Available online at www.sciencedirect.com
ScienceDirect
Journal of Chromatography A, 1153 (2007) 277–290
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Review

Romancing the "hidden proteome", *Anno Domini* two zero zero seven

Egisto Boschetti [a], Lee Lomas [a], Attilio Citterio [b], Pier Giorgio Righetti [b,*]

[a] *Ciphergen Biosystems, Fremont, California, USA*
[b] *Department of Chemistry, Materials and Chemical Engineering "Giulio Natta", Politecnico di Milano, Via Mancinelli 7, Milan 20131, Italy*

Available online 8 February 2007

Abstract

The mechanism of action and properties of a solid-phase ligand library made of hexapeptides, for capturing the "hidden proteome", i.e. the low- and very low-abundance proteins constituting the vast majority of species in any proteome, be it a cell or tissue lysate or a biological fluid, are here reviewed. Mechanisms of adsorption are evaluated, as well as different protocols for *en bloc* or sequential elution of the captured polypeptides. Examples are given of capture of proteins from serum, human platelet extracts, bacterial extract and egg white. The increment in detection of low-abundance species appears to be of at least four-fold as compared with untreated samples. One particular aspect of this capture is the adsorption of a high proportion of small peptides (in the Mr 600–8000 Da range) that are normally lost upon electrophoretic two-dimensional mapping. Such a peptide population, in human sera, may be of particular importance since it may contain protein cleavage products of diagnostic value.

Keywords: Ligand library; Peptide ligands; SELDI-mass spectrometry; Low-abundance proteome

Contents

1. Introduction

We do not know if this review on sample treatment prior to two-dimensional (2D) map analysis will merit the hall of fame like its, undoubtedly more famous, counterpart "Romancing the Stone", an adventure movie by Robert Zemeckis that made headlines in *Anno Domini* one nine eight four [1], but certainly it could offer an unexpected turn in proteome analysis, where scientists around the world have realized that not much progress has been made in the last few years and that the "hidden proteome" still remains a buried treasure to be mined.

The major problem is that, in most proteomes, as elegantly demonstrated by the Anderson's group [2–4], the dynamic concentration range could span some ten orders of magnitude, rendering thus the detection of the "rare" proteome (believed to comprise >50% of all proteins present in any cell lysate and/or biological fluid) highly problematic, considering that, e.g., 2D maps cannot span more than four orders of magnitude in protein

* Corresponding author.
E-mail address: piergiorgio.righetti@polimi.it (P.G. Righetti).

0021-9673/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2007.01.136

revelation. High-abundance proteins are noxious in more than one way e.g., in 2D maps, they obliterate the signals of the rare ones, which are often buried under the very large spots (and smears) of the abundant species. Additionally, when coupled to mass spectrometry (MS) analysis for the purpose of peptide mass fingerprinting, the high ion signal generated from their cleaved peptides masks the signal of the rare ones. One way to ameliorate such signals would be pre-fractionation, by which high-abundance species would be eluted in selected fractions in a chromatographic column (or captured by some affinity principles) or confined to some compartment in electrokinetic separation devices. A host of pre-fractionation protocols have been reported over the years, as reviewed in a number of articles [5–10], encompassing just about all possible chromatographic and electrophoretic techniques. In addition, immuno-depletion has been extensively used for subtracting the most abundant proteins, especially in biological fluids [3], either with IgGs or with IgYs (avian antibodies) [11]. Here, however, a number of problems have been outlined, particularly due to co-depletion, which might subtract thousands of minor species together with the target protein to be eliminated [12–15].

The approach here presented, the concept of "Equalizer Technology", might turn out to be an important step for bringing to the limelight the "hidden proteome" and discovering several new biomarkers of importance in routine clinical chemistry analysis. The basic article, outlining the synthesis of the "Equalizer beads" (or solid-phase combinatorial peptide ligand library) and some of their fundamental properties, has recently been published [16], together with reviews describing the very basic concepts [10,17,18]. Some applications, notably on urinary [19] and serum [20] protein analysis, as well as on tracking and "amplifying" impurities on r-DNA products [21,22] are available.

The combinatorial peptide bead library should not be confused with the technique of DNA and RNA aptamers. Aptamers are also used to bind specifically to target proteins with an affinity similar to that of antibodies [23]. But they typically are tailored to bind to one specific protein at a time and thus they are based on an exclusive or selective principle, similar to the immunoaffinity capturing protocols as discussed above. Combinatorial peptide ligand library beads are not meant to select a single protein like antibodies or protein families like lectins or to capture some specific components, but rather to embrace all proteins in a proteome and as such they are ecumenical (or at least they try to), i.e. they accept and adopt all faiths, colors, races and credos.

2. General properties and behavior of "Equalizer® beads"

Combinatorial peptide ligand library beads comprise a solid phase library of hexapeptides that are synthesized via a short spacer on a poly-(hydroxymethacrylate) substrate, according to a modified Merrifield approach [24], by using the split, couple and recombine method [25–27]. The ligands are represented throughout the beads porous structure and can achieve a ligand density of ca. 40–60 μmoles/mL of bead volume (average bead diameter of about 60 μm) of the same hexapeptide distributed throughout the core of the pearl. Each single bead, thus, has millions of copies of a single, unique ligand structure and each bead, potentially, has a different ligand from every other bead. Depending on the number of amino acids used, a hexapeptide library contains a population of dozen of million of different ligands (e.g., 11 millions for 15 different amino acids or 24 millions for 17 different amino acids or even 64 millions if the number of amino acids is 20).

The preparation of a peptide combinatorial library on a solid phase is made according to a relatively simple process. Briefly a batch of millions of microscopic porous chromatographic beads is split into several equal parts each representing a sub-reaction batch. The number of sub-reaction batches is the same as the number of building block used for the production of the ligand sequences (e.g., amino acids in the present case). Each bead sub-reaction batch receives a different building block which is chemically attached on the beads. The different sub-batches are mixed together, extensively washed and the entire batch split again into the same number of sub-reaction batches as before. The second building blocks are then attached at the extremity of the first according to the same scheme as above and the cycle continues until reaching the desired length of the combinatorial ligand (this process is described *in extenso* in Lam et al. [28]).

The vastly heterogeneous population of baits means that, in principle, an appropriate volume of beads could contain a partner able to interact with just about any protein present in a complex proteome (be it a biological fluid or a tissue or cell lysate of any origin). Lengthening the bait to a heptamer would generate many more diverse ligands, probably considerably more numerous than all the diverse proteins synthesized by all known living organisms.

How the "Equalization" process works is illustrated in Fig. 1. If one hypothesizes to have as many different solid-phase affinity columns as the proteins contained within a biological extract, each affinity column would capture independently its protein partner. If each individual affinity solid phase would work under the same physicochemical conditions, a mixed bed of all of them can be used instead of individual numerous batches or columns. In the present case, the affinity column is represented by a single bead, all beads being used all together with the same buffer. When a complex protein extract such as serum is exposed to such a ligand library, in large overloading conditions each bead with affinity to an abundant protein will rapidly become saturated and the vast majority of the same protein will remain unbound. In contrast, trace proteins will not be able to saturate all the beads that have an affinity for them unless the volume of the biological sample is very large. Thus, on the basis of the saturation-overloading chromatographic principle, a solid phase ligand library enriches for trace proteins while concomitantly reducing the concentration of abundant species. This principle is true if dissociation constants are compatible with the concentration of very dilute proteins and if the variety of ligands covers at least all necessary affinity interactions to capture all proteins. With a hexapeptide library the diversity of structures is so large (several dozen of millions) that the probability to form an affinity complex with each protein is relatively high. It is to be noted however that a given protein can adsorb on more than one