ELSEVIER

Journal of Chromatography A, 668 (1994) 129–137

JOURNAL OF CHROMATOGRAPHY A

Separation and purification of recombinant proteins from *Escherichia coli* with aqueous two-phase systems

J.A. Asenjo*, R.E. Turner, S.L. Mistry, A. Kaul

Biochemical Engineering Laboratory, University of Reading, Reading RG6 2AP, UK

Abstract

The partition of the protein thaumatin in the presence of *Escherichia coli* contaminant proteins has been studied. Extraction of thaumatin was followed by back-extraction of the product into a new phosphate phase and also back extraction combined with recycle of the top polyethylene glycol (PEG) phase. When partitioned in the absence or presence of insoluble cell debris (whole cell homogenate) little effect on the partitioning of thaumatin or the soluble *E. coli* protein was observed. A back extraction step that allowed for dilution of the NaCl was successful in extracting thaumatin back into a heavy phosphate phase.

The PEG top phase was recycled to the first extraction stage. The stability ratio (tie-line length) had an effect on the average partition coefficient (K_{app}) of *E. coli* soluble proteins in PEG–phosphate systems but in PEG sulphate systems K_{app} was independent of this ratio. An increase in NaCl resulted in an increase in K_{app} but this was always below 1. A mathematical model that describes the continuous steady-state operation of extraction and back extraction has been developed; it is based on steady state mass balances of the main components and phase equilibrium data and was successfully used to simulate the extraction and back extraction processes.

1. Introduction

The production of protein products in *Escherichia coli* is still a cornerstone of the biotechnology industry and thus much attention has been focused on the recovery of intracellular recombinant proteins [1]. Aqueous two-phase technology offers an attractive step in the separation and purification of proteins from their major contaminants. The main advantages over traditional adsorption separation processes lies in their ability to handle particulate material and to process large volumes in a continuous mode. Separation can be achieved whereby the particulates and major contaminants partition to the bottom phase and the protein product to the top phase. Recovery of the product can be achieved in a second stage where conditions for the elution and partition of the target protein into the bottom phase is obtained [2–5].

Aqueous two-phase systems (ATPSs) can be composed of either two water-soluble polymers usually polyethylene glycol (PEG) and dextran or a polymer (PEG) and salt usually phosphate, sulphate or citrate. There are many factors which influence partition in ATPSs. These factors are inherent to the system itself (*e.g.* choice of the system components, distance of the system from the binodial, polymer molecular mass, concentration of polymers and salts, ionic composition and strength and pH) and to the protein to be partitioned (*e.g.* molecular mass, charge and hydrophobicity). Thus selectivity of separation is based on manipulating these variables.

* Corresponding author.

0021-9673/94/$07.00
SSDI 0021-9673(94)00126-T

An important drawback of ATPSs is the possible instability of the two phases being formed when compared to conventional chromatographic processes, where one of the phases is solid. To allow for the application and scale up of ATPSs this problem has to be overcome. Systems have to be robust so that variations in process streams will not destroy the system by causing the formation of one phase. By developing a model to describe and predict phase separation in the extraction and back extraction stages, based on experimental data, it can be possible to simulate and investigate regions of potential instability and thus allow for a robust process to be achieved. Such a process can be operated under steady state conditions which will also minimize batch to batch variations.

Thaumatin [2], a flavour enhancer, has been cloned into *E. coli* and yeast and has a molecular mass of 22 200 and an isoelectric point of around pH 10.5–11.0. In this study we have investigated the partition, in PEG phosphate systems, of pure thaumatin in the presence of *E. coli* contaminant proteins; extraction was followed by back extraction of the product into a new phosphate bottom phase and also back extraction combined with the recycle of the top PEG phase. The data for this has been used in a mathematical model to describe a continuous extraction process. We have also investigated the effect of the systems robustness (*i.e.* the stability ratio), on the partition of the *E. coli* soluble protein matrix.

2. Materials and methods

2.1. Materials

PEG with M_r 6000, dipotassium hydrogenphosphate and sodium dihydrogen phosphate were obtained from BDH, Leicester, UK. PEG with M_r 4000 and magnesium sulphate were obtained from Fluka, Buchs, Switzerland. Thaumatin was kindly donated by Four F Nutrition, North Yorkshire, UK. All other chemicals were analytical grade.

2.2. Growth and lysis of E. coli

A homogenate of *E. coli* K12 (in 50 m*M* Tris–HCl pH 7), was prepared by sonication. After centrifugation the supernatant (soluble fraction) was removed and the remaining cell debris (insoluble fraction) was washed and resuspended with 50 m*M* Tris–HCl at pH 7.

2.3. Preparation of phase systems

Systems had a final mass of 1.5 g and were prepared from stock solutions of PEG (50%, w/w), potassium phosphate (40%, w/w), magnesium sulphate (23%, w/w) and sodium chloride (25%, w/w). pH was set at 7 and stocks were stored in the cold and prior to use temperature equilibrated by standing at room temperature (20°C). For back extraction thaumatin, *E. coli* soluble fraction and cell debris were added to give a final concentration of 1 mg/ml. For the study on stability ratio (Fig. 1), systems had a final mass of 1.8 g and *E. coli* soluble fraction was added to give a final concentration of 0.5 mg/ml.

Systems were mixed thoroughly and centrifuged (3000 rpm, *ca.* 500 *g*, 3 min) to assist phase separation. Samples of 0.1 ml were removed from the top and bottom phases, diluted with deionised water and assayed for protein

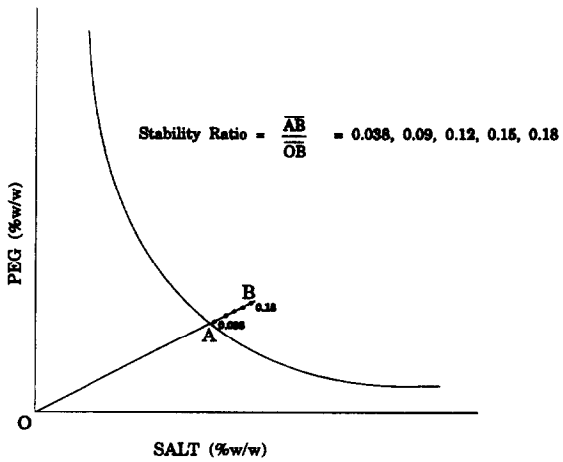

Fig. 1. Illustration of the concept of the stability ratio. B = System composition; A = point on the binodial; O = origin.

concentration. All experiments were carried out in duplicate.

2.4. Back extraction of thaumatin

A system with a phase composition of 12% (w/w) PEG, 13% (w/w) phosphate and 8.8% (w/w) NaCl and a phase volume ratio of 0.67 was used. Back extraction was achieved by the removal of the top phase and addition of a fresh bottom phase with no NaCl present. The back extraction was carried out three times into a fresh phosphate phase.

2.5. Back extraction and recycle of PEG

A system with a phase composition of 15.5% (w/w) PEG, 6% (w/w) phosphate and 8.8% (w/w) NaCl with a phase volume ratio of 0.2 was used. Recycling of PEG was achieved by removal of the top phase and addition of a fresh bottom phase with no NaCl present. The top phase was then removed and added to the original system where further NaCl and thaumatin, soluble fraction or insoluble fraction was added. This was repeated three times.

2.6. Protein determination

For the determination of protein concentration the modified Bradford assay (see ref. 6) was used. Standard curves of thaumatin and bovine serum albumin (BSA) were constructed which were linear within the range of protein concentration measured. Protein-free systems were used for control where the protein sample was replaced with 50 m*M* Tris–HCl at pH 7. The partition coefficient, *K*, corresponds to the ratio of protein concentration in the top and bottom phases.

3. Results and discussion

3.1. Partition of thaumatin and E. coli

When a particular ATPS has been selected for the separation and/or purification of a protein, the choice of the operating point in the phase diagram is important. If a chosen point is too far from the binodial, where concentrations of polymer and salt are high, the protein might precipitate out of solution. If the point is too near the binodial a small dilution of the system might cause a shift of the system composition to the left of the binodial and therefore the formation of a single phase where separation can not be achieved. The question of system robustness will depend to a large extent on the location of the operating point in relation to the binodial. This concept is reflected in the relative value of the tie-line length but, even simpler, in the value of the stability ratio as defined in Fig. 1. This ratio is conceptually similar to the tie-line length but is much easier to determine as it only requires a titration with water.

Cascone *et al.* [2] studied various systems for the separation of thaumatin from *E. coli* contaminant proteins. The results shown in Fig. 2 proved to have the best potential conditions to achieve a good separation. This particular system was therefore further analysed by reversed-phase chromatography to assess if the presence of *E. coli* proteins affects the partition of thaumatin. The *K* values obtained for thaumatin and *E. coli* proteins independently were maintained by the mixture of the proteins.

Fig. 2 shows the partitioning of pure

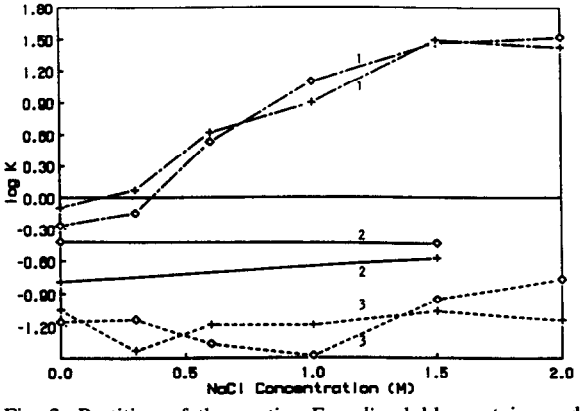

Fig. 2. Partition of thaumatin, *E. coli* soluble proteins and BSA in PEG 6000–phosphate systems as a function of NaCl concentration at (+) pH 7 and (◇) pH 9. 1 = Thaumatin; 2 = *E. coli*; 3 = BSA.

thaumatin, *E. coli* soluble proteins and BSA as a function of NaCl concentration at two pH values [2]. Log K for thaumatin increases dramatically as the NaCl concentration is increased to 1.5 M, while log K for *E. coli* soluble contaminants and BSA (a typical contaminant of mammalian cell culture), remains below 0. Exploiting this behaviour can allow for the selective separation of thaumatin at a high NaCl concentration where K for thaumatin is high ($K = 33$); further purification can be achieved by back extraction into a fresh phosphate bottom phase in the absence of NaCl thus decreasing the K value ($K = 0.53$). This can be achieved by diluting the NaCl from the top phase by using a large phosphate phase for back extraction. A similar but even more pronounced trend has recently been observed for the separation of α-amylase from its fermentation supernatant, where increasing the NaCl concentration increases K for α-amylase while K for the contaminants remains unaffected [7].

3.2. Partitioning of thaumatin and E. coli soluble and insoluble fractions

In order to assess the effect of processing a crude cell homogenate, without clarification of the cell debris, for the system described in Figs. 2 and 3 the protein partition coefficient at 8.8% NaCl (*ca.* 1.5 M) for the soluble, insoluble and combined fractions in the presence and absence of thaumatin (Table 1) was analysed. The partition coefficients for the *E. coli* fractions were all below 1 (log $K < 0$). In the presence of thaumatin, K for all fractions was somewhat lower than for thaumatin alone due to the presence of *E. coli* proteins, and in general little effect of the debris was observed.

3.3. Back extraction of thaumatin

To back extract thaumatin from the top PEG phase, dilution of NaCl present in this phase was necessary. A large phosphate bottom phase was used which lowered the concentration of NaCl in the new system and also allowed a large proportion of the thaumatin to be recovered (stage 2 in Fig. 3). Although K for thaumatin was relatively

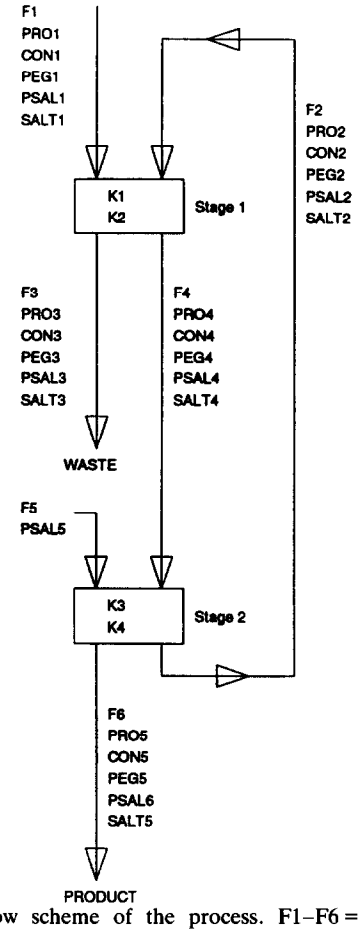

Fig. 3. Flow scheme of the process. F1–F6 = Flow-rates; PRO1–PRO5 = product protein concentrations; CON1–CON5 = contaminant protein concentrations; PEG1–PEG-5 = concentrations of PEG; PSAL1–PSAL6 = concentrations of phase forming salt (*e.g.* phosphate); SALT1–SALT5 = concentration of added salt (*e.g.* NaCl).

low, a proportion of thaumatin is still present in the top phase. Further purification was achieved in the following two back extraction steps (Table 2), where in the third back extraction step K for thaumatin was similar to that observed in Fig. 2 at 0% NaCl. This is evidently an effect of the decreasing concentration of NaCl in the second and third back extraction stages.

3.4. Recycle of PEG

The effect on the partition coefficients of pure thaumatin and protein measured in the presence of soluble and insoluble fractions when PEG was

Table 1
Partition of *E. coli* homogenate and thaumatin in a PEG 6000 (12%, w/w), potassium phosphate (13%, w/w) and NaCl (8.8%, w/w) system at pH 7

	Log *K*	
	+Thaumatin	−Thaumatin
Thaumatin only	1.5	–
Insoluble fraction	1.2	−0.4
Soluble fraction	1.3	−0.4
Combined fractions	1.3	−0.7

+Thaumatin = System with thaumatin; −Thaumatin = system without thaumatin.

recycled back to stage 1 (Fig. 3), is shown in Table 3. This was carried out batch-wise where three separations were performed. Manipulation of the phase volume ratio (0.2) enabled the dilution of NaCl from the system and a large proportion of the thaumatin to be recovered. It should be noted that the only loss of thaumatin is in the bottom phase at stage 1 as the thaumatin not recovered in stage 2, is recycled with the PEG phase. Table 3 shows that recycling of the PEG phase did not have a dramatic effect on the partition coefficients.

3.5. Partition of *E. coli* soluble proteins contaminants

To permit the selection of an ATPS for the separation of a protein cloned in *E. coli*, it was decided to characterise the behaviour of *E. coli* soluble proteins in polymer–salt ATPSs under different conditions; in particular the effect of

Table 2
Partition of thaumatin after back extraction in a PEG 6000 (12%, w/w), potassium phosphate (13%, w/w) and NaCl (8.8%, w/w) at pH 7

Log *K*			
Extraction, stage 1	B.E.1, stage 2	B.E.2, stage 3	B.E.3, stage 4
1.5	0.4	0.2	−0.3

B.E. = Back extraction.

Table 3
Partition of thaumatin and *E. coli* soluble and insoluble fractions after back extraction and recycle of the top PEG phase in a PEG 6000 (15.5%, w/w), potassium phosphate (6%, w/w) and NaCl (8.8%, w/w) system at pH 7

	Log *K*		
	Thaumatin	Soluble fraction	Insoluble fraction
Extraction	1.6	−0.3	0.2
B.E.1	−0.1	−0.1	0.0
Recycle	1.9	−0.2	−0.4
B.E.2	0.1	−0.1	−0.1
Recycle	1.7	−1.0	−0.2
B.E.3	0	0.1	0.1

NaCl and the stability ratio (Fig. 1). If known under which conditions the bulk of the *E. coli* matrix proteins partition into one phase (*e.g.* bottom phase), then a choice of systems will be available in which to test whether the target protein can be partitioned to the opposite phase (*e.g.* top phase).

Figs. 4 and 5 show the partition of *E. coli* soluble protein contaminants at different stability ratios as a function of NaCl concentration in PEG–phosphate and PEG–sulphate systems. In phosphate systems log K_{app} (K is "apparent" as it is an average K for *E. coli* soluble proteins) for the soluble proteins at 0% NaCl, increases as the stability ratio increases. At 4% and 8.8% NaCl,

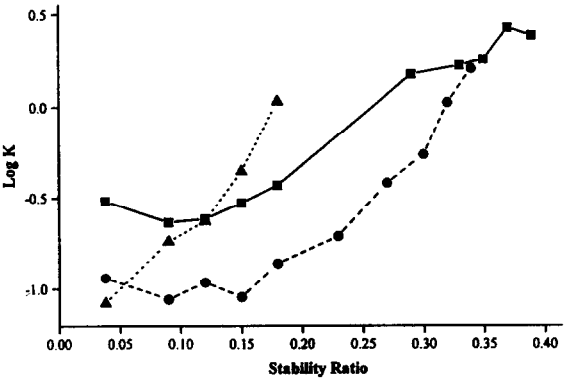

Fig. 4. Effect of the stability ratio on the partition of *E. coli* soluble proteins as a function of NaCl concentration in PEG 4000 phosphate systems at pH 7. ▲ = 0% NaCl; ● = 4% NaCl; ■ = 8.8% NaCl.

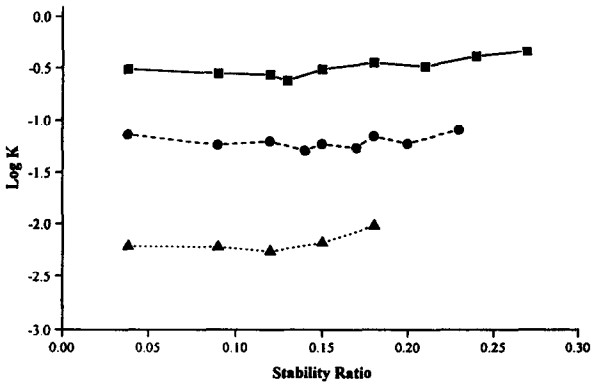

Fig. 5. Effect of the stability ratio on the partition of *E. coli* soluble proteins as a function of NaCl concentration in PEG 4000 sulphate systems at pH 7. Symbols as in Fig. 4.

log K_{app} remains low up to a stability ratio of 0.18 and then increases in both cases as the stability ratio increases. On the other hand, for sulphate systems (Fig. 5) log K_{app} for the soluble proteins at 0% NaCl remains relatively constant as the stability ratio is increased. The same trend is observed for 4 and 8.8% NaCl. In both cases NaCl causes an increase in K but in the sulphate systems the value of K is still well below 1 over the range of stability ratios studied. In these systems the stability ratio (tie-line length) seems to have little effect on K_{app} for *E. coli* soluble proteins. For both systems at 0, 4 and 8.8% NaCl the respective systems, at the same distance from the binodial, have approximately the same tie-line length. Log K_{app}, however, increases as a function of NaCl concentration suggesting that the presence of NaCl is changing the character of the phases and thus causing an increase in the value of K. In other systems this effect has been attributed to an increase in the hydrophobic character of the system [8]. These systems, with a partition coefficient lower than 1 (log $K < 0$) are all suitable for partitioning target recombinant proteins into the top PEG phase.

4. Mathematical modelling and simulation for process robustness and optimisation

A mathematical model has been developed to describe the continuous, steady state operation of the flow sheet described in Fig. 3. The model is based on fundamental mass balances of the main components and phase equilibrium data in the form of equations of the binodial and tie-lines. SPEEDUP, a powerful equation based simulator, was used to solve the model [9], which has been fitted to two sets of data: (i) separation of *E. coli* homogenate proteins, and (ii) separation of α-amylase from *Bacillus subtilis* supernatant. Here we will describe the application of the model to the former data.

4.1. Process description

The process illustrated in Fig. 3 is a two-stage continuous separation system based on the results for thaumatin discussed above. The first stage is the main separation step where F1 is *E. coli* lysate with or without cell debris together with the protein product. As above, the protein product is partitioned to the top phase and then back extracted into a fresh bottom phase, the top phase is recycled and added to a fresh bottom phase. Recycle of the PEG phase minimises product loss, which in turn, increases process yield in comparison to a single stage batch extraction.

4.2. Phase equilibrium

To allow for a robust separation process, prediction of the phase diagram is essential. Although precise equations do not exist, an exponential empirical curve (Eq. 1) can be fitted to the experimental data, as shown in Fig. 6.

$$Y = A \exp(BX) \quad (1)$$

where Y is the PEG concentration, X is the phosphate concentration and A and B are system constants. For a binodial of PEG 4000–phosphate $A = 79\%$ (PEG) and $B = -0.26$ (1/% salt).

4.3. Phase separation

Phase separation was modelled by calculating the total amount of each component entering a stage divided by the sum of the flow-rates. This provided the point in the phase diagram corre-

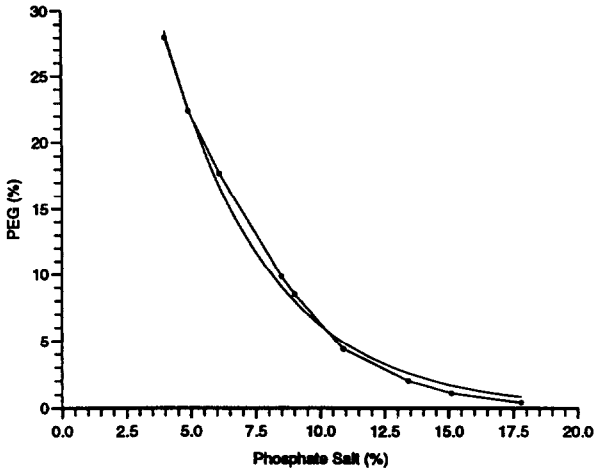

Fig. 6. Comparison graph of the binodial of the PEG 4000–phosphate system determined experimentally *versus* the exponential curve.

sponding to that stage. This point has an associated tie-line whose equation is calculated and then the intersections of the tie-line with the binodial curve are found, *i.e.* the composition of PEG in state 1 (Fig. 3), YM1 is:

$$YM1 = \frac{F1 \cdot \mathrm{PEG1} + F2 \cdot \mathrm{PEG2}}{F1 + F2} \quad (2)$$

where $F1$ and $F2$ are the flow-rates entering a stage and PEG1 and PEG2, the compositions of PEG in these streams.

From the point on the diagram, the associated tie-line is found which has the equation for a straight line:

$$Y = MX + C \quad (3)$$

where Y and X are the PEG and phase forming salt (*e.g.* phosphate) compositions, M is the slope and C is a constant [9].

4.4. Prediction of system performance

The model has been used to study the effects of input variables on system robustness, purity and yield over a range of conditions. The variables were increased/decreased until the model broke one of the following constraints: (1) the system entering the homogeneous region, left of the binodial curve; (2) the tie-line moving outside of the binodial curve (*i.e.* to the right). Examples of these simulations are shown in Fig. 7a and b.

4.5. System robustness

To study system stability and robustness (defined by operating at a predefined distance from the binodial), movements in the tie-line positions were studied for variations in PEG1, PSAL1, F5 and PSAL5 (see Fig. 3) over their working range. Fig. 8 represents a typical physico-chemi-

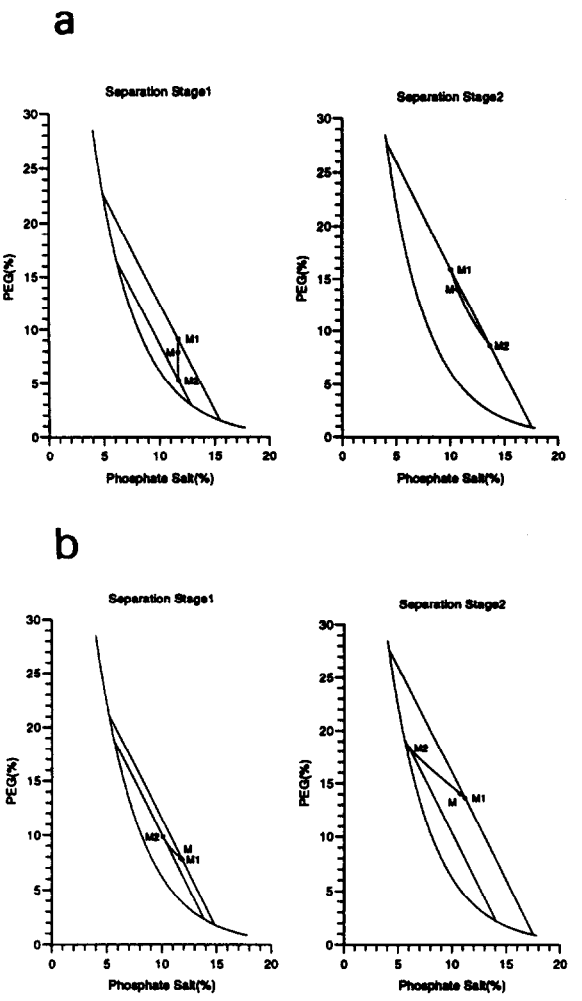

Fig. 7. (a) Tie-line movements about the basal value (M) for step change variations in the concentration of PEG into stage 1. (b) Tie-line movements about the basal value (M) for step change variations in the flow-rate of phosphate salt into stage 2.

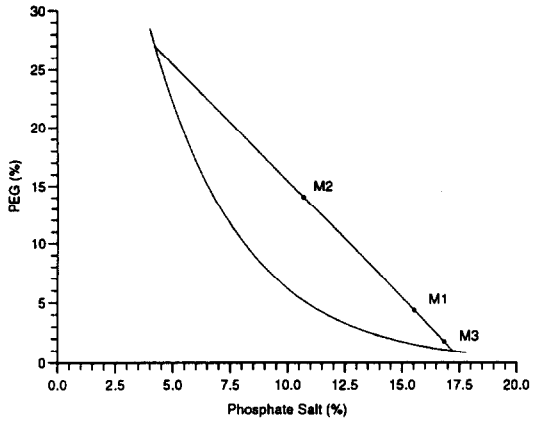

Fig. 8. Typical physico-chemical manipulation graph of the phase ratio.

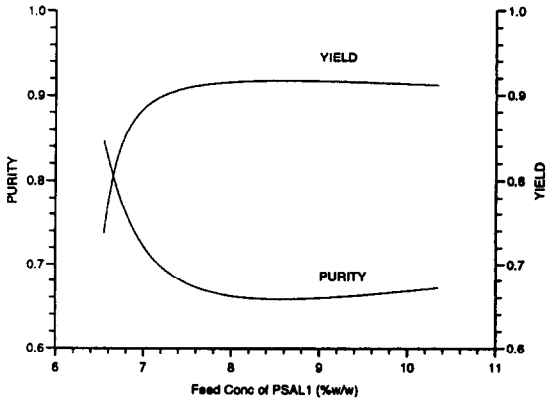

Fig. 9. The effect of varying the concentration of phosphate salt into stage 2 on the overall purity and yield of the process.

cal manipulation of the phase ratio within the same tie-line. This means equal composition of top and bottom phases but changing the phase ratio (equal value for $K_{\text{protein product}}$ and $K_{\text{contaminants}}$). The point M1, on Fig. 8 represents an "ideal" distance from the binodial. M2 is judged to be too far from the binodial; a high concentration of PEG and salt could lead to protein precipitation. Meanwhile M3 is "too close" to the binodial and the system is, therefore, not robust and small variations in system composition could lead to the system falling into the one phase region left of the binodial.

Fig. 7a illustrates the effect of varying the PEG1 concentration entering stage 1. Increasing the PEG 1 concentration (moving from M to M1) leads to a larger top phase and greater product separation in the first stage. This also leads to a smaller bottom phase in the second stage and lower product recovery. Moving from M to M2 represents decreasing the PEG1 concentration giving a smaller top phase in the separation. In the second stage the bottom phase becomes larger and product recovery is increased.

Illustrated in Fig. 7b is the effect of varying the F5 flow-rate (PO_4^{3-}) into stage 2. Increasing the flow-rate moves M to M1 representing a larger volume of PSAL in stage 2 and thus a greater bottom phase. The increased PSAL in stage 2 is recycled into stage 1 also giving a larger bottom phase. Decreasing the flow-rate F5 moves M to M2 which has a drastic effect in stage 2 leading to a very small bottom phase and hardly any product recovery. In stage 1 the bottom phase decreases due to lower PSAL recycle and aids product separation.

4.6. Purity and yield

The effect of varying the four variables, PEG1, PSAL1, F5 and PSAL5, on purity and yield in the overall process was studied. The effect of varying the concentration of phosphate salt into stage 2 is shown in Fig. 9. The purity and yield are fairly constant at high concentrations but there is a dramatic change at the lower range. The region below 7.5–8% PSAL1 concentration is very unstable and this can be attributed to the rapid decrease in tie-line length as we approach the binodial in stage 1. The high partition coefficient and minimal product loss (only in the bottom phase in stage 1) is the principal reason for the high yield observed (0.8–0.92). To improve product purity lowering the partition coefficient for the contaminants, has to be achieved experimentally.

5. Conclusions

The partition of the protein thaumatin in the presence of *E. coli* contaminant proteins has

been studied. Extraction of thaumatin was followed by back-extraction of the product into a new phosphate phase and also back extraction combined with recycle of the top PEG phase. When partitioned in the absence or presence of insoluble cell debris (whole cell homogenate) little effect on the partitioning of thaumatin or the soluble *E. coli* proteins was found. A back extraction step that allowed for dilution of the NaCl (necessary to obtain a high partition coefficient for thaumatin in the first extraction stage) was successful in extracting thaumatin back into a heavy phosphate phase. A second and third back extraction stage further increased this yield.

After back extraction the top PEG phase was successfully recycled to the first extraction. This minimised any loss of thaumatin as it is recycled together with the PEG. The stability ratio (tie-line length) had an effect on the average partition coefficient of *E. coli* soluble proteins in PEG/phosphate systems but in PEG sulphate systems K_{app} was independent of this ratio. An increase in NaCl resulted in an increase in K_{app} but this was always below 1 making these systems suitable for partitioning recombinant proteins into the top PEG phase. A mathematical model, that describes the continuous steady-state operation of extraction and back extraction has been developed: it proved to be an important tool to investigate regions of potential instability in the phase formation and to predict purity and yield after extraction and back extraction.

6. Acknowledgements

This work was partially supported by SERC and Smith, Kline Beecham to whom thanks are due.

7. References

[1] Y. Guan, X.-Y. Wu, T.E. Treffy and H. Lilley, *Biotechnol. Bioeng.*, 40 (1992) 517.
[2] O. Cascone, B.A. Andrews and J.A. Asenjo, *Enzyme Microb. Technol.*, 13 (1991) 629.
[3] J.A. Asenjo, T. Franco, A.T. Andrews and B.A. Andrews, in M.D. White, S. Reuveny and A. Shaffermanin (Editors), *Biologicals from Recombinant Microorganisms and Animal Cells: Production and Recovery*, VCH, Weinheim, 1991, p. 439.
[4] H. Hustedt, K.H. Kroner and N. Papamichael, *Process Biochem.*, 23 (1988) 129.
[5] M.R. Kula, in G. Durand (Editor), *Extraction Processes, Proceedings of the 8th International Biotechnology Symposium*, Vol. 1, Société Française de Microbiologie, Paris, 1988, p. 612.
[6] J.J. Sedmak and S.E. Grossberg, *Anal. Biochem.*, 79 (1977) 544.
[7] A.S. Schmidt, A.M. Ventom and J.A. Asenjo, *Enzyme Microb. Technol.*, 16 (1994) 131.
[8] J.A. Asenjo, A.S. Schmidt, F. Hachem and B.A. Andrews, *J. Chromatogr. A*, 668 (1994) 47.
[9] S.L. Mistry, J.A. Asenjo and C.A., Zaror, *Bioseparation*, 3 (1993) 343.