Protein Engineering vol.10 no.11 pp.1339–1345, 1997

High-level expression of bovine β-lactoglobulin in *Pichia pastoris* and characterization of its physical properties

Tae-Rak Kim, Yuji Goto[1], Nami Hirota[1], Kazuo Kuwata[2], Helen Denton[3], Su-Ying Wu[3], Lindsay Sawyer[3] and Carl A.Batt[4]

Department of Food Science, Cornell University, Ithaca, NY 14853, USA, [1]Department of Biology, Graduate School of Science, Osaka University, Toyonaka, Osaka 560, [2]Department of Physiology, School of Medicine, Gifu University, Gifu 500, Japan and [3]Structural Biochemistry Group, University of Edinburgh, Swann Building, King's Buildings, Mayfield Road, Edinburgh EH9 3JR, UK

[4]To whom correspondence should be addressed

Bovine β-lactoglobulin (BLG) variant A has been expressed in the methylotropic yeast *Pichia pastoris* by fusion of the cDNA to the sequence coding for the α-mating factor prepro-leader peptide from *Saccharomyces cerevisiae. P.pastoris* Mut$^+$ transformants were obtained by single cross-over integration of the BLG-containing vector into the AOX1 locus. In a fed-batch fermenter, a cell density of approximately 300 mg/ml was achieved by controlled glycerol feeding for a total of 24 h. After 72 h of methanol induction, the secreted BLG reached levels of >1 g/l. The secreted protein could be purified to homogeneity by ion-exchange chromatography. Amino-terminal sequencing of the secreted BLG revealed that the Glu–Ala spacer repeats inserted between the mature protein and the α-factor prepro-leader were still present. The purified protein was characterized by a number of methods, including CD spectroscopy, guanidine–HCl unfolding, crystallization and two-dimensional ^{1}H-NMR spectroscopy. By all of these measures, the physical characteristics of recombinant BLG were indistinguishable from those of the native purified bovine BLG, making it useful as a model for protein folding and other biophysical studies.

Keywords: bovine β-lactoglobulin expression/*Pichia pastoris*

Introduction

β-Lactoglobulin (BLG) is the major whey protein found in the milk of ruminants, including cows and sheep, and also monogastrics, e.g. pigs, horses, dogs and cats. It is not, however, found in the milk of humans (Hambling *et al.*, 1992). BLG is a globular protein with a molecular weight of ~18 kDa with each monomer of 162 amino acid residues. The secondary structure of bovine BLG is composed of 15% α-helix, 50% β-sheet and 15–20% reverse turn (Creamer *et al.*, 1983). The tertiary structure consists of nine strands of anti-parallel β-sheet, eight of which form a β-barrel that has the shape of a flattened cone or calyx (Papiz *et al.*, 1986; Brownlow *et al.*, 1997). There is a three-turn α-helix on the outer surface of the calyx. The interior of the β-barrel is hydrophobic, whereas the opening is lined with hydrophobic amino acids. The eight-stranded β-barrel is a major structural motif found in a family of proteins which have the ability to bind several amphiphilic or hydrophobic ligands (Flower *et al.*, 1993; Banaszak *et al.*, 1994) such as retinol (Futterman and Heller, 1972), long-chain fatty acids (Spector and Fletcher, 1970) and sodium dodecyl sulfate (McMeekin *et al.*, 1949; Seibles, 1969; Jones and Wilkinson, 1976; Lamiot *et al.*, 1994).

Recent studies by Goto and co-workers (Hamada *et al.*, 1995, 1996; Shiraki *et al.*, 1995; Kuroda *et al.*, 1996; Hamada and Goto, 1997) show that BLG is unique and an interesting model to elucidate the mechanism of protein folding. Whereas the native structure as determined by the free energy minimum is predominantly β-sheet, the local interactions between neighboring amino acid residues strongly favor the α-helical conformation. The results suggest a case of non-hierarchical protein folding in which the non-native α-helical structures play important roles in forming the native β-sheet, a striking contrast to the folding models suggested from studies of other proteins.

Examination of the structure–function relationships of a protein require an efficient expression system. Recombinant bovine BLG was first expressed in *Escherichia coli* (Batt *et al.*, 1990) and *Saccharomyces cerevisiae* (Totsuka *et al.*, 1990). More recently, it has also been expressed in *Kluyveromyces lactis* (Rocha *et al.*, 1996). We pursued an alternative host for high-level expression of recombinant bovine BLG, since when produced in *E.coli* it formed inclusion bodies, whereas in other systems the yields were not deemed sufficient for a number of biophysical studies. The methylotropic yeast *Pichia pastoris* was selected as the expression host (Cregg and Higgins, 1995; Romanos, 1995) because of its ability to grow to very high cell density (Cregg and Higgins, 1995) while producing alcohol oxidase up to 30% of its total soluble protein when fully induced (Couderc and Baratti, 1980; Cregg *et al.*, 1985). It also secretes very little of its own protein, simplifying purification of any heterologous secreted protein (Barr *et al.*, 1992). Prior reports documented the ability of *P.pastoris* to produce heterologous proteins at yields exceeding 5 g/l (Romanos, 1995).

Materials and methods

Strains and plasmids

E.coli TG1 was used as a host strain for constructing BLG/pPIC9. pPIC9 contains the alcohol oxidase (AOX) I promoter, His$^+$ selectable marker and prepro α-mating factor (α-MF prepro) secretion signal derived from *S.cerevisiae* (Invitrogen, San Diego, CA) (Figure 1). Plasmid pTTQ18BLG, the source of the BLG cDNA, was prepared from *E.coli* TG1 (Batt *et al.*, 1990). *P.pastoris* GS115 (Invitrogen) was selected as a host strain for expression (Cregg *et al.*, 1985).

DNA manipulation

Plasmid DNA was isolated using a Qiagen (Santa Clarita, CA) plasmid preparation kit. Restriction enzymes were purchased from New England Biolabs (Beverly, MA) and ligase from GIBCO BRL (Gaithersburg, MD). The purification and the ligation of digested plasmid or PCR products were performed by published procedures (Ausubel *et al.*, 1990).

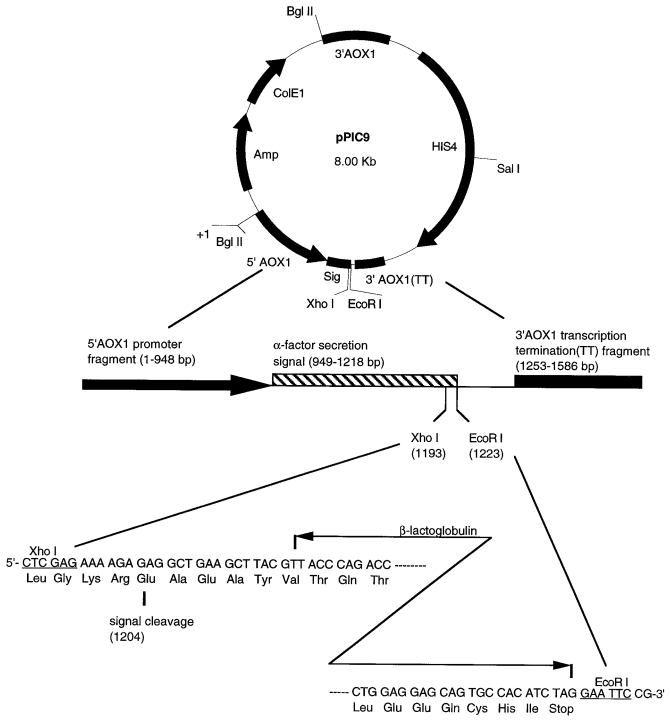

Fig. 1. Restriction map of BLG/pPIC9.

Medium

E.coli TG1 was grown in Luria broth (Ausubel *et al.*, 1990) containing ampicillin (100 mg/l). BMGY (buffered glycerol-complex medium; 1% yeast extract, 2% peptone, 0.1 M potassium phosphate, pH 6.0, 1.34% yeast nitrogen base, 4×10^{-5}% biotin and 1% glycerol) and BMMY (buffered methanol-complex medium; as BMGY except 0.5% methanol was used instead of glycerol) were used for growing *P.pastoris* and producing recombinant BLG, respectively. MD (Minimal Dextrose Medium; 1.34% yeast nitrogen base, 4×10^{-5}% biotin, 1% dextrose) agar was used for screening of His^+ transformants and MM (Minimal Methanol Medium; as MD except 0.5% methanol was added instead of dextrose) for Mut (Methanol Utilization) screening.

Construction of expression vector

A 480 bp DNA fragment encoding bovine BLG was amplified by PCR from pTTQ18BLG (Batt *et al.*, 1990), with the following oligonucleotide primers: 5′ccGCTCGAGAAAAGAGAGGCTGAAGCTTACGTTACCCAGACC-3 (5′*Xho*I) and 5′CGGAATTCCTAGATGTGGCACTGCTCCTCCAG-3 (3′*Eco*RI). The PCR was prepared as follows: 100 ng template (pTTQ18BLG), 1 μl primers (40 pmol), 12 μl magnesium chloride (25 mM), 10 μl reaction buffer (10×), 2.5 μl deoxynucleotides (2.5 mM), 0.5 μl *Taq* DNA polymerase (5 units/μl) (Perkin-Elmer Cetus, Norwalk, CT) in a final volume of 100 μl with dH_2O. This mixture was subjected to 1 cycle of 94°C for 5 min and 30 cycles of 94°C for 1 min (denaturation), 55°C for 1 min (annealing), 72°C for 1 min (extension) and 1 cycle of 72°C for 7 min using the 9600 DNA Thermal Cycler (Perkin-Elmer Cetus). The PCR product was purified using a Qiagen kit and after *Xho*I, *Eco*RI digestion it was cloned into the *Xho*I and *Eco*RI sites of pPIC9, generating BLG/pPIC9 (Figure 1). Portions of BLG/pPIC9 were sequenced with 5′*Xho*I primer and 5′AOX I primer (5′GACTGGTTCCAATTGACAAGC-3).

Transformation of P.pastoris

P.pastoris GS115(His^-) was transformed with BLG/pPIC9 digested with *Bgl*II. Approximately 0.3 μg of the linearized plasmid DNA was used for electroporation in 0.2 cm cuvettes at 1.5 kV, 25 mF, 400 W, using a Gene Pulser (Bio-Rad, Richmond, CA). Immediately after pulsing, 1 ml of cold 1 M sorbitol was added to the cuvette. Cells were plated on to MD agar for the selection of His^+ transformants. Transformants containing integrated BLG/pPIC9 were screened as follows: a His^+ transformant was inoculated into 4 ml of BMGY and incubated for 1 day at 30°C. A 1 μl aliquot from 20 μl of cell lysate was used for PCR with 5′*Xho*I and 3′*Eco*RI primers. To screen for methanol utilization, 2.5 μl of each culture were spotted on to MM agar plates and a Mut^+ control (GS115/His^+/β-galactosidase, Invitrogen) and Mut^s control (GS115/His^+/albumin, Invitrogen) were included for comparison. After 24–48 h, Mut^+ colonies were visibly larger than the Mut^s transformants.

Screening for bovine BLG expression

P.pastoris [BLG/pPIC9] were incubated for 2 days at 30°C with shaking and 3 ml of cells were collected by centrifugation and resuspended in 1 ml of BMMY. Methanol was added every 24 h to a final concentration of 0.5%. The secretion of BLG into the culture medium was monitored using a 12% SDS–PAGE and Coomassie Blue staining. Purified BLG (Sigma Chemical, St Louis, MO) was used as a reference standard.

Fermentation of P.pastoris [BLG/pPIC9]

The fermentation of *P.pastoris* [BLG/pPIC9] was carried out using a BioFlo 3000 fermenter (New Brunswick Scientific, Edison, NJ). The seed culture was prepared by inoculation from a frozen glycerol stock of GS115/His^+ BLG/Mut^+ into 5 ml of phosphate-buffered YNB (11.5 g/l KH_2PO_4, 2.66 g/l K_2HPO_4, 6.7 g/l yeast nitrogen base, pH 6.0) and 2% glycerol. After overnight incubation at 30°C, the culture was transferred into 300 ml of the same medium and grown to an OD_{600} of 2–6. The seed culture was transferred into 5 l of Basal Salts medium (BSM; 26.7 ml/l phosphoric acid, 0.93 g/l calcium sulfate dihydrate, 18.2 g/l potassium sulfate, 14.9 g/l magnesium sulfate heptahydrate and 4.13 g/l potassium hydroxide) plus 10 ml of PTM1 salts (6.0 g/l cupric sulfate pentahydrate, 0.08 g/l sodium iodide, 30 g/l manganese sulfate monohydrate, 0.2 g/l sodium molybdate dehydrate, 0.02 g/l boric acid, 0.5 g/l cobalt chloride, 20.0 g/l zinc chloride, 65.0 g/l ferrous sulfate heptahydrate, 0.2 g/l biotin and 5 ml/l sulfuric acid). The pH of the medium was adjusted to 5.0 wtih 28% ammonia solution. The culture was grown for ~14 h until the wet cell weight reached 90–150 g/l, and then 50% (w/v) glycerol was fed at a rate of about 8 ml/h.l for 12 h until the wet cell weight reached 300 g/l. To induce AOX1 transcription, methanol was added at a rate of 1 ml/h.l. The feeding rate of methanol was gradually increased to 8 ml/h.l over 10 h and then further increased to a maximum growth-limiting rate while monitoring DO spike and methanol accumulation. Samples were taken from the fermenter at different times and the cell density and protein level measured (Bio-Rad).

Purification of BLG

The culture supernatant was harvested 48–72 h after the start of methanol induction by centrifugation at 1500 *g* for 20 min and ~100 ml of supernatant was concentrated to 20 ml by pressure filtration using a YM10 membrane (Amicon, Beverly, MA). The retentate was washed with 50 ml of 50 mM

imidazole buffer (pH 6.7) and passed through a 2.5×25 cm Q-Sepharose Fast Flow column (Pharmacia, Piscataway, NJ) equilibrated with same buffer. The column was washed with 70 ml of same buffer and BLG was eluted with a linear gradient of 0 to 0.5 M NaCl in 50 mM imidazole–HCl buffer (pH 6.7). A 30 ml fraction of purified recombinant BLG was desalted by ultrafiltration using a YM10 membrane and then lyophilized.

Western blot analysis

Proteins in the cell-free supernatant and the purified fractions were analyzed using 12.5% SDS–PAGE (Laemmli) and electroblotted on to nitrocellulose membranes. Immunoreactive protein was detected using polyclonal antibodies against BLG raised in Flemish Giant rabbit (Batt *et al.*, 1990). Primary antibody/antigen complexes were detected using anti-goat anti-rabbit antiserum conjugated to horseradish peroxidase and developed using the Bio-Rad HRP substrate.

N-Terminal amino acid sequence

The secreted BLG in the cell-free supernatant was purified by 12% SDS–PAGE and electroblotted on to a PVDF membrane (Bio-Rad). The N-terminal sequence was determined by automated Edman degradation using an Applied Biosystems (Foster City, CA) gas-phase sequencer.

CD measurements

The CD spectra were measured using a Jasco Model J-720 spectropolarimeter at 20°C. The data were expressed as mean residue ellipticity, [θ]. The protein concentration was 0.2 mg/ml for both the far- and near-UV CD measurements. The pH was maintained at 3.0 with 20 mM glycine–HCl buffer. Cells with 1 and 10 mm lightpaths were used for the far- and near-UV CD spectra, respectively. The unfolding transition curves measured by the ellipticity at 218 nm were analyzed on the basis of a two-state transition mechanism between the native state (N) and the unfolded state (U), N ↔ U, and a linear dependence of the free energy change of unfolding (ΔG_U) on Gdn–HCl concentration, [Gdn–HCl]:

$$\Delta G_U = \Delta G_U(H_2O) - m[\text{Gdn–HCl}]$$

where $\Delta G_U(H_2O)$ is ΔG_U in the absence of denaturant and m provides a measure of the dependence of ΔG_U on Gdn–HCl concentration (Myers *et al.*, 1995). Adjusting the baselines manually for the native and unfolded states, the $\Delta G_U(H_2O)$ and m values were obtained by a least-squares curve fitting program.

Crystallization and X-ray measurements

Purified BLG was crystallized by the hanging drop method whereby a 5 μl drop of protein (15 mg/ml) in 0.2 M EPPS buffer, pH 7.9, was mixed with 5 μl of 2.35 M ammonium sulfate also in 0.2 M EPPS buffer from the well solution, over which the drop was equilibrated. Crystals grew at room temperature in several days and from their habit were tentatively identified as lattice Y (Papiz *et al.*, 1986).

X-ray measurements on a crystal ~0.1×0.1×0.2 mm were made at 100 K on a MarResearch image plate system mounted on an ENRAF-Nonius FR571 generator operating at 40 kV, 80 mA with a graphite monochromator. The 2° oscillations were collected at 600 s/° and the frames processed to obtain the unit cell dimensions with the program DENZO (Otwinowski and Minor, 1997).

NMR measurements

NMR spectra were measured using a Varian Inova 500 MHz spectrometer at 60°C. BLG solutions (10 mg/ml) were prepared by dissolving the lyophilized protein in 10% (v/v) H_2O and 90% H_2O. The pH was adjusted to 2.0 at room temperature by adding 1.0 M HCl. No additional salt was present. Water suppression was accomplished using a symmetrically shifted pulse (Smallcombe, 1993) with minimum presaturation. The 2D phase-sensitive COSY spectra were recorded by the hypercomplex method (States *et al.*, 1982) with the following parameters: spectral width 7500 Hz in each dimension and size 4096 complex points in t_2 with 512 hypercomplex t_1 increments. Each dimension was processed with sine-bell and Gaussian functions. The acquisition time was 1 day for each spectrum.

Results

Construction of pPIC9 expression vector

To facilitate the cloning of the BLG cDNA into pPIC9, a *Xho*I site was introduced at the end of α-MF prepro. This changed the Asp83, that is immediately prior to the cleavage site of the *KEX2* protease, to Glu. The cDNA for BLG was inserted in between the *Xho*I and *Sna*BI sites of pPIC9 3′ to the secretion signal sequence using the 5′*Xho*I primer. The cDNA was placed in-frame with the 89 amino acids of *S.cerevisiae* α-MF prepro (Kurjan and Herskowitz, 1982), producing a fusion between it and the sequence coding for the α-MF prepro peptide. Heterologous proteins fused to the sequence are cleaved between Arg85 and Glu86 in the sequence Glu–Lys–Arg–Glu–Ala–Glu–Ala by the *KEX2* endopeptidase which cleaves on the carboxyl side of dibasic residues (Julius *et al.*, 1984). The Glu–Ala spacer repeats at the N-terminus that provides a hydrophilic environment at the *KEX2* endopeptidase cleavage site are then removed by the *STE13* dipeptidyl aminopeptidase (Julius *et al.*, 1983; Bussey, 1988; Anna-Arriola and Herskowitz, 1994). The amino acid sequence (Tyr–Val) after Glu–Ala repeats was introduced by 5′*Xho*I primer, slightly altering the N-terminal amino acid sequence of BLG from ^{1}Leu–^{2}Ile–^{3}Val to ^{1}Ala–^{2}Tyr–^{3}Val. Nucleotide sequencing of BLG/pPIC9 with 5 AOX primer confirmed the junction between the secretion signal and BLG cDNA.

Transformation of GS115

BLG/pPIC9 was linearized by digestion with *Bgl*II to obtain primary transformants of *P.pastoris*. His^+ transformants (GS115/His^+) that appeared on minimal dextrose medium (MD) without histidine supplement were screened for the BLG gene by PCR using the 5′*Xho*I and 3′*Eco*RI primers. GS115/His^+/BLG transformants were spotted on an MM plate to score for Mut^+ and Mut^s. These phenotypes were confirmed by PCR with 5′ and 3′ AOX sequencing primers. Mut^+ strains had a 2.2 kb PCR product corresponding to the intact AOXI gene and 0.9 kb PCR product corresponding to the BLG insert, while in Mut^s only the 0.9 kb fragment is observed. Mut^s are the result of a double cross event between the AOXI and 3′ and 5′ AOXI regions on pPIC9. All of transformants also produced a 0.5 kb fragment when subjected to PCR with 5′*Xho*I and 3′*Eco*RI primers which corresponded to the BLG cDNA.

Expression of BLG

Small-scale cultures of Mut^+ and Mut^s strains of *P.pastoris* [BLG/pPIC9] were screened for their ability to secrete BLG. BLG was detected in Mut^s after 1 day of methanol

induction whereas Mut$^+$ took 2 days. SDS–PAGE revealed that the predominant protein in the culture broth had an apparent molecular weight of 18 kDa. Western blot analysis revealed that it reacted with antibodies against bovine BLG (data not shown). No other immunoreactive proteins or even any other proteins that stained with Coomassie Blue were apparent in the culture medium.

Some *P.pastoris* GS115 His$^+$ which initially appeared to carry BLG/pPIC9 when tested by PCR did not express BLG in small-scale culture. A second PCR assay was performed with cell lysates from these strains after they were grown in BMGY medium for 24 h. All failed to produce a 0.5 kb BLG fragment in the PCR assay with the 5′*Xho*I and 3′*Eco*RI primers. These putative PCR positives detected in the initial screening could have carried a linear or circular copy of BLG/pPIC9. Because this vector does not carry a replicon that functions in *P.pastoris*, it cannot be maintained without integration.

Fermentation

A Mut$^+$ strain of GS115 [BLG/pPIC9] was selected because it metabolizes methanol and continues to grow rapidly while being fed methanol. These Mut$^+$ strains are more tolerant of methanol and the level of methanol during the induction phase rarely reaches toxic levels compared with the Mut^s strains. Alcohol oxidase levels in methylotropic yeast grown in carbon-limited cultures are significantly higher than when grown with excess methanol (Veenhuis *et al.*, 1983). The cell density and BLG level of *P.pastoris* GS115 [BLG/pPIC9] during the fermentation are presented in Figure 2. During 12 h of glycerol fed-batch, the wet cell weight increased from 80 to 310 g/l. After a 1 h starvation period, methanol was fed to the culture as described in the Materials and methods. The cells continued to grow during the methanol fed-batch phase to final cell density of 455 g/l wet cells. BLG was initially observed 12 h after the start of methanol induction and its level rose from 0 to 1.5 g/l over a period of 3 days, after which the fermentation was terminated. The final yield of BLG based upon SDS–PAGE using standard amounts of purified BLG for comparison was 1.5 g/l (Figure 2B). The secreted BLG was the predominant if not the only observed protein in the culture medium.

Characterization of the recombinant BLG

Physical characterization. BLG was purified using Q-Sepharose chromatography and a single band was observed on Coomassie Blue stained SDS–PAGE. Even with significant overloading in SDS–PAGE, no extraneous proteins were observed. The purified BLG migrated at the same position as purified bovine BLG. A total of six amino acids were sequenced from the N-terminus of the secreted BLG (Glu–Ala–Glu–Ala–Tyr–Val) and they revealed that the α-MF prepro peptide was cleaved at the carboxyl side of the dibasic residue, Lys–Arg, but that the Glu–Ala spacer at the N-terminus remained.

CD spectra. Conformation of the recombinant BLG was studied by CD and NMR spectra. Whereas bovine BLG exists as a dimer at neutral pH, it dissociates into monomers below pH 3, but still retains a native conformation even in an acidic environment as low as pH 2 (Fugate and Song, 1980). The far-UV CD spectrum of the recombinant BLG at pH 3.0 and 20°C had a minimum at 217 nm, consistent with the abundance of β-sheet structures (Figure 3A). The near-UV CD spectrum showed sharp peaks at 286 and 294 nm, indicating unique tertiary structures (Figure 3B). These spectral features of the recombinant BLG were essentially the same as those of bovine

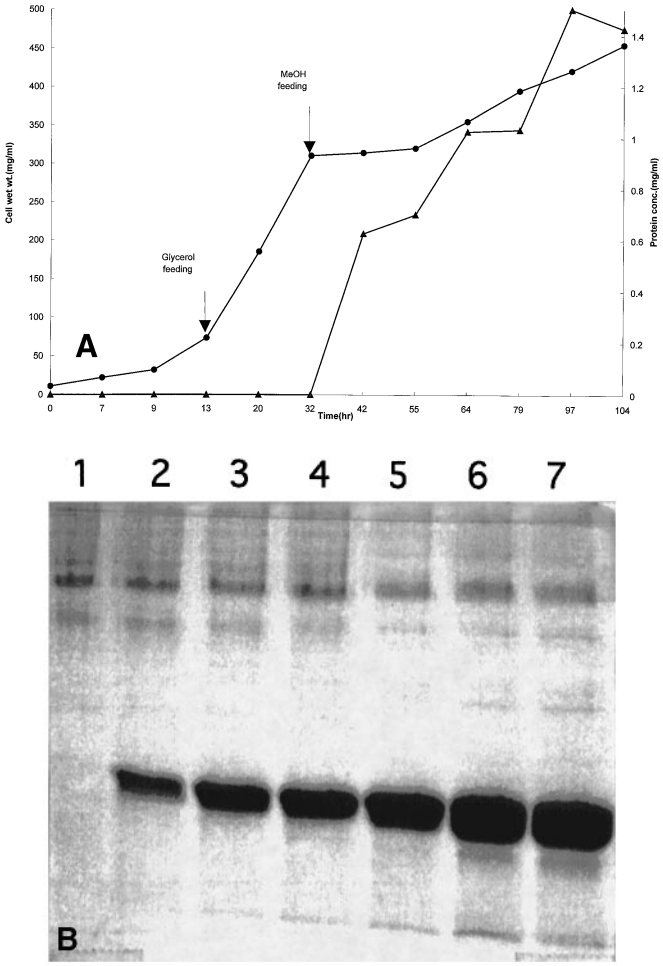

Fig. 2. Growth profile and β-lactoglobulin production by *P.pastoris* [BLG/pPIC9] in fermentation. (**A**) cell weight (●), BLG concentration (▲) and glycerol, methanol feeding times of fermentation culture grown at 30°C as described in Materials and methods. (**B**) SDS–PAGE analysis of culture broth after 0 h (lane 1), 10 h (lane 2), 23 h (lane 3), 32 h (lane 4), 47 h, (lane 5), 65 h (lane 6) and 72 h (lane 7).

BLG A purchased from Sigma. Hence the CD spectra indicate that the recombinant BLG at pH 3.0 assumes a unique native structure which is similar to that of bovine BLG. We also measured the CD spectra of the recombinant and bovine BLGs at pH 7. The spectra at pH 7 were similar to each other and were also similar to those at pH 3.0 (data not shown).

Gdn–HCl unfolding profiles. In order to compare the conformational stability, the unfolding transitions of the recombinant and bovine BLGs were measured by the ellipticity at 218 nm at pH 3.0 and 20°C (Figure 4). The unfolding transition curves of the two proteins were indistinguishable, again indicating the similarity of the two BLGs. For both proteins, the unfolding started at about 2 M Gdn–HCl and major transition ended at 4 M Gdn–HCl, the apparent mid-point of the transition being 3.2 M Gdn–HCl. The unfolding transition of the recombinant BLG at pH 3.0 was reversible (data not shown), as is the case with bovine BLG (Hamada *et al.*, 1996). The unfolding transition curves were analyzed on the basis of a two-state transition mechanism between the native state (N) and the unfolded state (U), and a linear dependence of the free energy change of unfolding (ΔG_U) upon Gdn–HCl concentration (see Materials and methods). The ΔG_U (H_2O) and m values were estimated to be 28.4 ± 0.1 kJ/mol and 8.8 ± 0.1 kJ/mol·(mol

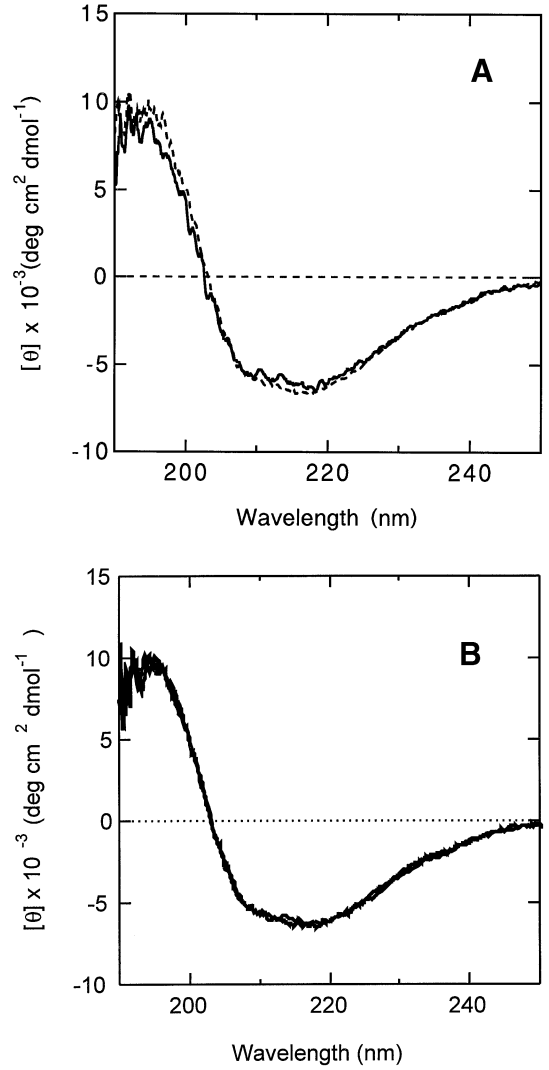

Fig. 3. Far- (**A**) and near-UV (**B**) CD spectra of the recombinant (solid lines) and bovine (broken lines) BLGs at pH 3.0 and 20°C.

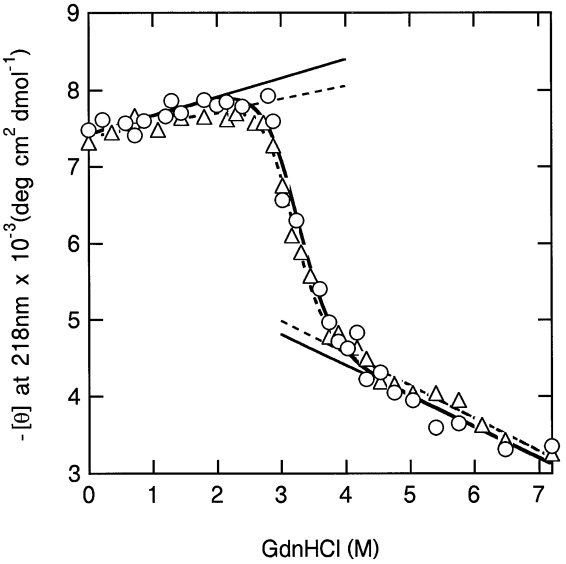

Fig. 4. Unfolding transitions of the recombinant (●) and bovine (▲) BLGs at pH 3.0 and 20°C. The curves indicate the theoretical curves on the basis of the two-state unfolding mechanism with the baselines (lines) and the parameters described in the text.

Table I. Unit cell parameters for the orthorhombic lattice Y or BLG, space group $C222_1$

Crystal	*T* (K)	*a* (Å)	*b* (Å)	*c* (Å)
Recombinant	100	54.64	80.06	66.97
Native	100	55.19	79.73	66.68
Native	293[a]	55.7	81.7	67.2

[a]Data from Aschaffenburg *et al.* (1965). Note that the space group given is $B22_12$, a permuted form of $C222_1$.

Gdn–HCl), respectively, for the recombinant BLG and 33.4 ± 0.1 kJ/mol and 10.6 ± 0.1 kJ/mol·(mol Gdn–HCl), respectively, for bovine BLG.

Crystallization. Crystals of *P.pastoris* produced BLG were obtained by the hanging drop method. These crystals were similar to one of the forms grown above pH 7 with the characteristic habit of orthorhombic lattice Y (Aschaffenburg *et al.*, 1965). While diffraction of these first crystals did not extend beyond 3 Å resolution, the unit cell parameters were determined to be close to those observed for native lattice Y crystals at 100 K and systematically shorter than those determined at room temperature (Table I).

NMR spectroscopy. To compare the structural details further, the 2D ^{1}H-NMR spectra of the recombinant and bovine BLGs were measured at pH 2.0 and 60°C (Figure 5). Measurements at a high temperature were necessary to obtain spectra with high peak resolution. The BLGs were still native at 60°C at pH 2.0 on the basis of the CD spectra (data not shown). In fact, for both BLGs, many αCH–NH cross peaks were clearly separated, consistent with the native structure with many β-sheets. The 162 amino acid residues of bovine BLG A give rise to 156 αCH–NH cross peaks in the COSY spectrum in H_2O, i.e. 162 (total amino acid residues) − 8 (Pro) − 1 (N-terminal) + 3 (Gly), and up to 36 additional cross peaks representing Arg and Lys side chain resonances. A total of 128 cross peaks are visible in Figure 5B. On the other hand, 165 amino acid residues of the recombinant BLG give rise to 159 αCH–NH cross peaks in the COSY spectrum, i.e. 165 (total amino acid residues) − 8 (Pro) − 1 (N-terminal) + 3 (Gly), and up to 36 additional cross peaks representing Arg and Lys side chain resonances. A total of 131 cross peaks are visible in Figure 5A. Hence we observed at least 80% of αCH–NH cross-peaks in the fingerprint regions of COSY spectra of both BLGs. Although the two spectra were very similar, several cross peaks were evidently different, demonstrating the high sensitivity of the NMR methodology. Although the assignments of the peaks were not performed, these differences probably come from the difference in amino acid composition at the N-terminal region and the differences in the three-dimensional structure around the corresponding region.

Discussion

BLG has been expressed from both recombinant prokaryotic and eukaryotic hosts. Initial levels of BLG expression in *E.coli* were low and the protein required solubilization followed by refolding. Typical yields were ~8–10 mg/l (Batt *et al.*, 1990). Parallel efforts to express BLG in *S.cerevisiae* using the glyceraldehyde 3-phosphate dehydrogenase promoter resulted in yields of 1.1 mg/l but, advantageously, this protein was secreted (Totsuka *et al.*, 1990). Finally, BLG has also been expressed in *K.lactis* using the phosphoglycerate kinase pro-

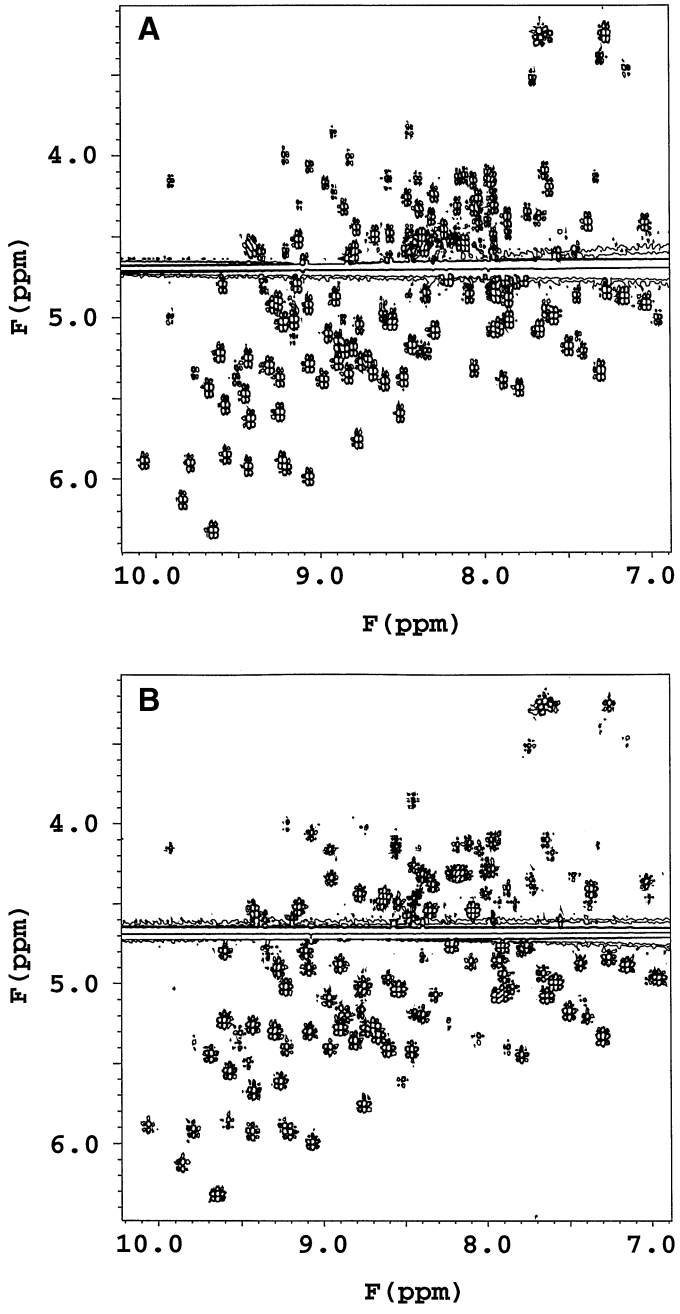

Fig. 5. Fingerprint regions of phase-sensitive COSY spectrum of the recombinant (**A**) and bovine (**B**) BLGs at pH 2.0 and 60°C. BLG solutions (10 mg/ml) were prepared by dissolving the lyophilized protein in 10% (v/v) H_2O and 90% H_2O.

moter and yields of 40–50 mg/l in secreted protein were reported (Rocha *et al.*, 1996). In this latter effort, the native BLG signal was used to direct secretion and proper processing observed. BLG was efficiently expressed from *P.pastoris* and copious amounts of fully folded protein in excess of 1 g/l could be recovered from the supernatant. A homolog to BLG, lipocalin, has also been expressed in *P.pastoris* using its native signal sequence with yields of 270 mg/l (Ferrari *et al.*, 1997). This recombinant lipocalin has been characterized with respect to its ability to bind odorant molecules and shown to be similar to the protein purified from its natural source.

Both Mut^s and Mut^+ *P.pastoris* recombinants were recovered after transformation with *Bgl*II linearized BLG/pPIC9 vector. Single-crossover AOX1 integrants occurred with *Bgl*II digested DNA due to the *in vivo* ligation of the linearized fragment (Clare *et al.*, 1991). The influence of the Mut phenotype on expression in *P.pastoris* has not been extensively examined. The faster growing Mut^+ strains are selected because they produce high levels of biomass in fermentation (Digan *et al.*, 1989) and, as noted for *S.cerevisiae*, secretion is localized to the growing bud tip (Schekman and Novick, 1982).

The preproα-MF leader sequence derived from *S.cerevisiae* is effective at secreting and directing the processing of the recombinant BLG in *P.pastoris*. Cleavage of the leader is apparently mediated through a *KEX2*-like activity in *P.pastoris*. In the secretory pathway of yeast, the signal peptide is removed by a peptidase (Blobel, 1977) and folding occurs in the endoplasmic reticulum with the assistance of accessory proteins including disulfide isomerase (Freedman, 1989).

The efficiency of the *STE13* cleavage depends on the heterologous protein being expressed (Brake *et al.*, 1984). The Glu–Ala spacer residues, which provide a hydrophilic environment at the *KEX2* cleavage site, are removed in certain cases, but not in others (Guisez *et al.*, 1991) suggesting an intrinsic role for the recombinant protein. Incomplete cleavage by the *STE13* gene product might also be due to the occlusion of the cleavage site on the secreted protein from the peptidase (Hinchliffe and Kenny, 1993). The N-terminus of the partially folded recombinant BLG may not be accessible to *STE13*. BLG exists as dimer at physiological pH (Pessen *et al.*, 1985) and the negatively charged N-terminus may be involved in dimer formation. Differences in the N-terminus of *P.pastoris* produced BLG as compared with native BLG probably do not have a significant effect on the overall structure and this region of the protein is not well defined in the crystal structure (Papiz *et al.*, 1986; Brownlow *et al.*, 1997). Certainly there is no obvious effect upon the crystallinity as reflected in the preliminary characterization of the crystals.

The level of *STE13* gene product normally present in the yeast may not be sufficient to process the high levels of expressed protein (Brake *et al.*, 1984). Barnes *et al.* (1982) reported that *STE13* gene on a multicopy plasmid increased complete processing of *MFα1* gene. No effort to overexpress STE13 in *P.pastoris* which might circumvent this problem has been reported.

All the CD, X-ray and NMR results indicate that the conformation and stability of the recombinant and bovine BLGs are practically the same, both in solution and in the crystalline state. In particular, the COSY spectra (Figure 5) provide convincing evidence of structural similarity as well as the slight difference probably arising from the N-terminal region. These efforts comprise a very complete characterization of the recombinant BLG as a prelude to using this system to study the structure–function of the protein.

Protein structures are determined by the local and non-local interactions. BLG is a unique example representing inconsistency of the local and non-local secondary structure preference (Hamada *et al.*, 1995, 1996; Shiraki *et al.*, 1995; Kuroda *et al.*, 1996; Hamada and Goto, 1997). Detailed analysis of the folding mechanism of BLG is of special importance for understanding the interplay between local and non-local interactions during protein folding. For such studies, heteronuclear NMR methods, which can provide the detailed structural information at the atomic level, will be essential. The COSY spectra with high peak resolution indicate that this protein is suitable for such NMR studies. The high level

expression of the native BLG in *P.pastoris* permits the efficient expression of the isotope-labeled proteins. Indeed, with the uniformly ^{15}N-labeled and ^{15}N- and ^{13}C-double-labeled BLGs, further NMR spectral studies are in progress, including the peak assignments and its three-dimensional structure determination in solution.

Acknowledgements

This work was supported by grants from the National Dairy Promotion and Research Board (C.A.B.), from the Ministry of Education, Science and Culture of Japan (Y.G.) and from the Biotechnology and Biological Sciences Research Council (L.S.). C.A.B. and L.S. are grateful to NATO for a travel award which facilitated aspects of this work. The authors also thank Lee Henderson of Viral Therapeutics (Ithaca, NY) for help with the fermentations.

References

Anna-Arriola,S.S. and Herskowitz,I. (1994) *Yeast*, **10**, 801–810.
Aschaffenburg,R., Green,D.W. and Simmons,R.M. (1965) *J. Mol. Biol.*, **13**, 194–201.
Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Smith,J.A., Seidman,J.G. and Struhl,K. (eds) (1990) In *Current Protocols in Molecular Biology.* Greene Publishing and Wiley-Interscience, New York.
Banaszak,L., Winter,N., Xu,Z., Bernlohr,D.A., Cowan,S. and Jones,T.A. (1994) *Adv. Protein Chem.*, **45**, 89–151.
Barnes,D., Blair,L., Brake,A., Church,M., Julius,D., Kunisawa,R., Lotko,J., Stetler,G. and Thorner,J. (1982) *Recent Adv. Yeast Mol. Biol.*, **1**, 295–305.
Barr,K.A., Hopkins,S.A. and Sreekrishna,K. (1992) *Pharm. Engng*, **12**, 48–51.
Batt,C.A., Rabson,L.D., Wong,D.W.S. and Kinsella,J.E. (1990) *Agric. Biol. Chem.*, **54**, 949–956.
Blobel,G. (1977) In Brinkley,B.R. and Porter,K.R. (eds), *International Cell Biology.* Rockefeller University Press, New York, p. 318.
Brake,A.J., Merryweather,J.P., Coit,D.G., Heberlein,U.A., Masiarz,F.R., Mullenbach,G.T., Urdea,M.S., Valenzuela,P. and Barr,P.J. (1984) *Proc. Natl Acad. Sci. USA*, **81**, 4642–4646.
Brownlow,S., Morais Cabral,J.H., Cooper,R., Flower,D.R., Yewdall,S.J., Polikarpov,I., North,A.C.T. and Sawyer,L. (1977) *Structure*, **5**, 481–495.
Bussey,H. (1988) *Yeast*, **4**, 17–26.
Clare,J.J., Rayment,F.B., Ballantine,S.P., Sreekrishna,K. and Romanos,M.A. (1991). *Bio/Technology*, **9**, 455–460.
Couderc,R. and Baratti,J. (1980) *Agric. Biol. Chem.*, **44**, 2279–2289.
Creamer,L.K., Parry,D.A.D. and Malcolm,G.N. (1983) *Arch. Biochem. Biophys.*, **227**, 98–105.
Cregg,J.M. and Higgins,D.R. (1995) *Can. J. Bot.*, **73**, S891–S897.
Cregg,J.M., Barringer,K.J., Hessler,A.Y. and Madden,K.R. (1985) *Mol. Cell Biol.*, **5**, 3376–3385.
Digan,M.E. *et al.* (1989) *Bio/Technology*, **7**, 160–164.
Ferrari,E., Lodi,T., Sorbi,R.T., Tirindelli,R., Cavaggioni,A. and Spisni,A. (1997) *FEBS Lett.*, **401**, 73–77.
Flower,D.R., North,A.C.T. and Attwood,T.K. (1993) *Protein Sci.*, **2**, 753–761.
Freedman,R.B. (1989) *Cell*, **57**, 1069–1072.
Fugate,R.G. and Song,P.-S. (1980) *Biochim. Biophys. Acta*, **625**, 28–42.
Futterman,S. and Heller,J. (1972) *J. Biol. Chem.*, **247**, 5168–5172.
Guisez,Y., Tison,B., Vandekerckhove,J., Demolder,J., Bauw,G., Haegeman,G., Fiers,W. and Contreras,R. (1991) *Eur. J. Biochem.*, **198**, 217–222.
Hamada,D. and Goto,Y. (1997) *J. Mol. Biol.*, **269**, 479–487.
Hamada,D., Kuroda,Y., Tanaka,T. and Goto,Y. (1995) *J. Mol. Biol.*, **254**, 737–746.
Hamada,D., Segawa,S.I. and Goto,Y. (1996) *Nature Struct. Biol.*, **3**, 868–873.
Hambling,S.G., McAlpine,A.S. and Sawyer,L. (1992) *Adv. Dairy Chem.*, **1**, 14l–190.
Hinchliffe,E. and Kenny,E. (1993) In Rose,A.H. and Harrison,J.S. (eds), *The Yeasts*. Academic Press, San Diego, pp. 325–352.
Jones,M.N. and Wilkinson,A. (1976) *Biochem. J.*, **153**, 713–718.
Julius,D., Blair,L., Brake,A., Sprague,G. and Thoner,J. (1983) *Cell*, **32**, 839–852.
Julius,D., Brake,A., Blair,L., Kunisawa,R. and Thorner,J. (1984) *Cell*, **37**, 1075–1089.
Kurjan,J. and Herskowitz,I. (1982) *Cell*, **30**, 933–943.
Kuroda,Y., Hamada,D., Tanaka,T. and Goto,Y. (1966) *Folding Desig.*, **1**, 243–251.
Lamiot,E., Dufour,E. and Haertle,T. (1994) *J. Agric. Food Chem.*, **42**, 695–699.
McMeekin,T.L., Polis,B.D., DellaMonica,E.S. and Custer,J.H. (1949) *J. Am. Chem. Soc.*, **71**, 3606–3609.
Myers,J.K., Pace,C.N. and Scholtz,J.M. (1995) *Protein Sci.*, **4**, 2138–2148.
Otwinowski,Z. and Minor,W. (1997) *Meth. Enzmol.* **276**, 307–326.
Papiz,M., Sawyer,L., Eliopoulos,E., North,A., Findlay,J., Sivaprasadarao,R., Jones,T., Newcomer,M. and Kraulis,P. (1986) *Nature*, **324**, 383–385.
Pessen,H., Purcell,J.M. and Farrell,H.M.,Jr (1985) *Biochim. Biophys. Acta*, **828**, 1–12.
Rocha,T.L., Paterson,G., Crimmins,K., Boyd,A., Sawyer,L. and Fothergill Gilmore,L.A. (1996) *Biochem. J.*, **313**, 927–932.
Romanos,M. (1995) *Curr. Opin. Biotechnol.*, **6**, 527–533.
Schekman,R. and Novick,P. (1982) In Strathern,J.N., Jones,E.W. and Broach,J.R. (eds), *The Molecular Biology of the Yeast Saccharomyces* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 361–393.
Seibles,T.S. (1969) *Biochemistry*, **8**, 2949–2954.
Shiraki,K., Nishikawa,K. and Goto,Y. (1995) *J. Mol. Biol.*, **245**, 180–194.
Smallcombe,S.H. (1993) *J. Am. Chem. Soc.*, **115**, 4776–4785.
Spector,A.E and Fletcher,J.E. (1970) *Lipids*, **5**, 403–411.
States,D.J., Haberkorn,R.A. and Ruben,D.J. (1982) *J. Magn. Reson.*, **48**, 386–393.
Totsuka,M., Katakura,Y., Shimizu,M., Kumagai,I., Miura,K.I. and Kaminogawa,S. (1990) *Agric. Biol. Chem.*, **54**, 3111–3116.
Veenhuis,M., van Dijken,J.P. and Harder,W. (1983) *Adv. Microb. Physiol.*, **24**, 1–82.

Received May 30, 1997; revised July 28, 1997; accepted August 4, 1997