PPR

PEPTIDE AND PROTEIN REVIEWS

Volume 1

Executive Editor

Milton T.W. Hearn

PEPTIDE AND PROTEIN REVIEWS

Volume 1

Executive Editor

MILTON T. W. HEARN

ST. VINCENT'S SCHOOL OF MEDICAL RESEARCH
VICTORIA PARADE
MELBOURNE, VICTORIA
AUSTRALIA

MARCEL DEKKER, INC. New York and Basel

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

ISBN 0-8247-7053-6

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE

Interest in all areas of peptide and protein chemistry has been expanding rapidly in recent years. It is no chance event that this has occurred. Progress in practically every area of the biological sciences is now dependent on advances made in peptide and protein chemistry. In fact, many of the health science desciplines, for example, immunology and endocrinology, are now so dependent on the capabilities elaborated by peptide and protein chemistry that advances in these fields of endeavor invariably follow and reflect earlier advances in peptide and protein chemistry. The resultant information gap and enormous growth in the published literature has tended to make communication between these various disciplines increasingly difficult. *Peptide and Protein Reviews* will serve as a forum to bridge these disciplines and bring increasing awareness of significant emerging areas in peptide and protein chemistry. As a consequence, *Peptide and Protein Reviews* should provide a cohesive vehicle for the exchange of ideas amongst all who have an interest in the techniques and concepts of peptide and protein chemistry.

To meet the more specialized needs of important new areas of development, *Peptide and Protein Reviews* will publish, from time to time, volumes devoted to a single topic or theme. In addition, *Peptide and Protein Reviews* will publish a calendar of events related to professional meetings, symposia, and specialized courses. I invite you to send this information to me for inclusion in subsequent volumes. I sincerely hope that you will make *Peptide and Protein Reviews* an integral part of your regular professional reading, in your undergraduate and graduate courses, and in the exchange of ideas with your scientific colleagues.

Milton T. W. Hearn

CONTRIBUTORS

David M. Abercromie Laboratory of Chemical Biology, National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland

Robert Anholt The Salk Institute for Biological Studies, San Diego, California

J. Claude Bennett Division of Clinical Immunology and Rheumatology, Department of Microbiology, University of Alabama in Birmingham, University Station, Birmingham, Alabama

Ajit S. Bhown Division of Clinical Immunology and Rheumatology, Department of Medicine, University of Alabama in Birmingham, University Station, Birmingham, Alabama

Irwin M. Chaiken Laboratory of Chemical Biology, National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland

Roberta F. Colman Department of Chemistry, University of Delaware, Newark, Delaware

M. J. S. De Wolf RUCA - Laboratory of Human Biochemistry, University of Antwerp, Antwerp, Belgium

W. S. H. Dierick UIA - Laboratory of Pathological Biochemistry, University of Antwerp, Antwerp, Belgium

H. J. J. Hilderson RUCA - Laboratory of Human Biochemistry, University of Antwerp, Antwerp, Belgium

Tatsuhiko Kanmera Laboratory of Chemical Biology, National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland

A. R. Lagrou RUCA - Laboratory of Human Biochemistry, University of Antwerp, Antwerp, Belgium

Jon Lindstrom The Salk Institute for Biological Studies, San Diego, California

Mauricio Montal Departments of Biology and Physics, University of California, San Diego, La Jolla, California

Yash P. Myer Institute of Hemoproteins, Department of Chemistry, State University of New York - Albany, Albany, New York

Reginald Paul Sequeira Laboratory of Chemical Biology, National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland

G. A. F. Van Dessel UIA - Laboratory of Pathological Biochemistry, University of Antwerp, Antwerp, Belgium

CONTENTS

RECENT ADVANCES IN THE MICROSEQUENCING OF PROTEINS AND PEPTIDES

Ajit S. Bhown and J. Claude Bennett

Division of Clinical Immunology and Rheumatology of the Department of Medicine, and Department of Microbiolgy, University of Alabama in Birmingham, University Station Birmingham, AL

INTRODUCTION

The ultimate goal in the study of a protein is to understand its biological function (1) expressed in terms of a unique molecular structure. This molecular organization is the biological evidence of possible genetic events (2,3) that may have resulted from experimental manipulations or that randomly occurred during evolution. In the broadest sense, the structure of a protein molecule is defined by the sequential order of amino acids along the polypeptide chain, this in turn dictates the folding of the chain responsible for its functional role.

The elucidation of the primary structure of complex protein molecules has been made possible by three major categories of experimental techniques; 1) physical 2) enzymatic and 3) chemical.

The physical methods include a) x-ray crystallography (4), which provides significant information about the folding of the polypeptide chain and the positions of the amino acid residues; b) nuclear magnetic resonance (NMR) (5), which evaluates the structure of proteins and peptides in solution by paramagnetic chemical shift; and c) gas chromatography-mass spectrometry (GS-MS) (6), which provides information on protein fragments generated by chemical or enzymatic cleavage. These physical methods offer great promise and utility, they will not be further discussed in this review. Our purpose here is to define and compare current methods for automated sequencing of polypeptides.

Both enzymatic and chemical methods are utilized in the production of appropriate fragments for amino acid sequence determination by the Edman method. Although sequential enzymatic degradation of proteins and peptides may sometimes be utilized, it is limited by the selectivity of available exopeptidases (7-9). These methods are highly dependent on the amino acid residues present in the polypeptide and sometimes may result in ambiguous information. In addition, commercial availability and stability of such enzymes in pure forms remain a major problem. However, their use in identifying amino acids located near the peptide termini may be helpful in planning future strategies of sequencing by chemical methods. Two aminopeptidases, leucine-aminopeptidase and aminopeptidase M have been usefully employed for sequence determination.

Amino acid sequence determination by chemical methods as originally employed by Sanger (10) and later extensively developed

TABLE 1

Historical Milestones in Chemical Degradation of Proteins and Peptides

INVESTIGATOR(S)	TECHNIQUE	YEAR	REF.
F. Sanger	DNP - Manual	1945	10
P. Edman	PITC - Manual	1949; 1950	11,12
P. Edman and G.S. Begg	PITC - Automated	1967	39
R. Laursen	Solid phase - Automated	1971	63
B. Wittmann-Liebold	PITC - Instrument modification with autoconverter	1973	13

by Edman (11,12) is based on the stepwise degradation of a polypeptide chain, releasing one amino acid at a time (Table I). This must be followed by an efficient identification of the liberated amino acid. These methodologies, have been extensively modified (13-16) in the last decade to meet the structural challenge posed by interesting macromolecules available in microquantities. Many of these proteins are keys to advances in biochemistry as they perform important biological functions. Recent technical developments have greatly improved our sequencing capabilities, both in terms of automation and sensitivity. Instrument modification, refinements in the Edman chemistry and progress in the methods for amino acid identification have all contributed towards the attainment of sequencing capabilities (Table II) at the level of 10^{-10} to 10^{-9} moles. These recent modifications that have made

TABLE II

Sequencing Capabilities Spanning Two Decades

	1960's	1980's
Edman Degradation	i) Manual	i) Automated a) Liquid Phase b) Solid Phase
Identification	i) Paper Chromatography ii) Back Hydrolysis iii) Gas chromatography	i) High Pressure Liquid Chromatography ii) Thin Layer Chromatography iii) Back Hydrolysis iv) Gas Chromatography
Sensitivity	Micromole	Picomole
Residues	10-20	50-70

sequencing on the micro scale possible, will be reviewed in this article. Although we have tried to include most of the recent literature, we must apologize to the workers in this field whom we have had to omit due to space limitations.

SEQUENCE ANALYSIS

Following the pioneer work of Sanger (10) in determining the primary structure of insulin, the most important contribution to amino acid sequence analysis was made by Edman (11,12,17) who introduced the phenylisothiocyanate (PITC), method for the sequential degradation of one amino acid at a time from the amino terminus of a polypeptide. This method has been widely employed to obtain meaningful information on the primary structures of proteins/peptides. Numerous other derivatives (18-24) and coupling agents (25-28) have been investigated but none seem to enhance significantly the level of sensitivity, that can be achieved by the PITC method.

Edman degradation has three operative steps, coupling, cleavage, and conversion (Fig. 1). The desire of the scientific community to explore the biological functions at the primary structure level of proteins available only in micro quantities has required the refining of Edman chemistry at each of these operative steps.

Coupling: The first step in the Edman degradation is coupling of PITC with the free α-amino group of the polypeptide chain at pH 9.0 and 56°C. Efficient and effective coupling is the limiting

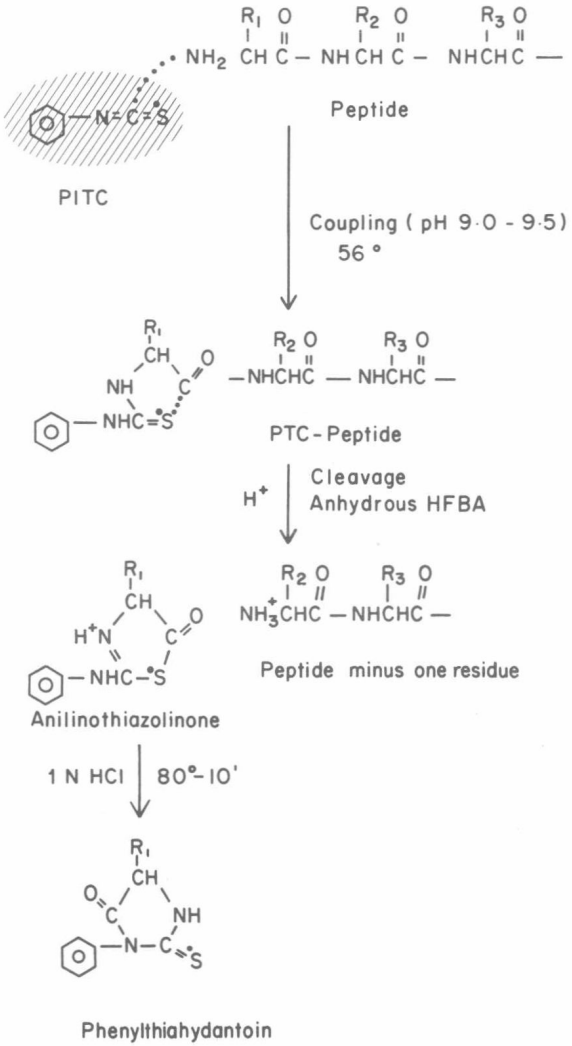

Figure 1. Chemistry of the Edman stepwise degradation. $^{o}S=^{35}S$.

factor in microsequencing. Some of the major factors which control this important initial reaction in stepwise degradation are: a) sample preparation; b) purity of reagents; and c) order of reagent addition.

a. Sample preparation: Quantitative coupling is dependent primarily on the availability of the free α-amino group of the protein molecule. The amino terminus of protein molecules have been found to exist as containing: i) a blocked α-amino-group; ii) a cyclized ring structure and iii) a amino terminus. When a free α-amino group is present the sequence chemistry proceeds smoothly. However, when present as the cyclized ring structure (Fig. 2a-c) or derivatized amino group (Fig. 2d) it becomes a formidable task to degrade the molecule from the amino terminus.

Polypeptides with blocked amino termini have been found to exist in native form, or they may develop during the course of purification. Acetylated or formylated α- amino groups have been found in nature (29-32). Certain compounds, notably aldehydes under optimum conditions of pH and temperature, irreversibly block the amino groups forming a Schiff's base (33) and thus affecting the sequencing efficiency.

Cyclization at the amino terminus may be observed in native polypeptides (32,34) or may be introduced during purification and handling. For example, exposure to: acidic conditions (35) or cyanocysteine cleavage (35) may cause such ring formation (Fig. 2a,b). Glutamine residues are particu-

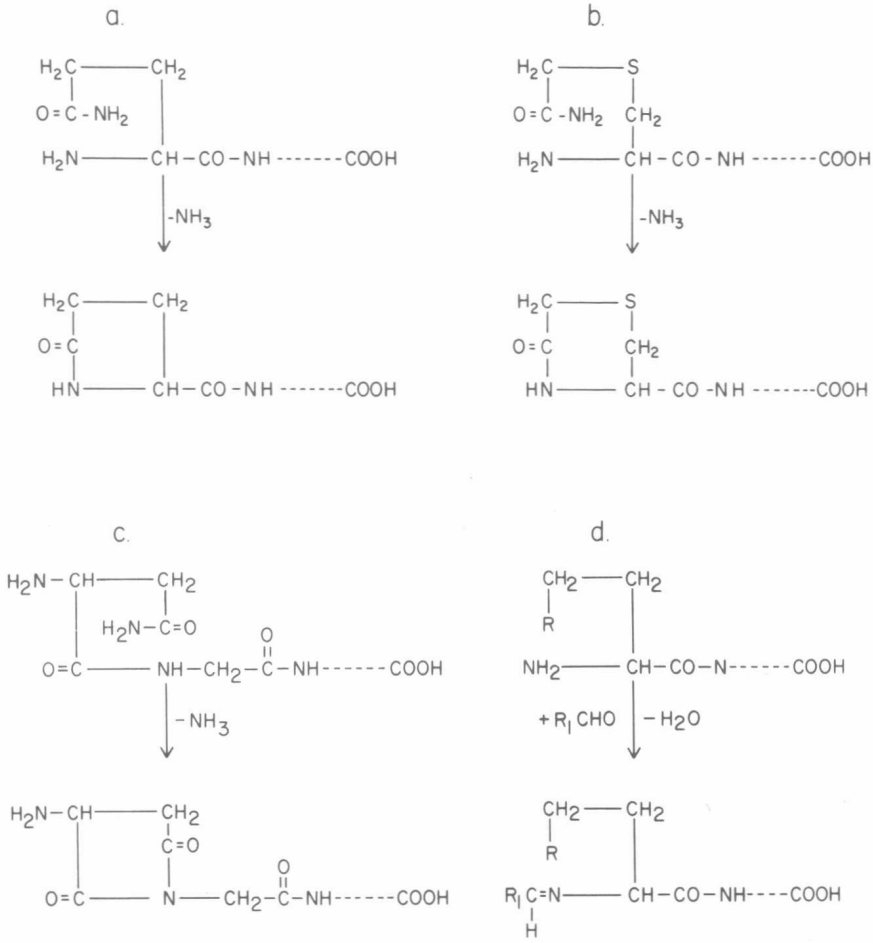

Figure 2. Chemistry of α-amino group blocking: a-Glutamine; b-cyanocysteine; c-Aspargine-Glycine (under acid condition); d-Derivatized α-amino group.

larly susceptible to ring formation. Similarly cleavage of an asparaginyl-glycine bond as a result of rearrangment may result in cyclization of the asparagine residue (35) (Fig. 2c).

Polypeptides with blocked amino termini cannot be degraded from their amino terminus by classical Edman chemistry. However, indirect methods have been employed to obtain sequence information (34). Both enzymatic (35,36) and chemical methods are now available to directly remove cyclized glutamine (pyrrolidone carboxylic acid) residues.

b. Purity of reagents: One of the most critical requirements for an efficient coupling is the purity of reagents. Impurities even in trace amounts can irreversibly compete with PITC and block free α-amino groups thus affecting the coupling yield. Automated Edman degradation utilizes two important reagents -- buffer and PITC -- to achieve the coupling step. The most commonly employed buffers i) quadrol and ii) dimethylallylamine (DMAA) or dimethylbenzylamine (DMBA) are commercially available in sufficently pure form for classical sequence determination. However, these reagents are not completely free from trace amounts of impurities which adversely affect the initial coupling step. This is particularly true for quadrol which undergoes auto-oxidation and generates aldehydes, which react with free α-amino groups of proteins and peptides. To achieve the desired level of coupling efficiency during microsequencing,

Wittmann-Liebold (13) and Hunkapiller and Hood (14,37) have recommended further purification of these chemicals. Frank (15) and Bhown et al. (16) have recently suggested the use of aminoethyl cellulose (15) and aminoethyl aminopropyl (AEAP) glass beads (16) to chemically adsorb the products of auto-oxidation of quadrol as they are formed. These methods of "in line" purification are simple, cheap and ensure delivery of an acceptable quadrol buffer. Since the introduction of polybrene and dilute quadrol (38) as buffer, this non-volatile system has become the one most commonly employed. Only on rare occassions will authors advocate DMAA or DMBA for peptide sequencing.

c. Order of reagent addition: The accumulation of background residues frequently has been the limiting factor to an extended automated Edman degradation (39). Several authors have attributed this to be due to a low degree of non-specific scission of peptide bonds during exposure to acidic conditions (40,41). Thomsen et al. (42) have proposed an additional mechanism involving seryl/threonyl residues in an N→O acyl shift, resulting in a cleavage of the peptide bonds on the amino side of these amino acids. In an effort to avoid such a shift they (42) have developed a program with a reversed delivery of PITC and quadrol to reverse the N→O acyl shift. Although the results of this program are encouraging, further confirmation is warranted.