Plant Molecular Biology **13**: 541–550, 1989.

Cloning and nucleotide sequence of the α-galactosidase cDNA from *Cyamopsis tetragonoloba* (guar)

Nico Overbeeke, Arthur J. Fellinger, Marjolein Y. Toonen, Dick van Wassenaar and C. Theo Verrips
Unilever Research Laboratorium Vlaardingen P.O. Box 114, 3130 AC Vlaardingen, Netherlands

Received 2 February 1989; accepted in revised form 19 June 1989

Key words: α-galactosidase, nucleotide sequence, *Cyamopsis tetragonoloba*, guar.

Abstract

Polyadenylated mRNA was purified from the aleurone cells of *Cyamopsis tetragonoloba* (guar) seeds germinated for 18 h and used for the construction of a cDNA library. Clones with the α-galactosidase encoding gene were identified using oligo-nucleotide mixed probes based on the NH_2 terminal amino acid sequence and on the sequence of an internal peptide. The nucleotide sequence of the cDNA clone showed that the enzyme is synthesized as a precursor with a 47 amino acid NH_2 terminal extension. This pre-sequence most likely functions to target the protein outside the aleurone cells into the endosperm. Based upon structural features, it is proposed to divide the precursor into a pre-(signal sequence) part and a glycosylated pro-part comparable with those of the yeast mat A/α factor and killer factor. A comparison of the derived amino acid sequence of this α-galactosidase from plant origin revealed significant stretches of homology with respect to the amino acid sequences of the enzymes from *Saccharomyces cerevisiae* and from human origin but only to a minor extent compared with the α-galactosidase from *Escherichia coli*.

Introduction

The endosperm of legumes like *Cyamopsis tetragonoloba* (guar) consists of an aleurone layer (cells) and of reserve material (mainly galactomannan), which is degraded during seed germination. During imbibition, enzymes responsible for hydrolysis of the galactomannan are leased into the endosperm. The main enzymes identified are α-galactosidase (EC 3.2.1.22), β-mannanase (EC 3.2.1.78) and β-mannoside mannohydrolase (EC 3.2.1.25). For the related legumes fennugreek and lucerne, there are indirect indications that these enzymes are synthesized in the aleurone cells (for a review, see [24]). For the α-galactosidase from guar, we have recently demonstrated this unambiguously by isolation of mRNA from aleurone cells and analysis by hybridization and *in vitro* translation [17].

α-Galactosidase enzymes are present in a wide variety of organisms (for a review, see [12]). In man, the deficiency of this enzyme results in X-linked sphingolipidosis known as Fabry's disease [11]. Enzyme preparations from microbial

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X14619 C tetragonoloba galactosidase alpha.

sources may be of interest commercially; examples are the hydrolysis of raffinose in sugar-beet processing, the hydrolysis of undesired galactose-containing sugars in soya milk and soya protein, etc. [15].

However, for each process, there are preferred specific α-galactosidase enzymes. The α-galactosidase enzyme from guar is particularly suitable for the hydrolysis of guar gum to a gum with improved gelling properties [10]. This α-galactosidase has been purified and studied in detail [23]. The primary structures of only three α-galactosidases have been reported. These were derived from cloned genes, namely the MEL1 gene from *Saccharomyces carlsbergensis* [19] and the gene from human origin [4, 33] and that from *Escherichia coli* [20].

The present work deals with the cloning and nucleotide sequence of the α-galactosidase enzyme from guar, being the first sequence of an α-galactosidase from plant origin. To this end, we used a recently developed method for the purification of RNA from aleurone cells and oligonucleotide mixed probes designed on established amino acid sequences [17].

Materials and methods

Construction of the cDNA library

Polyadenylated RNA was purified from aleurone cells of guar seeds germinated for 18 h as described by Hughes *et al.* [17].

The method for cDNA synthesis was based on the RNase H procedure described by Gubler and Hoffman [16] and modified by Safford *et al.* [31]. THe ds-cDNA was size-fractionated on a 2 ml Sepharose CL4B column (Pharmacia). Fractions larger than 500 bp were collected and dC-tailed, using 40 units of terminal transferase (Bethesda Research Laboratories). The dC-tailed ds-cDNA was annealed with a G-tailed vector (pBR322 cleaved with *Pst* I and dG-tailed from Dupont NEN). These samples were used to transform *E. coli* strain 192.

Colony hybridization of the cDNA library with oligonucleotide mixed probes

Samples of transformed *E. coli* cells were spread directly on nitrocellulose filters (Millipore, type HATF, 0.45 μm) placed on agar plates with tetracycline (10 μg/ml) such that each filter contained $3–5 \times 10^3$ colonies after overnight incubation at 37 °C. Colony hybridization was performed as described by the Bethesda Research Laboratories. Briefly, two identical replicas were made from one master filter and grown for 3–5 h at 37 °C until colonies became visible; these were transferred to L-broth agar plates containing chloramphenicol (150 μg/ml), and incubated overnight at 37 °C. The master filter was replaced on an agar plate containing tetracycline and stored at 4 °C.

The bacteria on the replica filters were lysed by placing them for 15 min on a stack of Whatmann 3MM papers saturated with 0.5 M NaOH and 1.5 M NaCl, neutralization was subsequently performed by placing the filters for 2–3 min on paper saturated with 1 M Tris-HCl (pH 7.0) and 1.5 M NaCl. Finally, the filters were dunked into 3 × SSC for 15–20 s, air-dried and baked at 80 °C under vacuum for 2 h. Prior to (pre)hybridization, the filters were washed extensively in 3 × SSC with 0.1% SDS at 65 °C for 16–24 h. Prehybridization was performed for 2 h at 37 °C in premix A (5 × SSC, 5 × Denhardt's solution, 0.1% SDS, 50 mM sodium phosphate pH 7.5, 1% glycine, 25 μg/ml calf thymus DNA, 75 μg/ml *E. coli* DNA (Sigma, type VIII; both DNA preparations were sheared and denatured), 500 μg/ml tRNA, 50% deionized formamide). Hybridization was performed in hybridization mix A (5 × SSC, 1 × Denhardt's solution, 0.1% SDS, 20 mM sodium phosphate pH 7.5, 25 μg/ml calf thymus DNA, 75 μg/ml *E. coli* DNA, 500 μg/ml tRNA, 50% deionized formamide) with the oligonucleotide probe MP44 labelled with γ-^{32}P-ATP essentially as described earlier [17]. Hybridization was performed overnight at 30 °C. Subsequently, the filters were washed for 3 × 15 min with 6 × SSC at room temperature, 15 min with 2 × SSC, 0.1% SDS, and finally 15 min at 37 °C

in prewarmed 0.1 × SSC with 0.1% SDS. The filters were dried and exposed to X-ray film overnight at − 70 °C. Colonies giving a positive signal on both replicas were picked from the master plate.

Plasmid purification and analysis

Plasmids were purified using the method of Birnboim and Doly [3] and analysed by restriction enzyme digestion performed according to the instructions of the supplier (Amersham, UK). DNA fragments were transferred to a Gene-Screen Plus membrane (Dupont NEN) by Southern blotting [21] and hybridized with oligonucleotide mixed probes essentially as described for the colony blot hybridization but omitting the pre-washing step.

Nucleotide sequence analysis

Single-stranded phage DNA was isolated and used for establishing the nucleotide sequence by the Sanger dideoxy chain termination procedure as modified by Biggin *et al.* [2], using α-^{35}S-dATP (2000 Ci/mmol) and Klenow enzyme (Amersham), ddNTP's (Pharmacia-PL Biochemicals), and dNTP's (Boehringer). The sequencing reaction products were separated on a denaturing polyacrylamide gel using a buffer gradient [2]. In one specific case, namely the region around the internal *Pst* I site, the super-coiled sequencing procedure [8] was used.

The nucleotide sequence was performed using a cascade-like approach (see Results and discussion). The oligonucleotide primers required were synthesized on an Applied Biosystems DNA synthesizer and purified as described earlier [17].

Strains, plasmids and growth conditions

E. coli strain 294 (endoI^-, B_I^-, r_k^-, m_k^+ [1]) was used for the construction of the cDNA library and transformed by the 'Hanahan' procedure [21]. *E. coli* strain JM 103 [25] was used as host for the vector M13mp18 (purchased from Bethesda Research Inc. and described by Norrander *et al.* [27]). This strain was transformed by the $CaCl_2$ procedure and plated on L-Broth with X-gal and IPTG [21]. *E. coli* strains were grown on L-Broth supplemented with antibiotics where appropriate.

Results and discussion

Construction and analysis of the cDNA library

Polyadenylated RNA from aleurone cells of guar seeds germinated for 18 h, was used for cDNA synthesis and a cDNA library was constructed, yielding more than 10^4 colonies. Analysis of the cDNA library for the presence of the α-galactosidase gene according to Hughes *et al.* [17] revealed that out of about 3000 colonies, 7 showed a positive reaction on both filters. From these positive clones, plasmid DNA was purified, digested with *Pst* I and analysed on agarose gels. Next to the 4.3 kb vector fragment, all clones except one had a 1250 bp *Pst* I fragment in common. All clones also had a smaller *Pst* I fragment ranging in size from 300 to 500 bp. These DNA fragments were subsequently transferred to nitrocellulose filters and hybridized with probes MP42, MP43, MP44, based on the NH_2 terminal amino acid sequence and MP33, based on an amino acid sequence from an internal peptide (Fig 1). The small *Pst* I fragment hybridized with probes MP43 and MP44 (Fig. 1), which means that, most probably, the part of the gene encoding the amino terminus (the 5′ part of the cDNA clone) is located on the small *Pst* I fragment. Probe MP33, based on an internal peptide, hybridized with the large *Pst* I fragment; this means that the carboxy terminal part of the protein is encoded by the large *Pst* I fragment. The variation in size of the small *Pst* I fragment has, therefore, most likely resulted from a preliminary termination of the first strand synthesis by reversed transcriptase. From the hybridization of the clones with three independent probes (MP42,

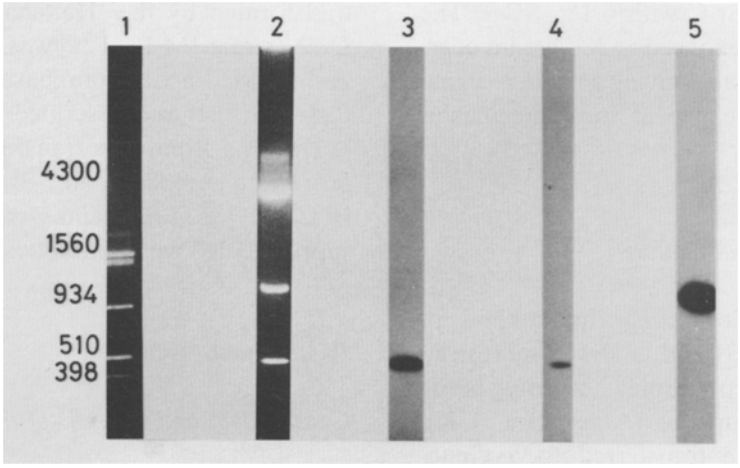

Fig. 1. Analysis of plasmid pUR2302 by hybridization with several oligonucleotide mixed probes. Plasmid pUR2302 was digested with *Pst* I (lane 2) and hybridized with NH_2 terminal probes MP44 (lane 3) and MP43 (lane 4) and with internal probe MP33 (lane 5). Molecular mass standard in base pairs: lane 1.

MP43 and MP33) we conclude we succeeded in cloning the cDNA encoding the α-galactosidase enzyme from guar.

The amount of the guar mRNA estimated from the number of positively-reacting clones (7 out of 3000) is between 0.3% and 0.1%. These figures have not been corrected for the size fractionation performed after cDNA synthesis nor for the presence of incomplete α-galactosidase clones. It remains questionable whether such corrections, if made, could account for the difference between the abundance of α-galactosidase mRNA expected (more than 1% based on the abundancy of the protein) and the amount observed. Furthermore, levels of 0.1–0.3% seem to be in agreement with estimations based on *in vitro* translation of purified mRNA [17]. However, it should be noted that all mRNA preparations were purified after a germination time of 18 h and a subsequent enzyme incubation time of 7 h, which procedure is necessary to remove all polysaccharides [17]. As these conditions make a kinetic study on the mRNA in guar seeds impossible at this moment, we can only give the following theoretical explanation for the discrepancy between abundancy of the α-galactosidase enzyme and the levels of mRNA.

The α-galactosidase gene is induced upon germination resulting in high levels of mRNA (more than 1%). However, the high quantity of mRNA has already passed its maximum and is going down at the moment we are able to purify mRNA from the seeds. In a comparable system, the α-amylase enzyme induced by gibberellic acid in the aleurone cells of barley, it is known that an induction of only a few hours results in a maximal level of α-amylase encoding mRNA [26, 29]. If the induction of α-galactosidase mimicks more or less the α-amylase induction, we can purify the RNA only significantly after reaching the maximum mRNA level. At that moment, the α-galactosidase mRNA level has already dropped considerably.

Nucleotide sequence analysis

From the plasmids harbouring the α-galactosidase cDNA the pUR2302 and pUR2314 were selected and digested with *Pst* I, resulting in fragments of 1250 bp + 500 bp and 1250 bp + 300 bp, respectively, in addition to the vector fragment. These four fragments were purified separately, ligated with vector M13mp18 and used to transform the *E. coli* strain JM103.

The nucleotide sequence was determined using a cascade-like approach (Fig. 2). The sequence at the borders of the subcloned fragments was first

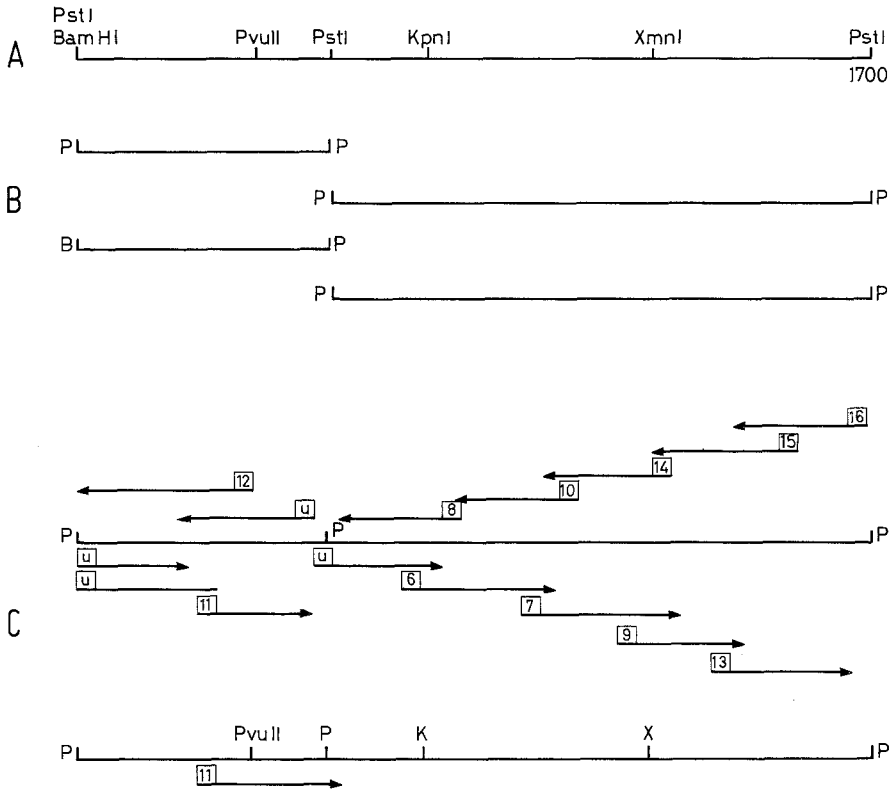

Fig. 2. Restriction map, subclones and sequencing strategy of plasmid pUR2302. A. Preliminary restriction map of the cDNA insert of plasmid pUR2302. B. The fragments from pUR2302 and pUR2314 as subcloned in the M13 vectors M13mp18 and M13mp19. C. Sequencing strategy used for establishing the sequence of pUR2302 and pUR2314. Arrows indicate the direction of sequencing. The boxes indicate the priming site for the sequencing reaction with the different oligonucleotides (U: universal M13 primer; numbers: oligonucleotide primers synthesized for the sequencing experiments). The lower line represents the super-coiled sequencing experiment. P = *Pst* I, K = *Kpn* I, X = *Xmn* I.

determined by using the universal M13 sequencing primer (Amersham). At the most distal position where the sequence data were still reliable, sequencing was continued using two new sequencing primers specially synthesized as described above; one for the continuation of the nucleotide sequence in the 3′-5′ direction and the other for sequencing the complementary strand. By this cascade-like approach, the complete nucleotide sequences for the subclones of the clones pUR2302 and pUR2314 were established. The nucleotide sequence surrounding the internal *Pst* I site was determined using the super-coiled sequencing procedure. Both *Pst* I fragments appeared to be joined directly.

The complete nucleotide sequence of the guar α-galactosidase cDNA clone is shown in Fig. 3. The nucleotide sequence starting with nucleotide 263 up to the poly(A) tail in plasmid pUR2302 is identical with the nucleotide sequence of the insert of plasmid pUR2314. Analysis of the nucleotide sequence shows an open translational reading frame encoding 411 amino acid residues with the first in frame ATG codon at nucleotide position 166 and two consecutive translation stop codon at nucleotide position 1399–1404. This ATG codon at position 166 is also preceded by an A/T rich region making it very likely that it is the translation start. The deduced amino acid sequence is shown in Fig. 3. The NH_2 terminal amino acid sequence of the α-galactosidase enzyme purified from guar seed has been identified as AENGLGQTPP... [17]. This amino acid sequence is encoded by nucleotides 307–336

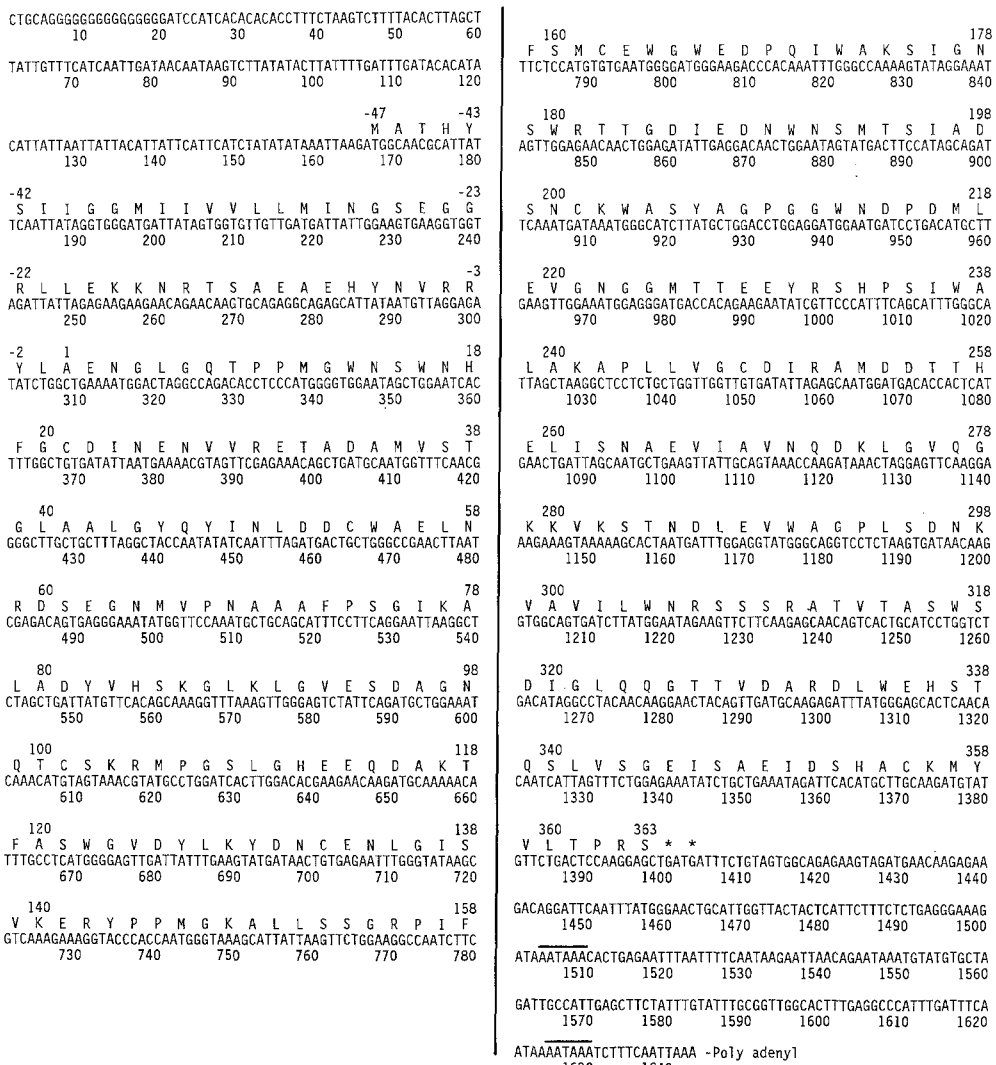

Fig. 3. Complete nucleotide sequence of the guar α-galactosidase cDNA clone pUR2302. The nucleotide sequence of the 5′-3′ strand is shown. The poly G/C tail synthesized in the cDNA cloning procedure is shown on the 5′ site. At the 3′ site it follows the poly(A) tail (not shown). Poly-adenylation consensus sequences [28] are marked with a line above the nucleotides. The amino acid sequence derived from the nucleotide sequence is shown in the one letter notation above the nucleotide sequence. The numbering of the amino acids starts at +1 for the mature protein. The amino acids in the pre-pro sequence are negatively numbered.

(Fig. 3). As the internal peptide DYLKYDN [17] was also found in the nucleotide sequence (nucleotides 680–701) in addition to the NH_2 terminal peptide, it is beyond any doubt that the mRNA for the α-galactosidase enzyme from guar was indeed cloned.

These findings also show that the mature α-galactosidase enzyme is composed of 364 amino acid residues having a calculated molecular mass of 39777 Da. This is in good agreement with the molecular mass of 40.5 kDa estimated by SDS-PAGE [23]. Furthermore, the natural enzyme is most likely synthesized in a precursor form with a 47 amino acid residue extension (numbered −47 to −1 in Fig. 3). This is in agreement with the observation that the *in vitro* translated mRNA encodes a protein with a molecular mass of about 44 kDa [17].

Analysis of the precursor sequence

The precursor sequence of the α-galactosidase is intriguing. Its most likely function will be the targeting of the protein outside the aleurone cells into the endosperm polysaccharide. This situation is very much comparable with that of the α-amylase enzymes from barley and wheat, which are secreted from the aleurone cells into the starch layer. However, the signal sequence from α-galactosidase is different from the signal sequences found for plant enzymes, for instance α-amylase from barley [7], α-amylase from wheat [30], and thaumatin [13]. The signal sequence of α-galactosidase from guar is significantly larger, namely 47 amino acid residues compared with about 24 residues found for the other sequences. Furthermore, at positions −16 and −27 the guar α-galactosidase precursor sequence has two concensus sequences where N-glycosylation may take place, Asn-X-Ser/Thr [32]. Applying the prediction of signal peptidase processing sites as described by Folz and Gordon [14], it is very likely there is a signal sequence processing site between residues −24 and −23. If processing takes place at this site, a prosequence of 23 amino acid residues with one site possible for N-glycosylation is left so that further processing will be necessary. It will be very interesting to know whether this guar α-galactosidase precursor sequence can also function in other plant species. Tobacco may be a good species because it is relatively easy to manipulate genetically.

Although the guar α-galactosidase secretion sequence seems to be different from other known plant secretion sequences, there is a resemblance with the precursor sequences from the mat A/α factor and the killer toxin in the yeast *S. cerevisiae* (for a review, see [6]). The model for processing these yeast precursor sequences is as follows: subsequent to the (assumed) processing of the signal sequence by a signal peptidase in the endoplasmic reticulum, the first steps in glycosylation are performed. Further proteolytic processing occurs in the Golgi by e.g. a protease encoded by the KEX2 gene (cleaving the carboxyl side at the second of two basic residues) and a peptidase encoded by the STE13 gene (a dipeptidyl-amino peptidase which cleaves Glu-Ala and Asp-Ala pairs). In the guar α-galactosidase there are two Arg residues (at −4 and −3), so that a 'KEX-like' protease may be involved. However, another dipeptide aminopeptidase than that in yeast should be required, removing Tyr and Leu in front of Ala being the first residue found in the mature protein. All these observations make it tempting to consider the precursor sequence of guar α-galactosidase as a pre-pro-sequences. Whether it is or not, it is structurally different from the pre-sequences of the α-amylase enzymes from barley and wheat, which are 24 and 23 residues, respectively [7, 30].

Although α-amylase and α-galactosidase are both induced upon germination in the aleurone layer and secreted into the reserve polysaccharide layer, their molecular mechanisms seem to be different. In contrast to α-galactosidase, the induction mechanism for α-amylase is dependent on the presence of gibberellic acid [26, 29]. Besides, the secretion mechanism seems to be different for the α-amylase, which has only a pre-sequence, whereas the α-galactosidase has a pre-pro-sequence. In this context it will also be interesting to establish whether the guar pre-pro-sequence indeed can be recognized by lower eukaryotes such as yeast, as has been reported for the wheat α-amylase signal sequence [30] and in plant species other than guar. One can anticipate a correct processing of the 'KEX-like site' at −4 and −3, but further processing is questionable, which might result in an uncorrectly processed enzyme.

Comparison of the guar α-galactosidase amino acid sequence with other α-galactosidases

A comparison of derived amino acid sequences of α-galactosidase from *S. cerevisiae* [19], human lysosomal [4, 33] and *E. coli* K-12 [20], with known nucleotide sequences, is depicted in Fig. 4. It shows there is a considerable degree of homology between plant α-galactosidase and the enzymes from *S. cerevisiae* and human lysozyme.

Fig. 4. Amino acid sequence homology between the different α-galactosidases. The sequences are aligned for optimal sequence similarity. Asterisks indicate identical residues while + indicate similarities between like residues (D, E; F, Y, W; I, L, V; K, R; Q, N and S, T).

Calculations show there is a 53% functional homology between plant and human α-galactosidase and only little less found for the *S. cerevisiae* enzyme. On the other hand, the α-galactosidase from *E. coli* shows only relatively little homology with respect to the other three enzymes. It also differs from the other α-galactosidases because of the requirement of Mn^{2+} and NAD as cofactors for biological activity [5]. All the enzymes catalyse the same reaction, i.e. splitting α-linked galactose moieties. Although their overall composition, pH optima, enzyme kinetics and their cellular location are different, a common active principle may be expected. Because the exact enzyme mechanism is not yet known it is tempting to speculate that some residues present in the homologous

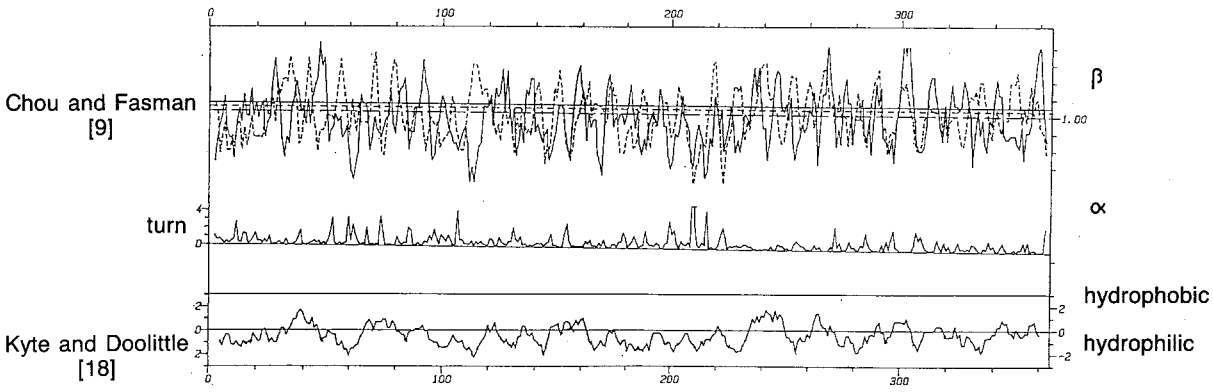

Fig. 5. Prediction of secondary structure elements of guar plant α-galactosidase. α-Helix and β-strand curves are the average of a residue-specific attribute over a window of four using the method of Chou and Fasman [9]. The hydrophobicity curve is the average of a specific hydrophobicity index over a window of nine residues according to Kyte and Doolittle [18].

regions, are part of an active site [12, 22]. An aspartic acid residue might be involved in the active site during the hydrolysis reaction. For the plant α-galactosidase, Asp272 might be such a residue as it is present in a region with a strong homology towards all the other α-galactosidase enzymes. Prediction of the secondary structural elements together with hydrophilicity profile (Fig. 5) reveals that this residue is most likely solvent-accessible due to its location on a predicted β-turn next to two β-strands in a relative hydrophilic area.

The glycosylation pattern of the above enzymes also differs from that of the yeast enzyme, which is heavily glycosylated and from that of the *E. coli* enzyme, which is not glycosylated. It would be interesting to establish whether this glycosylation has any effect on enzyme activity. Such a study requires the production of one enzyme in various hosts so that one can compare (next to e.g. the plant enzyme with plant-like glycosylation) the same enzyme with a yeast-type glycosylation, or the plant enzyme without any glycosylation. Experiments are in progress to express the guar enzyme in *S. cerevisiae* and *B. subtilis*, which might throw more light on the solution of these intriguing questions.

References

1. Backman K, Ptashne M, Gilbert W: Construction of plasmids carrying the cI gene of bacteriophage λ. Proc Natl Acad Sci USA 73: 4174–4178 (1976).
2. Biggin MD, Gibson TJ, Hong GF: Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc Natl Acad Sci USA 80: 3963–3965 (1983).
3. Birnboim HC, Doly J: A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7: 1513–1523 (1979).
4. Bishop DF, Calhoun DH, Bernstein HS, Hantropoulos P, Quinn M, Desnick, RJ: Human α-galactosidase A: Nucleotide sequence of a cDNA clone encoding the mature enzyme, Proc Natl Acad Sci USA 83: 4859–4863 (1986).
5. Burstein C, Kepes A α-Galactosidase from *Escherichia coli* K12. Biochim Biophys Acta 230: 52–63 (1971).
6. Bussey H, Santle D, Greene D, Tippen DI, Bostran KA: Secretion of *Saccharomyces cerevisiae* killer toxin: processing of the glycosylated precursor. Mol Cell Biol 3: 1362–1370 (1983).
7. Chandler PM, Zwar, JA, Jacobson JV, Higgins TJV, Inglis AS: The effects of gibberellic acid and abscisic acid on α-amylase mRNA levels in barley aleurone layers studies using an α-amylase cDNA clone. Plant Mol Biol 3: 407–418 (1984).
8. Chen EY, Seeburg PH: Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. DNA 4: 165–170 (1985).
9. Chou PY, Fasman GD: Prediction of the secondary structure of proteins from their amino acid sequence. Adv Enzymol Relat Areas Mol Biol 47: 45–148 (1978).
10. Critchley P: Commercial aspects of biocatalysis in low-water system, In: Laane C, Tramper J, Lilly MD (eds)

Biocatalysis in Organic Media, pp. 173–183. Elsevier, Amsterdam (1987).
11. Desnick RJ, Sweeley CC: α-Galactosidase A defiency. In: Stanbury JB, Wijngaarden JB, Frederickson DS, Goldstein JL, Brown MS (eds) The Metabolic Basis of Inherited Disease, 5th ed., pp. 906–944. McGraw-Hill, New York (1982).
12. Dey PM, Pridham JBC: Biochemistry of α-galactosidases. Adv Enzymol Relat Areas Mol Biol 36: 91–120 (1972).
13. Edens L, Heslinga L, Klok R, Ledeboer AM, Maat J, Toonen MY, Visser C, Verrips CT: Cloning of cDNA encoding the sweet-tasting plant protein thaumatin and its expression in *Escherichia coli*. Gene 18: 1–12 (1982).
14. Folz RJ, Gordon JI: Computer-assisted predictions of signal peptidase processing sites. Biochem Biophys Res Commun 146: 870–877 (1987).
15. Frost GM, Moss DA: Production of enzymes by fermentation. In: Kennedy JF, Rehm H-J, Reed G (eds) Biotechnology, vol. 7A, Enzyme Technology, pp. 66–155. VCH Verlagsgesellschaft mbH, Weinheim (1987).
16. Gubler U, Hoffman BJ: A simple and very efficient method for generating cDNA libraries. Gene 25: 263–269.
17. Hughes SG, Overbeeke N, Robinson S, Pollock K, Smeets FLM: Messenger RNA from isolated aleurone cells directs the synthesis of an α-galactosidase found in the endosperm during germination of guar (*Cyamopsis tetragonoloba*) seed. Plant Mol Biol 11: 783–789 (1988).
18. Kyte J, Doolittle RF: A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105–132 (1982).
19. Liljeström PL: The nucleotide sequence of the yeast MELI gene. Nucleic Acids Res 13: 7257-7268 (1985).
20. Liljeström PL, Liljeström P: The nucleotide sequence of the *melA* gene, coding for α-galactosidase in *Escherichia coli* K12. Nucleic Acids Res 15: 2213–2221 (1987).
21. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
22. Mathew CD, Balasubramaniam K: Mechanism of action of α-galactosidase. Phytochemistry 26: 1299–1300 (1987).
23. McCleary BV: Enzymic interactions in the hydrolysis of galactomannan in germinating guar: The role of exo-β-mannanase. Phytochemistry 22: 649–658 (1983).
24. Meier H, Reid JSG: Reserve polysaccharides other than starch in higher plants. In: Loewus FA, Tanner W (eds) Encyclopedia of Plant Physiology, New Series 13A, pp. 418–471. Springer Verlag, Heidelberg (1982).
25. Messing J, Crea R, Seeburg PH: A system for shotgun DNA sequencing. Nucleic Acids Res 81: 309–321.
26. Muthukrishnan S, Chandra GR, Maxwell ES: Hormonal control of α-amylase gene expression in barley. J Biol Chem 258: 2370–2375 (1983).
27. Norrander J, Kempe T, Messing J: Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene 26: 101–106 (1983).
28. Proudfoot NJ, Brownlee GG: 3′Non-coding region sequences in eukarotic messenger RNA. Nature 263: 211–214 (1976).
29. Rogers JC, Milliman C: Coordinate increase in major transcripts from the high pI α-amylase multigene family in barley aleurone cells stimulated with gibberellic acid. J Biol Chem 259: 12234–12240 (1984).
30. Rothstein SJ, Lazarus CM, Smith WE, Baulcombe DC, Gatenby AAC: Secretion of a wheat α-amylase expressed in yeast. Nature 308: 662–665 (1984).
31. Safford R, de Silva J, Lucas CM, Windust JHC, Shedden J, James J, Sidebottom CM, Slabas AR, Tombs MP, Hughes SG: Molecular cloning and sequence analysis of complementory DNA encoding rat mammary gland medium-chain *S*-acyl fatty acid synthetase thio ester hydrolase. Biochemistry 26: 1358–1364 (1987).
32. Struck DK, Lennarz WJ, Brew K: Primary structural requirements for the enzymic formation of the N-glycosidic bond in glycoproteins. Studies with α-lactalbumin. J Biol Chem 253: 5786–5794 (1978).
33. Tsuji S, Martin BM, Kaslow DC, Migeon BR, Choudary PV, Stubbleflied BK, Mayor JA, Murray GJ, Barranger JA, Gins EI: Signal sequence and DNA-mediated expression of human lysosomal α-galactosidase A. Eur J Biochem 165: 275–280 (1987).