Experimental studies and artificial neural networks (ANN) modeling of asparaginase partitioning in aqueous two-phase systems: Polyethylene glycol and inorganic salts

Mohsen Mesbahi-Nowrouzi¹, Soroush Sardari¹,*
¹Drug Design and Bioinformatics Unit, Medical Biotechnology Department, Biotechnology Research Centre, Pasteur Institute of Iran, Tehran, Iran
*Corresponding author: Soroush Sardari, Drug Design and Bioinformatics Unit, Medical Biotechnology Department, Biotechnology Research Centre, Pasteur Institute of Iran, Tehran, Iran. E-mail: sardari@pasteur.ac.ir
DOI: 10.22034/HBB.2023.03

Received: October 8, 2022; Accepted: October 26, 2022

ABSTRACT

This article studied applicabilities of Artificial Neural Networks (ANN), aiming at partitioning of Asparaginase (ASP). The first step investigated the partitioning of ASP using ATPS, consisting of polyethylene glycol and different inorganic salts. The studies of ASP partition were conducted by measuring the amounts of ASP in the upper and lower phase, followed by calculating the recovery. PEG of 6000 and (NH₄)₂SO₄ provided the best performance with recovery yield of 86 %. The partitioning behavior of the approach was investigated in the presence of an impurity kind of amylase. Related partition coefficient of 0.21 and recovery yield of ASP of 86 % were calculated. A Circular Dichroism (CD) spectroscopy compared the structure of standard and partitioned ASP, which remained stable. The relative activity was determined to be 94 %. A feed-forward neural network was applied to predict the best model to partition of ASP.

Keywords: Asparaginase; ATPS; artificial neural networks

INTRODUCTION

Nowadays, aging population and deficiency of traditional drugs increase the demand for new therapies such as biopharmaceutical products [1]. Biotechnology develops the upstream processing and improves their productions from various sources in higher yields. Nevertheless, the downstream processes are neglected because of less tendency of biopharmaceutical companies to replace
ANN modeling of asparaginase partitioning

close common processes with new strategies. The limit for this replacement is that these manufactures should allocate 80% of the total manufacturing cost to downstream processing. Now chromatography method is mainly practical in this step. However, it has restrictions such as complexity, operating, low capability in process integration, laborious processing cycles, economical aspects and environmental matters [2-4]. The researchers put significant efforts to present new non-chromatography strategies to develop the process throughput and decrease the processing time. Liquid–liquid extraction [5], membrane separation [6], high performance tangential flow filtration [7], affinity precipitation [8], high gradient magnetic fishing [9] and crystallization [10] are some common examples of these methods. Aqueous two-phase system is another qualified candidate in extraction, separation, purification, and enrichment of biomolecules [11]. These systems are formed by mixing two immiscible components for example two immiscible, above their critical concentrations [12,13]. ATPSs are easy to be operated, reliable, environmentally biocompatible, energy efficient and inexpensive. Moreover, it is one-step process and steps of clarification, concentration and purification can be integrated [14].

Up to now, different analytes were separated by this method. For example, a carbohydrate-based aqueous two-phase system separated and purified glatiramer acetate and amino acids [15]. Also, a mixture of PEG and DEX was used to remove soluble aggregates of IgG by ATPS [16]. In another research, Azevedo presented ATPS of PEG and dextran for the integrated purification of the anti-CD34 mAb [17]. Additionally, 15 proteins were partitioned in PEG–sodium sulfate ATPS successfully [18].

In this article, ATPSs, consisting of PEGs and inorganic salts, were applied to partition Asparaginase (ASP). ASP is a FDA-approved biopharmaceutical for Acute Lymphoblastic Leukemia (ALL). The phase diagrams such as binodal and tie-lines were determined, the effective parameters with respect to the kinds of polyethylene glycol, inorganic salts and pH were screened; afterwards, the partition coefficients and recoveries were reported. The partitioning of secondary protein (amylase) as an impurity was investigated. Structure and activity of ASP were analyzed and compared before and after partitioning with a standard sample. A
weakness of ATPSs application is their complexity so the partitioning mechanism is hard to be understood. In reference to, they are based on trial-and-error screening approaches. Artificial Neural Network (ANN) is an efficient approach to overcome this challenge, which models the partitioning ASP in ATPS in this study.

**MATERIALS AND METHODS**

Na$_3$C$_6$H$_5$O$_7$. 2H$_2$O, NaH$_2$PO$_4$, K$_2$HPO$_4$, (NH$_4$)$_2$SO$_4$, PEG (Mw: 4000, 6000 and 10000), L-Asparagine, Nessler’s reagent were purchased from Merck (Merck, Germany) in the analytical grade. The analytical grade of ASP, amylase and Trichloroacetic Acid (TCA) were prepared from Sigma-Aldrich. The double-distilled water is used in all experiments.

**Phase diagram determination**

The determination of phase diagram for the systems was conducted at room temperature (25 ± 1 °C) and atmospheric pressure by using the cloud-point method [19]. The system was prepared via adding exact amounts of salt solution to the polymer solution. At a critical point (the could-point), the mixture turn to a turbid sample and a two-phase system was achieved. The mixture composition was recorded, followed by providing a point on the binodal curve. Then, through adding distilled water drop by drop to the tube, a clear one-phase system was achieved and adding more salt solution turn it to a two-phase system. Merchuk’s equation (Eq. 1) was applied to fit the binodal curve [20].

\[
[Y]_r = A \times \exp\left((B \times [X]_r^{0.5}) - (C \times [X]^3)\right)
\]  

(1)

Where [Y] and [X] stand for polymer and salt weight percentages in binodal curves. The constants of A, B and C are found by regression of the experimental data. The components of each phase and the related tie-lines (TLs) were calculated using the lever-arm rule regarding Eq. (2) – (5) [21]:

\[
[Y]_r = A \times \exp\left((B \times [X]_r^{0.5}) - (C \times [X]^3)\right)
\]

(2)

\[
[Y]_B = A \times \exp\left((B \times [X]_B^{0.5}) - (C \times [X]^3)\right)
\]

(3)

\[
[X]_r = \frac{[Y]_m}{m} - \left(\frac{1-m}{m}\right) \times [Y]_B
\]

(4)

\[
[X]_r = \frac{[X]_m}{m} - \left(\frac{1-m}{m}\right) \times [X]_B
\]

(5)

In the equations of 2 to 5, M, T and B define the mixture point, top and bottom phase, respectively; m defines as following:

\[
m = \frac{M_{top}}{M_{total}}
\]

(6)

Eq. 7 determines the tie-line lengths (TLLs).
ANN modeling of asparaginase partitioning

\[ Y_T = \frac{[ASP]_T \times V_t}{[ASP]_i \times V_i} \]

(9)

In this equation, [ASP]_i and V_i are ASP concentration and volume of the phase containing more concentration of ASP, respectively. [ASP]_T expresses the concentration and V_T is the volume of the stock solution added to the system [23].

Analytical techniques section

Protein assay

The concentrations of protein were evaluated with Bradford approach that ASP was as the standard [24].

UV Circular Dichroism (CD) spectroscopy

The analysis started by adding (NH₄)₂SO₄ to the sample of ASP to increased its concentration and then centrifuged at 5000 rpm for 20 min and temperature of 25 ± 1 °C. Afterwards, the supernatant was thrown and the residual solution was added to a centrifugal filter unit with a 50 kDa molecular weight cut-off. The solution was centrifuged at 5000 rpm for 40 minutes at room temperature. The ASP structure was followed by a far UV CD spectroscopy (Jasco, J-810; Japan). The protein concentration was 0.5 mg mL⁻¹. J700 CD-JASCO software calculated the secondary structure parameters. Since ellipticity values depend on the ASP...
concentration, a same concentration of ASP was applied as treated and untreated samples, regarding Bradford method. Each experiment was conducted in triplicate and the related mean values were considered for comparison.

**ASP activity assay**
Wriston and Yellin’s method determines ASP activity, using nesslerization to measure the released ammonia concentration [25]. The sample (2 mL) containing of ASP (1.5 mL, 0.04 M), enzyme (0.5 mL) and Tris-HCl buffer (0.05 M, pH 8.6) was prepared, followed by being incubated at 37 °C for 30 min. TCA (0.5 ml, 1.5 M) ended the reaction.
The precipitated protein was removed by a centrifuge (5 min) and the top solution (0.5 ml) and distilled water (7 ml) was collected to quantify the released ammonia by adding Nessler’s reagent (1 ml). After 20 min, yellow color of mixture represents ammonia, which was measured by a UV-Visible spectrophotometer at the absorbance of 480 nm (Bio-Rad Laboratories, SmartSpec Plus, USA). The amount of released ammonia from the test sample was carried out by comparing the absorbance of solutions of NH₄Cl as the reference. One Unit (U) of ASP is the amount of enzyme, which released one μmole of ammonia at 37 °C at pH of 8.6 in one min. The activity ratio of treated ASP and untreated ASP was reported as relative activity.

**ANN modeling**
The ‘nntool’ in MATLAB toolbox (MATLAB R2020b, MathWorks, Natick, Mass., USA) analyzed the data in ANN model. The backpropagation algorithm performs learning on a two-layer feed-forward neural network. In this technique, a feed-forward algorithm was applied in order to investigate the output, along with backpropagation, related to the weights recursively. The Levenberg–Marquardt backpropagation Algorithm (LMA) was the best fitted algorithm to train the network (input: 15 parameters and output: one output (KASP)). Different networks with different numbers of hidden neurons were used and a one-hidden layer model was selected because it was effective on modeling the KASP of ASP in ATPS. The network optimization, such as optimum numbers of neurons in the hidden layer, plays important part in ANN modeling. Accordingly, different numbers, ranging from one to 10 were selected and the input and output variables were randomly divided to 70% for training and 15% for network validation. The remain 15% was
used for completely independent test of network generalization. Hyperbolic tangent sigmoid transfer function is used for the hidden and the output layers. During the learning process, the estimated output and the target were compared together, also the system error generates a model with a particular weight to express the selected system property.

Evaluation of the models was completed by comparison the model accuracy, separately for (i) training, (ii) validation, (iii) testing, and (iv) overall model, according to correlation coefficient (R), determination coefficient (R²) and mean squared error (MSE) (Eq. 11 – 13).

\[ MSE = \frac{1}{N} \sum_{i=1}^{N} (t_i - y_i)^2 \]  
(11)

\[ MSE = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{t_i - y_i}{t_i} \right| \]  
(12)

\[ R^2 = \frac{\sum_{i=1}^{N} (t_i - \bar{y})^2 - \sum_{i=1}^{N} (t_i - y_i)^2}{\sum_{i=1}^{N} (t_i - y_i)^2} \]  
(13)

In these equations, N, t and y are the number of data points, the target (experimental) data, and the anticipated value, respectively.

**RESULTS**

**Phase diagrams: Binodal curves and tie-lines**

Data on binodal curves of ATPSs under study shows areas where phase separation and consequently extraction could be happened. The binodal curves of ATPSs were plotted, eight tie lines were drawn and final compositions including salts, PEG and water were achieved by linear regression fitting (Figure 1 a-l). The constants and the correlation coefficients (R²) are calculated based on the mentioned equations 1-7 (Table 1).

**Preliminary studies on the partition behaviour of ASP**

In this study, the partition coefficients of ASP in an ATPS were determined. The maximum recovery (96 %) of ASP was obtained in the bottom phase consisting of PEG 4000 (15 %) and Na₃C₆H₅O₇. 2H₂O (25 %) at the pH of 6.8. The concentrations of PEG and salt impact on the ATPS recovery yield. However, the recovery of 88 %, which is related to sample of PEG 10000 (25 %) and K₂HPO₄ (16 %) shows a promising result.

**The optimal separation conditions**

The selectivity of the ASP separation procedure was investigated in the presence of an impurity kind of amylase (one mM)
and the feed (10 g). Table 2 summarizes the reports. In S1, $K_p$ amount is close to one (0.98), confirming same concentration of ASP in both phases and not being partitioned. This system is not recommended. The amounts of $K_p$ increased in systems S2 and S3, which show efficient ASP separation because ASP and amylase were partitioned in two different phases. S2 and S3 were selected for further experiments by CD.

**CD Spectroscopy & Steady-state kinetic study**

CD results of S2 were not satisfying; however, ASP in S3 was isolated from the bottom phase of the optimized feed, showing no significant alterations in the amounts of secondary structures compared to the ASP standard solution (0.5 mg mL$^{-1}$). The results of S2 and S3 are provided.

**ANN modeling of asparaginase partitioning**

Figure. 2 (a – d). The secondary structure estimation of standard ASP showed $\alpha$-Helix (23.5 %), $\beta$-Sheet (14.1 %), $\beta$-Turn (29.4 %) and random structure (33.1 %). These structures in S2 were characterized as 19.5 %, 33.7 %, 18.1 % and 28.7 % for the structures of $\alpha$-Helix, $\beta$-Sheet, $\beta$-Turn and random structure, respectively. But, the estimation of the secondary structure of isolated ASP from the optimum feed illustrates 24.6 % of $\alpha$-Helix, 18.4 % of $\beta$-Sheet, 23.4 % of $\beta$-Turn and 33.6 % of random structure. The standard and the optimum feed, in Figure 2, shows a minimum peak at 216 nm, attributing to the $\alpha$-helix secondary structure, no significant differences were observed.
Table 1 Merchuk’s parameters and $R^2$ obtained by fitting Eq. 1

<table>
<thead>
<tr>
<th>Salt</th>
<th>Polymer</th>
<th>$A$</th>
<th>$B$</th>
<th>$10^5C$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na$_2$C$_6$H$_5$O$_7$, 2H$_2$O</td>
<td>PEG 4000</td>
<td>60.1950</td>
<td>-0.3493</td>
<td>47.90513</td>
<td>0.995</td>
</tr>
<tr>
<td></td>
<td>PEG 6000</td>
<td>57.2989</td>
<td>-0.3590</td>
<td>58.97410</td>
<td>0.974</td>
</tr>
<tr>
<td></td>
<td>PEG 10000</td>
<td>81.0922</td>
<td>-0.4983</td>
<td>82.44736</td>
<td>0.989</td>
</tr>
<tr>
<td>NaH$_2$PO$_4$</td>
<td>PEG 4000</td>
<td>72.1761</td>
<td>-0.3678</td>
<td>24.37894</td>
<td>0.999</td>
</tr>
<tr>
<td>K$_2$HPO$_4$</td>
<td>PEG 6000</td>
<td>67.4228</td>
<td>-0.3213</td>
<td>27.27032</td>
<td>0.997</td>
</tr>
<tr>
<td></td>
<td>PEG 10000</td>
<td>79.8814</td>
<td>-0.4056</td>
<td>44.74472</td>
<td>0.998</td>
</tr>
<tr>
<td>(NH$_4$)$_2$SO$_4$</td>
<td>PEG 4000</td>
<td>65.6376</td>
<td>-0.4646</td>
<td>116.1371</td>
<td>0.974</td>
</tr>
<tr>
<td></td>
<td>PEG 6000</td>
<td>72.4897</td>
<td>-0.5296</td>
<td>110.3154</td>
<td>0.989</td>
</tr>
<tr>
<td></td>
<td>PEG 10000</td>
<td>128.4450</td>
<td>-0.7952</td>
<td>177.413</td>
<td>0.994</td>
</tr>
<tr>
<td></td>
<td>PEG 4000</td>
<td>309.5294</td>
<td>-1.0993</td>
<td>11.73635</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>PEG 6000</td>
<td>304.9436</td>
<td>-1.0747</td>
<td>10.82578</td>
<td>0.979</td>
</tr>
<tr>
<td></td>
<td>PEG 10000</td>
<td>110.4330</td>
<td>-0.6179</td>
<td>119.3288</td>
<td>0.989</td>
</tr>
</tbody>
</table>

Table 2 Asparaginase and Amylase partition coefficients and recovery yield parameters for systems based on polymer, salt and water

<table>
<thead>
<tr>
<th>system</th>
<th>PEG</th>
<th>Salt</th>
<th>pH</th>
<th>ASP $K_P$</th>
<th>ASP $Y_T$</th>
<th>Amylase $K_P$</th>
<th>Amylase $Y_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>S$_1$</td>
<td>PEG 4000 (15%)</td>
<td>Na$_2$C$_6$H$_5$O$_7$, 2H$_2$O (25%)</td>
<td>6.8</td>
<td>0.08</td>
<td>96</td>
<td>0.98</td>
<td>92</td>
</tr>
<tr>
<td>S$_2$</td>
<td>PEG 10000 (25%)</td>
<td>K$_2$HPO$_4$ (16%)</td>
<td>6.8</td>
<td>0.18</td>
<td>88</td>
<td>2.66</td>
<td>89</td>
</tr>
<tr>
<td>S$_3$</td>
<td>PEG 6000 (26%)</td>
<td>(NH$_4$)$_2$SO$_4$ (15%)</td>
<td>6.8</td>
<td>0.21</td>
<td>86</td>
<td>2.54</td>
<td>90</td>
</tr>
</tbody>
</table>
Figure 1. Phase diagrams comprising the binodal curves and respective experimental TLs for ATPSs composed of different polymers and salts.
Figure 2. CD spectrum of ASP standard solution and ASP sampled from the lower phase of the optimum feed (a and b related to \( S_2 \) and c and d related to \( S_3 \)).
Steady-state kinetic study compares specific activity of standard ASP and partitioned ASP in S3. Reaction time was achieved by calculating linear range of released ammonia from standard ASP and ASP of S3. Regarding Figure 3, up to 30 minutes, the quantity of released ammonia was linear and R² values for standard ASP and ASP related in S3 were 0.991 and 0.990, respectively. As a result, specific activity of ASP from the bottom phase of the optimum feed (S3) was found close to standard aqueous solution ASP (190 IU) and was around 179 IU.

**ANN Model**

For the selected models, the percentage of outliers using 95% confidence bounds was found. A diagram of the ANN model is presented in Figure 4. Figure 5 shows the R-value of dispersion of output data with respect to their corresponding target values. The high amount of R² (0.99159) confirms an efficient model, illustrating a real relationship between the independent variables. 99% of the variation in the response could be explained by the model. The deviation between network output and target value is more at the beginning stage of the training; however, by increasing the epochs, amounts of output and target were converged, confirming the learning relationship between input and output for the above network (Figure 6).

**DISCUSSION**

This article presented ATPS application to partition a biopharmaceutical, kind of ASP. The ATPS was consisted of three PEGs (Mw: 4000, 6000 and 10000) and inorganic salts (Na₃C₆H₅O₇·2H₂O, NaH₂PO₄, K₂HPO₄ and (NH₄)₂SO₄). In this study, the upper phase is PEG-rich phase and the lower phase is mainly composed of salt. However, due to viscosity of PEG, some samples were formed in gel and if this point is close to the critical point, the data is omitted. In this sample, the phases were converted. Phase separation depends on the type and concentration of salt and molecular weight, weight percent, volume ratio of the phase and equilibrium characteristics of polymer in the systems. According to data, by increasing weight of PEG, separation area of the ATPS increased so lower amounts of PEGs and salts are required for phase separation, justifying
by hydrophobicity of polymer, which increases in PEGs with more Mw [26,27]. Moreover, solute such as a salt on the miscibility of an aqueous solution (salting-out effect) is a basis of phase formation in ATPS. This effect is related to the hydration strength of the salt [28-35].

The partition coefficients of proteins in ATPS are dependable on two parameters. First the PEG exclusion effect, which changes the availability of free volume for proteins in the PEG rich phase. Secondly the PEG–protein binding depends on the hydrophobic area of the protein exposed to the solvent; hence, increasing this area rises the tendency of the proteins to be in the upper (PEG rich) phase. By increasing Mw of PEG, hydrophobicity grows; consequently, hydrophobic protein moves to the upper phase (PEG rich) and finally $K_{ASP}$ enhances [31]. Also, protein molecules were aggregated in the presence of PEG due to hydrophobic force, which is another reason of hydrophobic affinity of protein.

Charge of protein controls pH effect on $K_{ASP}$, which depends on pH of the solution and isoelectric point (pI) of the protein. The pH of ASP is about 9.8 [32]. In this study, effect of pH, ranging from 6.8 to 8, on $K_{ASP}$ in the presence of PEG 10000 is more significant that of PEG 4000; accordingly, ASP concentration with negative charge increases in PEG rich phase. The relationship between polymer concentration or Mw and the free volume between polymer networks is indirect, reducing the accessible space for the biomolecules in the upper (PEG rich) phase. This phenomenon is volume exclusion effect, causes more partitioning of ASP in the lower (salt rich) phase. In contrast, resulting from salting out effect, by increasing the salt concentration in the salt rich phase, the affinity of ASP in this phase decreases, and ASP concentration increases in the PEG rich phase [33]. Also, increasing the PEG molecular weight may partition the hydrophilic enzyme to the lower (salt-rich) phase, because of increasing the hydrophobicity of PEG with the chain length [34,35].
Figure 3. Bioactivity of (a) ASP standard (b) ASP sampled from the lower phase of the optimum feed related to S3.

Figure 4. nntool created feed-forward neural network.
Figure 5. The linear regression between the network output values (predictive) and the target values (experimental) for $K_{\text{ASP}}$. 
CONCLUSION

The results express that there is a common tendency for the majority of systems examined here and ASP was distributed mostly in the bottom phase (K<1). Furthermore, increasing the salt concentration have positive effect on ionic strength and enhance the protein movement to other phase (PEG rich phase) via electrostatic force so the $K_{ASP}$ increases. $K_{ASP}$ increases with the PEG concentration in upper (PEG rich) phase. Applicability of ATPS for separation depends on selective partition of molecules and impurities. Surface properties of proteins and phase properties effect on different phase preferences of ASP and amylase. Comparing the results indicate that in the pH of 6.8, ASP was partitioned to the lower phase, while amylase was partitioned to the upper phase. In the mentioned pH, ASP has positive charge, because its pH is about 9.8. Accordingly, the behavior of ASP separation depends on repulsion between
Sardari et al.  
the charge of ASP surface and hydrogen ions at pH=6.8. CD data for S3 confirm that the structure of ASP is unaffected and remain stable. Also, the activity was followed and the relative activity was reported. The ASP partitioning was modeled by a two-layer feed-forward neural network successfully.

REFERENCES

[6]. Chen X, et al. Separation of recombinant monoclonal antibodies

ANN modeling of asparaginase partitioning  


**ANN modeling of asparaginase partitioning**


Sardari et al.  


**ANN modeling of asparaginase partitioning**


