Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Chromatography B, 833 (2006) 33–40
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

# Reducing protein concentration range of biological samples using solid-phase ligand libraries☆

Luc Guerrier[a], Vanitha Thulasiraman[a], Annalisa Castagna[c], Frederic Fortis[a], Shanhua Lin[a], Lee Lomas[a], Pier Giorgio Righetti[b], Egisto Boschetti[a,*]

[a] *Ciphergen Biosystems Inc., Fremont, CA 94555, USA*
[b] *Department of Chemistry, Materials and Chemical Engineering – Giulio Natta, Politechnico di Milano, 20131 Milano, Italy*
[c] *University of Verona, Department of Clinical and Experimental Medicine, 37134 Verona, Italy*

Received 26 July 2005; accepted 18 December 2005
Available online 7 February 2006

**Abstract**

The discovery of specific polypeptides of diagnostic relevance from a biological liquid is complicated by the overall vast number and the large concentration range of all polypeptides/proteins in the sample. Depletion or fractionation methodologies have been used for selectively removing abundant proteins; however, they failed to significantly enrich trace proteins. Here we expand upon a new method that allows the reduction of the protein concentration range within a complex mixture, like neat serum, through the simultaneous dilution of high abundance proteins and the concentration of low abundance ones in a single, simple step. This methodology utilizes solid-phase ligand libraries of large diversity. With a controlled sample-to-ligand ratio it is possible to modulate the relative concentration of proteins such that a large number of peptides or proteins that are normally not detectable by classical analytical methods become, easily detectable. Application of this method for reducing the dynamic range of unfractionated serum is specifically described along with treatment of other biological extracts. Analytical surface enhanced laser desorption/ionization mass spectrometry (SELDI-MS) technology and mono- and two-dimensional electrophoresis (1-DE and 2-DE) demonstrate the increase in the number of proteins detected. Examples linking this approach with additional fractionation methods demonstrate a further increase in the number of detectable species using either the so-called "top down" or "bottom up" approaches for proteomics analysis. By enabling the detection of a greater proportion of polypeptides/proteins within a sample, this method may contribute significantly towards the discovery of new biomarkers of diagnostic relevance.


*Keywords:* Proteomics; Ligand library; Selective capture; Fractionation; Mass spectrometry; Electrophoresis; Dynamic range

## 1. Introduction

In proteomic investigations, major obstacles to resolve are around the discovery of specific, disease-related peptide/protein species that are present in trace amounts among a large background of non-relevant and/or abundant proteins. The situation is particularly complex in higher eukaryote organisms due to the large number of genes and gene splice-variants that encode proteins, as well as the number and extent of post-translational modifications (PTM) such as cleavage, phosphorylation, glycosylation, lipidation, etc. that can impart unique functions on a particular gene product, depending on the nature of the PTM. To this complexity one has to further consider that proteins within the sample can be present over a large concentration dynamic range. For example, in human serum it is estimated that the dynamic range of protein concentration is in excess of 10 orders of magnitude [1], and that the 50 most abundant proteins represent about 99% of the total amount of protein mass but only less than 0.1% in number [2]. This situation renders the discovery of peptides/proteins of diagnostic or therapeutic importance challenging; as a consequence, sample preparation strategies must be specifically conceived and/or optimized to complement the chosen method of detection.

☆ This paper was presented at the 2nd IPSo Congress on Proteomics and Genomics, Viterbo, Italy, 29 May to 1 June 2005.

* Correspondence to: Ciphergen Biosystems Inc., R&D, 48 Avenue des Genottes, 95800 Cergy Pontoise, France. Tel.: +33 134207806; fax: +33 134207878.
*E-mail address:* eboschetti@ciphergen.com (E. Boschetti).

1570-0232/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2005.12.048

Common methods used for proteome analysis include mono- and two-dimensional electrophoresis (1-DE, 2-DE), and mass spectrometry (MS); however, these methods have limitations due to the complex composition of samples [3]. In electrophoresis, for example, the protein bands/spot corresponding to the high abundance proteins can exhibit deformation and/or smearing during the electrophoretic separation due to protein overloading and thus obscure the detection of other proteins of similar mass and/or p*I*. Additionally, protein species below about 5 kDa or at the extreme p*I* ranges of the selected 2-DE gels are either lost during the separation or not resolved, and low abundance proteins of any mass and suitable p*I* range may not be visible because they fall below the sensitivity of the staining method. MS, both electrospray and laser desorption methods, can also suffer in analytical performance due to the nature and complexity of biological samples. To overcome many of these drawbacks, sample treatment strategies have been developed, ranging from simple clean-up methods to more complicated strategies of high-load 2-DE [4,5], prefractionation processes [6–9], abundant protein depletion [10–12], global digestion followed by MS analysis [13] and multidimensional chromatography followed by gel electrophoresis or MS analysis [14]. However, depending on the method used, specific cautions must be considered with respect to limited resin capacity, loss of low-abundance proteins during high-abundance protein depletion [15] or protein precipitation during sample treatment, inefficiencies in sample digestion and protein losses due to choices made in chromatographic separation modes.

In spite of these efforts, the ability to detect low abundant species still remains a critical challenge in deciphering complex proteomes and correlating proteome changes with metabolic events for diagnostic and therapeutic purposes. Recently the principle of a novel sample preparation approach that can decrease the protein concentration dynamic range without depletion has been described as it applies to a variety of proteomes [16,17]. This method is based on the selective adsorption of proteins on a solid phase combinatorial ligand library under capacity-limited binding conditions. In this paper the principle is briefly described, with additional emphasis on physicochemical parameters that are optimal for the detection of the maximum number of proteins from very complex mixtures.

## 2. Basis of the use of ligand libraries for the reduction of protein concentration differences

Solid phase affinity adsorption is a well-known chromatographic process for selectively capturing and concentrating a given protein. Its intrinsic limitation is the binding capacity of the sorbent; when the saturation is reached the excess of the protein in question cannot bind and is subsequently discarded in the flowthrough. Starting from this simple mechanism one can extend the phenomenon to a large number of different affinity ligands for a large number of different proteins. If such a diversity of affinity ligands is mixed together to form an affinity-ligand pool, and contacted with a diverse protein mixture, each unique affinity ligand beads within the pool will bind and concentrate its specific protein up to the point of ligand saturation and independently of all other affinity ligands and proteins. When the relative concentration from each species within the protein mixture forms a large dynamic range such that the high abundance proteins exceed the capacity and the low abundance proteins are below the capacity of their respective specific affinity ligand, the high abundance proteins will rapidly saturate their corresponding beads while low abundance ones will continue to adsorb as long as the sample is available. After removal of all proteins that are not bound, the composition of proteins retained by the beads will be defined by the presence of their specific affinity ligands, and the relative concentration of each retained protein species will be defined by the capacity and saturation degree of each of the affinity ligands and the relative starting concentration of each protein species.

This principle has been described using solid-phase peptide ligand libraries [16,17]. The library is generated using classical combinatorial synthesis methods, and is capable of producing tremendous ligand diversity where theoretically there is a ligand for every peptide and protein present in the starting material. For example, if the combinatorial synthesis for generating hexapeptide ligands is made using 20 amino acids, the total amount of ligands obtained is theoretically of 64 million, a number much larger than the expected number of different protein in biological samples. The use of such a highly diverse combinatorial library of affinity ligands under the described capacity limiting conditions results in a compression of the dynamic range of protein concentration (dilution of high-abundance proteins and concentration of low abundance proteins), while retaining representatives of all proteins within the mixture. Retained proteins can then be eluted in bulk or selectively from the affinity library using buffer modifiers such as ionic strength, pH, chaotropic agents or organic solvents with subsequent analysis by any number of analytical methods.

## 3. Materials and methods

### 3.1. Chemicals and biologicals

The solid-phase combinatorial hexapeptide library (Protein Equalizer™ beads) was supplied by Ciphergen Biosystems Inc., Fremont, CA; it was made using a previously described "split, couple and recombine" method [18,19]. By incorporating 20 different amino acids in the synthesis, the theoretical number of different ligand structures was $20^6$ or 64 million. Each bead of 65 μm average diameter carried about 50 pmol of hexapeptide. Urea, thiourea, tributylphosphine (TBP), glycine, sodium and lithium dodecyl sulfate (SDS and LDS), and 3-[3-cholamidopropyl dimethylammonio]-1-propansulfonate (CHAPS) were obtained from Fluka Chemie (Buchs, Switzerland). Ethanol, methanol, glycerol, sodium hydroxide, hydrochloric acid, acetone, and acetic acid were from Merck (Darmstadt, Germany). Bromophenol blue, agarose and carrier ampholytes (Pharmalyte) were from Pharmacia-LKB (Uppsala, Sweden). Linear Immobiline dry strips (pH gradient 3–10, 7 cm long) were from Bio-Rad Laboratories (Hercules, CA, USA). Protein molecular weight standards as well as frozen human serum were from Sigma Chemicals, St. Louis, MO.