

REVIEW

Genomic Analysis of *Koji* Mold *Aspergillus oryzae* and Investigation of Novel Peptidases by Post-genomic Approach

Ken-Ichi KUSUMOTO*

National Food Research Institute, National Agriculture and Food Research Organization (Tsukuba, Ibaraki 305–8642, Japan)

Abstract

Aspergillus oryzae is a filamentous fungus used as a starter microorganism in the process of manufacturing various kinds of traditional fermented food. It is also used as a producer of enzymes for both food processing and pharmaceuticals, such as proteases and taka-amylase. In order to clarify the function of this microorganism in fermentation industries and utilize the new function of *A. oryzae*, it was subject to genomic analysis. During the analysis, the telomere of *A. oryzae* exhibited a novel repeated sequence distinct from other fungi. Genomic analysis also revealed that the genome of *A. oryzae* consists of eight chromosomes containing 12,074 genes and is larger in size than *Aspergillus nidulans* and *Aspergillus fumigatus*. *A. oryzae* contains sequences specific for *A. oryzae* for genes involved in primary and secondary metabolism, when compared with *A. nidulans* and *A. fumigatus*. Among them, the function of one of the peptidase-like gene products was investigated with a post-genomic approach, and determined as an aminopeptidase specific for acidic amino acids (aspartate and glutamate) which was first identified in filamentous fungi. It also demonstrated that the addition of cobalt ion to the growth medium greatly improved the specific activity. Using the same approach, some of the novel aminopeptidases and carboxypeptidases of *A. oryzae* were identified and characterized.

Discipline: Biotechnology

Additional key words: enzyme production

Introduction

Aspergillus oryzae is a filamentous fungus used as *koji* mold (a starter microorganism mainly cultured on solid media such as steamed rice or soy beans etc.) in the process of manufacturing traditional fermented food, like *miso* (salted soy bean paste), *shoyu* (soy sauce) or *sake* (rice wine). It is also used as a producer of enzymes for both food processing and pharmaceuticals, such as proteases and taka-amylase. *Aspergillus flavus*, known to be a producer of aflatoxin, and *A. oryzae*, both belong to the *Flavi* section of the *Circumdati* subgenus of *Aspergillus*. Nevertheless, *A. oryzae* does not produce aflatoxin, and has a long history in the use of fermentation industry, hence its safety is affirmed as GRAS (Generally Recog-

nized As Safe) by the US Food and Drug Administration. During the fermentation process, those enzymes hydrolyze high molecular weight materials, such as starch from media to release sugar, and protein to release peptides and amino acids. As fermentation proceeds, these products are enriched to function as taste and flavor enhancers. For example, the peptides and amino acids in soy bean-derived fermented products are released from proteins by various *A. oryzae* end-type proteases. Subsequently, various aminopeptidases (exo-type protease) and dipeptidyl peptidase degrade the released peptides beginning at their amino-termini. At the same time, carboxypeptidases degrade the peptides from their carboxyl-termini, whereupon short peptides and amino acids are released.

The enzymatic research identified nearly twenty

This work was partly supported by the Program for Promotion of Basic Research Activities for Innovative Biosciences, Japan (PROBRAIN).

*Corresponding author: e-mail kusumoto@affrc.go.jp

Received 18 December 2010; accepted 14 April 2011.

proteases produced by *A. oryzae* during the manufacturing fermentation process food. During early research, the enzymes produced in bulk were purified and identified. Because they are mostly secreted into the media, the purification or identification is relatively straightforward. However, enzymes present in traces may also be considered involved in the material degradation process, leading the fermentation process to the matured product. Moreover, unknown enzymes may be hidden behind large amount of proteins, or unexpressed. This idea was the catalyst for the genome analysis of *A. oryzae* by a research consortium made up of several universities, companies, and national institutes.

Analysis of telomere sequence and the whole genomic sequence of *A. oryzae*

Accompanying the genome analysis of *A. oryzae*¹⁰, its telomere sequence had to be determined. Telomeres are the terminal sequence of the chromosome in eukary-

otes, and their sequences are mainly known as organism-specific repeat sequences, synthesized by a telomerase⁶. Since the genome sequence analysis was performed mainly by whole shot-gun sequencing, the sequence data assembling process was sometimes problematic in that certain repeat sequences continued at the presumed terminal region of each chromosome. Such presumed terminal region should be confirmed as the actual terminus of the chromosome for the purpose of complete genome analysis. Therefore, a DNA library for concentrating the terminal region of the chromosome was prepared (Fig. 1). The sequence analysis of the cloned DNA fragments revealed that the sequence of some clones contained repeats, the repeat unit of which was TTAGGGTCAACA⁷. The first six nucleotides in the repeat unit, TTAGGG, were the same as the repeat unit conserved in human, mice, slime molds, and several species of filamentous fungi^{2,5,9} (Fig. 2). Subsequently, those specific repeat sequences in *A. oryzae* were confirmed to telomere sequences by BAL31 exonuclease analysis, as the terminal

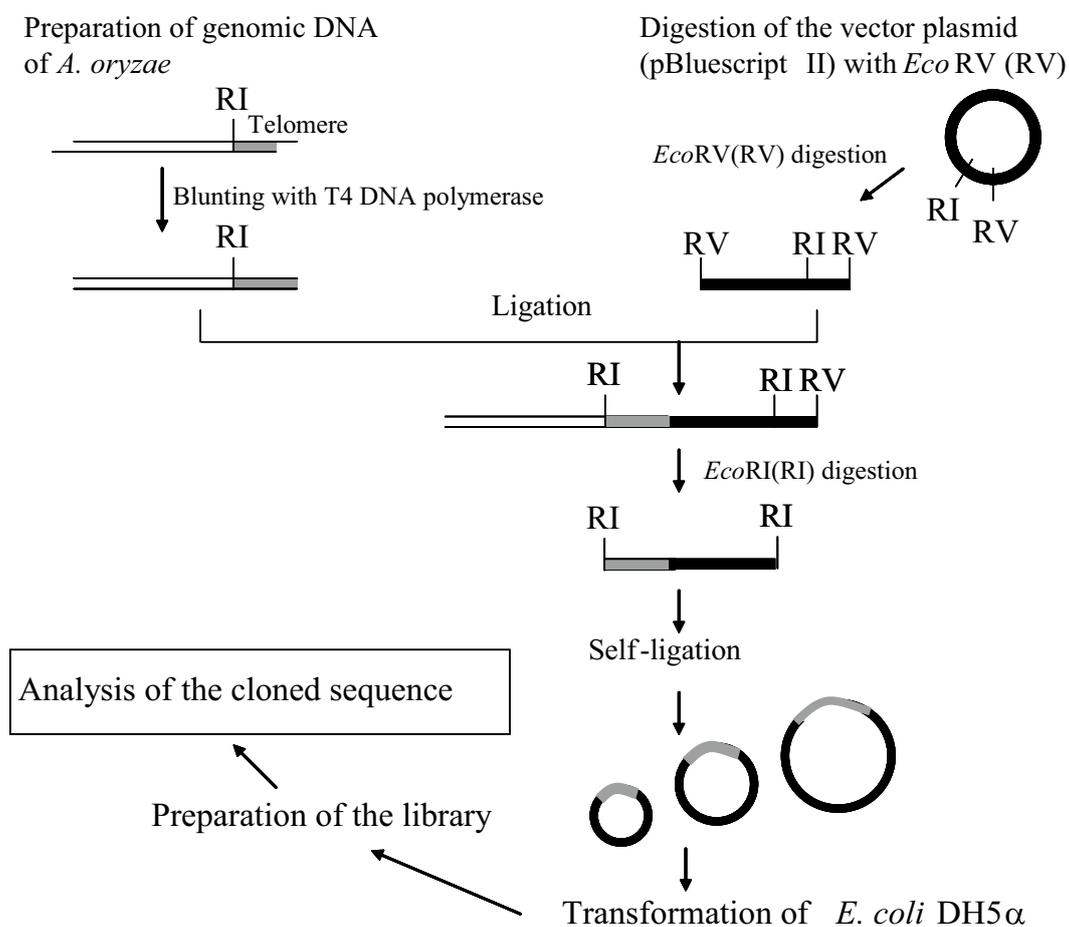
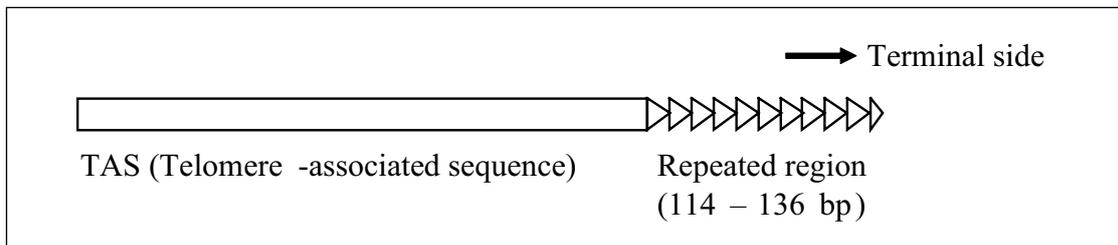


Fig. 1. Preparation of the DNA library of chromosomal termini from *A. oryzae*



Species	Repeat unit	No. of repeats
<i>Aspergillus oryzae</i> *	TTAGGGTCAACA	9 – 11 (114 – 136 bp)
<i>Aspergillus nidulans</i>	TTAGGG	4 – 22 (91 – 134 bp)
<i>Podospora anserina</i>	TTAGGG	(200 bp)
<i>Magnaporthe grisea</i>	TTAGGG	22 (132 bp)
<i>Fusarium oxysporum</i>	TTAGGG	18 (124 bp)
<i>Pestalotiopsis microspora</i>	TTAGGG	20 (120 bp)
<i>Saccharomyces cerevisiae</i>	TG ₁₋₃	(0.2 – 0.4 kb)
Mammals	TTAGGG	(10 – 150 kb)
<i>Tetrahymenathermophila</i>	TTGGGG	(300 bp)
<i>Kluyveromyces lactis</i>	TCG ₂ AT ₃ GAT ₂ AG ₂ TATGTG ₂ TGT	

* Kusumoto *et al.* Appl. Microbiol. Biotechnol. 2003

Fig. 2. Telomeric sequence of *A. oryzae* and other organisms

region of the chromosome was known to be sensitive to exonuclease digestion². Therefore, the telomere sequence of *A. oryzae* appeared to be a species-specific sequence, distinct from other fungal species.

Interestingly, the result of the terminal region sequence analysis, including telomeric repeat, revealed the length of the telomeric region in *A. oryzae* as 114-136 bp, which corresponds to 9-11 repeats of the twelve nucleotides. Instead, that of the telomeric regions of *A. nidulans* is 90-120 bp. The fact that fungal telomere of different species have similar lengths suggests the controlling mechanism is possibly conserved among *Aspergillus* species⁷.

Genome information analysis of *A. oryzae*

Sequence determination using the whole-genome shotgun method along with confirmation of the telomere and its neighbor sequences, gene finding, and gene annotation in *A. oryzae* finally achieved the completion of its genomic sequence analysis¹⁰. The genome consists of eight chromosomes, predicted to contain 12,074 genes encoding proteins consisting of over 100 amino acid resi-

dues. The genome size of *A. oryzae* is greater than *Aspergillus nidulans* and *Aspergillus fumigatus*, by about 29 and 34% respectively^{4,10,19}.

A. oryzae contains sequences specific for *A. oryzae* for genes involved in primary and secondary metabolism, compared with *A. nidulans* and *A. fumigatus*¹⁰. Considering fermentation in particular, *A. oryzae* contain 135 protease genes. Conversely, *A. nidulans* and *A. fumigatus* contain 90 and 99, respectively^{4,19}. All of the orthologues of the protease genes in *A. nidulans* and *A. fumigatus* are found in *A. oryzae*¹⁰. Moreover, *A. oryzae* has specific protease genes not found in the genomes of *A. nidulans* or *A. fumigatus*, whose optimum pH for activity is estimated to be acidic. These findings may support the idea that *A. oryzae* had been adapted, within an acidic pH¹⁰ environment.

Post-genomic approach of *A. oryzae* and identification of new enzymes

Examination of the genome information database of *A. oryzae* (http://www.bio.nite.go.jp/dogan/MicroTop?GENOME_ID=ao) revealed over 130 protease-like genes, about

80% of which are not enzymatically characterized¹⁰. Some of these unknown enzymes may be involved in the degradation of proteins from the material and release of amino acids during the fermentation process. Considering that the 'umami' taste in miso and shoyu is partly due to the released glutamate and aspartate, the peptidases involved in the release of these amino acids are important when manufacturing fermented food. If a new enzyme with novel function (i.e. new substrate specificity, strong stability, activity in the extreme condition such as a high or low temperature or high or low pH), is identified, it is possible to facilitate a fermentation process, or modify the taste of food.

During the characterization process for newly identified genes, some of such gene products were found to exhibit several kinds of substrate specificity previously unreported in research on enzymes from *A. oryzae*. Several aminopeptidases were so far enzymatically characterized by purification through several steps of column chromatography. Such enzymes include leucine aminopeptidases with wide substrate spectra^{15,16,17,18}. These were mainly purified from the solid culture extract called *koji* (steamed or cooked cereals such as rice or soy bean used as solid media of *koji* mold such as *A. oryzae*). They are mainly extracellular enzymes. When using the recently developed molecular genetics approach (from the late 1980s), the transformation technique and protein overproduction work in *A. oryzae* were developed, and some enzymes of *A. oryzae* were overproduced in *A. oryzae* as a host microorganism. These include alkaline protease, whose overproducing strain of *A. oryzae* secreted up to five times more enzyme than the host strain³.

Along with the progress of the genome analysis of *A. oryzae*, several research groups characterized the unidentified enzymes. Among these, one gene designated as *dapA* was identified as an orthologue of *Saccharomyces cerevisiae* aspartyl aminopeptidase, which specifically releases an acidic amino acid from amino termini of peptides. This gene product, DapA, was estimated to consti-

tute 498 amino acids with a molecular mass of 54 kDa. According to the significance of its potentially unique substrate selection, DapA was overproduced as a fusion protein with His-tag in its carboxy terminus in mycelia of *A. oryzae*, under the control of the Taka-amylase gene promoter in inducing a culture condition (Fig. 3)⁸. DapA production in the mycelia of overproducing strain was confirmed with Western blot analysis of the mycelia extract with an anti-His-tag antibody, and also with a preliminary enzymatic assay of the extract using a synthetic substrate, aspartyl-*para*-nitroanilide (Asp-*p*NA). Its hydrolyzing activity in the extract was high compared with that of the parent strain. The addition of a cobalt ion (cobalt chloride; CoCl₂) to the growth medium significantly improved the Asp-*p*NA hydrolyzing activity. Through several purification steps, including a MonoQ column and Ni-IMAC chromatography, the purified DapA was revealed to hydrolyze only Asp-*p*NA and Glu-*p*NA among the tested aminoacyl-*p*NA. When using angiotensin II as a peptide substrate of the purified DapA, hydrolyzing activity was detected for the release of its amino-terminal Asp from angiotensin II, and not detected for the next arginine residue (Fig. 4). Therefore, DapA is a novel aminopeptidase of *A. oryzae* specific for the acidic amino acid in the amino-termini of peptides. During the study of DapA⁸, the same protein was detected in the extract of solid-state culture of *A. oryzae* (steamed soy bean)²⁰. Both studies^{8,20} reported the importance of DapA, which possibly functions as a releasing enzyme for taste molecules (Asp, Glu) in the process of manufacturing fermented food.

Proline, the side chain of which includes a distinctive cyclic structure, is known for its sweet taste. Prolyl aminopeptidase specific for the proline in the amino-termini of peptides was identified in bacteria and filamentous fungi. Prolyl aminopeptidase from *Aspergillus niger* (PapA) was firstly identified as eukaryotic⁸ and the ortholog of the *A. niger* gene was annotated in the *A. oryzae* genome database. By the same strategy of *dapA*, the

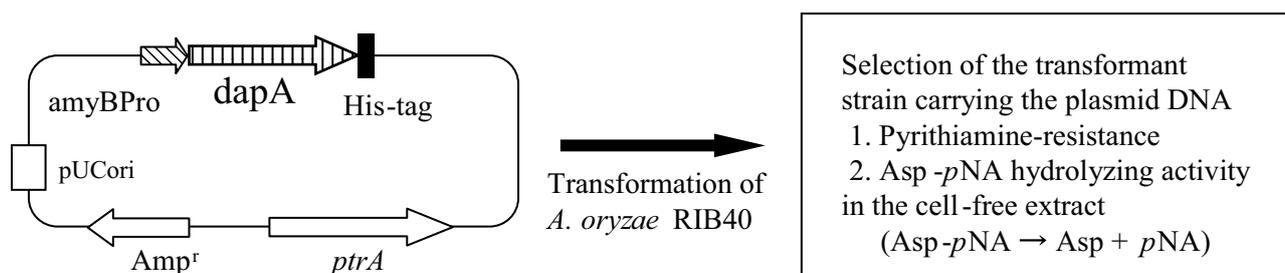
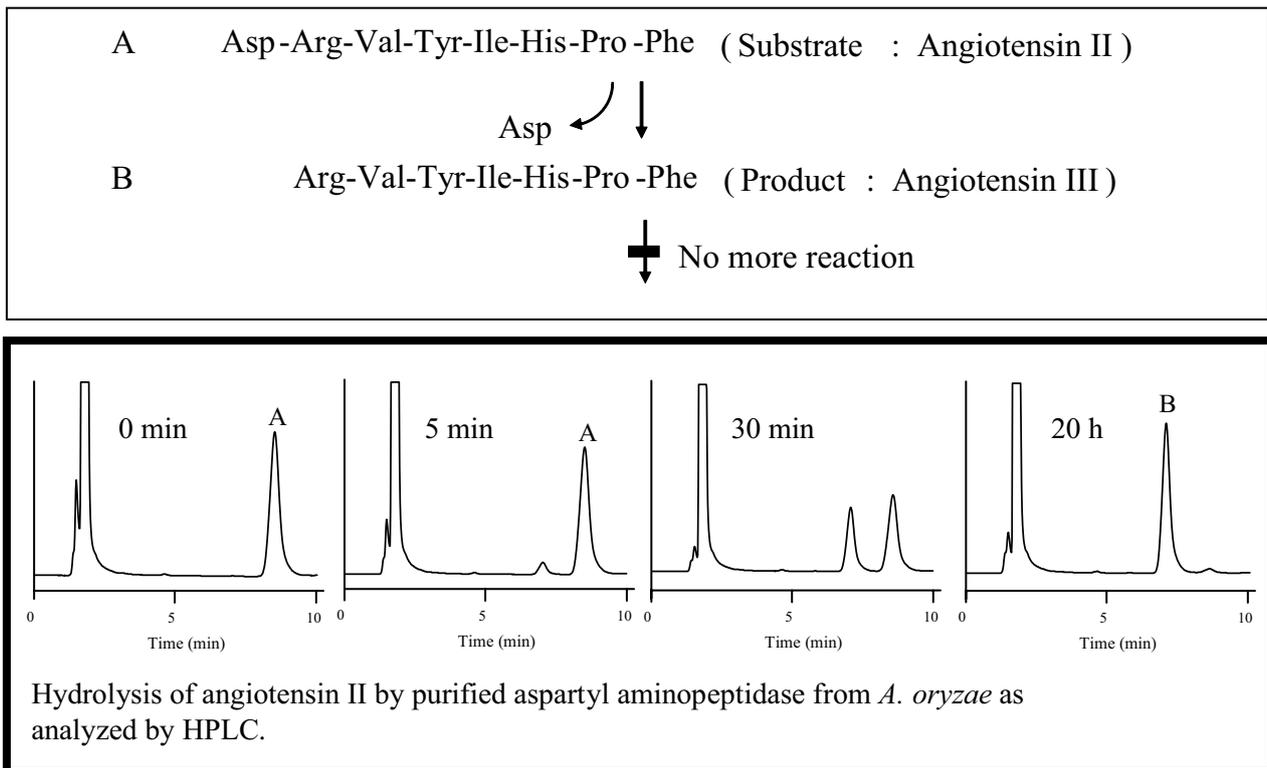


Fig. 3. Preparation of *dapA*-overexpressing strain of *A. oryzae*

amyBpro: promoter of *amyB* gene, His-tag: hexa-histidine-tag, *ptrA*, pyrithiamine-resistance gene: Amp^r, ampicillin-resistance gene, pUCori: replication origin from pUC.



Kusumoto et al. J. Appl. Microbiol. 2008

Fig. 4. Reactivity of DapA against natural peptide

ortholog designated as *pamA* in *A. oryzae* was overexpressed in the mycelia and its gene product was purified from cell-free extract¹¹. The specific activity of Pro-*p*NA hydrolysis in the cell-free extract was about twice as high with glucose as a carbon source for restrictedly inducing an *amyB* promoter compared to soluble starch. The purified PamA enzyme hydrolyzed only Pro-*p*NA among the thirteen aminoacyl-*p*NA substrates tested. Moreover, *A. oryzae* PamA hydrolyzed Pro-beta-naphthylamide was not hydrolyzed by *A. niger* PapA¹¹. The difference of the specificity against synthetic substrate may be due to the approx. 20% difference in the amino acid sequence. PamA was found to be active against peptides with an amino-terminal proline. It hydrolyzed di- and tripeptides almost equally, whereas octapeptide and dodecapeptide showed approx. one tenth the activity of the latter, hence this enzyme seemed to prefer short peptides. *A. oryzae* PamA appeared to be active with a high salt concentration. The enzyme activity at 1-4 M NaCl was mostly equal to the control condition (no NaCl), while its activity at 4.8 M NaCl was about half that of the control condition, meaning PamA might be applicable in the processing step of soy bean paste (about 6-13% NaCl) and soy sauce (about 12-16% NaCl) for the efficient hydrolyzation of peptides.

Consequently, the genome analysis of *A. oryzae* and its post-genome research revealed novel enzymes that may be involved in the degradation of proteins of the start material of fermented food. In the same approach, a novel leucine aminopeptidase from *A. oryzae*, which has broad substrate specificity mainly for the release of leucine from the amino-termini of peptides, were identified in *A. oryzae*¹². Some of the novel proteases are also identified^{13,14} in the same strategy. Such post-genomic approach in *A. oryzae* may develop new scope into the food and pharmaceutical industry in the near future.

Acknowledgments

This work was partly supported by the Program for Promotion of Basic Research Activities for Innovative Biosciences, Japan (PROBRAIN).

References

1. Basten, D. E. J. W. et al. (2005) Characterisation of *Aspergillus niger* prolyl aminopeptidase. *Mol. Genet. Genomics*, **272**, 673–679.
2. Bhattacharyya, A. & Blackburn, E. H. (1997) *Aspergillus nidulans* maintains short telomeres throughout develop-

- ment. *Nucleic. Acids. Res.*, **25**, 1426-1431.
3. Cheevadhanarak, S. et al. (1991) Cloning and selective overexpression of an alkaline protease-encoding gene from *Aspergillus oryzae*. *Gene*, **108**, 151-155.
 4. Galagan, J. E. et al. (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature*, **438**, 1105-1115.
 5. Garcia-Pedrajas, M. D. & Roncero, M. I. (1996) A homologous and self-replicating system for efficient transformation of *Fusarium oxysporum*. *Curr. Genet.*, **29**, 191-198.
 6. Greider, C. & Blackburn, E. H. (1995) Telomeres. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
 7. Kusumoto, K. I., Suzuki, S. & Kashiwagi, Y. (2003) Telomeric repeat sequence of *Aspergillus oryzae* consists of dodeca-nucleotides. *Appl. Microbiol. Biotechnol.*, **61**, 247-251.
 8. Kusumoto, K. I. et al. (2008) Efficient production and partial characterization of aspartyl aminopeptidase from *Aspergillus oryzae*. *J. Appl. Microbiol.*, **105**, 1711-1719.
 9. Long, D. M. et al. (1998) In vivo addition of telomeric repeats to foreign DNA generates extrachromosomal DNAs in the taxol-producing fungus *Pestalotiopsis microspora*. *Fungal. Genet. Biol.*, **24**, 335-344.
 10. Machida, M. et al. (2005) Genome sequencing and analysis of *Aspergillus oryzae*. *Nature*, **438**, 1157-1161.
 11. Matsushita-Morita, M. et al. (2010) Characterization of recombinant prolyl aminopeptidase from *Aspergillus oryzae*. *J. Appl. Microbiol.*, **109**, 156-165.
 12. Matsushita-Morita, M. et al. (2011) Overexpression and Characterization of an Extracellular Leucine Aminopeptidase from *Aspergillus oryzae*. *Curr Microbiol.*, **62**, 557-564.
 13. Morita, H. et al. (2009) Heterologous expression and characterization of Cpl, OcpA, and novel serine-type carboxypeptidase OcpB from *Aspergillus oryzae*. *Appl. Microbiol. Biotechnol.*, **85**, 335-346.
 14. Morita, H. et al. (2010) Molecular cloning of *ocpO* encoding carboxypeptidase O of *Aspergillus oryzae* IAM2640. *Biosci. Biotechnol. Biochem.*, **74**, 1000-1006.
 15. Nakadai, T., Nasuno, S. & Iguchi, N. (1973a) Purification and properties of leucine aminopeptidase I from *Aspergillus oryzae*. *Agric. Biol. Chem.*, **37**, 757-765.
 16. Nakadai, T., Nasuno, S. & Iguchi, N. (1973b) Purification and properties of leucine aminopeptidase II from *Aspergillus oryzae*. *Agric. Biol. Chem.*, **37**, 767-774.
 17. Nakadai, T., Nasuno, S. & Iguchi, N. (1973c) Purification and properties of leucine aminopeptidase III from *Aspergillus oryzae*. *Agric. Biol. Chem.*, **37**, 775-782.
 18. Nakadai, T. & Nasuno, S. (1977) Purification and properties of leucine aminopeptidase IV from *Aspergillus oryzae*. *Agric. Biol. Chem.*, **41**, 1657-1666.
 19. Nierman, W. C. et al. (2005) Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature*, **438**, 1151-1156.
 20. Watanabe, J. et al. (2007) Characterization of *Aspergillus oryzae* aspartyl aminopeptidase expressed in *Escherichia coli*. *Biosci. Biotechnol. Biochem.*, **71**, 2557-2560.