A computer program to determine the amino acid sequence of proteins by utilizing data obtained from peptide mixtures

Carlo Caporale\(^1\), Carla Caruso\(^2\), Pasquale Petrilli\(^2\), Ciro Sepè\(^2\), Elia Pecchio\(^2\), Vincenzo Buonocore\(^3\)

\(^1\) Dipartimento di Agrobiologia ed Agrochimica, Università della Tuscia, Via S. Camillo de Lellis, I-01100 Viterbo, Italy
\(^2\) Istituto di Industrie Agrarie, Università di Napoli, Portici, Italy

Received: 6 May 1993 / Accepted: 15 June 1993

**Abstract.** We have formulated a computer program with the aim of determining the primary structure of proteins using sequence data obtained from peptide mixtures. The protein under study is digested by different methods and the unfraccionated peptide mixtures are submitted to automatic sequence analysis. The data are cross-related to make exact peptide overlaps according to the strategy proposed by Gray [Nature (1968) 220:1300-1304]. To clarify possible doubts at some positions of the sequence, due to the simultaneous overlap of more than one residue, the output of the program also furnishes the masses (MH\(^{+}\)) of all possible peptides used for the sequence determination for each mixture. Thus, if existing, the doubts can be clarified performing a fast atom bombardment mass spectrometry (FAB-MS) analysis on the unfraccionated peptide mixture(s). To test the validity of the proposed method, the complete amino acid sequence of a trypsin inhibitor from wheat kernel (encoded WTI) was redetermined; no experimental FAB-MS data were necessary in the presented case.

**Introduction**

In 1968 Gray [1] proposed a strategy for determining protein sequences which limited the necessary peptide purification steps. The strategy consists in utilizing phenylthiobutyrazone (PTH)-amino acid data for the sequence analyses performed on the unfraccionated peptide mixtures obtained by digesting the protein by different methods. Obviously, several PTH-amino acid residues are identified at each step of the Edman degradation using this approach. The example Gray proposed to explain this strategy is shown in Fig. 1, where the amino acids released by ten steps of Edman degradation performed on CNBr (cleaving after Met) and clostripain (cleaving after Arg) digests of bovine pancreatic ribonuclease are listed [1]. It should be evident that, starting from the third step in the CNBr set and from the first step in the clostripain set, the unambiguous partial sequence Arg-Asn-Leu-Thr-Lys is compatible with both sets of PTH data. Similarly, starting from the sixth and ninth step in CNBr set and from the first step in clostripain set the partial sequences Arg-Glu-Thr-Gly-Ser and Arg-Cys are consistent with both sets of data, respectively. Moreover, several partial sequences showing doubts at some positions due to the overlap of more than one residue can be deduced. By means of this algorithm, consisting of determining the residue(s) common to each consecutive step of Edman degradation in all sets of data after their correct alignment, Gray concluded that the complete sequence of ribonuclease (124 residues) could be deduced if it were possible to specifically cleave the protein after Arg, Lys, Cys, Met, Tyr, His and to identify the first ten residues of each peptide [1]. Gray's strategy inspired some methods which use

**Abbreviations:** PyEt, pyridylethyl; FAB-MS, fast atom bombardment mass spectrometry; PTH, phenylthiobutyrazone; RP-HPLC, reverse phase high-performance liquid chromatography

**Enzymes:** Endopeptidase Lys-C; endopeptidase Asp-N

**Correspondence to:** C. Caporale

---

**Fig. 1.** Gray's strategy. Amino acids that would be released by ten steps of Edman degradation performed on (A) CNBr and (B) clostripain digests of bovine pancreatic ribonuclease. The unambiguous partial sequences Arg-Asn-Leu-Thr-Lys, Arg-Glu-Thr-Gly-Ser and Arg-Cys are consistent with both sets of data.
An algorithm to determine protein sequence alignment by utilizing data obtained from a peptide mixture and individual peptides

Carlo Caporale, Ciro Sepe, Carla Caruso, Pasquale Petrelli and Vincenzo Buonocore

Abstract

With the aim of limiting peptide purification steps and unambiguously ascertaining protein sequences, we have designed and implemented on a personal computer an algorithm to determine sequence alignment by utilizing data obtained from automatic Edman degradation performed on a single peptide mixture and individual peptides. The protein under study is digested by two different hydrolysis methods and fragments are just isolated from one mixture and sequenced, while the second mixture is submitted unfractonated to sequence analysis. The algorithm provides for the exact alignment of the individual peptides using the mixture data for the overlapping. We report an example of application of this approach by utilizing experimental data obtained from a protein of known sequence.

Introduction

An approach to limit peptide purification steps in assessing protein sequence was first proposed by Gray in 1968, who designed an algorithm working on phenylthiohydantoine (PTH) amino acid data deriving from sequence analysis of unfractonated peptide mixtures obtained digesting the protein by various hydrolysis methods (Gray, 1968). This algorithm, consisting of determining the residue(s) common to each consecutive step of Edman degradation in all sets of data after their correct alignment on the basis of the specificity of hydrolytic agents, suggested diverse approaches using data deriving from unfractonated protein digests (Fairwell et al., 1970; Cannon and Lovins, 1972; Biemann, 1980; Matsuo et al., 1981; Kitagishy et al., 1981; Shimonishi et al., 1981; Herlihy and Biemann, 1981; Kitagishy et al., 1982; Erickson and Jardine, 1986). These procedures, revised by Petrilli and Colosimo (Petrilli and Colosimo, 1990), are mainly based on the use of mass spectrometry data, owing to the feature of this technique of working on peptide mixtures. Furthermore, algorithms to determine sequence alignment utilizing Edman degradation data of individual peptides combined with amino acid composition data (Petrilli, 1985) or fast atom bombardment mass spectrometry (FAB-MS) data from peptide mixtures (Petrilli et al., 1991) have been reported.

Recently, we proposed an algorithm based on the original Gray's idea not requiring FAB-MS data (Caporale et al., 1993) and applied it in assessing the sequence of a trypsin inhibitor from wheat kernel coded WTI (Poerio et al., 1989; Poerio et al., 1994). We showed that the interpretation of the sequence data obtained from peptide mixtures deriving from the hydrolysis of the pyridylethylated protein (PyEt-WTI) by CNBr, endoproteinase Lys-C and endoproteinase Asp-N was very simple. The complete sequence of the protein was determined utilizing the above three sets of data and just one uncertainty was found at one position (Cys or Arg). The theoretically derived FAB-MS masses of all plausible peptides present in the mixtures and matching the output sequence were also furnished to the user in order to clarify doubts performing a FAB-MS analysis. In fact, the possible simultaneous overlap of more than one residue in the aligned sets of data generates uncertainties. However, these uncertainties can be also resolved by this algorithm supplying additional sequence data deriving from further hydrolysis method(s) (Caporale et al., 1993).

A different approach can be used to exclude all possibility of doubts. In fact, it is possible to reconstruct the entire amino acid sequence of a protein from sequencing data derived from a complete set of individual peptides together with the sequencing information from a mixture derived from another hydrolysis method. In this paper, we describe an algorithm implemented on a personal computer which is based on this idea allowing one to assess protein sequences unequivocally. Obviously, no doubt is possible when the primary structure of individual peptides is determined and data from a peptide mixture are utilized for the alignment in order to limit purification steps and sequence analyses. This algorithm represents an alternative to the previous one (Caporale et al., 1993) when FAB-MS data or numerous hydrolytic agents are not available. We also furnish an example of application using experimental data obtained from individual peptides and a mixture deriving from different digestions of PyEt-WTI.

1Dipartimento di Agrobiologia ed Agronomia, Università della Tuscia, via S. Camillo de Lellis, 01100, Viterbo, Italy and 2Istituto di Industria Agrarie, Università di Napoli, Portici, Italy

*To whom reprint requests should be sent
Assignment of protein disulphides by a computer method using mass spectrometric data

Carlo Caporale*, Ciro Sepe*, Carla Caruso*, Piero Pucci, Vincenzo Buonocore

*Dipartimento di Agrobiologia ed Agrochimica, Università della Tuscia, Via S. Camillo de Lellis, 01010 Viterbo, Italy
**Dipartimento di Chimica Organica e Biologica, Università Federico II di Napoli and Servizio di Spettrometria di Massa del CNR, Napoli, Italy

Received 18 July 1996; revised version received 2 August 1996

Abstract We designed a computer program for the assignment of protein disulphides using mass spectrometric data. All the theoretical linear peptides containing from one to three cysteines are generated on the basis of the protein sequence. Their combination ways are determined in order to create all the possible disulphide-bridged fragments containing from two to six cysteines and to calculate their molecular weight. One, two and three S-S bridges per linked fragment were considered, taking into account the possibility that the fragments contain unbridged residues. The mass data obtained from the spectral analysis of peptide mixtures of the digested protein are then associated to the fitting structures of disulphide-bridged peptides, giving rise to the primary output. This output can then be scored by using information on the specificity of the proteolytic agent(s) used and/or any further mass data provided by Edman degradation and/or carboxypeptidase treatment of the peptide mixtures. The need for such a computer aid is discussed and examples of application are given.

Key words: Disulphide bridge; Protein structure; Computer method; Mass spectrometry

1. Introduction

During the last decade, fast atom bombardment mass spectrometry (FAB-MS) has been widely used for the assignment of disulphide bridges in proteins [1-16]. The success of this procedure is mainly due to the advantage that the separation of complex peptide mixtures such as those generated by proteolytic digestion of native proteins is often unnecessary. Therefore, low amounts of the original protein are generally required and information for a rapid assignment can be obtained. The principle of the FAB-mapping strategy, first introduced by Morris and Pucci [1], consists in the direct identification of disulphide-bonded fragments in digestion mixtures. The attribution of the mass signals to the corresponding bridged peptides is usually confirmed by submitting the original peptide mixture to manual Edman degradation step(s) or carboxypeptidase treatment and observing the appearance of new signals in the mass spectra due to the expected loss of amino acids from the N- or C-terminus of the fragments. The focal point in this procedure consists in the unambiguous attribution of each mass signal to the corresponding structure of a cystine-linked fragment and/or cluster of peptides. The assignment procedure can be very complicated especially when specific proteases have to be used to digest the native protein with the aim of producing fragments containing a limited number of disulphide bridges (usually one or two). Moreover, depending on the size of the protein and the number of fragments, a single mass signal can be associated to many clusters of bridged peptides, with the ambiguity remaining still unsolved even after one step of chemical and/or enzymatic degradation [13].

This paper describes the design and application of a computer program analyzing the possible solutions and filtering the results on the basis of new data obtained after degradation and/or information on the specificity of the proteolytic agent(s). We also suggest that such a computer aid can be indispensable in many cases for the unambiguous assignment of disulphide bridges.

2. Materials and methods

The mass data and experimental procedures of the two proteins used in the application examples have already been reported [13,15]. Programs were written using Microsoft QuickBASIC (version 1.00b) and implemented on an Apple Macintosh LC 475 computer. The operating system was System 7.1. The compiled applications are compatible with all Apple Macintosh computers.

3. Results and discussion

3.1. Required data

The data required to run the program are the amino acid sequence of the target protein written in mono-literal code and the mass values corresponding to putative S-S bridged peptides, as provided by the mass analysis of protein digests. The specific algorithms described in this paper have been designed for the use of integer variables such as the nominal FAB-MS MH+ data, since, up to now, these are the most used data for the assignment of protein disulphide bridges [1-16]. However, the average mass data usually provided by conventional instruments equipped for electrospray ionisation (ESI) or matrix-assisted laser desorption/ionisation (MALDI) can also be processed by the method. The corresponding algorithms are very similar to those described in this paper; although the calculation speed is reduced, the running time strictly depends on the calculator used. All the structures of disulphide-bridged peptides whose molecular weights match the experimental mass values are identified by these data. Some additional information concerning the specificity of the proteolytic agent(s) and/or the mass data obtained following Edman degradation or carboxypeptidase treatment of the peptide mixtures must be provided in order to filter the global results. In particular, when possible, the user should indicate the expected cleavage sites on the protein sequence by indicating the amino acid(s) recognized by the proteolytic agent(s) and the side, whether N- or C-terminal, where the cleavage is expected to occur. All these data are input from the keyboard and stored on the disk by simple input/edit routines. The program is divided into three sections, each corresponding

*Corresponding author. Fax: (39) (761) 357242.
CysMap and CysJoin: Database and tools for protein disulphides localisation

Carlo Caporalea,*, Laura Bertinib, Piero Puccib, Vincenzo Buonocorea, Carla Carusoa

a Dipartimento di Agrobiologia ed Agrochimica, Università della Tuscia, via S. Camillo de Lellis, 01100 Viterbo, Italy
b Dipartimento di Clinica Organica e Biochimica, Università di Napoli Federico II, via Cinthia 6, 80126, Napoli, Italy

Received 28 February 2005; revised 30 March 2005; accepted 20 April 2005

Available online 5 May 2005

Abstract We have developed a computer program able to make user-customised databases derived from the public PIR non-redundant reference protein database. When the database of interest has been created, the user will generate the map of all the possible linear peptides containing one and two cysteines for each protein and combine them to calculate the mass of all the possible clusters of linear peptides linked by a disulphide bridge with a cysteine pair. It is also possible to create selected maps corresponding to peptides formed by the action of specific proteases. In this way, mass spectrometric data obtained from the hydrolysis of proteins of unknown sequence can be related to that contained in the database for quick disulphide assignment and protein identification. To confirm signal attribution, the program will also furnish the expected mass of cluster peptides after performing a cycle of Edman degradation. The utility of the program is discussed and examples of application are given.

© 2005 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Protein database; Mass spectrometry; Disulphide assignment; Software

1. Introduction

In the modern proteomics, mass spectrometry combined with electrophoresis is the elective technique for protein characterisation. In fact, the separation of proteins by electrophoresis followed by digestion with proteases is a common practice, the mass of the produced peptides being then determined by mass spectrometry analysis. These data are often used to perform similarity searches using public protein databases. To this purpose, many sites collecting useful links are present in the World Wide Web. As an example, the Expasy site (http://www.expasy.ch/tools) offers a series of informatic tools based on the analysis of linear peptide mass fingerprints [1–4]. Anyhow, no software dedicated to disulphide bridges assignment is available at the moment. Since 1985 [5] and up to recent years [6–22], mass spectrometry has been widely used for the assignment of disulphide bridges due to the advantage that the separation of the peptide mixtures generated by proteolytic digestion of native proteins is often unnecessary. The technique allows the direct identification of cluster peptides linked by disulphide bonds in the mass spectra of digestion mixtures by their unique mass values. The attributions of the mass signals to the corresponding bridged peptides can be confirmed by submitting the peptide mixture to manual Edman degradation step(s) and observing the appearance of new signals in the mass spectra due to the expected loss of amino acids from the N-terminus of the peptides [11,12]. The success of this procedure consists in the unambiguous attribution of a mass signal to the corresponding cystine-linked peptide containing a single S–S bridge, since the identification of cluster peptides containing more than two cysteines is not decisive for the disulphides assignment. In this case, further sub-digestions are necessary to produce useful peptides. In 1996, we developed a computer program generating all the possible pairings of Cys residues in a protein, followed by filtering the results on the basis of new data obtained after degradation and/or information on the specificity of the proteolytic agent(s) [23]. This program was developed for Macintosh computers and was able to work just on a single sequence inputted from the keyboard by the user. In this paper, we present a computer program able to work on entire databases extracted by the user from the PIR non-redundant reference protein database (PIR-NREF). In this way, mass spectrometric data obtained from the protein of interest can be used to perform similarity searches in the database for quick disulphide assignment and protein identification.

2. Materials and methods

Public PIR-NREF protein database was downloaded from the web site ftp://ftp.pir.georgetown.edu/pir_databases/nref. The source code of the program was written using Real Basic 5.5.2 from Active Software. The applications have been developed on a Macintosh computer equipped with a double G5 processor (2.5 GHz), 2.5 GB RAM and compiled for Windows, Mac OS and Linux. The operative system was Mac OS 10.3.5. All the programs work both on Apple Macintosh and Windows PC computers. The output files can be read by the text editor included in the program or by any word processor.

3. Results and discussion

3.1. User-customised databases

The PIR-NREF protein database includes all sequences from PIR-PSD, Swiss-Prot, TrEMBL, RefSeq, GenPept, and PDB. It is updated bi-weekly and grouped by sequence identity and taxonomy at the species level. At the moment, it contains 1,638,166c + 6 sequences. The file structure is very suitable for
A Computer Program to Compare Sequence Fingerprints of Homologous Proteins for the Rapid Assessment of Their Primary Structure Differences

Carlo Caporale,1,2 Ciro Sepe,1 Carla Caruso,1 and Vincenzo Buonocore1

Received July 20, 1998

We have developed a computer program for the rapid assessment of the primary structure differences between a protein of unknown sequence and a homologous known protein. Both proteins are reduced, alkylated, and digested with the same hydrolytic agent. The unfractionated peptide mixtures are submitted to automatic sequence analysis. Based on the knowledge of the reference sequence, the program utilizes the analysis data to identify all the potential peptides present in the two mixtures, determining their primary structure, homology degree, and molecular weight calculated both as integer MH+ and average mass variables. These fingerprints allow the user to easily identify the structural differences between the two proteins and clarify possible doubts by a mass spectrometric analysis of the two mixtures. In order to verify the utility of the program, we provide an application example using the already reported data of two homologous proteins.

KEY WORDS: Algorithm; computer; protein sequence; mass spectrometry; peptide mixtures.

1. INTRODUCTION

In recent years, sequence and mass spectrometric analyses have complemented each other in the rapid assessment of structural protein features since both techniques are able to analyze peptide mixtures. Automatic sequencers require a few picomoles of material to perform Edman degradation and problems related to different yields of phenylthiohydantoin (PTH)-amino acids and to the carryover of the reaction have been minimized in recent apparatus. Similarly, mass spectrometric analyses using modern ionization techniques are very sensitive. Of course, the data must be interpreted and many logical methods have been developed to solve specific problems quickly. The large spread of personal computers has inspired many authors to design algorithms for the interpretation and correlation of data. Methods based on knowledge of a sequence have been developed for the assignment of disulfide bridges (Caporale et al., 1996a) and the determination of the protein hydrolysis pathway (Caporale et al., 1996b, c; Petrelli et al., 1994), while different approaches to assessing protein sequences have been reported that depend on the degree of knowledge of the structure of the target molecule. These programs are useful for determining a completely unknown sequence (Johnson and Walsh, 1992; Caporale et al., 1993; Fernandez-de-Cossio et al., 1995), establishing the alignment of a partially known sequence (Petrelli et al., 1991; Caporale et al., 1994), or identifying new proteins by comparing their mass fingerprints with those stored in structure databases (Mann et al., 1993; Yates et al., 1993, 1995; James et al., 1993, 1994; Pappin et al., 1993; Henzel et al., 1993; Mann and Wilm, 1994; Cottrell, 1994; Griffin et al., 1995; Patterson and Aebersold, 1995; Mortz et al., 1996; Yates, 1996).

Noncomputerized strategies based on the analysis of peptide mixtures have been devised to determine the primary structure of a protein, given the sequence of a homologous protein (Caporale et al., 1991, 1996c). The methods described in these papers are very useful for...
An algorithm to analyse the hydrolysis pathway of peptides and proteins by sequence analyses of unfractionated digestion mixtures.

**Caporale C**, Sepe C, Caruso C, Garzillo AM, Buonocore V.

**Abstract**
We have designed and implemented on a personal computer a program for identifying and quantifying the fragments present in a peptide mixture obtained by hydrolysing a polypeptide of known sequence using digesting agents. The qualitative data utilized by the main algorithm consist of the target sequence of the intact molecule and the amino acid residues identified at each step of the automatic sequence analysis of the unfractionated digestion mixture. In this way, the sequence of each fragment present in the mixture is quickly reconstructed. Furthermore, if the quantitative data of the amino acid residues identified at each step of the sequence analysis are utilized, the program will correlate the sequence of each fragment to its amount. We furnish an example of the application intended for the rapid identification and characterization of the extracellular proteinases produced by a basidiomycete fungus, utilizing the bovine insulin beta-chain as target substrate. A variety of uses for the method are discussed.

PMID: 8744770 [PubMed - indexed for MEDLINE]
An algorithm to analyse the hydrolysis pathway of peptides and prot... - PubMed - NCBI

How to join PubMed Commons

Probing the Modelled Structure of Wheatwin 1 by Controlled Proteolysis and Sequence Analysis of Unfractionated Digestion Mixtures

Carlo Caporale,1,6* Carla Caruso,1 Angelo Facchiano,2 Monica Nobile,1 Luca Leonardi,1 Laura Bertini,1 Giovanni Colonna,2 and Vincenzo Buonocore1
1Dipartimento di Agrobiologia ed Agrochimica, Università della Tuscia, Viterbo, Italy
2CRISCEB-Centro di Ricerca Interdipartimentale di Scienze Computazionali e Biotecnologiche, II Università di Napoli, Napoli, Italy

ABSTRACT We set up a method to get rapid information on the three-dimensional structure of peptide and proteins of known sequence. Both native and alkylated polypeptide is hydrolyzed with a number of proteases at different digestion times and the resulting mixtures are compared by HPLC analysis to establish the differences in the hydrolysis pathways of the folded and unfolded molecule. Then, the unfractionated digestion mixtures of the native polypeptide are submitted to automatic sequence analysis to identify the hydrolysis sites. The sequence of each fragment present in the mixtures is reconstructed and its amount determined by quantitative data of the sequence analyses. We used this approach to determine the amino acid surface accessibility of wheat-win 1, a pathogenesis-related protein from wheat, and constructed a predictive three-dimensional model based on the knowledge of the tertiary structure of barley, a highly homologous protein from barley. The procedure allowed us to quickly identify and quantify the hydrolysis at the susceptible bonds which could be classified as exposed, partially hidden, or inaccessible. The results were useful to evidence and discuss concordances and differences between experimental and model predicted accessibilities of amino acid residues. Proteins 1999;36:192–204.

Key words: molecular modelling; peptide mixtures; protein structure; sequence analysis; surface accessibility

INTRODUCTION

The knowledge of the tertiary structure of proteins in solution is a fundamental goal in establishing their structure-function relationship. In the last years, NMR techniques have been used in tackling this problem [Ludvigsen and Poulsen,1 Lecompt,2 and Wutrich,3 and references therein]. This approach provides spatial atom coordinates which can be used by simulation procedures to determine the three-dimensional structure of homologous proteins.4–7 Furthermore, useful information on the global folding can be obtained by investigating the amino acid accessibility. This can be done by methods labeling the exposed residues such as isotopic exchange reactions8–9 and chemical modifications10–12 or identifying them by enzymatic hydrolysis of their peptide bonds.13–22 In general, the use of proteolytic enzymes has been shown to be an excellent tool in studying protein folding.26–27 Using this approach, peptic fragments are usually purified and successively identified. The identification can be accomplished by classical chemistry methodologies28–32 or by mass spectrometry analysis.33–42 In any case, the fragment quantification represents a problem since the yield of purification steps could be different for each product. On the other hand, mass spectrometry analysis, although being a rapid and sensitive method which can be directly performed on digestion mixtures, does not produce data which can be easily quantified. In this view, a fast and sensitive procedure not requiring purification steps and providing quantifiable data should be very useful when fragment isolation is not required for studying particular domains.

In the last years, the automatic sequence analysis of peptide mixtures by modern sequencers equipped on-line with phenylthiohydantoin (PTH) amino acid analyzers has been useful used to characterize proteins. Different strategies in assessing protein sequences have been reported43–49 together with methods for determining the hydrolysis pathway of polypeptides by unknown digesting agents.50–53 These approaches, which have been recently reviewed,54 do not require any purification step and are based upon the sequence determination of the fragments in mixture: moreover, sequence data being quantitative, the amount of each fragment can be evaluated.53–54 The method can also be used to get rapid information on the tertiary structure of proteins in solution. In fact, if a native protein of known sequence is digested by a specific protease, it is possible to follow the kinetics of the hydrolysis at the recognized sites by quantitative sequence analysis of the mixture products at different digestion times. In this way, each expected hydrolysis site can be classified as exposed, partially hidden,
Analytical techniques and computer algorithms combined for the rapi... - PubMed - NCBI
Analytical techniques and computer algorithms combined for the rapid characterization of structural peptide and protein features.

Caporale C.

Abstract
The most recent algorithms based on the use of modern analytical techniques for the assessment of structural peptide and protein features have been reviewed. No algorithm devoted to the realization of predictive models or statistical analysis has been discussed, but only methods furnishing information on the real structure of the molecules. In particular, the procedures designed for handling sequence and mass spectrometric data obtained from the analysis of unfractionated digestion mixtures allow the user to get rapid information on the structure of the target polypeptide. Two classes of methods are illustrated: the first regards the determination of the amino acid sequence, whereas the second used its knowledge to supply data on the localization, function and three-dimensional structure of disulphides.

PMID: 9894847 [PubMed - indexed for MEDLINE]