From DNA sequence to improved functionality: using protein sequence comparisons to rapidly design a thermostable consensus phytase

Martin Lehmann1, Dirk Kostrewa, Markus Wyss, Roland Brugger, Allan D’Arcy, Luis Pasamontes and Adolphus P.G.M. van Loon

F.Hoffmann-La Roche Ltd, Grenzacherstrasse 124, CH-4070 Basel, Switzerland

1To whom correspondence should be addressed; email: Martin.Lehmann@Roche.com

Naturally-occurring phytases having the required level of thermostability for application in animal feeding have not been found in nature thus far. We decided to de novo construct consensus phytases using primary protein sequence comparisons. A consensus enzyme based on 13 fungal phytase sequences had normal catalytic properties, but showed an unexpected 15–22°C increase in unfolding temperature compared with each of its parents. As a first step towards understanding the molecular basis of increased heat resistance, the crystal structure of consensus phytase was determined and compared with that of Aspergillus niger phytase. Aspergillus niger phytase unfolds at much lower temperatures. In most cases, consensus residues were indeed expected, based on comparisons of both three-dimensional structures, to contribute more to phytase stabilization than non-consensus amino acids. For some consensus amino acids, predicted by structural comparisons to destabilize the protein, mutational analysis was performed. Interestingly, these consensus residues in fact increased the unfolding temperature of the consensus phytase. In summary, for fungal phytases apparently an unexpected direct link between protein sequence conservation and protein stability exists.

Keywords: animal feed pelleting/consensus protein design/increased heat stability/phytase family/three-dimensional structure

Introduction

Phytases (EC 3.1.3.8) belong to the family of histidine acid phosphatases (Piddington et al., 1993; Mitchell et al., 1997) and are found primarily in microorganisms and plants. These enzymes catalyze the release of phosphate from phytic acid (myo-inositol hexakisphosphate), the major phosphorus storage form in plants. Since monogastric animals like poultry and pigs have very low or no phytase activities in their digestive tracts, inorganic phosphorus has to be added to the feed to meet the phosphorus requirements of these animals. Supplementation with inorganic phosphate, however, imposes ecological problems (eutrophication). This makes substitution of inorganic phosphate by phytase desirable, since it allows the use of the phytic acid already available in the feed. Because poultry and pig feed commonly is pelleted, a commercially attractive phytase should be able to withstand the temperatures that are reached temporarily during the pelleting process (60–90°C). Since all known phytases unfold at temperatures between 56 and 63°C (see below), we were interested in developing a rapid procedure to increase the unfolding temperature and thus the thermostability of phytases.

Increasing the thermostability of an enzyme usually requires combining multiple amino acid exchanges, each of which slightly increases the unfolding temperature of the protein. The main problem, however, is the identification of the relevant amino acid residues. In general terms mutations that increase thermostability may, for example, result in formation of hydrogen bonds, salt or disulphide bridges, increase the hydrophobic packing or the α-helix or β-sheet stability or stabilize β-turns or flexible termini or loops (for review see Jaenicke et al., 1996). Despite extensive knowledge of the general mechanisms governing protein stability (Dill et al., 1989; Dill, 1990; Fersht and Serrano, 1993; Matthews, 1993; Cordes et al., 1996) no rapid and reliable procedures are available for increasing the thermostability of a given protein. For successful thermostability engineering, many mutations have to be generated and analysed individually, followed by combination of the few stabilizing amino acid exchanges.

Three different approaches are commonly used to increase the unfolding temperature of a given protein: (i) random mutagenesis followed by selection; (ii) comparison of the amino acid sequence of the protein of choice with that of a thermostable and highly homologous counterpart, followed by replacement of selected amino acids by site-directed mutagenesis and (iii) introduction of possibly stabilizing amino acids based on the three-dimensional (3D) structure of the protein of interest. Each of these approaches has been used with success. Using approach (ii), for example, more stable variants of proteases and RNases were constructed (Imanaka et al., 1986; Serrano et al., 1993). Van den Burg et al. (1998) increased the thermostability of a thermolysin-like protease from Bacillus steaerotherophilus by 21°C by substitution of only eight amino acid residues. Six substitutions were based on a sequence comparison with its more stable counterpart thermolysin, while two mutations were selected by applying general rules for increasing thermostability. The resulting enzyme resisted boiling, while maintaining its proteolytic activity at 37°C.

Steipe et al. (1994) used another approach—they predicted stabilizing substitutions based on a comparison of the frequency of occurrence of amino acids in sequences of the immunoglobulin variable Vλ domain. Their work was based on the hypotheses that the immunoglobulin repertoire approximates a canonical ensemble of sequences, each derived from one of a set of germ-line sequences and selected in a process of random, independent mutations, and that destabilizing random mutations are highly probable, but are selectively neutral as long as the overall domain stability does not fall below a certain threshold. Conversely, random mutations resulting in increased thermostability are highly improbable in the absence of a positive selection. Consequently, they assumed that the most frequent amino acid at any position in an alignment of
The temperature optima for enzyme activity and the unfolding temperatures ($T_m$) determined by differential scanning calorimetry are shown. Phytases of other isolates of A. niger and A. fumigatus showed similar thermostability to those listed here (data not shown).

ND, not determined.

homologous immunoglobulin variable $V_c$ domains contributes most to the stability of the protein domain. They calculated a statistical ‘free energy’ from the frequencies of occurrence and predicted 10 stabilizing mutations for the immunoglobulin variable $V_c$ domain, of which six were indeed stabilizing.

We reasoned that it could be possible to obtain a thermostable phytase by construction of a consensus protein, containing at each position of the protein the amino acid occurring most frequently at that position in the fungal phytase family.

To date the sequences of 13 phytases from six different fungal species have been published (Piddington et al., 1993; Van Hartingsveld et al., 1993; Mitchell et al., 1997; Pasamontes et al., 1997a,b; van Loon et al., 1999). Despite the fact that several of the organisms used for their isolation were fact that several of the organisms used for their isolation were

Table I. Thermostability of phytases

<table>
<thead>
<tr>
<th>Phytase</th>
<th>Temperature optimum (°C)</th>
<th>$T_m$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consensus</td>
<td>71</td>
<td>78.0</td>
</tr>
<tr>
<td>Consensus Q27L</td>
<td>67</td>
<td>74.7</td>
</tr>
<tr>
<td>Consensus L276N</td>
<td>68.5</td>
<td>75.8</td>
</tr>
<tr>
<td>Consensus F298Y</td>
<td>70</td>
<td>ND</td>
</tr>
<tr>
<td>Consensus A384L</td>
<td>70</td>
<td>ND</td>
</tr>
<tr>
<td>A. niger NRRL3135</td>
<td>55</td>
<td>63.3</td>
</tr>
<tr>
<td>A. fumigatus ATCC 13073</td>
<td>55</td>
<td>62.5</td>
</tr>
<tr>
<td>A. fumigatus ATCC 13073$^\text{ thermo}$</td>
<td>60</td>
<td>67.0</td>
</tr>
<tr>
<td>A. terreus 9A1</td>
<td>49</td>
<td>57.5</td>
</tr>
<tr>
<td>A. terreus CBS</td>
<td>45</td>
<td>58.5</td>
</tr>
<tr>
<td>E. nidulans</td>
<td>45</td>
<td>55.7</td>
</tr>
<tr>
<td>M. thermophila</td>
<td>55</td>
<td>ND</td>
</tr>
<tr>
<td>T. thermophilus</td>
<td>45</td>
<td>ND</td>
</tr>
</tbody>
</table>

The temperature optima for enzyme activity and the unfolding temperatures ($T_m$) determined by differential scanning calorimetry are shown. Phytases of other isolates of A. niger and A. fumigatus showed similar thermostability to those listed here (data not shown).

ND, not determined.

the temperature optima, $T_m$'s and specific activities of phytases are affected by the host strain used for phytase production.

Consensus phytase (gene) construction and protein production

Calculation of the fungal phytase consensus sequence was performed as follows: all 13 known fungal phytase amino acid sequences were aligned using the program PILEUP from the GCG sequence analysis software package, release 9.0 (Devereux et al., 1984). The location of the gaps was refined using a text editor. Using this optimized alignment as input, the fungal phytase consensus sequence was calculated by the program PRETTY using the following conditions: plurality was set at 2.0, threshold at 2 and the scoring matrix used was prettypep.cmp which is also part of the GCG package. The first 26 amino acid residues of consensus phytase, which contain the signal sequence required for secretion, were replaced by the signal sequence of A. terreus CBS phytase (van Loon et al., 1999; Wyss et al., 1999a). The DNA sequence for the signal sequence was calculated using the approach of Purvis et al. (1987) and optimized for expression in S. cerevisiae. For the remainder of the protein, the codon frequency table of highly expressed S. cerevisiae genes (GCG package, release 9.0; Devereux et al., 1984) was used to design the DNA sequence.

The synthetic phytase consensus gene (fcp) was generated by dividing the calculated DNA sequence into oligonucleotides of 85 bp each (purchased from Microsynth, Balgach, Switzerland). Every oligonucleotide overlapped 20 bp with the previous oligonucleotide of the opposite strand. The PAGE-purified oligonucleotides were assembled to the whole fcp gene by three PCR reactions using the method of extensive overlap extension by PCR (Stemmer et al., 1995). Point mutations were introduced into the fcp gene using the pCR-Script™ Amp SK(+) Cloning Kit from Stratagene (La Jolla, CA). All other DNA manipulations were performed according to standard procedures (Sambrook et al., 1989).

The fcp gene was cloned into the EcoRI sites of the expression vector pFP (Gellissen et al., 1991) downstream of the FMD promoter (Hollenberg and Janowicz, 1989). The resulting plasmids were transformed into H. polymorpha strain RB11 (ura·). Transformation of H. polymorpha and expression of the consensus phytase gene were performed as described (Gellissen et al., 1996).

DNA sequencing of double-stranded DNA was done by the dideoxy method (Sanger et al., 1977) using the ABI PRISM Dye Terminator Cycle Sequencing Kit and the ABI PRISM 310 genetic analyzer as recommended by Applied Biosystems (Warrington, UK).

Purification and analysis of phytases

For protein purification the fermentation broths were centrifuged and the supernatant concentrated by ultrafiltration in Amicon 8400 cells using PM30 membranes. The concentrate (10 ml) was desalted on a 40 ml Sephadex G-25 Superfine column (Pharmacia Biotech, Uppsala, Sweden) with 10 mM sodium acetate, pH 5.0 as elution buffer. After addition of ammonium sulfate to a final concentration of 2 M, the sample was directly loaded onto a 1 ml Butyl Sepharose 4 Fast Flow hydrophobic interaction chromatography column (Pharmacia Biotech). Elution was done with a linear gradient from 2 to 0 M ammonium sulfate in 10 mM sodium acetate, pH 5.0. Phytase eluting at the beginning of the gradient was concentrated and then loaded on a 120 ml Sephacryl S-300 gel permeation chromatography
Designing a thermostable phytase

Phosphorylation was determined by the method of Brugger et al. (1999). One unit of enzyme activity was defined as the amount of enzyme that releases 1 µmol inorganic phosphate per minute at 37°C. Protein concentrations were not determined, using the extinction coefficient of 280 nm (for consensus phytase, 1.101) (Pace et al., 1995). For determination of the temperature optimum, enzyme and substrate solutions (100 µl each) were separately pre-incubated for 5 min at a range of predefined temperatures. Fifteen min after mixing both reagents, the reaction was stopped with trichloroacetic acid and the amount of phosphate released was determined by standard procedures (Wyss et al., 1999b).

To determine the unfolding temperature of phytases, differential scanning calorimetry was used as described (Brugger et al., 1999). The samples contained 50–60 mg/ml phytase (produced by recombinant H. polymorpha or A. niger strains and purified to homogeneity, see above) in 10 mM sodium acetate, pH 5.0. A constant heating rate of 10°C/min was applied up to 90°C.

Crystal structure determination

Crystals of consensus phytase were grown by vapour diffusion using the hanging drop technique. In a typical experiment 3 µl of the protein at 12 mg/ml was mixed with an equal volume of reservoir solution containing 12.5% PEG 3350 and 100 mM Li2SO4 in 50 mM Bis/Tris pH 5.5. The crystals were transferred to a cryoprotectant solution of 15% PEG 3350, 100 mM Li2SO4 and 30% glycerol in 50 mM Bis/Tris pH 5.5 prior to flash freezing in liquid nitrogen. Data were collected from a crystal, cryo-cooled at a temperature of 120 K using the Oxford Cryostream, on a MAR imaging plate with 18 cm diameter at the Swiss-Norwegian Beamline (SNBL) of the ESRF synchrotron facilities in Grenoble. Oscillation pictures were taken every 1°. The data set was processed with the programs DENZO/XDISP/SCALEPACK (Otwinowski and Minor, 1997). The space group is P212121 (No. 19) with cell dimensions a = 77.1 Å, b = 77.5 Å, c = 79.2 Å and one molecule in the asymmetric unit. The overall resolution range is 30.0–2.9 Å, with 10 817 unique reflections. The overall data (outer resolution shell 3.0–2.9 Å) statistics are as follows: redundancy, 3.1 (3.2); Rsym, 0.08 (0.37); completeness, 99.5% (99.6%) and I/σ(I), 14.1 (3.4). The structure was solved by molecular replacement with the CCP4 (Collaborative Computational Project Number 4, 1994) version of the program AMORE using the coordinates of A. niger phytase (Kostrewa et al., 1997), PDB (Bernstein et al., 1977) entry code 1HP, as a search model. A unique solution could be found with an initial R-factor of 0.43 and a correlation coefficient of 0.56 using data between 8.0–3.0 Å. Model building was done with the computer graphics program MOLOC (Gerber and Muller, 1995). Refinement was done against all data in the resolution range 30.0–2.9 Å with the program CNS 0.3c (β-test version, kindly provided by A.T. Brueger) using a maximum-likelihood target function (Pannu and Read, 1996) and conventional minimization with ideal stereochemistry parameters from Engh and Huber (1991). Iterative rounds of model building and refinement converged to a final model with an R-factor of 0.175, a free-R-factor (Brueger, 1992) of 0.250 and r.m.s. deviations from ideal geometry of 0.007 Å for bond lengths and 1.3° for bond angles. The crystal structure has been checked with PROCHECK (Laskowski et al., 1993) and WHAT_CHECK (Hooft et al., 1996).

Results

Design of a consensus protein

The sequence of the consensus protein was calculated using 13 phytase sequences obtained from six fungal species (see Introduction). The vote weight for each species was kept identical. For three of these species, phytase sequences from several different isolates were available [A. fumigatus, five sequences (Pasamontes et al., 1997b; van Loon et al., 1999); A. niger, three sequences (Piddington et al., 1993; Van Hartingsveldt et al., 1993; Mitchell et al., 1997) and A. terreus, two sequences (Mitchell et al., 1997; van Loon et al., 1999)], each of which was included in the calculation. Consequently, the vote weight per sequence was reduced to 0.2 (for each A. fumigatus phytase), 0.33 (for each A. niger phytase) and 0.5 (for each A. terreus phytase). In between the six species the phytases were 48.6–69.0% identical and showed similarities of 53.7–73.5%. For two species, isolates from different geography locations surprisingly showed up to 15% amino acid variation in phytase sequence (van Loon et al., 1999). Calculation of the consensus protein sequence was done using the PRETTY program (Devereux et al., 1984), resulting in the selection of the most frequently occurring amino acid for most but not all positions. At 18 positions adaptations from the consensus concept were required. For 10 of these positions, the program was unable to calculate a consensus amino acid. At eight ambiguous positions, the amino acid found in A. niger or A. fumigatus phytase was arbitrarily chosen instead of the amino acid proposed by the PRETTY program (Figure 1). The resulting consensus protein showed between 58.3 and 80.0% identity and between 64.5 and 83.9% similarity to each of its parents.

The consensus phytase displays drastically increased thermostability

The parent phytases had temperature optima between 45 and 55°C (Table I; Wyss et al., 1998; Brugger et al., 1999). In contrast to the purified consensus phytase had a temperature optimum of 71°C (Figure 2a; Table I). Thus, an increase in temperature optimum of 16–26°C was obtained in a single step. Using differential scanning calorimetry, an unfolding temperature (Tm) of 78.0°C was measured for the consensus phytase, which is 15–22°C higher than the Tm’s measured for the parent phytases (Figure 2b; Table I).

The consensus phytase has catalytic characteristics comparable with those of the A. emersonella and E. nidulans phytases

The substrate specificity of consensus phytase was most similar to those of the A. fumigatus and E. nidulans phytases (Figure 3a). For consensus phytase the highest specific activity was seen with p-nitrophenyl phosphate as substrate (78 U/mg), followed by fructose-1,6-bisphosphate (68 U/mg). The specific activities for phytate, glucose-6-phosphate and glycerol-2-phosphate were between 30 and 36 U/mg protein. With phytic acid as substrate, the A. fumigatus and E. nidulans phytases also showed specific activities of 25–30 U/mg protein, respectively (Wyss et al., 1999b). The pH-activity profile of consensus phytase with phytic acid as substrate was most similar to that of E. nidulans phytase. Both enzymes showed a pronounced

column (Pharmacia Biotech). Phytases eluted as homogeneous symmetrical peaks and were approximately 95% pure as shown by SDS–PAGE.
activity maximum around pH 6.5 (Figure 3b; Wyss et al., 1999b). The pH-activity profile and substrate specificity of consensus phytase differed most from those of the A.terreus and Mucorcheiophthora thermophila phytases (Mitchell et al., 1997; Wyss et al., 1999b). This correlated with the lower degree of homology between the consensus protein and the latter phytases (58.3% identity to M.thermophila phytase; 70.7% and 72.1% compared with the phytases of A.terreus strains 9A1 and CBS, respectively; 77.6% compared with E.nidulans phytase and between 78.2 and 80.0% compared with the phytases of the five A.fumigatus strains). Thus, although the total vote weight in the construction of the consensus protein was identical for each species, the subgroup consisting of A.fumigatus, E nidulans and to some extent A.niger phytases apparently dominated in its design.

For many thermophilic enzymes higher thermostability results from greater rigidity and thus decreased flexibility of the protein. As a result, proteins of thermophilic organisms frequently have decreased enzymatic activity at ambient temperatures compared with those of their mesophilic counterparts. Interestingly, consensus phytase, despite its drastically increased thermostability, had catalytic characteristics at 37°C comparable with those of two mesophilic phytases (Figure 3). Determination of the crystal structure of consensus phytase and comparison with the A.niger phytase structure

As a step towards understanding the increased stability of consensus phytase, its crystal structure was determined at 2.9 Å resolution. The overall fold of consensus phytase is shown in Figure 4a. Comparison of this structure with that of A.niger phytase (Wyss et al., 1999a) allowed analysis of the local interactions of each of the residues which differ between both phytases. Stabilization of a surface loop was observed in consensus phytase (Figure 4b). This surface loop is highly variable in the different phytases and has the consensus amino acid sequence (A.niger phytase residues in brackets) T249(S), S250(T), D251(V), A252(D), T253(T). T249 is the most conserved amino acid and is found in three of the six species, while two species have S249. This loop was not detectable in the electron density map of A.niger phytase, but is clearly seen in the electron density map of consensus phytase. Since for this part of the protein no contacts

Fig. 1. Alignment of phytases of six fungal species and of consensus phytase. Thirteen phytase sequences derived from six fungal species were aligned and for each position a consensus residue was calculated using the PRETTY program (see text). For 10 positions (*) no consensus residue was selected by the computer program under the conditions used. Either the most frequent or an arbitrarily chosen amino acid was introduced. At eight additional positions (Λ) the amino acid found in A.niger phytase (positions 135, 174, 204, 264 and 423), A.fumigatus phytase (positions 211 and 306) or both phytases (373) was used instead of the consensus residue calculated by the program. Phytase sequences shown are as follows: 1, A.terreus CBS; 2, A.niger NRRL3135; 3, A.fumigatus ATCC13073; 4, E.nidulans; 5, Talaromyces thermophilus; 6, M.thermophila; C, consensus phytase. The numbering starts with the N-terminal alanine of the mature A.niger phytase (Wyss et al., 1999a).
Designing a thermostable phytase

Fig. 2. Increased thermostability of purified consensus phytase. (a) Phytase activity was measured at the indicated temperatures for consensus phytase (□) and for the phytases of *E. nidulans* (○), *A. fumigatus* ATCC13073 (●) and *A. niger* NRRL3135 (▼). (b) Differential scanning calorimetry was used to determine the unfolding temperatures (Tm) for the indicated phytases.

with neighbouring consensus molecules in the crystal were present, stabilization results from intramolecular interactions. Based on the structural analysis, the most important single point mutation seems to be V251D. V251 is only found in *A. niger* phytases (three isolates), while all *A. terreus* (two isolates) and *A. fumigatus* phytases (five isolates) have D251. D251 stabilizes the loop by hydrogen bonds of D251:Oδ1 to T249:Oγ1 and to T253:Oγ1 (Figure 4b). Furthermore, the same aspartic acid residue forms a hydrogen bond from its Oδ2 atom to H6:Nδ1, thus also contributing to the stabilization of the N-terminus. The N-terminus was not visible in the electron density map of *A. niger* phytase, but is visible as a weak electron density in the map of consensus phytase. It is not clear how significant the stabilization by the D251:Oδ2–H6:Nδ1 hydrogen bond is, since in the crystal structure of consensus phytase, the N-terminus is further stabilized by two hydrogen bonds from the amino group of G3 (R) to a neighbouring phytase molecule in the crystal.

Effects of consensus residues on phytase stability: structural comparison and mutational analysis

Comparison of the crystal structures allowed clear predictions about the potential contributions to thermostability of only seven consensus residues; for all other consensus amino acids predictions using the 3D structures were ambiguous and thus could not be used to directly compare the consensus approach and the predictions based on the crystal structure comparisons. In those cases in which the use of both methods resulted in conflicting predictions, direct mutational analysis was performed (see below). Of the seven amino acids mentioned above, five (Y31, P225, M342, F346 and Y372) were expected to increase thermostability in consensus phytase, while two residues (F296, A384) should be destabilizing (see below). In one instance protein stabilization was apparently obtained by improving hydrophobic packing (compare Figure 4c with d). Three unambiguously predicted consensus amino acids are particularly interesting (*A. niger* phytase residues in brackets):
M342 (I) (M in four species and I and L in one species each), F346 (L) (F in four species and L in two species), and A384 (L) (A in four species and I and L in one species each). The first two consensus amino acids increase the size of the hydrophobic side chains, leading to tighter hydrophobic packing (Figure 4c and d), although F346 has an unfavourable $\chi_2$ torsion angle of $11^\circ$ ($\chi_2 = 70^\circ$). On the other hand, for position 384, the consensus amino acid would be expected to negatively affect temperature stability. A384 in consensus phytase creates an energetically unfavourable empty space in the hydrophobic core. In fact, computer homology modelling (MOLOC) showed that replacement of A384 in consensus phytase by L (L only found in A. niger phytase) should be possible with small local relaxation of the close contacts to F346, resulting in a better filling of a cavity in consensus phytase (Figure 4c and d) and should, thus, be energetically favourable. However, contrary to this, introduction of the A384L mutation into consensus phytase resulted in a 1°C decrease in thermostability (Table I). Thus in contrast to the prediction made based on the 3D structure, consensus residue A384 positively contributed to thermostability.

For Y31 (Y in three species, F in three species including
Designing a thermostable phytase

Fig. 4. Crystal structure of consensus phytase and a comparison of local interactions of selected residues in consensus phytase and A.niger phytase. All panels were prepared with the programs MOLSCRIPT (Kraulis, 1991) and RASTER3D (Merritt and Bacon, 1997). Stereo pictures are shown. The refined crystal structure includes amino acids 3–444 (A.niger phytase nomenclature), one short carbohydrate chain of two N-acetylglucosamine (NAG) residues N-linked to N184, two additional NAG residues N-linked to N316 and N353, two sulphates and 137 water molecules. (a) Backbone structure of consensus phytase. The green dashed circle indicates the active site. The catalytically essential H59 is shown in a green ball-and-stick representation. N- and C-, amino- and carboxyl-termini. Yellow sticks, disulphide bridges. (b) Consensus phytase surface loop 249–253. The main chain peptide connections of the loop to the remainder of the protein are indicated by grey dashed arrows. Small orange dots represent hydrogen bonds. The side chain of D251 (V251 in A.niger phytase) forms three hydrogen bonds to T249, T253 and H6, thus stabilizing the loop and the N-terminus. Here, it is assumed that the side chain of D251 is protonated at pH 5.5. Note that the proposed H6 Nδ-H-bond has poor geometry. Structural comparison of the hydrophobic environment of three residues in (c) A.niger phytase and (d) consensus phytase. The amino acids are shown in a ball-and-stick representation embedded in transparent van der Waals’ surfaces. The larger hydrophobic side chains of M342 and F346 in consensus phytase (shown in green) result in tighter hydrophobic packing of the core compared with I342 and L346 in A.niger phytase. In contrast, the smaller hydrophobic side chain of A384 in consensus phytase (shown in red) creates a ‘hole’ in the hydrophobic core when compared with L384 in A.niger phytase. The $\chi^2$ of F346 is unusual (11°), but real; above the noise there was no signal for a different side chain conformation. The map was very clear and the $B$-factors of the ring atoms were low ($B < 20$ Å$^2$). Comparison of the local environment of (e) Y296 in A.niger phytase with (f) P296 in consensus phytase. Hydrogen bonds are represented by orange dots. The tyrosine hydroxyl group is involved in two hydrogen bonds to Q54 and D335. D335 makes salt bridges to R136. P296 in consensus phytase lacks a hydroxyl group and, consequently, cannot form hydrogen bonds. Thus, in consensus phytase, a structural re-arrangement is seen of the side chains of Q54 and D335, with a distorted salt bridge of the latter amino acid to R136.
A. niger; not a typical consensus residue) and Y372 of consensus phytase (Y in five species, F in A. niger) stabilization is expected due to the formation of additional hydrogen bonds (data not shown). P225 (P in five species, S in A. niger) is predicted to stabilize a loop in consensus phytase (not shown). Based on computer homology modelling replacement of F296 in consensus phytase (F in four species, Y in A. niger and W in A. terreus) by the non-consensus amino acid Y296 should result in formation of two additional hydrogen bonds with the conserved residues D335 and Q54 (Figure 4c and f). Consensus residue F296 not only cannot form these hydrogen bonds but should also disturb the remaining hydrogen bond network of the surrounding amino acids (compare Figure 4e with f). Thus, based on the 3D structure, a stabilizing effect was expected upon introduction of the non-consensus residue Y296 into consensus phytase. However, mutation F296Y, which removed a clear consensus residue, also resulted in a decrease of the temperature optimum by 1°C (Table I). Thus, this consensus residue, despite the unfavorable prediction based on the 3D structure comparisons, in reality contributed to the increased stability of consensus phytase. This indicates that selecting consensus residues for the design of a novel phytase is a surprisingly rapid approach to stabilizing the resulting phytase.

Discussion

Thermostability is a prerequisite for successful application of enzymes in animal feeding because all components of animal feed, including enzymes, are exposed to temperatures between 60 and 90°C in the standard process used for feed pelleting (see Introduction). All known phytases, however, unfold at temperatures between 56 and 63°C. Only phytases from A. fumigatus refold efficiently after heat denaturation, resulting in increased pelleting resistance (Pasamontes et al., 1997b; Wyss et al., 1998). Further increase of the pelleting resistance may be obtained by increasing the unfolding temperature of phytase.

Here we describe the development of a consensus phytase with drastically increased heat resistance. By comparison of fungal phytase sequences a consensus protein sequence was constructed. The resulting consensus phytase showed between 58.3 and 80.0% amino acid identity to each of its parents. This constructed. The resulting consensus phytase showed between 58.3 and 80.0% amino acid identity to each of its parents. Previously, a similar approach was successfully applied to other protein families, including enzymes, are exposed to temperatures between 60 and 90°C in the standard process used for feed pelleting (see Introduction). All known phytases, however, unfold at temperatures between 56 and 63°C. Only phytases from A. fumigatus refold efficiently after heat denaturation, resulting in increased pelleting resistance (Pasamontes et al., 1997b; Wyss et al., 1998). Further increase of the pelleting resistance may be obtained by increasing the unfolding temperature of phytase.

By comparison of fungal phytase sequences a consensus protein sequence was constructed. The resulting consensus phytase showed between 58.3 and 80.0% amino acid identity to each of its parents. This constructed consensus phytase exhibited normal catalytic properties, but showed a remarkable increase of 16–26°C in temperature optimum and 15–22°C in unfolding temperature compared with each of its parents. Previously, a similar approach was successfully used for the prediction of individual stabilizing mutations in a small protein domain, the immunoglobulin variable V_{\kappa} domain (Steipe et al., 1994). Recently, hyperstable immunoglobulin V_{\kappa} domains were constructed by consensus sequence engineering and combining stabilizing point mutations (Ohage and Steipe, 1999). Our data show that it is possible to extend the use of the consensus approach for design and stabilization of an entire functional enzyme.

Theoretically, the increased temperature stability observed for consensus phytase compared with other phytases could be the result of only a single amino acid exchange which 'accidentally' happens to be present among the novel residues in consensus phytase. Precedence for this exists: an increase of 26°C in apparent melting temperature upon exchanging a single amino acid in Leishmania triosephosphate isomerase has recently been reported (Williams et al., 1999). However, for consensus phytase several lines of evidence indicate that its drastically increased thermostability may not result from exchanging a single amino acid. Despite extensive analysis by site-directed mutagenesis, no evidence exists currently to show that any amino acid exchange changes the thermosteractivity by more than a few degrees Celsius (see below). Firstly, two mutations in consensus phytase (F296Y, A384L), where consensus residues were replaced by non-consensus amino acids present in A. niger phytase, each decreased the temperature optimum of the enzyme by 1°C. Secondly, two other mutants of consensus phytase in which consensus residues Q27 (Q in five species, L in A. terreus phytase) and L276 (L in five species, V in M. thermophila phytase) were replaced by the respective non-consensus residue both showed a 2–3°C decrease in T_m. Meanwhile, more than 30 residues of consensus phytase were tested individually by site-directed mutagenesis (Lehmann, M. et al., manuscript in preparation). These substitutions affected thermostability usually by less than 1°C per mutation, or were neutral. Thirdly, six clear consensus residues (F31Y, V76I, F90Y, A220L, S242F and N271D) were simultaneously introduced into A. fumigatus phytase. This resulted in a 4.5°C increase in T_m (A. fumigatus phytase 1307°C; Table I). However, despite the extensive mutational analysis performed to date (see above), we cannot exclude the possibility that the increased thermostability of consensus phytase may result from the fortuitous combination of only a few of the many mutations simultaneously introduced into consensus phytase. These mutations could, for example, result in the stabilization of a loop usually triggering the unfolding process of the entire protein, thereby delaying the entire unfolding process. Our 3D structure comparisons showed that a surface loop and the N-terminus, both of which were invisible in the structure of the A. niger phytase, were clearly seen in the consensus phytase structure. These structures are thus prime candidates for potential dominant weak spots in the molecule, which could trigger protein unfolding of A. niger phytase. The importance of stabilization of these parts of the protein in consensus phytase on protein unfolding is, however, currently unclear. Further and more extensive mutational analysis is required, for example, on comparison of the 3D structures of consensus phytase and other phytases will be required to eventually exclude this possibility.

In summary, our approach allowed the rapid design of a fully functional phytase with drastically increased thermostability based only on a comparison of primary protein sequences. Should this approach also be applicable to other protein families, as can be expected, then a more productive use of the wealth of DNA sequence information which is currently available would be possible.

Acknowledgements

A. Kronenberger, C. Loch, L. Schooebelen, M. Tessier, R. Remy, D. Studer and D. Burger are gratefully acknowledged for excellent technical assistance. F. Mascarello and S. Augem for performing the differential scanning calorimetry measurements and Drs Fritz Winkler and Michael Hennig for valuable discussion and critical evaluation of the thermostability predictions using the 3D structure.

References


Designing a thermostable phytase


Received May 28, 1999; revised October 8, 1999; accepted October 27, 1999