Analytical Biochemistry 516 (2017) 23–36

Contents lists available at ScienceDirect

Analytical Biochemistry

journal homepage: www.elsevier.com/locate/yabio

Assessment of sample preparation methods for metaproteomics of extracellular proteins

Jutta Speda [a, 1], Mikaela A. Johansson [a, 1], Uno Carlsson [a], Martin Karlsson [a, b, *]

[a] Molecular Biotechnology, Dept. of Physics, Chemistry and Biology, Linköping University, SE-581 83 Linköping, Sweden
[b] InZymes Biotech AB, Gjuterigatan 1B, SE-582 73 Linköping, Sweden

ARTICLE INFO

Article history:
Received 16 August 2016
Received in revised form
6 October 2016
Accepted 8 October 2016
Available online 11 October 2016

Keywords:
Enzyme discovery
Microbial community
Metaproteome
Extracellular
Sample preparation
2D gel electrophoresis

ABSTRACT

Enzyme discovery in individual strains of microorganisms is compromised by the limitations of pure culturing. In principle, metaproteomics allows for fractionation and study of different parts of the protein complement but has hitherto mainly been used to identify intracellular proteins. However, the extracellular environment is also expected to comprise a wealth of information regarding important proteins. An absolute requirement for metaproteomic studies of protein expression, and irrespective of downstream methods for analysis, is that sample preparation methods provide clean, concentrated and representative samples of the protein complement. A battery of methods for concentration, extraction, precipitation and resolubilization of proteins in the extracellular environment of a constructed microbial community was assessed by means of 2D gel electrophoresis and image analysis to elucidate whether it is possible to make the extracellular protein complement available for metaproteomic analysis. Most methods failed to provide pure samples and therefore negatively influenced protein gel migration and gel background clarity. However, one direct precipitation method (TCA-DOC/acetone) and one extraction/precipitation method (phenol/methanol) provided complementary high quality 2D gels that allowed for high spot detection ability and thereby also spot detection of less abundant extracellular proteins.

1. Introduction

Enzymes are of great importance and represent a present and future potential as biocatalysts in sustainable industrial biotechnology [1–3]. Therefore, the identification of novel enzymes to improve economics of industrial biotechnology processes and to expand the use of enzymes into new areas is a central goal. Most industrially relevant enzymes to date have been isolated from pure cultured microorganisms. However, it is well established that only a small fraction of all existing microorganisms can be obtained in pure cultures. Therefore, the potential of finding novel enzymes by pure culturing is limited [4,5]. Furthermore, this also means that the industrial enzymes available today are very much biased to origin from microorganisms that simply have been obtainable in pure cultures, and not necessarily from those microorganisms that produce the best enzymes. The possibility to identify enzymes directly from more diverse microbial communities would therefore potentially give access to a vast number of novel enzyme candidates. In particular, studies of unconventional habitats hosting extremophiles to identify enzymes that are stable under different harsh conditions will give rise to new strategic opportunities of proteomic research [6–10]. This is especially so if unconventional habitats hosting extremophiles could be studied for identifying enzymes that are stable under different harsh conditions [9,10]. The highly diverse flora of microorganisms populating anaerobic environments is of particular interest, since most extreme environments are anoxic [11,12].

For this purpose metaproteomics, or “the large-scale characterization of the entire protein complement of environmental microbiota at a given point in time” [13] should be of great potential, especially if combined with metagenomics [14]. However, metaproteomics, similar to metagenomics, is largely applied for non-targeted studies and has hitherto mainly been used to understand microbial ecology and biological processes of microbial communities [15–17], focusing almost exclusively on the intracellular fraction of proteins [18,19]. In industrial biotechnology hydrolytic enzymes have the largest use, most of which are derived

* Corresponding author. Molecular Biotechnology, Dept. of Physics, Chemistry and Biology, Linköping University, SE-581 83 Linköping, Sweden.
E-mail address: marka@ifm.liu.se (M. Karlsson).
[1] These authors contributed equally.

http://dx.doi.org/10.1016/j.ab.2016.10.008
0003-2697/

from actively exported enzymes used by the microorganisms to obtain nutrients from the environment. Furthermore, extracellular free enzymes have potentially a larger industrial interest than intracellular enzymes, as they are adapted to be stable independently of the cell and are therefore often more stable than intracellular enzymes [20]. Thus, the investigation of extracellular proteins in diverse microbial communities from various environmental conditions should enable the discovery of hydrolytic enzymes which would be of outstanding interest for industrial biocatalysis [21]. In addition, it would appear essential to analyze the extracellular protein complement to better understand other important biological processes such as chemotaxis, biofilm formation, transport proteins, cell-cell and cell-host interactions. These important aspects of the extracellular environment of microbial communities and the wealth of important information therein are often emphasized in the literature [22–25]. However, and considering that 25–35% and 15–25% of the proteins of gram negative and gram positive bacteria, respectively, are estimated to be secreted [26], surprisingly little work has been addressed to make this information available.

Shotgun 2D-LC/MS proteomics is now the dominating technique as a result of the increasing availability of genomic sequence data and development of 2D liquid chromatography and advanced mass-spectrometry. However, the nature of metaproteomics means that the proteomic information come from microorganisms with largely unsequenced genomes. For these types of samples 2D gel-based methods still offer several advantages over gel free shotgun approaches, such as the possibility for *de novo* sequencing [27] of individual intact proteins identified by quantitation of different expression levels in two or more biologically relevant states [28]. In respect to bioprospecting for extracellular hydrolytic enzymes this is an important capability of gel based metaproteomics since it has been shown that microorganisms in communities react to external stimuli such as starvation or the limitation of certain nutrients with an increased production of hydrolytic enzymes [29,30]. These characteristics can therefore, potentially, be used to identify various enzyme activities whose expression is regulated in response to external stimuli [31]. In a recent study this attribute was exploited to show that it is possible to specifically suppress and independently induce cellulase and protease activity in a methanogenic microbial community in a constructed environment. This led to a manifold increase in the extracellular activity of the respective induced enzymes as compared to that of the pre-induced state [32]. It was also found that the induced cellulases have a very strong positive effect on the degradation rate of lignocellulosic material in biochemical methane potential (BMP) tests (unpublished data) and that proteases stable in the methanogenic environment can have a practical application in boosting biogas production of sludge [33]. Thus, this extracellular milieu of the induced microbial methanogenic community should contain induced enzymes of scientific and practical value. These enzymes could possibly be identified by comparing the distinct, but limited, differences in the metaproteomes of the induced and pre-induced state. By this differential metaproteomics approach, 2D gel electrophoresis could be especially suitable as the analysis method and the subsequent protein identification can be limited to only those extracellular proteins that are significantly up-regulated against a background of constitutively expressed proteins. This is, thus, in contrast to when metaproteomic studies are used to understand the overall microbial ecology and physiology, for which shotgun metaproteomics by 2D-LC/MS of the entire captured protein complement could be more informative.

A prerequisite for any metaproteomic analysis is that the sample preparation methods capture the widest array of proteins, since a protein missed in extraction will never be identified irrespective of selected down-stream analyses. However, the analysis depth of metaproteomics is low and it is expected that << 1% of the entire metaproteome of complex microbial communities can be resolved [28]. This is in large due to the limitations and biases of the often employed protein extraction methods [34]. Furthermore, differential metaproteomics requires that the true difference in protein expression levels can be identified over time, or against a reference. It is known that the sample preparation step is the most common source of variability [25], which further emphasizes the need for reliable and reproducible sample preparation methods. In addition, the sample preparation methods need to provide samples that are compatible with 2D gels and/or 2D-LC/MS. For sample preparation of intracellular proteins from microbial communities in natural environments these are all well recognized and addressed limitations, whereas information regarding sample preparation of extracellular proteins for metaprotemic studies is scarce.

The term metaproteomics further implies that the investigated proteins emanate from a mixed microbial community, in which different species can be expected to be represented to different degree. However, it cannot be taken for granted that the protein expression of the most abundant species differs the most between two different conditions, or produces the most interesting enzymes. It is, therefore, important that also those proteins that are produced in low concentrations by microorganisms that are less abundant in the microbial community can be detected. Detection of low abundance proteins in the large dynamic range of biological samples is a well-known limitation of e.g. gel-based proteomics since these can be masked by the presence of highly abundant proteins [35]. It is, therefore, important that the sample preparation method does not preferentially extract certain subsets of proteins and thereby produces an artificial effect of high abundance proteins. For detection of low abundance proteins by e.g. 2D gel electrophoresis it is not enough to capture all proteins, preferentially in their correct ratio. It is also important that the preparation method provides samples that are clean enough to produce high quality gels with low background, which could otherwise also mask low abundance proteins.

Thus, the investigation of the complete extracellular metaproteome is complex and will present problems in sample preparation that are different from those encountered in metaproteomic studies of the intracellular fraction. One obvious challenge to overcome in sample preparation of extracellular proteins is the expected low protein concentration as compared to the analysis of the intracellular metaproteome. When analyzing intracellular proteins, the protein concentration in the sample is increased due to centrifugation and cell pelleting in the cell washing steps before cell lysis. For analysis of the extracellular fraction the supernatant, on the contrary, needs to be collected after the centrifugation step. Thus, the first simple concentration step by cell pelleting is not applicable when working with the extracellular fraction and other sample concentration techniques need to be applied.

Another obvious complication is that the proteins in the extracellular environment will be mixed with the contaminants that are avoided in the analysis of the intracellular metaproteome by the cell washing steps. For the latter, several protocols have been assessed and published for gel based analyses [36–38], whereas the subject of sample preparation for the extracellular metaproteome has only been touched upon [39]. Samples collected from natural environments contain numerous contaminants which can hamper either or both the protein migration in the IEF and SDS-PAGE steps of 2D gel separation and proteomic image analysis. Contaminants can be interfering metabolites such as fatty acids in anaerobic populations or humic acids in soil samples. Hence, it is important that these contaminants interfering with the proteomic methods are removed in order to enable a reliable metaproteomic analysis.