. coli* and purified by Ni-NTA affinity column. The purified recombinant protein was tested for peptidase activity using the same synthetic substrates as used for *T. acidophilum* PIP characterization. We found that although this recombinant protein has H-Pro-AMC (30.1 μmol/h/mg) and H-Ala-AMC (12.8 μmol/h/mg) cleaving activities, the specific activities were much lower than those of *T. acidophilum* PIP (Table 1). Furthermore, *A. thaliana* PIP has neither prolyl-endopeptidase nor tri-peptide cleaving activity and none was generated by mixing it with *T. acidophilum* TRI.

4. Discussion

We have purified the TRI interacting factor 1 from *Thermoplasma acidophilum* as an enzyme generating Suc-LLVY-AMC hydrolyzing activity. Cloned F1 is homologous to proline iminopeptidases. Although proline iminopeptidases are regarded as highly specific against substrates with proline at the N-terminus [9], *T. acidophilum* PIP is able to hydrolyze a wide spectrum of substrates including H-Ala-AMC, H-Phe-AMC, H-Leu-AMC, H-Val-AMC, H-Try-AMC and H-Gly-AMC. Furthermore *T. acidophilum* rPIP hydrolyzed several chymotrypsin-like substrates (Table 1), but did not cleave polypeptides such as insulin B-chain and casein (data not shown). This suggests that the degradation products derived from TRI might be channeled to F1 for further hydrolysis.

Our sequence comparisons revealed that PIPs are members of the α/β hydrolase superfamily. The crystal structures of several α/β hydrolases have been determined and the catalytic triad has been identified [15]. A comparison between several PIPs and two α/β hydrolases, BPO-A2 and 2HAD, suggests that not only the topology but also the active site residues have been conserved [16,17] (Fig. 2). Our mutagenesis experiments with F1 confirmed that S105 and H271 act as the nucleophile and proton donor, respectively. In addition to these residues, D244 may be implicated in active site formation.

Upon mixing PIP with TRI, we found that at least six different activities were generated; Z-GGL-AMC, Suc-AAA-AMC, Suc-AAV-AMC, Suc-GPLGP-AMC, Suc-LLVY-AMC and H-AAF-AMC (also Suc-AAF-AMC) (data not shown). We have no evidence yet whether the released AMC from these peptides is due to an aminopeptidase or an endopeptidase activity. Experiments performed with inactive mutant rPIP ruled out that these peptidase activities were due to any contaminant peptidases.

Although preliminary evidence suggests that TRI and PIP

interact physically, the mode of this activation is unknown, and requires further characterization. To this end, all the components of this modular proteolytic system have to be isolated and characterized. We have already identified another factor, F2, which also generates several peptidase activities upon interaction with TRI distinct from these described here. F1 and F2 may be able to produce any length of peptide upon interacting with TRI.

We have preliminary evidence that low molecular mass proteins in rabbit reticulocyte lysate, fractionated on a 10–40% glycerol density gradient, generate a weak Suc-LLVY-AMC cleaving activity upon mixing with *T. acidophilum* TRI. This suggests that molecules functionally similar to *T. acidophilum* PIP exist in eukaryotic cells.

Acknowledgements: We are grateful to A. Lupas and M.A. Kania for critically reading the manuscript and for valuable suggestions, to M. Bocu for DNA sequencing. T. Tamura acknowledges a Postdoctoral Fellowship for Research Abroad from the Japan Society for the Promotion of Science. This work was also supported by a grant from the Human Frontier Scientific Program to W. Baumeister.

References

[1] Tamura, T., Tamura, N., Cejka, Z., Hegerl, R., Lottspeich, F. and Baumeister, W. (1996) Science (in press).
[2] Sarid, S., Berger, A. and Katchalski, E. (1959) J. Biol. Chem. 234, 1740–1746.
[3] Sarid, S., Berger, A. and Katchalski, E. (1962) J. Biol. Chem. 237, 2207–2212.
[4] Kitazono, A., Yoshimoto, T. and Tsuru, D. (1992) J. Bacteriol. 174, 7919–7925.
[5] Dudley, E.G. and Steele, J.L. (1994) FEMS Microbiol. Lett. 119, 41–46.
[6] Albertson, N.H. and Koomey, M. (1993) Mol. Microbiol. 9, 1203–1211.
[7] Klein, J.R., Schmidt, U. and Plapp, R. (1994) Microbiology 140, 1133–1139.
[8] Atlan, D., Gilbert, C., Blanc, B. and Portalier, R. (1994) Microbiology 140, 527–535.
[9] Gilbert, C., Atlan, D., Blanc, B. and Portalier, R. (1994) Microbiology 140, 537–542.
[10] Tamura, T., Nagy, I., Lupas, A., Lottspeich, F., Cejka, Z., Schoofs, G., Tanaka, K., De Mot, R. and Baumeister, W. (1995) Curr. Biol. 5, 766–774.
[11] Ochman, H., Medhora, M.M., Garza, D. and Hartl, D.L. (1990) in: PCR Protocols: A Guide to Methods and Applications (Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J., Eds.), pp. 219–227, Academic Press, San Diego, CA.
[12] Schoepfer, R. (1993) Gene 124, 83–85.
[13] Seemüller, E., Lupas, A., Zühl, F., Zwickl, P. and Baumeister, W. (1995) FEBS Lett. 359, 173–178.
[14] Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Methods Enzymol. 185, 60–89.
[15] Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., Sussman, J.L., Verschueren, K.H.G. and Goldman, A. (1992) Protein Engng. 5, 197–211.
[16] Hecht, H.J., Sobek, H., Haag, T., Pfeifer, O. and van Pée, K.-H. (1994) Nature Struct. Biol. 1, 532–537.
[17] Verschueren, K.H.G., Seljée, F., Rozeboom, H.J., Kalk, K.H. and Dijkstra, B.W. (1993) Nature 363, 693–698.