INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY

CAMPUS MONTERREY

DIVISIÓN DE INGENIERÍA Y ARQUITECTURA
PROGRAMA DE GRADUADOS EN INGENIERÍA

TECNOLÓGICO
DE MONTERREY®

RECOVERY AND CHARACTERIZATION OF PROTEINS FROM PLANTS: AQUEOUS TWO-PHASE SYSTEM-BASED BIOENGINEERING STRATEGIES

TESIS

PRESENTADA COMO REQUISITO PARCIAL PARA OBTENER
EL GRADO ACADÉMICO DE:

DOCTOR EN CIENCIAS DE INGENIERIA

POR:

OSCAR ALEJANDRO AGUILAR JIMÉNEZ

MONTERREY, N.L. **DICIEMBRE DE 2007**

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY

CAMPUS MONTERREY

DIVISIÓN DE INGENIERÍA Y ARQUITECTURA

PROGRAMA DE GRADUADOS EN INGENIERÍA

Los miembros del comité de tesis recomendamos que la presente disertación presentada por el M. en C. Oscar Alejandro Aguilar Jiménez sea aceptado como requisito parcial para obtener el grado académico de:

DOCTOR EN CIENCIAS DE INGENIERÍA

Comité de Tesis:

_____
Marco A. Rito-Palomares, Ph.D
Asesor

_____
Charles E. Glatz, PhD
Sinodal

_____
Blanca H. Lapizco Encinas, PhD
Sinodal

_____
Rocío I. Díaz de la Garza, PhD
Sinodal

_____
Manuel I. Zertuche Guerra, PhD
Sinodal

_____
Francisco Angel Bello, PhD.
Director del Programa de Graduados en Ingeniería

Diciembre de 2007

ABSTRACT

The main drawback for general acceptance of plants as economically viable production systems is the lack of efficient initial concentration and separation procedures. In order to facilitate the general acceptance of plants as bioreactors, the establishment of efficient downstream operations is critical. It has been established that with the general knowledge of the molecular properties of contaminant proteins, the selection and design of suitable downstream strategies for recombinant proteins can be improved. The present dissertation addresses the potential use of quantitative 2D electrophoresis (2-DE) coupled with hydrophobic partitioning in aqueous two-phase systems (ATPS) for three-dimensional characterization of proteins from plant extracts. The application of this experimental approach to soybean proteins resulted in molecular characterization of proteins. Molecular weight (MW), isoelectric point (pI) and hydrophobicity were measured simultaneously and demonstrated that this technique can be a valuable tool for predictive design of recovery steps for recombinant proteins from plants. The extension of this experimental approach in alfalfa green tissue extracts containing a model recombinant protein provided additional information on the molecular properties of the main host proteins that will allow the design of pre-fractionation and purification methods to facilitate its recovery from alfalfa extracts. As a result of the application of this three-dimensional characterization technique to soybean and alfalfa protein extracts, more efficient downstream strategies could be designed for recovery of recombinant proteins, facilitating the future adoption of plants as a production system.

ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude to many people who made this possible.

First to my family, for supporting all my decisions and for being there every time I needed help. For their support during all these long school years while I have been the eternal student. I know that you always trusted me and this degree is the best testimony I can give back to you to express my eternal gratitude.

To all the people that greatly influenced my academic development: Dr. Marco Rito-Palomares for his constant supervision, guidance and for providing me with the most powerful tool that a novel researcher could receive: freedom and support to explore new ideas without letting me lose the way.

To Dr. Carmen Hernandez for being part of my initial formation as a young scientist during my master's studies.

To Dr. Charles Glatz who gave me the chance of being part of his excellent research group in Iowa State University. I deeply appreciate this opportunity to work with such a great researcher and an excellent person.

To Dr. Manuel Zertuche, Dra. Blanca Lapizco and Dra. Rocio Diaz, the members of my doctoral committee for your valuable time and critical comments during the review of this document.

To all my friends at the Biotechnology lab, where I learned the real meaning of collaborative research. Dany, Irasema, Jenny, Denisse, Gaby, Mayra, Armando, Tanhia and Esther, thank you for making these years a pleasant journey. And especially to the Biotech Gang: Ily, George, Bertha, Santiago, Daris, Jane, Marcos, Laura, Perla and the recently added Pepe; thanks for those priceless adventures we had (and the ones coming) inside and outside the lab.

Your worst enemy cannot harm you as much
as your own unguarded thoughts.
- Buddha -

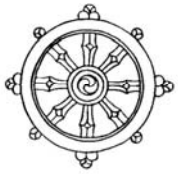

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

Page

Page

Page

LIST OF FIGURES

Page

CHAPTER FOUR

CHAPTER FIVE

LIST OF ABBREVIATIONS

°C	Celsius degrees
2-DE	Two-dimensional electrophoresis
3D	Three-dimensional
AMPD	2-amino-2-methyl-1,3-propanediol
ATP	Aqueous two-phase
ATPS	Aqueous two-phase systems
Ave	Avenue
BSA	Bovine serum albumin
CA	California
CHAPS	3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate
CHY	Chymotrypsin A
cm	Centimenter
Corp	Corporation
CYC	Cytochrome C
D	Dextro optical isomer
DIGE	2-D difference in-gel electrophoresis
dm	Decimeter
dpi	Dots per inch
DTT	Dithiothreitol
E.C.	Enzyme commission
EDTA	Ethylenediamine tetraacetic acid
Et al	From latin *et alii* “and others”
g	Gram
g	Gravities
GE	General Electric
GMP	Good manufacturing practices
GUS	β-glucuronidase
h	Hour
HIC	Hydrophobic interaction chromatography
i.e.	From latin *id est* “that is”
IEF	Isoelectric focusing
IPG	Immobilized pH gradient
kDa	Kilo Daltons
Kp	Partition coefficient
L	Levo optical isomer
LAC	α-Lactalbumin
log	Logarithm
LSBC	Large-Scale Biology Corporation
LYS	Lysozyme
M	Molar
mg	Milligram
min	Minute
mL	Milliliter
mM	Millimolar

MO	Missouri
Mr	Relative molecular weight
MW	Molecular weight
N.L.	Nuevo León
nm	Nanometer
NY	New York
PEG	Poly(ethyleneglycol)
pH	P function of the hydrogen ion concentration
Ph	Phone
pI	Isoelectric point
pmol	Picomol
PMSF	Phenylmethylsulfonyl fluride
PNPG	4-nitrophenyl-β-D-glucuronide
ppm	Parts per million
PVP	Polivinylpirrolidone
rhGM-CSF	Recombinant human granulocyte and macrophage colony stimulating factor
RNA	Ribonuclease A
RPC	Reverse phase chromatography
Rubisco	Ribulose-1,5-bis phosphate carboxylase
SDS	Sodium dodecyl sulfate
TBE	Tris-Borate-EDTA
TCA	Trichloroacetic acid
TLL	Tie-line length
TSP	Total soluble protein
U	Unit
US	United States
V	Volts
Vh	Volts*hour
Vr	Volume ratio
w/w	Weight ratio
μg	Microgram
μL	Microliter
μm	Micrometer

CHAPTER ONE – INTRODUCTION

1.1 PLANT-BASED PROTEIN PHARMACEUTICALS

Plants have long been used as a source of medicines, and almost a quarter of the medicines used today come from them (Fischer and Emans, 2000). Many therapeutically active proteins have been identified in the past decades, and their development in useful medicines has driven a demand for large-scale production. Molecular biotechnology has changed the nature of medicines that can be isolated from plants, through the development of the technology to produce recombinant proteins in crops (Fischer et al., 2002; Twyman et al., 2003). *Molecular farming* is defined as the large scale production of recombinant proteins in living cells or organisms; frequently applied to the use of crop plants or domestic animals as expression hosts because of the allusion to agriculture (Ma et al., 2003).

The use of vegetable cells as an expression system for human and therapeutic protein production is advantageous compared with human or animal fluids/tissues, recombinant microbes, transfected animal cell lines, or transgenic animals. These features were previously analyzed by Fischer and Emans (2000) and summarized in Table 1.1.

Daniell *et al*. (2001) reported that the cost of producing an IgG from alfalfa grown in 250 m^2 greenhouse is estimated to be 10 times lower compared with hybridoma cells. Two of the most expensive drugs, glucocerebrosidase and granulocyte-macrophage colony-stimulating factor have been successfully expressed in plant cells. It is clear that plants can be used as inexpensive factories of therapeutic proteins for humans (Kwon et al., 2003; James, et al., 2000; Giddings et al., 2000).

Table 1.1 Characteristics of different systems for recombinant protein production (Schillberg et al., 2003).

Factor	Transgenic plants	Plant cell cultures	Bacteria	Yeast	Mammalian cell culture	Transgenic animals
Costs						
Production costs	Low	Medium	Low	Medium	High	High
Time effort	High	Medium	Low	Medium	High	High
Scale-up costs	Low	High	High	High	High	High
Propagation	Easy	Easy	Easy	Easy	Limited	Possible
Productivity	High	Medium	Medium	Medium	Medium	High
Quality						
Product quality and homogeneity	High	High	Low	Medium	High	High
Glycosylation	Authentic	Authentic	Absent	Incorrect	Authentic	Authentic
Contamination risk	No	No	Yes	No	Yes	Yes
Practical application						
Data monitoring	Difficult	Easy	Easy	Easy	Easy	Difficult
Etical concerns	Medium	Low	Low	Low	Medium	High
GMP conformity	Difficult	Possible	Possible	Possible	Possible	Possible
Storage	Inexpensive /room temp.	Inexpensive /-20°C	Inexpensive /-20°C	Inexpensive /-20°C	Expensive/N_2	Expensive/N_2

Since 1997, Sigma Chemical Co. and ProdiGene Inc. have been commercializing introductory quantities of recombinant Avidin and β-glucuronidase (GUS) produced in transgenic maize, to Sigma-Aldrich customers in the research, industrial and pharmaceutical markets. In March 2004, Large Scale Biology Corp. and Sigma-Aldrich Fine Chemicals, announced commercial distribution of APRONEXIN® NP an animal-free recombinant plant-made bovine-sequence aprotinin produced using LSBC's proprietary GENEWARE® expression technology, a transient plant viral-vector-based gene expression technology (LSBC Homepage, 2006).

Despite the numerous economical and technical advantages of the use of crops for molecular farming compared with the most conventional systems, the main factor for the establishment of a plant-based production system is the ease of purification. Bioseparation strategies have a direct influence on the selection of the expression system (Larrick et al., 2001).

1.2 THE NEED FOR EFFICIENT DOWNSTREAM OPERATIONS

The general process for production of a recombinant protein starting from the plant material to the final formulation includes extraction, clarification, protein capture, purification and polishing (Menkhaus et al., 2004). Previous studies (Zhang and Glatz, 1999; Zhang et al., 2001) have shown the potential of traditional downstream processing operations applied to seed-produced recombinant proteins. Furthermore, the use of genetic engineering strategies for product recovery from bulk storage proteins where the product of interest is immersed has also been documented (James et al., 2000; Zhang et al., 2001). However, the main challenge is still the establishment of efficient recovery and purification steps (i.e. inexpensive processes with minimal losses of target protein) that allow the production of pharmaceutical proteins with advantages over traditional microbial/cell-based production systems.

In the case of pharmaceutical proteins that need to be purified before use, several strategies have been developed to reduce costs associated with downstream processing. General estimates agree that up to 95% of the production cost of a recombinant protein by molecular farming comes from extracting and purifying the protein from the plant source (Twyman et al., 2003; Evangelista et al., 1998). The overall production cost is mainly determined by the efficiency of the initial capture and purification steps where feed volumes are large. Volume

reduction is achieved once biomass solids and oils are removed and protein is concentrated (Menkhaus et al., 2004). An economic evaluation of the recovery of a transgenic protein from corn estimated 28% of the operative cost to be associated to initial capture step on DEAE-sepharose (Evangelista et al., 1998). The potential economical benefits of substitution of costly unit operations such as chromatography by aqueous two-phase systems (ATPS) without a decrease in the yield have been previously addressed (Aguilar et al., 2006; Benavides and Rito-Palomares, 2006). It is anticipated that similar strategies can be applicable for plant-made products (Gu and Glatz, 2007a). Aqueous two-phase partitioning has emerged as a practical technique that allows recovery and purification of biological products. This technique described first by Albertsson (1958) exploits mild hydrophobic interactions between proteins. ATPS are formed when two water-soluble polymers (polyethyleneglycol, dextran, etc.) or a polymer and a salt (potassium phosphate, sodium citrate, sodium sulfate, etc.) are mixed in aqueous solutions at a given proportion beyond the critical concentration. Separation of proteins is achieved by manipulating the partition coefficient of the proteins, varying the molecular weight of the polymers, the ionic strength of the salts, the relative proportion of each component, the pH, etc. (Rito-Palomares, 2004).

It has been stated previously that the establishment of adequate downstream strategies requires the characterization of the contaminant proteins from plant extracts (Asenjo and Andrews, 2004). It is clear that a better understanding of the molecular characteristics of the potential contaminants (MW distribution, hydrophobicity, pI, etc) will be beneficial for the selection, optimization and design of the downstream strategies.

1.3 TWO-PHASE PARTITIONING AND PROTEOMIC TOOLS FOR CHARACTERIZATION OF BIOLOGICAL PRODUCTS

1.3.1 Protein characterization using Aqueous Two-Phase Systems

Partitioning in aqueous two-phase systems (ATPS) is basically a surface-dependent phenomenon; it is a process where the exposed residues of the protein come in contact with the phase-forming components (Su and Chiang, 2006). Several protein properties influence this behavior: charge, size, hydrophobicity, and affinity. From these properties, hydrophobicity plays an important role during partitioning especially for systems with high molecular weight polymers or high salt content (Azevedo et al., 2007). Most of the reports on ATP partitioning describe it as a practical technique that allows recovery and purification of biological products. However, this technique is also useful in understanding the chemical properties and behavior of proteins in solution. The use of aqueous two-phase systems (ATPS) to measure the functional hydrophobicity of proteins has been reported by several authors (Andrews et al., 2005; Franco et al., 1996; Hachem et al., 1996) and it measures the surface hydrophobicity of a protein as the result of the interaction between the residues on the surface of the protein and the solvent.

It has been recently suggested that optimized ATPS can serve a presorting stage in proteomic studies since it can be customized for selective extraction and/or partition of large fractions of proteins from crude extracts (Roy et al., 2007).

1.3.2 Proteomic analysis of plant proteins

Proteomics, defined as the systematic analysis of the proteome, the protein complement of genome, allows the quantitative measurement of a large number of proteins that influence the cellular biochemistry. It provides an analysis of the cellular state during growth or development (Chen and Harmon, 2006). A proteomic analysis has usually two main steps: (i) protein separation and (ii) protein identification, which also includes the identification of post-translational modifications. So far, the two main approaches employed in proteomics are two-dimensional gel electrophoresis (extraction, destaining and protein digestion) followed by protein characterization by mass spectrometry (matrix-assisted laser desorption-ionization and time of flight; MALDI-TOF). Recently, proteomic analysis have been performed in "gel-less" condition, by using one- or two-dimensional liquid chromatography (2D-LC) for protein and peptide separation. It is usually combined with electrospray ionization (ESI) or tandem mass spectrometry (MS) for protein identification (Pirondini et al., 2006; Garbis et al., 2005).

The basic principles of two-dimensional electrophoresis (2-DE) have not undergone significant changes since first described 32 years ago by O'Farrell (1975). In this technique proteins are resolved in the first dimension along a pH gradient according to their isoelectric points. The gel is then transferred to the second dimension, typically sodium dodecyl sulfate-polyacryilamide gel electrophoresis (SDS-PAGE), where proteins are separated on the basis of their molecular weights (Ong and Pandey, 2001). One of the main advances that made more practical and reproducible the use of 2-DE was the advent of immobilized pH gradients stabilized on a plastic support strips by Bjellqvist (1982). Since then, numerous reports have

been published on the application of 2-DE for gene expression profiling, visualization of protein isoforms and post-translational modifications in numerous plant species (Natarajan et al., 2006; Imin et al., 2005; Incamps et al., 2005).

Analysis involving 2-DE includes: systematic separation, identification and quantitation of many proteins simultaneously from a single sample. This technique is unique in its ability to detect post and co-translational modifications, which cannot be predicted from the genome sequence (Berkelman and Stenstedt, 2004).

1.3.3 Multi-dimensional characterization of protein extracts

Coupling of 2-DE with other analytical techniques has been reported to overcome some of the drawbacks of SDS-PAGE (difficult to detect low abundant proteins, aggregation of proteins, specially hydrophobic proteins, reproducibility issues) by adding a preliminary analytical step. Some examples include a non-denaturing ion exchange chromatography prior to 2-DE (Butt et al., 2001) and a pre-fractionation/concentration step using affinity partitioning in ATPS prior to 2-DE–LC/MS (Schindler et al., 2006).

As previously mentioned, 2D gel technology offers protein visualization for a variety of purposes, in addition to allowing empirical estimation of molecular weight, pI and expression levels (Chen and Harmon, 2006). The use of this information coupled with aqueous two phase partitioning was recently reported by Gu and Glatz (2007) with the aim of obtaining a protein profile from corn endosperm. A three-dimensional pattern emerged from plotting MW, pI and a hydrophobicity scale obtained from ATP partitioned corn proteins. Valuable

information on the properties of potential contaminant proteins was obtained with this experimental approach on the way of facilitating further downstream purification. In this context, proteomics can provide tools to define properties of the recombinant protein that can be exploited to facilitate extraction and/or purification.

1.4 DISSERTATION OBJECTIVES

The general dissertation objective was established in response to the previously stated claim that a better knowledge of the potential contaminant proteins could facilitate downstream processing. Considering previous knowledge on aqueous two-phase partitioning, and recently acquired expertise on proteomic tools like 2D-electrophoresis, this research is focused on the integration of these two bioengineering approaches for molecular characterization and potential recovery of proteins expressed in plants. As a result of this research, knowledge will be generated to facilitate the predictive design of efficient downstream strategies for the recovery of recombinant proteins from plants.

In order to address this general objective, three particular objectives were established:

1. To evaluate the use of aqueous two-phase systems as a first step to establish a practical strategy for selective extraction of recombinant proteins from soybean.
2. To analyze the coupled application of aqueous two-phase partitioning and two-dimensional electrophoresis as a three-dimensional characterization method for proteins from soybean seed extracts.

3. To extend and evaluate the application of a three-dimensional technique for characterization of the potential contaminant proteins from alfalfa green-tissue protein extracts containing an artificially added recombinant protein.

1.5 EXPERIMENTAL MODELS

Three different experimental models were used in this research to address the general and particular objectives established for this dissertation.

- *Fractionated soybean extracts*

Obtained according to the experimental protocol reported by Thanh and Shibasaki (1976), isoelectric precipitation resulted in two protein fractions with different protein content (7S and 11S) that were used to prepare artificial mixtures containing recombinant β-glucuronidase as a model protein. These protein extracts were used to address the first particular objective.

- *Total soybean extract*

Proteins from defatted soybean flour were extracted according to the protocol reported by Gu and Glatz (2007) and produced a total soybean protein extract employed to address the second objective of this research.

- *Alfalfa green proteins containing a human cytokine (rGM-CSF)*

Three different extraction protocols were simultaneously evaluated for extraction of proteins from alfalfa green-tissue. The selected protein extract was used to prepare an artificial

mixture of alfalfa protein containing a recombinant model protein: human granulocyte-macrophage colony stimulating factor (rGM-CSF) used to address the third objective of this dissertation.

1.6 STRUCTURE OF THE DISSERTATION

All the experimental methods employed to reach the objectives of this dissertation are fully described in Chapter Two, including a brief background information of the main techniques used. The experimental results are presented from Chapter Three to Chapter Five. These chapters are presented in scientific manuscript format, following the instructions for authors reported in *Journal of Chemical Technology and Biotechnology* for Chapter Three: "**Processing of soybean extracts (*Glycine max*) in aqueous two-phase systems as a first step for the potential recovery of recombinant proteins**"; and *Biotechnology and Bioengineering* for Chapters Four and Five: "**Coupled application of aqueous two-phase systems and 2D-electrophoresis for three-dimensional characterization of soybean proteins**" and "**Application of a 3-D characterization technique to a green-tissue protein extract from alfalfa (*Medicago sativa*)**" respectively. Finally, the conclusions of this dissertation and suggestions for further work are presented in Chapter Six.

CHAPTER TWO – MATERIALS AND METHODS

2.1 MATERIALS

Poly(ethyleneglycol) MW 3350 (PEG), β-glucuronidase (GUS; E.C. 3.2.1.31), β-mercaptoethanol, DL-dithiothreitol (DTT), Tris(hydroxymethyl)aminomethane (Tris), Bradford reagent, and the selected model proteins: chymotrypsin A (CHY), cytochrome C (CYC), ribonuclease A (RNA), α-lactalbumin (LAC), bovine serum albumin (BSA) and lysozyme (LYS) were purchased from Sigma-Aldrich Chemicals (St Louis, MO). The ReadyPrep® rehydration buffer, 11 cm ReadyStrip® IPG strips (pH 3-10), iodoacetamide, Precision Plus® protein standards and Criterion® precast polyacrylamide gels (4-20%) were purchased from Bio-Rad (Hercules, CA). Coomassie Plus protein assay kit from Pierce Biotechnology, (Rockford, IL). Phenylmethylsulfonyl fluoride (PMSF) was purchased from Boehringer Mannheim Ltd (Germany). Commercial rhGM-CSF (Biofilgran®) was obtained from BioSidus S.A. (Buenos Aires, Argentina). All the other chemicals were purchased from Sigma-Aldrich Chemicals (St Louis, MO).

Plant materials.

Commercial soybean seeds (*Glycine max*) and alfalfa seeds (*Medicago sativa*) were kindly provided by the Agricultural Experimental Station of Tecnológico de Monterrey. Soybean seeds were briefly rinsed with tap water to remove soil and other materials and the whole grains were then ground to a fine powder using a household grinder in 20 s intervals to avoid heating. Ground powder was passed through mesh US No. 100 to obtain flour. Alfalfa seeds (*Medicago sativa*) were field grown during 4 weeks at 25°C. Aerial parts (stems and leaves) were harvested before flowering and immediately ground in liquid nitrogen with mortar and

pestle, adding powdered glass to improve cell wall breaking. Powder stocks were stored at -86°C for further use. Defatted soybean flour was kindly provided by Mark Reuber from the Center for Crops Utilization Research at Iowa State University.

2.2 PROTEIN EXTRACTION FROM PLANTS

A variety of different plant proteins have been analized for two-dimensional electrophoresis (Imin, et al., 2005; Incamps et al., 2005; Gu and Glatz, 2007; Natarajan et al., 2006) and despite the numerous advances on this technology, a bottleneck in proteome analysis remains at the procedure of protein extraction (Pirondini, et al., 2006). The main purpose of this stage is to release proteins from the plant materials into an aqueous environment for further purification operations (Menkhaus et al., 2004). Most of the proteomic analysis typically involves “one step” chemical extraction procedures with a solvent capable of disrupting protein aggregates and solubilizing significant amounts of proteins (Rose et al., 2004).

In the case of seed crops, dry particulate material is the preferred general source for extraction of proteins, faster extractions and higher protein yields are obtained using smaller particle sizes with aqueous extraction buffers (Bai and Nikolov, 2001). Wet grinding or tissue homogenization is another common extraction method for proteins especially from green tissues and it has been used for alfalfa and tobacco (Incamps et al., 2005; Platis and Labrou, 2006).

2.2.1 Extraction of proteins from soybean

Extraction and fractionation of soybean proteins

The seeds were briefly rinsed with tap water to remove soil and other materials and the whole grains were grounded to a fine powder using a household grinder in 20 s intervals to avoid over-heating. Ground powder was passed through mesh US No. 100 to obtain flour. Protein extracts were obtained according to the procedure reported by Thanh and Shibasaki (1976). Briefly, soybean flour was suspended in 0.03M Tris-HCl buffer containing 0.01M β-mercaptoethanol, pH 8 at a proportion of 1.0 gr solids/20 mL buffer. The slurry was stirred for 1 hr and centrifuged at 10000 x *g*, 20 min at room temperature to discard solids. The extract was carefully adjusted drop-wise to pH 6.4 with 2M HCl and centrifuged at 10000 x *g* for 20 min at 2-5°C, the precipitated collected was referred as 11S fraction or 11S fractionated soybean extract. The supernatant was further processed by isoelectric precipitation at pH 4.8 and the resulting precipitate was referred as 7S fraction (or 7S fractionated soybean extract). Each precipitate was re-suspended using the same initial volume 0.03M Tris-HCl buffer at pH to 7.5. Any trace of precipitate, if present, was removed by centrifugation (5 min at 10000 x *g*). These two fractionated extracts were used for partitioning experiments.

Extraction of total proteins from soybean

Soybean protein extract was obtained fresh according to the protocol reported by Gu and Glatz (2007). Briefly, defatted soybean flour, was suspended in 20 mM phosphate buffer, pH 7 at a proportion of 1.0 gr solids/10 mL buffer. The slurry was stirred for 1 hr with constant pH monitoring, then centrifuged (3000 *x g*, 30 min, room temperature; Centrifuge 5424,

Eppendorf, Hamburg, Germany) and decanted to eliminate waste solids. The supernatant was filtered using 0.22 mm syringe filter (mStar CA filter, Costar Corp., Corning, NY).

2.2.2 Extraction of proteins from alfalfa.

Commercial alfalfa seeds (*Medicago sativa*) kindly provided by the Agricultural Experimental Station of Tecnológico de Monterrey were briefly rinsed with tap water to remove antifungal protective coating and other solid materials. Seeds were field grown during 4 weeks at 25°C. Aerial parts (stems and leaves) were harvested before flowering and immediately ground with liquid nitrogen in 1.0 g samples using mortar and pestle. Powdered glass was added to improve cell wall breaking. Ground alfalfa stocks were stored at -86°C for further use. Three different protocols were used for protein extraction from alfalfa green biomass.

Sodium phosphate buffer

Alfalfa green biomass was suspended in 20 mM sodium phosphate buffer pH 7 containing 10mM EDTA at a proportion of 1.0 gr green biomass/10 mL buffer. The slurry was stirred for 1 hr at 25°C with constant pH monitoring, then centrifuged (10000 *x g*, 30 min, room temperature; Centrifuge 5804R, Eppendorf, Hamburg, Germany) and decanted to eliminate powdered glass and cell debris (Platis et al, 2006). The supernatant was filtered using 0.45 μm syringe filter (Corning Inc, U.S.A.) and used immediately.

$MgSO_4$-based buffer

Piridioni et al. (2006) reported the use of $MgSO_4$-based extraction for preparation of *Arabidopsis* protein crude extracts. In this research it was evaluated for alfalfa green tissue.

Buffer formulation included 50 mM Tris.HCl, 10 mM $MgSO_4$, 0.1% β-mercaptoethanol and 2mM phenylmethylsulfonyl fluoride adjusted to pH 7.8 at a proportion of 1.0 gr green biomass/10 mL buffer The slurry was stirred for 1 hr and centrifuged at 16,000 *x g* for 30 minutes at room temperature and then filtered using 0.45 μm syringe filter.

Tris-Borate-EDTA (TBE buffer)

The third buffer used for protein extraction was TBE buffer, composed of 0.45 M Tris HCl, 0.45 M H_3BO_3, 10 mM EDTA and adjusted to pH 8. It was used at a proportion of 1.0 g solids/10 mL buffer. The slurry was stirred for 1 hr, then centrifuged at 12000 *x g* for 10 minutes at room temperature and filtered using 0.45 μm syringe filter (Srere, 1969). After protein extraction an artificial mixture of alfalfa protein containing human rGM-CSF was prepared by adding an aliquot of the purified human cytokine to the selected alfalfa protein extract to have a final concentration of 63 μg of rGM-CSF /mL extract. This concentration of protein was found to be within the range of previously reported levels of this cytokine in plant cells (James et al., 2000).

2.3 AQUEOUS TWO-PHASE PARTITIONING

It is undoubtedly true that the proteome of plants is extremely complex, and the analysis of such complex protein mixtures requires a pre-fractionation or pre-sorting stage. A large number of proteins with a wide range of molecular weights and concentrations make the proteome of a living cell a dynamic system (Roy et al., 2007).

Aqueous two-phase partitioning emerged as a practical technique that allows recovery and purification of biological products, however this technique is also useful in understanding the chemical properties and behavior of proteins in solution (Rito-Palomares, 2004; Hachem et al., 1996). Partitioning in ATPS depends on unique physicochemical properties, such as size, surface charge, hydrophobicity, etc (Andrews et al., 2005). Consequently, it has been recently proposed as a valuable tool for protein characterization and pre-sorting stage in plant proteomics (Gu and Glatz, 2007b; Roy et al., 2007).

2.3.1 Partitioning of soybean protein extracts

Fractionated soybean extracts

Aqueous two-phase systems were selected based upon previous reports (Rito-Palomares and Lyddiatt, 2000; Rito-Palomares et al., 2000; Aguilar et al., 2006) to give a volume ratio of 1.0 and a fixed weight of 2.0 grams. The strategy behind the selection of the experimental systems is described elsewhere (Rito-Palomares, 2004). The system tie-line length (TLL), which represents the length of the line that connects the compositions of the top and bottom phases in a phase diagram for a defined system, was calculated according to the phase diagrams reported by Zaslavsky (1995). Four different MW of poly(ethyleneglycol) (PEG; Sigma Chemicals, MO, USA) were tested along with four different TLL values.

Predetermined quantities of stock solutions of potassium phosphate and PEG of nominal molecular weights 600, 1000, 1450 and 3350 g/mol were mixed with 0.2 grams of fractionated soybeans extracts (7S or 11S; with a protein concentration of 32 and 9.0 mg/mL, respectively) or 0.2 grams of purified β-glucuronidase solution with 30059 U/mL (GUS;

Sigma Chemicals) to give the desired PEG-salt composition as shown in Table 2.1. The total system weight was 2.0 grams (the amount of fractionated soybeans extracts or the purified protein solution added to the ATPS represented the 10% w/w of the total system). The amount of GUS in the stock solution (30059 U/mL) was defined to ensure enzyme activity detection in the ATPS.

Table 2.1 Systems selected for evaluation of the partition behaviour of proteins from fractionated soybean extracts.

System	Molecular weight of PEG (g/mol)	% PEG (w/w)	% Phosphate (w/w)	TLL (% w/w)
1	600	14.5	17.5	32.0
2		15.5	18.0	37.1
3		15.8	19.5	41.5
4		17.0	20.5	45.2
5	1000	15.6	18.0	47.2
6		17.6	18.0	49.9
7		19.8	18.5	53.6
8		22.2	23.0	67.7
9	1450	13.7	13.1	27.1
10		15.7	13.9	34.4
11		18.6	15.2	41.9
12		21.0	16.0	47.8
13	3350	16.9	14.5	42.3
14		18.7	15.0	46.2
15		21.0	15.7	51.3
16		22.1	17.0	56.2

Additional 2.0 grams spiking experiments were conducted, where the target protein (GUS) was artifially added to both fractionated extracts to simulate a protein extract from a transgenic plant. ATPS were formulated in which the stock solutions of PEG and salt were

mixed with 0.2 grams of spiked fractionated extracts (7S or 11S) containing a total protein concentration of 33 or 10 mg/mL, respectively and 60938 U/mL of β-glucuronidase. The stock solutions were gently mixed for 15 min. Adjustment of the pH was made by adding 1.0 M orthophosphoric acid or potassium hydroxide if needed. Complete phase separation was achieved by low-speed batch centrifugation at 1500 x *g* for 10 min.

Visual estimates of the volumes of top and bottom phases were made in graduated tubes. The volumes of the phases were then used to estimate the experimental volume ratio (Vr, defined as the ratio between the volume of the top phase and the bottom phase). Samples were carefully extracted from the phases (top and bottom phase) and analyzed. The top and bottom phase recovery was estimated as the amount of the target product present in the phase (volume of the phase x product concentration in the phase) and expressed relative to the original amount loaded into the system. Interface recovery was estimated as the necessary amount of the target product to complete the mass balance. Purification factor was estimated as the ratio of the relative purity of GUS after and before the extraction stage. Relative purity was defined as the amount of GUS relative to that of the proteins from the fractionated soybeans extracts. The results reported are the average of three independent experiments.

Total soybean extracts

Aqueous two-phase systems were formulated based upon previous reports on the partition of plant proteins (Gu and Glatz, 2007b) and the corresponding binodal curves reported by Zaslavsky (1995) to give a total weight of 5.0 g in the case of soybean protein extract and 2.0 g for the model proteins listed in Materials section (LYS, BSA, LAC, RNA CHY and CYC).

PEG 3350–potassium phosphate and PEG 3350–sodium sulfate systems were tested in order to compare protein partitioning among them. Predetermined quantities of stock solutions of PEG 3350, and sodium sulfate or potassium phosphate were mixed with solid NaCl and protein sample to give the following total composition: 15.7% (w/w) PEG 3350, 8.9% (w/w) Na_2SO_4, 3% (w/w) NaCl for sulfate-ATPS and 14.8% (w/w) PEG 3350, 10.3% (w/w) potassium phosphate, 3% (w/w) NaCl for phosphate-ATPS. Sample volume was such to give 1.0 mg protein/g ATPS for soybean extract and 0.35 mg each protein/g ATPS for model proteins previously listed.

Sodium chloride effect on protein partitioning was also evaluated using 0%, 1.5% and 3% (w/w). Adjustment of the pH to 7.0 was made by addition of 1 M orthophosphoric acid or potassium hydroxide if needed. After 1 hr of gentle mixing, complete phase separation was achieved by low-speed batch centrifugation at 1000 x *g* for 20 min. Visual estimates of the volumes of top and bottom phases were made in graduated tubes and used to estimate the experimental volume ratio (Vr). Top phase samples were carefully taken by pipetting, while bottom samples were extracted by piercing the bottom of the tube with a syringe. The interphase was left in the tube to avoid contamination. All partitioning experiments with soybean and model proteins were run by triplicate.

2.3.2 Partitioning of alfalfa protein extracts

The methodology used here was similar to that described in section 2.3.1 for total soybean extracts. However in this case, 2.0 g systems with a protein load of 1.0 mg/g ATPS were used for alfalfa extracts as well as for model proteins. Systems with different TLL values

were formulated based on binodal curves reported by Zaslavsky (1995). NaCl effect on protein partition was also evaluated varying from 0% to 9% w/w with the same protein load.

2.4 TRICHLOROACETIC ACID (TCA) PRECIPITATION

To eliminate interferences from phase-forming compounds and increase protein concentration, trichloroacetic acid (TCA) precipitation was performed to the top and bottom samples before isoelectric focusing. A volume of 0.8 mL of sample extracted from ATPS were thoroughly mixed with the same volume of ice cold 20% (w/v) TCA in a microcentrifuge tube and incubated on ice for 15 min before being centrifuged at 13,000 x *g* for 5 min at 4°C. The supernatant was decanted from the tube and the precipitate was washed once using ice-cold 10% (w/v) TCA, followed by three washes of ice-cold acetone, and centrifuged at 13,000 x *g* for 5 min at 4°C. All samples were incubated on ice for 5 minutes before centrifugation. In the case of model proteins mixture, sample was washed once with ice-cold 10% (w/v) TCA, and twice with ice-cold acetone. The final protein pellet was air-dried for 5 minutes, or until residual acetone was evaporated from the test tube (detected by the absence of solvent odor), being careful to not overdry the sample.

For soybean protein samples the protein pellet was completely re-dissolved using the lowest possible volume of 8 M urea according to the amount of protein present (0.12 mL for top phase and model proteins mixture samples and 0.24 mL and 0.48 mL for bottom phase). Protein concentration was measured at this step using Bradford assay with a 1:10 dilution of the sample to avoid interference of urea with Bradford reaction. An additional reference sample of non-partitioned soybean protein (1 mg/mL) was prepared by duplicate and

subjected to TCA treatment and the following steps in parallel with partitioned soybean samples. Protein samples in 8 M urea were diluted with rehydration buffer (8 M urea, 2% CHAPS, 50 mM DTT, 0.2% BioLyte®; BioRad) to a volume according to protein concentration to reach 200 μg of protein per strip (maximum load).

For alfalfa protein samples, in the cases where protein recovery from phases was low, three replicates were pooled at this step for top phase and two for bottom phase, and considered for final protein quantitation. The pellet was completely re-dissolved using 210 μL of rehydration buffer (8 M urea, 2% (w/v) CHAPS, 50 mM DTT, 0.2% (v/v) BioLyte®; Biorad) and protein concentration was measured at this step diluting samples 1:3 to avoid urea interfering with Bradford reagent (the highest concentration of urea allowed for Bradford reaction is 3 M). Appropriate volumes of these samples were used for strip re-hydration to reach a maximum of 200 μg of protein per strip.

2.5 TWO-DIMENSIONAL ELECTROPHORESIS

Two-dimensional electrophoresis (2-DE) has been for decades, the method of choice for parallel quantitative expression profiling of large sets of complex protein mixtures such as whole cell lysates (Chen and Harmon, 2006). 2-DE couples isoelectric focusing in the first dimension with SDS-PAGE in the second dimension, and allows separation and identification of complex mixtures of proteins according to an empirical estimation of their pI, MW, solubility and relative abundance (Görg et al., 2004). In addition, 2-DE not only provide information on the protein expression profiles, but also allows isolation of milligrams of protein for further analysis by mass spectrometry or sequencing. The major

steps of the 2-DE include: (i) sample preparation and protein solubilization; (ii) protein separation by 2-DE; (iii) protein detection and quantitation; (iv) computer assisted analysis of 2-DE patterns; (v) protein identification and characterization; (vi) 2-D protein database construction (Görg et al., 2004; Herbert et al., 2001).

2.5.1 Isoelectric focusing

The first-dimension isoelectric focusing (IEF) was performed using 11 cm pH 3-10 linear immobilized pH gradient strips (IPG ReadyStrip®, Bio-Rad) in an Ettan IPGphor apparatus (GE Healthcare). Strips were rehydrated according to manufacturer's recommendations with a volume of sample to have a maximum of 200 μg of protein per strip (11 cm) diluted in rehydration buffer (8 M urea, 2% (w/v) CHAPS, 50 mM DTT, 0.2% (v/v) BioLyte®; Biorad). A rehydration period of 16 h was used at room temperature using a reswelling tray (GE Healthcare). IEF run was optimized depending on the type of protein sample to be focused, the protocols for soybean and alfalfa are shown in Table 2.2

Table 2.2 Isoelectric focusing protocols used for protein samples.

Step	Mode	Voltage (V)	Time (h)
Soybean proteins			
1	Gradual	300	3
2	Gradual	1000	1
3	Gradual	8000	7
4	Hold	8000	12h 25'
Alfalfa proteins			
1	Hold	500	4
2	Gradual	1000	1
3	Gradual	8000	7
4	Hold	8000	2h 50'

According with this protocol, IEF was carried out for a total of 131,879 Vh for soybean protein samples and 50,250 Vh for alfalfa protein samples. The electric current was limited to 50 μA per strip.

2.5.2 Second-dimension electrophoresis

Soybean proteins

For the second dimension, the focused IPG strips were equilibrated with 6 M urea, 0.375 M Tris, 2% (w/v) SDS, 20% (v/v) glycerol, pH 8.8 and 2% (w/v) DTT for 15 minutes, and then acetylated using 2.5% (w/v) iodoacetamide instead of DTT for another 15 min. Strips were carefully placed onto 4-20% w/v gradient polyacrylamide gels (Criterion® precast gels, Bio-Rad) and covered with 0.5% agarose solution (BioRad Overlay Agarose) and the electrophoresis was performed at constant 200V during 65 minutes using a Criterion Dodeca Cell (Bio-Rad) unit. The gels were removed from the cassette and visualized by staining with 0.1% Coomassie Blue G-250 dissolved in 40% (v/v) methanol, 10% (v/v) acetic acid, 50% destilled water. After 1 h staining, gels were destained with the same methanol-acetic acid-water solution (without coomassie dye) for 3 h. Gels were scanned using a flatbed ImageScanner (GE Healthcare) at 600 dpi in transmissive mode and analyzed with PDQuest software from Bio-Rad. The mass of protein for individual spots was calculated from the spot densities (area multiplied by the pixel intensity) and related to the total amount of protein and total spot density for each gel. The non-partitioned sample from soybean extract was run in parallel with the samples from ATPS (top-bottom from sulfate and phosphate systems) and used as the reference to match spots from partitioned samples and calculate the mass balances from 2D gels. Protein concentrations from spot densities were used to calculate Kp

of individual proteins. Experiments with model proteins and soybean protein were run by duplicate.

Alfalfa proteins

For the second dimension, the focused IPG strips were equilibrated with 6 M urea, 75 mM Tris HCl, 2% (w/v) SDS, 29.3% (v/v) glycerol, 0.002% (w/v) bromophenol blue pH 8.8 and 2% (w/v) DTT for 15 minutes, and then acetylated using 2.5% (w/v) iodoacetamide instead of DTT for another 15 min. Strips were placed onto 12.5% (w/v) polyacrylamide gels prepared as described by Laemmli (1970) and covered with 0.5% agarose solution containin 0.002% bromophenol blue. The electrophoresis was performed at 200V during 4 h using a SE600 Ruby electrophoresis unit (GE Healthcare). The gels were removed from the cassette and stained with the same protocol described for soybean proteins. After destaining gels were scanned at 600-dpi resolution using a flat bed scanner in transmissive mode (Hewlett-Packard) and analyzed with PDQuest software (Bio-Rad). The mass of protein for individual spots was calculated from the spot densities (area multiplied by the pixel intensity) relative to the total amount of protein and total spot density for each gel. Protein concentrations from spot densities were used to calculate partition coefficient of individual proteins. All the experiments were run by duplicate.

2.6 ANALYTICAL TECHNIQUES

Protein measurements

Total protein determination for soybean and alfalfa protein extracts, phase samples and TCA precipitates was made using the microplate version of Bradford reaction. The procedure is based on the formation of a complex between the dye, brilliant blue G, and the proteins in

solution. The protein-dye complex causes a shift in the absorption maximum of the dye from 465 to 595 nm (Bradford, 1976; Sigma-Aldrich, Product Data Sheet). The Bradford reagent consists of brilliant blue G in a phosphoric acid and methanol solution. With the microplate version 5 μL of sample were placed into the well and 250 μL of Bradford reagent were added. After 15 minutes of incubation, microplate was read in a microplate reader (Bio-Tek Instruments, Vermont, U.S.A.) using BSA as standard (Bradford, 1976). Concentration of model proteins was measured spectrophotometrically at 280 nm (Pharmacia Biotech, Uppsala Sweden). All protein determinations included calibration curves using proper solvents and blank ATPS for correction of any interference from phase-forming components.

β-glucuronidase enzymatic assay

The enzymatic activity was estimated using the β-glucuronidase microassay adapted from Feurtado et al. (2005). Briefly, the modification included the use of 2-amino-2-methylpropanediol (AMPD) as enzyme inhibitor. The assay mixture consisted of 75 μL of McIlvine buffer pH 4.5, 15 μL of substrate (10 mM 4-nitrophenyl β-D-glucuronide (PNPG, Sigma-Aldrich)) in McIlvine buffer pH 5 and 60 μL of enzyme extract-sample. Microtitre-assay plate was incubated at 37°C for 15 minutes; reaction was stopped by adding 75 μL of 0.2 M AMPD (Sigma-Aldrich) and the yellow colour produced was measured at 405 nm in a microplate reader (Synergy HT, BioTek Instruments Inc., Vermont, U.S.A.). Blank wells involved adding aliquots from top or bottom phase of blank systems (ATPS without biological material) for zero calibration. One unit of enzyme activity was defined as the amount of enzyme producing 1.0 *p*mol of 4-nitrophenol from 4-nitrophenyl β-D-glucuronide per minute at pH 4.5 and 37°C.

Page left blank intentionally

CHAPTER THREE:

Processing of soybeans (*Glycine max.*) extracts in aqueous two-phase systems as a first step for the potential recovery of recombinant proteins

Oscar Aguilar and Marco Rito-Palomares*

Centro de Biotecnología, Departamento de Biotecnología e Ingeniería de Alimentos, Tecnológico de Monterrey. Campus Monterrey, Ave. Eugenio Garza Sada 2501 Sur, Monterrey, NL 64849, México.

*Author for correspondence: Marco Rito-Palomares, Ph: (+52) 81 8328-4132, Fax: (52) 81 8328-4136, *e-mail: mrito@itesm.mx*

Accepted for publication in *Journal of Chemical Technology and Biotechnology*.

Running title: Potential recovery of recombinant proteins from soybeans in ATPS

Abstract

BACKGROUND. The potential use of plants as production systems to establish bioprocesses has been established over the past decade. However, the lack of efficient initial concentration and separation procedures affect the generic acceptance of plants as economically viable systems. In this context the use of aqueous two-phase systems (ATPS) can provide strategies to facilitate the adoption of plants as a base for bioprocesses. Among the crops, soybeans (*Glycine max.*) represent an attractive alternative since potentially can produce high levels of recombinant protein. In this paper the processing of fractionated soybeans extracts using ATPS is evaluated as a first step to recover recombinant proteins expressed in plants, using β-glucuronidase (GUS; E.C. 3.2.1.31) as a model protein. **RESULTS**. The evaluation of the effect of systems parameters provided the conditions under which the contaminants proteins from fractionated soybeans extracts and GUS concentrated in opposite phases. PEG 600/phosphate system comprising 14.5% (w/w) polyethylene-glycol (PEG), 17.5% (w/w) phosphate, a volume ratio (Vr) equal to 1.0, a system pH of 7.0 resulted in the potential 83% recovery of GUS from the complex mixture and an increase in purity of 4.5 fold after ATPS. **CONCLUSIONS**. The findings reported here demonstrate the potential of ATPS to process fractionated soybeans extract as a first step to isolate and purify a recombinant protein expressed in soybeans. The proposed approach can simplify the way in which recombinant proteins expressed in plants can be recovered.

Key words: *Aqueous two-phase systems, Protein recovery, soybean extracts*

1. Introduction

There is a considerable interest in the development of biotechnological processes that exploit the use of plants as host for producing recombinant proteins. The potential use of plants as bioreactors has been established over the past decade. Transgenic plants are a potentially inexpensive system for the large-scale production of recombinant proteins for use in the pharmaceutical, agricultural and industrial sectors [1-3]. The advantages of using transgenic plants include low cost and flexibility in large-scale production, the presence of natural storage organs such as seeds and tubers, and existing technology to harvesting and processing of plant material [2]. However, the lack of efficient initial concentration and separation procedures affect the potential generic acceptance of plants as economically viable systems. One of the main drawbacks of the seed-based technology is the inherent need for the plant to generate the flower in order to produce the seed, rising bio-safety concerns to avoid the release of pollen to the environment and increasing the cost of production due to the need of containment and longer harvesting times [4]. The selection of the adequate crop for the production of recombinant proteins is affected by numerous aspects that contribute to the success and final cost of the products. The general alternative includes; maize (*Zea mays*), rice (*Oryza sativa*), canola (*Brassica sp*.), tobacco (*Nicotiana tabacum*), peas (*Pisum sativum*) and soybean (*Glycine max*). Canola, maize and soy have been regarded as the most likely production systems for commercial application of plants as bioreactors [5]. Maize is preferred over other plant hosts due to its lower hydro-soluble protein content. Extensive research on the potential recovery of recombinant proteins expressed in this particular crop has been documented [6, 7]. In the particular case of canola, compared to other crops, a greater fraction of its proteins are basic, thus recovery of anionic recombinant proteins could

be expected to be easier. Furthermore practical strategies for the recovery of recombinant proteins from canola have also been documented [8-10]. Soybeans (*Glycine max*) represent an attractive alternative since potentially can produce more recombinant protein that other crops. Soybeans contain approximately 40% of their weight as proteins compared to only 2% for potato tubers for example. Thus if the recombinant protein is expressed at an equal percentage of the total protein in these two crops, 20 times more product could be produced from soybean. However, the potential recovery of recombinant proteins from soybeans has not been widely studied.

The use of seeds for heterologous protein expression has proven to be an efficient system for both biopharmaceutical and industrial proteins [11, 12]. The use of seeds as host organ for recombinant protein production can be advantageous as they are natural storage organs with high expression levels, but this can be a challenge for downstream processing when the protein of interest has to be purified from the seed extract [8]. Soybean seeds contain more protein than any other commercial crop (up to 48%) consisting of a mixture of proteins (α-, β-, and γ-conglycinin, glycinin and other globulins) ranging in molecular weight from 140 to 300 kDa with different physicochemical properties allowing fractionation and differentiation [13-15]. The most popular and simple method for fractionation is the described by Thanh and Shibasaki [13] using simultaneous isoelectric precipitation to effectively isolate the two main soybean proteins 7S and 11S, named according to their Svedberg sedimentation coefficients. Two main globulins 7S and 11S, account for about 70-80% of the total soybean proteins (TSP). Of these two, the 7S fraction comprises about 35% of the TSP, where the majority of 7S proteins consist of the globulin called β-conglycinin, a trimeric glycoprotein with a

molecular weight of about 180 kDa that precipitates at pH 4.8 and makes up about 85% of the 7S fraction [13]. The 11S fraction comprises from 31 to 52% of the TSP, with about 85% being the 11S globulin called glycinin. The currently accepted model of glycinin is a hexamer with a molecular weight of about 360 kDa [14, 15]. Precipitation methods with variations are the most commonly used to recover soybeans proteins [16-18].

In general, protein recovery starting from plant material includes extraction, clarification, protein capture, purification and polishing. The overall production cost is mainly determined by the efficiency of the initial capture and purification steps where feed volumes are large until biomass solids and oils are removed and the protein is concentrated [5]. Therefore, the establishment of efficient primary recovery procedures for the recovery of recombinant proteins from transgenic crop is needed. In this context aqueous two-phase system (ATPS) represents an attractive alternative to facilitate the adoption of bioprocess based on plants as production systems. ATPS is a technique that has proved to have great potential for the recovery and purification of biological compounds [19-23].

Previous studies [9, 10] have shown the potential of traditional downstream processing operations to be applied to seed-produced recombinant proteins and the use of genetic engineering strategies to recover them from the bulk storage proteins where the product of interest is immersed. The potential economic benefits of substitution of costly unit operations, such as chromatography, by ATPS without a decrease in yield, have been previously addressed and the same strategies can be applicable for plant-made products [7, 24]. Recently, the potential applicability of ATPS for integrated extractive partitioning applied to the recovery of a model recombinant protein expressed in maize and tobacco has been

demonstrated [7, 25]. However, the potential application of ATPS to process soybeans extracts in route to the recovery of recombinant proteins needs to be addressed.

In the present study, the potential application of ATPS to process soybean extracts was evaluated as a first step to establish a practical strategy to recover recombinant proteins expressed in plants. Fractionated soybean protein extracts were obtained by simple scalable methods and used as an example of a complex mixture of contaminants where the recombinant proteins can be found. β-glucuronidase (GUS; E.C. 3.2.1.31) was chosen as a representative model protein to explore the potential use of ATPS to perform selective extraction of recombinant proteins from soybeans. GUS has been previously used as a model protein to be expressed in plant cells due to its high stability under a wide range of ionic forces and its absence in higher plants [10, 26]. The artificial mixture formed by adding GUS to fractionated soybeans extract served as an example to simulate the presence of a recombinant protein. GUS is a tetrameric acid hydrolase (68.2 kDa each) that catalyzes the cleavage of a wide variety of glycosides formed by β-D-glucosiduronic acid and residues with hydroxyl groups such as: steroids, phenols, antibiotics, and other metabolites; as a result it is easy to monitor by enzymatic assays. This enzyme maintains its hydrolytic capacity under very different environments, such as those found in ATPS [8, 10]. A practical approach to evaluate the effect of systems parameters such as polyethylene glycol (PEG) and phosphate concentration and nominal molecular weight of PEG upon the partition behaviour of proteins from the fractionated soybeans extracts and GUS was used. This approach was followed to establish the potential conditions under which the target protein (GUS) and the contaminant proteins from the fractionated soybean extract concentrated preferentially to

opposite phases. This practical strategy can be useful as a starting point to be applied in the recovery of recombinant proteins expressed in soybean.

2. Experimental

2.1 Protein extraction from soybeans.

Commercial soybean seeds (*Glycine max*) were kindly provided by the Agricultural Experimental Station of Tecnológico de Monterrey. The seeds were briefly rinsed with tap water to remove soil and other materials and the whole grains were then ground to a fine powder using a household grinder in 20 s intervals to avoid heating. Ground powder was passed through mesh US No. 100 to obtain flour. Soybeans rich protein extracts were obtained according to the procedure reported by Thanh and Shibasaki [13]. Briefly, soybean flour was suspended in 0.03M Tris-HCl buffer containing 0.01M β-mercaptoethanol, pH 8 at a proportion of 1.0 gr solids/20 cm^3 buffer. The slurry was stirred for 1 hr and centrifuged at 10,000 x *g*, 20 min at room temperature to discard solids. The extract is carefully adjusted drop-wise to pH 6.4 with 2M HCl and centrifuged at 10000 x *g* for 20 min at 2-5°C, the precipitate collected was referred in this study as 11S fraction or 11S fractionated soybean extract. The supernatant was further processed by isoelectric precipitation at pH 4.8 and the resulting precipitate was referred as 7S fraction (or 7S fractionated soybean extract). Both precipitates were re-suspended in 0.03M Tris-HCl buffer at pH to 7.5, any trace of precipitate, if present, was removed by centrifugation (5 min at 10000 x *g*). These two fractionated extracts were further used for partitioning experiments.

2.2 Aqueous two-phase experiments

Aqueous two-phase systems (see Table 1) were selected based upon previous experiences [19-21, 24] to give a volume ratio of 1.0 and a fixed weight of 2.0 grams. The strategy behind the selection of the experimental systems is well described elsewhere [19]. The system tie-line length (TLL), which represents the length of the line that connects the compositions of the top and bottom phases in a phase diagram for a defined system, was calculated as described before [24]. Predetermined quantities of stock solutions of potassium phosphate and poly(ethylene glycol) (PEG; Sigma Chemicals, St. Louis, MO, USA) of nominal molecular weights 600, 1000, 1450 and 3350 g mol^{-1} were mixed with 0.2 grams of fractionated soybeans extracts (7S or 11S; with a protein concentration of 32 and 9.0 mg cm^{-3}, respectively) or 0.2 grams of purified β-glucuronidase solution with 30059 U cm^{-3} (Sigma Chemicals) to give the desired PEG/salt composition (see Table 1) with a final weight of 2.0 grams (the amount of fractionated soybeans extracts or the purified protein solution added to the ATPS represent the 10% w/w of the total system). The amount of GUS in the stock solution (30059 U cm^{-3}) was defined to ensure enzyme activity detection in the ATPS. Additional, 2.0 grams spiking experiments were conducted, where the target protein (GUS) was added to both fractionated extracts to simulate a protein extract from a transgenic plant. ATPS were formulated in which the stock solutions of PEG and salt were mixed with 0.2 grams of spiked fractionated extracts (7S or 11S) containing a total protein concentration of 33 or 10 mg cm^{-3}, respectively and 60938 U cm^{-3} of β-glucuronidase. The stock solutions were gently mixed for 15 min. Adjustment of the pH was made by addition of 1 mol dm^{-3} orthophosphoric acid or potassium hydroxide if needed. Complete phase separation was achieved by low-speed batch centrifugation at 1500 x *g* for 10 min. Visual estimates of the

volumes of top and bottom phases were made in graduated tubes. The volumes of the phases were then used to estimate the experimental volume ratio (Vr, defined as the ratio between the volume of the top phase and the bottom phase). Samples were carefully extracted from the phases (top and bottom phase) and analyzed. The top and bottom phase recovery was estimated as the amount of the target product present in the phase (volume of the phase x product concentration in the phase) and expressed relative to the original amount loaded into the system. Interface recovery was estimated as the necessary amount of the target product to complete the mass balance. Purification factor was estimated as the ratio of the relative purity of GUS after and before the extraction stage. Relative purity was defined as the amount of GUS relative to that of the proteins from the fractionated soybeans extracts. Results reported are the average of three independent experiments.

2.3 Analytical techniques

The enzymatic activity was estimated using the β-glucuronidase microassay adapted from Feurtado *et al.* [27]. Briefly, the modification included the use of 2-amino-2-methylpropanediol (AMPD) as enzyme inhibitor. The assay mixture consisted of 75 µl of McIlvine buffer pH 4.5, 15 µl of substrate (10 mM 4-nitrophenyl β-D-glucuronide (PNPG, Sigma-Aldrich)) in McIlvine buffer pH 5 and 60 µl of enzyme extract-sample. Microtitre-assay plate was incubated at 37°C for 15 minutes; reaction was stopped by adding 75 µl of 0.2 M AMPD (Sigma-Aldrich) and the yellow color produced was measured at 405 nm in a microplate reader (BioTek Instruments Inc., Vermont, U.S.A.). Blank wells involved adding aliquots from top or bottom phase of blank systems (ATPS without biological material) for zero calibration. One unit of enzyme activity was defined as the amount of enzyme producing

1.0 *p*mol of 4-nitrophenol from 4-nitrophenyl β-D-glucuronide per minute at pH 4.5 and 37°C. Protein concentration from the phases was determined using the method of Bradford [28].

3. Results and Discussion

3.1 Processing of fractionated soybeans extracts in ATPS

In order to establish a purification strategy exploiting ATPS for the recovery of target products from a particular mixture, the behaviour of the major contaminants in the extraction stage must be characterised. In the particular case of the recovery of a recombinant protein expressed in soybean, the process conditions under which the product of interest and the contaminants concentrate in opposite phase must be established. For the design of a particular ATPS extraction stage, a strategy that involves the characterisation of the behaviour of the major contaminants proteins in selected ATPS can be adopted. In this study two fractionated extracts were obtained from soybean exploiting precipitation steps. The extraction of soybean proteins represents the first step to develop a primary protein recovery process. Thus the processing of the fractionated soybean extracts (identified as 7S and 11S fractions) in selected ATPS (Table 1) was initially attempted. In Table 2, the effect of system parameters (i.e. tie-line length and molecular weight of PEG) upon the potential recovery of proteins from fractionated soybeans extracts is illustrated. It is clear that proteins from both fractionated soybean extracts (7S and 11S) exhibited bottom and interface preference in the ATPS selected. In particular, proteins from the 7S fractionated soybean extract concentrated relatively equal in the interface and in the bottom phase. Increase in the TLL within the same molecular weight of PEG did not affect the protein behaviour (little or no change was

observed in partition coefficients of the same molecular weight of PEG); interface and bottom phase recovery remained relatively constant when TLL was increased. However, the increase in the molecular weight of PEG from 600 to 3350 g/mol caused a slight change in protein preference. Proteins from the 7S fractionated soybean extract increased preference for the interface particularly in PEG 3350 ATPS (systems 13 to 16 in Table 2). Such behaviour can be attributed to a potential migration of the proteins initially concentrated in the top phase at ATPS with low molecular weight of PEG to the interface. This change in protein phase preference can be explained by the reduction on the free available volume caused by the increase in the molecular weight of PEG in these types of ATPS [29, 30]. In the case of proteins from the 11S fractionated soybean extract, although these proteins also exhibited interface and bottom phase preference, an increase in the protein concentrated in the bottom phase was observed. As a result a decrease in the partition coefficients was observed when compared with those from the 7S fractionated soybean extract for the same molecular weight of PEG and TLL values (see Table 2). This increase in bottom phase preference of the proteins from the 11S fractionated soybean extract in comparison to that from the 7S can be attributed to the difference in the molecular size of the main proteins present in both fractionated extracts. The molecular weight of the majority of the proteins from the 11S (glycinin) and 7S (β-conglycinin) fractionated extracts is around 360 KDa and 180 KDa, respectively [13-16]. It has been reported [23, 29] that proteins with high molecular weight generally concentrate to the bottom phase. An increase in system TLL cause no effect on the partition behaviour of the proteins from the 11S fractionated soybean extract. Increasing molecular weight of PEG causes slight increase in the proteins concentrated to the bottom phase. ATPS with low molecular weight of PEG (i.e. 600 g/mol) and short TLL (i.e. 32% w/w; system 1 in Table 2) loaded with 11S fractionated soybean extract concentrated the

majority of the proteins in the bottom phase (i.e. 67%) and the rest in the top and interface (14% and 18%, respectively). It is important to note that system eight exhibited a strange behaviour, since the majority of the proteins, regardless of the fractionated soybean extract used, concentrated at the interface, 73% and 88% in ATPS loaded with 7S and 11S fractionated soybean extract, respectively. Such behaviour can be attributed to the fact that this system had the greatest TLL used (i.e. 67% w/w). A long tie-line length usually affects protein solubility and causes proteins present on the top and bottom phases to accumulate at the interface [23, 31].

The results obtained from the processing of fractionated soybeans extracts in ATPS may suggest several recovery strategies for the potential purification of a recombinant protein. It is clear that ATPS can be easily implemented for the recovery of recombinant proteins expressed in soybean that exhibited top phase preference (e.g. lysozyme and therapeutics proteins such as neuraminidase A and monoclonal antibodies against HIV [7, 25]). In this case it can be anticipated that the contaminant proteins (obtained from either fraction 7S or 11S) and the target products will concentrate in opposite phases. In the top phase the product of interest will be present with a reduce amount of contaminant proteins. In the case of target recombinant proteins that exhibited interface or bottom phase preference in ATPS, alternative strategies exploiting precipitation stages and process parameters need to be considered. Target protein can be recovered from the 7S or 11S fraction based on the particular partition behaviour of the protein of interest in ATPS. In order to further evaluate the potential application of ATPS for the recovery of recombinant proteins that can be expressed in soybean, β-glucuronidase (GUS) was selected as model protein. GUS was selected as a representative recombinant protein that can be expressed in plants (i.e. pea seeds, [26]) and

monitored in the presence of different solvents. As a first step, the partition behaviour of the model protein was evaluated in selected ATPS.

3.2 Partition of purified β-glucuronidase (GUS) in ATPS

The partition behaviour (expressed as product recovery) of glucuronidase in ATPS under increasing both TLL and molecular weight of PEG conditions is illustrated in Table 3. It is evident (based on mass balance analysis) that interface product accumulation accounted for the majority of the protein loaded to the systems. In ATPS of low molecular weight of PEG (i.e. 600 and 1000 g/mol; systems 1 to 8 in Table 3) GUS was not detected in the bottom phase. Enzyme activity was detected in the top phase and the rest of the product loaded to the ATPS was considered to be accumulated at the interface. The low product recovery from the top phase can be attributed to either product accumulation to the interface or loss of activity (or both). However, the presence of precipitate at the interface was visually observed in the ATPS. Increasing TLL caused no significant effect on product recovery. However, an increase in the molecular weight of PEG from 600 to 1000 g/mol caused GUS top phase recovery to slightly decrease. Further increase in molecular weight of PEG to 3350 g/mol resulted in an increase of GUS accumulation at the interface. To further evaluate the potential application of ATPS to recover GUS from fractionated soybean extracts, four systems (1 to 4 from Table 3) were selected. Systems in which GUS exhibited bottom phase preference (i.e. 9 and 10) were not considered due to the fact that additional contaminants such as debris tend to accumulate in this phase together with proteins from the fractionated soybean extracts as seen in Table 2.

3.3 Recovery of GUS from fractionated soybean extracts in ATPS

The potential recovery of GUS from fractionated soybean extract using ATPS was addressed using PEG 600-phosphate systems. In order to mimic the feedstock derived from the potential production of recombinant GUS expressed in soybeans, artificial mixtures containing purified GUS and fractionated extracts from soybeans were prepared and loaded into the ATPS. The original soybean extract was not used in this study since represents a very complex material that will challenge the efficiency and robustness of the extraction systems. Furthermore, the main reason for using fractionated extracts was to simulate a process where the target protein could be present in samples with different selective protein profiles. Isoelectric precipitation of spiked soybean extract (data not shown) yields fractions of the total extract where GUS is widely distributed from pH 3.5 to 6.5, being the fraction precipitated at 4.8 the one with the highest amount of GUS. The practical recovery and purification factor of GUS from ATPS loaded with 7S fractionated soybean extract is depicted in Table 4. A moderate top phase recovery for GUS was obtained (30 to 33% in Table 4), which was consistent with the results derived when ATPS were loaded with purified GUS (20 to 28%, see Table 3). However, the purity of GUS from the artificial mixture (in the presence of 7S fractionated soybean extract) increased up to 3 times after ATPS. The increase in purity can be explained by the poor top phase preference of the proteins from the 7S soybean extract (less than 11%; see Table 2). Potential recovery of GUS from the bottom phase was not evaluated since GUS did not exhibited preference for this particular phase and the contaminant proteins concentrated in the bottom phase and interface. Alternatively, the potential recovery of GUS from the interface can result in a product recovery of up to 67% (system 3 and 4 in Table 4). However, in this case the purity of GUS was not greatly benefited (purification factor of 1.3-1.5) due to the accumulation of the contaminant proteins at the interface. Furthermore, the potential recovery

of GUS from the interface may be questionable since the activity of the enzyme need to be evaluated. Additional solubilization stages may need to be considered. PEG 600/phosphate comprising 15.8% (w/w) PEG, 19.5% (w/w) phosphate, a volume ratio (Vr) equal to 1.0, a system pH of 7.0 and TLL of 41.5% loaded with GUS and 7S fractionated soybean extract achieved a recovery of GUS from the top phase of 33% and 67% from the interface with an increase in purity of 3.0 and 1.5 times, respectively. These results suggest the need for an alternative strategy. Furthermore, 7S fractionated soybeans extract is produced after two precipitation steps (see Experimental section) which necessarily increase the number of unit operations needed in the process. In contrast 11S fractionated soybean extract resulted from one single precipitation step at pH 6.4. Thus it was decided to evaluate the potential recovery of GUS from this type of fractionated extract. Table 5 illustrated the practical recovery and purification factor of GUS from ATPS loaded with 11S fractionated soybeans extract. In this particular case, the processing of the mixture of GUS and 11S fractionated soybean extract resulted in a moderate top phase recovery for GUS (i.e. 15% to 28%) with maximum purification factor of 3.0, similar to that obtained from the processing of 7S fractionated soybean extract added with GUS. However, the potential recovery of GUS from the interface was 82% (see system 1 in Table 5) and an increase in product purity was 4.5 times after ATPS. Such results can be explained by the fact that, for this particular ATPS, proteins from 11S fractionated soybean extract exhibited a stronger preference (compare to that of the proteins from the 7S extract) for the bottom phase and a moderate preference for the interface, 67% and 18%, respectively (Table 2). However, it is important to consider that the nature of the target product recovered from the interface need to be evaluated and additional re-solubilization stages may be required.

The general strategy proposed for the potential recovery of a GUS from fractionated soybean extracts is characterized by a two-stage process involving a precipitation stage prior to the ATPS extraction. PEG 600/phosphate system comprising 14.5% (w/w) PEG, 17.5% (w/w) phosphate, a volume ratio (Vr) equal to 1.0, a system pH of 7.0 and TLL of 32% w/w resulted in a potential strategy to recover 82% of the GUS protein from the interface. This strategy can result in a primary recovery process that increased the protein purity up to 4.5 times with respect to the purity of the initial crude extract loaded to the ATPS. Potential processing of the interface of ATPS for the recovery of biological products has been addressed before [23]. However, particular conditions to process the interface for the recovery of GUS from soybean proteins need to be established using conventional techniques such as precipitation or sequential ATPS extractions.

4. Conclusions

This paper reports a simplified strategy using aqueous two-phase systems to process fractionated soybean extracts as a first step for the potential recovery of recombinant proteins, particularly GUS. It has been shown that proteins from the 7S and 11S fractionated soybean extracts concentrated both at the bottom phase and interface. Changes in the phase preference were attributed to the nature of the proteins present in the fractionated extracts and the effect of systems parameters on the partition behaviour of the proteins. The selected model protein, GUS, accumulated predominantly at the interface and exhibited a top and bottom phase preference in ATPS with low and high molecular weight of PEG, respectively. Processing of the selected fractionated soybean extract with GUS by PEG 600-phosphate ATPS resulted in an increase in the purity of the target protein. The operating conditions established and derived from the proposed strategy resulted in the potential recovery of GUS from the interface. These conditions concentrated the contaminants proteins from the extracts preferentially to the top and bottom phase. Overall, the novel approach proposed here represents a practical strategy that can simplify the way in which recombinant proteins expressed in plants can be recovered. This research described an approach that is necessary as a starting point to establish a practical protocol to potentially isolate and purify a recombinant protein expressed in plants in general and in soybeans in particular.

Acknowledgements

The authors wish to acknowledge the financial support of Tecnológico de Monterrey, Biotechnology research chair (Grant CAT005). The technical assistance of Isabel Murillo is gratefully acknowledged.

References

1. Whitelam GC, Cockburn B, Gandecha AR and Owen MRL, Heterologous protein production in transgenic plants. *Biotechnol. Gen. Eng. Rev*. **11**:1-29 (1993).

2. Krebbers E, Bosch D and Vandekerckhove J, Prospects and progress in the production of foreign proteins and peptides in transgenic plants, in *Plant protein engineering*, ed. by Shewry PR and Gutteridges S. Cambridge University Press, London, pp. 315-325 (1992).

3. Austin S, Bingham ET, Koegel RG, Matthews DE, Shahan MN, Strab RJ and Burgess RR, An overview of a feasibility study for the production of industrial enzymes in transgenic alfalfa. *Ann. N. Y. Acad. Sci.* **721**:235-244 (1994).

4. Twyman RM, Stoger E, Schillberg S, Christou P and Fischer R, Molecular Farming in plants: host systems and expression technology. *Trends in Biotechnology*. **21** (12):570-578 (2003).

5. Menkhaus TJ, Bai Y, Zhang C, Nikolov ZL and Glatz CE, Considerations for the recovery of recombinant proteins from plants. *Biotechnol. Prog.* **20**:1001-1014 (2004).

6. Azzoni AR, Kusnadi AR, Miranda EA and Nikolov ZL, Recombinant aprotinin produced in transgenic corn seed: Extraction and purification studies. *Biotechnol. Bioeng.* **80** (3): 268-276 (2002).

7. Gu Z and Glatz CE, Aqueous two-phase extraction for protein recovery from corn extracts. *J Chromatogr B* **845** (1): 38-50 (2007).

8. Bai Y and Nikolov ZL, Effect of processing on the recovery of recombinant β-glucuronidase (rGUS) from transgenic canola. *Biotechnol Prog* **17**:168-174 (2001).

9. Zhang C and Glatz CE, Process engineering strategy for recombinant protein recovery from canola by cation exchange chromatography. *Biotechnol Prog* **15**:12-18 (1999).

10. Zhang C, Love RT, Jilka JM and Glatz CE, Genetic engineering strategies for purification of recombinant proteins from canola by anion exchange chromatography: An example of β-glucuronidase. *Biotechnol Prog* **17**:161-167 (2001).

11. Shyu DJH, Chou W-M, Yiu T-J, Lin CP and Tzen JTC, Cloning, functional expression, and characterization of cystatin in sesame seed. *J Agric Food Chem* **52**:1350-1356 (2004).

12. Zeitlin L, Olmsted SS, Moench TR, Co MS, Martinell BJ, Paradkar VM, Russell DR, Queen C, Cone RA and Whaley KJ, A humanized monoclonal antibody produced in transgenic plants for immunoprotection of the vagina against genital herpes. *Nature Biotechnol* **16**:1361-1364 (1998).

13. Thanh VH and Shibasaki K, Major proteins of soybean seeds. A straightforward fractionation and characterization. *J Agric Food Chem* **24**:1117-1121 (1976).

14. Friedman M and Brandon DL, Nutritional and Health Benefits of Soy Proteins. *J Agric Food Chem* **49**:1069-1086 (2001).

15. Natarajan SS, Xu C, Bae H, Caperna TJ and Garret WM, Characterization of storage proteins in wild (*Glycine soja*) and cultivated (*Glycine max*) soybean seeds using proteomic analysis. *J Agric Food Chem* **54**:3114-3120 (2006).

16. Golubovic M, van Hateren SH, Ottens M, Witkamp G-J and van der Wielen LAM, Novel method for the production of pure glycinin from soybeans. *J Agric Food Chem* **53**:5265-5269 (2005).

17. Lakemond CMM, De Jongh HHJ, Gruppen H and Voragen AGJ, Differences in denaturation of genetic variants of soy glycinin. *J Agric Food Chem* **50**:4275-4281 (2002).

18. Wu S, Murphy PA, Johnson LA, Fratzke MA and Reuber J, Pilot-plant fractionation of soybean glycinin and β-conglycinin. *J Am Oil Chem Soc* **76**:285-293 (1999).

19. Rito-Palomares M, Practical application of aqueous two-phase partition to process development for the recovery of biological products. *J. Chromatogr. B*. **807**: 3-11, (2004).

20. Rito-Palomares M, Dale C and Lyddiatt A, Generic application of an aqueous two-phase process for protein recovery from animal blood. *Process Biochem*. **35**: 665-673 (2000).

21. Rito-Palomares M and Lyddiatt A, Practical implementation of aqueous two-phase processes for protein recovery from yeast. *J. Chem. Technol. Biotechnol.* **75**: 632-638 (2000).

22. He C, Li S, Liu H, Li K and Liu F. Extraction of testosterone and epitestosterone in human urine using aqueous two-phase systems of ionic liquid and salt. *J. Chromatogr. A*. **1082**: 143-149 (2005).

23. Benavides J, Mena J, Cisneros M, Ramirez OT, Palomares LA and Rito-Palomares M, Rotavirus like-particles primary recovery from the insect cell in aqueous two-phase systems. *J. Chromatogr. B*. **842**, 48-57 (2006).

24. Aguilar O, Albiter V, Serrano-Carreon L and Rito-Palomares M, Direct comparison between ion-exchange chromatography and aqueous two-phase processes for the partial purification of penicillin acylase produced by *E. coli*. *J Chromatogr B* **835**:77-83 (2006).

25. Platis D and Labrou NE, Development of an aqueous two-phase partitioning system for fractionating therapeutic proteins from tobacco extract. *J Chromatogr A* **1128**:114-124 (2006).

26. Menkhaus TJ, Pate C, Krech A and Glatz CE, Recombinant protein purification from pea. *Biotechnol. Bioeng*. **86**:108-114 (2004).

27. Feurtado JA, Banik M and Bewley JD, The cloning and characterization of α-galactosidase present during and following germination of tomato (*Lycopersicon esculentum* Mill.) seed. *Journal of Experimental Botany* **52**:1239-1249 (2001).

28. Bradford MM, A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein–dye binding. *Anal Biochem* **72**:248–254 (1976).

29. Benavides J and Rito-Palomares M, Potential aqueous two-phase processes for the primary recovery of colored proteins from microbial origin. *Eng. Life Sci.* **5** (3), 259-266 (2005).

30. Rito-Palomares M and Hernandez M, Influence of system and process parameters on partitioning of cheese whey proteins in aqueous two-phase systems. *J.Chromatogr B*, **711**, 81-90, (1998).

31. Rito-Palomares M and Lyddiatt A, Process Integration using aqueous two-phase partition for the recovery of intracellular proteins. *Chem. Eng. J.* **87**, 313-319 (2002).

Tables

Table 1. Systems selected for evaluation of the partition behaviour of proteins from fractionated soybean extracts.

System	Molecular weight of PEG (g/mol)	% PEG (w/w)	% Phosphate (w/w)	TLL (% w/w)
1	600	14.5	17.5	32.0
2		15.5	18.0	37.1
3		15.8	19.5	41.5
4		17.0	20.5	45.2
5	1000	15.6	18.0	47.2
6		17.6	18.0	49.9
7		19.8	18.5	53.6
8		22.2	23.0	67.7
9	1450	13.7	13.1	27.1
10		15.7	13.9	34.4
11		18.6	15.2	41.9
12		21.0	16.0	47.8
13	3350	16.9	14.5	42.3
14		18.7	15.0	46.2
15		21.0	15.7	51.3
16		22.1	17.0	56.2

All 2.0 grams systems comprising PEG of four different molecular weights were selected to evaluate the impact of increasing TLL and molecular weight of PEG upon the partition behaviour of the proteins from the fractionated soybean extracts. The volume ratio (estimated from blank systems) and pH of the selected systems were kept constant and equal to 1.0 and 7.0 respectively. All systems were assembled and the TLL estimated as described in Experimental section.

Table 2. Effect of system tie-line lengths and molecular weight of PEG on the recovery of proteins from the fractionated soybeans extracts in ATPS

System	Proteins from 7S fractionated soybean extract				Proteins from 11S fractionated soybean extract			
	Top phase recovery (%)	Bottom phase recovery (%)	Interface Recovery (%)	Kp	Top phase recovery (%)	Bottom phase recovery (%)	Interface Recovery (%)	Kp
1	12±1	42±1	46±1	0.28 ±0.02	14±1	67±3	19±3	0.22 ±0.01
2	12±1	44±1	44±1	0.27 ±0.02	10±0.1	61±0.2	30±0.2	0.16 ±0.0
3	11±0.2	44±1	45±0.4	0.25 ±0.01	10±0.1	60±1	30±1	0.17 ±0.0
4	11±1	37±1	51±1	0.30 ±0.02	9±1	46±2	45±3	0.20 ±0.01
5	8±0.1	42±0.2	50±0.2	0.20 ±0.0	5±0.2	60±2	35±3	0.09 ±0.0
6	5±0.3	44±1	51±0.4	0.12 ±0.01	4±1	62±2	35±1	0.06 ±0.02
7	5±1	42±0.3	53±1	0.13 ±0.02	4±0.1	50±0.4	46±1	0.07 ±0.0
8	3±0.1	24±2	73±2	0.14 ±0.01	4±1	8±3	88±4	0.51 ±0.12
9	5±1	43±0.4	52±0.4	0.12 ±0.02	3±0.2	72±3	25±3	0.05 ±0.0
10	5±0.4	45±1	50±1	0.11 ±0.01	5±0.2	66±1	28±1	0.08 ±0.0
11	4±1	45±1	50±2	0.09 ±0.01	3±1	68±1	29±1	0.04 ±0.01
12	5±0.1	46±1	49±1	0.10 ±0.0	4±1	63±3	33±4	0.06 ±0.01
13	3±1	33±1	64±1	0.09 ±0.02	3±1	62±2	35±2	0.04 ±0.01
14	2±0.3	37±1	61±1	0.05 ±0.01	3±0.4	62±3	35±3	0.05 ±0.01
15	2±1	37±0.2	61±1	0.06 ±0.0	1±0.1	65±1	34±1	0.02 ±0.0
16	3±0.4	37±1	60±1	0.07 ±0.01	3±1	68±3	29±3	0.04 ±0.01

Compositions of the systems (1-16) are defined in Table 1. The concentration of the fractionated extracts in all the ATPS was 10% (w/w). The top and bottom phase protein recovery is expressed relative to the initial amount of proteins content in the soybeans extracts loaded into the systems. Interface recovery was estimated as the necessary amount of proteins to complete the mass balance. Protein partition coefficient (Kp) represents the ratio of the concentration of proteins in the phases.

Table 3. Effect of system tie-line lengths and molecular weight of PEG on the recovery of β–glucuronidase (GUS) in ATPS

System[a]	Top phase recovery (%)	Bottom phase recovery (%)	Interface recovery (%)
1	20±1	0.0	80±4
2	22±1	0.0	78±3
3	23±1	0.0	77±3
4	28±1	0.0	72±3
5	23±1	0.0	77±3
6	15±0.6	0.0	85±4
7	13±0.6	0.0	87±4
8	12±0.5	0.0	88±4
9	0.0	38±1.8	62±3
10	0.0	30±1.4	70±3
11	0.0	5.0±0.2	95±4
12	0.0	0.0	100.0
13	0.0	20±1	80±4
14	0.0	6±0.3	94±4
15	2±0.1	6±0.3	92±4
16	0.0	0.0	100.0

Compositions of the systems (1-16) are defined in Table 1. The concentration of the solution containing glucuronidase (30059 Uml^{-1}) in all the ATPS was 10% w/w. The top and bottom phase protein recovery is expressed relative to the initial amount of GUS loaded into the systems. Interface recovery was estimated as the necessary amount of protein to complete the mass balance.

Table 4. Recovery and purification factor of β–glucuronidase (GUS) from ATPS loaded with 7S fractionated soybeans extract

Systems	Top phase GUS recovery (%)	Top phase GUS purification factor	Bottom phase GUS recovery (%)	Bottom phase GUS purification factor	Interface GUS recovery (%)	Interface GUS purification factor
1	30±1.5	2.5	8±0.4	0.20	62±3	1.4
2	30±1.5	2.5	4±0.2	0.08	66±3	1.5
3	33±1.5	3.0	0.0	0.0	67±3	1.5
4	33±1.5	3.0	0.0	0.0	67±3	1.3

Compositions of the systems (1-4) are defined in Table 1. The concentration of the protein extract with glucuronidase (GUS) in all the ATPS was 10% w/w (12188 U per system). The top and bottom phase GUS recovery is expressed relative to the initial amount of the protein loaded into the systems. Interface recovery was estimated as the necessary amount of proteins to complete the mass balance. Purification factor is the ratio between the relative purity of GUS after and before the ATPS extraction. The relative purity of GUS in each phase is estimated as the amount of GUS relative to that of the proteins from the soybeans extract.

Table 5. Recovery and purification factor of β–glucuronidase (GUS) from ATPS loaded with 11S fractionated soybeans extract

Systems	Top phase GUS recovery (%)	Top phase GUS purification factor	Bottom phase GUS recovery (%)	Bottom phase GUS purification factor	Interface GUS recovery (%)	Interface GUS purification factor
1	15±0.7	1.0	3±0.1	0.04	82±4	4.5
2	20±1	2.1	4±0.2	0.06	76±3	2.6
3	23±1	2.2	0.0	0.0	77±3	2.6
4	28±1	3.0	0.0	0.0	72±3	1.6

Conditions as described in Table 4.

CHAPTER FOUR

Coupled application of aqueous two-phase systems and 2D-electrophoresis for three-dimensional characterization of soybean proteins.

Oscar Aguilar[1], Marco Rito-Palomares[1], Charles E. Glatz[2,*]

[1] Centro de Biotecnología, Departamento de Biotecnología e Ingeniería de Alimentos, Tecnológico de Monterrey. Campus Monterrey, Ave. Eugenio Garza Sada 2501 Sur, Monterrey, NL 64849, México.

[2] Departament of Chemical and Biological Engineering, 2114 Sweeney Hall, Iowa State University, Ames, Iowa 50011-2230, U.S.A.

* Author for correspondence: Charles E. Glatz, Ph.: +1 515 294 8472; Fax +1 515 294 2689. *e-mail address*: cglatz@iastate.edu

ABSTRACT

General knowledge of the molecular properties of the potential contaminant proteins benefit the selection and design of suitable strategies for the recovery of recombinant proteins. Alternatively, such knowledge can be exploited for the selection of an adequate host by considering the molecular properties of the target protein. A recently reported experimental approach that resulted from the combination of quantitative 2D electrophoresis (2-DE) with hydrophobic partitioning in aqueous two-phase systems (ATPS) was applied for the three-dimensional characterization of soybean proteins. The three dimensional scatter plots formed by molecular weight (MW), isoelectric point (pI) and surface hydrophobicity (log Kp) were obtained using two different ATPS compositions (PEG 3350 (15.7%)-sodium sulfate (8.9%)-NaCl (3%) and PEG 3350 (14.8%)-potassium phosphates (10.3%)-NaCl (3%)) to obtain hydrophobicity data. Six model proteins (lysozyme, bovine serum albumin, cytochrome C, α-lactalbumin, ribonuclease A and chymotrypsin A) with known hydrophobicities were used to validate ATPS and provide an accurate hydrophobicity scale. The two ATPS evaluated produced different hydrophobicity values for the same protein extract, suggesting a high influence of the phase-forming salt on the partition behavior of soybean proteins. Molecular properties of soybean proteins were obtained (MW, pI and log Kp) simultaneously using two different ATPS combined with 2-DE and can be a valuable tool for predictive design of recovery steps for recombinant proteins from plants.

Keywords: *2D-electrophoresis; aqueous two-phase systems; soybean; protein characterization.*

INTRODUCTION.

During the last decades, plants have emerged as serious competitive hosts for large-scale production of recombinant proteins (Twyman et al., 2003). Despite the numerous economical and technical advantages of the use of crops as biorreactors, the main factor for the establishment of a production system is the ease of purification, which influences directly the selection of the expression system (Larrick et al., 2001). As much as 95% of the costs of the production of a recombinant protein by molecular farming come from extracting the protein from the plant and purifying it. Although the production yields are important, the development of efficient extraction and purification methods is of primary commercial interest (Twyman et al., 2003; Evangelista et al., 1998). Several attempts have been reported on the use of alternative extraction and purification methods for recombinant proteins (i.e. ATPS, expanded bed adsorption, etc) regardless of the expression system (Benavides et al., 2006; Menkhaus and Glatz, 2005), however, the main challenge is still the establishment of efficient recovery and purification steps.

Aqueous two-phase systems have emerged as a practical technique that allows recovery and purification of biological products. They are formed when two water-soluble polymers or a polymer and a salt are mixed in aqueous solutions at a given proportion beyond the critical concentration at which two inmiscible phases are formed (Albertsson, 1958). ATP partitioning has been used to define strategies for primary recovery and purification of proteins from plants (Rito-Palomares, 2004; Gu and Glatz, 2007), and to measure protein hydrophobicity (Hachem et al., 1996).

The establishment of adequate strategies demands the characterization of the contaminant proteins from plant extracts. It is clear that a better understanding of the molecular characteristics of the potential contaminants (MW distribution, hydrophobicity, pI, etc) will benefit the process of selection, optimization and design of the downstream strategies (Asenjo and Andrews, 2004). In this context, proteomics can provide tools to define properties of the recombinant protein that can be exploited to facilitate extraction and/or purification. 2-D electrophoresis (2-DE) enables the separaration of complex mixtures of proteins and provide valuable information on p*I*, *MW*, and relative abundance, and it can be currently used for quantitative protein profiling of large sets of complex protein mixtures, like soybean protein extracts (Görg et al., 2004; Natarajan et al., 2006).

The hydrophobic properties of a protein play a fundamental role defining its behavior in solution and how the protein relates to other biomolecules. It is, in fact the ruling principle in hydrophobic interaction chromatography (HIC), a separation technique being used in most industrial processes for protein purification as well as in laboratory scale (Salgado et al., 2005). Several methods have been widely used to measure hydrophobicity of proteins, namely HIC, RPC, $(NH_4)_2SO_4$ precipitation, etc; however the use of aqueous two-phase systems (ATPS) for measuring the functional hydrophobicity of proteins, previously reported by several authors (Andrews et al., 2005; Franco et al., 1996; Hachem et al., 1996) measures the surface hydrophobicity of a protein as the result of the interaction between the residues on the surface of the protein and the solvent.

Gu and Glatz (2007) first reported the integration of 2-DE and ATPS for the characterization

of proteins from plant origin, establishing the protocol for a three-dimensional analysis of corn proteins based on their molecular properties. This method of 3D mapping consisted on the use ATP partitioning for measuring hydrophobicity of the proteins, followed by 2D electrophoresis to estimate MW and pI. The three-dimensional properties of the proteins could allow a better visualization of the surrounding environment for a recombinant protein given its molecular properties (MW, pI, hydrophobicity), thus allowing a better design of a downstream strategy or even the selection of an adequate host for expression.

The aim of this research was to extend this three-dimensional approach previously reported for corn germ protein to a soybean protein extract, selected for being a prolific source of protein that has been previously reported as host for recombinant protein production (Martinell et al., 2002; Hinchee et al., 1988). Two parallel ATPS with different phase-forming salt were formulated in order to evaluate the hydrophobic partitioning of soybean proteins. The characterization of soybean proteins in 2-DE (MW and pI), together with ATPS composed by PEG 3350 (14.8%)-potassium phosphates (10.3%)-NaCl (3%) and the previously reported system used for corn proteins PEG 3350 (15.7%)-Na_2SO_4 (8.9%)-NaCl (3%) were used to simultaneously measure partition coefficient of soybean proteins and obtain three-dimensional scatter plots. Application of this 3D technique to soybean proteins, will serve as a basis for the extension of this technique to other potential hosts for recombinant proteins and represents the first step in the route to the predictive design of recovery strategies for proteins expressed in plants.

MATERIALS AND METHODS

Materials

Defatted soybean flour, kindly provided by Mark Reuber from the Center for Crops Utilization Research at Iowa State University. Poly(ethyleneglycol) MW 3350, DL-dithiothreitol (DTT) and the selected model proteins: chymotrypsin A (CHY), cytochrome C (CYC), ribonuclease A (RNA), α-lactalbumin (LAC), bovine serum albumin (BSA) and lysozyme (LYS) were purchased from Sigma-Aldrich Chemicals (St Louis, MO). The ReadyPrep® rehydration buffer, 11 cm ReadyStrip® IPG strips (pH 3-10), iodoacetamide, Precision Plus® protein standards and Criterion® precast polyacrylamide gels (4-20%) were purchased from Bio-Rad (Hercules, CA). Coomassie Plus protein assay kit from Pierce Biotechnology (Rockford, IL).

Protein extraction.

Soybean protein extract was obtained fresh according to the protocol reported by Gu and Glatz (2007). Briefly, defatted soybean flour, was suspended in 20 mM phosphate buffer, pH 7 at a proportion of 1.0 g solids/10 mL buffer. The slurry was stirred for 1 h with constant pH monitoring, and then centrifuged (3000 *x g*, 30 min, room temperature; Centrifuge 5424, Eppendorf, Hamburg, Germany) and decanted to eliminate waste solids. The supernatant was filtered using 0.22 mm syringe filter (mStar CA filter, Costar Corp., Corning, NY). Total protein determination for soybean extract, phase samples and TCA precipitates was made using microplate Bradford reaction with BSA as standard (EL340, Bio-Tek Instruments, Winooski, VT) (Bradford, 1976). Concentration of model proteins was measured spectrophotometrically at 280 nm (Pharmacia Biotech, Uppsala Sweden). All protein

determinations included calibration curves using proper solvents and blank ATPS for correction of any interference from phase-forming components.

Aqueous two-phase partitioning

Aqueous two-phase systems were formulated based upon previous reports on the partition of plant proteins (Gu and Glatz, 2007b) and the corresponding binodal curves reported by Zaslavsky (1995) to give a fixed weight of 5.0 g in the case of soybean protein extract and 2.0 g for model proteins previously listed. PEG 3350–potassium phosphate and PEG 3350–sodium sulfate systems were tested in order to compare protein partitioning among them. Predetermined quantities of stock solutions of PEG 3350, and sodium sulfate or potassium phosphate were mixed with solid NaCl and protein sample to give the following total composition: 15.7% PEG 3350, 8.9% Na_2SO_4, 3% NaCl for sulfate-ATPS and 14.8% PEG 3350, 10.3% potassium phosphate, 3% NaCl for phosphate-ATPS. Protein load was 1 mg/g ATPS for soybean extract and model proteins mixture and 0.35 mg/g ATPS for individual proteins. Sodium chloride effect on protein partition was also evaluated using 0%, 1.5% and 3%. All concentrations were calculated in a weight basis, adjustment of the pH to 7 was made by the addition of 1 M orthophosphoric acid or potassium hydroxide if needed. After 1 hr of gentle mixing, complete phase separation was achieved by low-speed batch centrifugation at 1000 x *g* for 20 min. Visual estimates of the volumes of top and bottom phases were made in graduated tubes and used to estimate the experimental volume ratio (Vr). Top phase samples were carefully taken by pipetting, while bottom samples were extracted by piercing the bottom of the tube with a syringe. The interphase was left in the tube to avoid contamination. All partitioning experiments with soybean and model proteins

were run by triplicate. For a fourth partitioning experiment, samples of model proteins and soybean extract in each ATPS phase were analyzed by 2-DE with replicate gels of each. A non-partitioned total extract sample (1 mg/mL) was run as a reference with replicate gels.

TCA precipitation

To eliminate interferences from phase-forming compounds and increase protein concentration, trichloroacetic acid (TCA) precipitation was performed to the top and bottom samples before isoelectric focusing. Equal volumes of sample and 20% (w/v) TCA were thoroughly mixed in a microcentrifuge tube and incubated on ice for 15 min before being centrifuged at 13,000 x *g* for 5 min. The supernatant was decanted from the tube and the precipitate was washed once using ice-cold 10% (w/v) TCA, followed by three washes of ice-cold acetone, centrifuging at 13,000 x *g* for 5 min. All samples were incubated on ice for 5 minutes before centrifugation. In the case of model proteins mixture, sample was washed once with ice-cold 10% (w/v) TCA, and twice with ice-cold acetone. All steps were preformed in a cold room at 4°C. The final protein pellet was air-dried for 30-45 minutes, or until residual acetone was evaporated being careful to not overdry the sample. The pellet was completely re-dissolved using the lowest possible volume of 8 M urea according to the amount of protein present; 0.12 mL for top phase and model proteins mixture samples and 0.24 mL and 0.48 mL for bottom phase and total protein samples respectively. Protein concentration was measured at this step.

Isoelectric Focusing

After TCA precipitation, protein samples in urea were diluted with rehydration buffer (8M urea, 2% CHAPS, 50mM DTT, 0.2% BioLyte) to a volume according to its protein concentration to reach 200 μg of protein per strip (maximum load). Immobilized pH gradient strips (11 cm, pH 3-10, ReadyStrip®, Bio-Rad) were rehydrated with 185 mL of this protein solution for 16 h at room temperature. Isoelectric focusing (IEF) was carried out in an Amersham Ettan IPGphor II® IEF cell for a total of 131,879 Vh.

Second dimension electrophoresis

For the second dimension, the focused IPG strips were equilibrated with 6 M urea, 0.375 M Tris, 2% SDS, 20% glycerol, pH 8.8 and 2% w/v DTT for 15 minutes, and then acetylated using 2.5% w/v iodoacetamide instead of DTT for another 15 min. Strips were placed onto 4-20% w/v gradient polyacrylamide gels (Criterion® precast gels, Bio-Rad) and the electrophoresis was performed using a Criterion Dodeca Cell (Bio-Rad) unit. The gels were visualized by staining with Coomassie Blue G-250, and scanned using a flatbed ImageScanner (GE Healthcare) at 600 dpi in transmissive mode and analyzed with PDQuest (Bio-Rad). The mass of protein for individual spots was calculated from the spot densities (area multiplied by the pixel intensity) relative to the total amount of protein and total spot density for each gel. The non-partitioned sample from soybean extract was run in parallel with the samples from ATPS (top-bottom from sulfate and phosphate systems) and used as the reference to match spots from partitioned samples and calculate the mass balances from 2D gels. Protein concentrations from spot densities were used to calculate Kp of individual proteins.

RESULTS AND DISCUSSION

Protein extraction

Native proteins from defatted soybean flour were extracted at pH 7 using the protocol described by Gu and Glatz (2006) yielding protein extracts with an average of 30.7 mg/mL as total protein, this high concentration (0.31 ±0.01 g protein/g solids) was consistent with the high protein content of soybean reported previously (up to 48% w/w) (Thanh and Shibasaki, 1976; Aguilar and Rito-Palomares, 2007).

Protein partition experiments

Since the surface of a protein is made of several types of amino acids, two properties of the protein surface, namely its charge and hydrophobicity, play an important role in protein partitioning when ATPS are used (Hachem et al., 1996). The composition of the PEG-sulfate system evaluated was selected based on previous reports by Gu and Glatz (2007) for the three-dimensional characterization of corn proteins. Additionally, in order to study the effect of a different phase-forming salt on the partition of soybean protein, the composition of the PEG-phosphate system was defined based on the experiences of Aguilar and Rito-Palomares (2007). The criteria to evaluate the choice of a particular ATPS composition are defined by: (1) a partition coefficient (Kp) closest to 1.0, in order to obtain the highest number of proteins in both phases; (2) the highest % of recovery, to keep most of the proteins soluble and not at the interface; and (3) keep an intrinsic hydrophobicity that allows accurate hydrophobicity measurements.

Table 1 shows the effect of NaCl on the hydrophobic partitioning of the soybean native proteins. In the case of sulfate system, there is a small increase in Kp of soybean proteins when the sodium chloride increases from 0% to 3%. This behavior may be explained by the increase in the hydrophobic difference between the two phases, increasing the hydrophobic interaction between the hydrophobic proteins and the polymer in the top phase (Hachem et al., 1996; Franco et al., 1995). In the case of phosphate system, NaCl addition is not effective for Kp modification, this may be due to the trivalent nature of the phosphate ion, which may dominate the ionic forces that define intrinsic hydrophobicity of the system, which drive partition of soybean proteins. The ionic force of salt-rich phase in PEG-phosphate ATPS is relatively high (Azevedo et al., 2007) explaining the minor effect of the addition of a neutral salt on protein partitioning. This insignificant effect of NaCl on the partition of soybean proteins for the selected ATPS may be explained in terms of the composition of the total extract used. Two main soybean proteins, named in terms of their Svedverg sedimentation coeffcient, 7S and 11S, together comprise about 70% of total soybean proteins, any effect of NaCl on the partition of these two main proteins dominate the total effect on Kp masking any possible effect on other proteins present in small amounts. In this case, small changes in Kp of soybean proteins (0.06 to 0.09) were only evident for the PEG 3350-sulfate system. Considering this small increase in Kp, and previous experiences reported for corn protein (Gu and Glatz, 2007a), the system with 3% NaCl content was selected for hydrophobic partitioning of soybean proteins.

Hydrophobic partitioning of soybean proteins

The use of two-phase partitioning to estimate hydrophobicity has been previously addressed by Hachem et al. (1996) introducing the 1/m* parameter to measure the functional hydrophobicity of a protein. The m* value, derived from protein precipitation curves, represents the salt concentration at which the protein starts to precipitate (given a initial fixed concentration) and it is a good parameter to measure solubility or hydrophilicity of the protein. Thus, the inverse of this value (*1/m**) can be used as an indicator of surface hydrophobicity in salt solutions. They found a good linear relationship between the hydrophobicity of several model proteins measured as a function of their solubility and the partition coefficient in systems PEG-salt with a high concentration of NaCl. This methodology was employed to test the capacity of ATPS formulated with phosphate and sulfate salt to provide an accurate hydrophobic measurement for model proteins correlating their Kp value with the corresponding 1/m* parameter.

The partitioning experiments of the six model proteins in PEG/sulfate and PEG/phosphate systems resulted in important differences in the log Kp values, as can be seen in Table 2, special attention should be paid in the case of proteins like lysozyme, where the difference between both systems is almost one unit. BSA also shows this behavior, with a differential of 0.85 units in log Kp. These two proteins are very well known for their top and bottom preference respectively, in hydrophobic ATPS (Gu and Glatz, 2007a,b). Considering that these two systems were formulated in order to have similar TLL values, this protein behavior suggests that the ionic force of the phase-forming salt employed may help to define the intrinsic hydrophobicity of the two-phase system. The effect of the type of salt employed on

the functional or surface hydrophobicity of each protein seems to be very different. Unlike the previous experiments with complex soybean protein, in this case the effect of ionic force of the phase-forming salt on the hydrophobicity of the system is more evident. For model proteins like LYS, small changes in hydrophobicity readily affect protein-polymer and protein-protein interactions (Su and Chiang, 2006). The NaCl content also affects this intrinsic hydrophobicity. As previously reported, the use of NaCl in aqueous two-phase systems increases the hydrophobic difference between the phases, thus increasing the interaction between proteins and top-phase polymer (Hachem et al., 1996).

The log Kp values measured in ATPS for these model proteins were correlated with the hydrophobicity calculated from the precipitation curves (m* values) reported by Hachem et al. (1996) and Gu and Glatz (2007b). The two different salts used to formulate the systems resulted in ATPS where the model proteins have different Kp values. Figure 1 (a,b) shows the hydrophobicity scale obtained for the six different model proteins, for both systems the consecutive order obtained remains unaltered, despite the different hydrophobicity values obtained (log Kp). The good correlation factor (R^2= 0.96) obtained between 1/m* and Log Kp measured in PEG/phosphate system demonstrates that this system provides a reliable method to measure the functional hydrophobicity of proteins. The system used by Gu and Glatz (2007b) for corn protein formulated with PEG/sulfate gives a relatively good linear correlation (R^2= 0.69), although it is not as high as the PEG/phosphate system, is good enough to provide the same hydrophobicity scale with the model proteins. In both cases the use of the aqueous two-phase systems with the composition described here, demonstrated to be a reliable technique to measure hydrophobicity. However, it was clear that the linear

relationship between log Kp and log (1/m) was highly dependent on the selection of the phase-forming salt to formulate the systems. Regarding the criteria defined previously for the selection of an adequate system for hydrophobic partitioning, the requirements of high recoveries and intrinsic hydrophobicity are fulfilled for both cases. However, higher Kp values would be desirable in order to increase the number of protein spots to be detected in top phase.

Application of 3D method to a model proteins mixture

In order to assess the feasibility to obtain a 3D map of a simple mixture of well-known proteins, the three-dimensional technique was first tested using three model proteins (BSA, LAC, and LYS). Figure 2 shows the result of the application of this technique to the artificial mixture of proteins fractionated in ATPS. It is clear the differences obtained for the particular case of LAC and BSA. These differences obtained from the two ATPS used may be attributed to the different intrinsic hydrophobicity of the systems that resulted in different protein partition behavior. Kp for each model protein was calculated by the ratio of the spot densities obtained from the corresponding gel (top, or bottom). Multiple spots for a single protein are a common problem presented when working with model proteins, they could be attributed to impurities in protein standards. Additionally, LYS tends to form aggregates, which were detected as vertical spots sharing the same pI values (Su and Chiang, 2006). Regardless the system used, LYS and its aggregates have a preference for the top phase, this higher hydrophobicity may be explained in terms of a higher surface hydrophobicity (Su and Chiang, 2006), since LAC with a similar MW (~14 kDa) presented a different partition behavior also depending on the system used. The different hydrophobicity values obtained

for LAC and BSA by the 3D method and during validation of the systems using individual proteins may be explained by the low recovery percentages obtained after TCA precipitation (Table 3) affecting the mass balances for individual proteins and thus the Kp estimation by spots densitometry.

Despite of the differences in Kp obtained between the two different ATPS used, the MW and pI of the spots detected were similar to that reported by Gasteiger *et al*. (2005) for these proteins. With both ATPS it was demonstrated the feasibility of this 3D technique for simultaneous determination of the molecular characteristics of a protein mixture, PEG-phosphate.

Molecular characterization of proteins from soybean

Removal of the phase-forming components before performing IEF represents one of the main drawbacks encountered from the combination of ATP partitioning with 2-DE. In the case of top phase, polymer from this phase was removed as it could potentially interact with proteins and interfere with protein mobility on polyacrylamide gels; additionally the high salt concentration in the bottom phase makes it impossible to use for IEF runs, because of the high electric current that it could cause. In this context, TCA precipitation was used as an alternative to isolate proteins from phase-forming components. However, the protein loss during the precipitation step is an inevitable consequence. In Table 3, a quantification of protein loss was evaluated for the different samples treated (i.e. soybean protein sample, three model protein mixture). It is clear that protein losses up to 68% where observed after TCA precipitation from top-phase samples, and 60% for bottom-phase samples. The low

recoveries obtained for top-phase samples demonstrate again the strong interaction between PEG and proteins. TCA precipitation step yielded in some cases up to 32% recovery of the originally measured protein from the previous ATPS step. These protein losses were further considered, when protein concentration was estimated from spot densities. In most of the cases, the mass balances performed after every ATPS step (Table 3), resulted in recovery percentages higher than 90%. These findings differ from those reported for soybean proteins partitioned in similar hydrophobic systems (Aguilar and Rito-Palomares, 2007). Such behavior or differences can be attributed to the increased TLL used in the previously reported study (42-56% w/w) compared to that used here (i.e. 27%). It has been reported that an increase in TLL results in biomass accumulation at the interphase (Rito-Palomares et al., 2001).

This characterization technique was first established using corn germ protein (Gu and Glatz, 2007b) in order to provide a better understanding of the potential contaminants that would be present if a recombinant protein is produced using corn as host. Therefore, the expansion of this concept to soybean (a system with a different protein profile) seemed to be the suitable next step to prove the potential generic application of this procedure. The two ATPS formulated were tested with the soybean protein extracts obtained at pH 7 in order to provide different conditions for hydrophobic partitioning of soybean proteins. As a result, Table 4 shows the list of proteins that were detected in both top and bottom phase gels for the two different ATPS. The full lists of protein spots detected in 2-DE gels are available as supplementary materials. From the total list of spots detected, only a few were found to be present in both phases and matched as the same protein. As a result of this low number of

spots, the three-dimensional properties reported in Table 4 correspond to all the spots found in both phases. In the case of the PEG 3350-sulfate system a total of 72 protein spots were detected from PEG-sulfate system, and quantified according to the spot density and relative to the original amount of protein loaded to IPG strip as described in Materials and Methods section. Only 15 spots were detected in both top and bottom gels and their 3D properties could be determined (Table 4). These spots comprise 44% of total protein detected in both phases (from 72 spots). In a similar way, a total of 68 spots were detected and quantified from PEG-phosphate system, where only 11 spots were found to be present in both top and bottom phases. These spots comprised only 25% of total protein detected in both phases (68 spots). It was evident the presence of dominant proteins from gel images (Figures 3 and 4) where the main spots detected mainly at bottom and total phase gels, probably corresponded to glycinin and β-conglycinin subunits. The reported pI range of these subunits (pI 4.7-5.4 for the acidic subunits and 8-8.5 for the basic subunits) corresponded to the main spots detected in gels. The overall partitioning result was influenced by the partition behavior of these storage proteins that showed a marked hydrophilicity leading to higher bottom phase preference. Increasing the loading of gels could improve the detection of low abundant proteins in top and bottom phase gels.

In comparison with the previously reported experiments with corn protein (Gu and Glatz, 2007b) soybean analysis yielded a lower number of protein spots, probably as a result of the loss of proteins at the TCA precipitation step, especially at the top phase where more than 50% of the proteins were lost during washing steps. It is also important to consider the different protein profile in soybean, as already mentioned, two main proteins glycinin and

conclycinin (11S and 7S respectively) account for up to 70% of total soluble protein in soybeans, thus these two proteins (or their subunits) proportionally represent the majority of the proteins loaded and detected in 2D gels. Protein accumulation at the interphase is not a factor of considerable losses in these cases as can be seen in Table 3, presumably the relatively low TLL value of these systems is accounting for this behavior.

The differences in protein profiles obtained can be explained by the use of two different ATPS that rendered two different subsets of proteins partitioned. This is more evident when the 3D properties of these proteins are plotted. Figure 3 shows the scatter plot for the proteins partitioned using PEG 3350-phosphate system. The fraction of proteins that remained at the interphase according to the mass balance was lower than in the case of PEG 3350-sulfate system (see Table 3), however, in spite of the low number of protein spots detected at the top phase (only 11) and the limited molecular weight range (13-31 kDa, see Figure 3a) more spots were obtained with positive log Kp values, indicating top phase preference for these proteins in PEG 3350-phosphate ATPS. In contrast, when PEG 3350-sulfate system was used for partitioning, a wider MW range of proteins was detected (15-70 kDa). However, most of them resulted in negative log Kp values, indicating a clear bottom preference. It is important to consider that this behavior is the result of a limited top-bottom matching, with many proteins probably lost during TCA precipitation step or by interface accumulation especially in the case of PEG-sulfate system. An alternative strategy for phase-forming components removal needs to be considered to raise the potential of the proposed approach for the full characterization of the host proteins.

For both ATP systems, recovery percentages for specific proteins were based on the total amount of the same protein matched in the extract gel. The extract sample was subjected to TCA precipitation, the percentage of protein loss during this step was evaluated and accounted to calculate the initial protein concentration in soybean. For some proteins, the recoveries were higher than 100% (see Table 4), such results can be explained in the light that probably some of the proteins may be affected in a different way during the TCA procedure. A common assumption when using this TCA procedure is that all of the proteins are equally precipitated (Gu and Glatz, 2007); however the different affinity of the proteins for the polymer and/or salt may affect their response to the precipitating agent employed. Another alternative explanation may involve overlapping of spots with very similar MW-pI properties. In all of the cases, the spots matching is the most frequent source of errors, due to the light variations presented between replicate gels or the low resolution of the running gels. These errors were partially corrected by manually selecting landmarks during the image processing step, however miss-matching, over-detection of big spot groups or non-detected faint spots are the main source of errors.

A final comparison between the Kp values obtained by a common chemical method of measuring protein concentration (Bradford reagent) and spot densitometry is presented in Table 5. Considering the PEG 3350-sulfate system, it can be seen that there is no significant difference between the Kp values obtained using either method, both values are coincident for both, model proteins and soybean extracts. In the case of PEG 3350-phosphate system the Kp value measured by Bradford reaction resulted to be higher than that measured by 2-DE method, this behavior of Kp may be influenced by the number of experimental steps that

cause losses of the sample in bottom phase, such as TCA precipitation (see Table 3) or re-disolution of sample into rehydration buffer. Even though there are small differences in Kp for some cases, 2-DE represents a reliable method for the overall measurement of partition coefficient.

The findings reported here evidenced the importance of a careful selection of the partitioning ATP system to estimate Kp; a previous screening step is recommended in order to formulate a system with a PEG-salt composition that maximizes top phase partitioning without compromising the hydrophobic resolution of the system. However, the objective to establish the general applicability of this technique can be compromised by the evident need to perform this screening step to customize the protein partitioning case by case. The most immediate application of the information generated from the integration of these two experimental approaches (ATPS and 2-DE) could be the design or selection of more suitable downstream purification steps for recombinant proteins from soybean.

CONCLUSIONS

The application of 3D characterization approach resulted in valuable information about the molecular nature of the potential contaminant proteins when a target protein is expressed in soybean. A more detailed and complete protein profile could be obtained if some of the drawbacks of the hydrophobic ATP partitioning can be overcome. The composition of the PEG-sulfate system (PEG 3350 15.7%, Na_2SO_4 8.9%, NaCl 3%) resulted in a higher number of spots detected from both phases in 2DE gels, compared with the PEG-phosphate system (PEG 3350 14.8%, potassium phosphate 10.3%, NaCl 3%). It was established the

applicability of this novel characterization technique to soybean extracts. However, it was also evident the need to perform a careful selection of the partitioning system to be used to fulfill the established criteria for an adequate ATP system for 3D characterization: a Kp the closest to one, the highest protein recoveries and proved hydrophobicity as the driving principle for partitioning. Given these criteria, PEG 3350-phosphate system, resulted to be better for 3D protein characterization. It was also demonstrated that the phase-forming salt has a great influence on the intrinsic hydrophobicity of the ATP system used for Kp determination when model proteins and soybean extracts were used.

The future optimization of the hydrophobic ATP system as well as the recovery steps needed to isolate proteins from phase-forming components could contribute to have a more complete image of proteins from crude extracts. This information could serve as the initial step for better selection of the downstream strategy or the even most adequate host, given the molecular properties of the target molecule and the host proteins.

ACKNOWLEDGEMENTS.

Authors wish to acknowledge the financial support of Tecnológico de Monterrey, Biotechnology Research Chair (Grant CAT005), and Chemical and Biological Engineering Department at Iowa State University.

REFERENCES.

Aguilar O, Rito-Palomares M. 2007. Processing of soybeans (*Glycine max.*) extracts in aqueous two-phase systems as a first step for the potential recovery of recombinant proteins. J Chem Technol and Biotechnol DOI:10.1002 (*In Press*).

Albertsson PA. 1958. Partition of proteins in liquid polymer-polymer two-phase systems. Nature 182:709-711.

Andrews BA, Schmidt AS, Asenjo JA. 2005. Correlation for the partition behavior of proteins in aqueous two-phase systems: Effect of surface hydrophobicity and charge. Biotechnol Bioeng 90:380-390.

Asenjo JA, Andrews BA. 2004. Is there a rational method to purify proteins? From expert systems to proteomics. J Mol Recognit 17:236-247.

Azevedo AM, Rosa PAJ, Ferreira IF, Aires-Barros MR. 2007. Optimisation of aqueous two-phase extraction of human antibodies. J Biotechnol 132:209-217.

Benavides J, Mena J, Cisneros M, Ramirez OT, Palomares LA and Rito-Palomares M. 2006. Rotavirus-like particles primary recovery from the insect cell in aqueous two-phase systems. J Chromatogt B 842:48-57.

Bradford MM, 1976. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248-254.

Evangelista RL, Kusnadi AR, Howard JA, Nikolov ZL. 1998. Process and economic evaluation of the extraction and purification of recombinant beta-glucuronidase from transgenic corn. Biotechnol Prog 14:607-614.

Franco TT, Andrews AT, Asenjo JA. 1996. Use of chemically modified proteins to study the effect of a single protein property on partitioning in aqueous two-phase systems: Effect of surface hydrophobicity. Biotech Bioeng 49:300-308.

Friedman M, Brandon DL. 2001. Nutritional and Health Benefits of Soy Proteins. J Agric Food Chem 49:1069-1086.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel R.D., Bairoch A. 2005. Protein Identification and Analysis Tools on the ExPASy Server. *In*: John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607. http://www.expasy.org/tools/protparam.html.

Görg A, Weiss W, Dunn MJ. 2004. Current two-dimensional electrophoresis technology for proteomics. Proteomics 4:3665-3685.

Gu Z, Glatz CE. 2007 (a). Aqueous two-phase extraction for protein recovery from corn extracts. *J Chromatogr B* 845:38–50.

Gu Z, Glatz CE. 2007 (b). A method for three-dimensional protein characterization and its application to a complex plant (corn) extract. Biotechnol Bioeng 97:1158-1169.

Hachem F, Andrews BA, Asenjo JA. 1996. Hydrophobic partitioning of proteins in aqueous two-phase systems. Enzyme Microb Tech 19:507-517.

Hinchee MA, Connor-Ward D. October 20, 1998. Method for soybean transformation and regeneration. U.S. Patent Number 5824877.

Larrick JW, Yu L, Naftzger C, Jaiswal S, Wycoff K. 2001 Production of secretory IgA antibodies in plants. Biomol Eng 18:87–94.

Martinell BJ, Julson LS, Emler CA, Yong H, McCabe DE, Williams EJ. May 7, 2002. Soybean *Agrobacterium* transformation method. U.S. Patent Application No. 6384301.

Menkhaus TJ, Glatz CE. 2005. Antibody Capture from Corn Endosperm Extracts by Packed Bed and Expanded Bed Adsorption. Biotechnol Prog 21:473-485.

Natarajan SS, Xu C, Hanhong B, Caperna TJ, Garret WM. 2006. Characterization of storage proteins in wild (*Glycine soja*) and cultivated (*Glycine max*) soybean seeds using proteomic analysis. J Agric Food Chem 54:3114-3120.

Rito-Palomares M, Negrete A, Miranda L, Flores C, Galindo E, Serrano-Carreón L. 2001. The potential application of aqueous two-phase systems for in situ recovery of 6-pentyl--pyrone produced by *Trichoderma harzianum*. Enzyme Microb Technol 28:625-631.

Rito-Palomares M. 2004. Practical application of aqueous two-phase partition to process development for the recovery of biological products. J Chromatogr B 807:3-11.

Salgado JC, Rapaport I, Asenjo JA. 2005. Is it possible to predict the average surface hydrophobicity of a protein using only its amino acid composition? J Chromatogr A 1075:133-143.

Su C-K, Chiang BH. 2006. Partitioning and purification of lysozyme from chicken egg white using aqueous two-phase system. Process Biochem 41:257-263

Thanh VH, Shibasaki K. 1976. Major proteins of soybean seeds. A straightforward fractionation and characterization. J Agric Food Chem 24:1117-1121.

Twyman RM, Stoger E, Schillberg S, Christou P, Fischer R. 2003. Molecular farming in plants: host systems and expression technology. Trends in Biotechnol 21:570-578.

Zaslavsky A. 1995. Aqueous two-phase partitioning. Physical Chemistry and Bioanalytical Applications. New York; Marcel Dekker Inc. 660 p.

Legend of Figures.

Figure 1. Correlation between surface hydrophobicity measured as log Kp in ATPS with 1/m*. Kp is the partitioning coefficient of protein in PEG-salt system at pH 7 and protein loading of 0.4 mg/g ATPS each, and m* is the concentration of salt at the discontinuity point of the protein in $(NH_4)_2SO_4$ precipitation curve, reported by Hachem *et al*. (1996) for selected model proteins using an initial concentration of 2 mg/mL. **a**: ATPS: PEG 3350 (14.8%), potassium phosphate (10.3%), NaCl (3%). **b**: ATPS: PEG 3350 (15.7%), Na_2SO_4 (8.9%), NaCl (3%).

Figure 2. 2D gels and 3D map of artificial mixture of three model proteins fractionated in ATPS. PEG-salt system at pH 7 and protein loading of 0.35 mg/g ATPS for each protein. **a**: ATPS: PEG 3350 (14.8%), potassium phosphate (10.3%), NaCl (3%). **b**: ATPS: PEG 3350 (15.7%), Na_2SO_4 (8.9%), NaCl (3%).

Figure 3. 2D gels and 3D scatter plot of proteins from soybean using PEG 3350-phosphate system. ATPS: PEG 3350 (14.8%), potassium phosphate (10.3%), NaCl (3%) at pH 7 and 1 mg protein/g ATPS. **a**: top phase gel, **b**: bottom phase gel, **c**: initial soybean extract gel, **d**: 3D scatter plot of soybean proteins detected in both phases (see data in Table 4).

Figure 4. 2D gels and 3D scatter plot of proteins from soybean using PEG 3350-Na_2SO_4 system. ATPS: PEG 3350 (15.7%), potassium phosphate (8.9%), NaCl (3%) at pH 7 and 1 mg protein/g ATPS. **a**: top phase gel, **b**: bottom phase gel, **c**: initial soybean extract gel, **d**: 3D scatter plot of soybean proteins detected in both phases (see data in Table 4).

Table 1. Effect of NaCl addition on the partition coefficient of soybean proteins in ATPS

	NaCl content (w/w)[a]		
	0%	**1.5%**	**3.0%**
TLL: 32% w/w PEG 3350 (15.7% w/w) Na_2SO_4 (8.9% w/w)	0.06± 0.01	0.06±0.003	0.09±0.01
TLL: 27% w/w PEG 3350 (14.8% w/w) Phosphate (10.3% w/w)	0.08±0.0001	0.08±0.001	0.08±0.003

[a]Systems formulated at pH 7 and 25°C. Sample load was 1.0 mg protein/g ATPS.

Table 2. Differential observed on the hydrophobic partitioning of selected model proteins in ATPS by changing the phase-forming salt.

	M.W.	**pI**	**|Δ(log Kp)|[a]**
α-Lactoalbumin	14.0	4.5	0.13
Bovine Serum Albumin	66.5	4.9	0.85
Chymotrypsin	25.0	8.8	0.15
Ribonuclease A	13.7	9.6	0.19
Cytochrome C	12.3	10.2	0.61
Lysozyme	14.3	10.3	0.92

[a]Absolute values calculated from: $|\Delta(\log Kp)| = |(\log Kp_{phosphate}) - (\log Kp_{sulfate})|$, where $Kp_{phosphate}$ is the partition coefficient calculated from PEG 3350/potassium phosphate system and $Kp_{sulphate}$ is the partition coefficient calculated from PEG 3350/Na_2SO_4; both systems at pH 7 and 25°C.

Table 3. Mass balances obtained after sequential ATP partitioning and TCA precipitation of different protein samples.

Type of sample partitioned:	PEG 3350-PO_4 system[a]		PEG 3350-Na_2SO_4 system[b]	
	% Protein recovered after:			
	ATPS[c]	TCA Precip.[d]	ATPS[c]	TCA Precip.[d]
Soybean Protein				
Top phase	4.9	46.5	5.4	32.0
Bottom phase	81.7	89.2	51.2	104.0
Model proteins mixture[e]				
Top phase	15.1	31.9	14.6	74.3
Bottom phase	84.6	41.3	78.8	39.9

[a]ATPS is PEG 3350 (14.8%)-potassium phosphates (10.3%)-NaCl (3%) at pH 7.
[b]ATPS is PEG 3350 (15.7%)-Na_2SO_4 (8.9%)-NaCl (3%) at pH 7.
[c]Based on the initial amount of protein added to the system (1.0 mg/g ATPS).
[d]Based on the protein concentration measured after ATPS.
[e]Mixture containing LAC-BSA-LYS as model proteins.

Table 4. Three-dimensional properties and mass balance of selected soybean proteins partitioned in ATPS.

ATPS[a]	MW (kDa)	pI	log Kp	Yield[b]
PEG 3350-SO_4 system	24.68	4.68	-0.761	37.3
15.7% PEG 3350	26.04	4.69	-0.688	58.9
8.9% Na_2SO_4	17.54	4.84	-1.017	31.5
3.0% NaCl	28.03	4.99	-1.998	52.8
	59.36	5.24	-2.746	39.5
TLL 32%	27.46	5.35	-2.268	65.5
Vr ~ 1.0	33.7	5.4	-2.874	50.6
	14.5	5.58	-1.722	26.0
	27.9	5.6	-2.687	68.1
	29.56	6.23	-1.995	15.7
	25.18	6.25	-1.234	192.8
	24.96	6.33	-0.699	28.9
	30.96	6.65	0.160	78.4
	16.56	6.68	-1.164	33.9
	31.3	6.8	N.D. bot	640.8
	30.2	6.91	N.D. bot.	255.5
	26.05	7.3	N.D. bot.	28.8
	69.76	7.54	-1.400	165.6
PEG 3350-PO_4 system	13.5	4.91	-0.361	58.2
14.8% PEG 3350	24.99	5.35	-2.409	79.7
10.3% Phosphate	21.75	5.48	N.D. bot.	374.0
3.0% NaCl	22.46	5.65	-1.469	91.6
	23.4	5.77	0.164	83.8
TLL 30%	24.12	6	0.353	44.5
Vr ~ 1.25	27.26	6.17	-0.386	54.6
	25.89	6.9	N.D. bot.	1.1
	29.69	7.14	1.295	139.3
	12.9	7.06	-0.144	5.3
	16.21	7.93	-2.168	320.2
	30.15	7.78	1.064	1901.2
	30.955	9.46	0.386	36.3

[a]All data are the average of duplicate experiments run at pH 7 and 25°C. Load of protein was 1.0 mg/g ATPS.
[b]Calculated using the sum of top + bottom divided by the total amount of the same protein spot matched in the gel where no partition experiment was performed (total extract).

Table 5. Comparison between the Kp obtained simultaneously by 2DE spot densitometry and a chemical assay using Bradford reagent.[a]

Method	Sample[b]	ATPS[c]	
		PEG 3350-potassium phosphate	PEG 3350-sodium sulfate
Kp-Bradford assay	Soybean Protein	0.08±0.003	0.09±0.01
	Model mixture[d]	0.13±0.003	0.13±0.001
Kp-2DE[e]	Soybean Protein	0.04±0.01	0.11±0.03
	Model mixture	0.11±0.001	0.13±0.004

[a]All data are the average of duplicate experiment in the case of 2DE and triplicate for protein assay.
[b]Protein load was 1 mg/g ATPS for soybean samples and model proteins mixture.
[c]ATPS: PEG 3350 (14.8%)-potassium phosphate (10.3%)-NaCl (3%) and PEG 3350 (15.7%)-Na_2SO_4 (8.9%)-NaCl (3%) at pH 7.
[d]Mixture: lysozyme, bovine serum albumin and α-lactoalbumin.
[e]Calculated as the summatory of proteins detected by spot densitometry considering average losses for each sample.

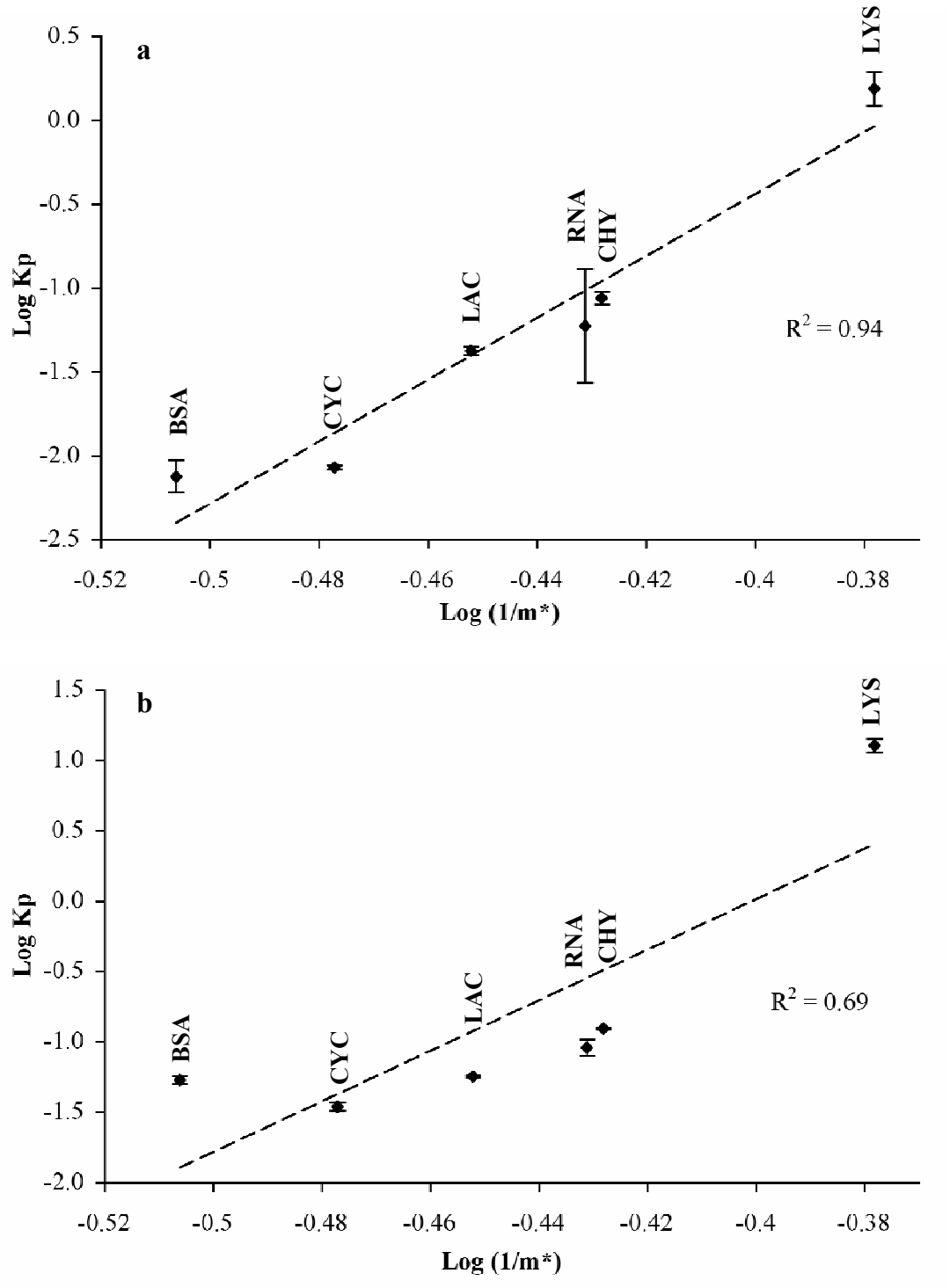
a
b
Log Kp
Log (1/m*)
BSA
CYC
LAC
RNA
CHY
LYS
R2 = 0.94
R2 = 0.69
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
1.5
1.0
-0.52
-0.5
-0.48
-0.46
-0.44
-0.42
-0.4
-0.38

Figure 1

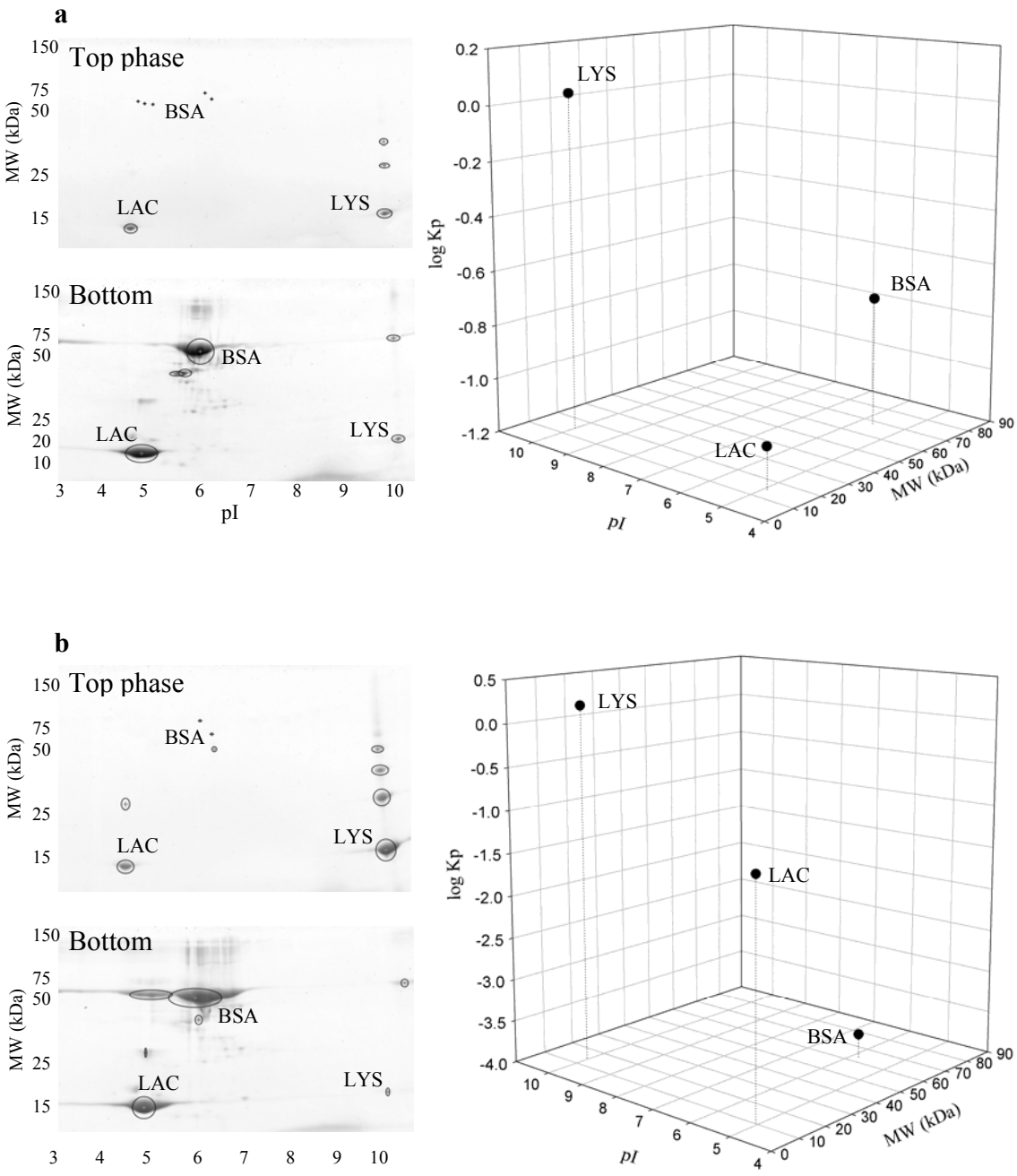
a
Top phase
BSA
LAC
LYS
MW (kDa)
Bottom
BSA
LAC
LYS
pI
log Kp
LYS
BSA
LAC
pI
MW (kDa)
b
Top phase
BSA
LAC
LYS
MW (kDa)
Bottom
BSA
LAC
LYS
pI
log Kp
LYS
LAC
BSA
pI
MW (kDa)

Figure 2

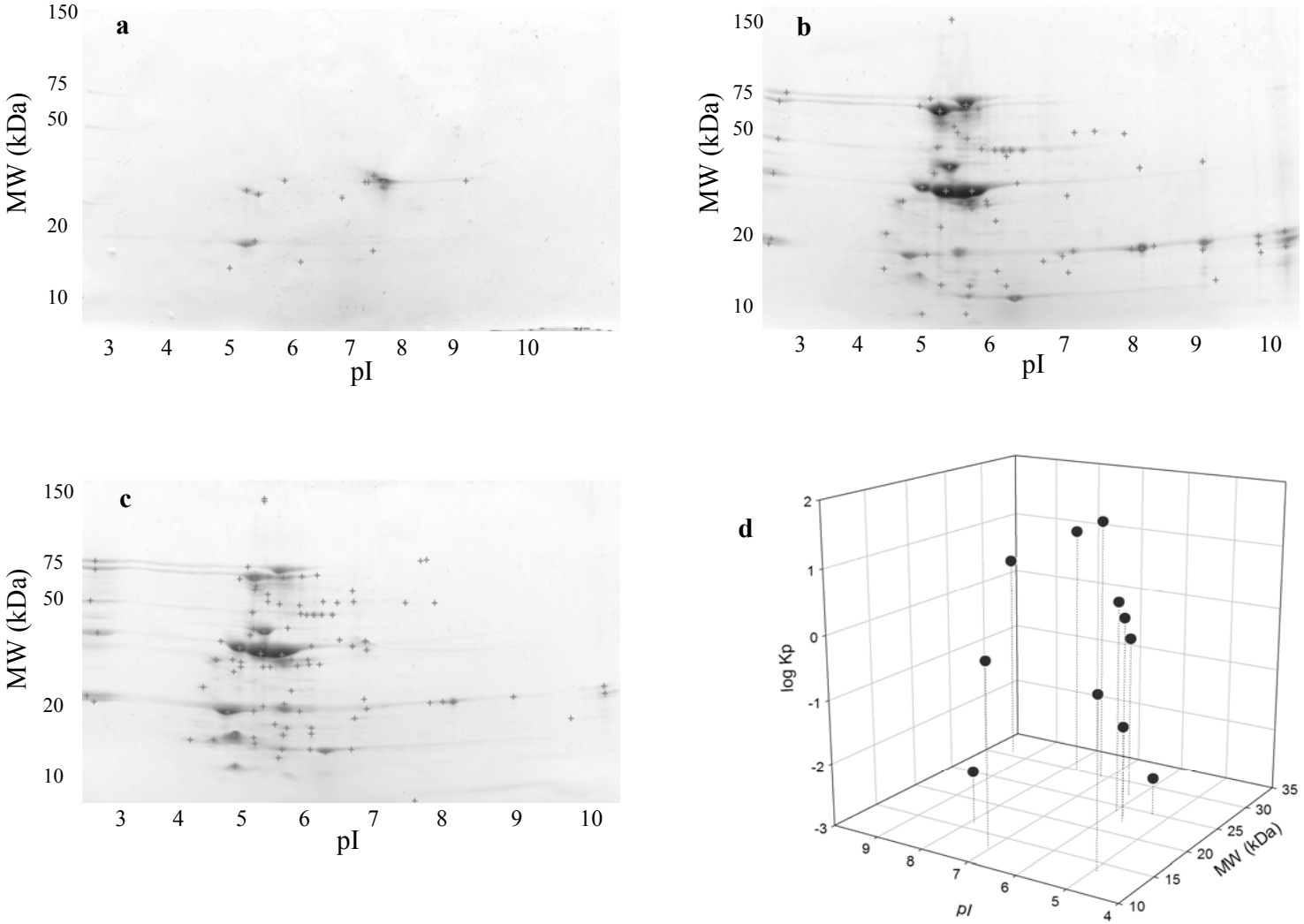
a
b
c
d
MW (kDa)
150
75
50
20
10
3
4
5
6
7
8
9
10
pI
log Kp
2
1
0
-1
-2
-3
pI
MW (kDa)
35
30
25
20
15
10

Figure 3

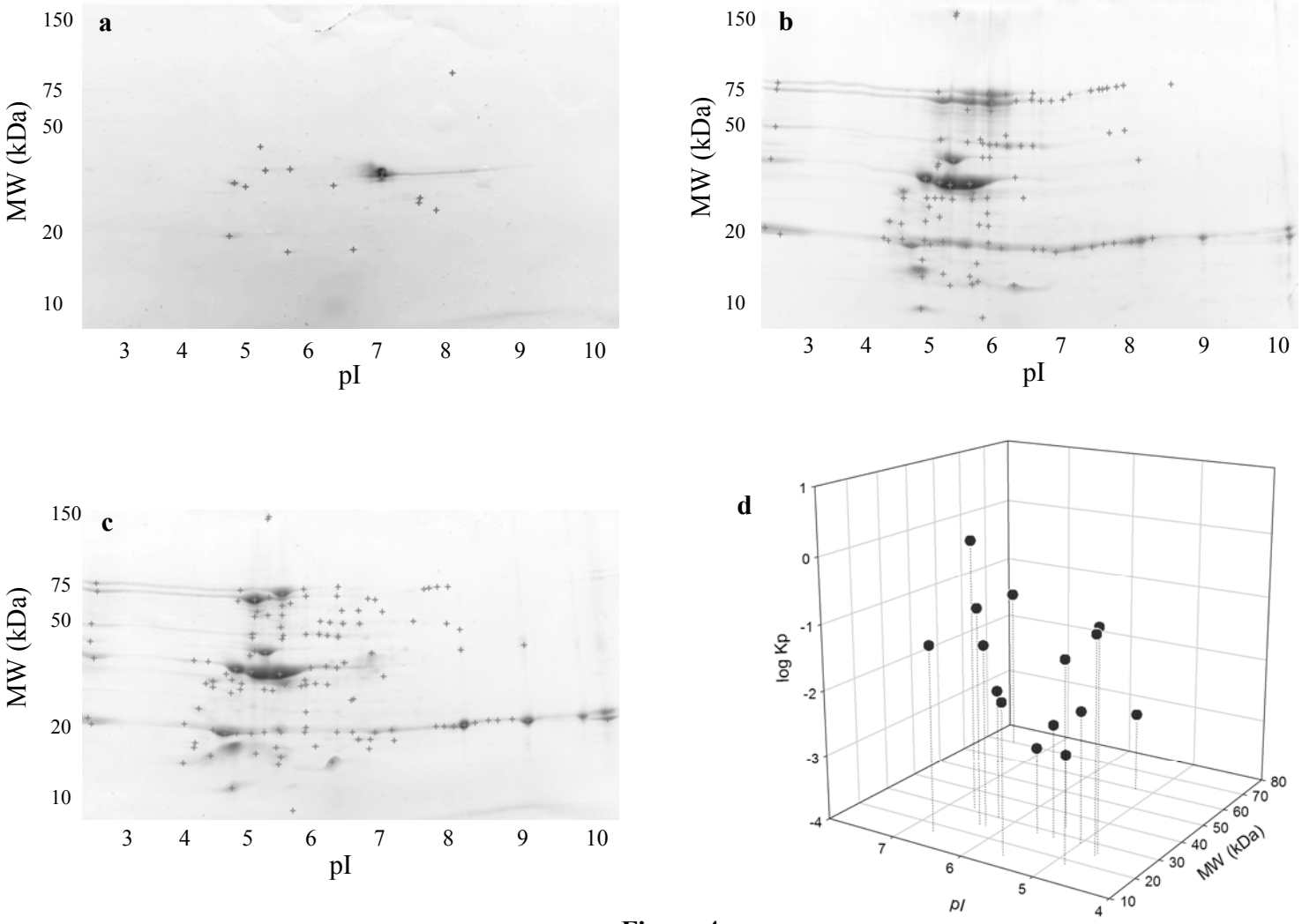
a
b
c
d
MW (kDa)
pI
150
75
50
20
10
3
4
5
6
7
8
9
10
log Kp
1
0
-1
-2
-3
-4
pI
MW (kDa)
80
70
60
50
40
30
20
10

Figure 4

Supplementary Material.

Table 1. Total protein spots detected and quantified from 2D gels after partitioning in PEG 3350-Na_2SO_4 system.

		Protein Concentration in ATPS, ppm		
M.W.	**pI**	**TOP PHASE**	**BOTTOM PHASE**	**TOTAL SOYBEAN EXTRACT**
19.9	3.0		21.59	107.30
37.5	3.0			8.75*
19.0	3.0		5.26	18.71
32.2	3.1		26.70	45.89
76.7	3.1		3.00	34.63
67.9	3.1		9.34	63.99
13.3	4.3			20.92*
15.5	4.4			11.14*
16.2	4.4		1.80	8.31
21.2	4.5		5.00	34.16
25.4	4.6		1.59	11.23
13.3	4.6			118.64*
24.7	4.7	0.30	1.70	15.54
18.1	4.7		15.67	38.05
26.0	4.7	2.68	13.03	77.31
17.5	4.8	2.82	36.40	356.98
10.2	4.9		18.00	129.06
13.5	5.0		48.74	224.13
28.0	5.0	0.70	69.53	361.16
17.4	5.1		7.54	48.54
24.8	5.1		1.97	36.16
71.5	5.1		7.77	35.01
12.6	5.2		2.80	38.08
32.2	5.2		5.19	33.07
14.0	5.2			48.20*
40.3	5.2		7.87	46.87
59.4	5.2	0.11	68.32	447.53
21.4	5.3		2.32	17.33
34.4	5.3		12.42	82.05
267.3	5.3		1.98	5.18
304.2	5.3		1.24	2.04
27.5	5.4	0.65	115.29	650.58
17.9	5.4		27.17	55.24
33.7	5.4	0.06	46.46	245.55

Table 1. ***(continuation)***

		Protein Concentration in ATPS, ppm		
MW	**pI**	**TOP PHASE**	**BOTTOM PHASE**	**TOTAL SOYBEAN EXTRACT**
61.2	5.4		13.07	117.89
14.5	5.6	0.18	13.44	163.2
27.9	5.6	0.35	169.31	739.39
42.7	5.6			24.33*
54.5	5.6		14.41	56.64
65.1	5.6		51.45	262.78
25.2	5.7		11.86	117.28
9.0	5.7		4.18	23.12
18.3	5.9		39.09	202.85
24.8	5.9		6.60	30.63
69.5	5.9		39.10	24.46
26.2	5.9			51.30*
20.7	6.0		1.88	7.08
40.1	6.0		6.81	33.12
63.2	6.0		44.38	34.72
40.0	6.0		3.75	14.12
15.7	6.1			18.31*
39.9	6.2		14.75	71.69
63.8	6.2		5.06	6.36
29.6	6.2	0.04	4.11	70.93
25.2	6.3	0.46	22.28	82.60
44.8	6.3			11.49*
25.0	6.3	0.22	1.05	13.08
39.2	6.4		6.27	45.17
72.6	6.4		2.28	4.26
44.4	6.5			17.08*
30.7	6.5		3.00	15.85
17.7	6.6		3.31	17.14
31.0	6.7	7.26	5.03	47.98
16.6	6.7	0.20	2.93	25.03
44.2	6.7			17.83*
31.3	6.8	79.29		40.52
17.2	6.8		17.04	67.54
59.8	6.9		2.24	5.31
30.2	6.9	30.16		39.35
18.1	7.2		19.17	47.80
26.1	7.3	3.65		57.35
45.5	7.4		0.89	6.71

Table 1. ***(continuation)***

MW	pI	Protein Concentration in ATPS, ppm		
		TOP PHASE	**BOTTOM PHASE**	**TOTAL SOYBEAN EXTRACT**
18.5	7.7		4.52	11.18
69.8	7.5	0.12	2.68	6.99
18.6	8.0		21.00	63.49
34.5	8.0		1.77	8.16
18.8	8.1		16.33	159.66
19.2	8.2		5.87	33.62
19.7	8.7		4.49	45.41
19.5	8.9		20.22	134.91
20.1	9.7		9.33	130.34
20.7	10.0		6.48	119.43
19.8	10.0		7.78	81.30
Total Protein Concentration, ppm		**129.23**	**1214.89**	**6903.99**
Total spots count:	**83**	*11 spots not found either phase		
Partition coefficient:	**0.11**			

All data are the average of two experiments using two-phase system: 15.7% PEG 3350, 8.9% Na_2SO_4, 3% NaCl, pH 7 and 25°C. Load of protein was 1.0 mg/g ATPS. Protein concentration is expressed in parts per million considering the average loss during TCA step for each phase.

Table 2. Total protein spots detected and quantified from 2D gels after partitioning in PEG 3350-phosphate system.

		Protein Concentration in ATPS, ppm		
MW	**pI**	**TOP PHASE**	**BOTTOM PHASE**	**TOTAL SOYBEAN EXTRACT**
16.0	3.0		17.60	167.06
42.3	3.0		9.34	58.60
15.1	3.1		2.36	28.67
62.8	3.1		7.07	64.01
71.4	3.1		9.80	51.77
30.7	3.1		11.67	55.48
10.0	4.4		1.74	8.68
17.6	4.6		5.89	25.32
10.0	4.7			87.77*
23.4	4.8		19.47	74.84
28.2	4.8			2.80*
13.5	4.9	23.77	55.25	387.64
23.6	5.0			20.98*
10.6	5.0		15.07	200.00
20.8	5.0			6.24*
7.5	5.0		6.63	129.29
26.6	5.1		104.29	435.35
54.8	5.1		14.99	0.58
21.7	5.1			11.18*
22.9	5.1			17.17*
13.1	5.1		5.34	12.57
65.0	5.2		7.44	27.67
30.1	5.2		9.67	32.77
37.6	5.2		15.84	40.06
13.7	5.3			50.42*
9.9	5.3		2.99	60.81
31.8	5.3			43.07*
55.7	5.3		272.54	495.08
48.7	5.3		36.45	86.59
25.0	5.4	1.19	249.96	767.51
14.3	5.4			5.37*
21.7	5.4			6.43*
27.1	5.4			133.48*
30.9	5.4		128.00	269.01

Table 2. ***(continuation)***

MW	pI	Protein Concentration in ATPS, ppm		
		TOP PHASE	BOTTOM PHASE	TOTAL SOYBEAN EXTRACT
250.0	5.4		2.82	5.25
65.3	5.4			19.36*
43.5	5.5		10.11	14.68
21.8	5.5	3.65		3.34
31.5	5.5			21.29*
10.9	5.5		0.75	9.82
14.0	5.6			15.87*
8.3	5.6		0.38	19.44
37.3	5.6		4.59	10.55
40.7	5.6		10.89	20.91
61.9	5.6		149.21	292.77
9.0	5.6		18.95	36.51
10.5	5.7		4.56	34.07
22.5	5.7	1.14	68.17	194.76
24.8	5.7		375.38	467.16
55.5	5.7		51.05	137.53
14.4	5.7		51.43	114.84
11.4	5.7			29.70*
32.1	5.7			5.13*
17.0	5.8			3.99*
23.4	5.8	3.17	9.69	20.81
22.0	5.8		12.02	32.28
40.4	5.9			2.48*
36.8	5.9		3.55	14.69
56.5	5.9		10.93	32.71
36.7	6.0		5.19	8.71
24.1	6.0	8.85	11.72	157.32
11.0	6.0		3.04	19.67
36.7	6.1		7.19	35.37
57.9	6.1			5.00*
22.4	6.2			6.41*
27.3	6.2	2.14	5.21	18.53
36.7	6.2		21.12	26.98
42.0	6.2			7.85*
9.2	6.2		62.59	74.99
36.7	6.3		29.28	51.74
41.7	6.4		6.55	14.30
30.1	6.4		5.66	10.96

Table 2. ***(continuation)***

MW	pI	Protein Concentration in ATPS, ppm		
		TOP PHASE	**BOTTOM PHASE**	**TOTAL SOYBEAN EXTRACT**
25.9	6.9	0.14		40.59
47.0	6.6		6.79	6.83
41.8	6.6		0.31	15.40
12.5	6.6		4.62	22.38
13.7	6.8		21.06	43.48
55.2	6.8			2.01*
29.7	7.1	18.78	0.48	50.64
12.9	7.1	0.69	0.48	33.24
29.8	7.2			2.08*
44.8	7.3		1.94	2.44
65.9	7.3			3.36*
68.3	7.4			6.24*
93.6	7.5			0.81*
71.4	7.5			2.54*
72.7	7.6			3.06*
14.9	7.6		16.90	2.33
41.3	7.7		1.32	4.13
16.2	7.9	0.48	80.62	64.06
30.2	7.8	29.23	1.26	2.72
15.0	8.0		5.61	172.11
15.8	8.8		71.08	136.69
31.0	9.5	2.30	0.47	12.44
12.5	9.6			1.03*
16.5	9.6		19.42	64.58
17.7	10.0		34.53	113.54
Total Protein Concentration, ppm		**95.52**	**2218.27**	**6670.77**
Total spots count:	**97**	*29 spots not found in either phase		
Partition coefficient:	**0.04**			

All data are the average of two experiments using two-phase system: 14.8% PEG 3350, 10.3% potassium phosphate, 3% NaCl, pH 7 and 25°C. Load of protein was 1.0 mg/g ATPS. Protein concentrations are expressed in parts per million in the ATPS considering the % of protein loss for each phase during TCA step.

CHAPTER FIVE

Application of a 3-D characterization technique to a green-tissue protein extract from alfalfa (*Medicago sativa*).

Oscar Aguilar[1], Charles E. Glatz[2], Marco Rito-Palomares[1]*

[1] Centro de Biotecnología, Departamento de Biotecnología e Ingeniería de Alimentos, Tecnológico de Monterrey. Campus Monterrey, Ave. Eugenio Garza Sada 2501 Sur, Monterrey, NL 64849, México.

[2] Departament of Chemical and Biological Engineering, 2114 Sweeney Hall, Iowa State University, Ames, Iowa 50011-2230, U.S.A.

* Author for correspondence: Marco Rito-Palomares, Ph.: +52 81 8328-4132; Fax +52 81 8328-4136. *e-mail address*: mrito@itesm.mx

ABSTRACT

There is a growing interest of pharmaceutical companies for plant-based production systems. To facilitate the general acceptance of plants as bioreactors, the establishment of efficient downstream operations is critical. It has been proposed that a better understanding of the properties of the contaminant proteins can benefit downstream processing design and operation. The coupled application of 2-D electrophoresis (2-DE) with aqueous two-phase partitioning has been suggested as a practical 3D method to characterize potential contaminant proteins from plant extracts. The application of this novel 3D approach to a complex protein extract from alfalfa (*Medicago sativa*) containing a model recombinant protein (human granulocyte-macrophage colony stimulating factor; GMCSF) resulted in the quantification of 55 protein spots. The three-dimensional properties (MW, pI and Kp) obtained for 17 proteins comprising 69% of the alfalfa proteins, allowed the proposal of a pre-fractionation step as well as the identification of the target molecule (rGM-CSF) from bulk of alfalfa proteins. The information obtained from this experimental approach was useful for identification of the potential contaminant proteins that will occur in alfalfa when this plant is used as a host for recombinant proteins. Additionally, this method will be useful on the design of adequate purification strategies for recombinant proteins expressed in alfalfa green tissue.

Keywords: *2D-electrophoresis; aqueous two-phase systems; alfalfa protein; GM-CSF; proteomics.*

INTRODUCTION

A wide number of pharmaceutical proteins have been produced in a variety of plant species (including tobacco, potato, rice, soybean alfalfa, tomato, lettuce, etc) reflecting the interest of pharmaceutical companies to benefit from the advantages of plant-based production systems (Ma et al., 2003, Stoger et al., 2002). A main aspect that requires attention when a recombinant protein production system is defined involves the design of efficient extraction and purification methods to maximize recovery of target protein. Downstream processing costs typically contribute with more than 80% of the total, and efficient and robust processing strategies are therefore essential (Abranches, et al., 2005). In this context, aqueous two-phase systems (ATPS) based strategies have been proved to have potential for recovery and purification of biological compounds (Gu and Glatz, 2007a; Platis and Labrou, 2006; Benavides et al., 2006). ATP partitioning has been used also for understanding the chemical properties and behavior of proteins in solution (Berggren et al., 2002; Hachem et al., 1996).

A better understanding of the properties of the contaminant proteins can benefit downstream processing design and operation (Gu and Glatz, 2007b; Asenjo and Andrews, 2004). Proteomic tools like mass spectrometry and two-dimensional electrophoresis (2-DE) have become common techniques for accurately detect and examine protein composition from a variety of plant hosts. These techniques provide useful information on the molecular properties of complex mixtures that can be exploited for optimization and better design of downstream strategies (Natarajan et al., 2006).

A three-dimensional technique for molecular characterization of corn germ protein extracts was recently reported by Gu and Glatz (2007b) and it was based on the coupled application of ATP partitioning to measure hydrophobicity in terms of Kp, and two-dimensional electrophoresis (2-DE) to evaluate molecular weight (MW) and isoelectric point (pI) of individual proteins. The three-dimensional information obtained for each protein (MW, pI and hydrophobicity) was used as a basis for better visualization of the molecular properties of the surrounding environment where a recombinant protein could be eventually produced. A recent extension of this technique to a different plant protein extract (soybean) was recently addressed by Aguilar et al. (2007). However, in order to establish the generic application of this experimental approach, alternative plant systems need to be considered.

The aim of this work was to evaluate the application of the tree-dimensional technique for characterization of alfalfa protein extract containing an artificially added recombinant model protein. Human granulocyte-macrophage colony stimulating factor (rGM-CSF) was selected as an example of a recombinant product that could be readily produced in alfalfa cells at low quantities. Colony-stimulating factors are glycoproteins which act on hematopoietic cells by binding to specific cell surface receptors and stimulating proliferation, differentiation commitment, and some end-cell functional activation (Neupogen® Data Sheet). Previous reports have demonstrated the feasibility for recombinant human cytokines production in genetically modified plant cells (Kwon et al., 2003; James et al., 2000). One of the most relevant characteristics of proteins produced in plants is the possibility of production of glycosylated forms of a recombinant protein. Any improvement in the function or life of the drug will have a remarkable impact for patients with cancer or bone marrow transplantation

(Hoglund, 1998). The possibility of producing glycoforms of GM-CSF in a commercially viable plant system has an enormous potential that can be explored, considering the increase in the biological activity of the molecule and the reduction in the production costs. However, such experimental model represents a real challenge for downstream processing given the high concentration of contaminant proteins that would be present.

In this research, the effect of several ATPS parameters (NaCl content, TLL and phase-forming salt) was evaluated in order to maximize alfalfa protein partitioning in hydrophobic two-phase systems. The approach proposed here involving the coupled application of ATPS and 2-DE was for the first time explored for green-tissue proteins extracted from aerial parts of alfalfa.

MATERIALS AND METHODS.

Materials.

Poly(ethyleneglycol) MW 3350, β-mercaptoethanol, DL-dithiothreitol (DTT), Tris(hydroxymethyl)aminomethane (Tris), Bradford reagent and the selected model proteins: ribonuclease A (RNA), α-lactalbumin (LAC), bovine serum albumin (BSA) and lysozyme (LYS) were purchased from Sigma-Aldrich Chemicals (St Louis, MO). The ReadyPrep® rehydration buffer, 11 cm ReadyStrip® IPG strips (pH 3-10), iodoacetamide, and Precision Plus® protein standard plugs were purchased from Bio-Rad (Hercules, CA). Phenylmethylsulfonyl fluoride (PMSF) was purchased from Boehringer Mannheim Gmbh (Germany). Commercial rGM-CSF (Biofilgran®) was obtained from BioSidus S.A. (Buenos

Aires, Argentina). All the other chemicals were purchased from Sigma-Aldrich Chemicals (St Louis, MO).

Plant materials.

Commercial alfalfa seeds (*Medicago sativa*), kindly provided by the Agricultural Experimental Station of Tecnológico de Monterrey were field grown during 4 weeks at 25°C. Aerial parts (stems and leaves) were harvested before flowering and immediately ground in liquid nitrogen with mortar and pestle, adding powdered glass to improve cell wall breaking. Powder stocks were stored at -86°C for further use.

Alfalfa green-tissue protein extraction.

Three different protocols (listed in Table 1) were evaluated for protein extraction from powdered alfalfa green tissue at a proportion of 1.0 gr solids/10 mL extraction buffer. Protocol A: Green biomass was suspended in sodium phosphate buffer (Buffer A) and stirred for 1 hr with constant pH monitoring, centrifuged (10,000 *x g*, 30 min, room temperature; Centrifuge 5804R, Eppendorf, Hamburg, Germany) and decanted to eliminate waste solids. The supernatant was filtered using 0.45 μm syringe filter (Corning Inc, U.S.A.) and used immediately (Platis et al, 2006). Protocol B: Protein extraction with $MgSO_4$-based buffer (Buffer B) was previously reported by Piridioni et al. (2006) in preparation of *Arabidopsis* protein crude extracts and it was evaluated for alfalfa green tissue. The slurry was stirred for 1 hr and centrifuged at 16,000 *x g* for 30 minutes at room temperature and the supernatant filtered using 0.45 μm syringe filter. Protocol C: TBE extraction buffer (Buffer C) was evaluated using the same biomass:buffer proportion. The slurry was stirred for 1 hr, then

centrifuged at 12,000 *x g* for 10 minutes at room temperature and the supernatant filtered using 0.45 μm syringe filter (Srere, 1969).

An artificial mixture of alfalfa protein containing human rGM-CSF was prepared by adding an aliquot of the purified human cytokine to the selected alfalfa protein extract (derived from protocol A, B or C) to have a final concentration of 63 μg of rGM-CSF /mL extract. This concentration of protein was found to be within the range of previously reported levels of this cytokine in plant cells (James et al., 2000).

Total protein determination for alfalfa extracts, phase samples and TCA precipitates was made using microplate Bradford reaction with BSA as standard (Bradford, 1976). Concentration of model proteins was measured spectrophotometrically at 280 nm using a microplate reader (Synergy HT, BioTek Instruments Inc., Vermont, U.S.A). All protein determinations included calibration curves using proper solvents and blank ATPS for correction of any interference from phase-forming components.

Aqueous two-phase partitioning.

ATP systems were formulated according to the corresponding binodal curves reported by Zaslavsky (1995) and the systems reported by Gu and Glatz (2007b) and Aguilar and Rito-Palomares (2007) to give a fixed weight of 2.0 g for alfalfa protein extracts and model proteins. PEG 3350–potassium phosphate and PEG 3350–sodium sulfate systems were tested in order to compare protein partitioning among them. Predetermined quantities of stock solutions of PEG 3350, and sodium sulfate or potassium phosphate were mixed with solid

NaCl and protein sample to give the following total composition: 15.7% PEG 3350, 8.9% Na_2SO_4, 3% NaCl for sulfate-ATPS and 14.8% PEG 3350, 10.3% potassium phosphate, 3% NaCl for phosphate-ATPS. Sample was added according to the protein concentration to have 1 mg protein/g ATPS for alfalfa protein extracts and for individual model proteins. NaCl effect on protein partition was also evaluated from varying from 0% to 9% with the same protein load. All concentrations were calculated in a weight basis, pH of the systems and protein extracts was adjusted to 7.0 by addition of 1 M orthophosphoric acid or potassium hydroxide if needed. After 1 hr of gentle mixing at 25°C, complete phase separation was achieved by low-speed batch centrifugation at 1000 x *g* for 20 min. Visual estimates of the volumes of top and bottom phases were made in graduated tubes and used to estimate the experimental volume ratio (Vr). Top phase samples were carefully taken by pipetting, while bottom samples were extracted by piercing the bottom of the tube with a syringe. The interphase was left in the tube to avoid contamination. All partitioning experiments with alfalfa protein extracts and model proteins were run by triplicate.

TCA precipitation

To eliminate interferences from phase-forming compounds and increase protein concentration, trichloroacetic acid (TCA) precipitation was performed to the top and bottom samples before isoelectric focusing. Equal volumes of sample and ice cold 20% (w/v) TCA were thoroughly mixed in a microcentrifuge tube and incubated on ice for 15 min before being centrifuged at 13,000 x *g* for 5 min at 4°C. The supernatant was decanted from the tube and the precipitate was washed once using ice-cold 10% (w/v) TCA, followed by three washes of acetone (previously stored at -20 for 30 minutes), and centrifuged at 13,000 x *g* for

5 min at 4°C. All samples were incubated on ice for 5 minutes before centrifugation. The final protein pellet was air-dried for 5 minutes, or until residual acetone was evaporated being careful to not overdry the sample. In the cases where protein recovery from phases was low, multiple replicates were pooled at this step and considered for final protein quantitation. The pellet was completely re-dissolved using 210 μL of rehydration buffer (8M urea, 2% (w/v) CHAPS, 50mM DTT, 0.2% (v/v) BioLyte from Biorad) and used for first-dimension isoelectric focusing. Protein concentration was measured at this step diluting samples 1:3 to avoid highly concentrated urea interfere with Bradford reagent. All calibration equations for protein measurement were obtained using the proper solvents and dilutions for correction of any interference.

Isoelectric focusing

The first-dimension isoelectric focusing (IEF) was performed using 11 cm pH 3-10 linear immobilized pH gradient strips (IPG ReadyStrip®, Bio-Rad) in an Ettan IPGphor3 apparatus (GE Healthcare). Strips were rehydrated according to manufacturer's recommendations with 200 μL of sample to a maximum of 200 μg of protein per strip during 16 h at room temperature. IEF was carried out for a total of 50,250 Vh.

Second-dimension electrophoresis

For the second dimension, the focused IPG strips were equilibrated with 6 M urea, 75mM Tris HCl, 2% (w/v) SDS, 29.3% (v/v) glycerol, 0.002% (w/v) bromophenol blue pH 8.8 and 2% (w/v) DTT for 15 minutes, and then acetylated using 2.5% (w/v) iodoacetamide instead of DTT for another 15 min. Strips were placed onto 12.5% (w/v) linear polyacrylamide gels

prepared as described by Laemmli (1970) and electrophoresis was performed using a SE600 Ruby electrophoresis unit (GE Healthcare). The gels were visualized by staining with Coomassie Blue G-250, and scanned at 600-dpi resolution using a flat bed scanner in transmissive mode (Hewlett-Packard). Spot densitometry (area multiplied by the pixel intensity) was performed using PDQuest software (Bio-Rad). The mass of protein for individual spots was calculated from the spot density and relative to the total amount of protein loaded to the gel. Protein concentrations from spot densities were used to calculate partition coefficient (Kp, the rate between top and bottom concentrations) of individual proteins. All the experiments were run by duplicate.

Protein assay

Total protein determination for alfalfa extracts, phase samples and TCA precipitates was made using microplate Bradford reaction with BSA as standard (Synergy HT, Bio-Tek Instruments Inc., Vermont, USA) (Bradford, 1976). Concentration of model proteins was measured at 280 nm in microplate reader. All protein determinations included calibration curves using proper solvents and blank ATPS for correction of any interference from phase-forming components.

RESULTS AND DISCUSSION.

Protein extraction

In order to minimize protein degradation during grinding steps, liquid nitrogen was used to freeze-dry the stems and leaves immediately after harvesting. Three different extraction methods were tested for extraction of alfalfa proteins from green tissue, the results are shown

in Table 1. To avoid proteolysis during extraction at room temperature, all the extraction buffers included a protease inhibitor in the formulation either EDTA for sequestering metal ions needed for metalloproteases activity, or PMSF reported to be an irreversible inhibitor for serine and cysteine proteases (Berkelman and Stenstedt, 2004). The amount of protein extracted per gram of fresh alfalfa was found to be similar between $MgSO_4$-based and sodium phosphate-based buffers, around 50% of the total protein reported for alfalfa leaves (30 mg/g fresh weight; Abranches et al., 2005). However, Tris-borate-EDTA (TBE) buffer showed almost double capacity to solubilize leaf proteins (27.1 g protein/g fresh alfalfa) recovering up to 90% of the proteins, probably explained by the higher salt concentration. TBE buffer was further used for protein extraction in partition experiments and 3D analysis.

Aqueous two-phase partitioning experiments.

Partitioning experiments performed using the same system compositions previously reported by Gu and Glatz (2007) for corn protein extracts resulted in low partition coefficients, as can be seen in Table 2. These hydrophobic ATP systems produced low Kp values showing bottom phase preference (0.03 for PEG 3350-phosphate and 0.05 for PEG 3350-sulfate) however, total recovery of soluble proteins was low (37%). The hydrophobic nature of these systems caused an elevated accumulation of proteins at the interface confirmed visually and by the mass balances. The partition of the proteins in systems with high molecular weight of polymer and NaCl is dominated by hydrophobic interactions. An explanation for the partition behavior of alfalfa proteins in these ATP systems may involve the hydrophilic nature of these proteins that were extracted from the green tissue. Consequently, low preference for both hydrophobic phases was shown by the extracted proteins and accumulated at the interface.

The criteria to evaluate the choice of a particular ATPS composition for 3D characterization are defined by: (1) a partition coefficient (Kp) closest to 1.0, in order to obtain the highest number of proteins in both phases; (2) the highest % of recovery, to keep most of the proteins soluble and not at the interface; and (3) keep an intrinsic hydrophobicity that allows accurate hydrophobicity measurements.

The first attempt to improve alfalfa protein partitioning included formulation of ATP systems with different NaCl composition to evaluate the effect of this salt on the distribution of alfalfa proteins in hydrophobic systems. It has been reported that increasing NaCl content also increases the hydrophobic differences between the two phases, favoring the polymer-protein interactions (Hachem et al., 1996). Table 3 shows that increasing NaCl from 0 to 3% in both PEG 3350-phosphate and PEG 3350-sulfate systems causes a decrease in Kp for total alfalfa protein. In the case of PEG 3350-phosphate systems, a higher NaCl content increases Kp values. Even when these values could be considered as favorable for top phase protein content, total recoveries (Table 3) showed that this increase on Kp values with NaCl is actually explained by a decrease in protein content of the bottom phase, not for a real increase of protein to the top phase.

For both systems (PEG 3350-phosphate and PEG 3350-sulfate) the highest amount of protein recovered was observed with 0% NaCl. Total protein recovery for PEG 3350-phosphate system increase up to 58%. In the case of PEG 3350-sulfate system, addition of 3% NaCl

resulted in higher total recovery but only 1.5% at the top phase. For both systems, the absence of NaCl resulted in better protein distribution between the phases.

The effect of NaCl added to ATPS has been documented (Cascone et al., 1991; Hachem et al., 1996). Such studies concluded that while some model proteins are dramatically affected by NaCl addition, some others remain unaffected. In the case of complex protein mixtures, such as alfalfa extracts, some of the changes on Kp and recoveries observed in Table 3 can be explained in terms of changes in the solubility of proteins that precipitate at the interface.

Tie-line length (TLL) was also evaluated as an attempt to improve partition and recovery of protein in ATPS. It has been reported that the increase in TLL generally causes an increment in partition coefficient (salting out effect). However in some cases it also results in massive precipitation at the interface (Huddleston et al., 1991; Rito-Palomares, 2004; Platis and Labrou, 2006). Table 4 shows the two different TLL's tested for each system. In the case of PEG 3350-phosphate systems, the increase in TLL resulted in an increase in partition coefficient from 0.14 to 4.1. This dramatic increase could be attributed to a lower recovery at the bottom phase, probably caused by precipitation at the interface, also reflected on the decrease in total recovery from 58.3% to 23.3%. On the other hand, PEG 3350-sulfate system presented the opposite effect, a minimal increase in total recovery (29.4% to 31.6%) caused by the increase in TLL.

Considering these results and the previous criteria for the selection of the most adequate partitioning system, a system with lower concentrations of PEG 3350-phosphate (TLL=30%)

and no NaCl added was preferred for partitioning of proteins from alfalfa green tissue. The system composed with PEG 3350 (14.8%) and potassium phosphate (10.3%) resulted in a Kp of 0.14 and % of recovery of 58.3%. This system was selected to perform ATP partitioning prior to 2-D electrophoresis in order to obtain the highest number of proteins from both phases. This information was used for the three-dimensional characterization of alfalfa proteins.

Partition of model proteins in ATPS

The use of two-phase partitioning to estimate hydrophobicity of proteins has been addressed before (Andrews et al., 2005; Hachem et al., 1996; Franco et al. 1996). To test the ability of these ATPS to provide an accurate hydrophobicity measurement, several partition experiments were performed using four model proteins (LYS, BSA, RNA and LAC) with known hydrophobicity values. The parameter 1/m* reported by Hachem et al. (1996) for several model proteins measures the hydrophobicity of a protein derived from precipitation curves in $(NH_4)_2SO_4$. This parameter represents the salt concentration at which the protein starts to precipitate (given a initial fixed concentration). The functional hydrophobicity of the protein in solution is measured by 1/m* as the result of interactions between protein surface and the solvent. These hydrophobicity values have been reported for several model proteins and were used to validate our ATPS as an accurate method to estimate hydrophobicity. A correlation was performed between 1/m* parameter and the corresponding Kp value for selected model proteins partitioned in PEG 3350-phosphate system. The high linear correlation factor obtained of 0.96 (see Figure 1) between Kp measured in ATPS and the parameter 1/m* demonstrated that the system composed by PEG 3350 (14.8%) potassium

phosphate (10.3%) without NaCl added, provided a reliable method to measure the functional hydrophobicity of proteins.

Application of 3-D characterization method to green-tissue alfalfa extracts containing a recombinant protein.

As described in materials and methods section, proteins from aerial parts of alfalfa (stems and leaves) were extracted with TBE buffer at pH 8; protein extracts with an average 2.71 mg protein/mL were obtained. In order to simulate a complex protein extract where a recombinant protein would be the target of the downstream processing, an artificial mixture containing purified human rGM-CSF was prepared and loaded into the ATPS. Previous reports indicate that this cytokine has been successfully produced in a variety of plant cells (i.e. tobacco, and tomato) at concentrations ranging from 0.1 to 0.5% of total soluble protein (TSP) (James et al., 2000; Kwon et al., 2003). Artificial alfalfa protein mixtures containing 63 μg of rGM-CSF/mL extract, were used for 3D characterization. This higher concentration equivalent to 20 μg of rGM-CSF/mg alfalfa soluble protein (or 2% TSP) was used with the final purpose of facilitating protein identification after the 3D technique.

After ATP partitioning, two-dimensional gels of top and bottom phases were run under the same electrophoretic conditions, resulting in 55 protein spots with a wide range of molecular weight and pI distributions among the phases, as can be seen in Figure 2. Removal of phase-forming components using TCA precipitation implied certain protein losses that were quantified for each phase and were considered for calculation of yields for individual

proteins. From the total amount of proteins, only 17 proteins detected on 2-D gels were found in both top and bottom phases.

The overall partition coefficient estimated by spot densitometry was 0.13±0.01 (from Table 5), and resulted to be statistically equal to that measured by the Bradford method (0.14±0.02; Table 4). Although these results demonstrate that both methods (2-DE gels and Bradford) can be readily used to estimate Kp obtaining similar values, it is important to consider the large amount of protein not solubilized by this ATPS and that still remains at the interface (aprox. 42%).

As can be seen Table 5, most of proteins that were not detected in top phase are those with a higher value of pI, specifically those with pI>6. Additionally, except for a couple of spots, most of the proteins that were detected only at the bottom phase have higher MW values, between 40-60 kDa. This behavior can be attributed to a limited "free" volume on top phase to accommodate bigger proteins, keeping most of them dissolved at the interphase or bottom phase (Benavides et al., 2006; Huddleston et al., 1991). Although the number of protein spots detected in both phases was similar, the concentration of such proteins was generally higher for bottom phase proteins. Kp<1 indicated a clear preference of the proteins for the this phase.

The extension of this application to the green-tissue proteins of alfalfa resulted in molecular characterization of 17 proteins listed in Table 6. These protein spots were those detected in both top and bottom phases and they together comprised 69% of the total proteins detected

on 2-D gels. This means that 31% of the proteins showed exclusive affinity for one of the two phases, and Kp could not be calculated in these cases. Figure 2 illustrate the scatter plot obtained from three-dimensional properties of proteins from Table 6. Except for a couple of spots, most of the proteins showed bottom phase preference, confirmed by the negative values for Log Kp axis. This subset of proteins showed a limited pI range as the result of the preference of proteins with pI values higher than 6.5 for the bottom phase only.

Protein spot identification can be partially addressed by the information obtained from 2D gels (MW, pI), in this case, spiked samples containing rGM-CSF were used in order to simulate a recombinant extract. MW and pI information are usually not conclusive to establish the chemical identity of protein spots detected in 2-DE, a more detailed analysis using mass spectrometry could confirm the identity. However, in the case of alfalfa and the majority of green-tissue protein extracts, the major contaminant protein that can be found is ribulose-1,5-bis phosphate carboxylase (Rubisco, E.C. 4.1.1.39). This tetrameric enzyme (and its subunits) accounts for 30-50% of total protein from plant tissues (Platis et al., 2006). Rubisco from alfalfa has a reported molecular weight of approximately 48-52 kDa and pI values of 5-5.5 (Gasteiger et al., 2005). Although it has been reported that rubisco precipitate at the interphase of PEG-phosphate systems (Platis et al., 2006), the Figure 2b shows spots corresponding to the predicted MW and pI values for this enzyme. The spot with 49.9 kDa and pI 5.7 observed in Table 6 probably corresponded to fractions of rubisco in solubilized form, mostly at the bottom phase. Four subunits of approximately 14 kDa could explain the couple of spots obtained at 14.9 kDa and pI 5.8, and 14.2 kDa and pI 6.5. This set of spots together accounted for 71% of the subset of proteins detected in both phases.

Regarding the granulocyte-macrophage colony stimulating factor (rGM-CSF) added to alfalfa protein extract, the reported MW and pI parameters are 14.4 kDa and 5.2, respectively (Gasteiger et al., 2005). Preliminary partition experiments for this protein using the same hydrophobic ATPS (data not shown) showed a higher preference for bottom phase, this hydrophilic behavior agrees with its GRAVY value of -0.448 (Gasteiger et al., 2005). The grand average of hydropaticity scores (GRAVY) provide a picture of the hydrophobicity of the whole protein, usually ranging from 2 to -2, where positive scores indicate hydrophobic, and negative scores hydrophilic proteins (Garbis et al, 2005).

Considering this apparent hydrophilic behavior and the molecular properties reported for this cytokine, only one spot (15.1 kDa and pI 5.6; in Table 6) could be potentially identified as the rGM-CSF added to alfalfa extract. With the low relative amount of this protein (20 μg of rGM-CSF/mg alfalfa soluble protein) added to the extract, further identification methods have to be performed (mass spectrometry, western blot, etc.) to positively conclude the chemical identity of this protein spot.

The results reported here evidenced the importance of case-by-case optimization of ATP partitioning in order to maximize partition to top and bottom phases, and minimize protein precipitation at the interphase. The lower the number of proteins found at the interphase phase, the higher the number of proteins that can be 3D-characterized, and consequently a more complete protein profile can be obtained. It was found that proteins extracted from alfalfa green tissues tend to accumulate at the interphase of hydrophobic ATP systems,

resulting in a limited protein profile that could be obtained from the 3D characterization technique. However, a tentative identification was performed in the case of the model recombinant protein added to the extract (rGM-CSF). For most of the green-tissue protein extracts, the main potential contaminant that can be found, rubisco could be potentially removed by harvesting the interphase of an optimized ATP system specially formulated with that purpose without compromise recovery of target protein. This pre-fractionation strategy proposed could overcome one of the main disadvantages of 2-D electrophoresis regarding detection of low abundant proteins when masked with more abundant proteins.

CONCLUSIONS.

A three-dimensional characterization method was applied to alfalfa green-tissue proteins, providing information on the molecular properties of a large number of host proteins. The ATP system composed of PEG 3350 (14.8%) and potassium phosphate (10.3%) resulted in 69% of total proteins partitioned in both phases from which 3D properties were calculated. In the case of rGM-CSF, a further identification method (like western blot or mass spectrometry) would be needed in order to confirm the chemical identity of the protein spot tentatively identified as this human cytokine.

It was evident that a careful selection of the ATP partitioning system has to be done prior to 2-DE to optimize protein partitioning. This is a necessary step in the case of highly hydrophilic proteins added into ATP systems where hydrophobicity is the ruling principle for partitioning. However, as a result of the application of the 3D characterization technique to alfalfa protein extracts, the molecular properties obtained for the main contaminant proteins

will allow the design of pre-fractionation and purification methods to facilitate downstream processing of a recombinant alfalfa extract.

ACKNOWLEDGEMENTS.

Authors wish to acknowledge the financial support of Tecnológico de Monterrey, Biotechnology Research Chair (Grant CAT005) and BioMaP REU Program 2007 for the technical assistance of Mattan Rojstaczer during this research.

REFERENCES.

Abranches R, Sylvain M, Arcalis E, Altmann F, Fevereiro P, Stoger E. 2005. Plants as biorreactors: A comparative study suggests that *Medicago truncatula* is a promising production system. J Biotechnol 120:121-134.

Aguilar O, Rito-Palomares M, Glatz CE. 2007. Coupled application of aqueous two-phase systems and 2D-electrophoresis for three-dimensional characterization of soybean proteins (*In preparation*).

Aguilar O, Rito-Palomares M. 2007. Processing of soybeans (*Glycine max.*) extracts in aqueous two-phase systems as a first step for the potential recovery of recombinant proteins. J Chem Technol and Biotechnol DOI:10.1002 (*In Press*).

Andrews BA, Schmidt AS, Asenjo JA. 2005. Correlation for the partition behavior of proteins in aqueous two-phase systems: Effect of surface hydrophobicity and charge. Biotechnol Bioeng 90:380-390

Asenjo JA, Andrews BA. 2004. Is there a rational method to purify proteins? From expert systems to proteomics. J Mol Recognit 17:236-247.

Benavides J, Mena JA, Cisneros-Ruiz M, Ramírez OT, Palomares LA, Rito-Palomares M. 2006. Rotavirus-like particles primary recovery from insect cells in aqueous two-phase systems. J Chromatogr B 842:48-57.

Berggren K, Wolf A, Asenjo JA, Andrews BA, Tjerneld F. 2002. The surface exposed amino acid residues of monomeric proteins determine the partitioning in aqueous two-phase systems. Biochim Biophys Acta 1596:253-268.

Berkelman T, Stenstedt T. 2004. 2-D Electrophoresis. Principles and Methods. GE Healthcare. Buckinghamshire, UK.

Cascone O, Andrews BA, Asenjo JA. 1991. Partitioning and purification of thaumatin in aqueous two-phase systems. Enzyme Microb Technol 13:629-635.

Franco TT, Andrews AT, Asenjo JA. 1996. Use of chemically modified proteins to study the effect of a single protein property on partitioning in aqueous two-phase systems: Effect of surface hydrophobicity. Biotech Bioeng 49:300-308.

Garbis S, Lubec G, Fountoulakis M. 2005. Limitations of current proteomics technologies. J Chromatogr A 1077:1-18.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM, editor. The Proteomics Protocols Handbook, Humana Press. pp. 571-607. http://www.expasy.org/tools/protparam.html.

Gu Z, Glatz CE. 2007 (a). Aqueous two-phase extraction for protein recovery from corn extracts. *J Chromatogr B* 845:38–50.

Gu Z, Glatz CE. 2007 (b). A method for three-dimensional protein characterization and its application to a complex plant (corn) extract. Biotechnol Bioeng 97:1158-1169.

Hachem F, Andrews BA, Asenjo JA. 1996. Hydrophobic partitioning of proteins in aqueous two-phase systems. Enzyme Microb Tech 19:507-517.

Hoglund M. 1998. Glycosylated and non-glycosylated recombinant human granulocyte colony-stimulating factor (rhG-CSF)--what is the difference? Med Oncol 15:229-33.

Huddleston JG, Ottomar KW, Ngonyani DM, Lyddiatt A. 1991. Partitioning and purification of α-amylase in aqueous two-phase systems. Enzyme Microb Technol 16:131-142.

James EA, Wang C, Wang Z, Reeves R, Han Shin J, Magnuson NS, Lee JM. 2000. Production and characterization of biologically active human GM-CSF secreted by genetically modified plant cells. Prot Express Purif 19:131-138.

Kwon T-H, Kim Y-S, Lee J-H, Yang M-S. 2003. Production and secretion of biologically active human granulocyte-macrophage colony stimulating factor in transgenic tomato suspension cultures. Biotechnol Lett 25:1571-1574.

Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685.

Ma JK-C, Drake PMW, Christou P. 2003. The production of recombinant pharmaceutical proteins in plants. Nat Rev Genet 4:794–805.

Natarajan SS, Xu C, Hanhong B, Caperna TJ, Garret WM. 2006. Characterization of storage proteins in wild (*Glycine soja*) and cultivated (*Glycine max*) soybean seeds using proteomic analysis. J Agric Food Chem 54:3114-3120.

Neupogen® Data Sheet. 2004. Amgen Inc. Thousand Oaks, CA.

Pirondini A, Visioli G, Malcevschi A, Marmiroli N. 2006. A 2-D liquid-phase chromatography for proteomic analysis in plant tissues. J Chromatogr B 833:91-100.

Platis D, Labrou NE. 2006. Development of an aqueous two-phase partitioning system for fractionating therapeutic proteins from tobacco extract. J Cromatogr A 1128:114-124.

Rito-Palomares M. 2004. Practical application of aqueous two-phase partition to process development for the recovery of biological products. J Chromatogr B 807:3-11.

Srere PA. 1969. Citrate Synthase. In: Packer L. editor. Methods in Enzymology Vol. 13. London. Academic Press pp. 3-11.

Stoger E, Sack M, Fischer R, Christou P. 2002. Plantibodies: applications, advantages and bottlenecks. Curr Opin Biotechnol 13:161–166.

Twyman RM, Stoger E, Schillberg S, Christou P, Fischer R. 2003. Molecular farming in plants: host systems and expression technology. Trends in Biotechnol 21:570-578.

Zaslavsky A. 1995. Aqueous two-phase partitioning. Physical Chemistry and Bioanalytical Applications. New York; Marcel Dekker Inc. 660 p.

Legend of Figures

Figure 1. Correlation between surface hydrophobicity measured as log Kp in ATPS with 1/m* parameter. Kp is the partitioning coefficient of protein in PEG 3350 (14.8%), potassium phosphate (10.3%) system at pH 7, 25°C and protein loading of 1.0 mg/g ATPS each protein, and m* is the concentration of salt at the discontinuity point of the protein in $(NH_4)_2SO_4$ precipitation curve, reported by Hachem *et al*. (1996) for selected model proteins using an initial concentration of 2 mg/mL.

Figure 2. 2D gels and 3D scatter plot of green-tissue proteins from alfalfa using PEG 3350-phosphate system. ATPS: PEG 3350 (14.8%), potassium phosphate (10.3%) at pH 7, 25°C and 1.0 mg protein/g ATPS. **a**: top phase gel, **b**: bottom phase gel, **c**: 3D scatter plot of alfalfa proteins detected in both phases (see full data in Table 6).

Table 1. Efficiency of alfalfa protein extraction with three different buffers[a].

Extraction Buffer	Extracted protein, mg prot./g fresh alfalfa[a]
Protocol A	
20 mM sodium phosphate 10 mM EDTA pH 7	15.4±0.5
Protocol B	
50 mM Tris HCl 10 mM MgSO4 0.1% β-mercaptoethanol 2 mM phenylmethylsulphonifluoride pH 7.8	14.6±1
Protocol C	
0.45 M Tris.HCl 0.45 M H_3BO_3 10 mM EDTA pH 8	27.1±2

[a] Alfalfa green-tissue ground on liquid N_2 and extracted for 1 h at 25°C with constant pH monitoring with a 1:10 plant material:buffer proportion.
[b] Expressed as the average of triplicate experiments. Protein concentration measured by Bradford reaction.

Table 2. Partition coefficient of alfalfa proteins in selected ATPS.

ATP system	Kp[a]	% Total protein recovery[b]
PEG 3350-PO_4 system		
14.8% (w/w) PEG 3350 10.3% (w/w) Phosphate 3.0% (w/w) NaCl	0.03±0.01	37.3±2.0
PEG 3350-SO_4 system		
15.7% (w/w) PEG 3350 8.9% (w/w) Na_2SO_4 3.0% (w/w) NaCl	0.05±0.001	37.4±1.0

[a] Partition coefficient (Kp) and % recovery data expressed as the average of triplicate experiments at 25°C. pH of the systems and protein samples was previously adjusted to 7.0. Load of alfalfa protein was 1.0 mg protein/g ATPS.

[b] Expressed as the summation of top and bottom phase recovery. Interface precipitation estimate to be (100-x)% of the loaded protein.

Table 3. Effect of NaCl addition on the % recovery and Kp of alfalfa proteins in two different ATPS[a].

System[b] (w/w)		NaCl content (w/w)				
		0%	**1.5%**	**3.0%**	**6.0%**	**9.0%**
TLL: 30% PEG 3350: 14.8% Phosphate: 10.3%	Top phase recovery (%)	8.1±0.9	0.8±0.6	2.9±0.9	5.6±3.6	4.6±2.7
	Bottom phase recovery (%)	50.2±0.4	39.4±0.1	34.4±1.9	15.3±1.1	1.4±0.5
	Kp	0.14±0.02	0.04±0.02	0.03±0.01	0.49±0.3	5.06±2.0
TLL: 32% PEG 3350: 15.7% Na_2SO_4: 8.9%	Top phase recovery (%)	13.4±2.4	6.2±1.3	1.5±0.9	0.0	1.2±0.2
	Bottom phase recovery (%)	16.0±2.4	15.2±1.3	35.9±0.9	12.1±1.9	8.2±2.8
	Kp	0.99±0.3	0.79±0.1	0.05±0.001	N.D. top[c]	0.25±0.1

[a] Partition coefficient (Kp) and % recovery data expressed as the average of triplicate experiments at 25°C. % Interphase precipitation estimated to be the complement to meet 100% of the loaded protein.

[b] pH of the systems and protein samples was previously adjusted to 7.0. Load of alfalfa protein was 1.0 mg protein/g ATPS.

[c] N.D. top: no protein detected in top phase.

Table 4. Effect of tie-line length[a] on the partition coefficient and total recovery of alfalfa proteins in ATPS.

ATP system[b]	PEG 3350 (w/w)	Salt (w/w)	TLL (%)	Kp[c]	% Total protein recovery[c]
PEG 3350-phosphate, short TLL	14.8%	10.3%	30	0.14±0.02	58.3±0.8
PEG 3350-phosphate, large TLL	16.9%	14.5%	42	4.1±3.5	23.3±1.2
PEG 3350-Na2SO4, short TLL	15.7%	8.9%	32	0.99±0.3	29.4±0.001
PEG 3350-Na2SO4, large TLL	17.9%	12.9%	49	0.12±0.03	31.6±1.5

[a] Tie-line length calculated from corresponding binodal curves and equations reported by Zaslavsky (1991).
[b] Systems formulated at pH 7 and 25°C. Sample load was 1.0 mg protein/g ATPS.
[c] Partition coefficient (Kp) and %Recovery expressed as result of triplicate experiments. Interface precipitation estimate to be (100-x)% of the loaded protein.

Table 5. Total protein spots detected from 2D gels after ATP partitioning of alfalfa protein extracts[a].

Molecular weight (kDa)	pI	Protein concentration (ppm)[b]		%Yield[c]
		Top phase	Bottom phase	
17.71	3	1.82	2.81	0.32*
54.74	3.1		25.40	1.78
57.68	3.15		2.43	0.17
57.96	3.2		45.96	3.22
22.83	3.2	0.37		0.03
18.43	3.27		0.32	0.02
13.74	3.34	0.28		0.02
25.02	3.6	1.85		0.13
25.12	3.81	3.52		0.25
13.36	3.95	2.61	9.11	0.82*
26.54	3.97	16.96	5.78	1.59*
16.36	4.05	0.11		0.01
16.89	4.11	2.85		0.20
29.1	4.18	1.31		0.09
34.66	4.23	10.86		0.76
18.76	4.3	2.69	3.02	0.40*
16.81	4.49	13.90		0.97
50.05	4.84	0.18	0.35	0.04*
16.59	4.98	33.15		2.32
49.88	5.03	0.92		0.06
36.52	5.08	4.40	2.89	0.51*
26.1	5.13	0.26	9.59	0.69*
34.2	5.29	12.42		0.87
26.08	5.3	1.22	8.75	0.70*
33.75	5.5	12.40		0.87
50.3	5.5	0.96		0.07
26.38	5.51	0.31		0.02
27.18	5.53	0.69	4.06	0.33*
33.79	5.54	4.52		0.32
33.36	5.58	2.24		0.16
13.96	5.61	1.79		0.13
18.09	5.61	0.10		0.01
72.4	5.62	0.09	0.54	0.04*
15.12	5.63	6.61	57.99	4.53*
33.46	5.64	4.61		0.32
49.9	5.7	9.50	166.93	12.36*

Table 5. ***(continuation)***

Molecular weight (kDa)	pI	Protein concentration (ppm)[b]		%Yield[c]
		Top phase	Bottom phase	
25.88	5.73	1.27	8.24	0.67*
18.44	5.74	0.35		0.02
14.86	5.84	6.89	227.82	16.45*
26.49	5.84	0.36	35.96	2.55*
46.3	5.87		5.23	0.37
50.26	5.93	0.82	96.16	6.80*
53.16	6.07		82.29	5.77
55.98	6.26		4.33	0.30
14.18	6.45	4.30	283.87	20.19*
63.38	6.96		3.39	0.24
25.53	6.98		27.32	1.91
54.52	6.98		14.13	0.99
55.57	7.38		19.87	1.39
58.43	7.54		3.99	0.28
14.61	7.55		39.78	2.79
25.49	7.71		35.65	2.50
55.36	7.93		5.70	0.40
54.67	9.75		5.49	0.38
58.2	10.0		12.47	0.87
	Total	**169.48**	**1257.62**	**100.0**

[a] All data are the average of two experiments using two-phase system: 14.8% PEG 3350, 10.3% potassium phosphate, pH 7 and 25°C. Load of protein was 1.0 mg/g ATPS.
[b] Protein concentration is expressed in parts per million considering the average protein loss during TCA steps for each phase.
[c] Proteins marked with * were selected for 3D characterization for being present at both phases.

Table 6. Three-dimensional properties and content of selected alfalfa proteins partitioned in ATPS.

ATPS[a]	M.W. (kDa)	pI	Log Kp	% subset of proeins[b]
	17.7	3.0	-0.189±0.1	*0.5*
PEG 3350 (14.8% w/w)	13.4	4.0	-0.527±0.2	1.2
Potassium phosphate	26.5	4.0	0.467±0.1	2.3
(10.3% w/w)	18.8	4.3	-0.055±0.1	0.6
	26.1	5.1	-1.562±0.01	1.0
TLL 30%	50.1	4.8	-0.298±0.1	0.1
Vr = 1.25	27.2	5.5	-0.766±0.01	0.5
	72.4	5.6	-0.766±0.01	0.1
Total count: 55 spots	36.5	5.1	0.181±0.1	0.7
17 spots contain 69%	15.1	5.6	-0.936±0.1	6.6
total protein.	26.1	5.3	-0.858±0.01	1.0
	49.9	5.7	-1.245±0.01	17.9
	14.9	5.8	-1.520±0.01	23.8
	25.9	5.7	-0.813±0.01	1.0
	50.3	5.9	-2.070±0.01	9.9
	26.5	5.8	-1.997±0.01	3.7
	14.2	6.5	-1.820±0.01	29.3

[a] All data are the average of duplicate experiments run at pH 7 and 25°C. Protein load was 1.0 mg/gr ATPS.
[b] Calculated from the summation of the individual masses of the spots included in this table.

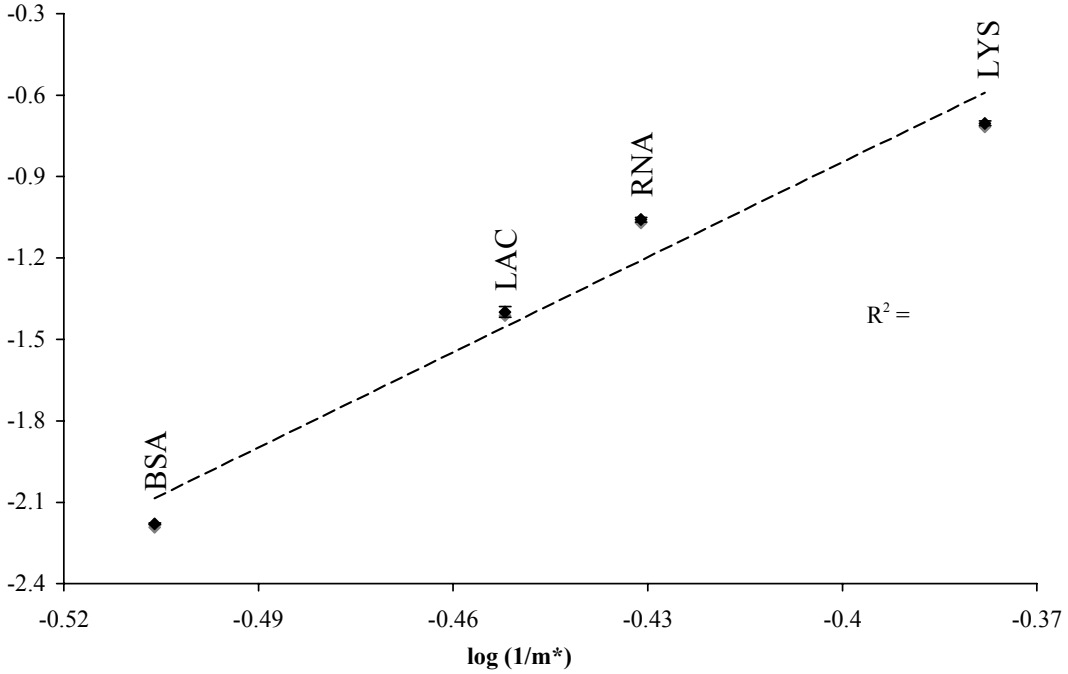
-0.3
-0.6
-0.9
-1.2
-1.5
-1.8
-2.1
-2.4
-0.52
-0.49
-0.46
-0.43
-0.4
-0.37
log (1/m*)
BSA
LAC
RNA
LYS
R2 =

Figure 1.

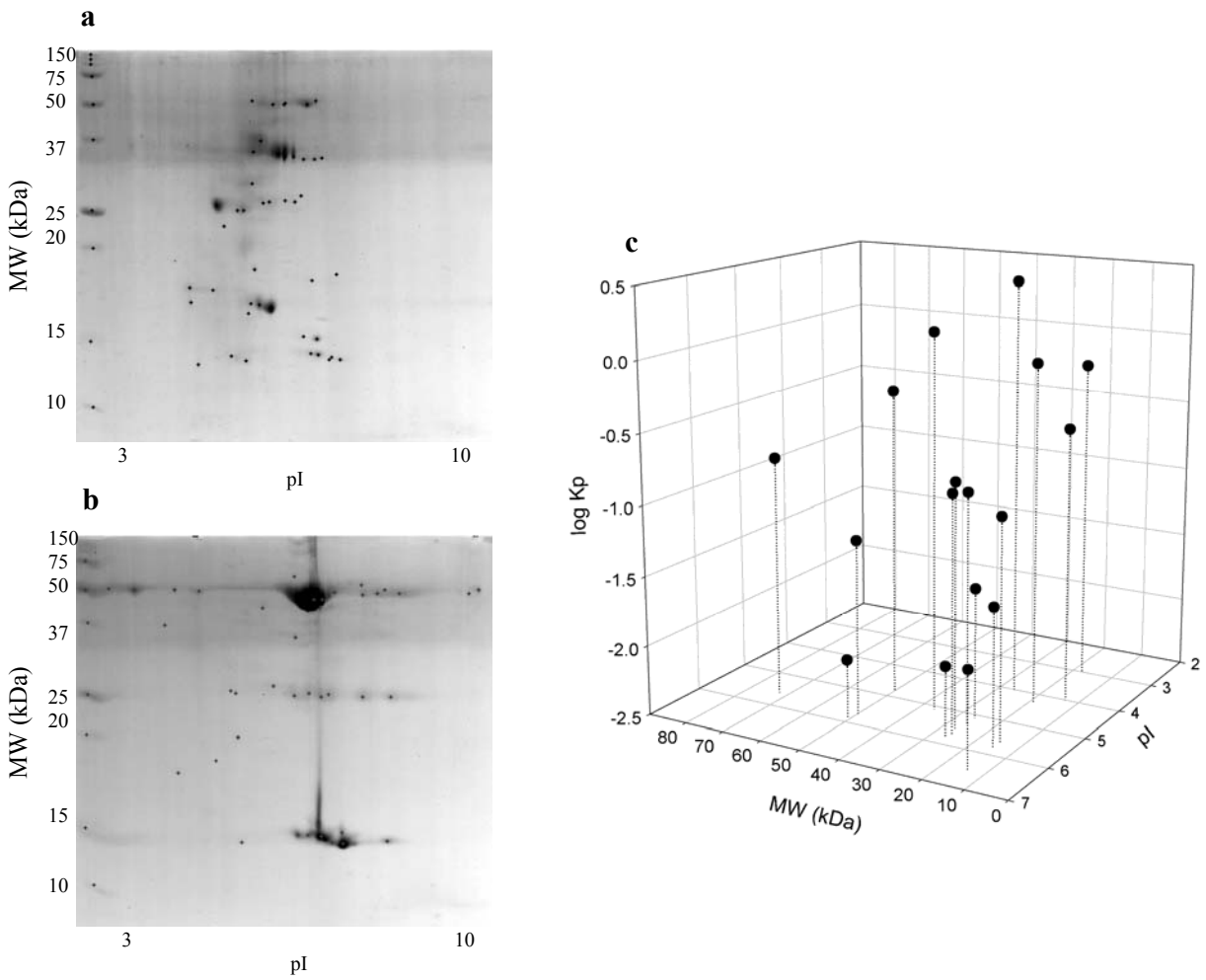
a
150
75
50
37
25
20
15
10
MW (kDa)
3
10
pI
b
150
75
50
37
25
20
15
10
MW (kDa)
3
10
pI
c
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
log Kp
80
70
60
50
40
30
20
10
0
MW (kDa)
2
3
4
5
6
7
pI

Figure 2

CHAPTER SIX – CONCLUSIONS AND SUGGESTIONS

6.1 CONCLUSIONS

Several common engineering strategies such as process integration and process intensification have been addressed to face the challenge of making more efficient the downstream processing for recombinant proteins having a direct impact on the final cost of the purified recombinant product. This dissertation addressed the use of a bioengineering strategy based on aqueous two-phase partitioning (ATPS) for the potential recovery of proteins from plants, and its integration with the proteomic tool of two-dimensional electrophoresis (2-DE) for multidimensional characterization of proteins from plant origin. This strategy is supported on the premise that a better knowledge of the contaminant proteins facilitate the design of downstream processes, given the molecular properties of the target protein.

First, the potential of ATPS-based processes as a primary recovery step was demonstrated for a recombinant protein from soybean protein extracts. The selected model protein, β-glucuronidase (GUS), contained in fractionated soybean extracts 7S and 11S was processed in PEG-phosphate ATPS. The target protein accumulated predominantly at the interphase, but showed top and bottom preference depending on the molecular weight of the PEG. The use of PEG 600 resulted in conditions at which GUS could be recovered from top phase or harvested from the interface with higher purification factors. The findings reported here represent a starting point in the way to establish a practical strategy for the recovery of recombinant proteins expressed in plants.

It was demonstrated the applicability of a three-dimensional technique for molecular characterization of soybean proteins providing valuable information on the molecular properties of these proteins. Given the case of purifying a recombinant protein from soybean extracts, this information will be useful for the design of downstream strategies, as these proteins will be considered as potential contaminants. The importance of selecting the phase-forming salt was clearly evident to define intrinsic hydrophobicity of the ATP system used for Kp measurement. It was demonstrated the influence of the type of salt on the profile and the number of proteins that can be characterized using hydrophobic ATPS.

The extension of the three-dimensional approach to protein extracts derived from alfalfa green tissues added a different perspective to the generic application of this experimental approach. It was evidenced the need of a previous partitioning study in order to customize the hydrophobic system according to the nature of the proteins to be partitioned. The alfalfa protein extract including human granulocyte-macrophage colony stimulating factor (rGM-CSF) as model recombinant protein was partitioned in PEG-phosphate system before 2-DE. The molecular properties obtained for a set of proteins allowed the potential identification of the target protein added to alfalfa extract, but also provided tools for the future design of pre-fractionation steps to eliminate main contaminants from alfalfa extracts.

The generic application of the three dimensional technique for protein characterization is subjected to find an hydrophobic ATP system where partitioning is ideally favoured to both phases and proteins remain soluble in order to obtain the most complete protein profile.

In the route to the predictive design of downstream strategies, the three-dimensional characterization of the selected plant proteins performed in this research, provided the tools for the generic application of this technique to different expression systems with the final purpose of facilitating the downstream processing of recombinant proteins from plants.

6.2 SUGGESTIONS FOR FUTURE WORK

From the results presented here, some suggestions for future work are described as follows. A good alternative that can be explored to solve the problem of low top phase partitioning of proteins would be the use of some protein solubilization agents such as polyvinylpirrolidone (PVP). The purpose of these experiments would be to diminish the amount of material that precipitate at the interphase, consequently more proteins could be characterized. It was also demonstrated the primary effect that the selection of the phase-forming salt has on the protein profiles obtained, a simple alternative to improve protein partition would be try different salts to diminish precipitation at the interphase in the case of alfalfa.

With the final purpose of increasing the Kp of proteins, another interesting possibility would be to explore the lower limit of PEG molecular weight that can be used in ATPS. Lowering the molecular weight of PEG could increase partitioning to the top phase but keeping an intrinsic hydrophobicity in ATPS large enough to still obtain an accurate estimation of hydrophobicity to fulfill the selection criteria for an ATP system for 3D characterization.

A good alternative to elucidate the problem One of the most time consuming steps on the whole process of running 2-DE is probably the gel analysis. It has been documented to be a

source of errors such as miss-matching or false spot identification caused by experimental variations between repetitions. A good alternative to solve this problem would be the use of fluorescent staining for protein detection. Fluorescent staining has the ability to detect picograms of protein, thus allowing detection of low abundant proteins. This could be useful in the case of recombinant proteins with low expression levels. 2-D Difference in-Gel Electrophoresis system (DIGE) is a methodology where up to three samples stained with different fluorescent Cy dyes can be run on the same gel, avoiding gel-to-gel variation during matching. Besides this advantage, an accurate comparison between protein profiles could be obtained for top and samples just by overlapping images from the same gel. As a result, large sets of protein spots could be discriminated based on the affinity for either top or bottom phase.

Regarding protein spot identification, a more detailed method is needed in order to confirm the chemical identity of the recombinant protein added to extracts. A western blot protocol would be probably the most adequate for rhGM-CSF identification. In the case of rubisco subunits or some other proteins like the soybean protein subunits, mass spectrometry would be the most adequate method if protein identification were needed.

In order to improve detection of low abundant proteins in the presence of dominant spots in a gel, a pre-fractionation step is suggested for removal of abundant proteins like rubisco. This strategy can be explored to enrich alfalfa extracts by reducing the amount of this abundant protein. It has been reported that this enzyme had a tendency to precipitate when PEG-phosphate ATPS were used with tobacco extracts. The same strategy could be used to

remove it from alfalfa extracts or deplete its concentration being careful of not affecting the target protein. The final purpose would be to increase the amount of proteins that can be detected in 2D gels and thus 3D-characterized to have a more complete protein profile.

REFERENCES

Abranches R, Sylvain M, Arcalis E, Altmann F, Fevereiro P, Stoger E. 2005. Plants as biorreactors: A comparative study suggests that *Medicago truncatula* is a promising production system. J Biotechnol 120:121-134.

Aguilar O, Albiter V, Serrano-Carreon L, Rito-Palomares M. 2006. Direct comparison between ion-exchange chromatography and aqueous two-phase processes for the partial purification of penicillin acylase produced by E. coli. J Chromatogr B 835:77-83.

Aguilar O, Rito-Palomares M, Glatz CE. 2007. Coupled application of aqueous two-phase systems and 2D-electrophoresis for three-dimensional characterization of soybean proteins (*In preparation*).

Aguilar O, Rito-Palomares M. 2007. Processing of soybeans (*Glycine max.*) extracts in aqueous two-phase systems as a first step for the potential recovery of recombinant proteins. J Chem Technol and Biotechnol DOI:10.1002 (*In Press*).

Albertsson PA. 1958. Partition of proteins in liquid polymer-polymer two-phase systems. Nature 182:709-711.

Andrews BA, Schmidt AS, Asenjo JA. 2005. Correlation for the partition behavior of proteins in aqueous two-phase systems: Effect of surface hydrophobicity and charge. Biotechnol Bioeng 90:380-390

Asenjo JA, Andrews BA. 2004. Is there a rational method to purify proteins? From expert systems to proteomics. J Mol Recognit 17:236-247.

Austin S, Bingham ET, Koegel RG, Matthews DE, Shahan MN, Strab RJ, Burgess RR. 1994. An overview of a feasibility study for the production of industrial enzymes in transgenic alfalfa. Ann NY Acad Sci 721:235-244.

Azevedo AM, Rosa PAJ, Ferreira IF, Aires-Barros MR. 2007. Optimisation of aqueous two-phase extraction of human antibodies. J Biotechnol 132:209-217.

Azzoni AR, Kusnadi AR, Miranda EA, Nikolov ZL. 2002. Recombinant aprotinin produced in transgenic corn seed: Extraction and purification studies. Biotechnol Bioeng 80: 268-276.

Bai Y, Nikolov ZL. 2001. Effect of processing on the recovery of recombinant b-glucuronidase (rGUS) from transgenic canola. Biotechnol Prog 17:168-174.

Benavides J, Mena JA, Cisneros-Ruiz M, Ramírez OT, Palomares LA, Rito-Palomares M. 2006. Rotavirus-like particles primary recovery from insect cells in aqueous two-phase systems. J Chromatogr B 842:48-57.

Benavides J, Rito-Palomares M. 2005. Potential aqueous two-phase processes for the primary recovery of colored proteins from microbial origin. Eng Life Sci 5:259-266.

Benavides J, Rito-Palomares M. 2006. Simplified two-stage method to B-phycoerythrin recovery from Porphyridium cruentum. J Chromatogr B 844:39-44.

Berggren K, Wolf A, Asenjo JA, Andrews BA, Tjerneld F. 2002. The surface exposed amino acid residues of monomeric proteins determine the partitioning in aqueous two-phase systems. Biochim Biophys Acta 1596:253-268.

Berkelman T, Stenstedt T. 2004. 2-D Electrophoresis. Principles and Methods. GE Healthcare. Buckinghamshire, UK.

Bjellqvist B, Ek K, Righetti PG, Gianazza E, Görg A, Westermeier R, Postel W. 1982. Isoelectric focusing in immobilized pH gradients: Principle, methodology and some applications. J Biochem Biophys Methods 6:317-339.

Bradford MM, 1976. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248-254.

Butt A, Davison MD, Smith GJ, Young JA, Gaskell SJ, Oliver SG, Beynon RJ. 2001. Chromatographic separations as a prelude to two-dimensional electrophoresis in proteomics analysis. Proteomics 1:42-53.

Cascone O, Andrews BA, Asenjo JA. 1991. Partitioning and purification of thaumatin in aqueous two-phase systems. Enzyme Microb Technol 13:629-635.

Chen S, Harmon AC. 2006. Advances in plant proteomics. Proteomics 6:5504-5516.

Daniell H, Streatfield SJ, Wycoff K. 2001. Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. Trends Plant Sci 6:219-226.

Evangelista RL, Kusnadi AR, Howard JA, Nikolov ZL. 1998. Process and economic evaluation of the extraction and purification of recombinant beta-glucuronidase from transgenic corn. Biotechnol Prog 14:607-614.

Feurtado JA, Banik M, Bewley JD. 2001. The cloning and characterization of a-galactosidase present during and following germination of tomato (Lycopersicon esculentum Mill.) seed. J Exp Bot 52:1239-1249.

Fischer R, Emans N, Schillberg S. 2002. The Production of Recombinant Antibodies in Plants and Plant Cells In: Erickson et al., editors. Molecular Farming of Plants and Animals for Human and Veterinary Medicine. The Netherlands. Kluwer Academic Publishers. pp. 259-285.

Fischer R, Emans N. 2000. Molecular farming of pharmaceutical proteins. Transgen Res 9:279-299.

Franco TT, Andrews AT, Asenjo JA. 1996. Use of chemically modified proteins to study the effect of a single protein property on partitioning in aqueous two-phase systems: Effect of surface hydrophobicity. Biotech Bioeng 49:300-308

Friedman M, Brandon DL. 2001. Nutritional and Health Benefits of Soy Proteins. J Agric Food Chem 49:1069-1086.

Garbis S, Lubec G, Fountoulakis M. 2005. Limitations of current proteomics technologies. J Chromatogr 1077:1-18.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM, editor. The Proteomics Protocols Handbook, Humana Press. pp. 571-607. http://www.expasy.org/tools/protparam.html.

Giddings G, Allison G, Brooks D, Carter A. 2000. Transgenic plants as factories for biopharmaceuticals. Nature Biotechnol 18:1151-1155.

Golubovic M, van Hateren SH, Ottens M, Witkamp G-J and van der Wielen LAM, Novel method for the production of pure glycinin from soybeans. J Agric Food Chem 53:5265-5269 (2005).

Görg A, Weiss W, Dunn MJ. 2004. Current two-dimensional electrophoresis technology for proteomics. Proteomics 4:3665-3685.

Gu Z, Glatz C.E. 2007 (a). Aqueous two-phase extraction for protein recovery from corn extracts. J Chromatogr B 845:38–50.

Gu Z, Glatz C.E. 2007 (b). A method for three-dimensional protein characterization and its application to a complex plant (corn) extract. Biotechnol Bioeng 97:1158-1169.

Hachem F, Andrews BA, Asenjo JA. 1996. Hydrophobic partitioning of proteins in aqueous two-phase systems. Enzyme Microb Tech 19:507-517.

He C, Li S, Liu H, Li K, Liu F. 2005. Extraction of testosterone and epitestosterone in human urine using aqueous two-phase systems of ionic liquid and salt. J Chromatogr A 1082:143-149.

Herbert BR, Harry JL, Packer NH, Gooley AA, Pedersen SK, Williams KL. 2001. What place for polyacrylamide in proteomics? Trends Biotechnol 19:3-9.

Hinchee MA, Connor-Ward D. October 20, 1998. Method for soybean transformation and regeneration. U.S. Patent Number 5824877.

Hoglund M. 1998. Glycosylated and non-glycosylated recombinant human granulocyte colony-stimulating factor (rhG-CSF)--what is the difference? Med Oncol 15:229-33.

Huddleston JG, Ottomar KW, Ngonyani DM, Lyddiatt A. 1991. Partitioning and purification of α-amylase in aqueous two-phase systems. Enzyme Microb Technol 16:131-142.

Imin N, Nizamidin M, Daniher D, Nolan KE, Rose RJ, Rolfe BG. 2005. Proteomic analysis of somatic embryogenesis in *Medicago truncatula*. Explant cultures grown under 6-benzylaminopurine and 1-naphtaleneacetic acid treatments. Plant Physiol 137:1250-1260.

Incamps A, Hély-Joly F, Chagvardieff P, Rambourg J-C, Dedieu A, Linares E, Quéméneur E. 2005. Industrial process proteomics: Alfalfa protein patterns during wet fractionation processing. Biotechnol Bioeng 91:447-459.

James EA, Wang C, Wang Z, Reeves R, Han Shin J, Magnuson NS, Lee JM. 2000. Production and characterization of biologically active human GM-CSF secreted by genetically modified plant cells. Prot Express Purif 19:131-138.

Krebbers E, Bosch D, Vandekerckhove J. 1992. Prospects and progress in the production of foreign proteins and peptides in transgenic plants. In: Shewry PR and Gutteridges S, editor. Plant protein engineering. London. Cambridge University Press. 315-325 pp.

Kusnadi A, Nikolov ZL, Howard JA. 1997. Production of recombinant proteins in transgenic plants: practical considerations. Biotechnol Bioeng 56:473-484.

Kwon T-H, Kim Y-S, Lee J-H, Yang M-S. 2003. Production and secretion of biologically active human granulocyte-macrophage colony stimulating factor in transgenic tomato suspension cultures. Biotechnol Lett 25:1571-1574.

Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685.

Lakemond CMM, De Jongh HHJ, Gruppen H, Voragen AGJ. 2002. Differences in denaturation of genetic variants of soy glycinin. J Agric Food Chem 50:4275-4281.

Larrick JW, Yu L, Naftzger C, Jaiswal S, Wycoff K. 2001 Production of secretory IgA antibodies in plants. Biomol Eng 18:87–94.

Larrik JW, Yu L, Naftzger C, Jaiswal S, Wycoff K. 2002. Human Pharmaceuticals Produced in Plants In: Hood EE, Howard JA editors. Plants as Factories for Protein Production. The Netherlands. Kluwer Academic Publishers. pp. 79-101.

Ma JK-C, Drake PMW, Christou P. 2003. The production of recombinant pharmaceutical proteins in plants. Nat Rev Genet 4:794–805.

Martinell BJ, Julson LS, Emler CA, Yong H, McCabe DE, Williams EJ. May 7, 2002.

Soybean *Agrobacterium* transformation method. U.S. Patent Application No. 6384301.

Menkhaus TJ, Bai Y, Zhang C, Nikolov ZL, Glatz CE. 2004. Considerations for the recovery of recombinant proteins from plants. Biotechnol Prog 20:1001-1014.

Menkhaus TJ, Glatz CE. 2005. Antibody Capture from Corn Endosperm Extracts by Packed Bed and Expanded Bed Adsorption. Biotechnol Prog 21:473-485.

Menkhaus TJ, Pate C, Krech A, Glatz CE. 2004. Recombinant protein purification from pea. Biotechnol Bioeng 86:108-114.

Natarajan SS, Xu C, Bae H, Caperna TJ, Garret WM. 2006. Characterization of storage proteins in wild (Glycine soja) and cultivated (Glycine max) soybean seeds using proteomic analysis. J Agric Food Chem 54:3114-3120.

Neupogen® Data Sheet. 2004. Amgen Inc. Thousand Oaks, CA.

O'Farrell PH. 1975. High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250: 4007-4021.

Ong S-E, Pandey A. 2001. An evaluation of the use of two-dimensional electrophoresis in proteomics. Biomol Eng 18:195-205.

Pandey A, Mann M. 2000. Proteomics to study genes and genomes. Nature. 405:837-846.

Pirondini A, Visioli G, Malcevschi A, Marmiroli N. 2006. A 2-D liquid-phase chromatography for proteomic analysis in plant tissues. J Chromatogr B 833:91-100.

Platis D, Labrou NE. 2006. Development of an aqueous two-phase partitioning system for fractionating therapeutic proteins from tobacco extract. J Chromatogr A 1128:114-124.

Rito-Palomares M, Dale C, Lyddiatt A. 2000. Generic application of an aqueous two-phase process for protein recovery from animal blood. Process Biochem 35: 665-673.

Rito-Palomares M, Hernandez M. 1998. Influence of system and process parameters on partitioning of cheese whey proteins in aqueous two-phase systems. J Chromatogr B, 711:81-90.

Rito-Palomares M, Lyddiatt A. 2000. Practical implementation of aqueous two-phase processes for protein recovery from yeast. J Chem Technol Biotechnol 75: 632-638.

Rito-Palomares M, Lyddiatt A. 2002. Process Integration using aqueous two-phase partition for the recovery of intracellular proteins. Chem Eng J 87:313-319.

Rito-Palomares M, Negrete A, Miranda L, Flores C, Galindo E, Serrano-Carreón L. 2001. The potential application of aqueous two-phase systems for in situ recovery of 6-pentyl--pyrone produced by *Trichoderma harzianum*. Enzyme Microb Technol 28:625-631.

Rito-Palomares M. 2004. Practical application of aqueous two-phase partition to process development for the recovery of biological products. J Chromatogr B 807:3-11.

Rose JKC, Bashir S, Giovannoni JJ, Jahn MM, Saravanan RS. 2004. Tackling the plant proteome: practical approaches, hurdles and experimental tools. Plant J 39:744-751.

Roy I, Mondal K, Gupta MN. 2007. Leveraging protein purification strategies in proteomics. J Chromatogr B 849:32-42.

Salgado JC, Rapaport I, Asenjo JA. 2005. Is it possible to predict the average surface hydrophobicity of a protein using only its amino acid composition? J Chromatogr A 1075:133-143.

Schillberg S, Fischer R, Emans N. 2003. Molecular farming of recombinant antibodies in plants. Cell Mol Life Sci 60:433-445.

Schindler J, Lewandrowski U, Sickmann A, Friauf E, Nothwang HG. 2006. Proteomic analysis of brain plasma membranes isolated by affinity two-phase partitioning. Mol Cel Prot 5:390–400.

Shyu DJH, Chou W-M, Yiu T-J, Lin CP, Tzen JTC. 2004. Cloning, functional expression, and characterization of cystatin in sesame seed. J Agric Food Chem 52:1350-1356.

Srere PA. 1969. Citrate Synthase. In: Packer L. editor. Methods in Enzymology Vol. 13. London. Academic Press pp. 3-11.

Stoger E, Sack M, Fischer R, Christou P. 2002. Plantibodies: applications, advantages and bottlenecks. Curr Opin Biotechnol 13:161–166.

Su C-K, Chiang BH. 2006. Partitioning and purification of lysozyme from chicken egg white using aqueous two-phase system. Process Biochem 41:257-263.

Thanh VH, Shibasaki K. 1976. Major proteins of soybean seeds. A straightforward fractionation and characterization. J Agric Food Chem 24:1117-1121.

Twyman RM, Stoger E, Schillberg S, Christou P, Fischer R. 2003. Molecular Farming in plants: host systems and expression technology. Trends in Biotechnol 21:570-578.

Whitelam GC, Cockburn B, Gandecha AR, Owen MRL. 2003. Heterologous protein production in transgenic plants. Biotechnol Gen Eng Rev 11:1-29.

Wu S, Murphy PA, Johnson LA, Fratzke MA, Reuber J. 1999. Pilot-plant fractionation of soybean glycinin and b-conglycinin. J Am Oil Chem Soc 76:285-293.

Zaslavsky A. 1995. Aqueous two-phase partitioning. Physical Chemistry and Bioanalytical Applications. New York; Marcel Dekker Inc. 660 p.

Zeitlin L, Olmsted SS, Moench TR, Co MS, Martinell BJ, Paradkar VM, Russell DR, Queen C, Cone RA, Whaley KJ. 1998. A humanized monoclonal antibody produced in

transgenic plants for immunoprotection of the vagina against genital herpes. Nature Biotechnol 16:1361-1364.

Zhang C, Glatz CE. 1999. Process engineering strategy for recombinant protein recovery from canola by cation exchange chromatography. Biotechnol Prog 15:12-18.

Zhang C, Love RT, Jilka JM, Glatz CE. 2001. Genetic engineering strategies for purification of recombinant proteins from canola by anion exchange chromatography: An example of b-glucuronidase. Biotechnol Prog 17:161-167.

http://www.LSBC.com

http://www.phytomedics.com

http://www.prodigene.com

Page left blank intentionally