Electronic Journal of Biotechnology ISSN: 0717-3458 Vol.7 No.3, Issue of December 15, 2004
© 2004 by Pontificia Universidad Católica de Valparaíso -- Chile
Received May 5, 2004 / Accepted September 28, 2004

RESEARCH ARTICLE

Controlling proteolytic degradation of the methionine enriched MB-1Trp protein

Maxime Sasseville
Groupe de recherche en énergie et information biomoléculaire
Département de Chimie-Biologie
Université du Québec à Trois-Rivières
Trois-Rivières, Québec, Canada
Tel: 1 819 376 5011 #3354
Fax: 1 819 376 5084
E-mail: maxime_sasseville@uqtr.ca

Catherine St-Louis
Groupe de recherche en énergie et information biomoléculaire
Département de Chimie-Biologie
Université du Québec à Trois-Rivières
Trois-Rivières, Québec, Canada
Tel: 1 819 376 5011 #3354
Fax: 1 819 376 5084
E-mail: catherine_stlouis@yahoo.ca

Habib Khoudi
Groupe de recherche en énergie et information biomoléculaire
Département de Chimie-Biologie
Université du Québec à Trois-Rivières
Trois-Rivières, Québec, Canada
Tel: 1 819 376 5011 #3354
Fax: 1 819 376 5084
E-mail: habib_khoudi@uqtr.ca

Marc Beauregard*
Groupe de recherche en énergie et information biomoléculaire
Département de Chimie-Biologie
Université du Québec à Trois-Rivières
Trois-Rivières, Québec, Canada
Tel: 1 819 376 5011 #3354
Fax: 1 819 376 5084
E-mail: marc_beauregard@uqtr.ca

Financial support: This work was supported by grants from CORPAQ (Gouvernement du Québec) and NSERC (Canada) awarded to M.B.

Keywords: agro-biotechnology, essential amino acids, protein design, protein engineering, proteolysis.

Abbreviations: A: alanine;
ANSA: anilinonaphthalene sulfonic acid;
BCA: bicinchoninic acid;
D: aspartic acid,
E: glutamic acid;
EAA: essential amino acids;
F: phenylalanine;
H: histidine;
K: lysine;
L: leucine;
M: methionine;
MBP: maltose-binding protein;
N: asparagine;
SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis;
V: valine.

Protein design is currently used for the creation of new proteins with desirable traits, which include a superior nutritional value. One of the challenges of protein design in this area is to achieve the production of stable native-like proteins that resist the proteolytic pressure of the organism used for its production (the bioreactor). We report here the identification of a specific peptide bond sensitive to *E. coli* proteolysis in the designer protein MB-1Trp. In an attempt to reduce proteolysis, we have created a MB-1TrpHis gene library in which the two amino acids surrounding the peptide bond, N44 and L45, were randomized using degenerated

*Corresponding author

oligonucleotides. The initial characterization of MB-1TrpHis N44E/L45V and MB-1TrpHis N44E/L45M, 2 variants of the library that were more resistant than the parent protein, was performed in order to investigate the nature of the mutants' resistance. Our results suggest that the mutants behaved like MB-1Trp regarding folding and thermal stability, and that proteolytic resistance is due to the elimination of the protease recognition site.

The control of proteolysis is crucial in a number of protein applications and in several cellular events (Pipe and Kaufman, 1997; Morrow et al. 2002). In agro-biotechnology, BT toxin which is expressed in insect resistant crops has been engineered to resist a specific inactivating proteolysis by insect gut proteases (Audtho et al. 1999). Similarly, Wyss and coworkers have engineered a longer half-life phytase by decreasing its proteolytic susceptibility, allowing a better assimilation of phytic acid by monogastric farm animals (Wyss et al.1999).

The expenses related to feed additives used for animal production has led to the study of intracellular production of high-quality proteins by transgenic crops and other organisms as a means of obtaining efficient and less costly sources of essential amino acids (EAA) (Dyer et al. 1993; Ealing et al. 1994; Williamson et al. 1996; Forano et al. 2000). Previous attempts to express methionine-rich proteins in different crops achieved various levels of success (Kirihara et al. 1988; Guerche et al. 1990; Altenbach et al. 1992; Aragao et al. 1992; Saalbach et al. 1994; Chaudhuri et al. 1995; Dyer et al. 1995; Teuber et al.1998; Alcocer et al. 2002). Interesting results have been reported in the manipulation of EAA contents through metabolic engineering (Galili and Hofgen, 2002). Approaches that combine both strategies (protein production and metabolic engineering) have been reported in recent patents (Falco et al. 1996; Beach and Tarczynski, 2000).

The issue of proteolytic stability is also of paramount importance for the production of such high value proteins. In this regard, *de novo* design of proteins offers a powerful way to deliver high quality nutrients for animal nutrition, but earlier attempts lead to proteins that were marginally stable and/or had a poor behaviour *in vivo* (Doel et al. 1980; Jaynes et al. 1985). In the early 1990s, we met the stability challenge inherent to designer proteins and created a *de novo* designed sink protein, MB-1, that was shown to have a folded core and a low affinity for 8-anilino-1-naphthalenesulfonic acid (ANSA) (Beauregard et al. 1995). Its behaviour and expression levels *in vivo* were found to be far superior to those obtained from earlier attempts in this field of research (MacCallum et al. 1997). MB-1 was demonstrated to work efficiently as a sink for EAA, with an important advantage over its natural counterpart: the controlled proportion of each EAA included in its sequence.

Despite this initial success, the designer protein MB-1 had a low conformational stability and was only marginally resistant to proteases at physiological temperature (McCallum et al 1997). Such results suggest that MB-1 would not accumulate in bioreactors such as transgenic organisms, where it would be exposed to endogenous proteases (Liao, 1993).

A second generation mutant, named MB-1Trp, was generated and was significantly more thermoresistant than its parent protein MB-1, by up to 16ºC (Gagnon et al. 2000). This increase in thermostability was associated with a general increase in resistance against a non-specific mix of proteases and Pronase E. In spite of this improvement, specific degradation products of our protein were detectable after the purification from *E. coli*, which pointed to existence of particularly sensitive sites. This was also observed in MB-1His and was attributed to specific protease cleavage (Grundy et al. 1998).

In order to improve MB-1Trp resistance to proteolytic degradation, the localization and the modification of sensitive residues was undertaken. Using a combinatorial approach, the residues flanking the targeted bond were substituted by similar residues based on their predicted positions in the protein. Two stabilized mutants of the library were further characterised *in vitro* in order to understand the effect of the mutations on MB-1Trp stability and behaviour.

MATERIALS AND METHODS

MB-1Trp expression and purification

MB-1Trp was expressed and purified according to the method described previously (Gagnon et al. 2000). Protein samples intended for mass spectrometry were purified on a DEAE-Sepharose column, equilibrated with 10 mM Tris-1 mM EDTA containing 0.1% Tween-20. The protein was eluted using a continuous NaCl gradient (0 to 1000 mM). Fractions were analysed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) using 16% polyacrylamide-tricine gels and colored by silver nitrate staining (Amersham Biosciences). The proteins were concentrated and dialysed as described previously (Gagnon et al. 2000).

Matrix Assisted Laser Desorption Ionisation-Time of Flight (MALDI-TOF) mass spectrometry

The size of the fragments was determined using an Applied Biosystems MALDI-TOF Mass Spectrometer Voyager System 2016 with an α-cyano-4-hydroxycinnamic acid matrix. WinPep software (Hennig, 1999) was used to calculate the mass of the fragments as deduced from sequence.

MB-1TrpHis construction, expression and purification

Table 1. Mass spectrometry analysis of MB-1Trp proteolytic fragments.

Peak no.	Mass (Da)	Relative Intensity (a.u.)	Fragment	Calculated Mass (Da)
1	10674	65	M1-N93	10663
2	10318	30	M1-H90	10308
3	5002	75	M1-N44	5000
4	4645	25	M1-H41	4644
5	6775	15	L42-A100	6765
6	5331	20	L45-H90	5326

To allow a C-terminal specific purification of MB-1Trp and its fragments, and detection with a commercially available anti-His-Tag monoclonal antibody, a poly-histidine tag (His-tag) was ligated to the C-terminal end of MB-1Trp. This was achieved by constructing a chimeric plasmid from the original pMal-c2 MB-1Trp and pMal-c2 MB-1His plasmids (Grundy et al. 1998). Protein samples for N-terminal sequencing were purified by amylose affinity chromatography followed by an immobilized metal affinity chromatography (IMAC), as described earlier (Grundy et al. 1998).

N-terminal sequencing

Proteins were transferred to a polyvinylidene fluoride (PVDF) membrane as described by Geisow and Aitken (1989). The proteins were then sequenced using an ABI model 473A Protein Sequencer (Service Protéomique de l'Est du Québec, Sainte-Foy, Canada).

Preparation of MB-1TrpHis mutants

Substitutions of asparagine in position 44 and leucine in position 45 of MB-1TrpHis were performed using the "Site-Directed Mutagenesis Kit" (Stratagene). The asparagine codon (AAC) was replaced by a GA(A/T) codon, encoding for E or D. The leucine codon (CTG) was replaced by a (A/T/G)T(G/C) codon, encoding V, M, I, L and F in the following proportion: 2:1:1:1:1 respectively. A total of ten different mutation sets were therefore available with this strategy. The mutagenic oligonucleotides, purchased from Biocorp (Montréal, Canada), are shown below along with the corresponding MB-1TrpHis original sequence. Oligonucleotides were purified using polyacrylamide gel electrophoresis and phosphorylated.

MB-1TrpHis:

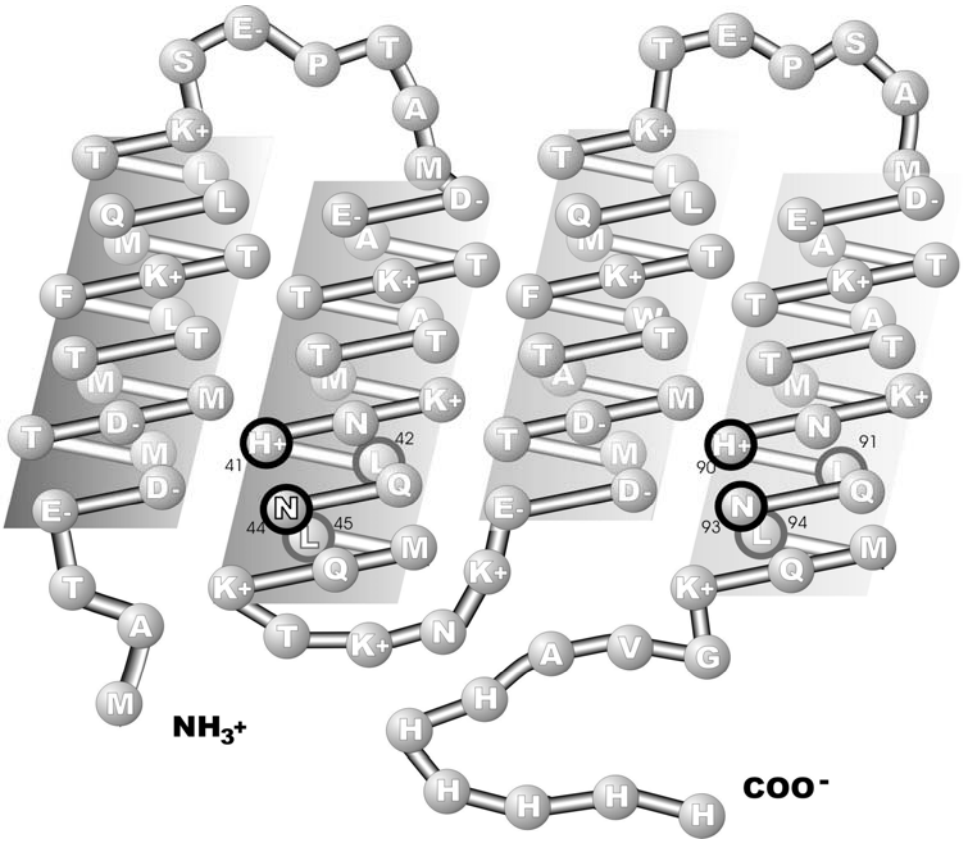

Figure 1. Flattened model of MB-1TrpHis. The illustration shows the predicted position of buried residues in the background and the exposed residues of the four helices in one plane at the forefront. Residues flanking the protease sensitive peptide bonds are indicated by thick circles.

Table 2. Mutations distribution in the MB-1TrpHis mutants library.

Mutations in position 44-45	Clones sequenced
E – V	1
E – M	2
E – F	1
D – V	3
D – M	2
D – F	3
D – I	2

5'-CAATG AAG AAT CAT CTT CAA **AAC CTG** ATG CAG AAG ACT AAG AAC-3'

Oligo NL#1:

5'-CAATG AAG AAT CAT CTT CAA **GAW DTS** ATG CAG AAG ACT AAG AAC-3'

Oligo NL#1-R:

5'-GTT CTT AGT CTT CTG CAT **SAH WTC** TTG AAG ATG ATT CTT CATTG-3'

where **W** stand for A or T; **D** stand for A, T or G; **S** stand for C or G; and **H** stand for A, C, or T.

The mutations were confirmed by dideoxynucleotide DNA sequencing. The mutated genes were expressed in fusion with the Maltose-Binding Protein (MBP) in pMal-c2 expression vectors (New England Biolabs).

Proteolytic resistance screening by Western blotting

E. coli XL-1Blue cells (*rec*A1, *end*A1, *gyr*A96, *thi-1*, *hsd*R17(r_K^-,m_K^+), *sup*E44, *rel*A1, *lac*$^-$, [F' *pro*AB, *lac*$I^qZ\Delta$ *M15*, Tn10(Tet^r)]) (Stratagene, Cedar Creek, USA) harbouring pMal-c2 MB-1TrpHis mutant plasmids were grown in 15 mL of LB Miller medium (EM Science, Hawthorne, USA) supplemented by 100 mg l-1 ampicillin (Sigma) to an optical density of 0.3 at 600 nm. Transcription was induced by adding 1 mM isopropylthio-b-D-galactoside (IPTG) and cells were grown for an additional 3 hrs. Cells were then centrifuged at 4000 g for 10 minutes and resuspended in 5 ml cleavage buffer (20 mM Tris, 100 mM NaCl and 3 mM $CaCl_2$). A 30 seconds sonication (60% output control on a Branson Sonifier 250) was performed to break the cell wall. The solutions were then incubated at 37ºC in order to reactivate *E. coli* proteases and challenge the mutants' resistance to proteolytic degradation. After incubation at 37ºC (0 to 30 min), factor Xa (New England Biolabs) and cleavage buffer were added, followed by an overnight incubation at 4ºC in order to cleave MB-1 variants from their MBP fusion. Alternatively, the cleavage with factor Xa was performed directly after sonication, and the proteins resistance to proteolytic degradation was challenged after separation of the fusion partner MBP from the designer proteins MB-1TrpHis (and mutants). For that, 7 ml of freshly prepared non-transformed *E. coli* extract was added to the mixture before incubation at 37ºC for 0 to 45 min.

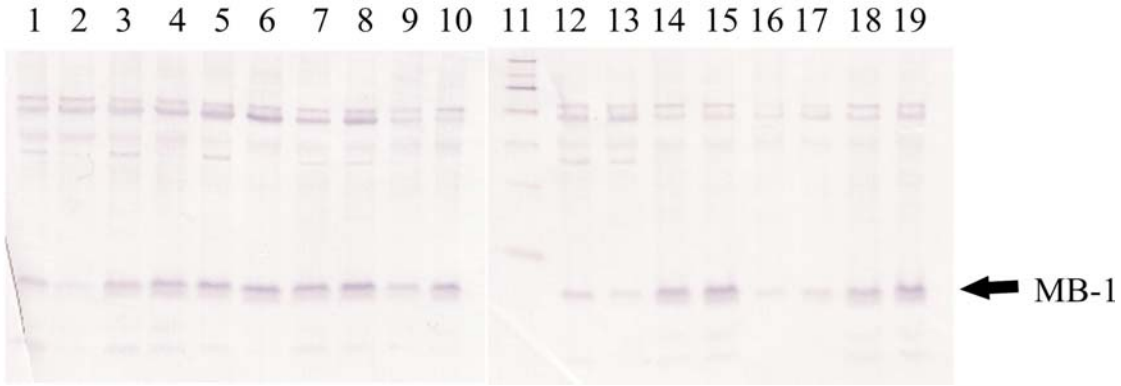

Figure 2. Western blot analysis of the proteolytic resistance of the protein and mutants performed after 0 and 30 min exposure to *E. coli* proteases.

Lane 1: MB-1TrpHis 0 min;
Lane 2: MB-1TrpHis 30 min;
Lane 3: E44V45 0 min;
Lane 4: E44V45 30 min;
Lane 5: E44M45 0 min;
Lane 6: E44M45 30 min;
Lane 7: D44I45 0 min;
Lane 8: D44I45 30 min;
Lane 9: D44F45 0 min;
Lane 10: D44F45 30 min;
Lane 11: protein Western markers band of 15, 25, 35, 50, 75, 100 and 150kDa;
Lane 12: MB-1TrpHis 0 min;
Lane 13: MB-1TrpHis 30 min;
Lane 14: D44M45 0 min;
Lane 15: D44M45 30 min;
Lane 16: E44F45 0 min;
Lane 17: E44F45 30 min;
Lane 18: D44V45 0 min;
Lane 19: D44V45 30 min.
Proteolysis was performed before cleavage of MBP-MB-1Trp mutant fusions.

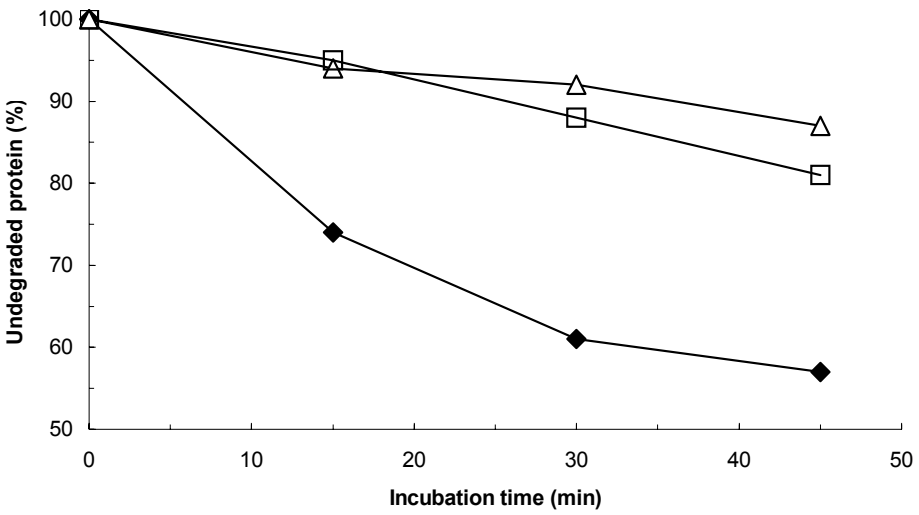

Figure 3. Proteolytic degradation of MB-1TrpHis and mutants by *E. coli* proteases. At variance with results in Figure 2, proteolysis was performed after separating MB-1TrpHis mutants from its fusion partner MBP.

Symbols: ♦ MB-1TrpHis; Δ E44V45; □ E44M45.

The average standard deviation for undegraded protein measurements is 5%.

Following incubation with *E. coli* proteases, the solutions were denatured using 2% SDS and heated at 95ºC for 5 min. Samples were analysed by a SDS-PAGE followed by an overnight transfer onto a nitrocellulose membrane (#162-0146, Bio-Rad, Hercules, USA) following the manufacturer's instructions. The Western blotting was performed using a mouse anti-His-Tag monoclonal antibody (Novagen, Darmstadt, Germany) and a goat anti-mouse secondary antibody conjugated to alkaline phosphatase. The revelation of the binding was performed by a standard colorimetric Western blotting procedure (Novagen). Alternatively, the Western blot experiment was performed using a rabbit anti-MB-1 polyclonal antibody followed by reaction with goat anti-rabbit HRP-conjugated (Sigma) and chemi-luminescent detection reagents (Roche). After exposure to nitrocellulose membrane, Kodak films were scanned and protein band densities were evaluated using Scion Image software.

Expression and purification of MB-1TrpHis mutants

All mutants were expressed and purified as described in Gagnon and coworkers (Gagnonet al. 2000) with a second chromatography involving metal affinity purification using the modifications cited above.

Protein quantification and purity evaluation

For all experiments, protein concentration was determined by bicinchoninic acid (BCA) assay (Sigma) using bovine serum albumin as a standard. SDS-PAGE experiments were conducted prior to measurements to confirm protein purity.

Conformational investigation by Circular Dichroism (CD)

CD measurements were carried out as described by Gagnon and coworkers (Gagnon et al. 2000). Experiments were conducted at a protein concentration of 0.6 mg ml-1 in a phosphate buffer (128 mM NaH_2PO_4, pH6.8). CD signal was converted to $\Delta\varepsilon_{MRW}$ which is the difference in extinction coefficient per mole of residues. Secondary structures were calculated using CDSSTR (Curtis Johnson, 1999) with default settings.

Thermal denaturation

The samples were prepared as described in the previous section and conformational changes were monitored by CD. To measure thermostability, temperature was increased from 5 to 98ºC at a rate of 20ºC h-1 using a Neslab RTE-111 controlled by Jasco spectropolarimeter software. CD spectra were collected every 5ºC from 190 to 260 nm, at a speed of 20 nm min-1, and CD signals at 222 nm were collected every 1ºC. Thermal stability was calculated assuming a unimolecular, two-state process as previously described in Pace et al.1989. The CD signal at 222 nm measured at various temperatures was used as the property (y) indicative of the extent of a conformational change assumed to be unfolding. This model is accurate for MB-1Trp since 1- it was shown to be monomeric at room

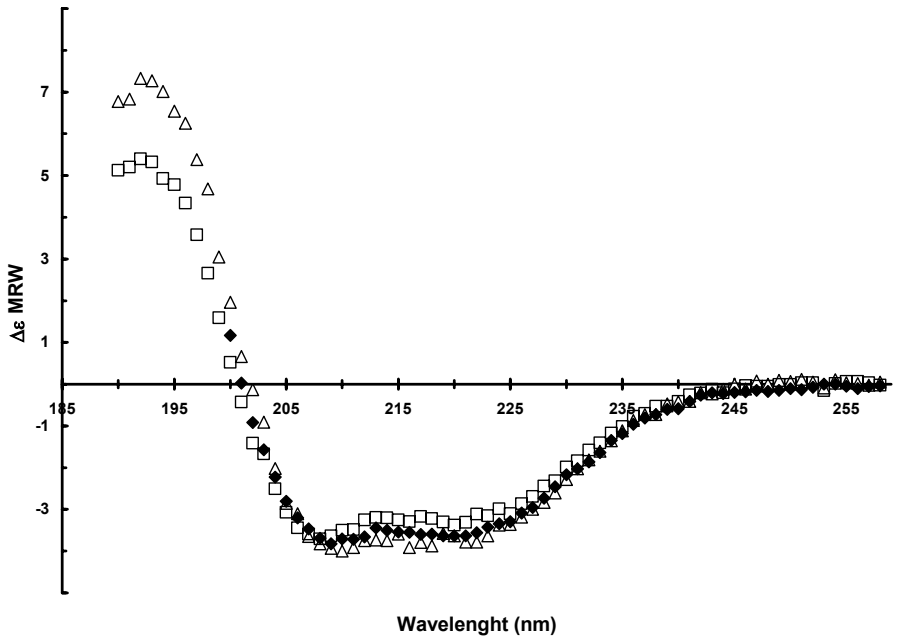

Figure 4. CD spectra of MB-1Trp and MB-1TrpHis evolved mutants. All spectra have the usual features observed for helical proteins.

Symbols: ♦ MB-1TrpHis; Δ E44V45; □ E44M45.

temperature, 2- the protein can be reversibly unfolded (Gagnon et al. 2000) and 3- at room temperature MB-1Trp was shown to be highly helical. In this "folded state", the parameter $y = y_f$ and the fraction of folded protein f_f is equal to 1. When the designer protein has changed its conformational state as evidenced by the loss of helical content and precipitation when incubated at high temperature, it is assumed to be in an unfolded state, typical of a denatured natural protein. The parameter $y = y_u$, and the fraction of unfolded protein f_u is equal to 1. For intermediate states, y is given by $y_f f_f + y_u f_u$. Thus, by measuring y, we can calculate the fraction of unfolded protein: $f_u = (y_f - y)/(y_f - y_u)$. The equilibrium constant for the unfolding process is $K_u = f_u/(1-f_u)$ and melting temperatures (T_m) are obtained at $K_u = 1$ (Pace et al. 1989).

RESULTS

Identification of the protease sensitive peptide bonds

Mass spectrometry was performed on MB-1Trp fragments purified using anionic exchange purification (as detailed in Grundy et al. 1988). The assignment of mass spectrometry results is described in Table 1. Seven fragments were tentatively assigned, the ones with the largest signal being M1-N93 (peak no. 1) and M1-N44 (peak no. 3). For some fragments slight differences in calculated and measured masses were found. They might be ascribed to mass changes caused by ionisation and/or post-translational modifications, such as methionine oxidation, which is not unexpected considering the high methionine contents of MB-1Trp (Fontana et al. 1997; Hollemeyer et al. 2002). The assignment of peaks no. 1 and 3 was confirmed by N-terminal sequencing. To facilitate the fragment identification process, N-terminal sequencing was performed on MB-1 TrpHis fragments recovered from an IMAC column. Only fragments that carry the intact C-terminus are retained under such conditions. After recovery from *E. coli,* several fragments were obtained, the two most important bearing the sequence: LMQKTKNKE and LMQKGVAH at their N-terminus respectively. Such sequences correspond to amino acids 45-53 and 94-101 respectively. Therefore, both approaches (mass spectrometry and N-terminal sequencing) point to a cleavage between N44 and L45, and a second site at N93 and L94. The detection of the second site is not unexpected, because in MB-1Trp, helices 2 and 4 are predicted equivalent regarding both conformation and sequence (per design, Figure 1). Furthermore, the correlation between fragments size and N-terminal sequencing results suggest that adding the His tag to MB-1Trp did not change its sensitivity to proteases.

Another minor fragment had the amino-terminus sequence LQNLMQ, which corresponds to either segment 42-47 associated to peak no. 5 (L42 – A 100) or 91-96, associated to peak no. 2 and 6 (M1-H90 and L45-H90 respectively). The population of fragment that has been cut using the site N44-L45 (Fragment M1-N44, peak no. 3) appeared more important than the ones being cut at the site H41-L42 (Fragment M1-H41 peak no. 4 and fragment L42-A100,

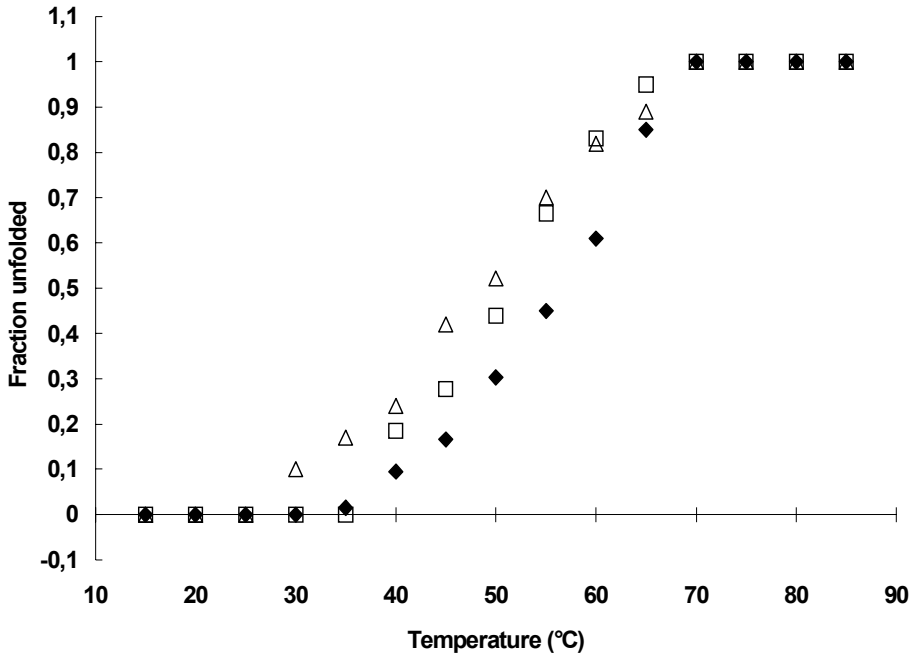

Figure 5. Denaturation curves calculated from CD data obtained at various temperatures.

Symbols: ♦ MB-1TrpHis; Δ E44V45; □ E44M45.

peak no. 5). Similarly, the fragment population of M1-N93 (peak no. 1) was more important than the M1-H90 population (peak no. 2). Thus, the N-L bonds appear to be more susceptible or accessible than H-L bonds, probably because they are more exposed as predicted per design (Figure 1). Among the two N-L bonds identified, the N44-L45 bond was selected for our study. A cleavage right in the middle of the protein was believed to have a more important impact on protein integrity and "residual" resistance after a first cleavage.

Combinatorial mutagenesis

The predicted positions of the two amino acids targeted by proteolysis (N44, L45) are shown in Figure 1. It was hypothesised that changing the residues directly involved at this peptide bond would decrease sensibility to a putative specific protease of *E. coli*. The N44 residue was allowed to be mutated to D or E, two acidic residues. Interactions with the neighbouring basic residues, H41 and K48, could promote the formation of stabilizing intrahelical salt-bridges. The structural stabilisation of this region of helix 2 could also contribute to prevent proteolytic activity. The L45 residue was allowed to be mutated to the hydrophobic amino acids, I, V, M, F and the wild-type L. With this strategy, we hoped that the most accommodating residues would result in more stable mutants while providing a change in the amino acid sequence at the target site. Since L45 is predicted to be buried in the hydrophobic core, where mutations are usually not well tolerated, a wide variety of non-polar amino acids was allowed (Sanchez-Ruiz and Makhatadze, 2001).

Clones generated by the mutagenesis reaction were sequenced to assess the distribution of the mutations in our bank. Seven out of 10 possible combinations were recovered from the mutant library as shown in Table 2. Interestingly, both possible mutant with a conserved L in position 45 were absent from the clones sampled for sequencing. There was no obvious bias in the distribution of the mutations.

Screening mutants for proteolytic resistance

Randomly selected mutants were screened for their proteolytic resistance by exposing them to proteolytic degradation assays, followed by an analysis by Western blotting with an anti-His-Tag monoclonal antibody. For degradation assays, MB-1TrpHis (containing a His tag) was used as the "wild type" protein instead of MB-1Trp. Proteins were exposed to bacterial proteases before and after the proteins were separated by factor Xa from their fusion partner MBP. As can be seen on Figure 2, several mutants outperformed their parental counterpart MB-1TrpHis when exposed to *E. coli* proteases: The disappearance of MB-1TrpHis after a 30 min incubation at 37ºC is evidenced by the comparison of the band intensities in lanes 1 vs 2, and 12 vs 13. Most of the 7 mutants analysed outperformed the protein MB-1TrpHis in terms of relative amount of protein left after incubation. The presence of proteolytic products in all mutant preparation, the absence of inclusion bodies, and the comparative yields

in MBP-MB-1 mutant fusion protein obtained during purification (data not shown) all indicate that the differences in proteolytic resistance are not due to differences in solubility among the proteins analysed.

Careful examination of Figure 2 allows the detection of several protein bands that differ from MB-1TrpHis and its mutants (band position shown by an arrow). In the upper part of the gels, the first band at the top is the MBP-MB-1TrpHis fusion protein that escaped cleavage with factor Xa, migrating at 54 kDa. The next band is the large cleavage product of the fusion: MBP, running at 43 kDa. Then, few bands in the mid-range (25-35 kDa) are believed to be *E. coli* proteins with histidines at their surface, resulting in some affinity for the antibody used in this experiment.

The screened clones were sequenced to assess the distribution of the stabilizing mutations. The most stabilized mutants carry the mutation combinations E-V, E-M, D-I, D-V and D-M. The least stable mutants carry the mutation combinations E-F (lanes 16-17) and D-F (lanes 9-10). These results could be explained by the fact that the hydrophobic core might not be able to harbour the large phenylalanine side-chain in position 45, which could promote local unfolding of the protein and expose protease targets (Goldberg et al. 1978; Betz et al. 1993).

The impact of the fusion partner on the mutant resistance to proteolysis was investigated in Figure 3. For these results, the fusion proteins MBP-MB-1Trp mutants were cleaved prior to incubation with bacterial proteases. Figure 3 shows a comparison of MB-1TrpHis with the most stable mutants (MB-1TrpHis E44V45, referred to as "E44V45", and MB-1TrpHis E44M45, renamed "E44M45"). While 40% of MB-1TrpHis has disappeared after 40 min incubation with proteases, only 10-15% of both mutants initial populations was lost to proteolysis. Such results confirm the intrinsic stabilisation of the nutritive protein MB-1TrpHis by the mutations performed here and the absence of interference due to some protective effect from the fusion partner MBP. Note that Figure 2 was generated using antibody to the His tag, while Figure 3 describes results obtained using antibody to MB-1Trp. Therefore, the specificity of the detection protocol has no impact on the apparent stability of the designer proteins under study.

Smaller protein fragments (probably MB-1TrpHis proteolytic products) were detected in the parent protein electrophoretic pattern and for most mutants. They appear to be less important for the mutants however, when considering the relative amount of full length product vs degradation bands.

Effect of mutations on secondary structure

Secondary structures of the two most stable mutants MB-1TrpHis N44E/L45V and MB-1TrpHis N44E/L45M were analysed by CD. The spectra measured for the mutants were typical of helical proteins and equivalent to MB-1Trp (Gagnon et al. 2000) as shown in Figure 4. The two mutants' secondary structures were calculated using CDSSTR algorithm (data not shown). No significant changes in the secondary structures were detected, which indicates that proteolytic resistance is not the consequence of conformational changes that would promote hindrance for protease access to its target.

Effect of mutations on conformational stability

Thermal denaturation experiments were carried out on the two purified mutants using CD as described above in Materials and Methods. Their melting temperatures (*Tm*) are 55ºC for MB-1Trp, 52ºC for E44M45, and 53.5ºC for E44V45. Denaturation curves shown in Figure 5 indicate that each mutants were slightly less stable than the mother protein, with a similar degree of cooperativity for the transitions studied.

DISCUSSION

The design of novel proteins with native-like properties is, thus far, still a challenging issue. Many early attempts have given ambiguous results demonstrating molten globule states, loosely packed proteins and polydisperse protein populations (Hill et al. 2000). In particular, β-based structures have proven difficult to design in a highly soluble state or without aggregation problems (Beauregard et al. 1991; Jin et al. 2003). Recent improvements in core modelling and packing algorithm have lead to engineering of well behaved proteins (Dahiyat et al. 1997). It is worth mentioning that such computer generated cores do not include significant amount of Met, if any, and their AA composition is not among the criteria used for their design.

Obviously, designer proteins did not go through billions of years of optimisation via evolution like natural proteins did. This in part may explain the poor characteristics of several *de novo* designed proteins. Another factor is our limited comprehension of all the rules that govern protein folding and dynamics. Nutritious designer proteins are not exempt from such limitations, and early attempts have revealed low stability and low protein yield *in vivo* (Doel et al. 1980; Jaynes et al. 1985).

Instead of modifying their protein of interest, other groups have modified the production host by disabling the specific protease that digest the protein (Heo et al.2002). Obviously, such a strategy cannot be of application for all available bioreactors. Further, the use of nutritive proteins for ruminants requires a high degree of intrinsic resistance to proteases for them to behave as by-pass proteins and have an optimal impact on animal performances (Stern et al. 1985).

We reported here a continuation of our rational approach to designing stable proteins with the additional constraint

represented by a biased content in selected EAA (M, T, K and L). The mutations introduced to the protein MB-1Trp were shown to have an important impact on its resistance to *E. coli* proteases (Figure 3). This alone is a major improvement in MB-1Trp behaviour. Enhancement of resistance to proteolysis afforded by specific interactions with the fusion partner MBP was considered and ruled out: the evolved proteins studied here were more resistant before and after cleavage from MBP. In order to understand the nature of the stabilization brought by the mutations, their effect on MB-1Trp conformation was assessed. As shown by CD and thermal denaturation experiments, these contributions appeared marginal. This implies that the improvement probably came from the simple deletion of the targeted sequence recognised by the protease(s). The improvement in resistance achieved by modifying residues 44 and 45 supports the validity of our strategy in selecting this particular area as described above.

The sites of proteolysis could be informative on the secondary structure locations. As proteases cut in loop or accessible structures preferably (Yoo et al. 2001; Heo et al. 2002), one could speculate that the region around position 44-45 would not be in a helical conformation as predicted, but would rather be in unordered or flexible conformation. This would be in agreement with the helical content of MB-1Trp, which is lower than expected per design (55% vs 75%). Premature ending of the second and fourth helices could lead to exposure of these regions. Further experiments including non-specific proteinase K degradation will help us to locate more precisely the helices, as they are not cleaved under those conditions (in globular protein in native state, short incubation times and low protease concentration; Polverino de Laureto et al. 1995).

However, as the targeted bonds were identified to be the same in second and fourth helices, it gives us evidence that the structure around the homologous site Asn93-Leu94 could be quite similar and that our strategy, applied to these two positions, would similarly improve the behaviour of MB-1Trp, at least for the yield of full length protein.

Our strategy of protease target location reveals itself to be an effective and simple way to increase MB-1Trp half-life in a normal cellular environment. It would be interesting to see how this set of mutations will affect the accumulation of our protein in plant cytosol for example. Our goal is to improve the stability of EAA enriched proteins to a level comparable to that of natural crop proteins.

ACKNOWLEDGMENTS

The authors are indebt to Dr Catherine Servis at Institut de Biochimie, Université de Lausanne (Epalinges, Switzerland) who performed the mass spectrometry measurements in collaboration with Simon Gaudrault. Our thanks to François O. McDuff for skilful preparation of figures.

REFERENCES

ALCOCER, Marcos J.C.; MURTAGH, Gareth J.; BAILEY, Kevin; DUMOULIN, Mireille; MESEGUER, Amparo Sarabia; PARKER, Martin J. and ARCHER, David B. The disulphide mapping, folding and characterisation of recombinant Ber e 1, an allergenic protein, and SFA8, two sulphur-rich 2S plant albumins. *Journal of Molecular Biology*, November 2002, vol. 324, no. 1, p. 165-175.

ALTENBACH, Susan B.; KUO, C.C.; STARACI, L.C.; PEARSON, K.W.; WAINWRIGHT, C.; GEORGESCU, A. and TOWNSEND, J. Accumulation of a Brazil nut albumin in seeds of transgenic canola results in enhanced levels of seed protein methionine. *Plant Molecular Biology*, January 1992, vol. 18, no. 2, p. 235-245.

ARAGAO, F.J.; DE SA, M.F.; ALMEIDA, E.R.; GANDER, E.S. and RECH, E.L. Particle bombardment-mediated transient expression of a Brazil nut methionine-rich albumin in bean (Phaseolus vulgaris L.). *Plant Molecular Biology*, October 1992, vol. 20, no. 2, p. 357-359.

AUDTHO, Mongkon; VALAITIS, Algimantas P.; ALZATE, Oscar and DEAN, Donald H. Production of chymotrypsin-resistant Bacillus thuringiensis Cry2Aa1 delta-endotoxin by protein engineering. *Applied and Environmental Microbiology*, October 1999, vol. 65, no. 10, p. 4601-4605.

BEACH, Larry and TARCZYNSKI, Mitchell C. Methods of increasing accumulation of essential amino acids in seeds. United States Patent #6,127,600, (2000).

BEAUREGARD, M.; GORAJ, K.; GOFFIN, V.; HEREMANS, K.; GOORMAGHTIGH, E.; RUYSSCHAERT, J.M. and MARTIAL, J.A. Spectroscopic investigation of structure in octarellin (a de novo protein designed to adopt the alpha/beta-barrel packing). *Protein Engineering*, October 1991, vol. 4, no. 7, p. 745-749.

BEAUREGARD, Marc; DUPONT, Claude; TEATHER, Ron M. and HEFFORD, Mary A. Design, expression and initial characterization of MB-1, a de novo protein enriched in essential amino acids. *Biotechnology* (N.Y.), September 1995, vol. 13, no. 9, p. 974-981.

BETZ, Stephen F.; RALEIGH, Daniel P. and DEGRADO, William F. De novo protein design: from molten globules to native-like states. *Current Opinion in Structural Biology*, August 1993, vol. 3, no. 4, p. 601-610.

CHAUDHURI, Sumita and MESSING, Joachim. RFLP mapping of the maize dzr1 locus, which regulates methionine-rich 10 kDa zein accumulation. *Molecular and General Genetics*, March 1995, vol. 246, no. 6, p. 707-715.

CURTIS JOHNSON, W. Analyzing protein circular dichroism spectra for accurate secondary structures.

Proteins, May 1999, vol. 35, no. 3, p. 307-312.

DAHIYAT, Bassil I.; SARISKY, Catherine A. and MAYO, Stephen L. De novo protein design: Toward fully automated sequence selection. *Journal of Molecular Biology*, November 1997, vol. 273, no. 4, p. 789-796.

DOEL, M.T.; EATON, M.; COOK, E.A.; LEWIS, H.; PATEL, T. and CAREY, N.H. The expression in E. coli of synthetic repeating polymeric genes coding for poly(L-aspartyl-L-phenylalanine). *Nucleic Acids Research*, October 1980, vol. 8, no. 20, p. 4575-4592.

DYER, J.M.; NELSON, J.W. and MURAI, N. Extensive modifications for methionine enhancement in the beta-barrels do not alter the structural stability of the bean seed storage protein phaseolin. *Journal of Protein Chemistry*, November 1995, vol. 14, no. 8, p. 665-678.

DYER, J.M.; NELSON, J.W. and MURAI, N. Strategies for selecting mutation sites for methionine enhancement in the bean seed storage protein phaseolin. *Journal of Protein Chemistry*, October 1993, vol. 12, no. 5, p. 545-560.

EALING, P.M.; HANCOCK, K.R. and WHITE, D.W. Expression of the pea albumin 1 gene in transgenic white clover and tobacco. *Transgenic Research*, November 1994, vol. 3, no. 6, p. 344-354.

FALCO, Saverio C.; KEELER, Sharon J. and RICE, Janet A. Synthetic storage proteins with defined structure containing programmable levels of essential amino acids for improvement of the nutritional value of plants. United States Patent 5,559,223, (1996).

FONTANA, Angelo; ZAMBONIN, Marcello; POLVERINO DE LAURETO, Patrizia; DE FILIPPIS, Vincenzo; CLEMENTI, Angelica and SCARAMELLA, Elena. Probing the conformational state of apomyoglobin by limited proteolysis. *Journal of Molecular Biology*, February 1997, vol. 266, no. 2, p. 223-230.

FORANO, E. and FLINT, H.J. Genetically modified organisms: consequences for ruminant health and nutrition. *Annales de Zootechnie*, August 2000, vol. 49, no. 4, p. 255-271.

GAGNON, Mylène Claude; WILLIAMS, Martin; DOUCET, Alain and BEAUREGARD, Marc. Replacement of tyr62 by trp in the designer protein milk bundle-1 results in significant improvement of conformational stability. *FEBS Letters*, November 2000, vol. 484, no. 2, p. 144-148.

GALILI, Gad and HOFGEN, Rainer. Metabolic engineering of amino acids and storage proteins in plants. *Metabolic Engineering*, January 2002, vol. 4, no. 1, p. 3-11.

GEISOW, M.J. and AITKEN, A. Gas- or pulsed liquid-phase sequence analysis. In: FINDLAY, J.B.C. and GEISOW, M.J. eds. *Protein Sequencing A Practical Approach*. IRL Press, Oxford, 1989, p. 85-98.

GOLDBERG, A.L.; KOWIT, J.; ELTINGER, J. and KLEMES, Y. Selective degradation of abnormal proteins in animal and bacterial cells. In: SEGAL, H.L. and DOYLE, D.L. eds. Protein Turnover and lysozyme Function. *Academic Press*, New York, 1978, p. 171-196.

GRUNDY, Jean E.; WIRTANEN, Liz Y. and BEAUREGARD, Marc. Addition of a poly-(6X) His tag to Milk Bundle-1 and purification using immobilized metal-affinity chromatography. *Protein Expression and Purification*, June 1998, vol. 13, no. 1, p. 61-66.

GUERCHE, P.; DE ALMEIDA, E.R.; SCHWARZTEIN, M.A.; GANDER, E.; KREBBERS, E. and PELLETIER, G. Expression of the 2S albumin from Bertholletia excelsa in Brassica napus. *Molecular and General Genetics*, May 1990, vol. 221, no. 3, p. 306-314.

HENNIG, L. WinGene/WinPep: User-friendly software for the analysis of amino acid sequences. *BioTechniques*, June 1999, vol. 26, no. 6, p. 1170-1172.

HEO, Joo-Hyung; WON, Hye Soon; KANG, Hyun A.H.; RHEE, Sang-Ki and CHUNG, Bong Hyun. Purification of recombinant human epidermal growth factor secreted from the methylotrophic yeast Hansenula polymorpha. *Protein Expression and Purification*, February 2002, vol. 24, no. 1, p. 117-122.

HILL, R. Blake; RALEIGH, Daniel P.; LOMBARDI, Angela and DEGRADO, William F. De novo design of helical bundles as models for understanding protein folding and function. *Accounts of Chemical Research*, November 2000, vol. 33, no. 11, p. 745-754.

HOLLEMEYER, Klaus; HEINZLE, Elmar and THOLEY, Andreas. Identification of oxidized methionine residues in peptides containing two methionine residues by derivatization and matrix-assisted laser desorption/ionization mass spectrometry. *Proteomics*, November 2002, vol. 2, no. 11, p. 1524-1531.

JAYNES, J.; LANGRIDGE, P.; ANDERSON, K.; BOND, C.; SANDS, D.; NEWMAN, C. and NEWMAN, R. Construction and expression of synthetic DNA fragments coding for polypeptides with elevated levels of essential amino acids. *Applied Microbiology and Biotechnology*, July 1985, vol. 21, no. 5, p. 200-205.

JIN, W., KAMBARA, O., SASAKAWA, H., TAMURA, A. and TAKADA, S. De novo design of foldable proteins with smooth folding funnel: Automated negative design and experimental verification. *Structure* (Camb.), May 2003, vol. 11, no. 5, p. 581-590.

KIRIHARA, J.A.; HUNSPERGER, J.P.; MAHONEY, W.C. and MESSING, J.W. Differential expression of a gene for a methionine-rich storage protein in maize.

Molecular and General Genetics, March 1988, vol. 211, no. 3, p. 477-484.

LIAO, H.H. Thermostable mutants of kanamycin nucleotidyl-transferase are also more stable to proteinase-K, urea, detergents, and water-miscible organic solvents. *Enzyme and Microbial Technology*, April 1993, vol. 115, no. 4, p. 286-292.

MACCALLUM, Jillian D.; HEFFORD, Mary A.; OMAR, Semir and BEAUREGARD, Marc. Prediction of folding and degradation of the de novo designed protein MB-1 in cow rumen. *Applied Biochemistry Biotechnology*, April 1997, vol. 66, no. 1, p. 83-89.

MORROW, Julie A.; HATTERS, Danny M.; LU, Bin; HOCHTL, Peter; OBERG, Keith A.; RUPP, Bernhard and WEISGRABER, Karl H. Apolipoprotein E4 forms a molten globule. A potential basis for its association with disease. *Journal of Biological Chemistry*, December 2002, vol. 277, no. 52, p. 50380-50385.

PACE, C.N.; SHIRLEY, B.A. and THOMSON, J.A. Measuring the conformational stability of a protein. In: CREIGHTON, T.E. ed. *Protein Structure - A Practical Approach*. IRL Press, Oxford, 1989, p. 311-330.

PIPE, Steven W. and KAUFMAN, Randal J. Characterization of a genetically engineered inactivation-resistant coagulation factor VIIIa. *Proceedings of the National Academy of Science U.S.A.*, October 1997, vol. 94, no. 22, p. 11851-11856.

POLVERINO DE LAURETO, P.; TOMA, S.; TONON, G. and FONTANA, A. Probing the structure of human growth hormone by limited proteolysis. *International Journal of Peptide and Protein Research*, February 1995, vol. 45, no. 2, p. 200-208.

SAALBACH, I.; PICKARDT, T.; MACHEMEHL, F.; SAALBACH, G.; SCHIEDER, O. and MUNTZ, K. A chimeric gene encoding the methionine-rich 2S albumin of the Brazil nut (Bertholletia excelsa H.B.K.) is stably expressed and inherited in transgenic grain legumes. *Molecular and General Genetics*, January 1994, vol. 242, no. 2, p. 226-236.

SANCHEZ-RUIZ, J.M. and MAKHATADZE, G.I. To charge or not to charge? *Trends in Biotechnology*, April 2001, vol. 19, no. 4, p. 132-135.

STERN, M.D.; SANTOS, K.A. and SATTER, L.D. Protein degradation in rumen and amino acid absorption in small intestine of lactating dairy cattle fed heat-treated whole soybeans. *Journal of Dairy Science*, January 1985, vol. 68, no. 1, p. 45-56.

TEUBER, S.S.; DANDEKAR, A.M.; PETERSON, W.R. and SELLERS, C.L. Cloning and sequencing of a gene encoding a 2S albumin seed storage protein precursor from English walnut (Juglans regia), a major food allergen. *Journal of Allergy and Clinical Immunology*, June 1998, vol. 101, no. 6, part 1, p. 807-814.

WILLIAMSON, M. Can the risks from transgenic crop plants be estimated? *Trends in Biotechnology*, December 1996, vol. 14, no. 12, p. 449-450.

WYSS, M.; PASAMONTES, L.; FRIEDLEIN, A.; REMY, R.; TESSIER, M.; KRONENBERGER, A.; MIDDENDORF, A.; LEHMANN, M.; SCHNOEBELEN, L.; OTHLISBERGER, U.; KUSZNIR, E.; WAHL, G.; MULLER, F.; LAHM, H.W.; VOGEL, K. and VAN LOON, A.P. Biophysical characterization of fungal phytases (myo-inositol hexakisphosphate phosphohydrolases): molecular size, glycosylation pattern, and engineering of proteolytic resistance. *Applied and Environmental Microbiology*, February 1999, vol. 65, no. 2, p. 359-366.

YOO, J.H.; CHENG, O.H. and GERBER, G.E. Determination of the native form of FadD, the Escherichia coli fatty acyl-CoA synthetase, and characterization of limited proteolysis by outer membrane protease OmpT. *Biochemical Journal*, December 2001, vol. 360, part 3, p. 699-706.

Note: Electronic Journal of Biotechnology is not responsible if on-line references cited on manuscripts are not available any more after the date of publication.
Supported by UNESCO / MIRCEN network.