Article:
Chandler, E., Cordiner, J. and Brown, S. orcid.org/0000-0001-8229-8004 (2021) Accounting for interface behaviour in multi-stage aqueous two-phase extraction. Chemical Engineering Science, 230. 116172. ISSN 0009-2509

https://doi.org/10.1016/j.ces.2020.116172

Accounting for Interface Behaviour in Multi-Stage Aqueous Two-Phase Extraction

Emma Chandler[a], Joan Cordiner[a], Solomon Brown[a,*]

[a] *Department of Chemical and Biological Engineering, University of Sheffield, S1 3JD*

Abstract

Aqueous two-phase extraction (ATPE) is an alternative, cheaper and continuous protein purification technique. For ATPE to be used in industry, it must compete with current batch purification procedures, which can be achieved using multi-stage operation. To design extraction processes using ATPE appropriate process models must be developed. In this study, experimentally determined single-stage equilibrium data was used to generate a model to describe the behaviour of protein in multi-stage counter-current ATPE. Two distribution systems are considered: liquid-liquid (LL) and liquid-interface-liquid (LIL) distribution. The LIL model considers material which precipitates at the interface of the system. These models were then both compared against a three-stage experimental case study ATPE. The LIL model described the case study system better than the LL model, reducing the error from 40% to 11%. With this significant increase in accuracy, the LIL model represents an important tool with which to design multi-stage ATPE processes.

Keywords: Multi-stage extraction, Continuous protein purification, McCabe Thiele, Aqueous two-phase systems

1. Introduction

Aqueous two-phase extraction (ATPE) has been demonstrated to be a valuable alternative protein purification technique to chromatography (Rosa et al., 2010). ATPE, in comparison to chromatography, is relatively cheap and is ideal for continuous operation. A switch from batch to continuous manufacturing in bio-processing has been encouraged by regulatory bodies, including the FDA, because it can reduce processing costs while increasing manufacturing capacity and the consistency of the product quality (Konstantinov and Cooney, 2015). However, this move will require modification

*Corresponding Author. Tel.: +44 114 222 7597
Email address: s.f.brown@sheffield.ac.uk (Solomon Brown)

or replacement of current techniques while still retaining the high purity standards required in this industry (Azevedo et al., 2009). Aside from having the advantage of being a continuous manufacturing technique, ATPE is capable of handling large amounts of varying and crude material and is a very mild, low shear technique which is unlikely to damage the target protein (Asenjo and Andrews, 2011). The process has been demonstrated with a number of products, including: enzymes, monoclonal antibodies, protease inhibitors and serum albumin (Andrews et al., 1996, Harris et al., 1997, Rito-Palomares et al., 2000, Rito-Palomares and Lyddiatt, 2000). However, there are still challenges for ATPE to overcome in order to be commercially viable: the low resolution and a lack of understanding of the phase forming mechanisms (Asenjo et al., 1994, 2002, Rito-Palomares, 2004, Ruiz-Ruiz et al., 2012).

The low resolution of ATPE can be overcome either through the use of multi-stage operation or affinity ligands. The latter has been used to increase purity and yield to over 90% in a single step (Azevedo et al., 2009). Multi-stage ATPE was used by Rosa et al. (2013) to purify IgG from CHO cell media to 99% purity and 80% yield. Even when accounting for the added cost of a multi-stage operation, ATPE is cheap when compared with other purification steps, for instance chromatography. Further cost reductions can be made as it has been demonstrated that the more expensive phase forming polymer can be recycled using multi-stage ATPE (Rosa et al., 2013).

Because of the lack of understanding of the phase forming mechanisms, process design is reliant on both trial and error and individual expertise; this is both time consuming and expensive (Rito-Palomares, 2004). In ATPE system selection, parameters which need to be considered can include: choice of type and concentration of phase forming constituents, use of additional constituents such as NaCl and ligands, system conditions such as pH and temperature. As a result of the number of parameters, a trial and error approach is not guaranteed to yield the required results. For instance, Rito-Palomares and Middelberg (2002) screened a number of systems to extract a recombinant viral coat protein from *E Coli*, varying both PEG type and phase forming component concentration; however the most successful system removed 55% of contaminants and recovered 55% of the target protein. In order to generate experimental expertise in the area, Benavides and Rito-Palomares (2008) released a 'practical guide' which can aid in generating a starting point in system selection; however, there are still many variables to be tested. Bensch et al. (2007) suggested that high throughput robotic aided strategies could aid in rapid system selection; however, this would drastically increase the cost. Recently, there have been efforts to reduce material costs generated in extensive experimental system screening through

the use of miniaturized microfluidic platforms, these can evaluate multiple system's partition coefficient at the µL scale Bras et al. (2017).

Another approach to design and optimise ATPE is to use modelling strategies; this approach has been shown to be economically advantageous and is capable of reducing experimental workload while increasing the number of systems evaluated (Biegler and Grossmann, 2004). Modelling approaches in ATPE have been used to identify and describe the effect of key process parameters and select robust systems which can handle feedstock variations (Patel et al., 2018). Predictive modelling has been identified by Soares et al. (2015) and Torres-Acosta et al. (2019) as an area of ATPE which would benefit from further development. One road block is the large variety which exists between different systems and shifts in behaviour because of changes to system conditions. Currently, partitioning for a specific system and protein is predicted using physiochemical properties, including: amino acid composition and structural features of the protein and the affect of salt in the system (Salgado et al., 2008, Ferreira et al., 2015b,a). Ultimately, the aim would be predicting the binodal curve of a system and then subsequently the partitioning of protein, this has been attempted through molecular dynamic simulations (Dismer et al., 2013).

One way to combine experimental and modelling strategies to improve the performance is to utilise modelling techniques in multi-stage process design; this has been applied to both ATPE and the general chemical processing industry. In industry, multi-stage techniques have been used for processes that are simple and cost effective but provides a low purity and/or yield; the most well known is example is binary distillation. A common technique to optimise the number of stages in a process is to use the McCabe Thiele method which has been used extensively for binary distillation (Richardson et al., 2002). The McCabe Thiele method has been demonstrated to be suitable for ATPE when it was adapted and applied by Rosa et al. (2009a,b) Liu et al. (2018) and Chandler et al. (2019) for a stage wise optimisation of IgG, tetracycline and C-Phycocyanin respectively. For the method to be adapted to LLE and ATPE, equilibrium curves using experimental data of each of the contaminant(s) and target protein being considered must be constructed; these show the distribution of the contaminant(s) or target protein between the top and bottom phases of the selected system at a range of concentrations.

In the multi-stage ATPE models previously developed, only the protein partitioning into the top and bottom phases of the system is considered. Albertsson (1971) described two distribution regimes: liquid-liquid (LL) and liquid-interface distribution. This means that there are three regions in a

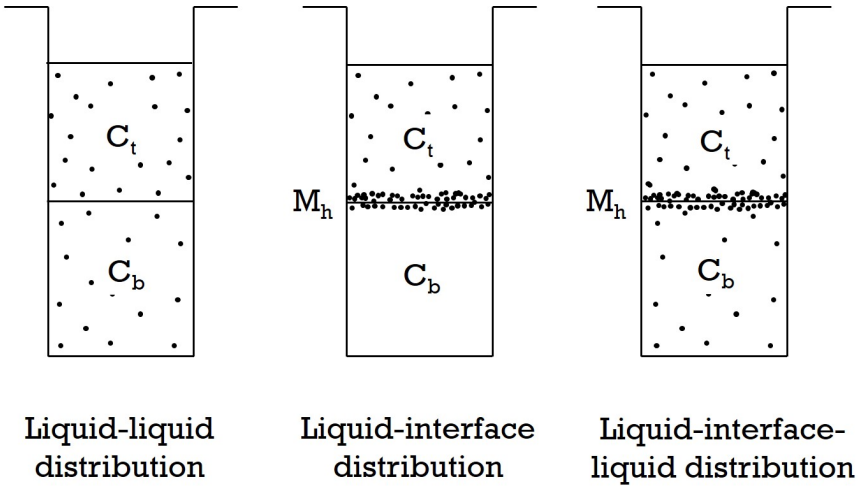


Figure 1: A diagram to show the different options for distribution of material in an ATPE, adapted from Albertsson (1971)

system which protein can partition: the top and bottom phase and the horizontal interface (HI). Where material in the top and bottom phases is dissolved in the bulk phase and material in the HI is precipitated. As a result, there are multiple possible distribution regimes which are demonstrated in Figure 1. Subsequently, the previously developed models are only capable of describing LL distribution regimes. Aside from different partitioning regimes, it is also known that LL distribution systems become saturated at very high protein concentrations, at which point protein partitions into the HI (Albertsson, 1971). Mündges et al. (2015) reported IgG aggregation with high CHO cell media loading which influenced the yield and purity. Andrews and Asenjo (1996) investigated partitioning of protein across the three regions described, with precipitation of protein into the HI of a system (Asenjo and Andrews, 2011). They found that in six systems tested, with three different proteins: Amyloglucosidase, subtilisin and Trypsin inhibitor, that systems rarely conform to a true LL distribution until phase saturation. Instead their data showed that above relatively low protein concentrations and the formation of solid phase which can settle into the HI was a common occurrence in systems.

This work aims to build upon previous work using an adapted McCabe Thiele method to describe multi-stage ATPE by considering HI as well as top and bottom phase partitioning. The work firstly screens a number of PEG-1500 and potassium phosphate buffer at a pH of 8.0 through experi-

mentation to select a suitable system. Single stage equilibrium data of the selected system was presented showing haemoglobin partitioning into the top, bottom and HI across a range of concentrations. This equilibrium data was used to develop models to describe haemoglobin behaviour in multi-stage counter-current ATPE with both LL distribution and liquid-interface-liquid (LIL) distribution. These models were then both compared against an experimental case study of a 3 stage counter-current ATPE which was spiked with haemoglobin in the waste stream.

2. Materials and Methods

2.1. Materials

Polyethylene Glycol (PEG) with an average molecular weight of 1500 Da was obtained from Sigma Aldrich, potassium phosphate dibasic, K_2HPO_4, was obtained from Acros organics, potassium dihydrogen phosphate, H_2KPO_4, was obtained from Alfa Aesar and haemoglobin was obtained from Sigma Aldrich.

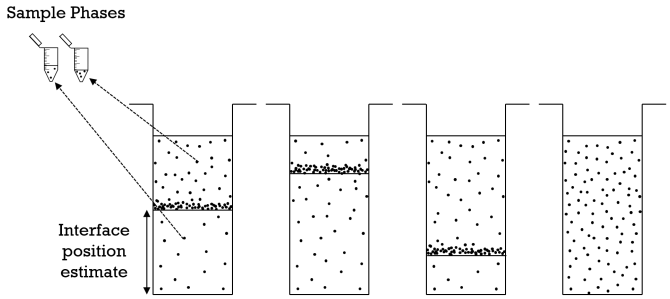
2.2. Experimental Methods

Phase systems were prepared using stock solutions of potassium phosphate pH 8.0 (20% w/w), and PEG 1500 (50% w/w). The phosphate buffer stock solution consisted of potassium phosphate dibasic and potassium dihydrogen phosphate. The stock solutions of PEG 1500 and phosphate buffer were stored at room temperature. Haemoglobin stocks were prepared to concentrations of 10 and 30 mg/mL and stored at 4°C.

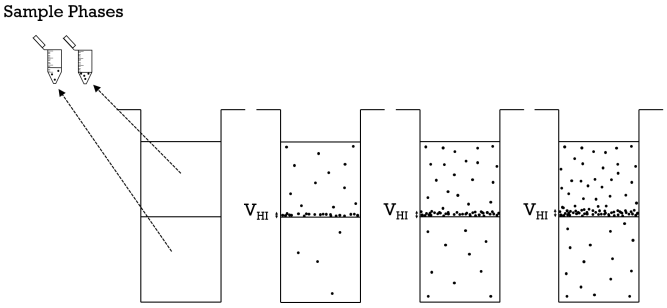
The system screening protocol is illustrated in Figure 2a and 2d. 1.5 mL system, at the required concentrations of PEG 1500 and phosphate buffer were prepared in 2 mL centrifuge tubes. Systems were then each spiked with 100 µL of 30 mg/mL haemoglobin. The systems were mixed thoroughly for 30 seconds by vigorous shaking and inversion of the centrifuge tube. Systems were centrifuged at 800 g for 30 minutes to ensure that systems had reached equilibrium. Approximate partitioning volumes were then determined using the graduations on the centrifuge tube. Known volumes of the top and bottom phase were then taken carefully using a micro-pipette, ensuring that no material was taken from the interface region. Samples were then diluted with deionised water by a known amount of 1 - 4 times their volume after which they were vortexed for 30 seconds to ensure they were thoroughly mixed. Samples were then analysed in a UV-Vis at 405 nm and haemoglobin concentration was determined through the use of standard curves and the Beer-Lambert law. As only an estimated concentration was required for the

system screening, the standard curves were constructed in water, rather than each individual system tested, to save time. The amount of material in the HI was calculated from the total material added and the material measured in the top and bottom phases.

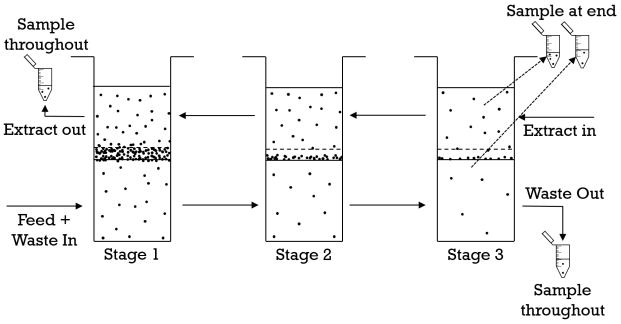
Haemoglobin was chosen as a model protein because of its pigment. Multiple systems were screened to evaluate protein partitioning into the top, bottom and HI of systems as well as system volume ratios. This was carried out in order to find a system suitable for multi-stage extraction experiments and evaluate partitioning into the HI across a range of systems. The work of Kan and Lee (1994) was used as a basis for choosing system conditions of a PEG 1500 and potassium phosphate buffer (pH 8.0). Table 1 shows the concentrations systems screened.



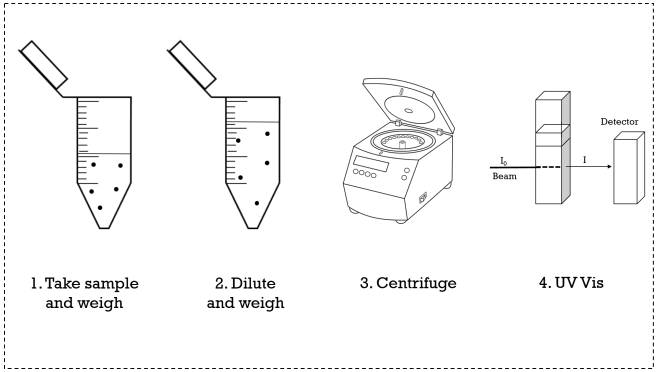
(a) Screening of volume ratio and protein partitioning for a range of system conditions



(b) Equilibrium curve construction for a range of protein concentrations in a system



(c) Schematic of the counter-current multi-stage extraction used in the experimentation



(d) Sample analysis method schematic

Figure 2: A Schematic to show an overview of the experimentation carried out in this study.

One of the systems screened, a 13.0% w/w PEG 1500 and 11.2% w/w potassium phosphate buffer pH 8.0 system, was used in further experimental investigations to study multi-stage ATPE. For this system equilibrium curves were constructed, phase volumes were determined and a multi-stage extraction was carried out; this is depicted in Figure 2b and 2d. For the equilibrium curve, the systems were prepared and then spiked with a known volume of haemoglobin, up to 10% of the total volume. Systems were thoroughly mixed and then centrifuged at 800 g for 30 minutes to ensure phases had reached equilibrium. 300 µL of the top phase and 500 µL of the bottom phase was carefully extracted using a micro-pipette. Particular care was taken to not disrupt the material at the HI. The top phase was then diluted by 900 µL and the bottom phase sample by 700 µL of deionised water. Samples were then vortexed for 30 seconds to ensure they were thoroughly mixed. Calibration curves were constructed for the top and bottom phase which were prepared as described in this paragraph and then spiked with haemoglobin stock. Samples were then analysed in a UV-vis at 405 nm and haemoglobin concentration was determined through the use of standard curves and the Beer-Lambert law. The amount of material in the HI was calculated from the total material added and the material measured in the top and bottom phases. Experimentation was carried out in triplicate.

The size of the HI and the material gathered there changes with respect to the total protein concentration; an experiment was carried out to approximate this. The determination of the HI volume vs protein concentration was carried out as follows. A 6 mL system of a 13.0% w/w PEG 1500 and 11.2% w/w potassium phosphate buffer pH 8.0 was prepared by weight, and spiked with haemoglobin stock (up to 10% volume of the total system volume). Systems were then allowed to separate until equilibrium. Phase heights of the bottom phase, top phase, and HI material were then determined using a digital micrometer along with the falcon tube diameter. The phase heights and the falcon tube diameter were then used to determine phase volumes. Experimentation was carried out in triplicate.

For the multi-stage extraction, three counter-current stages were used; this is depicted in Figure 2c and 2d. Larger (*ca.* 15 mL) volumes of a 13.0% w/w PEG 1500 and 11.2% w/w potassium phosphate buffer pH 8.0 system were prepared by weight in 15 mL falcon tubes. These were mixed thoroughly for 30 seconds and allowed to separate overnight. The top and bottom phases were then collected using a hypodermic needle and syringe, and stored as stocks for the top and bottom phase for use in the multi-stage extraction. Care was taken to avoid the material surrounding the HI which was discarded.

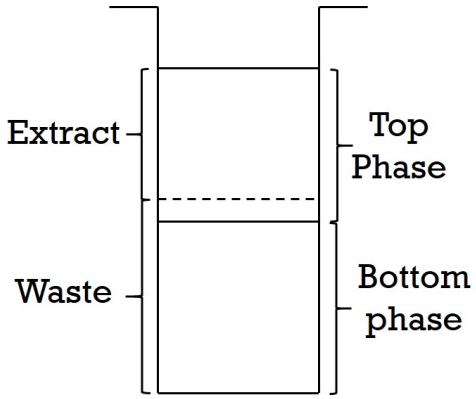


Figure 3: A schematic representation of the extract and waste phases for the multi-stage extraction.

The HI was measured to be at 52% of the total system height. The total volumes of the systems were 1500 mL. The extract consisted of the 650 µL of the top phase, and the waste phase consisted of 780 µL of the bottom phase stock and 70 µL of the top phase stock. 54.6 µL of haemoglobin stock was added to the waste feed at stage 1.

For each run of the multi-stage extraction, the extract and waste phases were transferred across the stages as demonstrated in Figure 2c. Once phases were transferred across the stages systems were shaken vigorously for at least 30 seconds to mix and allowed to separate to equilibrium. The system was run for 15 runs, the extract and waste output for each run was collected. After 15 runs the waste and extract for each stage was collected for analysis which is depicted in Figure 2d. All samples were weighed and then diluted by a known volume. Samples were then vortexed for 30 seconds to ensure they were properly mixed. Samples were then analysed in a UV-vis at 405 nm and haemoglobin concentration was determined through the use of standard curves and the Beer-Lambert law. Standard curves were constructed for both the waste and extract phases. The multi-stage extraction was carried out in triplicate.

2.3. Computational Methods

In this work, we compare two models; the first is based on the traditional method of modelling ATPE and assumes the distribution of material is between the two regions: the top and bottom phases of the system. The model developed in this paper assumes material is partitioned between three regions: the top and bottom phase and the HI of the system. Both use

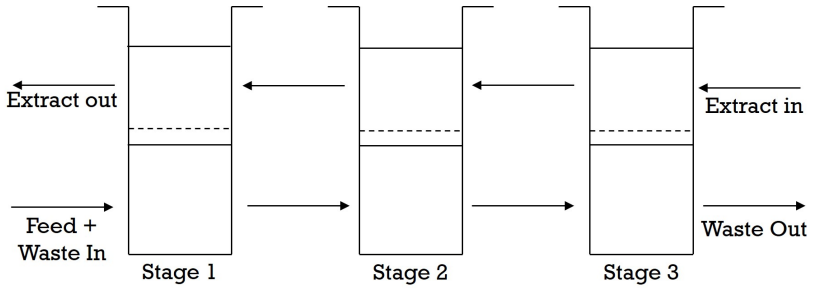


Figure 4: A schematic to illustrate the multi-stage extraction.

single-stage equilibrium behaviour as a basis for modelling the multi-stage ATPE.

The model firstly assumes that the phases are immiscible, this is a common assumption for ATPE despite that phases are semi-miscible. The process is isothermal, so an energy balance is not used (Mistry et al., 1996, Rosa et al., 2009b,a). It was assumed that perfect mixing and equilibrium for phase separation was achieved at each stage. In ATPE several components must be considered, these can be divided into the phase forming materials (polymer(s), salt(s) and water) and the material to be partitioned (the contaminants and target protein). The partitioning of the protein is assumed to be a result of the phase forming materials. The concentration of the phase forming components and the phase volumes are assumed to be constant across the system (Rosa et al., 2009b,a). It was lastly assumed that the HI was flat.

The mass balances in the model assume a semi-continuous multi-stage operation with rapid batch separations which are moved counter-currently upon separation. The model utilises the same system conditions at each stage with only the protein concentration varying; this similar to Rosa et al. (2009a,b). In comparison,Mistry et al. (1996) simulated a continuous multi-stage operation where the extraction conditions changed at each stage to account for extraction and back extraction; this requires more experimental data to generate adequate mass balances which cover all components within the system. While the modelling in the present paper assumes and accounts for partitioning between three regions (the top, the bottom, and the HI of the system), the simulations from Rosa et al. (2009a,b) and Mistry et al. (1996) assume partitioning between only the top and bottom phase. Mistry et al. (1996) presents the partitioning of protein as a constant value, whereas the present paper and Rosa et al. (2009a,b) account for changes in partitioning

as target concentration changes.

For the multi-stage distribution, if V is the phase volume (mL), C is the concentration of protein (mg/mL), x is the phase which protein is being moved from, y is the phase protein is being moved to, and N is the total number of stages, the mass balance is described by:

$$V_x C_{x,0} + V_y C_{y,N+1} = V_x C_{x,N} + V_y C_{y,1} \tag{1}$$

Likewise, material balances for stage n can be written as:

$$V_x C_{x,n-1} + V_y C_{y,n+1} = V_x C_{x,n} + V_y C_{y,n} \tag{2}$$

Using equations (1) and (2) we can pose the stage-wise optimisation of the multi-stage ATPE as follows:

$$\min N \tag{3}$$

subject to:

$$C_{x,N} \leq C_{Target} \tag{4}$$

Where C_{Target} is the required concentration in phase x. If optimising number of stages for yield, x is the target protein concentration in the waste phase. If optimising number of stages for purity, x is contaminant protein concentration in the extract phase. In this paper, we consider the yield of the system. To determine the first step, a mass balance across the entire system is used:

$$C_{y,1} = C_{y,n+1} - \frac{V_x}{V_y}(C_{x,n} - C_{x,0}) \tag{5}$$

Each stage is then evaluated as follows:

$$\left.\begin{aligned} C_{x,n} &= f(\text{equilibrium line}), \\ C_{y,n+1} &= C_{y,n} - \frac{V_x}{V_y}(C_{x,n} - C_{x,n-1}), \end{aligned}\right\} \qquad n = 1, \ldots, N \tag{6}$$

Where f is the function determined for the equilibrium line. The line is determined either by a piecewise linear function or a curve of the equilibrium points determined experimentally. This collection of the experimental data is described in the previous section. In this paper, the equilibrium line is described by a curve, shown below in equations 7 and 8.

The partitioning in the LL distribution model is constructed as follows. Based on the experimental data, quadratic relationships were assumed between the total concentration and the concentration in the top and bottom

phases of the system. As a result, the partitioning in each stage is determined by:

$$C_{Bottom} = k_{b1}C_{Total}^2 + k_{b2}C_{Total} \tag{7}$$

$$C_{Top} = k_{t1}C_{Total}^2 + k_{t2}C_{Total} \tag{8}$$

The coefficients k_{b1}, k_{b2}, k_{t1} and k_{t2} were then determined by regression against the experimental data points from the equilibrium curve; k_{b1} and k_{t1} have the units mL/mg, k_{b2} and k_{t2} are dimensionless. There is a range that a linear relationship can be seen between the total and phase concentration for the top and bottom phases; however, the equilibrium curve includes data points at which the phases have reached a level of saturation (Albertsson, 1971, Mistry et al., 1996). As such, a linear relationship was no longer appropriate and a second order polynomial was used to describe the top and bottom phase concentrations in relation to the total system concentration.

For the phase volumes, V_{Bottom} and V_{Top} were determined experimentally and remained constant regardless of protein concentration.

In the LIL distribution model, the material in the HI was also considered. In the single-stage equilibrium curve, the amount of protein in the HI was calculated using a mass balance:

$$M_{HI} = C_{Total}V_{Total} - C_{Bottom}V_{Bottom} - C_{Top}V_{Top} \tag{9}$$

The concentration, C_{HI} of mass in the HI, M_{HI} (mg) was calculated with respect to the total system volume, rather than with respect to the volume measured at the HI. This is to avoid potentially large errors with volume measurement being propagated through the model. If a more accurate volume measurement of the interface could be used, it could be calculated as a true concentration. Based on the experimental data a linear relationship was assumed between the total concentration and the concentration in the HI of the system. As a result, the partitioning in each stage is determined by:

$$\frac{M_{HI}}{V_{Total}} = k_{HI}C_{Total} \tag{10}$$

The coefficient k_{HI} is dimensionless. While the volume of the phase at the HI was not used in concentration calculations to avoid error propagation, it was used to determine whether the mass at the HI was part of the 'extract' or 'waste'. Larger HI regions partitioned into the extract and smaller

third phase regions partitioned into the waste. Ideally material in the HI should remain in the waste phase; however, partitioning of HI material into the extract occurs at higher protein concentrations and therefore large HI volumes. The volume of the HI, V_{HI}, was considered as an insoluble phase at the HI over the top of the liquid phases which is described by:

$$\%V_{HI} = alog_{10}\left(\frac{C_{Total} + b}{c}\right) \tag{11}$$

$$V_{HI} = \%V_{HI}V_{Total} \tag{12}$$

The coefficients a (dimensionless), b (mg/mL) and c (mL/mg) were determined using the data points determined experimentally from the HI volume curve using least squares regression.

It was assumed that the mass in the HI was all in the extract or the waste phase. If the volume of the HI was below $V_{Critical}$ then the HI material partitioned into the waste phase. If the volume of the HI was above $V_{Critical}$ then the HI material partitioned into the extract phase. $V_{Critical}$ is dependant on the extraction method and materials used in the experimental protocol.

$V_{Critical}$ is defined as:

$$V_{Critical} = E_H \pi r^2 \tag{13}$$

Where E_H is the extraction height (cm) above the HI and r is the radius of the test tube (cm). The the waste, W and extract, E, phase protein concentration (mg/mL) for each stage, n, are defined as follows:

If $V_{HI} < V_{Critical}$:

$$C_{W,n} = V_{Bottom}C_{Bottom,n} + (V_{Top} - V_E)C_{Top,n} + C_{HI,n}V_{Total} \tag{14}$$

$$C_{E,n} = V_{E,n}C_{Top,n} \tag{15}$$

If $V_{HI,n} \geq V_{Critical}$:

$$C_{W,n} = V_{Bottom}C_{Bottom,n} + (V_{Top} - V_E)C_{Top,n} \tag{16}$$

$$C_{E,n} = V_E C_{Top,n} + C_{HI,n}V_{Total} \tag{17}$$

Then for the stage-wise balance:

$$V_W C_{W,n-1} + V_E C_{E,n+1} = V_W C_{W,n} + V_E C_{E,n} \tag{18}$$

We can then use equations 1 to 6 to calculate how the system behaves in a multi-stage process. For this paper, as the yield is being considered, the material is being moved from the waste phase to the extract phase. Therefore y is the extract phase and x is the waste phase.

3. Results and Discussion

The first step in using ATPE involves screening for systems which have suitable extraction conditions. For single-stage ATPE, often many systems have to be screened to find desirable extraction conditions. For multi-stage ATPE, less desirable extraction conditions with lower resolutions can be accepted because the resolution can be improved by altering the number of stages, and so the screening is less extensive. Screening is still necessary in multi-stage extraction as non-optimal conditions could result in losses between stages which are higher than that of single-stage extraction. While there are many parameters which can be varied in order to alter partitioning, in this study the chosen system conditions were based upon the work of Kan and Lee (1994).

Table 1 shows the screening concentrations and position of the HI and 5 shows the amount of protein in the top, bottom and HI of each system. Table 1 shows that three systems did not contain high enough concentrations of phase forming material (PEG 1500 and potassium phosphate buffer) to form two-phase systems. Systems A, B, C and J all partitioned so that one phase consisted of around 31% of the volume and the other around 69% of the volume. While it is possible to operate LLE with a large volume ratio, it is preferable to use systems with more or less equal volumes such as those shown by systems E, K and L.

Volume ratios are also important to consider in system selection for three reasons: they have a relationship with the phase separation kinetics, systems with a very small phase reach protein saturation faster, and separator design is dependant on volume ratios. For the phase separation kinetics of PEG-salt systems, systems which have a large top phase have continuous top PEG-rich phases, meaning they have much slower phase separation kinetics than systems with continuous bottom salt-rich phase (Albertsson, 1971, Kaul et al., 1995). The fastest phase separations are observed at intermediate positions on a binodal where the phase volumes are approximately equal (Salamanca et al., 1998). This difference is seen because of the much larger viscosities of the PEG-rich phase compared to the salt-rich phases. Slow separation kinetics would become problematic upon scale up, reducing the amount of material processed. For protein accumulation in the HI, if one of the phases

Label	PEG 1500 (% w/w)	Phosphate Buffer (% w/w)	HI position (Estimated) (μl)
A	9.7	12.6	1100
B	9.5	14.4	1100
C	9.4	16.2	1100
D	13.2	9.3	No Phase Formation
E	13.0	11.2	750
F	12.8	13.1	1000
G	12.6	15.0	1000
H	16.9	5.8	No Phase Formation
I	16.6	7.9	No Phase Formation
J	16.3	9.9	500
K	16.0	11.8	750
L	15.8	13.7	800

Table 1: The concentration and position of the HI in a 1600 µL sample of the PEG 1500 - potassium phosphate pH 8.0 systems screened

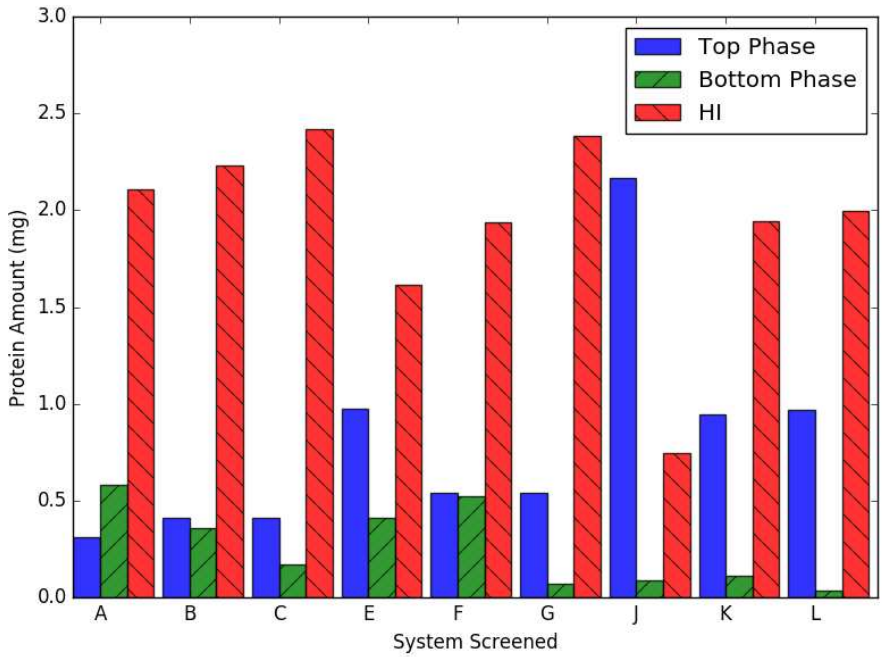


Figure 5: Distribution of 3 mg of haemoglobin in different 1600 µL PEG 1500 - potassium phosphate pH 8.0 systems screened.

is too small, only a small amount of protein will be able to partition into it, even if the phase can handle a high concentration of protein, until the phase becomes saturated (Albertsson, 1971). Systems with a saturated phase will result in accumulation of material in another phase, resulting in either material in the HI or a less desirable partition coefficient. Lastly, with separator designs, the equipment dimensions are often dependant on the volume rations. A large volume ratio would require non-traditional settler designs which could increase the complexity and therefore cost of the equipment.

For an extraction, it is preferable to maximise the amount of dissolved protein as opposed to precipitated protein. While the precipitated protein can still theoretically redissolve, it can become damaged, aggregate or be difficult to redissolve (Albertsson, 1971). Theoretically, systems have either a LL or liquid-interface distribution; because a large amount of haemoglobin partitioned into the top and HI of the systems, these systems would be considered to have liquid-interface distribution. However, Andrews and Asenjo (1996) demonstrated that true two-phase partitioning regimes are uncommon and often partitioning is seen across the top and bottom phases with precipitated material in the HI of a system, even at low protein concentrations. Figure 5 showing the partitioning of systems screened, demonstrates there is protein in both phases and the HI for all two-phase systems tested. It should be noted that relatively high concentrations of protein (2 mg/mL) were used in the screening. As higher concentrations are reached, more material tends to partition into the HI. The point at which more material tends towards the HI is dependant on the system conditions and the protein. Haemoglobin has likely become saturated in the top and bottom phase meaning partitioning into the HI is likely exaggerated by the high protein concentration.

For further testing, system E was chosen for study as it had a suitable distribution of material into the top phase, as well as volume ratio close to 1. System J had better distribution into the top phase; however, had less suitable volume ratios with the bottom phase only forming 500 µL of the 1600 µL system. Figure 6 shows the partitioning of haemoglobin in system E into the top and bottom phases and the HI with respect to the total system concentration. In this graph, the material in the HI is shown as a pseudo concentration of the total system volume as the volume of the HI is very small, and the top and bottom phases are shown as a concentration of their respective phases. When considering material in the HI, Andrews and Asenjo (1996) also considered the HI material as a pseudo concentration of the total system volume. Figure 6 shows that the haemoglobin favours the top phase, followed by the HI and then the bottom phase. The material in the top and bottom phases was dissolved and the material in the HI is a precipitated mass

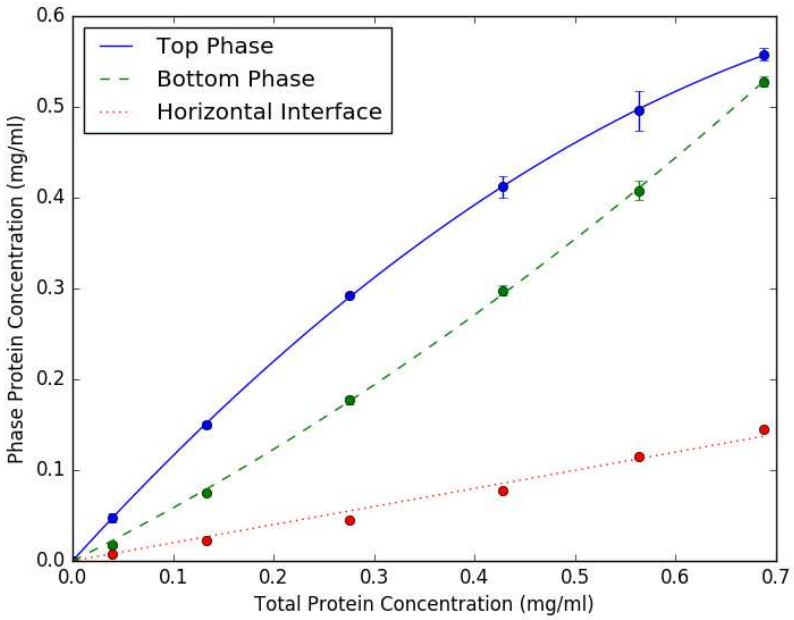


Figure 6: The partitioning of haemoglobin in a 13.0% w/w PEG 1500 and 11.2% w/w potassium phosphate buffer at pH 8.0 system.

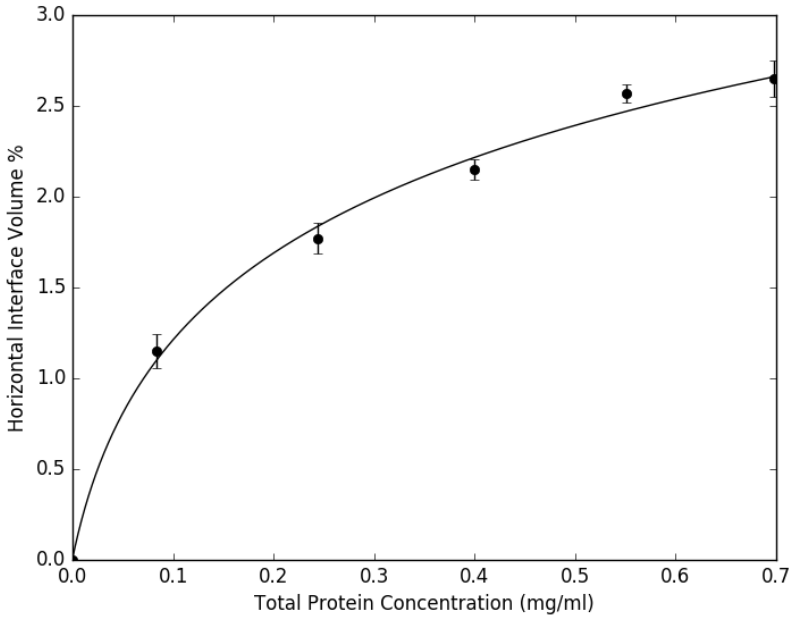


Figure 7: Volume of the HI as a percentage of the total volume in a 13.0% w/w PEG 1500 and 11.2% w/w Potassium phosphate buffer at pH 8.0 system.

which sits at the HI. As the total protein concentration increases, the mass in the HI grows into the top phase. Once a total haemoglobin concentration of greater than approximately 0.45 mg/mL is reached, the top phase begins to become saturated and an increasing amount of haemoglobin partitions towards the bottom phase. If higher concentrations of protein were used, the bottom phase would also become saturated and more material would partition into the HI. It can be seen that material immediately partitions into the top, bottom and HI. This contradicts the assumption that systems partition either as a liquid-liquid distribution or as a liquid-interface distribution until the system is saturated and supports the findings of Andrews et al. (1996), who found that protein often partitioned into the top, bottom and HI of systems.

The experimental data in Figure 6 and Figure 7 were both used in the model. The equilibrium curves were used to predict the partitioning of material of between the phases of a system given a total protein concentration, and the volume curve was used to predict where the material at the HI went, the waste or extract phase, given a total protein concentration for each stage.

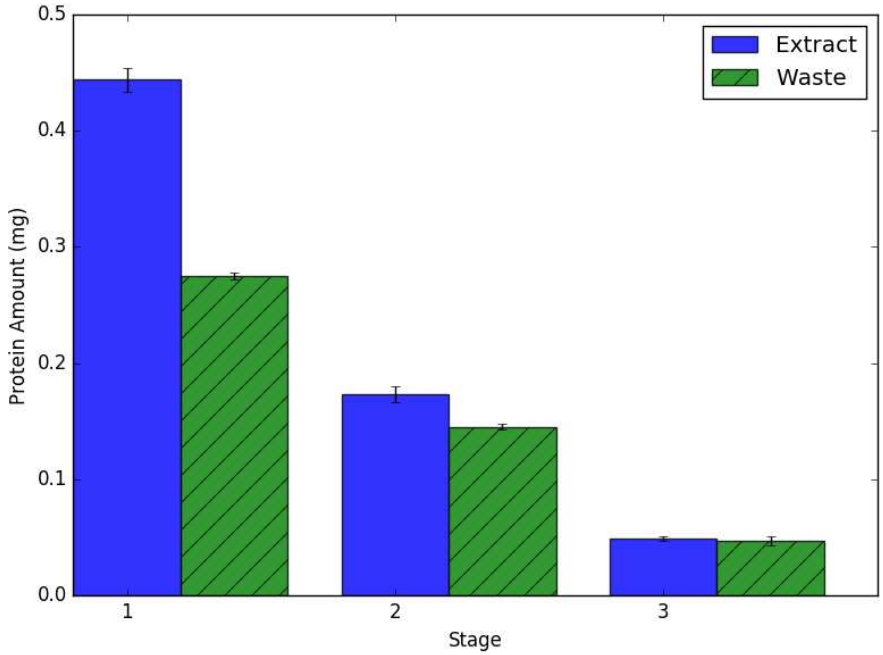


Figure 8: The amount of protein in the extract and waste phases of a 3 stage counter-current 13.0% w/w PEG 1500 and 11.2% w/w potassium phosphate buffer at pH 8.0 ATPE

Stage	Phase	Experiment Amount (mg)	LL model Amount (mg)	LL model Error (%)	LIL model Amount (mg)	LIL model Error (%)
1	Extract	0.444	0.244	-40.0	0.436	-1.6
	Waste	0.275	0.248	-5.4	0.295	4.0
2	Extract	0.173	0.127	-9.2	0.228	11.0
	Waste	0.145	0.107	-7.6	0.129	-3.2
3	Extract	0.049	0.048	-0.2	0.057	1.6
	Waste	0.047	0.037	-2.0	0.073	5.2

Table 2: Comparison of the amount of protein in each phase between the experimental data, the LL distribution model and the LIL distribution model.

The model was then tested against a case study system. Figure 8 shows the amounts of haemoglobin in the extract and waste phases of each stage of a 3 stage counter-current ATPS. Using three stages of system E the yield of haemoglobin ATPE improved from 61.8% in a single stage to 85.3% in three stages, drastically improving the yield of ATPE using a multi-stage extraction. Table 2 shows the results for the LL and the LIL distribution models in comparison to the experimental results. The table shows the amount (mg) of haemoglobin in the extract and waste phases of each stage, as well as the percentage difference between the experimental data and the LL and LIL distribution models with respect to the amount of protein in the feed.

The LL distribution model assumes protein partitions into just the top and bottom phases and is the methodology previously used to describe partitioning behaviour of multi-stage ATPE as it is commonly assumed in single-stage systems (Rosa et al., 2009b,a, Liu et al., 2018). The LL model can be seen describe the case study extraction in stages with less material in the system; however, for the extract in stage one, with the highest concentration of haemoglobin, the model has a 40% error. This is because it does not account for the protein in the HI meaning that all of the protein in the HI was lost. While the rest of the model reflects the experimental data well, should the LL distribution model be used to look at any stages placed in front of stage 1 (to look at contaminant partitioning), this error would carry into these stages. This would also cause more problems if a higher feed concentration was used. The experimental data shows the total concentration in stage 1 is 0.462 mg/ mL. Figure 6 showed the top phase starts becoming saturated at concentrations of >0.45 mg/mL. This means that the LL distribution model is showing a 40% error in stage 1, despite only just reaching saturation. It

is likely that at larger concentrations, the model would deviate more.

The material in the HI could be considered as part of the waste or extract phase, and an equilibrium curve could be constructed using these phases. However, this would give less flexibility in the modelling as the extract and waste phases would have a set size. By constructing an equilibrium curves considering the top, bottom and HI, it allows the model to consider different extract and waste phase sizes; thereby allowing for different extraction methods, equipment and scale to be considered.

The behaviour of the HI material is dependant on both the volume of the mass at the interface and the extraction methodology. Practically speaking, its preferable to avoid systems which precipitate a target protein into the interface. However, Andrews and Asenjo (1996) identified that very few systems exist which do not have protein partitioning into the HI, even at low concentrations. It has known that it is common larger contaminants, cells and cell debris to experience partitioning to the HI; therefore, in systems with a large number of contaminants in the HI it would be important to avoid taking this region into the extract as it would reduce purity. As a result, having models which predict for HI partitioning, would allow users account for this behaviour and including the HI in the extract phase can be avoided or minimised, without changing the feed characteristics. For instance this could be achieved by changing equipment to avoid disturbing material at the HI, decreasing the extract size, adding extra stages, choosing a more suitable system, or decreasing the amount of protein in the feed.

This model is only used to predict the behaviour of a single protein; in future work it could be used to evaluate both a target and contaminants separately. As contaminants are likely to consist of a range of different protein and material, often they are grouped together, either completely or as two groups: high and low molecular weight components (Rosa et al., 2009a,b). There are other divisions possible, and a stubborn contaminant could be evaluated on its own. Generally for extraction steps, the yield of the system can be evaluated by looking at reducing the amount of target phase in the waste stream. The purity can be improved by reducing the contaminants in the extract stream. For a back extraction step, only the target movement out of the polymer phase and into the salt phase needs to be considered. Lastly for a washing step, removal of any remaining material from the polymer phase should be considered. Using this model as a basis, evaluating different protein components in turn the model can be used to build up multi-stage process flow sheets for ATPE which can greatly improve the performance of the extraction. The added consideration of the HI partitioning could also be integrated into other modelling approaches such as the work of Mistry et al.

(1996) and Rosa et al. (2009a,b).

In this work, only a single protein was considered and the above approach considered each protein in turn in isolation. In practice, while the partitioning of the proteins will mostly be a result of the system conditions, different proteins in a system together will have an impact on each other's partitioning behaviour. This is particularly likely to be the case in systems which have a saturated phase, and a build-up of material in the HI. As a result, future work should look at experimentally evaluating multi-protein mixtures and partitioning in terms of the HI, top and bottom phase. Further future work in modelling could involve adopting a multi-component feed approach as is used in distillation and in more traditional LLE. In this approach, the components are pseudo separated and parameters are assigned determine their interaction, for processes where component partitioning are strongly dependant on phase compositions, as is the case in ATPE, a modified Rachford-Rice algorithm could be applied (Seader et al., 2006). While there are likely a large number of components in a biological separation, similar contaminants could be grouped together and major or stubborn contaminants could be evaluated separately (Rosa et al., 2009a,b). As feeds in bioprocessing are complex and variable, future work should also look at robustness testing of the model, evaluating how feed variation affects model output.

Ideally, this model would be extended to work with single stage predictive behaviour models, so as to reduce experimental work load and increase the number of systems considered when picking optimal conditions. However, currently single stage models and experimental work defines partitioning in terms of top and bottom phase partitioning. For more accurate results, partitioning into the HI could also be considered.

4. Conclusions

The use of multi-stage ATPE is capable of increasing the performance of the process, the yield of a model protein, haemoglobin, in a model multi-stage extraction was improved from 61.8% in a single step to 85.3% in three stages. This work describes a model to describe multi-stage ATPE which is based on mass balances between the stages, with each stage having three regions in which protein can partition: the top, bottom and HI. The model requires two types of single-stage equilibrium experimental data: protein partitioning data and the volume of precipitation of protein at the HI over a range of protein concentrations. In order to account for the material the only extra experimental data collection required is volume of the material

precipitating at the HI, which is easy to obtain. The amount of material at the interface is determined via calculation.

Generally partitioning of protein with systems is considered to be between two regions: the top and bottom phase. The main finding of this paper is that to successfully describe multi-stage ATPE the protein needs to be considered. Across 9 systems screened haemoglobin was found to partition into all three regions when systems were spiked with 2 mg/mL of haemoglobin: the HI, top and bottom phases. In the model system which was studied in more detail, it was found that protein partitioned into all three regions immediately, not just after the top and bottom phases were saturated. As a result, the model presented in this paper which accounts for the behaviour of protein in the system at the HI, reduced the maximum error experienced to 11% from 40% compared with a model which described only the top and bottom phase protein partitioning.

5. Acknowledgements

This research did not receive any grant specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

P.-Å. Albertsson. *Partition of cell particles and macromolecules in polymer two-phase systems.* Wiley, USA, 2 edition, 1971.

B. Andrews and J. Asenjo. Protein partitioning equilibrium between the aqueous poly(ethylene glycol) and salt phases and the solid protein phase in poly(ethylene glycol)-salt two-phase systems. *Journal of Chromatography B*, 685(11):15–20, 1996. doi: 10.1016/0378-4347(96)00134-X.

B. A. Andrews, S. Nielsen, and J. A. Asenjo. Partitioning and purification of monoclonal antibodies in aqueous two-phase systems. *Bioseparation*, 6 (5):303–13, 1996. ISSN 0923-179X.

J. a. Asenjo and B. a. Andrews. Aqueous two-phase systems for protein separation: A perspective. *Journal of Chromatography A*, 1218(49):8826–8835, 2011. ISSN 00219673. doi: 10.1016/j.chroma.2011.06.051.

J. A. Asenjo, A. S. Schmidt, F. Hachem, and B. A. Andrews. Model for predicting the partition behaviour of proteins in aqueous two-phase systems. *Journal of Chromatography A*, 668(1):47–54, 1994. ISSN 00219673. doi: 10.1016/0021-9673(94)80090-1.

J. a. Asenjo, S. L. Mistry, B. a. Andrews, and J. C. Merchuk. Phase separation rates of aqueous two-phase systems: Correlation with system properties. *Biotechnology and Bioengineering*, 79(2):217–223, 2002. doi: 10.1002/bit.10273.

A. M. Azevedo, P. A. Rosa, I. F. Ferreira, and M. R. Aires-Barros. Chromatography-free recovery of biopharmaceuticals through aqueous two-phase processing. *Trends in Biotechnology*, 27(4):240–247, apr 2009. ISSN 01677799. doi: 10.1016/j.tibtech.2009.01.004. URL https://linkinghub.elsevier.com/retrieve/pii/S0167779909000353.

J. Benavides and M. Rito-Palomares. Practical experiences from the development of aqueous two-phase processes for the recovery of high value biological products. *Journal of Chemical Technology and Biotechnology*, 83(2):133–142, 2008. ISSN 02682575. doi: 10.1002/jctb.1844.

M. Bensch, B. Selbach, and J. Hubbuch. High throughput screening techniques in downstream processing: Preparation, characterization and optimization of aqueous two-phase systems. *Chemical Engineering Science*, 62(7):2011–2021, 2007. doi: 10.1016/j.ces.2006.12.053.

L. T. Biegler and I. E. Grossmann. Retrospective on optimization. *Computers & Chemical Engineering*, 28(8):1169–1192, jul 2004. doi: 10.1016/J.COMPCHEMENG.2003.11.003.

E. J. Bras, R. R. Soares, A. M. Azevedo, P. Fernandes, M. Arévalo-Rodríguez, V. Chu, J. P. Conde, and M. R. Aires-Barros. A multiplexed microfluidic toolbox for the rapid optimization of affinity-driven partition in aqueous two phase systems. *Journal of Chromatography A*, 1515:252–259, sep 2017. doi: 10.1016/j.chroma.2017.07.094.

E. Chandler, R. Falconer, and S. Brown. Optimisation of Aqueous Two-Phase Systems. In *Proceedings of the 29th European Symposium on Computer Aided Process Engineering*, pages 631–636, 2019. doi: 10.1016/B978-0-12-818634-3.50048-5.

F. Dismer, S. Alexander Oelmeier, and J. Hubbuch. Molecular dynamics simulations of aqueous two-phase systems: Understanding phase formation and protein partitioning. *Chemical Engineering Science*, 96:142–151, jun 2013. ISSN 00092509. doi: 10.1016/j.ces.2013.03.020.

L. A. Ferreira, O. Fedotoff, V. N. Uversky, and B. Y. Zaslavsky. Effects of osmolytes on protein-solvent interactions in crowded environments: Study

of sucrose and trehalose effects on different proteins by solvent interaction analysis. *RSC Advances*, 5(34):27154–27162, 2015a. ISSN 20462069. doi: 10.1039/c5ra02997j.

L. A. Ferreira, P. P. Madeira, V. N. Uversky, and B. Y. Zaslavsky. Analyzing the effects of protecting osmolytes on solute-water interactions by solvatochromic comparison method: I. Small organic compounds. *RSC Advances*, 5(74):59812–59822, jul 2015b. ISSN 20462069. doi: 10.1039/c5ra08610h.

D. P. Harris, A. T. Andrews, G. Wright, D. L. Pyle, and J. A. Asenjo. The application of aqueous two-phase systems to the purification of pharmaceutical proteins from transgenic sheep milk. *Bioseparation*, 7(1):31–7, 1997. doi: 10.1023/A:1007908703773.

P. Kan and C. J. Lee. Application of aqueous two-phase systems in separation/purification of stroma free hemoglobin from animal blood. *Artificial cells, blood substitutes, and immobilization biotechnology*, 22(3):641–9, 1994. ISSN 1073-1199. doi: 10.3109/10731199409117894.

A. Kaul, R. a. M. Pereira, J. A. Asenjo, and J. C. Merchuk. Kinetics of phase separation for polyethylene glycol-phosphate two-phase systems. *Biotechnology and Bioengineering*, 48(3):246–256, 1995. doi: c.

K. B. Konstantinov and C. L. Cooney. White Paper on Continuous Bioprocessing. May 20-21, 2014 Manufacturing Continuous Symposium. *Journal of pharmaceutical sciences*, 104(3):813–820, 2015. doi: 10.1002/jps.24268.

Y. Liu, L. Chen, J. Zhou, and Z. Yan. Extraction of Tetracycline using the Ionic Liquid-Based Aqueous Two-Phase Systems:Single-Stage Versus Multi-Stage. *Chemical Industry and Chemical Engineering Quarterly*, 24 (4):387–397, 2018. doi: 10.2298/CICEQ170904011L.

S. L. Mistry, A. Kaul, J. C. Merchuk, and J. A. Asenjo. Mathematical modelling and computer simulation of aqueous two-phase continuous protein extraction. *Journal of Chromatography A*, 741(2):151–163, 1996. doi: 10.1016/0021-9673(96)00179-3.

J. Mündges, J. Zierow, and T. Zeiner. Experiment and simulation of an aqueous two-phase extraction process for the purification of a monoclonal antibody. *Chemical Engineering and Processing: Process Intensification*, 95:31–42, sep 2015. doi: 10.1016/j.cep.2015.04.013.

N. Patel, D. G. Bracewell, and E. Sorensen. Dynamic modelling of aqueous two-phase systems to quantify the impact of bioprocess design, operation and variability. *Food and Bioproducts Processing*, 107:10–24, jan 2018.

J. F. J. F. Richardson, J. H. J. H. Harker, J. R. Backhurst, and J. M. J. M. Coulson. *Coulson and Richardson's chemical engineering. Vol. 2, Particle technology and separation processes.* Butterworth-Heinemann, 2002. ISBN 9780080490649.

M. Rito-Palomares. Practical application of aqueous two-phase partition to process development for the recovery of biological products. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, 807(1):3–11, 2004. doi: 10.1016/j.jchromb.2004.01.008.

M. Rito-Palomares and A. Lyddiatt. Practical implementation of aqueous two-phase processes for protein recovery from yeast. *Journal of Chemical Technology & Biotechnology*, 75(7):632–638, jul 2000. doi: 10.1002/1097-4660(200007)75:7<632::AID-JCTB248>3.0.CO;2-7.

M. Rito-Palomares and A. P. Middelberg. Aqueous two-phase systems for the recovery of a recombinant viral coat protein fromEscherichia coli. *Journal of Chemical Technology & Biotechnology*, 77(9):1025–1029, sep 2002. ISSN 0268-2575. doi: 10.1002/jctb.673.

M. Rito-Palomares, C. Dale, and A. Lyddiatt. Generic application of an aqueous two-phase process for protein recovery from animal blood. *Process Biochemistry*, 35(7):665–673, 2000. doi: 10.1016/S0032-9592(99)00119-3.

P. Rosa, A. Azevedo, I. Ferreira, S. Sommerfeld, W. Bäcker, and M. Aires-Barros. Downstream processing of antibodies: Single-stage versus multi-stage aqueous two-phase extraction. *Journal of Chromatography A*, 1216 (50):8741–8749, dec 2009a. doi: 10.1016/j.chroma.2009.02.024.

P. A. J. Rosa, A. M. Azevedo, S. Sommerfeld, A. Mutter, M. R. Aires-Barros, and W. Bäcker. Application of aqueous two-phase systems to antibody purification: a multi-stage approach. *Journal of biotechnology*, 139(4):306–13, feb 2009b. doi: 10.1016/j.jbiotec.2009.01.001.

P. a. J. Rosa, I. F. Ferreira, a. M. Azevedo, and M. R. Aires-Barros. Aqueous two-phase systems: A viable platform in the manufacturing of biopharmaceuticals. *Journal of Chromatography A*, 1217(16):2296–2305, 2010. ISSN 00219673. doi: 10.1016/j.chroma.2009.11.034.

P. a. J. Rosa, A. M. Azevedo, S. Sommerfeld, M. Mutter, W. Bäcker, and M. R. Aires-Barros. Continuous purification of antibodies from cell culture supernatant with aqueous two-phase systems: From concept to process. *Biotechnology Journal*, 8(March):352–362, 2013. ISSN 18606768. doi: 10.1002/biot.201200031.

F. Ruiz-Ruiz, J. Benavides, O. Aguilar, and M. Rito-Palomares. Aqueous two-phase affinity partitioning systems: Current applications and trends. *Journal of Chromatography A*, 1244:1–13, 2012. doi: 10.1016/j.chroma.2012.04.077.

M. Salamanca, J. Merchuk, B. Andrews, and J. Asenjo. On the kinetics of phase separation in aqueous two-phase systems. *Journal of Chromatography B: Biomedical Sciences and Applications*, 711(1-2):319–329, 1998. doi: 10.1016/S0378-4347(98)00173-X.

J. C. Salgado, B. A. Andrews, M. F. Ortuzar, and J. A. Asenjo. Prediction of the partitioning behaviour of proteins in aqueous two-phase systems using only their amino acid composition. *Journal of Chromatography A*, 1178(1-2):134–144, jan 2008. ISSN 00219673. doi: 10.1016/j.chroma.2007.11.064.

J. D. Seader, E. J. Henley, and Roper. *Separation process principles: Chemical and Biochemical operations*. Wiley, 2006. ISBN 9780470481837.

R. R. G. Soares, A. M. Azevedo, J. M. Van Alstine, and M. R. Aires-Barros. Partitioning in aqueous two-phase systems: Analysis of strengths, weaknesses, opportunities and threats. *Biotechnology Journal*, 10(8):1158–1169, aug 2015. doi: 0.1002/biot.201400532.

M. A. Torres-Acosta, K. Mayolo-Deloisa, J. González-Valdez, and M. Rito-Palomares. Aqueous Two-Phase Systems at Large Scale: Challenges and Opportunities. *Biotechnology Journal*, 14(1):1800117, jan 2019. ISSN 18606768. doi: 10.1002/biot.201800117.