Biochimica et Biophysica Acta 1864 (2016) 945–951

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Diagonal chromatography to study plant protein modifications☆

Alan Walton [a,b,c,d], Liana Tsiatsiani [e,f], Silke Jacques [a,b,c,d], Elisabeth Stes [a,b,c,d], Joris Messens [g,h,i], Frank Van Breusegem [c,d], Sofie Goormachtig [c,d], Kris Gevaert [a,b,*]

[a] Department of Medical Protein Research, VIB, 9000 Ghent, Belgium
[b] Department of Biochemistry, Ghent University, 9000 Ghent, Belgium
[c] Department of Plant Systems Biology, VIB, 9052 Ghent, Belgium
[d] Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium
[e] Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Padualaan 8, 3584 CH Utrecht, the Netherlands
[f] Netherlands Proteomics Centre, Padualaan 8, 3584 CH Utrecht, the Netherlands
[g] Structural Biology Research Center, VIB, 1050 Brussels, Belgium
[h] Brussels Center for Redox Biology, 1050 Brussels, Belgium
[i] Structural Biology Brussels, Vrije Universiteit Brussel, 1050 Brussels, Belgium

ARTICLE INFO

Article history:
Received 22 October 2015
Received in revised form 1 December 2015
Accepted 4 January 2016
Available online 7 January 2016

Keywords:
Diagonal chromatography
COFRADIC
ChaFRADIC
Protease
Protein processing
Protein ubiquitination
Methionine sulfoxide
Methionine oxidation

ABSTRACT

An interesting asset of diagonal chromatography, which we have introduced for contemporary proteome research, is its high versatility concerning proteomic applications. Indeed, the peptide modification or sorting step that is required between consecutive peptide separations can easily be altered and thereby allows for the enrichment of specific, though different types of peptides. Here, we focus on the application of diagonal chromatography for the study of modifications of plant proteins. In particular, we show how diagonal chromatography allows for studying proteins processed by proteases, protein ubiquitination, and the oxidation of protein-bound methionines. We discuss the actual sorting steps needed for each of these applications and the obtained results. This article is part of a Special Issue entitled: Plant Proteomics— a bridge between fundamental processes and crop production, edited by Dr. Hans-Peter Mock.

1. A brief history of contemporary proteomics

In data-dependent acquisition mode, mass spectrometers are currently capable of generating fragmentation data for up to a few tens of peptide ions per second (see for example Ref. [1]). In sharp contrast, less than a decade ago, the hybrid quadrupole time-of-flight instrument in the Gevaert lab needed one second to generate a fragmentation spectrum of a single peptide [2], which was close to the limit of obtaining high quality spectra. This illustrates how peptide mass spectrometry has evolved over a short period of time and points to the limits in peptide sequencing speed that mass spectrometry-driven proteomics faced in the early days. As a result, for older mass spectrometers, it is impossible to identify all peptides derived from the tryptic digest of a full proteome. Indeed, we can now expect expression of about 10,000 different proteins in human cell lines, as shown a couple of years ago in papers from the Mann and Aebersold labs [3,4]. And, we previously in silico calculated that, upon trypsin digestion of the human proteome and allowing for a maximum of one missed cleavage event, a human protein will, on average, yield 20 peptides that are identifiable by mass spectrometry as these peptides had a mass between 600 and 4000 Da, resembling the mass range of MS-identified tryptic peptides [5]. Thus, without taking into consideration protein modifications or different protein forms from alternative splicing events, a given human cell line would then theoretically yield 200,000 different tryptic peptides that need to be identified by mass spectrometry to cover as many aspects of proteins (e.g. modifications) as possible. This figure of 200,000 peptides clearly is an underestimate, but is already far beyond the reach of older instruments.

Abbreviations: Boc-GLy, *N*-(tert-butoxycarbonyl)glycine; ChaFRADIC, charge-based fractional diagonal chromatography; COFRADIC, combined fractional diagonal chromatography; GST, glutathione S-transferase; ICAT, isotope-coded affinity tag; Msr, methionine sulfoxide reductases; MudPIT, multidimensional protein identification technology; PICS, proteomic identification of protease cleavage sites; ROS, reactive oxygen species; TAILS, terminal amine isotopic labelling of substrates; TNBS, 2,4,6-trinitrobenzenesulfonic acid.

☆ This article is part of a Special Issue entitled: Plant Proteomics— a bridge between fundamental processes and crop production, edited by Dr. Hans-Peter Mock.

* Corresponding author at: Department of Medical Protein Research, VIB, 9000 Ghent, Belgium.
E-mail address: kris.gevaert@vib-ugent.be (K. Gevaert).

http://dx.doi.org/10.1016/j.bbapap.2016.01.001
1570-9639/

Therefore, the question of how to cope with such enormously complex peptide mixtures arose, due to limitations of these early instruments. Seminal work was done in the Yates and Aebersold labs starting in the late 1990s. The Yates lab introduced different and consecutive chromatographic separation steps of peptides prior to analysis by mass spectrometry. By separating peptides according to charge (ion exchange chromatography) followed by separation according to hydrophobicity (reverse-phase chromatography) [6], their MudPIT approach (multidimensional protein identification technology [7]) allowed for a comprehensive analysis of proteomes, albeit at the expense of an increased analysis time. A different route was taken by the Aebersold lab, who introduced the isotope-coded affinity tag or ICAT approach [8]. Here, instead of analysing the whole peptide mixture, cysteine-containing peptides were specifically modified with a biotin label, allowing for their enrichment prior to mass spectrometric analysis. Given that cysteine is a rather rare amino acid, though well-distributed over a given proteome (implying that most proteins do hold at least one cysteine residue), these cysteine-containing peptides served as proxies for the original proteins. As a bonus, the ICAT molecules came in variants, each containing different stable isotopes, thus allowing for quantitative proteome analysis.

The combined fractional diagonal chromatography or COFRADIC approach introduced by Gevaert and Vandekerckhove in 2002 [9] followed to some extent the ICAT concept; i.e. specific peptide families rather than whole proteome digests were analysed by mass spectrometers. In the next paragraphs, we will first introduce COFRADIC and discuss how this technology has provided insights in different modifications occurring on plant proteins.

2. Diagonal chromatography and mass spectrometry-driven proteomics — history repeating itself?

Back in 1966, Brown and Hartley introduced diagonal paper electrophoresis for isolating cysteine-containing peptides [10]. Following a first separation of peptides, the essential step was the oxidation of cysteines to cysteine sulfonic acids using vapours of performic acid. The extra negative charge that the cysteine-containing peptides obtained caused them to migrate outside of the diagonal line upon repeating electrophoresis in a direction perpendicular to the first dimension. Later, Cruickshank and colleagues introduced the term diagonal chromatography when describing their work on isolating tyrosine-containing peptides [11].

We replaced paper electrophoresis/chromatography by reverse-phase high-performance liquid chromatography for peptide separation [9]. As a consequence, the peptide modification step now had to alter the hydrophobicity (and thus the column retention) rather than the overall charge of a targeted set of peptides, such that these peptides (or the non-modified peptides, see below) can be isolated during the second peptide separation step. Further, by combining peptides fractionated during the first chromatographic step (Fig. 1), in total less chromatographic separation time was required, explaining the COFRADIC acronym as combined fractional diagonal chromatography.

Originally, hydrogen peroxide was introduced as the reagent that specifically oxidizes the side-chain of methionine to a methionine sulfoxide under well-controlled reaction conditions. The net result was that peptides carrying the latter amino acid became less hydrophobic were less retained by the hydrophobic reverse-phase resin and could therefore be isolated from peptides not containing methionine during a series of peptide separation steps [9].

It is worth mentioning that methionines were chosen for similar reasons as to why cysteines were chosen as targets for ICAT because, like cysteines, methionines are rare though well-distributed amino acids and the methionine side-chain allows for specific chemical modification. On the other hand, a clear advantage of COFRADIC over ICAT is its versatility as different peptide modification steps can be used, allowing the isolation of different sets of peptides. As a result, we were able to introduce a variety of COFRADIC methods, each targeting a specific set of peptides. Examples include methods that isolate amino- or carboxy-terminal peptides [12,13] that, amongst others, allow the identification of substrates of proteases (see below), and methods by which glycopeptides can be studied upon specific enzymatic deglycosylation used as the sorting reaction in diagonal chromatography [2,14].

Every proteomics method comes with its own advantages and disadvantages, the latter for COFRADIC being the different protein and peptide chemistries that are necessary and often considered as difficult, and the rather lengthy chromatographic steps. Nevertheless, diagonal chromatography has now been fairly broadly adopted by various groups for their own proteome research such as for isolating phosphorylated peptides [15], cross-linked peptides [16], for studying newly synthesized proteins [17] and tyrosine nitration events [18]. Other groups introduced variants of the original COFRADIC method by changing the amino-terminal peptide sorting reagent such as in the recent proteogenomics study of Bland and co-workers on the *Roseobacter denitrificans* bacterium [19]. Finally, the concepts of the COFRADIC approach for studying protein ubiquitination [20] were used in a recent study on protein sumoylation [21].

All these examples nicely illustrate that the concept of diagonal chromatography for mass spectrometry-driven proteomics was taken up by different labs leading towards the further expansion of the diversity of proteomic aspects that can be tackled by such techniques.

3. Diagonal chromatography applied to study plant protease substrates

Proteases cleave peptide bonds and in this way irreversibly modify their substrates. Protein processing is known to lead to the activation as well the inactivation of substrates, and is involved in various biological processes including highly controlled signalling events during organ development for protein catabolism and amino acid recycling by proteasomes (see for example Ref. [22]).

According to the latest release of the MEROPS database [23], the genome of the model plant *Arabidopsis thaliana* encodes for 745 different proteases. However, the in vivo substrates for the majority of these proteases are unknown, despite the key importance of this information for the unravelling of how these proteases function [24]. Proteases create new protein termini after cleaving a substrate. Therefore, proteomic strategies that specifically isolate protein termini and further distinguish between termini formed by protease activity from those present irrespectively of the protease studied, are ideal to identify native protease substrates and even the exact sites of cleavage. The Gevaert lab pioneered these so-called protease degradome studies [12,25], and since then, several such proteomic technologies became available (recently reviewed in Ref. [26]). Protein processing creates new, primary α-amino groups that are typically not further modified by enzymes other than (exo)peptidases. Note here that the main modification occurring on protein N-termini, N-terminal acetylation, is a typical co-translational modification (reviewed in Ref. [27]) that is thus not expected to modify and block N-termini posttranslationally introduced by proteases. Nonetheless, following import in chloroplasts and removal of the chloroplast transit peptide, such newly generated protein N-termini may get posttranslationally acetylated as was recently demonstrated in a large-scale study by the van Wijk lab [28]. Different stable isotopes can be chemically introduced at protease-generated N-termini to flag these new N-termini either when comparing different setups or within a single setup. In N-terminal COFRADIC, N-terminal peptides are isolated by a series of steps, including stable isotope flagging of protein N-termini, strong cation exchange at acidic pH and the actual COFRADIC chromatographic steps in which the sorting agent 2,4,6-trinitrobenzenesulfonic acid (TNBS) is used for modifying peptide N-termini generated upon trypsin digestion (for technical details, we refer the reader to reference [29]).