ELSEVIER

Journal of Chromatography A, 668 (1994) 47–54

JOURNAL OF
CHROMATOGRAPHY A

Model for predicting the partition behaviour of proteins in aqueous two-phase systems

J.A. Asenjo *, A.S. Schmidt, F. Hachem, B.A. Andrews
Biochemical Engineering Laboratory, University of Reading, Reading RG6 2AP, UK

Abstract

The effect of protein hydrophobicity, charge, molecular mass and concentration has been studied in poly-(ethylene glycol) (PEG)–phosphate and PEG–dextran aqueous two-phase systems in the presence and absence of NaCl for several model proteins. The surface hydrophobicity of the protein measured by precipitation correlated well with the partition coefficient in PEG–salt systems at high levels of NaCl. The charge of proteins also has an important effect on partition; this is expected to be more pronounced at lower NaCl concentrations. For molecular mass a tendency was found in PEG–dextran systems at low NaCl concentrations. No clear tendency was observed in the PEG–salt systems. The solubility of the protein in the phases also affects its partition behaviour. This behaviour was fitted to a saturation type equation for α-amylase in each of the phases of a PEG–phosphate system

1. Introduction

Aqueous two-phase systems (ATPSs) have important potential in downstream processing as a large-scale continuous operation for the separation of proteins and removal of contaminants from fermentations as they can process streams continuously and rapidly produce an initial separation including separation of whole cells or cell debris as in the case of intracellular proteins. However, the wider implementation in many applications has been restrained in part by the limited predictability of the partition behaviour of proteins in ATPSs. Factors and mechanisms that cause the uneven distribution of biomolecules are poorly understood. Fundamental theories of protein partition derived from classical polymer solution thermodynamics are being developed [1–4]. Although the present models provide important information, no comprehensive theory currently exists to guide the design of systems for separation of specific mixtures of proteins. The value of the partition coefficient, K (defined as the ratio of the protein concentration in the lighter phase to that in the heavier phase) relies on the physico-chemical properties of the target protein and contaminants (*e.g.* hydrophobicity, charge, molecular mass) and their interactions with those of the chosen system (*e.g.* composition, ionic strength, addition of specific salt ions, pH). Depending on the manipulation of these system parameters the target protein and contaminants can be partitioned selectively.

This paper describes preliminary work on a model to correlate the physico-chemical properties of a protein to its partition coefficient in ATPSs using a modified group contribution approach:

* Corresponding author.

0021-9673/94/$07.00 © 1994 Elsevier Science B.V. All rights reserved
SSDI 0021-9673(94)00102-F

$$K = K_{hphob} \cdot K_{el} \cdot K_{size} \cdot K_{sol} \cdot K_{aff} \quad (1)$$

where K_{hphob}, K_{el}, K_{size}, K_{sol} and K_{aff} denote the contribution to the overall partition coefficient by hydrophobicity, electrostatic forces, size, solubility and affinity, respectively. Changes in conformation will usually affect K_{el}, K_{hphob}, K_{sol} or even K_{aff}. Not all the properties are equally important, this depends on the type of system used. The initial results of our investigation into the effects of hydrophobicity, charge, molecular mass and solubility are described.

2. Materials and methods

2.1. Materials

Poly(ethylene glycol) (PEG) with molecular masses of 4000 and 8000 was obtained from Fluka. Dextran T500 was purchased from Pharmacia, Sweden. All other chemicals were analytical grade. Thaumatin was a gift from Four-F Nutrition, Northallerton, UK. Subtilisin and α-amylase (both from *Bacillus subtilis*) were obtained from Boehringer Mannheim, Germany. Other proteins were purchased from Sigma. α-Chymotrypsinogen A was from bovine pancreas, α-lactalbumin and β-lactoglobulin from bovine milk, invertase from Bakers yeast and lysozyme from chicken egg white. Molecular mass and p*I* values of the proteins used in this study are shown in Table 1.

2.2. Preparation of phase systems

Phase systems were prepared from stock solutions of PEG (50%, w/w), phosphate (40%, w/w) and NaCl (25%, w/w). The phosphate stock solution consisted of a mixture of K_2HPO_4 and NaH_2PO_4 at the appropriate pH. Stock solutions were stored at 4°C. Before use the temperature of all solutions was equilibrated by standing at room temperature. All partition experiments were done at 20°C. Total protein was added to the systems at a final concentration of 1 g/l unless otherwise stated. Samples of top and bottom phase were assayed for protein concentration.

2.3. Protein assays

Total protein was measured by the modified Lowry assay [5] using bicinchoninic acid (BCA) at 37°C or by the modified Bradford dye-binding assay [6]. The BCA assay kit was supplied by Pierce Europe, Netherlands, standard curves were prepared for each protein. Blank systems without protein were used as reference and no interference from phase components was observed.

α-Amylase concentration was assayed as activity as described previously [7].

Table 1
Molecular mass and p*I* of the model proteins

Proteins	Abbreviation	M_r	p*I*
α-Amylase	Amy	50 000	5.0
Bovine serum albumin	BSA	67 000	4.7
α-Chymotrypsinogen A	Chy	23 600	8.9
Conalbumin	Con	77 000	5.9
Invertase	Inv	270 000	3.4
α-Lactalbumin	Lac	17 400	5.1
β-Lactoglobulin A	Lag	37 100	5.1
Lysozyme	Lys	14 300	10.3
Subtilisin	Sub	27 500	8.4
Thaumatin	Tha	28 000	10.8

2.4. *Reversed-phase chromatography*

A Nucleosil 300-7 C_{18} column (25 cm × 4.6 mm I.D.) was used with a Perkin-Elmer Binary LC pump and a Rheodyne Model 7125 valve. Gradient elution at a flow-rate of 1 ml/min was used. Further experimental details have been previously described [8].

2.5. *Hydrophobic interaction chromatography (HIC)*

A fast protein liquid chromatography (FPLC) system (Pharmacia) was used with an HR 5/5 (1 ml) column with fast flow Phenyl-Sepharose (Pharmacia). Further details can be found in ref. 8.

2.6. *Precipitation of proteins*

Proteins were dissolved in sodium phosphate buffer (0.05 *M*, pH 7) at a concentration of 2 g/l. To a measured volume of each solution solid ammonium sulphate was added slowly with continuous stirring. The solutions were left to equilibrate for 15 min at 24°C. After centrifugation (25 000 *g*, 20 min) a sample was taken from the supernatant and assayed for protein concentration. More ammonium sulphate was added to the rest of the supernatant making the saturation higher; after each addition the supernatant was assayed for protein concentration. This was repeated until no protein could be precipitated. A blank was prepared and treated in the same way to account for the interference of ammonium sulphate in the protein assays.

3. Results and discussion

The systems chosen for this study, PEG 4000–phosphate, PEG 8000–phosphate and PEG 4000–dextran, were selected on their ability to illustrate differences in physico-chemical properties of proteins. Initially PEG 8000–phosphate systems at several NaCl concentrations were used to correlate the surface hydrophobicity of proteins measured by different techniques, RP-HPLC, HIC and ammonium sulphate precipitation. Five levels of NaCl were added to the systems [0, 0.3, 4.8, 9.6 and 17.6% (w/w)] as very different partition behaviours have been observed at these concentrations [7,9]. The difference in partition behaviour observed at high NaCl concentrations in PEG–phosphate systems [7,9] has been attributed to hydrophobicity. Proteins were chosen with clear differences in hydrophobicity, molecular mass (14 300–270 000) and surface charge (p*I* 3.4–10.8).

3.1. *Hydrophobicity*

In order to develop possible correlations for prediction of partitioning behaviour of proteins five model proteins (α-lactalbumin, β-lactoglobulin A, bovine serum albumin, conalbumin and lysozyme) with different physico-chemical properties were initially chosen. Evaluation of the relationship between hydrophobicity and behaviour of the proteins in different aqueous two-phase systems was investigated in detail by measuring surface hydrophobicity by HIC and precipitation/solubility in addition to reversed-phase chromatography (RP-HPLC) which although separating proteins and peptides based on their hydrophobicity can denature larger proteins. Molecular masses and p*I* values of these proteins are shown in Table 1.

For the five proteins a poor correlation was found between the hydrophobicity measurements evaluated as retention time in an RP-HPLC column and HIC and log *K* in all aqueous two-phase systems investigated which included PEG–dextran and PEG–phosphate with and without the addition of NaCl [8]. Some of the better correlations found (in PEG–phosphate systems with NaCl) are shown in Figs. 1 and 2 for RP-HPLC and HIC; however, this correlation was still rather poor. The hydrophobicity, log *P*, measured by RP-HPLC is expressed as the log of the volumetric fraction of acetonitrile (*P*) at which the proteins elute from a C_{18} column. The hydrophobicity measured by HIC is expressed as the log of (1/molar fraction of ammonium sulphate) at which the proteins elute from a Phenyl-Sepharose column using a linear

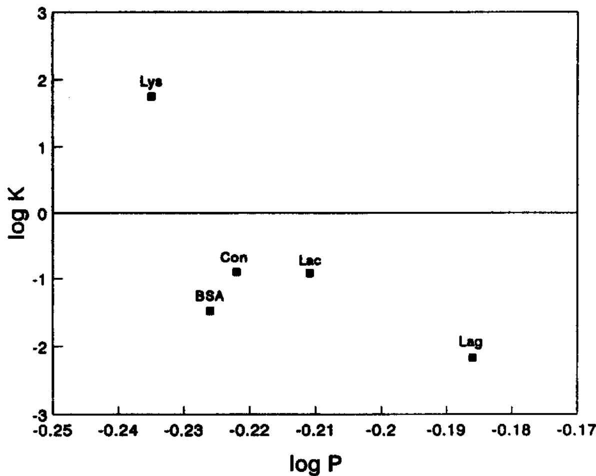

Fig. 1. The relationship between the partition coefficient (log K) of several model proteins in PEG 8000 (8%, w/w)–phosphate (12%, w/w) system with 9.6% (w/w) NaCl at pH 7 and the hydrophobicity of the proteins (log P) measured by RP-HPLC on a C_{18} column as the volumetric fraction of acetonitrile at which they eluted. Abbreviations as in Table 1.

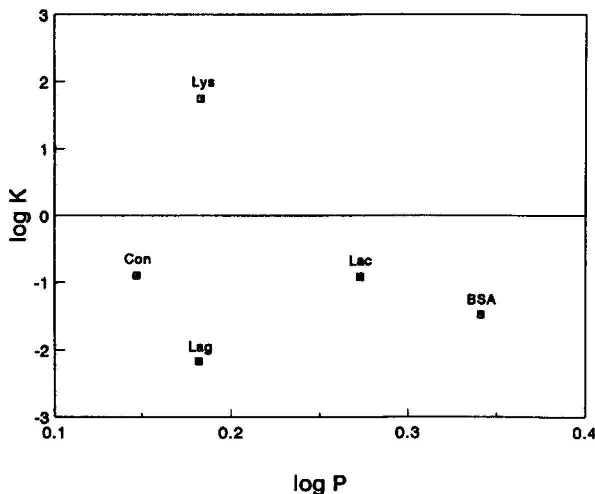

Fig. 2. The relationship between the partition coefficient (log K) of several model proteins in PEG 8000 (8%, w/w)–phosphate (12%, w/w) system with 9.6% (w/w) NaCl at pH 7 and the hydrophobicity of the proteins (log P) measured by HIC–FPLC on a Phenyl-Sepharose columns. The hydrophobicity is expressed as the log of (1/molar fraction of ammonium sulphate) at which the proteins eluted from the column using a linear gradient from 1.5 to 0.0 M ammonium sulphate. Abbreviations as in Table 1.

gradient from 1.5 to 0.0 M ammonium sulphate in 0.05 M sodium phosphate buffer (pH 7).

An alternative method to evaluate a proteins hydrophobicity in solution is by measuring the point at which precipitation begins in a typical ammonium sulphate precipitation curve [10,11]. In a typical protein precipitation graph, as shown in Fig. 3 for β-lactoglobulin, the solubility of the protein (and thus its hydrophilicity) can be expressed by point m* which is the point at which the protein starts to precipitate. Thus the hydrophobicity, P, can be represented as 1/m*. Fig. 4 shows the relationship between log K and the hydrophobicity of the five model proteins measured by ammonium sulphate precipitation (1/m*). It is clear from Figs. 1, 2 and 4 that the inverse of the solubility (point 1/m*) gave a good correlation of log K with log P (log P = log 1/m*) whereas the hydrophobicity evaluated by RP-HPLC and HIC did not. The reason for this seems to be that solubility represents a measurement of a protein's "average" surface hydrophilicity (or hydrophobicity) in "bulk solution" similar to that measured in certain ATPSs whereas in HIC the hydrophobicity/hydrophilicity distribution on the protein's surface (*e.g.* hydrophobic patches) probably plays a greater role in the interaction of the protein with the

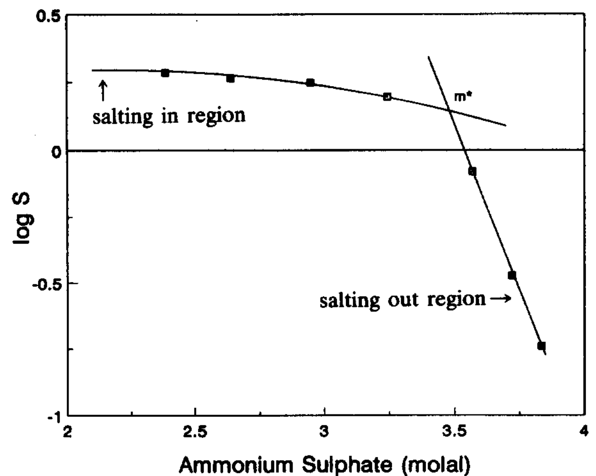

Fig. 3. The solubility/precipitaion curve in ammonium sulphate of β-lactoglobulin A at 25°C. The intercept of the two fitted lines x value represent the solubility, m*.

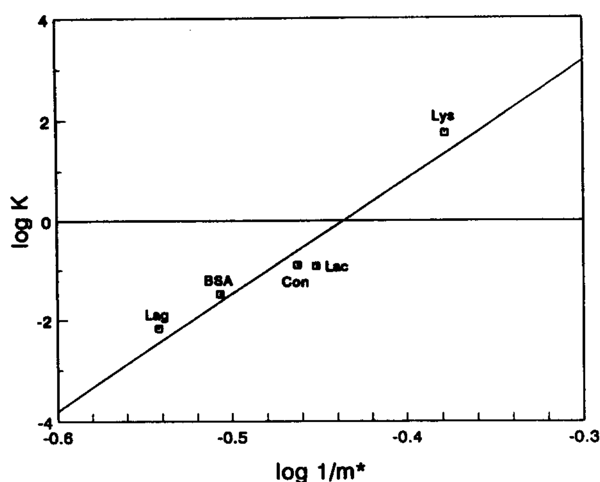

Fig. 4. The relationship between the partition coefficient (log K) of several model proteins in PEG 8000 (8%, w/w)–phosphate (12%, w/w) system with 9.6% (w/w) NaCl at pH 7 and the hydrophobicity of the proteins (log P) measured by ammonium sulphate precipitation [log $P = \log (1/m^*)$] ($r = 0.957$). Abbreviations as in Table 1.

hydrophobic matrix. In the case of RP-HPLC this is probably even more evident as the solvent used may partially denature larger proteins exposing some of the internal amino acids and thus affecting the hydrophobic/hydrophilic behaviour.

A model for the prediction of partitioning of proteins in certain ATPSs that mainly exploit hydrophobicity has thus been proposed [4,8]:

$$\log K = D\Delta W \log(P/P_0) \quad (2)$$

or

$$\log K = D\Delta W \log P - D\Delta W \log P_0 \quad (3)$$

in this equation $D\Delta W$ describes its resolution power in which ΔW represents the difference in concentration of one component between the top and bottom phases (*e.g.* PEG; this takes into account the effect of the tie-line length). D has been called the discrimination factor. This equation is similar to that proposed for smaller molecules and peptides [4]. According to this model, PEG–phosphate systems with added NaCl mainly exploit the surface hydrophobicity for partitioning of different proteins. Table 2

Table 2
The calculated values of $D\Delta W$ and the intrinsic hydrophobicity (log P_0) of PEG 8000 (8%, w/w)–phosphate (12%, w/w) systems with varying concentration of NaCl at pH 7

NaCl (%, w/w)	$D\Delta W$	log P_0
0	5.4	−0.23
0.48	8.3	−0.29
4.8	14.9	−0.38
9.6	22.4	−0.45
17.6	22.8	−0.45

shows the calculated values of $D\Delta W$ and the intrinsic hydrophobicity, log P_0, of the ATPSs used at 20°C. Clearly the systems with higher concentrations of NaCl give a higher resolution to exploit the proteins hydrophobicity in partitioning which is given by the value of the slope, $D\Delta W$, in Fig. 4.

This behaviour is presently being investigated for a larger number of proteins (listed in Table 1). Fig. 5 shows the partition behaviour (log K) and hydrophobicity ($1/m^* = P$) of these proteins as measured by solubility (m^*) during ammonium sulphate precipitation in a PEG 4000–phosphate system with 8.8% NaCl.

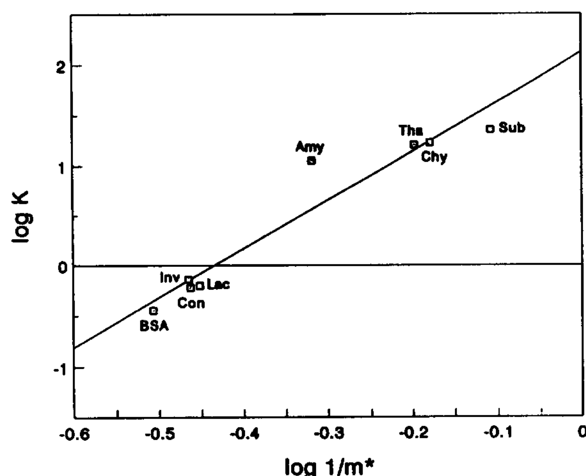

Fig. 5. The relationship between the partition coefficient (log K) of several model proteins in PEG 4000 (13%, w/w)–phosphate (10.7%, w/w) system with 8.8% (w/w) NaCl at pH 7 and the hydrophobicity of the proteins (log P) measured by ammonium sulphate precipitation [log $P = \log (1/m^*)$] ($r = 0.965$). Abbreviations as in Table 1.

3.2. Charge

Electrical potential differences between the phases of an aqueous two-phase system will affect the partition coefficient of proteins according to their charge. This has been demonstrated in PEG–dextran systems with addition of small concentrations of salts whose charged ions will partition between the phases (*e.g.* 0.1 *M* NaCl or 0.05 *M* Na_2SO_4) and also by manipulating the pH of the system [12,13]. In PEG–salt systems a change in the pH can affect the value of *K*. Fig. 6 shows the increase in partition coefficient with pH of α-amylase in a PEG–phosphate system. As the bottom phase has a very high concentration of salt, it appears that it is the pH of the top phase that will determine partitioning due to electrostatic effects. Theory has been recently developed to describe this effect as a function of the pH difference between the phases [14]. α-Amylase has a p*I* of 5. As pH was increased the protein becomes more negatively charged following its titration curve. In Fig. 6 there was a small increase in tie-line length as pH was increased but the effect shown is mainly due to pH change. This clearly suggests that the top phase has a higher density of positive charges. This behaviour is presently being investigated following titration curve analysis in several ATPSs for most of the proteins shown in Table 1.

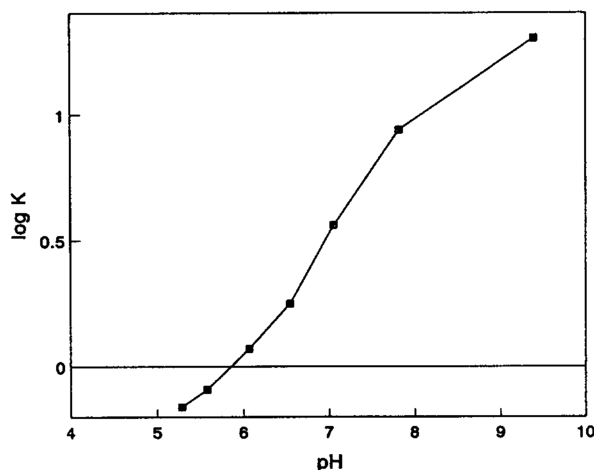

Fig. 6. Effect of system pH on partition coefficient (log *K*) of α-amylase in PEG 4000 (10%, w/w)–phosphate (11.5%, w/w) systems.

3.3. Molecular mass

The effect of protein molecular mass on its partition coefficient has been described in PEG–dextran systems with a low concentration of added NaCl [1,15]. This work is now being extended to include different ATPSs and proteins with a wider range of molecular masses. A correlation has been obtained at low concentrations of NaCl [1].

$$\log K_{size} = \beta + \alpha M_r \quad (4)$$

where α and β are constants depending on system properties.

Fig. 7A and B show the correlation of molecular mass with partition coefficient for proteins in PEG–dextran and PEG–phosphate systems, respectively. In the PEG–salt system there was no apparent correlation of molecular mass and log *K*. However, the correlation is far better in the PEG–dextran system as has been previously observed [1,15].

3.4. Protein concentration

As the concentration of protein in an ATPS is increased, a solubility limit in each of the phases is reached often leading to saturation in that phase [7]. True partitioning is only observed at low protein concentration (*e.g.* <2 g/l) depending on the protein [9]. This behaviour results in changes in the value of *K* at higher concentrations. Precipitation of the protein at the interphase is also usually observed. This phenomenon is demonstrated in Fig. 8 which shows the concentration of α-amylase in each of the phases as well as the partition coefficient as a function of overall concentration in the system. It can be seen that the α-amylase reaches saturation in the heavy, phosphate-rich bottom phase at a relatively low protein concentration (*ca.* 2 g/l), far lower than in the top, PEG phase (*ca.* 4–5 g/l). This results in a change in the "appar-

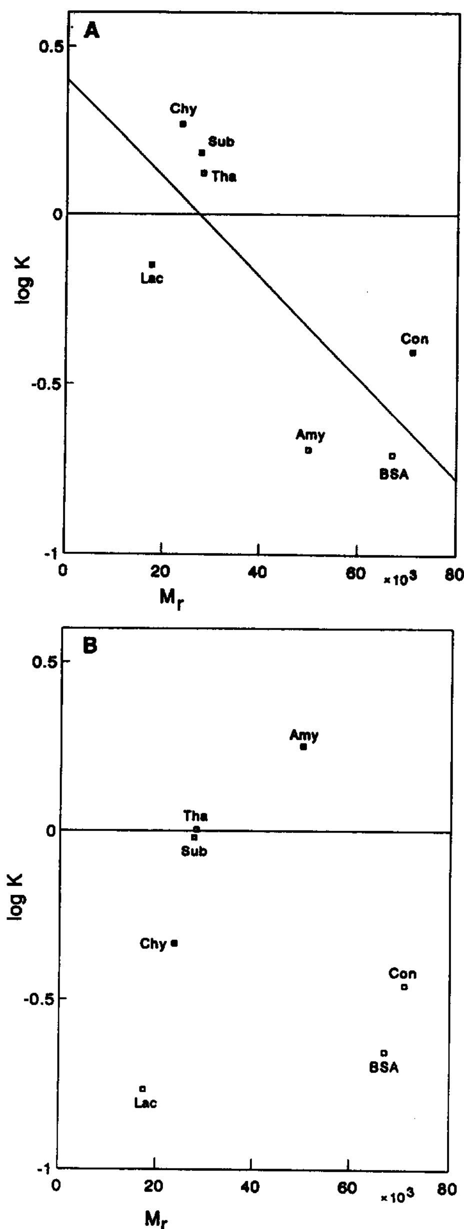

Fig. 7. Effect of protein molecular mass on partition behaviour of proteins in ATPS with 0.6% (w/w) NaCl at pH 7. (A) PEG 4000–dextran ($r = 0.8$); (B) PEG 4000–phosphate.

ent" partition coefficient (higher at higher protein concentration).

The concentration of protein in each phase has been satisfactorily modelled with a "saturation" type curve for α-amylase (Fig. 8).

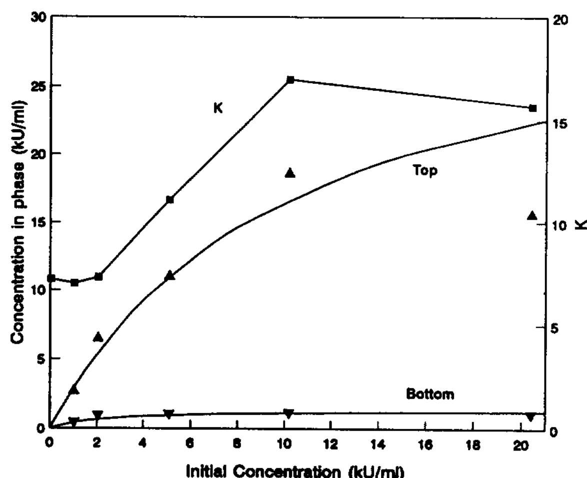

Fig. 8. Effect of overall (initial) system concentration of α-Amylase (1800 U/mg) on the saturation in the two phases (top and bottom) and partition coefficient, K (■) in PEG 4000 (10%, w/w)–phosphate (11.5%, w/w) system with 3% (w/w) NaCl at pH 7. 1 kU = 1000 U.

$$C_{\text{phase}} = C_M C_i/(C_0 + C_i) \quad (5)$$

where C_{phase} and C_i denote the concentration in the phase (U/l) and the overall protein concentration (U/l) respectively. C_M and C_0 are constants. The partition behaviour of α-amylase can thus be expressed as $K = C_{\text{top phase}}/C_{\text{bottom phase}}$ using Eq. 5 for the concentration in each of the phases.

4. Conclusions

The effect of protein hydrophobicity, charge, molecular mass and concentration has been studied in PEG–phosphate and PEG–sulphate ATPSs in the presence and absence of NaCl for several model proteins. The proteins' hydrophobicity evaluated both by RP-HPLC and HIC did not give a good correlation with partition coefficients in several ATPSs. The surface hydrophobicity of the protein measured by solubility correlated well with the partition coefficient in PEG–salt systems, particularly at high levels of NaCl (5–18%). The charge of proteins also has an important effect on partition; this is expected

to be more pronounced at lower NaCl concentrations. As pH was increased (5.5–9.5) the partition coefficient of α-amylase also increased suggesting a higher concentration of positive charges in the top phase. For molecular mass a tendency was found in PEG–dextran systems at low NaCl concentrations. No clear tendency was observed in the PEG–salt systems. The solubility of the protein in the phases affects its partition behaviour. It was found that α-amylase reaches saturation in the heavy, phosphate-rich phase at relatively low protein concentration, far lower than in the top, PEG-rich phase, results in an increase in the "apparent" partition coefficient. This behaviour was fitted to a saturation type equation for α-amylase in each of the phases of a PEG–phosphate system.

5. Acknowledgements

This work was partially supported by a grant from the European Community to whom thanks are due. Financial assistance from the Hariri Foundation (F.H.) and the European Science Foundation (A.S.S.) is also gratefully acknowledged.

6. References

[1] J.N. Baskir, T.A. Hatton and U.W. Suter, *Biotechnol. Bioeng.*, 34 (1989) 541.
[2] A.D. Diamond and J.T. Hsu, *Biotechnol. Bioeng.*, 34 (1989) 1000.
[3] N.L. Abbott, D. Blankschtein and T.A. Hatton, *Bioseparation*, 1 (1990) 191.
[4] M.A. Eiteman and J.L. Gainer, *Bioseparation*, 2 (1991) 31.
[5] P.K. Smith, R.I. Krohn, G.T. Hermanson, A.K. Mallia, F.H. Gartner, M.D. Provenzano, E.K. Fujimoto, N.M., Goeke, B.J., Olson and D.C. Klenk, *Biochemistry*, 150 (1985) 76.
[6] J.J. Sedmak and S.E. Grossberg, *Anal. Biochem.*, 79 (1977) 544.
[7] A.S. Schmidt, A.M. Ventom and J.A. Asenjo, J.A., *Enzyme Microb. Technol.*, 16 (1994) 131.
[8] F.M. Hachem, *Ph.D. Thesis*, University of Reading, Reading, 1992.
[9] O. Cascone, B.A. Andrews and J.A. Asenjo, *Enzyme Microb. Technol.*, 13 (1991) 629.
[10] W. Melander and Cs. Horváth, *Arch. Biochem. Biophys.*, 183 (1977) 200.
[11] T.M. Przybycien and J.E. Bailey, *Enzyme Microb. Technol.*, 11 (1989) 264.
[12] G. Johansson, *J. Chromatogr.*, 322 (1985) 425.
[13] J.A. Asenjo, T. Franco, A.T. Andrews and B.A. Andrews, in M. White, S. Reuveny and A. Shafferman (Editors), *Biologicals from Recombinant Microorganisms and Animal Cells: Production and Recovery*, VCH, Weinheim, 1991, p. 439.
[14] M.A. Eiteman, *J. Chromatogr.*, 668 (1994) in press.
[15] Sasakawa, S. and H. Walter, *Biochemistry*, 11 (1972) 2760.