JOURNAL OF PROTEOMICS XX (2014) XXX–XXX

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jprot

Review

Laying the groundwork for proteomics: Mass spectrometry from 1958 to 1988☆,☆☆

Klaus Biemann

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

ARTICLE INFO

Keywords:
Mass spectrometry
GCMS
Peptides
Proteins
DNA sequencing

ABSTRACT

The development of mass spectrometric methods in peptide and protein chemistry in the author's laboratory is reviewed, from the first determination of the amino acid sequence of small peptides in the late 1950s to its use for the determination of the primary structure of large proteins by a combination of mass spectrometry and DNA sequencing in the late 1980s. This article is part of a Special Issue entitled: 20 years of Proteomics.

Contents

1. Introduction

The year 1953, only 60 years ago, was an important one for biochemistry and biology. Fred Sanger had completed the first primary structure of a protein, insulin [1] and James Watson and Francis Crick had proposed the double helical structure for DNA [2]. The unraveling of the genetic code was still years away. Mass spectrometry was an important tool for hydrocarbon analysis in the petroleum industry, but practically unknown to biochemists, with the exception of its use for isotope ratio measurements of gases, like CO_2 and N_2, introduced in 1939 by Schoenheimer and Rittenberg for the study of metabolisms [3]. Chromatography was limited to preparative procedures such as column chromatography on alumina or silica gel, or paper chromatography for analytical separations.

☆ This article is part of a Special Issue entitled: 20 years of Proteomics.
☆☆ This is a personalized, partly condensed, partly expanded version of an article that appeared in the International Journal of Mass Spectrometry 2007;259:1–7.

E-mail address: kbiemann@mit.edu.

1874-3919/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jprot.2014.01.008

Please cite this article as: Biemann K, Laying the groundwork for proteomics: Mass spectrometry from 1958 to 1988, J Prot (2014), http://dx.doi.org/10.1016/j.jprot.2014.01.008

For the sequencing of the two amino acid chains of insulin, which could be separated upon cleavage of the connecting two disulfide bonds, Sanger had developed a procedure which involved partial acid hydrolysis of the oligopeptide chain into small peptides, separating these by paper chromatography, visualizing them with ninhydrin, cutting out each band and treating each peptide with dinitro-fluorobenzene, a reagent he had developed for this purpose. It reacts with the N-terminal amino group of a peptide. The product was then acid hydrolyzed and again subjected to paper chromatography. Upon spraying with ninhydrin, free amino acids could be identified by their positions on the paper strip and the dinitrophenyl(DNP)-amino acid showed up as a yellow band. Thus, for a dipeptide, the sequence was unambiguous, while there were two possible sequences for a tripeptide, but the latter still could provide overlaps. In this tedious way the complete amino acid sequence of the two chains, 21 and 30 amino acids, respectively, long, was determined. The arrangement of the two disulfide bonds was deduced from the amino acid composition of peptides obtained by partial acid hydrolysis of intact insulin.

1.1. *Peptide sequencing by mass spectrometry*

My training at the University of Innsbruck, Austria, had been in synthetic organic chemistry, but then in 1957 I was appointed instructor – at the time the lowest rung of the academic ladder – at the Massachusetts Institute of Technology, Cambridge, MA, in the analytical division of the department of chemistry, I had to think about a field of research with an analytical ring. One of the ideas was to develop a method for the labeling of a small peptide at its C-terminus, to complement Sanger's strategy. It would uniquely determine the sequence of a tripeptide, and limit that of a tetrapeptide to two possibilities. This would greatly simplify the experiments and make the combined procedure more efficient. The reaction involved the conversion of the C-terminal carboxyl group into an amino-triazole ring, a reaction I had worked out while still at the University of Innsbruck [4]. I applied successfully for an NIH grant to carry out this work.

While an undergraduate student, Robert Goldman, B.S.'58, was working out this reaction on a small scale with small peptides, as his senior thesis, I learned, entirely by accident, as described elsewhere [5], about mass spectrometry, an instrumental method I had never heard of before. From what was known at the time, I quickly realized that it is particularly informative for linear molecules, and my interest in methods for the sequencing of peptides, and ultimately proteins, came to mind. So I switched my approach from a chemical to an instrumental one.

At that early time, the only way to obtain a mass spectrum was to ionize a molecule in the gas phase by an electron beam (electron ionization, EI). However, peptides cannot be vaporized without thermal decomposition. Here again, my background in organic synthesis reactions came in handy. Since the reason for the non-volatility of peptides is their zwitterionic character, and the hydrogen bonds between the carbonyl groups and amino group of the peptide bonds, it was necessary to eliminate both, while retaining the linear backbone of the molecule. I knew about Paul Karrer's work at the University of Zurich, Switzerland, who had shown that carbonyl groups (free carboxylic acids, esters or amides) could be reduced to CH_2 groups using lithium aluminum hydride ($LiAlH_4$). The only problem was that the chemistry department at MIT did not have a mass spectrometer. But when I described my plans to the head of the department, Professor Arthur C. Cope, he promised to raise the money ($ 50,000, at that time more than any instrument used by organic chemists), if I promise that the instrument would not collect dust. We both kept our promise.

The NIH grant I had received not only permitted switching to a more promising approach, but also contained funds for a postdoctoral associate, so I offered this position to Dr. Josef Seibl, who had recently obtained his Ph.D. in organic chemistry at the same department as I had a few years earlier, and whom I knew well. He arrived at MIT a few days before the mass spectrometer, a CEC 21-103C model (standard in the petroleum industry) was delivered. An NSF grant I had received at the same time, provided for a second postdoctoral position, which I offered to another recent Ph.D. from Innsbruck, Fritz Gapp, who arrived a few months later. The three of us got to work on the reduction of peptides to polyamino alcohols (Fig. 1). Lithium aluminum deuteride was used to differentiate aspartic and glutamic acid from serine and threonine. As I had predicted, these molecules cleaved upon electron impact preferentially at the –CH(R) – CH_2-bond flanked by the two nitrogens producing sequence specific ions of significant abundance (Fig. 2). This work, published as a short, one page communication in the Journal of the American Chemical Society in 1959 [6] represented the first use of mass spectrometry for peptide (and eventually protein) structure analysis. Word about further work got around quickly among the then still small peptide sequencing and synthesis community which invited me to present it at the European Peptide Symposium held in Basle, Switzerland, August 1960 (see Fig. 3).

Demonstrating the principle was, however, far from using it to deduce the primary structure of even a small protein. The separation of the complex mixtures of peptides produced upon partial acid hydrolysis by tedious paper chromatography, as Sanger had done, would be incompatible with the speed of the mass spectrometer. Fortunately, James and Martin in the UK had, in the meantime, developed gas chromatography for the separation of volatile organic molecules [7]. Equally fortunately, the polyamino alcohols derived from small peptides were sufficiently volatile to be amenable to this separation technique (Fig. 4) [8].

1.2. *Sequencing of intact proteins*

It took almost two decades since these first experiments for our technique to mature to the point that it could be used to sequence a protein exclusively by mass spectrometry. The direct coupling of the gas chromatograph to the mass spectrometer [9] greatly facilitated and speeded up the data acquisition, and computer algorithms aided in their interpretation [10]. Further derivatization of the polyaminoalcohols by trifluoroacetylation [11] rather than acetylation and silylation of the hydroxyl groups [12] extended the size of peptide derivatives that could be analyzed. The first complete protein sequenced exclusively by mass spectrometry was that of subunit I of monellin, a polypeptide of 44 amino acids in length [13].

Please cite this article as: Biemann K, Laying the groundwork for proteomics: Mass spectrometry from 1958 to 1988, J Prot (2014), http://dx.doi.org/10.1016/j.jprot.2014.01.008