The folding and activity of the extracellular lipase of *Rhizopus oryzae* are modulated by a prosequence

Hans-Dietmar BEER*§, Gerd WOHLFAHRT*, Rolf D. SCHMID† and John E. G. MCCARTHY‡||

Departments of *Enzymology and ‡Gene Expression, National Biotechnology Research Centre (GBF), Mascheroder Weg 1, D-38124 Braunschweig, Federal Republic of Germany, and †Institute of Technical Biochemistry, Universität Stuttgart, Allmandring 31, D-70569 Stuttgart, Federal Republic of Germany

The fungus *Rhizopus oryzae* synthesizes an extracellular lipase precursor bearing N-terminal pre- and pro-sequences. Our studies in *Escherichia coli* and using recombinant lipase *in vitro* indicate that the prosequence of 97 amino acids has at least two functions. First, it modulates the enzyme activity of the lipase so that this enzyme can initially be synthesized in a non-destructive form. Direct synthesis of the mature form of the lipase in the cell has toxic consequences, at least partly because of phospholipase activity that is suppressed in the proprotein. Secondly, it supports

folding of the lipase via a pathway influenced by a single cysteine residue at position –68. Mutational analysis of the prosequence demonstrates not only the key role of this cysteine residue but also the importance of the neighbouring amino acids. In particular, Arg-69 probably enhances the leaving group character of Cys-68. We propose a model in which Cys-68 acts as an intramolecular thiodisulphide reagent, playing a catalytic role in the folding of the enzyme. The prosequence is capable of performing the described functions both in *cis* and in *trans*.

INTRODUCTION

A wide range of proteins are initially synthesized as preproteins bearing N-terminal regions that are cleaved off to yield mature, active products. Presequences of 20–30 amino acids are widely used to direct membrane translocation in prokaryotes and eukaryotes [1]. In contrast, a number of hydrolases, growth factors and hormones possess a further type of N-terminal extension, called a prosequence, which is generally also proteolytically removed, but can manifest diverse functions. For example, proregions play an important role in zymogens, which are inactive precursors whose cleavage results in a conformational change that generates an active enzyme. Many biologically important processes, such as the blood clotting cascade, are critically dependent on this type of conversion from precursor to active enzyme [2]. Hepatocyte growth factor/scatter factor is activated in a similar way by urokinase, but the processing step generates a disulphide-linked α , β -heterodimer without releasing a propeptide [3]. Propeptides can also contain signals that direct the protein into distinct cell compartments. C-terminal H/KDEL sequences prevent the secretion of luminal ER proteins [4], whereas certain N-terminal sequences can target proteins to the vacuole [5].

A more recently discovered function of prosequences relates to protein folding. Since the demonstration that protein folding can be catalysed by molecular chaperones [6], attention has turned to the potential roles of proregions in controlling this process. Investigations of a number of hydrolases have indicated that their prosequences not only inhibit enzyme activity but also support folding (reviewed in [7,8]). Activin A and transforming growth factor β 1 are members of a group of structurally related proteins that are involved in differentiation and have endocrine effects. This family comprises disulphide-linked homodimers derived from a larger precursor. The proregions of these precursors are essential *in vivo* for dimerization and secretion of biologically active homodimers, suggesting that they play a role in folding and the formation of disulphide bonds [9].

The only example in which the role of a prosequence in the folding of a precursor has been explored in any detail is that of pro-bovine pancreatic trypsin inhibitor (BPTI) [10]. The limiting factor in the folding pathway of this small protein is the formation of three disulphide bridges. The cysteine residue of the prosequence, or a single cysteine residue tethered to the C-terminal end of BPTI via a flexible linker, is sufficient to assist disulphide formation. These cysteine residues seem to facilitate folding by providing a solvent-accessible, intramolecular thiol-disulphide reagent. However, Creighton et al. [11] interpreted the small

Figure 1 Schematic representation of the *Rhizopus oryzae* lipase and the derivatives used in this work

The diagrams indicate the relative positions of the cysteine residues (thin vertical lines) and the mutations (asterisks) in the amino acid (aa) sequence of the lipase (ROL) and its proregion (PRO).

Abbreviations used: BPTI, bovine pancreatic trypsin inhibitor; GSH, glutathione; GSSG, glutathione disulphide; ProROL, *Rhizopus oryzae* prolipase; ROL, *Rhizopus oryzae* lipase.

§ Present address: AG Werner, MPI-Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, Federal Republic of Germany.

|| To whom correspondence should be addressed.

Figure 2 Structure of the polypeptide backbone of ROL

The lid region and cysteine residues are highlighted.

increase in the rate of the rearrangement events as a result of the prosequence as not significant under physiological conditions.

Proteolytic processing of a precursor yields equal amounts of two protein segments. In prokaryotic cells, in contrast, a 1:1 ratio of gene products can also be achieved by translational coupling of a bicistronic messenger RNA (see, for example, [12]). This principle seems to be the basis for generating equimolar amounts of a lipase and a downstream gene product in certain *Pseudomonas* species. The downstream gene is necessary for the production of an enzymically active lipase [13–15], suggesting the same physiological role for protein folding as some prosequences.

The gene encoding the extracellular lipase from *Rhizopus* sp. comprises a signal sequence of 26 amino acids, a prosequence of 97 amino acids, and the region of the mature lipase of 269 amino acids [16,17] (Figures 1 and 2). Here we demonstrate that the prosequence of the lipase from *Rhizopus oryzae* is able to facilitate folding both *in vitro* and in *Escherichia coli* cells. Our results indicate that this kinetic effect is linked to a cysteine residue in the prosequence that supports the formation of the three disulphide bonds in the active lipase, the limiting event in the folding pathway. Moreover, the prosequence modulates the enzyme activity so that the lipase can be overexpressed in *E. coli* without causing damage to the host, an effect comparable to that observed with other zymogens.

EXPERIMENTAL

Strains and media

E. coli strain BL21(DE3) was used for the expression experiments with plasmid pET11a [18], and BL321 [19] was used with pCYTEXP1 [20,21]. Strain TG1 [22] was used for plasmid propagation. *E. coli* was grown in Luria–Bertani medium [22]. PCR was performed with a GeneAmp-PCR Reagent Kit and a DNA Thermal Cycler, both from Perkin-Elmer Cetus.

Plasmid constructions and mutagenesis

All DNA subcloning steps and transformations of *E. coli* were performed with standard procedures [22]. The generation of N-terminal or C-terminal mutations was achieved by means of one-step PCR reactions. A two-step PCR procedure was used to generate point mutations within the reading frame [23]. After PCR the whole subcloned fragment was sequenced. Synthetic annealed oligonucleotides encoding the preOmpA sequence were introduced into the *Nde*I restriction site at the start codon of the reading frame of the chosen gene in pCYTEXP1. Bicistronic vectors were constructed by first excising the prosequence together with the preOmpA sequence by using *Xho*I and *Sal*I. The resulting fragments were then ligated into the *Sal*I site of the pCYTEXP1 vector containing preOmpAROL or preOmpAProROL.

Expression experiments

Freshly transformed cultures that had reached a D_{600} value of 0.5 were shifted from 30 to 39.4 °C. For strain BL21(DE3), protein expression was induced by the addition of 1 mM isopropyl β -D-thiogalactoside. Samples were taken at hourly intervals without removing shake-flasks from the water bath. After a short centrifugation, pellets were frozen in liquid nitrogen and stored at –25 °C until being assayed for lipase activity. The frozen pellets were resuspended in L-buffer [50 mM Tris/1mM EDTA (pH 8)] and sonicated in plastic tubes on ice for 30 s. The activities determined with the resulting extracts are expressed as lipase units per ml of the original culture.

Purification of recombinant lipase

Overexpression was performed with strain BL321 containing pCYTEXP1 carrying the desired gene plus the preOmpA sequence. Cultures reaching a D_{600} value of 0.5–1 were shifted from 30 to 42 °C for 3 h. The cells of two 250 ml cultures were resuspended in 16 ml of L-buffer and sonicated for 7 min (Branson Sonifier 250). After centrifugation for 20 min at 17 000 g (4 °C) the pellet was resuspended in 16 ml of buffer I [100 mM NaCl/ 50 mM Tris buffer/1 mM EDTA (pH 8)] and incubated for 20 min at 20 °C in the presence of DNase (30 μ g/ml). After centrifugation, the pellet was washed with buffer II [100 mM NaCl/50 mM Tris buffer/10 mM EDTA/0.5% Triton X-100 (pH 8)] and again with buffer I. The resulting pellet was resuspended in 16 ml of 5 M guanidinium hydrochloride/50 mM Tris buffer (pH 8.5)/30 mM dithiothreitol and incubated at room temperature for 30 min. After centrifugation for 30 min at 4 °C and 17 000 g, the supernatant was diluted in 1 litre of refolding buffer [50 mM Tris buffer (pH 8.5)/20 mM CaCl_2 /10% (v/v) glycerol/2 mM glutathione (GSH)/1 mM glutathione disulphide (GSSG)]. After 36 h at 10 °C the solution was dialysed against 10 mM CaCl_2 , centrifuged for 20 min at 4 °C and 10 000 g and then filtered (filter pore size 0.45 μ m). The solution was applied to a column of S-Sepharose (Pharmacia) and eluted with a linear NaCl gradient of 0–500 mM. Fractions containing lipase were pooled and, after dialysis against deionized water, freeze-dried in aliquots. The freeze-dried lipase was stored at –25 °C.

Folding experiments *in vitro*

Purified enzymes were denatured for 30 min in 7 M guanidinium hydrochloride/50 mM Tris buffer (pH 8.5)/30 mM dithiothreitol at 20 °C. The protein solution was then diluted 1:100 in degassed

200 mM Tris buffer, pH 8.5, containing the indicated concentrations of GSH and GSSG (final protein concentration 10 μ g/ml).

Characterization of the recombinant lipase preparations

The freeze-dried enzyme was dissolved in 20 mM Tris buffer, pH 8, and after centrifugation the supernatant was measured for protein content with the bicinchoninic acid kit (Sigma). The specific activities were measured by titrimetry [24]. Lipase was added to the emulsion prepared by mixing (Ultra-Turrax, IKA-Labortechnik) 14 g of triolein (Fluka) and 200 ml of 2% (w/v) gum arabic (Merck) and 0.5% $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ for 2 min. The release of free fatty acids was monitored continuously by means of titration with 10 mM NaOH on a pH-stat (Radiometer Copenhagen) at pH 8.3. Determination of phospholipase activity was performed with 1-palmitoyl-2-[1- ^{14}C]linoleoyl-L-3-phosphatidylethanolamine, (Amersham) in accordance with standard procedures [25–27].

RESULTS

Expression and translocation of enzymically active lipase in *E. coli*

The lipase from *R. oryzae* contains three disulphide bonds, one of which is formed between the N- and C-terminal cysteine residues (residues 29 and 268 in Figure 2) [17,28]. The fact that Cys-268 is adjacent to one of the component amino acids of the catalytic triad (His-257) means that this disulphide bond is essential for the stability and catalytic activity of the protein. The measured lipase activity, which can be determined with a sensitive and rapid assay [24], can therefore be taken as an indicator of the amount of correctly folded protein.

To examine whether translocation to the periplasm is a prerequisite for the expression of active disulphide-bonded lipase [29,30], we initially constructed four different plasmids (Fig. 1). The genes encoding mature lipase (ROL) and the prolipase (ProROL) were cloned with and without the preOmpA-sequence [31] into the expression vector pET11a [18]. Readily detectable (and similar) levels of expression of all of the constructs, except that encoding preOmpAROL, were observed in strain BL21(DE3). The fact that the preOmpA sequence was correctly removed indicates that ProROL can be translocated to the periplasm (Figure 3A and results not shown). Analysis of the soluble ROL activity of the different constructs after breaking open the respective cell types revealed marked differences. Cytoplasmic ROL was found to be insoluble and inactive (Figure 3B). Cytoplasmic ProROL showed similar behaviour, although some enzyme activity could be measured. Periplasmic ProROL was in part soluble and active (Figure 3B), in that the enzyme activity was found solely in the soluble fraction (results not shown). This experiment indicates that the folding of ProROL is in fact limited by the formation of disulphide bonds. ProROL can fold to its active conformation only in the oxidizing environment of the periplasmic space, which promotes disulphide bonding [29,30].

The insufficient repressibility of the pET-type vectors described above prevented us from achieving stable synthesis of the periplasmic form of ROL. Thus to compare the effects of translocation of ROL and ProROL into the periplasmic space, we cloned each of the genes together with the preOmpA signal sequence into the strongly regulated pCYTEXP1 vector [20,21]. Overexpression induced by a temperature shift to 42 °C resulted in the accumulation of the proteins as insoluble aggregates. Both fusion proteins were correctly processed, indicating that they

Figure 3 Expression of cytoplasmic ROL and ProROL constructs and translocation of ProROL into the periplasm

(A) SDS/PAGE of the cell extracts derived from cells carrying ROL and ProROL constructs transcribed from the *lac* promoter. Equal amounts of total cell protein from the respective extracts were loaded on to the gel. The positions of the recombinant ROL and ProROL proteins are indicated by arrowheads. Abbreviation: kD, kDa. (B) The enzyme activities of ROL (cytoplasmic) and ProROL (cytoplasmic and periplasmic) are expressed in lipase units (LU; μ mol of free fatty acids released/min) per ml of culture broth. The genes encoding ProROL and ROL were cloned into the vector pET11a with and without the OmpA signal sequence. Expression of the gene products was induced in the strain BL21(DE3) by the addition of isopropyl β -D-thiogalactoside. ROL could not be stably expressed in combination with the preOmpA sequence when transcribed from the *lac* promoter.

were actually translocated to the periplasmic space (results not shown). Reducing the induction temperature resulted in a decreased expression level from the heat-inducible lambda promoter. Simultaneously, the amount of ProROL that could fold correctly to form active enzyme increased, suggesting that the yield of active ProROL is a complex function of the expression level. Figure 4 shows the results of expression at 39.4 °C. Whereas ROL activity and attenuation of the preOmpA-ProROL strain increased in a similar way to that observed with the equivalent pET11a plasmid derivative, active ROL could be obtained only in small amounts from the preOmpA-ROL strain. The decreased yield of active ROL in comparison with ProROL might be explained either by a lower level of folding or by decreased thermal stability of correctly folded ROL or of a possible folding intermediate. A striking difference between the consequences of expression of ROL and ProROL is evident from the growth behaviour of the cells (Figure 4). Cells synthesizing ProROL grew further after induction, whereas the synthesis of ROL had

Figure 4 Translocation of ROL and ProROL into the periplasm

The genes encoding ROL and ProROL were N-terminally fused with the OmpA signal sequence and expressed by using the pCYTEXP1 vector. Expression in the strain BL321 was induced by shifting the temperature from 30 to 39.4 °C. The plots show values of D_{600} and lipase activity measured at time points subsequent to the initiation of induction. Abbreviation: LU, lipase units (μmol of free fatty acids released/min).

Figure 5 Growth and lipase activity of cells expressing mutant lipase genes

Expression of the constructs indicated was initiated once the cultures had reached a D_{600} value of 0.5. The D_{600} values and lipase activities determined 5 h later are compared here. They allow a rapid assessment of the relative activities and toxicities of the respective lipase forms. Abbreviation: LU, lipase units (μmol of free fatty acids released/min).

a severely toxic effect on *E. coli*, the D_{600} value decreasing after 2 h of induction.

The toxic effects of ROL expression

Whereas the expression of active ProROL has no toxic effects on *E. coli*, the expression of ROL causes cell lysis, despite the fact that the amount of soluble and active enzyme is smaller than that of ProROL by a factor of more than 100. To clarify the issue of whether the enzyme activity of ROL damages the cells, we

performed an expression experiment with inactive mutants of ROL and ProROL. Mutation of the aspartic acid residue belonging to the catalytic triad (Asp-203) to alanine decreases the specific activity by a factor of 2000 [17]. This mutation in ProROL was found to have no effect on the growth curve of cells during expression (Figure 5). In comparison, cells expressing inactive ROL grow even better than those expressing active or inactive ProROL. From this experiment it is obvious that the enzyme activity of ROL damages cells even at low concentrations, whereas the activity of ProROL has no such effect.

It has been demonstrated that prosequences can facilitate folding of the corresponding mature protein in *trans* both *in vivo* and *in vitro* [32–35]. We investigated the influence of *in trans* synthesis of the prosequence of ROL (Figure 5; ROL-Pro). Translated from the same mRNA, the prosequence abolished the toxic effect of expression of ROL alone and considerably enhanced the yield of active ROL. Thus the prosequence is likely to have an affinity for both the unfolded and the correctly folded forms of ROL. The effect of the facilitation of folding is likely to be due to an affinity for a folding intermediate, whereas suppression of the toxic effect of ROL is probably due to an affinity for correctly folded ROL. Co-expression of the prosequence together with ProROL evidently had little effect on this protein's properties (Figure 5, ProROL-Pro).

The effects of modifications in the prosequence

To examine which parts of the prosequence are necessary for obtaining active protein and suppressing the toxicity of ROL, we generated three different mutants of ProROL (Figure 1). *Rhizopus niveus* secretes a lipase that possesses the last 28 amino acids of the propeptide at its N-terminus and is converted to the smaller lipase form via limited proteolysis catalysed by a specific extracellular protease [36]. We therefore named the corresponding recombinant protein Niveus. In Pro68ROL the 29 N-terminal codons of the prosequence were deleted, so that the first amino acid after the initial methionine residue is Cys-68. In ProROL(Cys-68 → Ser), the only cysteine residue of the prosequence has been mutated to an isosteric serine residue. Whereas Niveus is correctly processed by leader peptidase, Pro68ROL accumulates in two forms in the periplasmic space. One third of the protein starts with a methionine residue followed by the cysteine, and two-thirds have been processed behind the cysteine residue (results not shown). Niveus is not as toxic for the cells as ROL; growth is inhibited but cells do not lyse. Cells expressing the ProROL(Cys-68 → Ser) construct grow like wild-type ProROL, and *E. coli* synthesizing Pro68ROL grows even better than when it synthesizes ProROL (Figure 6A). The mutation of the cysteine residue in the prosequence results in a decrease in enzyme activity to one-fifth. Niveus shows more activity *in vivo* than ROL but significantly less than ProROL(Cys-68 → Ser) or Pro68ROL (Figure 6B). This experiment suggests that the whole prosequence, and particularly the cysteine residue, are necessary to support folding. However, 28 C-terminal amino acids of the prosequence of Niveus are enough to modulate the enzyme activity so that the protein can be expressed without lysing the cells.

If the cysteine residue at position –68 in the prosequence is important for folding, it would be expected that its performance as a mediator of disulphide bridge formation would be influenced by its neighbours in the amino acid sequence. Examination of the sequence context of Cys-68 reveals the existence of a number of charged amino acids (Ala-Leu-Pro-Pro-Leu-Ile-Ser-Ser-Arg-Cys-Ala-ProPro-Ser-Asn-Lys-Gly-Ser-Lys). We constructed further mutants in which Arg-69, Lys-62 and Lys-59 were each

Figure 6 Expression of derivatives of ProROL

The plots show the growth (A) and lipase activity (B) of strains carrying expression vectors directing the synthesis of ROL and ProROL and their derivatives. The *x*-axes give the time after initiation of induction by temperature shift. Abbreviations: LU, lipase units (μmol of free fatty acids released/min).

Table 1 Refolding of different lipase forms

The refolding was assessed on the basis of measured lipase activity. The strains expressed the respective lipase derivatives to the same levels.

Construct	Refolding (% related to ProROL)	
	<i>In vivo</i>	<i>In vitro</i>
ProROL	100	100
ROL	0	19
Pro68ROL	18	21
ProROL(Cys-68 → Ser)	20	19
ProROL(Glu-265 → Gln)	50	53

converted to alanine residues. The Arg-69 → Ala mutant lipase was inactivated to a similar degree to the Cys-68 → Ser mutant (i.e. approx. 5-fold), whereas the other two mutant lipases showed decreases in activity compared with the wild-type enzyme of between 10 % and 30 % (Table 1). An important feature of the

Figure 7 Purification of recombinant lipase proteins from *E. coli*

SDS/PAGE demonstrating the purity of the respective enzymes: Lane 1, ROL; lane 2, Pro68ROL; lane 3, ProROL; lane 4, ProROL(Cys-68 → Ser); lane 5, ProROL(Glu-265 → Gln). The molecular masses (in kDa) of standards are indicated at the left.

mature lipase is the interaction between the N-terminal and C-terminal regions of the protein, including a disulphide bond between Cys-29 and Cys-268. The latter residue has two charged amino acid residues, Glu-265 and Asp-262, in its immediate vicinity. In the light of the effects of mutations of charged amino acids near Cys-68, and for comparative purposes, we created the derivatives Glu-265 → Gln and Asp-262 → Gly (Figure 1). These mutant ProROL derivatives showed activities equivalent to approx. 50 % and 70 % of the wild-type level (Table 1).

Characterization of lipase folding and activity *in vitro*

The experiments *in vivo* raise the question of how the lipase activity of ROL can be toxic to *E. coli* whereas the activity of ProROL is not. Another question is whether the observed differences in the yields of enzyme activity are due to folding effects or differences in thermal stability. To be able to address these issues we prepared recombinant lipase proteins from *E. coli* (Figure 7). The specific activities of ROL, ProROL, Pro68ROL and ProROL(Cys-68 → Ser) are all between 10000 and 13000 units/mg, suggesting that the catalytic site is in a fully functional conformation in all proteins. Renaturation was performed under conditions allowing complete refolding of the respective proteins (see also the Discussion section). The specific activity of ProROL is approx. 30 % higher than that of ROL, although the molar mass is approx. one-third higher. This is in agreement with earlier results [37]. Figure 8 shows the activities of ROL and ProROL as a function of the substrate concentration. Whereas there is no difference in the decrease of activity for ROL and ProROL at high substrate concentrations, the curve of ProROL drops sharply to zero at concentrations below the apparent K_m . This is in contrast with ROL, which possesses considerable enzymic activity at triolein concentrations between 0.1 and 1 mM. A second surprising difference is the fact that ROL was found to have a low level of phospholipase C activity, whereas ProROL does not. This might provide an explanation of why the expression of ROL results in cell lysis whereas ProROL can be expressed without damage to the host. ROL has a significantly broader and higher pH optimum in comparison to ProROL (results not shown). The pH dependence curve of Pro68ROL lies between those of ROL and ProROL at high pH. These results demonstrate that the enzymic properties of ROL and ProROL are significantly different.

Investigations of the thermal stability of the purified proteins (Figure 9) revealed that ProROL and its point-mutant derivatives have more or less the same inactivation temperature, which is considerably reduced for ROL. The stability of Pro68ROL more

Figure 8 Lipase activity of ROL and ProROL as a function of substrate concentration

The enzymic activity of purified ROL and ProROL was determined by using varying initial concentrations of substrate and a constant amount of emulsifier. The maximum variation in the measured values was 15%. Abbreviation: LU, lipase units (μmol of free fatty acids released/min).

Figure 9 Thermal stability of ROL, Pro68ROL, ProROL(Cys-68 → Ser), ProROL(Glu-265 → Gln) and ProROL

The purified proteins (25 $\mu\text{g}/\text{ml}$ in 20 mM Tris buffer, pH 8) were incubated for 30 min at the indicated temperatures and assayed for lipase activity. The plotted values are those of ROL (●), ProROL (○), Pro68ROL (▼), ProROL(Glu-265 → Gln) (▽) and ProROL(Cys-68 → Ser) (■).

closely resembles that of ProROL than that of ROL. Because ROL is not stable at the induction temperature, correctly folded and active protein can be expected to denature and aggregate. This explains why expression of ROL leads to the generation of only small amounts of active protein. On the other hand, the point-mutant forms of ProROL have the same thermal stability as ProROL itself, thus indicating that the different levels of enzyme activity might be due to folding effects. For comparison we also investigated a derivative of ProROL whose ability to fold was reduced by a mutation of one of the charged amino acids close to Cys-268 (Glu-265 → Gln).

To test the latter hypothesis, we performed folding experiments *in vitro* with purified proteins. At 12 °C, ROL and ProROL folded with similar kinetics and gave yields close to 100 %

Figure 10 Folding of ROL and ProROL *in vitro* at 12 °C

Purified and denatured ROL and ProROL were renatured in 200 mM Tris buffer, pH 8.5, supplemented with 2 mM GSH and 1 mM GSSG. The plots show the recovery of activity with time.

(Figure 10). The curve for ROL is more sigmoid at the beginning of the refolding than that of ProROL. This kinetic effect suggests that there is an inhibitory effect in the refolding of ROL at this temperature that is absent in the folding of ProROL. However, this experiment cannot explain why active ROL and ProROL and the other mutant forms accumulate at different levels *in vivo*. Indeed, overall the results obtained at 12 °C seem to be consistent with the thermodynamic hypothesis of protein folding, which states that the final three-dimensional structure of a protein is ultimately determined by its sequence [38].

However, folding experiments performed at 25 °C, which is more similar to the situation *in vivo*, leads to totally different results (Figure 11). The recovery of correctly folded protein is decreased for both ROL and ProROL, but to different degrees. The folding of ProROL is more or less independent of the redox potential determined by different concentrations of GSH and GSSG, leading to recoveries of about 50 %, except for folding in the absence of GSH/GSSG or in the presence of only small concentrations of GSSG (Figure 11A). In contrast, the refolding of ROL is strongly dependent on the redox potential. The recovery varies from 25 % to less than 10 %. Higher concentrations of GSH, which favour rearrangements, decrease the yield of correctly folded protein (Figure 11B).

The ratio of GSH to GSSG within the secretory pathway in eukaryotic cells ranges from 1:1 to 3:1 in the millimolar range [39]. The molecular components contributing to the redox buffer in the periplasm of *E. coli* are unknown, but it is likely that the redox state is dictated by external medium conditions [40]. Low concentrations of GSH, which in the case of ROL lead to better refolding yields *in vitro*, are not representative of the situation in the natural host. We therefore compared the refolding of the five recombinant proteins *in vitro* at a more physiological concentration of GSH and GSSG (2 mM; Figure 12). There are only three levels of enzyme activity, the highest being that of ProROL with about 50 %, followed by the activity of ProROL(Glu-265 → Gln) and a low activity of 10 % for ROL, Pro68ROL and ProROL(Cys-68 → Ser). The apparent degrees of refolding (relative to ProROL) achieved *in vitro* reflect those obtained *in vivo* (Table 1). The only exception is ROL, which is not stable at the induction temperature. Overall, these results indicate that the

Figure 11 Folding of ROL and ProROL at 25 °C and various concentrations of GSH and GSSG

Purified and denatured ProROL (A) and ROL (B) were renatured in 200 mM Tris buffer, pH 8.5, supplemented with the following concentrations of GSH and GSSG: no added GSSG/GSH (●); 0.1 mM GSSG (○); 1 mM GSSG (▼); 2 mM GSSG (▽); 10 mM GSSG (■); 2 mM GSSG, 2 mM GSH (□); 2 mM GSSG, 0.1 mM GSH (◆); 2 mM GSSG, 10 mM GSH (◇).

Figure 12 Folding of ROL (●), Pro68ROL (▼), ProROL(Cys-68 → Ser) (▲), ProROL(Glu-265 → Gln) (◆) and ProROL (■) at 25 °C and 2 mM GSH and GSSG

cysteine residue of the prosequence facilitates the folding of ProROL.

DISCUSSION

In the present work we demonstrate that the prosequence of the extracellular lipase from *R. oryzae* has at least two functions. It modulates the enzyme activity of the mature lipase so that it can be synthesized without damaging the host, and it supports folding via a pathway that involves a single cysteine residue. In the sense that the prosequence protects the host from damaging properties of the mature enzyme, ProROL seems to have a function analogous to that of other zymogens. The proteolytic activation of zymogens generally produces only very limited changes in the overall folding of the corresponding protein. In phospholipase A, the α -amino group of Ala-1, which appears on activation, is essential for enzyme activity [41]. For other proteolytic enzymes it was shown that their precursors, initially thought to be inert molecules, have inherent proteolytic activity [42]. In pig pepsinogen, inhibition of the enzyme is caused by the formation of salt bridges that stabilize the positioning of the N-terminal proenzyme segment across the active site. At low pH these salt bridges are disrupted, releasing the N-terminal segment by intramolecular cleavage and thereby exposing the catalytic apparatus and the substrate-binding sites [43].

However, our results also reveal that certain properties of ProROL are apparently atypical. In particular, ProROL possesses enzyme activity that is even higher than that of the mature enzyme, thus demonstrating that the mature lipase part of the proenzyme is already properly folded and active. In spite of the low level of enzyme activity of ROL in *E. coli*, synthesis of this enzyme *in vivo* causes cell lysis, whereas ProROL has no toxic effect (Figures 3 and 4). The active site in lipases is completely buried beneath a short helical segment. In the presence of a substrate that generates an oil/water interface this lid moves, exposing the active site, in a process known as interfacial activation [44]. ROL and ProROL are likely to have different affinities for their substrates (Figure 8), the active pH range of ProROL is restricted, and ROL has phospholipase activity whereas ProROL does not. A central issue is whether the altered enzymic behaviour of ProROL is due to a direct interaction of the prosegment with the lid region. It is evident from our protein model of ROL [17] that the N-terminus of the mature lipase is on the opposite side of the lid (Figure 2). In the absence of sufficiently detailed three-dimensional data on the ProROL structure, we cannot unequivocally exclude the possibility that there is a direct interaction of the prosequence with the lid. However, it is unlikely that a direct interaction of this kind is responsible for the modulation of lipase activity. This conclusion derives from the observation that the shortened prosequences in Pro68ROL and Niveus, which we predict are even less likely to be able to interact with the lid, still limit the toxicity of the lipase in *E. coli*. It therefore seems likely that not only the lid, but also additional structural elements, are responsible for the interfacial activation process. We suggest that, in addition to lid movement, lipase activity is regulated by the enzyme's ability to penetrate the oil/water interphase and, by analogy, the membrane. Decreased penetration by ProROL might be the reason why this form of the enzyme can be expressed without causing serious damage to the host.

ProROL is significantly more stable against thermal inactivation than ROL (Figure 9). This contrasts with the properties of pro-proteases [45] and is inconsistent with the reason for the existence of these segments proposed by Baker et al. [7]. For example, the folding reactions of subtilisin and α -lytic protease

Scheme 1 Proposed function of the cysteine residue in the prosequence of *R. oryzae* lipase

are clearly under kinetic control [34,46]. Moreover, some of our results are consistent with the operation of simple thermodynamic principles of control in the folding of ROL and ProROL [38,47]. Thus ROL can fold with similar kinetics to ProROL and give a yield of approx. 100 % at low temperature (Figure 10). At first sight this seems to indicate that the reduced yield of active ROL *in vivo* is not due to a reduced ability to fold but solely to decreased thermal stability. However, our results with the different mutants of the prosequence clearly demonstrate that the folding of the lipase *in vivo*, and at physiological temperatures *in vitro*, is under kinetic control. The thermal stabilities of the point mutations of ProROL are as high as in the wild-type (Figure 9), but the yields of lipase activity are decreased to the same levels both *in vivo* and *in vitro* at a more physiological temperature (Table 1). We therefore conclude that whereas folding at low temperature is slow and may be subject to thermodynamic control, folding at physiological temperatures is much faster and primarily under kinetic control.

The mechanism by which prosequences can facilitate folding at the molecular level remains generally unclear. The model proposed for ProBPTI [10], involving a cysteine residue proposed to facilitate disulphide exchange, has been called into question [11]. Our results indicate that the cysteine residue of the lipase prosequence plays a key role in facilitating the folding of the enzyme. Mutation of the cysteine residue to a serine results in a protein with the same thermal stability, but the yield of lipase activity both *in vivo* and *in vitro* is decreased to one-fifth (Figure 12 and Table 1). The cysteine residue of the prosequence is surrounded by proline residues (Pro-Pro-Leu-Ile-Ser-Ser-Arg-Cys-Ala-Pro-Pro), suggesting that this stretch of amino acids forms a flexible surface loop. The whole prosequence contains 9 proline, 8 glycine, 19 serine, 6 aspartic acid and 6 asparagine residues, constituting a content of approx. 50 % of amino acids showing enhanced probability of causing β -turns. This indicates that the prosequence is probably very flexible, as was found for carboxypeptidase Y [48]. Pro68ROL has similar properties to ProROL(Cys-68 $\rightarrow$ Ser), despite the fact that one-third of the prosequence has been deleted. Only one-third of the processed protein carries the cysteine residue, owing to multi-site processing by leader peptidase. This suggests that the cysteine residue in this

one-third of processed Pro68ROL does not function as in the wild-type enzyme. This is not surprising, given that our mutational analysis has indicated that the neighbouring arginine residue at position -69 is necessary for the cysteine residue to be able to facilitate folding (Figure 5).

One explanation of the key role of the cysteine residue of the prosequence is that it provides a solvent-accessible, intramolecular thiol-disulphide reagent, analogously to the role suggested previously for the cysteine residue in the prosequence of BPTI [10]. According to this model, the cysteine residue has to react with a molecule of GSSG (Scheme 1). The reactive species in this step is thought to be the deprotonated cysteine residue, which forms a mixed disulphide via nucleophilic attack on GSSG [49]. The equilibrium of this reaction is dependent on the pK_a of the cysteine, which in turn is influenced by the nature of the surrounding amino acids in the protein sequence. In this context it is important to note that mutation of the proline residue in the active site of thioredoxin (Cys-Gly-Pro-Cys) to histidine has been shown to yield a cluster that mimics the active site of protein disulphide-isomerase. The mutated thioredoxin was found to have an increased redox potential and showed characteristics typical of protein disulphide-isomerase [50]. The relatively small effect of the prosequence in BPTI [10] on this enzyme's folding may also be due at least partly to a relatively high pK_a , because the cysteine residue in the prosequence is located next to an aspartic acid residue. In contrast, the cysteine residue of ProROL is an ideal candidate for the reaction with GSSG, which yields a protein-GSH disulphide. This mixed disulphide can either form a disulphide bond with another cysteine residue of the protein or transfer the GSH to a further cysteine residue (Scheme 1). Cysteine residues with a low pK_a are good leaving groups [51], making a transfer of GSH (Scheme 1) more likely. That the environment of a cysteine residue can strongly influence its ability to form structurally important disulphide bridges is underlined by the results obtained with the mutants Glu-265 $\rightarrow$ Gln and Asp-262 $\rightarrow$ Gly.

In conclusion, we propose a model in which the cysteine residue in the prosequence of ProROL facilitates the folding of the enzyme by acting as an intramolecular thiol-disulphide reagent acting analogously to the role of a catalytic cysteine proposed by Weissman and Kim [10]. However, in contrast with the model proposed for BPTI, we suggest that the key function of Cys-68 in ProROL is greatly dependent on its immediate environment in the polypeptide chain. In the absence of the appropriate context, this residue would be unable to perform its catalytic role in the process of folding. This might mean that protein structure can exert quite subtle control on the folding process via the environment of such cysteine residues in prosequences.

We thank R. Gieringer for performing the phospholipase assay, and Professor Robert Freedman for his comments on the manuscript. This work was supported in part by a European Community BRIDGE project grant.

REFERENCES

- Randall, L. L. and Hardy, S. J. S. (1989) *Science* **243**, 1156–1159
- Neurath, H. (1989) *Trends Biochem. Sci.* **14**, 268–271
- Naldini, L., Tamagnone, L., Vigna, E., Sachs, M., Hartmann, G., Birchmeier, W., Daikuhara, Y., Tsubouchi, H., Blasi, F. and Logmoglio, P. M. (1992) *EMBO J.* **11**, 4825–4833
- Munro, S. and Pelham, H. R. B. (1987) *Cell* **48**, 899–907
- Valls, L. A., Winther, J. R. and Stevens, T. H. (1990) *J. Cell Biol.* **111**, 361–368
- Ellis, R. J. and van der Vies, S. M. (1991) *Annu. Rev. Biochem.* **60**, 321–347
- Baker, D., Shiau, A. K. and Agard, D. A. (1993) *Curr. Opin. Cell Biol.* **5**, 966–970
- Shinde, U. and Inouye, M. (1993) *Trends Biochem. Sci.* **18**, 442–446

- 9 Gray, A. M. and Mason, A. J. (1990) *Science* **247**, 1328–1330
- 10 Weisman, J. S. and Kim, P. S. (1992) *Cell* **71**, 841–851
- 11 Creighton, T. E., Bagley, C. J., Cooper, L., Darby, N. J., Freedman, R. B., Kemmink, J. and Sheikh, A. (1993) *J. Mol. Biol.* **232**, 1176–1196
- 12 Rex, G., Surin, B., Besse, G., Schneppe, B. and McCarthy, J. E. G. (1994) *J. Biol. Chem.* **269**, 18118–18127
- 13 Hobson, A. H., Buckley, C. M., Aamand, J. L., Jorgensen, S. T., Diderichsen, B. and McConnell, D. J. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5682–5686
- 14 Frenken, L. G. J., Bos, J. W., Visser, C., Müller, W., Tommassen, J. and Verrips, C. T. (1993) *Mol. Microbiol.* **9**, 579–589
- 15 Frenken, L. G. J., de Groot, A., Tommassen, J. and Verrips, C. T. (1993) *Mol. Microbiol.* **9**, 591–599
- 16 Haas, M. J., Allen, J. and Berka, T. R. (1991) *Gene* **109**, 107–113
- 17 Beer, H. D., Wohlfahrt, G., McCarthy, J. E. G., Schomburg, D. and Schmid, R. D. (1996) *Protein Eng.* **6**, 507–517
- 18 Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W. (1990) *Methods Enzymol.* **185**, 60–89
- 19 Studier, F. W. (1975) *J. Bacteriol.* **124**, 307–316
- 20 Belev, T. N., Singh, M. and McCarthy, J. E. G. (1991) *Plasmid* **26**, 147–150
- 21 McCarthy, J. E. G. (1991) in *Advances in Gene Technology*, vol. 2 (Greenaway, P. J., ed.), pp. 145–175, JAI Press, London
- 22 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- 23 Mikaelian, I. and Sergeant, A. (1992) *Nucleic Acids Res.* **20**, 376
- 24 Sidebottom, C. M., Charton, E., Dunn, P. P. J., Mycock, G., Davies, C., Sutton, J. L., Macrae, A. R. and Slabas, A. R. (1991) *Eur. J. Biochem.* **202**, 485–491
- 25 Chalifa, V., Möhn, H. and Liscovitch, M. (1990) *J. Biol. Chem.* **265**, 17512–17519
- 26 Chattopadhyay, J., Natarajan, V. and Schmid, H. H. O. (1991) *J. Neurochem.* **57**, 1429–1436
- 27 Kobayashi, M. and Kanfer, J. N. (1987) *J. Neurochem.* **48**, 1597–1603
- 28 Derewenda, U., Swenson, R., Green, R., Wei, Y., Yamaguchi, S., Joerger, R., Haas, M. J. and Derewenda, Z. S. (1994) *Protein Eng.* **7**, 551–557
- 29 Derman, A. I., Prinz, W. A., Belin, D. and Beckwith, J. (1993) *Science* **262**, 1744–1747
- 30 Wülfing, C. and Plückthun, A. (1994) *Mol. Microbiol.* **12**, 685–692
- 31 Movva, N. R., Nakamura, K. and Inouye, M. (1980) *J. Biol. Chem.* **255**, 27–29
- 32 Zhu, X., Ohta, Y., Jordan, F. and Inouye, M. (1989) *Nature (London)* **339**, 483–484
- 33 Ohta, Y., Hojo, H., Aimoto, S., Kobayashi, T., Zhu, X., Jordan, F. and Inouye, M. (1991) *Mol. Microbiol.* **5**, 1507–1510
- 34 Baker, D., Silen, J. L. and Agard, D. A. (1992) *Nature (London)* **356**, 263–265
- 35 Winther, J. R., Sorensen, P. and Kielland-Brand, M. C. (1994) *J. Biol. Chem.* **269**, 22007–22013
- 36 Kohno, M., Kugimiya, W., Hashimoto, Y. and Morita, Y. (1994) *Biosci. Biotechnol. Biochem.* **58**, 1007–1012
- 37 Benzonana, G. (1974) *Lipids* **9**, 166–172
- 38 Anfinsen, C. B. (1973) *Science* **181**, 223–230
- 39 Hwang, C., Sinskey, A. J. and Lodish, H. F. (1992) *Science* **257**, 1496–1502
- 40 Wunderlich, M. and Glockshuber, R. (1993) *J. Biol. Chem.* **268**, 24547–24550
- 41 Abita, J.-P., Lazdunski, M., Bonsen, P. P. M., Pieterse, W. A. and de Haas, G. H. (1972) *Eur. J. Biochem.* **30**, 37–47
- 42 Kassel, B. and Kay, J. (1973) *Science* **180**, 1022–1027
- 43 James, N. N. G. and Sielecki, A. R. (1986) *Nature (London)* **319**, 33–38
- 44 Brzozowski, A. M., Derewenda, U., Derewenda, Z. S., Dodson, G. G., Lawson, D. M., Turkenburg, J. P., Bjorkling, F., Høge-Jensen, B., Patkar, S. A. and Thim, L. (1991) *Nature (London)* **351**, 491–494
- 45 Winther, J. R. and Sorensen, P. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 9330–9334
- 46 Eder, J., Rheinhecker, M. and Fersht, A. R. (1993) *Biochemistry* **32**, 18–26
- 47 Baker, D. and Agard, A. (1994) *Biochemistry* **33**, 7505–7509
- 48 Sorensen, P., Winther, J. R., Kaarsholm, N. C. and Poulsen, F. M. (1993) *Biochemistry* **32**, 12160–12166
- 49 Wunderlich, M., Otto, A., Seckler, R. and Glockshuber, R. (1993) *Biochemistry* **32**, 12251–12256
- 50 Krause, G., Lundström, J., Barea, J. L., de la Cuesta, C. P. and Holmgren, A. (1991) *J. Biol. Chem.* **266**, 9494–9500
- 51 Gilbert, H. F. (1990) *Adv. Enzymol.* **63**, 69–172