This is to certify that the

dissertation entitled

Novel Methodologies to Study Cysteine Status in Proteins by Chemical Cleavage and Mass-Mapping by MALDI-TOF Mass Spectrometry

presented by

Jiang Wu

has been accepted towards fulfillment of the requirements for

Ph. D. degree in Chemistry

Major professor

Date May 16, 1997

NOVEL METHODOLOGIES TO STUDY CYSTEINE STATUS IN PROTEINS BY CHEMICAL CLEAVAGE AND MASS-MAPPING BY MALDI-TOF MASS SPECTROMETRY

By

Jiang Wu

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry

1997

ABSTRACT

NOVEL METHODOLOGIES TO STUDY CYSTEINE STATUS IN PROTEINS BY CHEMICAL CLEAVAGE AND MASS-MAPPING BY MALDI-TOF MASS SPECTROMETRY

By

Jiang Wu

Current methodology for characterizing free sulfhydryls and disulfide bonds in proteins involves the modification of cysteine residues, enzymatic digestion of protein chain for cleavage between half cystinyl residues, HPLC fractionation, and sequence determination of peptide fragments. This strategy is tedious, cumbersome, and can promote disulfide exchange.

A simple methodology has been developed to localize free cysteine groups in peptides and proteins. This new approach employs a specific reaction between free sulfhydryls and 2-nitro-5-thiocyanobenzoic acid (NTCB) or 1-cyano-4-dimethylamino-pyridinium tetrafluoroborate (CDAP) to specifically cyanylate cysteine thiols. The N-terminal peptide bond of the cyanylated cysteinyl residue can then be cleaved under alkaline conditions to form an amino-terminal peptide and a series of 2-iminothiazolidine-4-carboxylyl peptides which can be mapped to the sequence by MALDI-MS. The cleavage conditions have been systematically investigated using a number of peptides containing different amino acids adjacent to the N-terminal side of cysteine residues. While the cleavage reaction was traditionally performed in pH 8.5-9.5 buffer at 37-65°C for 12-80 hours, optimal results have been obtained in 1M ammonium solution in which

the cleavage is complete within an hour at ambient temperature. This improvement also minimizes side reactions resulting from prolonged exposure to the alkaline medium.

A novel strategy is described for assignment of disulfide pairings in proteins. A denatured protein is subjected to limited reduction in acidic solution to produce a mixture of partially reduced protein isomers; the nascent sulfhydryls are immediately cyanylated by CDAP under the same buffered conditions. The cyanylated protein isomers, separated by and collected from reversed-phase HPLC, are subjected to the cleavage reaction to form truncated peptides, which, after further reduction of the remaining disulfide bonds, can be mass-mapped by MALDI-MS. This simple, fast, and sensitive strategy avoids disulfide scrambling and is applicable to disulfide characterization in proteins containing adjacent cysteines. Several proteins with various disulfide structures have been studied to demonstrate the feasibility of the methodology.

Two techniques are introduced for the study of disulfide structures of protein folding intermediates. The cyanylation of thiol groups by the CDAP under acidic conditions has been applied to trapping folding intermediates of recombinant human epidermal growth factor (hEGF). Disulfide structures of seven CDAP-trapped intermediates have been identified by the approach described above. Both native and non-native intermediates were found in the folding process. The analysis of the disulfide intermediates provides new insight into the folding pathway of hEGF.

AC

I would like to express my si

Watson, for his guidance, encourage

study at MSU. I would also like to t

and support, and for serving as a sec

Eugene LeGoff for serving on my co

I want to acknowledge Drs

knowledge in protein chemistry an

members, past and present, thank yo

many thanks to other members of the

Bev, Mel. Without their technical sup

And last, but not the least, I

daughter, Mengyu, for their love, supp

also very grateful to my parents, moth

their caring, encouragement, and und

State University.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my research advisor, Dr. J. Throck Watson, for his guidance, encouragement, and financial support throughout my graduate study at MSU. I would also like to thank Dr. Doug Gage for his continual encouragement and support, and for serving as a second reader of my dissertation; Drs. John Allison and Eugene LeGoff for serving on my committee.

I want to acknowledge Drs. Eugene Zaluzec, and Zhi H. Huang for sharing their knowledge in protein chemistry and organic chemistry. To all the Watson group members, past and present, thank you for your friendship and cooperation. I also owe many thanks to other members of the Mass Spectrometry Facility: Melinda, Paochi, Gary, Bev, Mel. Without their technical support, my graduate study would have been difficult.

And last, but not the least, I would like to thank my wife, Ping Gong, and my daughter, Mengyu, for their love, support, and patience that make my life enjoyable. I am also very grateful to my parents, mother-in-law, brothers and sister, and sister-in-laws, for their caring, encouragement, and understanding during my graduate study at Michigan State University.

TAB

TABLE OF CONTENTS

L

LIST OF TABLES

L

LIST OF FIGURES

Figure 3.4 HPLC chromatograms
EKPLQNFTLCFR in (
1 M NH_4OH, rt, 1h, re

Figure 3.5 HPLC chromatograms
MHRQEAVDCLKKFN
18h, and (B) 1 M NH_4C

Figure 3.6 MALDI spectra of the c
after cleavage in (A) pH
solution, rt, 1h, respecti
represent fragment I, fra
uncleaved peptide, and i

Figure 3.7 MALDI spectra of the cy
DRVYIHPCHLLYYS, a
18h and (B) 1 M NH_4OH
Symbol ++ represents a c

Figure 3.8 MALDI spectra of the cy
after cleavage in (A) pH
solution, rt, 1h, respectiv

Figure 3.9 MALDI spectra of the cya
MHRQEAVDCLKKFNA
37°C, 18h and (B) 1 M NH

Figure 3.10 MALDI spectra of spinach
cyanylation by the CDAP
pH 9.0 buffer for 18 hours
solution for one hour at ro

C

INT

I. Introduction

Since its discovery in 1912, ma

broad and diverse disciplines. Howeve

of biomolecules stems in large par

techniques in the early 1980s (1, 2).

sciences have led to a dramatic incr

information of biologically active ma

new discipline-biological mass spectr

structural issues of biomedically impor

has accelerated dramatically, due to

ionization techniques: matrix-assisted

electrospray ionization (ESI) (4). Com

routine the analysis of high mass comp

natural products, and drug metabolites

MS has been revolutionary in spurrin

and biotechnology, and most recently i

Protein analytical techniques a

structure as well as detecting and defi

that the complete structure of a protein

CHAPTER 1

INTRODUCTION

I. Introduction

Since its discovery in 1912, mass spectrometry (MS) has provided key insight for broad and diverse disciplines. However, the involvement of mass spectrometry in studies of biomolecules stems in large part from the discovery of desorption ionization techniques in the early 1980s (1, 2). At the same time, rapid advances in biological sciences have led to a dramatic increase in the demand for chemical and structural information of biologically active materials. The combination led to the creation of a new discipline-biological mass spectrometry-which addresses the challenging unsolved structural issues of biomedically important molecules. In the past decade, this expansion has accelerated dramatically, due to the discovery of two new mass spectrometry ionization techniques: matrix-assisted laser desorption/ionization (MALDI) (3) and electrospray ionization (ESI) (4). Commercial availability of these instruments has made routine the analysis of high mass compounds including proteins, peptides, carbohydrates, natural products, and drug metabolites with picomole to femtomole sensitivity. Today, MS has been revolutionary in spurring research in protein biochemistry, glycobiology, and biotechnology, and most recently in DNA sequencing (5).

Protein analytical techniques are essential for determining a protein's primary structure as well as detecting and defining posttranslational modifications of proteins so that the complete structure of a protein can be obtained (6, 7). Nucleotide sequences of

genes embody the information requ
sequence) of proteins but they do no
chemically modified *in vivo* after tra
Defining posttranslational modificati
are destined for use as pharmaceuti
same chemical structure as their natu
Unacceptable structures could b
pharmacological or immunological a

Defining the locations and str
is one of the most important contribu
to make in the field of protein bioc
evident when it is used in conjuncti
protein sequence and amino acid an
which identification is based entirely
20 commonly occurring amino acids
molecular mass, an intrinsic physical
of choice for characterization of post
well-developed, accurate, and sensi
available.

Disulfide bond formation af
posttranslational modifications frequ
Mass spectrometry, combined with
extensively applied to the recogniti

genes embody the information required to deduce the primary structure (amino acid sequence) of proteins but they do not reveal whether the side chains of amino acids are chemically modified *in vivo* after translation of genes or the extent of any modifications. Defining posttranslational modifications is essential in the case of expressed proteins that are destined for use as pharmaceuticals. This is required to ensure that they have the same chemical structure as their natural counterparts or an acceptable alternate structure. Unacceptable structures could be detrimental to recipients due to unwanted pharmacological or immunological activities.

Defining the locations and structures of posttranslational modifications in proteins is one of the most important contributions that mass spectrometry has made and continues to make in the field of protein biochemistry (8). The unique potential of MS is most evident when it is used in conjunction with conventional methodologies like automated protein sequence and amino acid analysis. In contrast to the latter two approaches in which identification is based entirely on chromatographic retention relative to one of the 20 commonly occurring amino acids or a derivative thereof, MS relies on measurement of molecular mass, an intrinsic physical property. Mass spectrometry is clearly the method of choice for characterization of posttranslationally modified proteins, since in most cases well-developed, accurate, and sensitive chemical or biochemical approaches are not available.

Disulfide bond formation after gene expression is one of the most important posttranslational modifications frequently encountered in protein characterization (9, 10). Mass spectrometry, combined with other chemical/biochemical techniques, has been extensively applied to the recognition of both free sulfhydryls and disulfide bonds in

proteins. In this dissertation, novel m
disulfide bonds, and disulfide struc
greater detail in the chapter 2, 4, and
introduce basic aspects of MALDI as
status in proteins.

II. Characterization of Cysteine St

Among the 20 amino acids th
The reduced form of cysteine contain
cystine, is formed by linking two sul
contributes to protein biological func
active site for enzyme catalysis such
metals (11). Sulfhydryl groups c
physiological and biochemical proce
division, oxidative phosphorylation a
band, comprises the major covalent
chain in nature (Figure 1.1). *Intra*-ch
in ribonuclease A, serve to confer
chain. Additionally, by limiting or
the correct orientation of the amino
antibodies and other biologically
functional in maintaining the quatern
only linkage between subunits or pro

proteins. In this dissertation, novel methodologies for recognizing free sulfhydryl groups, disulfide bonds, and disulfide structures of folding intermediates will be discussed in greater detail in the chapter 2, 4, and 5, respectively. The objective of this chapter is to introduce basic aspects of MALDI and current methodologies for characterizing cysteine status in proteins.

II. Characterization of Cysteine Status in proteins

Among the 20 amino acids that compose proteins, cysteine has unique properties. The reduced form of cysteine contains a sulfhydryl (thiol) group while its oxidized form, cystine, is formed by linking two sulfhydryls together to form a disulfide bond. Cysteine contributes to protein biological functions by using its free sulfhydryl (-SH) group as the active site for enzyme catalysis such as in cysteine proteases, and as the chelating site for metals (11). Sulfhydryl groups can also play an important role in a variety of physiological and biochemical processes such as muscular contraction, nerve activity, cell division, oxidative phosphorylation and photosynthesis (11). Disulfide bond, on the other hand, comprises the major covalent cross-linkage in proteins and may be *intra-* or *inter-*chain in nature (Figure 1.1). *Intra*-chain disulfide bonds, such as the four disulfide bonds in ribonuclease A, serve to confer conformational stability on the folded polypeptide chain. Additionally, by limiting or directing the folding these bonds may contribute to the correct orientation of the amino acid residues that form the active sites of enzymes, antibodies and other biologically active proteins. *Inter*-chain disulfide bonds are functional in maintaining the quaternary structure of multi-chain proteins, serving as the only linkage between subunits or providing covalent stability to structures otherwise

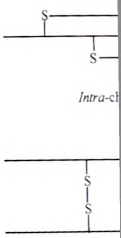


Figure 1.1. *Intra-* and

maintained by non-covalent forces

proteins play a unique role in its act

the mechanism of activity in glutath

that cysteine sulfhydryls and cysti

functions, the recognition of cyste

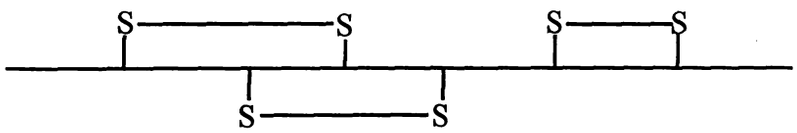
characterization (9, 11).

For a new protein with unk

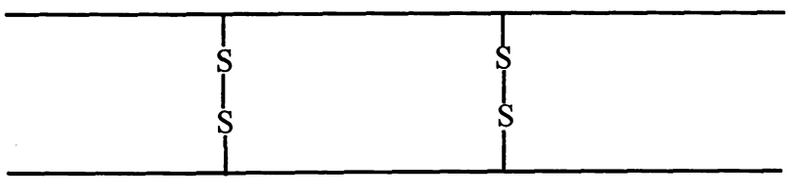
cysteine content is very beneficial to

the amino acid sequence of the pro

cDNA sequence, differentiating bet



Intra-chain disulfide bonds



Inter-chain disulfide bonds

Figure 1.1. *Intra*- and *Inter*-chain disulfide bonds in proteins.

maintained by non-covalent forces (e. g., insulin, IgG). Disulfide bonds in particular proteins play a unique role in its activity. For example, disulfide bonds are involved in the mechanism of activity in glutathione reductase and thioredoxin. Due to diverse roles that cysteine sulfhydryls and cystine disulfide bonds play in protein structures and functions, the recognition of cysteine status is a very important aspect of protein characterization (9, 11).

For a new protein with unknown sequence, the quantitative determination of cysteine content is very beneficial to further characterization of the protein. Even though the amino acid sequence of the protein can be easily deduced from the corresponding cDNA sequence, differentiating between free cysteine residues and those involved in

disulfide bonding (cystines) is also
the cysteine residues after translati
primary structure. Characterization
bonds in proteins has become a cruci
because molecules having incorrec
biological activity than that of the
cysteine status in protein chemistry
cystine; (2) localization of free sulf
linkage.

A. Quantitative Determination of

The cysteine sulfhydryl grou
acids by its high reactivity and by th
they participate, such as alkylatio
exchange, charge-transfer complexe
(11). Sulfhydryls can dissociate into
sulfhydryl groups is attributable to t
ions, which exist at reasonable con
(pK_a 8-10.5). Thiolate-dependent re
at moderate speed in the pH range 6-

RSH

disulfide bonding (cystines) is also very essential because the state and connectivity of the cysteine residues after translation can not be predicted readily from the protein's primary structure. Characterization of the location of sulfhydryl groups and disulfide bonds in proteins has become a crucial part of the analysis of recombinant DNA products, because molecules having incorrect disulfide bond linkages may have much lower biological activity than that of the desired product. The problems associated with cysteine status in protein chemistry include: (1) quantitative evaluation of cysteine and cystine; (2) localization of free sulfhydryl groups; and (3) assignment of disulfide bond linkage.

A. Quantitative Determination of SH and S-S groups in proteins

The cysteine sulfhydryl group is distinguished from other side chains of amino acids by its high reactivity and by the exceptionally diverse chemical reactions in which they participate, such as alkylation, acylation, oxidation, sulfhydryl-disulfide bond exchange, charge-transfer complexes, reactions with organic heavy metal compounds (11). Sulfhydryls can dissociate into a proton and a thiolate anion. The high reactivity of sulfhydryl groups is attributable to the high nucelophilicity of the corresponding thiolate ions, which exist at reasonable concentrations at neutral to weakly alkaline pH-values (pK_a 8-10.5). Thiolate-dependent reactions thus proceed readily at pH 8 and above, and at moderate speed in the pH range 6-8.

$$RSH \longrightarrow RS^- + H^+$$

A variety of methods, based
disulfide bonds participate, have be
disulfide bond content in proteins. R
chapters (11-13).

I. Methods Based on Alkylation

Figure 1.2. The modifica
S-carboxymethylcysteine

Alkylation (S-carboxymethyl
is the most frequently used reaction
be accomplished with compounds c
iodoacetamide, bromoacetate, etc., a
and specificity to SH (Figure 1.2)
electrochemical method (14) or abso
by ^{14}C-labeled iodoacetate permits th
radioactivity of the product (16,

A variety of methods, based on numerous reactions in which sulfhydryls and disulfide bonds participate, have been studied for the determination of sulfhydryl and disulfide bond content in proteins. Reviews on this subject can be found in several book chapters (11-13).

1. Methods Based on Alkylation

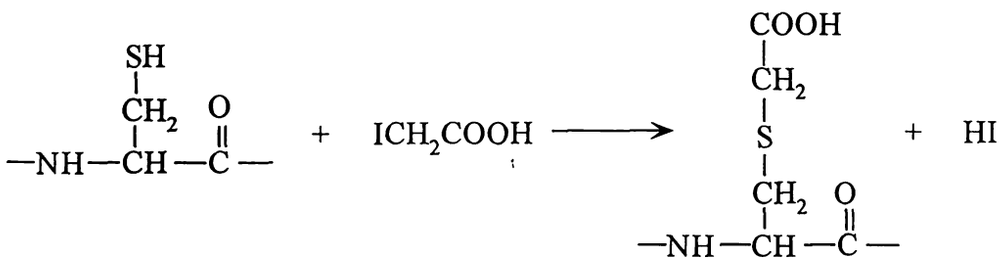


Figure 1.2. The modification of cysteine with iodoacetic acid to form S-carboxymethylcysteine.

Alkylation (S-carboxymethyltion) of SH groups to form S-carboxymethylcysteine is the most frequently used reaction in protein characterization (11-13). Alkylation can be accomplished with compounds containing an alkyl halide group such as iodoacetate, iodoacetamide, bromoacetate, etc., among which, iodoacetamide shows highest reactivity and specificity to SH (Figure 1.2). The iodide liberated can be determined by an electrochemical method (14) or absorbance at 226 nm (15). As an alternative, alkylation by ^{14}C-labeled iodoacetate permits the number of thiol groups to be determined from the radioactivity of the product (16, 17). To avoid side reactions connected with the

formation of iodine by photooxida

performed in the dark.

Based on the same principl

developed in order to improve the

alkylation by iodoacetate is still t

sulfhydryls because of its simplicity

1. Methods Based on Mercaptidat

Organomercurial compounds

SH groups (Figure 1.3). In con

reversible. Among various organo

compounds have widely been e

mercuribenzoate proposed by Boye

sensitivity, selectivity, and precisi

measurement of the increase in abso

binding of *p*-mercuribenzoate (*p*-ME

only the number of SH groups in a

of loss of enzymatic activity and th

rate and the extent of the reaction of

pH 4.6 than at pH 7-8 because of the

result, the reactivity of the reagent in

reagents, such as 1-(4-chloromercur

and p-(hydroxymercuri)-benzoic a

formation of iodine by photooxidation of iodide ions, the carboxymethylation was performed in the dark.

Based on the same principle, a number of other alkylating reagents have been developed in order to improve the sensitivity and specificity of the reaction (18). But alkylation by iodoacetate is still the most popular derivatization reaction to modify sulfhydryls because of its simplicity and specificity.

2. Methods Based on Mercaptidation

Organomercurial compounds are high affinitive and specific reagents for protein SH groups (Figure 1.3). In contrast to alkylation, the mercaptidation reaction is reversible. Among various organomercurial reagents, monofunctional organomercurial compounds have widely been exploited. Spectrophotometric titration with *p*-mercuribenzoate proposed by Boyer stands in first place because of its simplicity, high sensitivity, selectivity, and precision (19). This classical method is based on the measurement of the increase in absorbance in the 250-255 nm region that occurs upon the binding of *p*-mercuribenzoate (*p*-MB) to SH groups. It provides a means to measure not only the number of SH groups in a protein, but also the relationship between the degree of loss of enzymatic activity and the number of blocked SH groups. Interestingly, the rate and the extent of the reaction of SH groups of many proteins with *p*-MB is higher at pH 4.6 than at pH 7-8 because of the high affinity of *p*-MB for hydroxide ions (20). As a result, the reactivity of the reagent increases as the pH of the solution is decreased. Other reagents, such as 1-(4-chloromercuriphenylazo)-napphthol-2 (CMPN; mercury-orange), and p-(hydroxymercuri)-benzoic acid (21), were also proposed to improve the

spectrophotometric detection of S

carboxyhemoglobin (23).

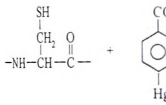


Figure 1.3. Chemical reaction be

3. Methods Based on Sulfhydryl/I

$$R_1SSR_1 + R$$

$$R_1SSR + RS$$

Among oxidants of SH grou

reactions with sulfhydryls are abs

sulfhydryl/disulfide exchange. As c

of two steps of nucleophilic substitu

intermediate stage.

spectrophotometric detection of SH groups in lysine monoxygenase (22) and in carboxyhemoglobin (23).

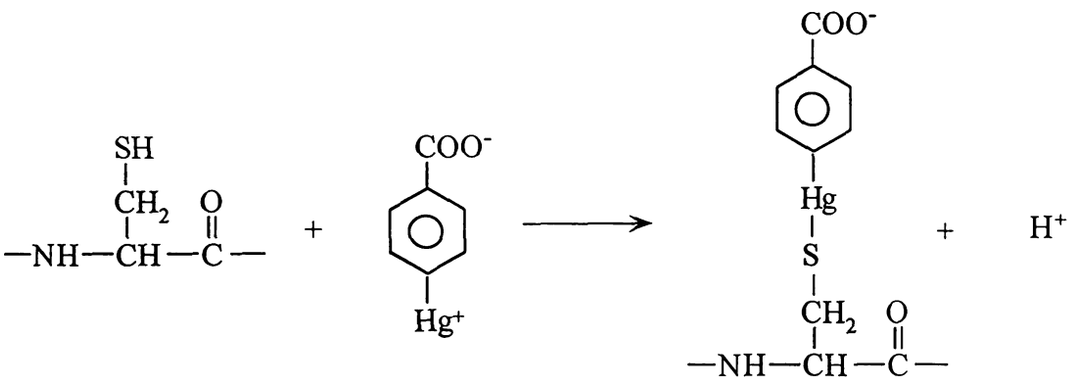


Figure 1.3. Chemical reaction between *p*-mercuribenzoate and sulfhydryl group.

3. Methods Based on Sulfhydryl/Disulfide Exchange

$$R_1SSR_1 + RSH \longrightarrow R_1SH + R_1SSR$$

$$R_1SSR + RSH \longrightarrow R_1SH + RSSR$$

Among oxidants of SH groups, disulfides occupy a special position since their reactions with sulfhydryls are absolutely specific. This reaction is referred to as sulfhydryl/disulfide exchange. As can be seen from the equations, this reaction consists of two steps of nucleophilic substitution with the formation of a mixed disulfide at the intermediate stage.

Among the various disulfi

nitrobenzoic acid), DTNB, has rec

thiolate anion with excess Ellm

stoichiometric formation of TNB

disulfide (RS-TNB) (Figure 1.4).

tris buffer, pH 8, at a concentrati

containing 1mM EDTA in a five t

groups (EDTA is included not only

metal ions can affect the developme

= 13,600 $M^{-1}cm^{-1}$), very fast and st

of SH groups in both native and der

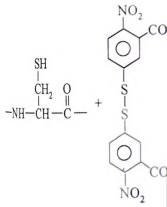


Figure 1.4. The reaction of Ellr

with cysteinyl residues in prot

Among the various disulfides proposed, Ellman's reagent, 5,5'-dithio-*bis*-(2-nitrobenzoic acid), DTNB, has received the widest application (24). The reaction of thiolate anion with excess Ellman's reagent at pH 7-8 is favored toward the stoichiometric formation of TNB thiolate (5-thio-2-nitrobenzoate, TNB^-) and a mixed disulfide (RS-TNB) (Figure 1.4). Ellman's reagent usually is dissolved in phosphate or tris buffer, pH 8, at a concentration of 10 mM. It is added to a solution of protein containing 1mM EDTA in a five to ten-fold excess with respect to the number of SH groups (EDTA is included not only to protect SH groups from oxidation but also because metal ions can affect the development of the color). The method is highly sensitive (ε_{412nm} = 13,600 $M^{-1}cm^{-1}$), very fast and strictly specific, and may be used for the determination of SH groups in both native and denatured proteins (25-27).

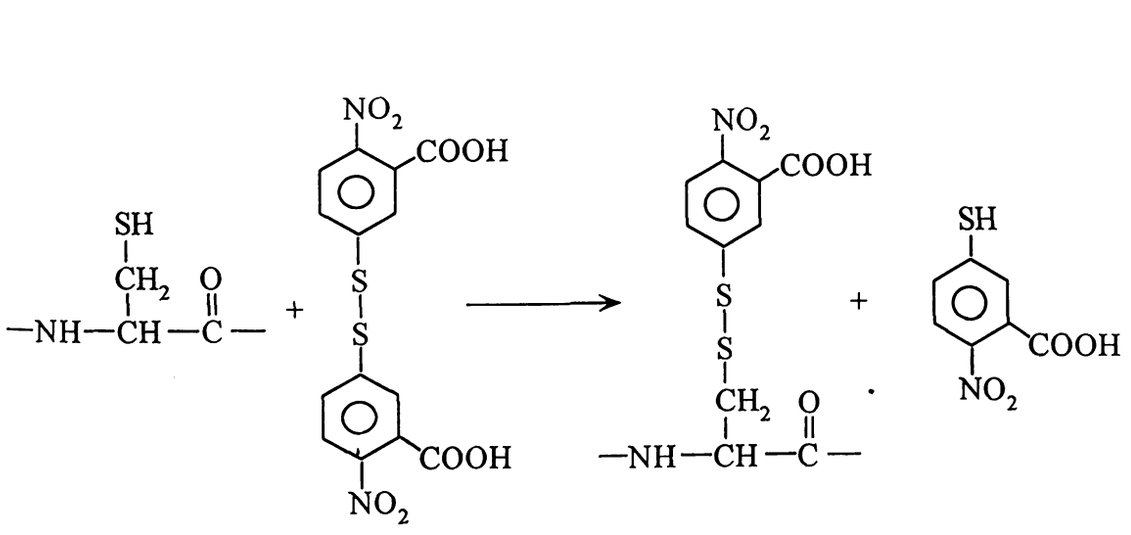


Figure 1.4. The reaction of Ellman's reagent, 5,5'-dithio-bis-(2-nitrobenzoic acid), with cysteinyl residues in proteins.

Grassetti and Murray (28,

dithiodipyridine were good substitue

stronger reactivity to SH groups in

widely used as the binding bed o

proteins (31).

4. Methods Based on Electrochem

One of the important proper

each other though oxidation/reducti

Under mild conditions (e.g., in the

bonds. Further oxidation results in

reducing agent, disulfide bonds can

$$2RSH \rightleftharpoons$$

Those specific oxidation/red

the basis of amperometric titration

the first portion of $AgNO_3$ binds t

small current to increase. As soo

appear in solution and are reduced

current which is proportional to the

Grassetti and Murray (28, 29) demonstrated 2, 2'-dithiodipyridine or 4, 4'-dithiodipyridine were good substituents for Ellman's reagent. These compounds showed stronger reactivity to SH groups in a wide pH range (pH 1-8) (30) and are therefore widely used as the binding bed of affinity chromatography for sulfhydryl-containing proteins (31).

4. Methods Based on Electrochemical Titration

One of the important properties of sulfhydryl/disulfide is that they can convert to each other though oxidation/reduction reactions. Cysteine is very sensitive to oxidation. Under mild conditions (e.g., in the air), SH groups undergo oxidation to form disulfide bonds. Further oxidation results in the formation of cysteic acid. Under the presence of reducing agent, disulfide bonds can be reduced to form sulfhydryls.

$$2RSH \underset{[R]}{\overset{[O]}{\rightleftharpoons}} RSSR + 2H^+ + 2e^-$$

Those specific oxidation/reduction properties of a sulfhydryl/disulfide pair form the basis of amperometric titration of sulfhydryls with silver nitrate (32). In this method, the first portion of $AgNO_3$ binds to the SH groups of the protein and do not cause the small current to increase. As soon as all the SH groups are blocked, free metal ions appear in solution and are reduced on the platinum electrode, resulting in a diffusion current which is proportional to the concentration of the metal ions released to solution.

The amperometric titration
sulfhydryl groups. With prior re
extended to quantitation of disulfide
(33) successfully measured the con
purified pig kidney Na^+, K^+-ATPa
ammonium nitrate buffer (pH 7.6).

Recently Sun *et al* (34)
simultaneously both sulfhydryls an
detector system actually has five ele
steel auxiliary electrode, and three
used to distinguish between sulfhy
chromatographic analysis, as illustra
and partially reduced bovine insulin.

Although the advantage
spectrophotometry lies in the possi
solutions, in enzymological studies,
a usage as the spectrophotometric
sensitivity, the necessity of special a

5. Methods based on Oxidation

Shipton *et al* (35) found tha
selectively and quantitatively oxidi

The amperometric titration method was first proposed for the measurement of sulfhydryl groups. With prior reduction of disulfide bonds, this method was also extended to quantitation of disulfide bonds in proteins. For example, Gevondyan *et al* (33) successfully measured the content of free SH groups and disulfide bonds in the purified pig kidney Na^+, K^+-ATPase by amperometric titration with silver nitrate in ammonium nitrate buffer (pH 7.6).

Recently Sun *et al* (34) developed an electrochemical detector to detect simultaneously both sulfhydryls and disulfides in peptides. The three-electrode EC detector system actually has five electrodes, one Ag/AgCl reference electrode, a stainless steel auxiliary electrode, and three Hg/Au amalgam working electrodes, and thus can be used to distinguish between sulfhydryl- and disulfide-containing peptides in a single chromatographic analysis, as illustrated with proteolytic digests of bovine αA-crystallin and partially reduced bovine insulin.

Although the advantage of electrochemical titration in comparision with spectrophotometry lies in the possibility of carrying out analysis in cloudy or colored solutions, in enzymological studies, the electrochemical method has not received as wide a usage as the spectrophotometric method. This is evidently because of its lower sensitivity, the necessity of special apparatus.

5. Methods based on Oxidation

Shipton *et al* (35) found that benzofuroxan (benzo-2-cxa-1, 3-diazole N-oxide) selectively and quantitatively oxidizes the SH groups of proteins, the reagent being

reduced to *o*-benzoquinone dioxime.

absorbance at wavelengths far remo

Alternatively, proteins are treated

unstable residues of cysteine into cys

then be reliably determined by amino

performic acid leads to destruction

quantitative conversion of methionin

is therefore limited.

6. Methods Based on Enzyme Ass

RSH + Papain-S-SCH

Singh *et al* (37) developed

group determination using an inact

The sulfhydryl/disulfide exchange r

SCH_3, results in the stoichiometri

reactivated papain catalyzes the hy

amplified spectrophotometric signal

is about 100-fold more sensitive

sulfhydryls, e. g., glutathione, cyst

determined by this approach. Howe

reduced to *o*-benzoquinone dioxime. The reduction is accompanied by a large increase in absorbance at wavelengths far removed from the region where proteins absorb strongly. Alternatively, proteins are treated with performic acid (36). This is to convert the unstable residues of cysteine into cysteic acid, which withstands acid hydrolysis and may then be reliably determined by amino acid analysis. However, treatment of a protein with performic acid leads to destruction of tryptophan, partial destruction of tyrosine, and quantitative conversion of methionine into methionine sulfone. The practical application is therefore limited.

6. Methods Based on Enzyme Assay

$$RSH + \text{Papain-S-}SCH_3 \longrightarrow RS\text{-}SCH_3 + \text{Papain-SH}$$

Singh *et al* (37) developed a sensitive spectrophotometric assay for sulfhydryl group determination using an inactive disulfide derivative of papain (papain-S-SCH_3). The sulfhydryl/disulfide exchange reaction of a protein cysteinyl residue with papain-S-SCH_3 results in the stoichiometric formation of active papain (papain-SH). The reactivated papain catalyzes the hydrolysis of a chromogenic substrate, resulting in an amplified spectrophotometric signal proportional to the initial amount of SH. The assay is about 100-fold more sensitive than that using Ellman's reagent. A variety of sulfhydryls, e. g., glutathione, cysteamine, penicillamine, etc. have been successfully determined by this approach. However, the feasibility of the method for the measurement

of thiols in large proteins remain to
proteins, the accessibility of protein

7. Methods Based on Mass Spectr

Mass spectrometry is be
determining protein cysteine/cysti
limitations of sensitivity, multiple c
many classical methods (38-40).

Feng *et al* (38) demonstrate
ESI-MS for rapidly counting cyste
proteins. In this analysis, a peptid
iodoacetic acid in the absence of a
sulfhydryl group(s). Parallel analys
data from which the number of fre
molecular weight of the alkylated d
case for iodoacetic acid, a mass shift
protein. A complementary experim
protein to a disulfide reducing re
differentiation of the total number
Recently Sun *et al* applied this pro
lactoglobulin B (39). However, the
shift from alkylation by iodoacetic a

of thiols in large proteins remain to be tested. Because of the steric hindrance of large proteins, the accessibility of protein sulfhydryls to papain might be a problem.

7. Methods Based on Mass Spectrometry

Mass spectrometry is becoming an alternative means of quantitatively determining protein cysteine/cystine content in a way that promises to overcome limitations of sensitivity, multiple derivatization, and protein contamination that burden many classical methods (38-40).

Feng *et al* (38) demonstrated a method that combines reduction/alkylation and ESI-MS for rapidly counting cysteines, free sulfhydryl groups, and disulfide bonds in proteins. In this analysis, a peptide or denatured protein is first allowed to react with iodoacetic acid in the absence of a disulfide reducing agent to selectively label any free sulfhydryl group(s). Parallel analyses using underivatized protein provide mass spectral data from which the number of free sulfhydryls can be determined by the shift in the molecular weight of the alkylated derivative relative to that of the native protein. In the case for iodoacetic acid, a mass shift of 59 u is observed for each derivatized cysteine of a protein. A complementary experiment involving the same reaction, after exposure of the protein to a disulfide reducing reagent, provides mass spectral data that allows for differentiation of the total number of cysteine(s) and/or cystine(s) originally present. Recently Sun *et al* applied this procedure to ribonuclease A, lysozyme, insulin, and β-lactoglobulin B (39). However, the above method is limited to small proteins as the mass shift from alkylation by iodoacetic acid is small and the limited mass resolution available

could potentially prevent small ma

drawback, Zaluzec *et al* (40) u

hydroxymercuribenzoate (*p*HMB)

derivatization, the mass shift of 321

mass shift minimizes the error which

to poor resolution of the instrument

protein samples and provide an e

cysteine content even in a protein m

8. Determination of S-S Content

Conventional methodologie

reductive cleavage of the S-S bonds

the methods mentioned above. Th

deduced.

A sensitive and quantitative

thiosulfobenzoate (NTSB) has dra

actually composed of two sequentia

disulfide bond with sodium sulfite:

$$RSSR' + SO_3^{2-}$$

could potentially prevent small mass shift from being detected. To circumvent the drawback, Zaluzec *et al* (40) used monofunctional organomercurial reagent (*p*-hydroxymercuribenzoate (*p*HMB) for selective derivatization of thiol groups. After derivatization, the mass shift of 321 u is obtained for each derivatized cysteine. The large mass shift minimizes the error which might occur during mass spectrometric analysis due to poor resolution of the instrument. Both MALDI and ESI can detect low picomole of protein samples and provide an excellent method for the quantitative assessment of cysteine content even in a protein mixture.

8. Determination of S-S Content

Conventional methodologies for S-S counting in proteins are based on the reductive cleavage of the S-S bonds to form cysteinyl residues that can be determined by the methods mentioned above. The content of disulfide bonds can then indirectly be deduced.

A sensitive and quantitative method involving the use of the reagent 2-nitro-5-thiosulfobenzoate (NTSB) has drawn great attention (41, 42). The NTSB assay is actually composed of two sequential reactions. The first reaction is the cleavage of a disulfide bond with sodium sulfite:

$$RSSR' + SO_3^{2-} \longrightarrow RSSO_3^- + R'S^-$$

The second reaction (Figur

produced in reaction on NTSB to

thiobenzoate (NTB), the latter can b

The reaction of NTSB with t

of both dissolved oxygen and the

inconvenience associated with the n

well as the need to remove the reduc

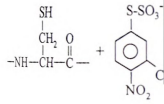


Figure 1.5. Modification of cystei

Hirose *et al* (43) descri

polyacrylamide gel electrophoresi

quantitative determination of in

electrophoresis is one of the most

biological systems, since it requi

visualized by selective methods of

numbers of cleaved disulfide bonds

The second reaction (Figure 1.5) involves nucleophilic attack of the thiolate produced in reaction on NTSB to yield 1 mol each of a thiosulfonate and 2-nitro-5-thiobenzoate (NTB), the latter can be measured by the absorbance at 412 nm.

The reaction of NTSB with thiols and disulfides can be carried out in the presence of both dissolved oxygen and the reducing agent which eliminates the inaccuracy and inconvenience associated with the necessity to work under an oxygen-free atmosphere as well as the need to remove the reducing agent.

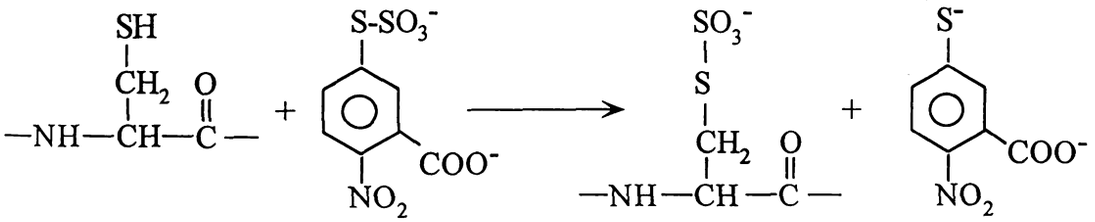


Figure 1.5. Modification of cysteine sulfhydryl by 2-nitro-5-thiosulfobenzoate (NTSB).

Hirose *et al* (43) described another novel method that makes use of polyacrylamide gel electrophoresis (PAGE) following two-step alkylation for the quantitative determination of intramolecular disulfide bonds in proteins. Gel electrophoresis is one of the most powerful techniques for protein analysis in complex biological systems, since it requires very small amounts of proteins which can be visualized by selective methods of gel staining. In this method, proteins with different numbers of cleaved disulfide bonds are alkylated with iodoacetic acid or iodoacetamide

as the first step. The remaining di

and the newly generated free sulfh

(iodoacetamide, iodoacetic acid, or

acid-urea PAGE, different intermed

bands, depending on differences

alkylation used in combination w

analysis of disulfide bonds in protei

B. Localization of Sulfhydryl Gro

Localization of cysteine resi

bonds (see section C) in protei

manipulations in protein chemistry

been tested and used both in the re

available SH groups, and there is no

yet.

I. Classical Approach

The general strategy for

conventional methods involves seve

groups, usually, by an irreversible r

prevent or minimize sulfhydryl/disu

enzymes or chemical reagents betwe

is separated by chromatography. Fi

as the first step. The remaining disulfide bonds were reduced by excess dithiothreitol, and the newly generated free sulfhydryls were alkylated with the reagent not yet used (iodoacetamide, iodoacetic acid, or vinylpyridine) as the second step. By subsequent acid-urea PAGE, different intermediates formed in two steps can be resolved into distinct bands, depending on differences in their net charge and conformation. Two-step alkylation used in combination with autoradiography was especially useful for the analysis of disulfide bonds in proteins synthesized in complex biological systems.

B. Localization of Sulfhydryl Groups

Localization of cysteine residues and their possible pairings in forming disulfide bonds (see section C) in proteins involves some of the oldest and best studied manipulations in protein chemistry. A large number of reagents and procedures have been tested and used both in the reduction of disulfide bonds and in the reaction of the available SH groups, and there is no indication that the search for better methods is over yet.

1. Classical Approach

The general strategy for locating free sulfhydryl groups in proteins by conventional methods involves several steps. The first step is to modify free sulfhydryl groups, usually, by an irreversible reaction such as alkylation, under conditions that can prevent or minimize sulfhydryl/disulfide exchange. Second, the protein is cleaved by enzymes or chemical reagents between cysteine residues. Third, the mixture of the digest is separated by chromatography. Finally, the derivatized sulfhydryl-containing peptides

are recognized, mapped to sequen
and related to specific segments of

A key to the success of thi
reaction and reagent for sulfhydryl
medium and label sulfhydryl gro
conditions (ideally under weak a
exchange which is minimized at
strong UV or fluorescent absor
absorption of proteins, or easily att
of derivatized peptides after HPLC
be distinguishable from other amin
the identification of amino acids
derivatives.

While iodoacetate and pyri
has been far greater interest in the u
larger molecule which can serve a
2,1,3-benzoxadiazole (47), 7-fluor
(aminosulfonyl)-7-fluoro-2,1,3-ban
assay of sulfhydryls by reason of
The adducts exhibit fluorescence a
shift corresponding to environment
under borate buffer, pH 9.5, co
oxidation of the sulfhydryls. For

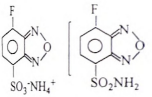
are recognized, mapped to sequence by the Edman technique and/or mass spectrometry, and related to specific segments of the protein.

A key to the success of this approach is to choose an appropriate derivatization reaction and reagent for sulfhydryl groups. The reagent should be soluble in the reaction medium and label sulfhydryl groups selectively, rapidly, and irreversibly under mild conditions (ideally under weak acidic conditions to avoid sulfhydryl/disulfide bond exchange which is minimized at pH 2-6.5). Furthermore, the reagent should possess strong UV or fluorescent absorption which does not overlap with the maximum absorption of proteins, or easily attach a radioactive element to facilitate the recognition of derivatized peptides after HPLC separation. Finally, the derivative of cysteine should be distinguishable from other amino acids by the Edman degradation technique in which the identification of amino acids exclusively relies on the retention time of PTH-derivatives.

While iodoacetate and pyridylethylation are still extensively used (44-46), there has been far greater interest in the use of this chemistry as a mechanism for introducing a larger molecule which can serve as a sulfhydryl probe. Two derivatives of 7-fluoro-2,1,3-benzoxadiazole (47), 7-fluoro-2,1,3-benzoxadiazole 4-sulfonate (SBD-F) and 4-(aminosulfonyl)-7-fluoro-2,1,3-banzoxadiazole (ABD-F), have proven useful for the assay of sulfhydryls by reason of their formation of fluorescent products (Figure 1.6). The adducts exhibit fluorescence at a long wavelength (~515 nm) and show a spectral shift corresponding to environmental hydrophobicity (48, 49). The reaction is performed under borate buffer, pH 9.5, containing 1 mM EDTA to prevent metal-catalyzed oxidation of the sulfhydryls. For the derivatization of sulfhydryls in large proteins, the

proteins have to be denatured to en

51).



SBD-F ABD-F

Figure 1.6. Derivatization

Chin and Wold (52) emplo

reduce disulfide bonds and ABD-

sulfhydryls and disulfide bonds i

apparently do not react with each o

general method for the complete

Using the similar procedure, Kirle

bonds and one free sulfhydryl group

Another reagent, 5-[2-((

(IAEDANS) (Figure 1.7) has been

Recently, Sturrock *et al* (58) used

testicular angiotensin-converting en

from enzymatic digests by HPLC,

proteins have to be denatured to ensure that every SH can be attacked by the reagents (50, 51).

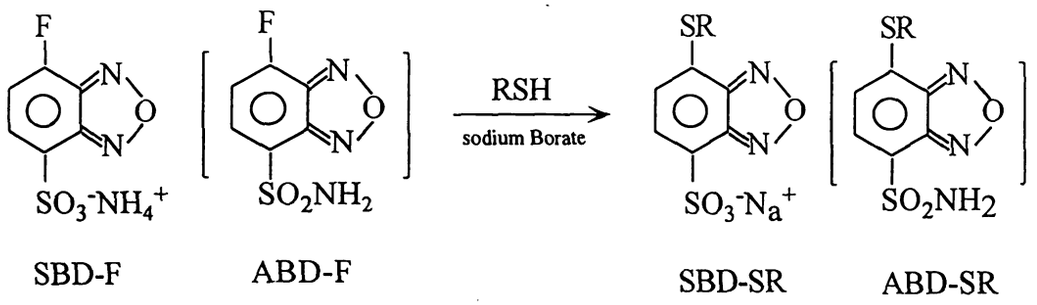


Figure 1.6. Derivatization of the cysteinyl group by SBD-F or ABD-F.

Chin and Wold (52) employed the combination of tributylphosphine (Bu_3P) to reduce disulfide bonds and ABD-F to block free sulfhydryls to characterize both free sulfhydryls and disulfide bonds in a number of proteins. Since the two reagents apparently do not react with each other, their combination offers a convenient and quite general method for the complete characterization of free and cross-linked cysteines. Using the similar procedure, Kirley (53, 54) determined the location of three disulfide bonds and one free sulfhydryl group in the β-subunit of (Na^+,K^+)-ATPase.

Another reagent, 5-[2-((iodoacetyl)amino)-ethyl]naphthalene-1-sulfonic acid (IAEDANS) (Figure 1.7) has been proven powerful for sulfhydryl derivatization (55-57). Recently, Sturrock *et al* (58) used this reagent to label a free cysteine residue in human testicular angiotensin-converting enzyme (tACE). After isolating the fluorescent peptide from enzymatic digests by HPLC, the sequence of the fluorescent peptides was mapped

by MALDI-PSD. The sequence

tACE.

Figure 1.7. 5-[2-((iodoacetyl)am

2. Affinity Chromatography

The only functional group

a stable covalent bond which can

The sulfhydryl-containing protei

interaction (59). After enzymati

peptides do not bind to the colum

containing fragments are then r

structural characterization by Edm

interactions include (60): (i) bind

binding to affinity media with rea

exchange, and (iii) binding to af

methods were originally used to pu

by MALDI-PSD. The sequence data established that Cys496 is a free thiol form in tACE.

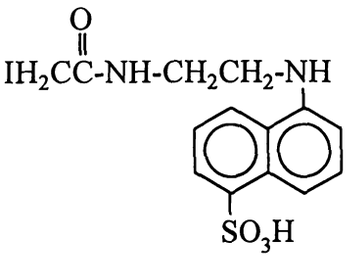


Figure 1.7. 5-[2-((iodoacetyl)amino)-ethyl]naphthalene-1-sulfonic acid (IAEDANS).

2. Affinity Chromatography

The only functional group for which conditions are available for the formation of a stable covalent bond which can be split under mild conditions is the sulfhydryl group. The sulfhydryl-containing protein is attached to an affinity column by covalent interaction (59). After enzymatic digestion on column, the nonsulfhydryl-containing peptides do not bind to the column and can be washed away. The attached sulfhydryl-containing fragments are then removed from the column and subjected to further structural characterization by Edman degradation and/or mass spectrometry. The affinity interactions include (60): (i) binding to affinity media that contain heavy metals, (ii) binding to affinity media with reactive disulfide that undergo facile sulfhydryl/disulfide exchange, and (iii) binding to affinity media that contain chelated zinc. While those methods were originally used to purify sulfhydryl-containing proteins, this approach was

recently used to locate cysteine
sulfhydryl-containing peptides (61
groups in high mass proteins becau
But conditions have to be establi
groups can be attached to the colum

3. **Cleavage at Cysteine Residues**

The third method utilizes a
of an SH group into an SCN gr
treatment of a protein with Ellman
showed that 2-nitro-5-thiocyanobe
sulfhydryls, which subsequently cle
residues under mildly alkaline conc
of 2-iminothiazolidine-4-carboxyly
residues, the cleavage reaction resu
alignment of which indicates the
potentially a very useful method,
because all but the N-terminal p
carboxylyl group, which is not am
convenient method to remove the
Because cysteines are relatively sc
produces large fragments which
polyacrylamide gel electrophoresis

recently used to locate cysteine residues in proteins by mass mapping of bound sulfhydryl-containing peptides (61). This approach is powerful for locating sulfhydryl groups in high mass proteins because the HPLC separation of the digests is unnecessary. But conditions have to be established and optimized under which all the sulfhydryl groups can be attached to the column.

3. Cleavage at Cysteine Residues

The third method utilizes a cleavage reaction at cysteine residues. The conversion of an SH group into an SCN group was first achieved in two stages by successive treatment of a protein with Ellman's reagent and then with cyanide (62, 63). Stark (64) showed that 2-nitro-5-thiocyanobenzoic acid (NTCB) specifically cyanylates cysteine sulfhydryls, which subsequently cleave at the N-terminal side of the cyanylated cysteinyl residues under mildly alkaline conditions to form an amino-terminal peptide and a series of 2-iminothiazolidine-4-carboxylyl (ITC) peptides. If a protein contains **n** cysteine residues, the cleavage reaction results in the formation of **n+1** peptide fragments, mass alignment of which indicates the number and location of cysteine residues. While potentially a very useful method, it has seldom been used for sequence determination because all but the N-terminal peptide become blocked by the iminothiazolidine-carboxylyl group, which is not amenable to Edman degradation. Thus far, there is no convenient method to remove the ITC group from the cleavage products (65, 66). Because cysteines are relatively scarce in proteins, cleavage at these residues usually produces large fragments which can be mass mapped by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (63, 64, 67). However, peptide

assignments using this approach are

5%) (64, 67). Papayannopoulos

spectrometry to sequence the NTC

isolated from *Sarcophaga bullata*.

mass assignment and can be applie

the NTCB approach has primarily

proteins although selective identif

bonds was also reported (68-73).

The cleavage reaction men

mass spectrometry are the theme o

the cleavage reaction will be discus

C. Disulfide Bond Assignment

Although the amino acid s

corresponding cDNA, modification

predicted accurately. Disulfide

posttranslational modifications, is es

As a result, most disulfide-containi

bonding is correct. Improved analy

of S-S are thus important to ma

However, although there is good m

assignments using this approach are often complicated by its poor mass accuracy (error > 5%) (64, 67). Papayannopoulos and Biemann (68) first used CID tandem mass spectrometry to sequence the NTCB cleavage reaction products of a protease inhibitor isolated from *Sarcophaga bullata*. Mass spectrometry provides much more accuracy for mass assignment and can be applied to the sequencing of ITC blocked peptides. To date, the NTCB approach has primarily been employed to locate total cysteines in various proteins although selective identification of free cysteine in the presence of disulfide bonds was also reported (68-73).

The cleavage reaction mentioned above and mass mapping of the products by mass spectrometry are the theme of this dissertation. The chemistry and mechanism of the cleavage reaction will be discussed in detail in chapter 2.

C. Disulfide Bond Assignment

Although the amino acid sequences of proteins are readily deduced from the corresponding cDNA, modifications occurring to the protein after translation can not be predicted accurately. Disulfide bonding, one of the most frequently encountered posttranslational modifications, is essential for stabilization of protein's tertiary structure. As a result, most disulfide-containing enzymes have maximum activity only if disulfide bonding is correct. Improved analytical methods for the rapid and definitive assignment of S-S are thus important to many areas of biochemical research and technology. However, although there is good methods for quantifying the number of disulfide bonds

in proteins (41, 42), the unambiguo

task which is sometimes only resol

Two frequently used ter

exchange and disulfide bond scra

sulfhydryl/disulfide exchange (suc

both sulfhydryl/disulfide exchange

mildly alkaline media when stru

disulfide bond structures. Figure

protein.

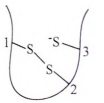


Figure 1.8. Illustration o

Sanger and his coworkers

insulin, recognized that an intercha

protein was subjected to partial

formation of mixed disulfides. Su

disulfide exchange, disulfide bond

was minimal over the range of pH 2

in proteins (41, 42), the unambiguous determination of disulfide bonds remains a difficult task which is sometimes only resolved by a main commitment of time and sample.

Two frequently used terms in cysteine chemistry are sulfhydryl/disulfide exchange and disulfide bond scrambling (or disulfide bond interchange). Although sulfhydryl/disulfide exchange (such as Ellman's reagent) is the most specific reaction, both sulfhydryl/disulfide exchange and disulfide bond scrambling in proteins occur in mildly alkaline media when structures permit and result in mismatched nonnative disulfide bond structures. Figure 1.8 shows the sulfhydryl/disulfide exchange in a protein.

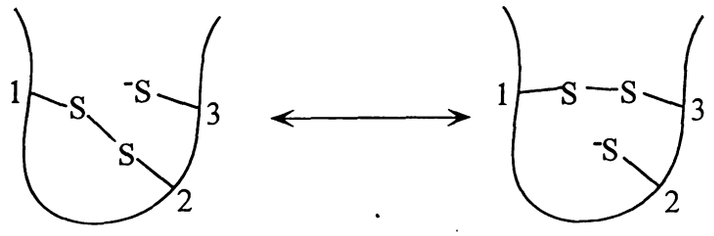


Figure 1.8. Illustration of sulfhydryl/disulfide exchange in proteins.

Sanger and his coworkers (74-76), in their pioneering work on the structure of insulin, recognized that an interchange reaction took place with disulfide bonds when the protein was subjected to partial hydrolysis in cold concentrated HCl, leading to the formation of mixed disulfides. Subsequent studies demonstrated that , like sulfhydryl/ disulfide exchange, disulfide bond scrambling also occurs in slightly alkaline buffers but was minimal over the range of pH 2 to pH 6.5 (77, 78).

R'S-SR' + R"S

In alkaline and neutral mec

from the hydrolytic cleavage of dis

presence of alkylating agents, such

1. **Classical Strategy**

Multiple step en

Fracti

Re

Modific

Separati

Sequenc

Figure 1.9. Classical strateg

$$R'S\text{-}SR' + R''S\text{-}SR'' \longrightarrow 2R'S\text{-}SR''$$

In alkaline and neutral media, the reaction is catalyzed by thiols which can arise from the hydrolytic cleavage of disulfides. Disulfide bond interchange is inhibited by the presence of alkylating agents, such as iodoacetate (79).

1. Classical Strategy

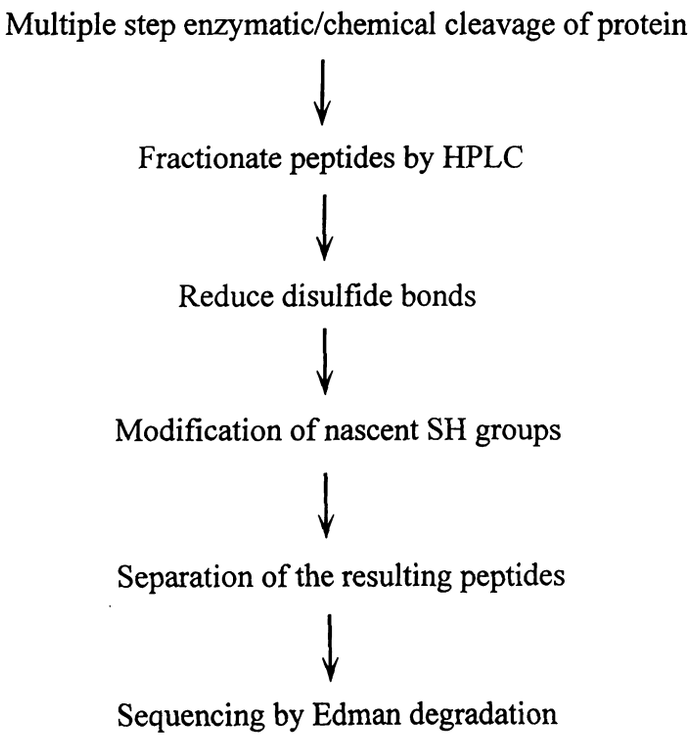


Figure 1.9. Classical strategy for disulfide bond assignment of a protein.

As illustrated in Figure
recognizing disulfide bond structu
similar to the assignment of free
enzymes or chemical reagents bet
avoid or minimize disulfide scram
reversed-phase HPLC. Third, the
peptides are determined using Ed
identified peptides are related to sp
the methodology focused on the
especially various mass spectrom
characteristic aspects of the strateg

2. Cleavage of Cystine Peptides

The choice of enzymes for
specificity of the enzyme to pro
identification process; the condi
capability of cleaving between eve
no more than one disulfide bond. U
factors. The identification of dis
specific cleavage reagent is used (
with a variety of enzymes, such as
or less specifically, chymotrypsin.
highly specific for Arg and Lys resi

As illustrated in Figure 1.9, a well established approach (65, 80, 81) for recognizing disulfide bond structure of proteins involves several steps which are very similar to the assignment of free sulfhydryl groups. First, a protein is cleaved by enzymes or chemical reagents between half-cystinyl residues under conditions that can avoid or minimize disulfide scrambling. Second, the mixture of digests is separated by reversed-phase HPLC. Third, the amino acid sequence or molecular masses of these peptides are determined using Edman degradation or mass spectrometry. Finally, the identified peptides are related to specific segments of the protein. Recent development of the methodology focused on the identification of the disulfide-containing peptides, especially various mass spectrometric techniques are more and more involved. The characteristic aspects of the strategy will be discussed below.

2. Cleavage of Cystine Peptides

The choice of enzymes for cleavage is largely dictated by three requirements: the specificity of the enzyme to produce well-defined fragments that can simplify the identification process; the conditions to minimize disulfide interchange; and the capability of cleaving between every half-cystinyl residue to obtain peptides containing no more than one disulfide bond. Usually compromise has to be made among these three factors. The identification of disulfide-containing peptides is facilitated greatly if a specific cleavage reagent is used (10, 39, 82, 83). Proteins may be cleaved selectively with a variety of enzymes, such as trypsin, *staphyococcus aureus* V8, cyanogen bromide, or less specifically, chymotrypsin. Trypsin has been used most frequently because it is highly specific for Arg and Lys residues. Unfortunately, trypsin has maximum activity at

pH 8.3, and is not active in acid.

enzymolysis. However, in spite of

published papers on the disulfide

cleavage reagent, among which

Chymotrypsin is less specific tha

conditions. *staphyococcus aureus*

4.0 and 7.8, cleaving on the C-t

Cyanogen bromide is also used fr

Met. This reaction is very attracti

of side reactions, and because th

volatile. In addition, disulfide bo

proteins in their native state are of

is used to open up or unfold the pr

is not a particularly abundant cons

large peptides. Subsequent digesti

peptides (82).

Unfortunately, if proteins c

must be cleaved by non-specific r

thermolysin, and partial acid hydrol

specific proteases it is difficult

segments of the protein simply by

case, multiple-step Edman degradat

of the disulfide-bridged peptides.

pH 8.3, and is not active in acid. As a result, disulfide scrambling may occur during enzymolysis. However, in spite of the jeopardy, our survey indicated that among over 60 published papers on the disulfide assignment in various proteins, 70% used trypsin as a cleavage reagent, among which 80% of the digestion was performed at pH>8. Chymotrypsin is less specific than trypsin, and likewise active only under alkaline conditions. *staphyococcus aureus* V8 is useful because it has maximum activity at pH 4.0 and 7.8, cleaving on the C-terminal side of Glu or Glu and Asp, respectively. Cyanogen bromide is also used frequently to cleave proteins on the C-terminal side of Met. This reaction is very attractive as it is highly specific for Met and is generally free of side reactions, and because the reagents (cyanogen bromide and formic acid) are volatile. In addition, disulfide bonds are stable during treatment with CNBr. Since proteins in their native state are often resistant to enzymatic attack, cleavage with CNBr is used to open up or unfold the protein, rendering it susceptible to enzymolysis. As Met is not a particularly abundant constituent of proteins, cleavage with CNBr usually gives large peptides. Subsequent digestion by other enzyme(s) is required to produce smaller peptides (82).

Unfortunately, if proteins can not be cleaved by specific cleavage reagents, they must be cleaved by non-specific reagents. Pepsin (maximum activity around pH 3.0), thermolysin, and partial acid hydrolysis are useful for this purpose. However, using non-specific proteases it is difficult to relate the disulfide-bridged peptides to specific segments of the protein simply by determining their molecular masses by MS. In such a case, multiple-step Edman degradation and/or MS/MS is a useful tool for characterization of the disulfide-bridged peptides.

3. Purification of Cystine-Contai

Following digestion of a

original disulfide bonds remain inta

to identify, by N-terminal amino

disulfide bonded peptide. Reverse

acetonitrile in 0.1% trifluoroacetic

rapid and employing conditions

usually detected by UV at 215 nm

from high-mass proteins is still a

HPLC columns are used, because

hydrolysis.

An alternative method

identification of disulfide-bonded

based on the chemical modificatio

or paper electrophoresis is carried

performic acid vapour. After th

repeated in the perpendicular di

migrate along a line at 45° to ea

formed cysteic acid peptides migra

then eluted and analyzed by amino

method is diagonal electrophoresis

3. Purification of Cystine-Containing Peptides

Following digestion of a protein to give a mixture of peptides in which the original disulfide bonds remain intact, the next step is to purify each of these peptides and to identify, by N-terminal amino acid and if necessary, partial sequence analysis, each disulfide bonded peptide. Reversed-phase HPLC under acidic conditions (a gradient of acetonitrile in 0.1% trifluoroacetic acid, for example), is ideal for this purpose, being both rapid and employing conditions favoring stability of disulfide bonds. Peptides are usually detected by UV at 215 nm with acceptable sensitivity. The purification of digests from high-mass proteins is still a challenging task even though microbore or capillary HPLC columns are used, because so many peptide fragments might be produced after hydrolysis.

An alternative method to reversed-phase HPLC for the separation and identification of disulfide-bonded peptides is diagonal electrophoresis. This method is based on the chemical modification of cysteinyl residues to cysteic acids. The thin-layer or paper electrophoresis is carried out in the horizontal direction, and then exposed to the performic acid vapour. After the performic acid is removed, the electrophoresis is repeated in the perpendicular direction. All peptides unaffected by performic acid migrate along a line at 45° to each direction of the electrophoresis, while the newly formed cysteic acid peptides migrate off the diagnal. Peptides containing cysteic acid are then eluted and analyzed by amino acid composition or sequences. An example of this method is diagonal electrophoresis map of a peptic digests of α-chymotrypsin.

4. Identification of Disulfide Pep

The conventional approach
after HPLC separation is very ted
made by a comparison between ch
the disulfide bonds. Peptide peaks
S-S; for each such peak, two new p
S, it is likely that there will be at le
However, this strategy depends to
before and after reduction, and is
reduced species are alkylated and
Performic acid oxidation (36) cou
modification has the disadvantag
residues will also have their elutio
half-cystine residues may be identi

The advent of modern mass
disulfide bonds in proteins more
bombardment (FAB), matrix-as
electrospray ionization (ESI) have
mass spectrometry, as shown in Fi
their unique masses and by compa
the signal corresponding to the
corresponding to the respective thi

4. Identification of Disulfide Peptides

The conventional approach for the identification of disulfide-containing peptides after HPLC separation is very tedious and time consuming. The identification can be made by a comparison between chromatograms of peptides before and after reduction of the disulfide bonds. Peptide peaks that disappear upon reduction are presumed to contain S-S; for each such peak, two new peaks are formed. If a single peptide has an internal S-S, it is likely that there will be at least a small shift in the elution position upon reduction. However, this strategy depends too much on the variation of hydrophobicity of peptides before and after reduction, and is therefore not always reliable. More effectively, the reduced species are alkylated and re-purified by HPLC followed by Edman sequencing. Performic acid oxidation (36) could also be used to identify peptides with S-S, but this modification has the disadvantage that peptides containing methionine and trytophan residues will also have their elution positions altered. Alternatively, peptides containing half-cystine residues may be identified colorimetrically (52-54).

The advent of modern mass spectrometric techniques has made the task of locating disulfide bonds in proteins more tractable (10, 83). In the last decade, fast atom bombardment (FAB), matrix-assisted laser desorption/ionization (MALDI), and electrospray ionization (ESI) have widely been used for disulfide bond assignment. In mass spectrometry, as shown in Figure 1.10, disulfide linked peptides are identified from their unique masses and by comparison with the spectrum of reduced samples in which the signal corresponding to the S-S linked peptide(s) is replaced by two signals corresponding to the respective thiol peptide components, if *inter*-bridged, or shifted by

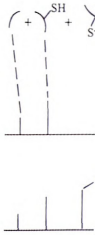


Figure 1.10. Current protocol

two mass units (dithiol) if *intra*-b

for identifying peptides because it

that becomes important when th

hydrolysis of large proteins wh

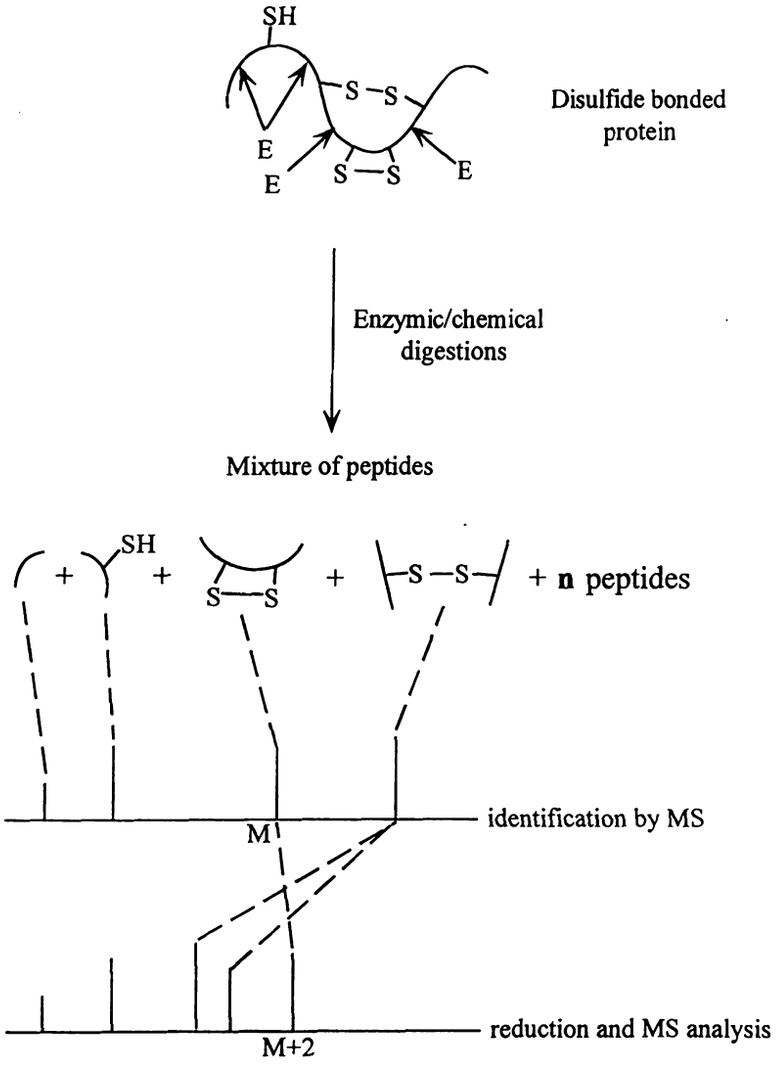


Figure 1.10. Current protocol for disulfide bond assignment by mass spectrometry.

two mass units (dithiol) if *intra*-bridged. Mass spectrometry is a particularly good tool for identifying peptides because it does not require rigorously purified peptides, a feature that becomes important when the amount of protein is small or when investigating hydrolysis of large proteins where purification of peptides to homogeneity may be

difficult. Furthermore, modern r
MALDI-PSD, FAB/CID/MS/MS,
provides confirmation of a specific

Takao *et al* (84) first re
disulfide-containing peptides from
be determined. Almost at the s
strategy for the recognition of dis
became an accepted method and
inhibitor of amylase (86) and huma
group subsequently demonstrated
ribonuclease A and lysozyme.
methodologies based on mass spe
containing peptides (39, 83).

Peptide-mapping by ESI (
increasingly in the characterization
used here is exactly same as in
sensitivity (low picomole to high
HPLC has made the purification
MALDI-MS provides the capacity
proteins and peptides. It is particu
mapping without the previous isolat
interchain disulfide-containing pep
excellent tool for the direct diagno

difficult. Furthermore, modern mass spectrometric techniques, such as ESI/MS/MS, MALDI-PSD, FAB/CID/MS/MS, permit the sequence determination of peptides which provides confirmation of a specific peptide segment in a protein.

Takao *et al* (84) first recognized the potential of FAB-MS for identifying disulfide-containing peptides from which the location of disulfide bonds in proteins may be determined. Almost at the same time, Morris and Pucci (85) reported the same strategy for the recognition of disulfide bond structure in insulin. Their protocol soon became an accepted method and successfully applied to proteins such as bacterial inhibitor of amylase (86) and human tissue inhibitor of metalloproteinases (87). Smith's group subsequently demonstrated the feasibility of the FAB/MS for S-S assignment in ribonuclease A and lysozyme, and further developed several supplementary methodologies based on mass spectrometry to facilitate the identification of disulfide-containing peptides (39, 83).

Peptide-mapping by ESI (88, 89) or MALDI-MS (90, 91) are being used increasingly in the characterization and identification of disulfide bridges. The strategy used here is exactly same as in the FAB-MS. However, they provide much more sensitivity (low picomole to high femtomole). Besides, the easy interface of ESI to HPLC has made the purification of the disulfide-containing peptides an easy task. MALDI-MS provides the capacity of rapidly analyzing small quantities of mixtures of proteins and peptides. It is particularly well suited to the determination of peptide mass mapping without the previous isolation or fractionation. The *in situ* disulfide cleavage of interchain disulfide-containing peptide during the analysis by MALDI-MS provides an excellent tool for the direct diagnosis of disulfide bonded peptides (90, 91). The poor

resolution of both ESI and MALI
identification of intrachain disulf
reduction of the S-S bond.

Bean and Carr (92) used
electronvolt) collision-activated di
masses up to 2,000 Da containing
cleavage at the disulfide bridge wi
fragment, giving a characteristic tr
32 u. These peaks are remarkabl
products of a disulfide bond cleava

5. Other Methodologies for S-S

Despite their many successe
locating disulfide bond linkages.
in mild alkaline solutions. The er
because the requirement to find a
always of priority and most of the
Another problem is that the cleava
the locations of S-S can be deduce
separated by the amino acids re
reactions. This is especially the c
example, trypsin is useful only if t
Even when the half-cystinyl residu

resolution of both ESI and MALDI instruments, however, could be a problem for the identification of intrachain disulfide bond which has only two mass units shift after reduction of the S-S bond.

Bean and Carr (92) used tandem mass spectrometry with high-energy (kilo-electronvolt) collision-activated dissociation (CAD) to examine a number of peptides of masses up to 2,000 Da containing a single *inter*-chain disulfide bridge. They observed cleavage at the disulfide bridge with retention of zero, one, or two sulfurs on the charged fragment, giving a characteristic triplet of peaks in the mass spectrum with separation of 32 u. These peaks are remarkably intense, so they can provide ready identification of products of a disulfide bond cleavage.

5. Other Methodologies for S-S Assignment

Despite their many successes, two problems remain in the traditional methods for locating disulfide bond linkages. One problem is the disulfide scrambling which occurs in mild alkaline solutions. The endeavor to minimize the scrambling are often fruitless because the requirement to find a specific enzyme to produce recognizable fragments is always of priority and most of the specific enzymes only work under alkaline conditions. Another problem is that the cleavage reaction does not always give peptides from which the locations of S-S can be deduced. This may be because half-cystinyl residues are not separated by the amino acids required by specific chemical or enzymatic cleavage reactions. This is especially the case for a protein containing adjacent cysteines. For example, trypsin is useful only if the half-cystinyl residues are separated by Lys or Arg. Even when the half-cystinyl residues are separated by amino acids suitable for cleavage

by specific enzymes, the reaction r
highly folded, as it is when disulfi
for the reduced and carboxymethy
94).

If a protein fails to produc
method, partial acid hydrolysis, wa
mixture of disulfide-containing pe
deduced (95). Partial acid hydro
temperature, etc.) under which on
hydrolysis is attractive as disulfid
Steric effects play less of a role i
under acid conditions, and becaus
enzymes. However, due to the lo
even a medium-sized protein (e. g.
of peptide fragments. Many overla
the cleavage pattern is not predicta
identify peptide fragments, com
assisted data interpretation have to

Another problem occurs whe
any enzyme. That is, it is unlike
cleave between the two cysteine re
one disulfide bond. The resultin
structure and there are two possible

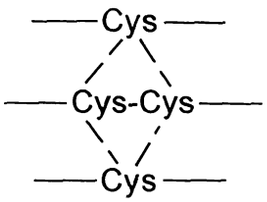
by specific enzymes, the reaction may not proceed at an appreciable rate if the protein is highly folded, as it is when disulfide bonds are intact. Hence, cleavage that occur readily for the reduced and carboxymethylated protein may not occur in the native protein (93, 94).

If a protein fails to produce desirable fragments after digestion, an alternative method, partial acid hydrolysis, was proposed to replace enzymatic digestion to produce a mixture of disulfide-containing peptides from which the disulfide connections may be deduced (95). Partial acid hydrolysis is performed under controlled conditions (time, temperature, etc.) under which only limited peptide fragments are formed. Partial acid hydrolysis is attractive as disulfide bonds are particularly stable in dilute or weak acids. Steric effects play less of a role in acid hydrolysis because most proteins are denatured under acid conditions, and because the catalyst, H_3O^+, is much smaller than proteolytic enzymes. However, due to the low specificity of partial acid hydrolysis, the digests of even a medium-sized protein (e. g., hen egg-white lysozyme) are a very complex mixture of peptide fragments. Many overlapped disulfide-containing peptides were produced and the cleavage pattern is not predictable. Therefore, even if mass spectrometry was used to identify peptide fragments, complicated chromatographic separation and computer-assisted data interpretation have to be applied to the definitive identification.

Another problem occurs when a protein has a -Cys-Cys- structure that is resistant to any enzyme. That is, it is unlikely, if impossible, to find a cleavage reagent that can cleave between the two cysteine residues and produce peptide fragments that only contain one disulfide bond. The resulting peptide always contains two disulfide bonds in its structure and there are two possible isomers (suppose they are *inter*-chain disulfides).

Different approaches have b
used the combination of organi
containing peptide and high-perfo
assign the disulfide structure of r
large investment of time and mat
strategy takes advantage of multi
(99) demonstrated that during sequ
no phenylthiohydantoin (PTH) can
presumably remains attached to th
the second Cys, a PTH appea
dehydroalanine, which can easi
sequencer. Zhang and Liang (100)
Huwentoxin-I that contains two ad

Gray (103-105) developed a
cysteine-rich small proteins. This
disulfide bonds and sequence anal
subjected to limited degree of re



Different approaches have been reported to solve the problem. Stults *et al* (96) used the combination of organic synthesis of the possible isomers of a disulfide-containing peptide and high-performance tandem mass spectrometry in their attempt to assign the disulfide structure of recombinant relaxin. This method obviously requires large investment of time and material, and is therefore not widely accepted. Another strategy takes advantage of multiple step Edman degradation (97-102). Nokihara *et al* (99) demonstrated that during sequence analysis of a peptide containing a disulfide bond, no phenylthiohydantoin (PTH) can be seen after cleavage of the first Cys residue, since it presumably remains attached to the second Cys by the disulfide bond. After cleavage of the second Cys, a PTH appears as a cysteine and a dithiothreitol adduct of dehydroalanine, which can easily be distinguished from other PTH-AAs on the sequencer. Zhang and Liang (100) used this strategy for studying the disulfide linkage of Huwentoxin-I that contains two adjacent cysteine residues.

Gray (103-105) developed a new approach for disulfide mapping of tightly folded, cysteine-rich small proteins. This approach utilized a combination of partial reduction of disulfide bonds and sequence analysis of alkylated peptides. In this strategy, a protein is subjected to limited degree of reduction under acidic conditions to form a series of

partially opened disulfide bonds
sulfhydryls, followed by sequencer
bond(s) that had been reduced. T
and has been successfully applied
However, one limit on the applica
fragments after digestion of the
demonstrated by Ishibashi *et al* (10

In all, disulfide linkage assig
Two troublesome problems, the d
Cys-containing protein and the pre
present. There is a considerable
techniques for disulfide bond assi
areas in studying protein structur
answers to the problems in dete
protein chemists.

III. Matrix-Assisted Laser Deso

Since the 1960s mass spectr
peptide and protein analysis, but
compounds real progress in the fie
ionization methods; plasma deso
bombardment mass spectrometry
techniques, matrix-assisted laser

partially opened disulfide bonds that are separated by HPLC. Alkylation of free sulfhydryls, followed by sequencer analysis, provided explicit assignment of the disulfide bond(s) that had been reduced. This approach can minimize disulfide bond scrambling and has been successfully applied to a number of proteins containing adjacent cysteines. However, one limit on the application of the method concerns peptide size. Analysis of fragments after digestion of the protein can hopefully overcome the drawback, as demonstrated by Ishibashi *et al* (106).

In all, disulfide linkage assignment in proteins is still challenging protein chemists. Two troublesome problems, the determination of disulfide bond arrangement in a -Cys-Cys-containing protein and the prevention of disulfide scrambling, remain bothersome at present. There is a considerable interest in the protein society in developing novel techniques for disulfide bond assignment. Mass spectrometry is one of the most active areas in studying protein structure. It is hoped that mass spectrometry can provide answers to the problems in determination of disulfide linkage that remain bothering protein chemists.

III. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry (MALDI-MS)

Since the 1960s mass spectrometry has been considered a promising technique for peptide and protein analysis, but owing to the size, polarity and low volatility of these compounds real progress in the field was only achieved after discovery of the desorption ionization methods; plasma desorption mass spectrometry (PDMS) (1) and fast atom bombardment mass spectrometry (FAB MS) (2). The introduction in 1988 of two new techniques, matrix-assisted laser desorption/ionization (MALDI) (3) and electrospray

ionization (ESI) (4), has extended
science far beyond the scope of PD
very different ideas and princip
difficulties with MS methods. i.
biomolecules, and shown consi
accurate determination of molecu
combination with biochemical tech

The first attempts to ge
desorption/ionization date back to
which can be desorbed and ion
molecules are analyzed, more ener
energy flux leads to pyrolytic
breakthrough came in the late
experimented with the use of a m
desorption/ionization of large, non
obtained when such samples were
large molar excess of an energy-ab
not strongly absorb at the wavelen
this technique was conducted with
nm. The ejected ions were analyze
demonstrated that the ions generat
67,000 daltons could be formed
Despite lack of an established io

ionization (ESI) (4), has extended the applicability of mass spectrometry in the biological science far beyond the scope of PDMS and FAB-MS. Although MALDI and ESI utilized very different ideas and principles, both new techniques have overcome the main difficulties with MS methods, i. e., the desorption and ionization of large and labile biomolecules, and shown considerable promise in characterizing biomolecules by accurate determination of molecular mass up to a few hundred thousand daltons. In combination with biochemical techniques, numerous applications have been achieved (5).

The first attempts to generate ions of organic molecules by direct laser desorption/ionization date back to early 1970s (107). However, the size of the analytes which can be desorbed and ionized was limited to ~1,000 daltons. When larger molecules are analyzed, more energy is required to desorb them. However, such a high energy flux leads to pyrolytic or photochemical decomposition of analytes. The breakthrough came in the late 1980s when Hillenkamp and Karas successfully experimented with the use of a matrix, nicotinic acid (3, 108). It was discovered that desorption/ionization of large, nonvolatile molecules such as proteins could be similarly obtained when such samples were irradiated by a laser after being codeposited with a large molar excess of an energy-absorbing "matrix" material, even though the analyte did not strongly absorb at the wavelength of the laser radiation. The first experiment with this technique was conducted with frequency-quadrupled Nd-YAG laser operated at 266 nm. The ejected ions were analyzed by a time-of-flight (TOF) mass spectrometer. It was demonstrated that the ions generated from proteins with molecular mass range 10,000-67,000 daltons could be formed and detected from picomolar amounts of analyte. Despite lack of an established ionization mechanism, the dependence on the matrix

naterial in the desorption/ionizati

laser desorption/ionization (MALI

powerful technique for accurate an

greater than 200,000 Da. With the

(PSD) and delayed extraction (DE

for structural elucidation of biopol

The majority of the work

instruments being available in th

(Chapter 2) was conducted on a

Biosystem Voyager Elite, which c

spectra for structural analysis; thi

extraction hardware and softwar

however, are based on the same pri

A. The Matrix

Matrix is the key to the suc

discovery of nicotinic acid as

compounds have been investigated

applicable. Among them, sinap

cinnamic acid, and α-cyano-4-hy

(111-113). It was considered imp

maximum of the matrix compound.

material in the desorption/ionization process gave rise to the terminology matrix-assisted laser desorption/ionization (MALDI). Today MALDI has developed into an extremely powerful technique for accurate and sensitive analysis of molecular masses extending to greater than 200,000 Da. With the introduction of two new techniques, post-source decay (PSD) and delayed extraction (DE), MALDI-TOF MS is also becoming an effective tool for structural elucidation of biopolymers.

The majority of the work that will be presented was performed on two MALDI instruments being available in the MSU/NIH mass spectrometry facility. Early work (Chapter 2) was conducted on a Vestec 2000. Others were gained on a PerSeptive Biosystem Voyager Elite, which can be operated in the reflectron mode to acquire PSD spectra for structural analysis; this instrument has recently been equipped with delayed extraction hardware and software for improving resolution. The two instruments, however, are based on the same principle.

A. The Matrix

Matrix is the key to the successful analysis of high-mass biomolecules. Since the discovery of nicotinic acid as the first matrix, several hundred different organic compounds have been investigated (109, 110). However, only a few of these are widely applicable. Among them, sinapinic acid (SA), 2,5-dihydroxybenzoic acid (DHB), cinnamic acid, and α-cyano-4-hydroxycinnamic acid (α-CHCA) are extremely useful (111-113). It was considered important that the laser wavelength match the absorption maximum of the matrix compound. All the matrices above have strong UV absorption in

the 320-350 nm range and can be

frequency-doubled Nd-YAG lase

include miscibility with the analy

required for the dissolution of the

chemical composition that promot

protons to the analyte, nonreactivi

as the low heat of sublimation and

The physicochemical even

and their ionization in MALDI ha

believed to serve to minimize samp

incident laser energy, resulting in t

gas phase. One model for the me

of matrix are induced to undergo

subsequent expansion of these n

isolated protein molecules into the

protein undergoes ionization throu

processes that remain to be explain

B. Sample Preparation

Proper sample preparation i

(118). Matrix solutions are typi

mixtures at a concentration of

the 320-350 nm range and can be used with a much cheaper nitrogen laser (337 nm) or frequency-doubled Nd-YAG lasers (355 nm). Other important matrix characteristics include miscibility with the analyte in the solid phase, solubility in the same solvents required for the dissolution of the analyte, vacuum compatibility (low vapor pressure), a chemical composition that promotes the ionization of matrix substituents that can donate protons to the analyte, nonreactivity with the analyte, and other physical properties such as the low heat of sublimation and a capacity to crystallize readily (114).

The physicochemical events leading to the transfer of proteins to the gas phase and their ionization in MALDI have not yet been fully elucidated (115). The matrix is believed to serve to minimize sample fragmentation from the laser beam by absorbing the incident laser energy, resulting in the sample and matrix molecules being ejected into the gas phase. One model for the mechanism (116, 117) assumes that the uppermost layers of matrix are induced to undergo a phase transition from the solid to the gas phase. The subsequent expansion of these matrix molecules into the vacuum drags the matrix-isolated protein molecules into the gas phase. During the transfer to the gas phase, the protein undergoes ionization through proton transfer reactions with the matrix by reaction processes that remain to be explained.

B. Sample Preparation

Proper sample preparation is also critical for successful analysis by MALDI-MS (118). Matrix solutions are typically prepared in water/ethanol or water/acetonitrile mixtures at a concentration of 5-10 μg/μl (~40 mM), depending on the solubility

properties of the matrix. The ana

solvent that is miscible with the m

trifluoroacetic acid (TFA) is freque

matrix and analyte solutions are

~5,000:1 to 10,000:1, and final co

respectively. An aliquot of ~1µl

(VT2000), and allowed to dry by

warm air, or under vacuum. Instea

the sample plate holder used in Vo

permit a much higher sample t

codeposits from solution with the

chamber by a probe and a vacuum

hand" in the Voyager Elite MAL

duration (1 to 10 ns) pulses of an U

To date, the chemistry of

the choice of matrix and method o

different matrices and solvents(11

122) affects the resolution and se

analyses under different conditions

properties of the matrix. The analyte is prepared at a concentration of ~0.1 μg/μl in a solvent that is miscible with the matrix solution (for peptides and proteins, aqueous 0.1% trifluoroacetic acid (TFA) is frequently used to maintain the acidity of the solution). The matrix and analyte solutions are mixed to give a final matrix:analyte molar ratio of ~5,000:1 to 10,000:1, and final concentration of ~40 nmol of matrix/ ~1pmol of analyte, respectively. An aliquot of ~1μl of the mixture is applied to MALDI-MS probe tip (VT2000), and allowed to dry by either ambient evaporation, heating with a stream of warm air, or under vacuum. Instead of the stainless steel probe tips used in the VT2000, the sample plate holder used in Voyager Elite system can accommodate 100 samples and permit a much higher sample throughput. During the drying process, the matrix codeposits from solution with the analytes. The sample is inserted into the vacuum chamber by a probe and a vacuum lock in the VT2000, and by an "automatic mechanical hand" in the Voyager Elite MALDI-TOF mass spectrometer, and irradiated with short duration (1 to 10 ns) pulses of an UV laser beam.

To date, the chemistry of "matrix assistance" remains incompletely understood; the choice of matrix and method of application is still empirical. Sample preparation by different matrices and solvents(119), matrix additive (120), and evaporation rate (121, 122) affects the resolution and sensitivity of MALDI. Optimal results require parallel analyses under different conditions.

C. Instrumentation

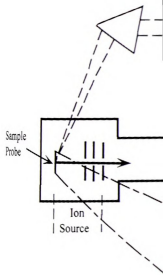
In MALDI (Figure 1.11) t
laser which is directed and focu
controlled by an attenuator and is
interaction of photons with the ma
of the cocrystallized sample/matr
spread of ions generated by MA
accelerating voltage or a reflectro
resolution. Experimentally, a stati
sample by application of a two-sta
with regard to a closely spaced ac
the same kinetic energy (assuming
toward a long (1-2 m) field-free tir

MALDI is typically used ir
because TOF is simple, cheap, and
MALDI. It has virtually no upper
which can produce very high m/z
its capacity to generate the entire
information, in contrast to a scan
traverse the flight tube (TOF) is
relationship: TOF = L/υ = L(m/2z
velocity, m is the mass of the ion,

C. Instrumentation

In MALDI (Figure 1.11) the cocrystal of sample/matrix is irradiated by a pulsed laser which is directed and focused by a prism and optical lens. The irradiance is controlled by an attenuator and is increased gradually until the threshold is reached. The interaction of photons with the matrix and protein results in the desorption and ionization of the cocrystallized sample/matrix from a metal surface. The initial kinetic energy spread of ions generated by MALDI is large, so either a linear TOF with a high accelerating voltage or a reflectron with an ion mirror (or both) is used to improve mass resolution. Experimentally, a static electric field is imposed upon ions generated from the sample by application of a two-stage high voltage (typically ± 25 kV) to the sample probe with regard to a closely spaced accelerating electrode. The ions are thus accelerated to the same kinetic energy (assuming the initial kinetic energy is zero) by the electric field toward a long (1-2 m) field-free time-of-flight (TOF) analyzer.

MALDI is typically used in conjunction with a time-of-flight mass analyzer (123), because TOF is simple, cheap, and well-suited to the pulsed nature of laser desorption in MALDI. It has virtually no upper mass range and is therefore compatible with MALDI, which can produce very high m/z ions. Another advantage of the TOF mass analyzer is its capacity to generate the entire spectrum from every single laser shot without losing information, in contrast to a scanning mass analyzer. The time required for ions to traverse the flight tube (TOF) is dependent on their masses and is described by the relationship: $TOF = L/\upsilon = L(m/2z\ eV)^{1/2}$, where L is the length of flight tube, υ is the ion velocity, m is the mass of the ion, and V is the acceleration potential. Thus, low-mass



Probe Ti

Figure 1.11. Scheme of matrix-as
spectrometer (MALDI-TOF MS)

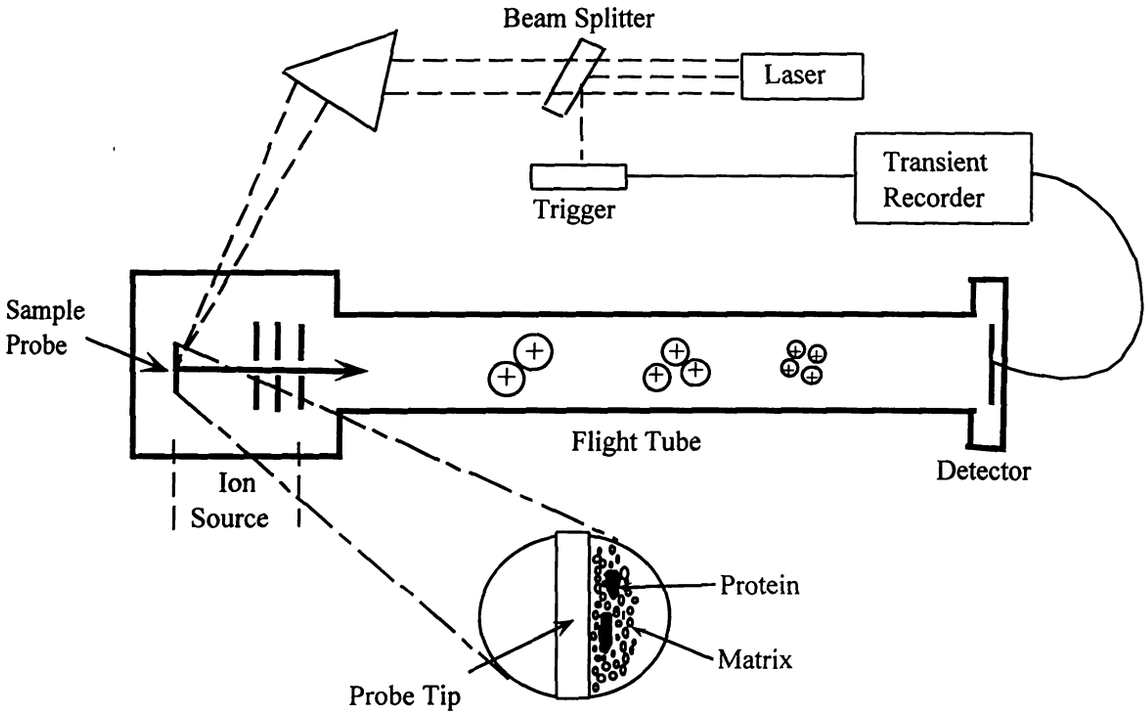


Figure 1.11. Scheme of matrix-assisted laser desorption/ionization time-of-flight mass spectrometer (MALDI-TOF MS)

ions have a shorter flight time t
spatially discrete individual ion pa
m/z ratio. A detector positioned
as each ion packet strikes it. A
yields a TOF spectrum. The di
pulse and common to all ions, an
proportional to $(m/z)^{1/2}$ and can b
time, tof) into a m/z ratios axis
highly efficient because all ions c
measured; they simply arrive at
recorded for a particular ion refle
ion source (such as the time/l
distributions). The poor resolving
with the mass of the ionized ma
isotopic masses are utilized in the
detect certain protein modificatio
masses. In spite of this, it is p
proteins with molecular masses
accuracy for proteins above 40 kD

D. Characteristic Features of M

One of the important fea
MALDI is a typical soft ionizatio

ions have a shorter flight time than heavier ions. They are separated into a series of spatially discrete individual ion packets, each traveling with a velocity characteristic of its m/z ratio. A detector positioned at the end of the field-free flight-tube produces a signal as each ion packet strikes it. A recording of the detector signal as a function of time yields a TOF spectrum. The difference between the start time, triggered by the laser pulse and common to all ions, and the arrival time of an individual ion at the detector is proportional to $(m/z)^{+1/2}$ and can be used to convert the x-axis of the spectrum (ion arrival time, tof) into a m/z ratios axis (a conventional mass spectrum). The tof analyzer is highly efficient because all ions of different m/z ratio arising from a single laser shot are measured; they simply arrive at the ion detector at different times. However, The tof recorded for a particular ion reflects many different initial conditions experienced in the ion source (such as the time/location of ion formation and initial kinetic energy distributions). The poor resolving power is reflected by peak broadening which increases with the mass of the ionized macromolecules. Therefore, average masses rather than isotopic masses are utilized in the MALDI-TOF measurement. This limits the capacity to detect certain protein modifications and protein sequence variations, especially at high masses. In spite of this, it is possible to achieve a mass accuracy of 0.1-0.01% for proteins with molecular masses between 1 and 40 kDa, and with somewhat poorer accuracy for proteins above 40 kDa (123).

D. Characteristic Features of MALDI-TOF MS

One of the important features of MALDI spectra is its simplicity (116, 118). MALDI is a typical soft ionization technique that yields little fragmentation. Unlike ESI

spectra which are dominated by r

in MALDI is the singly protona

also observed to varying degrees.

or analyte/alkali metal interaction

Compared to other ioniz

relatively tolerant to contamina

common in biological samples

unseparated, heterogeneous com

mixtures or synthetic reaction mi

The most striking feature

has been used to measure protei

500,000 Da (124). MALDI is als

Spectra with as little as 1 femt

reported (112). The analysis at l

obtained.

E. Novel Techniques in MALD

MALDI has been a superi

and peptides, and other biologica

because it is a soft ionization te

introduction of post-source deca

technique (125, 126) is based o

collision-induced decay taking pl

spectra which are dominated by multiply charged species, the predominant analyte signal in MALDI is the singly protonated species, though the doubly protonated molecule is also observed to varying degrees. Other adducts arising from analyte/matrix interactions or analyte/alkali metal interactions are of relatively low abundance.

Compared to other ionization techniques such as ESI or FAB, MALDI-MS is relatively tolerant to contaminants, such as buffers, salts, and denaturants which are common in biological samples (116). It has therefore the capacity to analyze unseparated, heterogeneous complex mixtures, such as enzymatic or chemical digest mixtures or synthetic reaction mixtures.

The most striking feature of MALDI-MS is its very large practical mass range. It has been used to measure proteins and glycoproteins with molecular weight as high as 500,000 Da (124). MALDI is also a very sensitive technique for protein characterization. Spectra with as little as 1 femtomole of protein applied on the probe tip have been reported (112). The analysis at low picomole to high femtomole range can be routinely obtained.

E. Novel Techniques in MALDI-TOF MS

MALDI has been a superior method for molecular mass determination of proteins and peptides, and other biological materials, but lacks the capacity for structural analysis, because it is a soft ionization technique. This deficiency has been overcome since the introduction of post-source decay (PSD) analysis of MALDI-generated ions. The PSD technique (125, 126) is based on mass analysis of product ions from unimolecular or collision-induced decay taking place in the field-free region between the ion source and

the reflectron. In contrast to a lir
energy analyzer to differentiate i
velocity in the linear mode. Ions
accelerated precursors are detecte
precursor in the linear mode of
because the product ions have l
reflectron mode of the TOF ins
PSD, like MS/MS techniques, ca
sized peptides (<2500 Da). Sinc
PSD has quickly evolved into a p
at the low picomole scale.

Another limitation of
distribution of ions in the MALD
nearly independent of mass and
the analyte. In addition, when
presumably lost by collision wit
energy dispersion. To attain hig
ion extraction. In contrast to
generated by the laser beam nea
potential, in delayed extraction, a
ionization and ion extraction eve
is field-free during the delay. F
repeller. Application of the ap

the reflectron. In contrast to a linear instrument, reflection instruments can be used as an energy analyzer to differentiate ions that are otherwise detected as species with the same velocity in the linear mode. Ions formed as a result of metastable decomposition of fully accelerated precursors are detected at the same arrival time (same apparent mass) as their precursor in the linear mode of TOF because they have the same velocity. However, because the product ions have lower kinetic energy, they can often be resolved in the reflectron mode of the TOF instrument by lowering potential of reflector. Therefore, PSD, like MS/MS techniques, can provide full or partial sequence information of medium sized peptides (<2500 Da). Since its commercialization about three years ago, MALDI-PSD has quickly evolved into a powerful technique for sequence determination of peptide at the low picomole scale.

Another limitation of traditional MALDI-TOF MS is the broad energy distribution of ions in the MALDI source. The initial velocity of desorbed analyte ions is nearly independent of mass and the initial kinetic energy is proportional to the mass of the analyte. In addition, when desorption occurs in a strong electric field, energy is presumably lost by collision with the neutral plume, resulting in further mass-dependent energy dispersion. To attain high resolution, a delayed extraction technique is used for ion extraction. In contrast to conventional MALDI instruments in which the ions generated by the laser beam near the surface of the sample probe are extracted by a dc potential, in delayed extraction, a short time delay (<300 ns) is inserted between the laser ionization and ion extraction event. The region between the repeller and extraction grid is field-free during the delay. Following the delay, a pulsed potential is applied to the repeller. Application of the appropriate pulse voltage provides the energy correction

necessary to simultaneously dete
initial energy. The initially less
pulse is applied and traverse a l
ions. An energy/spatial correcti
reach the detector plane simultar
accuracy, and the quality of MA
reducing chemical noise, and m
The potential of delayed extrac
peptides (127-130) as well as for

F. Application of MALDI-TO

Over the past few years,
yet available for macromolecula
achievements in protein, nucleoti

MALDI has been extens
structures of proteins derived fro
provides routine and reliable mea
typically have a high degree of
extraneous compounds has beer
malaria parasite. This remarkabl
now within the reach of mass spe

MALDI is an excellent me
modifications. Covalently modi

necessary to simultaneously detect all ions of the same mass/charge regardless of their initial energy. The initially less energetic ions are closer to the repeller at the time the pulse is applied and traverse a longer segment of the electric field than more energetic ions. An energy/spatial correction is thus provided such that all ions of the same m/z reach the detector plane simultaneously. Delayed extraction improves resolution, mass accuracy, and the quality of MALDI mass spectra by suppressing matrix background, reducing chemical noise, and minimizing the effect of laser intensity on performance. The potential of delayed extraction MALDI has been demonstrated for proteins and peptides (127-130) as well as for oligonucleotides (131).

F. Application of MALDI-TOF MS

Over the past few years, MALDI has become among the most powerful methods yet available for macromolecular characterization of living systems. A wide range of achievements in protein, nucleotide, and glycobiology have been attained in a short time.

MALDI has been extensively used for the determination of primary covalent structures of proteins derived from both natural and recombinant sources (132). MALDI provides routine and reliable means to analyzing tryptic digests and glycoproteins which typically have a high degree of heterogeneity (133). The high tolerance of MALDI to extraneous compounds has been used to study the degradation of hemoglobin by a malaria parasite. This remarkable achievement shows direct analysis of cell contents is now within the reach of mass spectrometry (134).

MALDI is an excellent method of choice for dealing with protein posttranslational modifications. Covalently modified protein N- and C- termini (135), disulfide bonds

(136, 137), phosphorylation (13
DNA interactions (141) have all b

A dramatic demonstratio
samples is shown with "protein
analysis of a mixture of pepti
degradation (142). Each of the
can be identified from the mass di

Another highly promisin
identification by MALDI of Coo
electrophoresis (143-145). This
protein spot identification as the
protein quantities in 2D gel sp
proteins, poor protein recoveries

In additional to the use
analysis of combinatorial librarie
et al (146) used MALDI-MS for
peptides isolated from support-b
monoclonal antibody was screen
peptides isolated were shown to p

The capacity for mass sp
lagged behind that of oligopeptid
last few years (147, 148), especi

(136, 137), phosphorylation (138), glycosylation (139), lipidation (140), and protein-DNA interactions (141) have all been studied by MALDI.

A dramatic demonstration of the ability of MALDI to analyze heterogeneous samples is shown with "protein ladder sequencing" which involves the simultaneous analysis of a mixture of peptide/proteins that have undergone a stepwise Edman degradation (142). Each of the fragments differs from the next by one amino acid and can be identified from the mass difference between successive peaks.

Another highly promising technique in the last three years is the unambiguous identification by MALDI of Coomassie-stained protein spots from two dimensional gel electrophoresis (143-145). This technique is superior to Edman microsequencing for protein spot identification as the latter often suffers from problems such as insufficient protein quantities in 2D gel spots, widespread occurrence of N-terminally blocked proteins, poor protein recoveries from gels, and other unknown factors.

In additional to the use of MALDI as a sequencing tool, its ability toward the analysis of combinatorial libraries has also been explored. As an early example, Keough *et al* (146) used MALDI-MS for the rapid sequence determination of biologically active peptides isolated from support-bound combinatorial peptide libraries. An anti-gp120 monoclonal antibody was screened against a hexapeptide library and six of the eight peptides isolated were shown to possess the exact recognition sequence for the antibody.

The capacity for mass spectrometric analysis of oligonucleotides has generally lagged behind that of oligopeptides. However, this position has changed markedly in the last few years (147, 148), especially with the development of new matrices (149, 150).

Investigation of a 50-mer is now

reports of the analysis of much lar

It is believed that MAL

biological studies. As is often

exciting breakthroughs may, in th

this writing.

Investigation of a 50-mer is now somewhat routine (151, 152) and there are individual reports of the analysis of much larger RNAs (153).

It is believed that MALDI-MS will play more and more important roles in biological studies. As is often true in a new and rapidly developing field, the most exciting breakthroughs may, in the end, occur in areas not even anticipated at the time of this writing.

IV. References

1. Thorgersen, D. F., Skowrons

Commun., **60**, 616-620(1974)

2. Barber, M., Bordoli, R. S., S

(1981).

3. Karas, M., and Hillenkamp, F

4. Meng, C. K., Mann, M., ar

Spectrometry and Allied Top

5. Burlingame, A. L., Boyd, R

(1996).

6. Carr, S. A., Roberts, G. D.,

Materials, C. N. McEwen a

(1990), Ch.3, p87.

7. Yan, S. C. B., Grinnell, B. W

8. Carr, S. A., and Biemann, K.

9. Creighton, T. E., in *Mechani*

(1992).

10. Hirayama, K., and Akashi, S

Edited by Matsuo, T., Capri

Sons Ltd, (1994).

11. Torchinsky, Y. M., in *Surfu*

Oxford, (1981).

IV. References

1. Thorgersen, D. F., Skowronski, R. P., and Macfarlane, R. D., *Biochem. Biophys. Res. Commun.*, **60**, 616-620(1974).
2. Barber, M., Bordoli, R. S., Sedgwick, R. D., and Tyler, A. N., *Chem. Commun.*, 325 (1981).
3. Karas, M., and Hillenkamp, F., *Anal. Chem.*, **60**, 2299-2301 (1988).
4. Meng, C. K., Mann, M., and Fenn, J. B., Proc. 36th ASMS Conference on Mass Spectrometry and Allied Topics, San Francisco, CA, (1988), p771.
5. Burlingame, A. L., Boyd, R. K., and Gaskell, S. J., *Anal. Chem.*, **68**, 599R-651R (1996).
6. Carr, S. A., Roberts, G. D., and Hemling, M. E., in *Mass Spectrometry of Biological Materials*, C. N. McEwen and B. S. Larsen (Editors), Marcel Dekker, New York, (1990), Ch.3, p87.
7. Yan, S. C. B., Grinnell, B. W., and Wold, F., *Trends Biochem. Sci.*, **14,** 264 (1989).
8. Carr, S. A., and Biemann, K., Methods Enzymol., 106, 29-58(1984).
9. Creighton, T. E., in *Mechanism of Protein Folding*, Edited by Pain, R. H., IRL Press, (1992).
10. Hirayama, K., and Akashi, S., in *Biological Mass Spectrometry: Present and Future*, Edited by Matsuo, T., Caprioli, R. M., Gross, M. L., and Seyama, Y., John Wiley & Sons Ltd, (1994).
11. Torchinsky, Y. M., in *Surfur in Protein;* (Metzler, D., Translation Ed.), Pergamon, Oxford, (1981).

12. Jocelyn, P. C., *Biochemistry o*

13. Jocelyn, P. C., in *Methods in*

W. B., Griffith, O. W. Inc., (1

14. Stark, G. R., and Stein, W. H.

15. Swoboda, G., and Hasselbac

(1973).

16. Oliveira, B., and Lamm, M. E

17. Gall, W. E., Cunningham, B.

M., *Biochemistry*, **7**, 1973-19

18. Lundblad, R. L., in *Techniqu*

19. Boyer, P. D., *JACS*, **76**, 4331

20. Kochkina, V. M., and Torchi

(1975).

21. Sakai, H., *Anal. Biochem.*, **26**

22. Yamuchi, T., Yamamoto, S.,

23. Gibson, Q. H., *J. Biol. Chem.*

24. Ellman, G. L., *Arch. Biochem*

25. Riddles, P. W., Blakeley, R.

(1983).

26. Gilbert, H. F., *Adv. Enzymol.*,

27. Huxtable, R. J., *Biochemistry*

28. Grassetti, D. R., and Murray,

12. Jocelyn, P. C., *Biochemistry of the SH Group*, Academic Press, London, (1974).

13. Jocelyn, P. C., in *Methods in Enzymol.*, Vol. 143, Academic Press, edited by Jakoby, W. B., Griffith, O. W. Inc., (1987).

14. Stark, G. R., and Stein, W. H., *J. Biol. Chem.*, **239**, 3755 (1964).

15. Swoboda, G., and Hasselbach, W., *Hoppe-Seyler's Z. Physiol. Chem.*, **354**, 1611 (1973).

16. Oliveira, B., and Lamm,,M. E., *Biochemistry*, **10**, 26-31 (1971).

17. Gall, W. E., Cunningham, B. A., Waxdal, M. J., Konigsberg, W. H., and Edelman, G. M., *Biochemistry*, **7**, 1973-1982 (1968).

18. Lundblad, R. L., in *Techniques in Protein Modification*, CRC Press, (1995).

19. Boyer, P. D., *JACS*, **76**, 4331 (1954).

20. Kochkina, V. M., and Torchinsky Yu, M., *Biochem. Biophys. Res. Communs.* **63**, 392 (1975).

21. Sakai, H., *Anal. Biochem.*, **26**, 387 (1968).

22. Yamuchi, T., Yamamoto, S., and Hayaishi, O., *J. Biol. Chem.*, **250**, 7127(1975).

23. Gibson, Q. H., *J. Biol. Chem.*, **248**, 1281 (1973).

24. Ellman, G. L., *Arch. Biochem. Biophys.*, **82**, 70 (1959).

25. Riddles, P. W., Blakeley, R. L., and Zerner, B., in *Methods in Enzymol.*, **91**, 49-60 (1983).

26. Gilbert, H. F., *Adv. Enzymol.*, **63**, 69-172 (1990).

27. Huxtable, R. J., *Biochemistry of Sulfur*, Plenum, New York, 1986.

28. Grassetti, D. R., and Murray, J. F., *Arch. Biochem. Biophys.*, **119**, 41 (1967).

29. Grassetti, D. R., and Murray,.

30. Stuchbury, T., Shipton, M., N

J. A. L., and Suschitzky, H., *E*

31. Janson, J.; and Ryden, L., *Pro*

and Applications. VCH Press,

32. Gruen, L. C., and Harrap, B. S

33. Gevondyan, N., M., Gevond

FEBS Letters, **255**(2), 265-26

34. Sun, Y., Smith, D. L., and Sh

35. Shipton, M., and Brocklehurs

36. Sun, Y., and Smith, D. L., *An*

37. Singh, R., Blattler, W. A., and

38. Feng, F., Bell, A., Dumas,

Conference on Mass Spectron

39. Sun, Y., Bauer, M. D., Keo

Biology, Vol: 61: Protein an

Chapman, J. R., Humana Pres

40. Zaluzec, E. J., Gage, D. A., a

366 (1994).

41. Thannhauser, T. W., Konish

(1984).

42. Thannhauser, T. W., McWhe

322 (1985).

29. Grassetti, D. R., and Murray., J. F. *J. chromatogr.* **41**, 121 (1969).

30. Stuchbury, T., Shipton, M., Norris, R., Malthouse, J. P. G., Brocklehurst, K., Herbert, J. A. L., and Suschitzky, H., *Biochem. J.*, **151**, 417 (1975).

31. Janson, J.; and Ryden, L., *Protein Purification: Principles, High Resolution Methods, and Applications,* VCH Press, (1993), p252.

32. Gruen, L. C., and Harrap, B. S., *Anal. Biochem.*, **42**, 377 (1971).

33. Gevondyan, N., M., Gevondyan, V. S., Gavrilyeva, E.,E., and Modyanov, N., N., *FEBS Letters*, **255**(2), 265-268 (1989).

34. Sun, Y., Smith, D. L., and Shoup, R. E., *Anal. Biochem.*, **197**, 69-76 (1991).

35. Shipton, M., and Brocklehurst, K., *Biochem. J.*, **167**, 799 (1977).

36. Sun, Y., and Smith, D. L., *Anal. Biochem.*, **172**, 130-138 (1988).

37. Singh, R., Blattler, W. A., and Collinson, A. R.,, *Anal. Biochem.*, **213**, 49-56(1993).

38. Feng, F., Bell, A., Dumas, F., and Konishi, Y., Proceedings of the 38th Annual Conference on Mass Spectrometry and Allied Topics, Tucson, AZ, (1990), p273.

39. Sun, Y., Bauer, M. D., Keough, T. W., and Lacey, M. P., *Methods in Molecular Biology,* Vol: 61: Protein and Peptide Analysis by Mass Spectrometry, edited by: Chapman, J. R., Humana Press Inc., Totowa, NJ, (1996).

40. Zaluzec, E. J., Gage, D. A., and Watson, J. T., *J. Amer. Soc. Mass Spectrom.*, **5**, 359-366 (1994).

41. Thannhauser, T. W., Konishi, Y., and Scheraga, H. A., *Anal. Biochem.*, **138**, 181 (1984).

42. Thannhauser, T. W., McWherter, C. A., and Scheraga, H. A., *Anal. Biochem.*, **149**, 322 (1985).

43. Hirose, M., Takahashi, N., Oe

44. Makinen, A. L., and Nowak, T

45. Wang, Z. X., Preiss, B., and T

46. Mak, A. S., and Jones, B. L.,

47. Toyo'oka, T., and Imai, K., *A*

48. Toyo'oka, T., and Imai, K., *A*

49. Sueyoshi, T., Miyata, T., I

(Tokyo), **97**, 1811 (1985).

50. Toyo'oka, T., and Imai, K., *A*

51. Kobayashi, T., Kudo, I., Kar

Biochem.(Tokyo), **98**, 1017 (

52. Chin, C. C. Q., and Wold, F.,

53. Kirley, T. L., *J. Biol. Chem.*,

54. Kirley, T. L., *Anal. Biochem.*

55. Pardo, J. P., and Slayman, C.

56. Miyanishi, T., and Borejdo, J

57. Bishop, J. E., Squier, T. C.,

(1988).

58. Sturrock, E. D., Yu, X. C., W

9560-9566 (1996).

59. Janson, J.; and Ryden, L., *Pr*

and Applications, VCH Press

43. Hirose, M., Takahashi, N., Oe, H., and Doi, E., *Anal. Biochem.*, **168**, 193-201 (1988).

44. Makinen, A. L., and Nowak, T., *J. Biol. Chem.*, **264**, 12148 (1989).

45. Wang, Z. X., Preiss, B., and Tsou, C. L., *Biochemistry*, **27**, 5095 (1988).

46. Mak, A. S., and Jones, B. L., *Anal. Biochem.*, **84**, 432 (1978).

47. Toyo'oka, T., and Imai, K., *Anal. Chem.*, **56**, 2461-2464 (1984).

48. Toyo'oka, T., and Imai, K., *Analyst*, **109**, 1003 (1984).

49. Sueyoshi, T., Miyata, T., Iwanaga, S., Toyo'oka, T., and Imai, K., *J. Biochem. (Tokyo),* **97**, 1811 (1985).

50. Toyo'oka, T., and Imai, K., *Anal. Chem.*, **57**, 1931 (1985).

51. Kobayashi, T., Kudo, I., Karasawa, K., Mizushima, H., Inoue, K., and Nojima, S., *J. Biochem.(Tokyo)*, **98**, 1017 (1985).

52. Chin, C. C. Q., and Wold, F., *Anal. Biochem.*, **214**, 128-134 (1993).

53. Kirley, T. L., *J. Biol. Chem.*, **264**, 7185-7192 (1989).

54. Kirley, T. L., *Anal. Biochem.*, **180**, 231-236 (1989).

55. Pardo, J. P., and Slayman, C. W., *J. Biol. Chem.*, **264**, 9373 (1989).

56. Miyanishi, T., and Borejdo, J., *Biochemistry*, **28**, 1287 (1989).

57. Bishop, J. E., Squier, T. C., Bigelow, D. J., and Inesi, G., *Biochemistry*, **27**, 5233 (1988).

58. Sturrock, E. D., Yu, X. C., Wu, Z., Biemann, K., and Riordan, J. F., *Biochemistry*, **35**, 9560-9566 (1996).

59. Janson, J.; and Ryden, L., *Protein Purification: Principles, High Resolution Methods, and Applications,* VCH Press, (1993), p252.

60. Kalef, E.; Walfish, P. G.; and

61. Zaluzec, E. J., Gage, D. A.,

Conference on Mass Spectror

62. Catsimpoolas, N., and Wood,

63. Degani, Y.; and Patchornik, A

64. Stark, G. R., *Methods in Enz*

65. Biemann, K., in *Methods in*

(1993), p119.

66. Schaffer, M. H., and Stark, G

(1976).

67. Darbre, A., *Practical Protei*

p149.

68. Papayannopoulos, I.A., and F

69. Mahboub, S., Richard, C., De

(1986).

70. Altamirano, M. M., Plumbrid

(1992).

71. Tang, C., Yuksel, K. U., Jac

283, 12-19 (1990).

72. Nefsky, B., and Bretscher, A

73. Minard, P., Desmadril, M.,

Biochem., **185**, 419-423 (198

60. Kalef, E.; Walfish, P. G.; and Gitler, C., *Anal. Biochem.*, 212,325-334(1993)

61. Zaluzec, E. J.; Gage, D. A.; and Watson, J. T., in *Proceeding of the 42nd ASMS Conference on Mass Spectrometry and allied Topics,* Chicago, (1994), p20.

62. Catsimpoolas, N., and Wood, J. L., *J. Biol. Chem.*, **241**, 1790 (1966).

63. Degani, Y.; and Patchornik, A., *Biochemistry* 13, 1 (1974).

64. Stark, G. R., *Methods in Enzymology,* vol. 47, 129-132 (1977).

65. Biemann, K., in *Methods in Protein Sequence Analysis*, Plenum Press, New York, (1993), p119.

66. Schaffer, M. H., and Stark, G. R., *Biochem. Biophys. Res. Commun.*, **71**, 1040-1047 (1976).

67. Darbre, A., *Practical Protein Chemistry--A Handbook,* John Wiley & Sons, (1985), p149.

68. Papayannopoulos, I.A., and Biemann, K., *Protein Sci.* **1**, 278-288 (1992).

69. Mahboub, S., Richard, C., Delacourte, A., and Han, K., *Anal. Biochem.*, **154**, 171-182 (1986).

70. Altamirano, M. M., Plumbridge, J. A., and Calcagno, M. L., *Biochem.*, **31**, 1153-1158 (1992).

71. Tang, C., Yuksel, K. U., Jacobson, T. M., and Gracy, R. W., *Arch. Biochem. Phys.*, **283**, 12-19 (1990).

72. Nefsky, B., and Bretscher, A., *Proc. Natl. Acad. Sci. USA*, **86**, 3549-3553 (1989).

73. Minard, P., Desmadril, M., Ballery, N., Perahia, D., and Mouawad, L., *Eur. J. Biochem.*, **185**, 419-423 (1989).

74. Sanger, F., *Nature (london)*, 1

75. Ryle, A. P., and Sanger, F., *B*

76. Ryle, A. P., Sanger, F., Smith

77. Lumper, L., and Zahn, H., *Ad*

78. Spackman, D. H., Stein, W. H

79. Marti, T., Rosselet, S. J., Tit

(1987).

80. Staswick, P. E., Hermodson,

(1984).

81. Haeffner-Gormley, L., Paren

83 (1985).

82. Yazdanparast, R., Andrews,

2507-2513 (1987).

83. Smith, D. L., and Zhou, Z., *M*

84. Takao, T., Yoshida, M., Hon

Spectrom., **11**, 549-556 (198

85. Morris, H. R., and Pucci, P.,

86. Akashi, S.; and Hirayama, K

87. Williamson, R. A., Marston,

88. Krishnamurthy, T.; and Haue

89. Bauer, M.; Sun, Y.; Degenh

764 (1993).

74. Sanger, F., *Nature (london)*, **171**, 1025-1026 (1953).

75. Ryle, A. P., and Sanger, F., *Biochem. J.*, **60**, 535-540 (1955).

76. Ryle, A. P., Sanger, F., Smith, L. F., and Kitai, R., *Biochem. J.*, **60**, 541-556 (1955).

77. Lumper, L., and Zahn, H., *Adv. Enzymol.*, **27**, 199 (1965).

78. Spackman, D. H., Stein, W. H., and Moore, S., *J. Biol. Chem.*, **235**, 648-659 (1960).

79. Marti, T., Rosselet, S. J., Titani, K., and Walsh, K. A., *Biochemistry*, **26**, 8099-8109 (1987).

80. Staswick, P. E., Hermodson, M. A., and Nielsen, N. C., *J. Biol. Chem.*, **259**, 13431 (1984).

81. Haeffner-Gormley, L., Parente, L., and Wetlaufer, D. B., *Int. J. Pept. Protein Res.*, **26**, 83 (1985).

82. Yazdanparast, R., Andrews, P. C., Smith, D. L., and Dixon, J. E., *J. Biol. Chem.*, **262**, 2507-2513 (1987).

83. Smith, D. L., and Zhou, Z., *Methods in Enzymol.*, Vol. 193, 374-389 (1990).

84. Takao, T., Yoshida, M., Hong, Y. M., Aimoto, S., and Shimonishi, Y., *Biomed. Mass. Spectrom.*, **11**, 549-556 (1984).

85. Morris, H. R., and Pucci, P., *Biochem. Biophys. Res. Commun.* **126**, 1122 (1985).

86. Akashi, S.; and Hirayama, K., *Biomed. Environ. Mass Spectrom.*, **15**,541-546 (1988).

87. Williamson, R. A., Marston, F. A.; and Angal, S., *Biochem. J.*, **268**, 267-274 (1990).

88. Krishnamurthy, T.; and Hauer, C. C., *Biol. Mass Spectrom.* 23, 719-726 (1994).

89. Bauer, M.; Sun, Y.; Degenhardt, C.; and Kozikowski, B., *J. Protein Chem.* **12**, 759-764 (1993).

90. Patterson, S.; and Katta, V., *A*

91. Crimmins, D. L., Saylor, M.,

361 (1995).

92. Bean, M. F.; and Carr, S. A.,

93. Chung, D., Sairom, M. R., *a*

(1975).

94. Thannhauser, T. W., Konishi

115-119 (1987).

95. Zhou, Z., and Smith, D. L., *J*

96. Stults, J. T., Bourell, J. H., C

J. W., Griffin, P. R., Rinder

Spectrom., **19**, 655-664 (199

97. Fox, J. W., Elzinga, M., and

98. Serensen, H. H., Thomsen, J

19, 713-720 (1990).

99. Nokihara, K., Morita, N., Ya

(1992).

100. Zhang, D., and Liang, S., *J*

101. Krishnamurthy, T., Hauer,

Biol. Mass Spectrom., **23**, 71

102. Tang, Y., and Selsted, M. E

90. Patterson, S.; and Katta, V., *Anal. Chem.* **66**, 3727-3732 (1994).

91. Crimmins, D. L., Saylor, M., Rush, J., and Thoma, R. S., *Anal. Biochem.*, **226**, 355-361 (1995).

92. Bean, M. F.; and Carr, S. A., *Anal. Biochem.*, **201**, 216-226 (1992).

93. Chung, D., Sairom, M. R., and Li, C. H., *Int. J. Peptide Protein Res.*, **7**, 487-493 (1975).

94. Thannhauser, T. W., Konishi, Y., and Scheraga, H. A., *Methods Enzymol.*, Vol: 143, 115-119 (1987).

95. Zhou, Z., and Smith, D. L., *J. Protein Chem.*, **9**, 523-532 (1990).

96. Stults, J. T., Bourell, J. H., Canova-Davis, E., Ling, V. T., Laramee, G. R., Winslow, J. W., Griffin, P. R., Rinderknecht, E., and Vandlen, R. L., *Biomed. Environ. Mass Spectrom.*, **19**, 655-664 (1990).

97. Fox, J. W., Elzinga, M., and Tu, A. T., *Biochemistry*, **18**, 678-684 (1979).

98. Serensen, H. H., Thomsen, J., and Bayne, S., Biomed. And Environ. Mass Spectrom., **19**, 713-720 (1990).

99. Nokihara, K., Morita, N., Yamaguchi, M., and Watanabe, M., *Anal. Lett.*, **25**, 513-533 (1992).

100. Zhang, D., and Liang, S., *J. Protein Chem.*, **12**, 735-740 (1993).

101. Krishnamurthy, T., Hauer, C. R., Prabhakaran, M., Freedy, J. G., and Hayashi, K., *Biol. Mass Spectrom.*, **23**, 719-726 (1994).

102. Tang, Y., and Selsted, M. E., *J. Biol. Chem.*, **268**, 6649-6653 (1993).

103. Gray, W. R., Luque, F. A., C
Reyes, A., Alford, J., McInt
Biochemistry, **23**, 2796 (1984
104. Gray, W. R., *Protein Sci.*, **2**,
105. Gray W. R., *Protein Sci.*, **2**,
106. Ishibashi, J., Kataoka, H.,
Mizoguchi, A., Ishizaki, H., a
107. Posthumus, M. A., Kistema
Chem., **50**, 985-991 (1978).
108. Karas, M., Bachmann, D., I
Processes, **78**, 53-68 (1987).
109. Ehring, H., Karas, M., an
(1992).
110. Fitzgerald, M. C., Parr, G. I
111. Beavis, R. C., and Chait, B.
112. Strupat, K., Karas, M., and
111, 89-102 (1991).
113. Beavis, R. C., Chaudhary,
(1992).
114. Juhasz, P., Costello, C. E.,
409 (1993).
115. Chait, B. T.; and Kent. S. B

103. Gray, W. R., Luque, F. A., Galyean, R., Atherton, E., Sheppard, R. C., Stone, B. L., Reyes, A., Alford, J., McIntosh, M., Olivera, B. M., Cruz, L. J., and Rivier, J., *Biochemistry*, **23**, 2796 (1984).

104. Gray, W. R., *Protein Sci.*, **2**, 1732-1748 (1993).

105. Gray W. R., *Protein Sci.*, **2**, 1749-1755 (1993).

106. Ishibashi, J., Kataoka, H., Isogai, A., Kawakami, A., Saegusa, H., Yagi, Y., Mizoguchi, A., Ishizaki, H., and Suzuki, A., *Biochemistry*, **33**, 5912-5919 (1994).

107. Posthumus, M. A., Kistemaker, P. G., Meuzelaar, H. L. C., and Brauw, M. C., *Anal. Chem.*, **50**, 985-991 (1978).

108. Karas, M., Bachmann, D., Bahr, U., and Hillenkamp, F., *Int. J. Mass Spectrom. Ion Processes,* **78**, 53-68 (1987).

109. Ehring, H., Karas, M., and Hillenkamp, F., *Org. Mass Spectrom.*, **27**, 472-480 (1992).

110. Fitzgerald, M. C., Parr, G. R., and Smith, L. M., *Anal. Chem.*, **65**, 3204-3211 (1993).

111. Beavis, R. C., and Chait, B. T., *Rapid Commun. Mass Spectrom.*, **3**, 432-435 (1989).

112. Strupat, K., Karas, M., and Hillenkamp, F., *Int. J. Mass Spectrom. Ion Processes*, **111**, 89-102 (1991).

113. Beavis, R. C., Chaudhary, T., and Chait, B.T., *Org. Mass Spectrom.*, **27**, 156-158 (1992).

114. Juhasz, P., Costello, C. E., and Biemann, K., *J. Am. Soc. Mass Spectrom.*, **4**, 399-409 (1993).

115. Chait, B. T.; and Kent. S. B. H., *Science,* 257, 1885-1894 (1992).

116. Siuzdak, G., *Proc. Natl. Acad.*

117. Busch, K. L., *J. Mass Spectro*

118. Zaluzec, E. J.; Gage, D. A.; a

6, 109-123 (1995).

119. Cohen, S. L., and Chait, B. T.

120. Preston, L. M., Murray, K. K

550 (1993).

121. Vorm, O., and Roepstorff, P.

122. Vorm, O., and Mann, M., *J.*

123. Beavis, R. C., and Chait, B.

124. Hillenkamp, F., Karas, M.,

1198A (1991).

125. Kaufmann, R.; Kirsch, D.;

355-385 (1994).

126. Kaufmann, R.; Spengler, B

7, 902-910 (1993).

127. Brown, R. S., and Lennon,

128. Colby, S. M., King, T. B.,

868 (1994).

129. Vestal, M. L., Juhasz, P.,

1044-1050 (1995).

130. Whittal, R. M., and Li, L.,

116. Siuzdak, G., *Proc. Natl. Acad. Sci. USA,* **91**, 11290-11297 (1994).

117. Busch, K. L., *J. Mass Spectrom.*, **30**, 233-240 (1995).

118. Zaluzec, E. J.; Gage, D. A.; and Watson, J. T., *Protein Expression and Purification,* **6**, 109-123 (1995).

119. Cohen, S. L., and Chait, B. T., *Anal. Chem.*, **68**, 31-37 (1996).

120. Preston, L. M., Murray, K. K., and Russell, D. H., *Biol. Mass Spectrom.*, **22**, 544-550 (1993).

121. Vorm, O., and Roepstorff, P., and Mann, M., *Anal. Chem.*, **66**, 3281-3287 (1994).

122. Vorm, O., and Mann, M., *J. Am. Soc. Mass Spectrom.*, **5**, 955-958 (1994).

123. Beavis, R. C., and Chait, B. T., *Proc. Natl. Acad. Sci. USA*, **87**, 6873-6877 (1990).

124. Hillenkamp, F., Karas, M., Beavis, R. C., and Chait, B. T., *Anal. Chem.*, **63**, 1193-1198A (1991).

125. Kaufmann, R.; Kirsch, D.; and Spengler, B., *Int. J. Mass Spectrom. Ion Proc.* **131**, 355-385 (1994).

126. Kaufmann, R.; Spengler, B.; and Lutzenkirchen., *Rapid Commun. Mass Spectrom.*, **7**, 902-910 (1993).

127. Brown, R. S., and Lennon, J. J., *Anal. Chem.*, **67**, 1998-2003 (1995).

128. Colby, S. M., King, T. B., and Reilly, J. P., *Rapid Commun. Mass Spectrom.*, **8**, 865-868 (1994).

129. Vestal, M. L., Juhasz, P., and Martin, S. A., *Rapid Commun. Mass Spectrom.*, **9**, 1044-1050 (1995).

130. Whittal, R. M., and Li, L., *Anal. Chem.*, **67**, 1950-1954 (1995).

131. Juhasz, P., Roskey, M. T., Sm
A., *Anal. Chem.*, **68**, 941-946 (
132. Carr, S. A., in *Mass Spectro*
Carr, S. A., Eds.; Humana Pres
133. Billeci, T. M., and Stults, J. T
134. Goldberg, D. E., Slater, A.
Henderson, G. B., *J. Exp. Med*
135. Specht, B., Oudenampsen-K
and Spener, F., *J. Biotechnol.*,
136. Robertson, J. G., Adams, C
Villafranca, J. J., *Biochemistr*
137. Chang, J. Y., Schindler, P.
(1995).
138. Liao, P-C., Leykam, J., And
219, 9-20 (1994).
139. Burlingame, A. L., *Curr. Op*
140. Wilcox, M. D., Schey, K
Biophys. Res. Commun., **212**
141. Zhao, Y., and Chait, B. T.,
142. Chait, B. T., Wang, R., Bea
143. Patterson, S. D., and Aeber
144. Erdjument,-Bromage, H.,
Protein Sci., **3**, 2435-2446 (

131. Juhasz, P., Roskey, M. T., Smirnov, I. P., Haff, L. A., Vestal, M. L., and Martin, S. A., *Anal. Chem.*, **68**, 941-946 (1996).

132. Carr, S. A., in *Mass Spectrometry in the biological sciences*, Burlingame, A. L., Carr, S. A., Eds.; Humana Press: Totowa, NJ, (1996)

133. Billeci, T. M., and Stults, J. T., *Anal. Chem.*, **65**, 1709-1716 (1993).

134. Goldberg, D. E., Slater, A. F. G., Beavis, B. T., Chait, B. T., Cerami, A., and Henderson, G. B., *J. Exp. Med.*, **173**, 961 (1991).

135. Specht, B., Oudenampsen-Kruger, E., Ingendoh, A., Hillenkamp, F., Leaius, A. G., and Spener, F., *J. Biotechnol.*, **33**, 259-269 (1994).

136. Robertson, J. G., Adams, G. W., Medzihradszky, K. F., Burlingame, A. L., and Villafranca, J. J., *Biochemistry*, **33**, 11563-11575 (1994).

137. Chang, J. Y., Schindler, P., and Chatrenet, B., *J. Biol. Chem.*, **270**, 11992-11997 (1995).

138. Liao, P-C., Leykam, J., Andrew, P. C., Gage, D. A., and Allison, J., *Anal. Biochem.*, **219**, 9-20 (1994).

139. Burlingame, A. L., *Curr. Opin. Biotechnol.*, **7**, 4-10 (1996).

140. Wilcox, M. D., Schey, K. L., Busman, M., and Hildebrandt, J. D., *Biochem. Biophys. Res. Commun.*, **212**, 367-374 (1995).

141. Zhao, Y., and Chait, B. T., *Anal. Chem.*, **66**, 3723-3726 (1994).

142. Chait, B. T., Wang, R., Beavis, R. C., and Kent, S. B., *Science*, **262**, 89-92 (1993).

143. Patterson, S. D., and Aebersold, R., *Electrophoresis*, **16**, 1791-1814 (1995).

144. Erdjument,-Bromage, H., Lui, M., Sabatini, D. M., Snyder, D. H., and Tempst, P., *Protein Sci.*, **3**, 2435-2446 (1994).

145. Clauser, K. R., Hall, S. C., Sr
Epstein, L. B., and Burlingam
(1995).
146. Youngquist, R. S., Fuentes
Commun. Mass Spectrom., **8**,
147. Nordhoff, E., Kirpekar, F.,
(1996).
148. Limbach, P. A., Crain, P. F.
102 (1995).
149. Fitzgerald, M. C., Parr, C
(1993).
150. Wu, K. J., Shaler, T. A., and
151. Nordhoff, E., Kirpekar, F.
Kristiansen, K., Roepstorff, I
152. Nordhoff, E., Karas, M.,
Lezius, A., Muth, J., Meier,
(1995).
153. Kirpekar, F., Nordhoff, E.
Karas, M., and Hillenkamp,

145. Clauser, K. R., Hall, S. C., Smith, D. M., Webb, J. W., Andrews, L. E., Tran, H. M., Epstein, L. B., and Burlingame, A. L., *Proc. Natl. Acad. Sci. USA*. **92**, 5072-5076 (1995).

146. Youngquist, R. S., Fuentes, G. R., and Lacey, M. P., and Keough, T., *Rapid Commun. Mass Spectrom.*, **8**, 77-81 (1994).

147. Nordhoff, E., Kirpekar, F., and Roepstorff, P., *Mass Spectrom. Rev.*, **15**, 67-138 (1996).

148. Limbach, P. A., Crain, P. F., and McCloskey, J. A., *Curr. Opin. Biotechnol.*, **6**, 96-102 (1995).

149. Fitzgerald, M. C., Parr, G. R., and Smith, L. M., *Anal. Chem.*, **65**, 3204-3211 (1993).

150. Wu, K. J., Shaler, T. A., and Becker, C. H., *Anal. Chem.*, **66**, 1637-1645 (1994).

151. Nordhoff, E., Kirpekar, F., Karas, M., Cramer, R., Hahner, S., Hillenkamp, F., Kristiansen, K., Roepstorff, P., Lezius, A., *Nucleic Acids Res.*, **22**, 2460-2465 (1994).

152. Nordhoff, E., Karas, M., Cramer, R., Hahner, S., Hillenkamp, F., Kirpekar, F., Lezius, A., Muth, J., Meier, C., and Engles, J. W., *J. Mass Spectrom.*, **30**, 99-112 (1995).

153. Kirpekar, F., Nordhoff, E., Kristiansen, K., Roepstorff, P., Lezius, A., Hahner, S., Karas, M., and Hillenkamp, F., *Nucleic Acids Res.*, **22**, 3866-3870 (1994).

A STRATEGY TO LOCA

BY SPECIFIC CHEMICA

ASSISTED LASER DESO

MA

I. Introduction

Cysteine sulfhydryl grou

using its free sulfhydryl (-SH) g

cysteine proteases, as the chelati

reshuffling enzymes. Pinpointi

becomes an important strategy f

large majority of new protein se

cloned cDNA (or, in the case

fragments), differentiating betwe

(disulfide bonds) is still cumbers

conventional techniques, cannot

Classical approaches for

modification of free sulfhydryls

derivatized protein chemically

by HPLC fractionation of the p

(1, 2). Recently, peptide mapp

CHAPTER 2

A STRATEGY TO LOCATE CYSTEINE RESIDUES IN PROTEINS BY SPECIFIC CHEMICAL CLEAVAGE FOLLOWED BY MATRIX-ASSISTED LASER DESORPTION/IONIZATION TIME-OF-FLIGHT MASS SPECTROMETRY

I. Introduction

Cysteine sulfhydryl group (thiol) contributes to protein biological functions by using its free sulfhydryl (-SH) group in the active site for enzyme catalysis such as in cysteine proteases, as the chelating site for metal ions, and as the active site of disulfide reshuffling enzymes. Pinpointing the number and location of cysteine residues thus becomes an important strategy for determining protein structure. While it is true that the large majority of new protein sequences are now deduced from nucleotide sequences of cloned cDNA (or, in the case of prokaryotic proteins, from cloned genomic DNA fragments), differentiating between free cysteine residues (sulfhydryl groups) and cystines (disulfide bonds) is still cumbersome and tedious, because the cDNA technique, like other conventional techniques, cannot identify posttranslational modifications.

Classical approaches for localizing protein free sulfhydryl groups usually involve modification of free sulfhydryls before and after reduction of a protein, degradation of the derivatized protein chemically or enzymatically into smaller peptide fragments, followed by HPLC fractionation of the peptides and Edman degradation of the derivatized peptides (1, 2). Recently, peptide mapping by mass spectrometry of proteolytic digests before and

after HPLC fractionation also has

the sulfhydryl derivatization und

disulfide bond exchange is not a

be cleaved at both sides of cys

cysteine. Multiple-step enzymat

case which is tedious and requi

derivatizing reagents have to be

and to avoid the retention time c

an Edman sequencer. Althoug

chromatography of sulfhydryl

feasibility remains to be tested.

II. Chemical Cleavage at Cyst

Cleavage at cysteine re

cyanolysis of disulfide bonds

However, because of the severa

and the elimination of thiocyan

not acceptable to protein chemis

Jacobson *et al* (7) s

specifically cyanylates cysteine

side of the cyanylated cystein

amino-terminal peptide and a

(Figure 2.1). If a protein conta

after HPLC fractionation also has been reported (3-5). However, this procedure requires the sulfhydryl derivatization under a carefully controlled conditions so that sulfhydryl-disulfide bond exchange is not an issue. Secondly, this approach requires that a protein be cleaved at both sides of cysteine residues to give peptides that contain only one cysteine. Multiple-step enzymatic or chemical digestions usually are performed in this case which is tedious and requires a sample size on the order of nanomoles. Thirdly, derivatizing reagents have to be chosen to facilitate HPLC separation with UV detection and to avoid the retention time overlap of derivatized cysteine and other amino acids on an Edman sequencer. Although other approaches such as methods based on affinity chromatography of sulfhydryl groups have also been proposed (see chapter 1), their feasibility remains to be tested.

II. Chemical Cleavage at Cysteine Residues

Cleavage at cysteine residues of peptide chains under alkaline conditions after cyanolysis of disulfide bonds was first observed by Catsimpoolas and Wood (6). However, because of the several side reactions, such as the reversibility of the cyanolysis and the elimination of thiocyanylate, the yields were low and the cleavage reaction was not acceptable to protein chemists.

Jacobson *et al* (7) showed that 2-nitro-5-thiocyanobenzoic acid (NTCB) specifically cyanylates cysteine thiols. Subsequent cleavage occurs on the N-terminal side of the cyanylated cysteinyl residue under mildly alkaline conditions to form an amino-terminal peptide and a series of 2-iminothiazolidine-4-carboxylyl (ITC) peptides (Figure 2.1). If a protein contains **n** cysteine residues, the cleavage reaction results in the

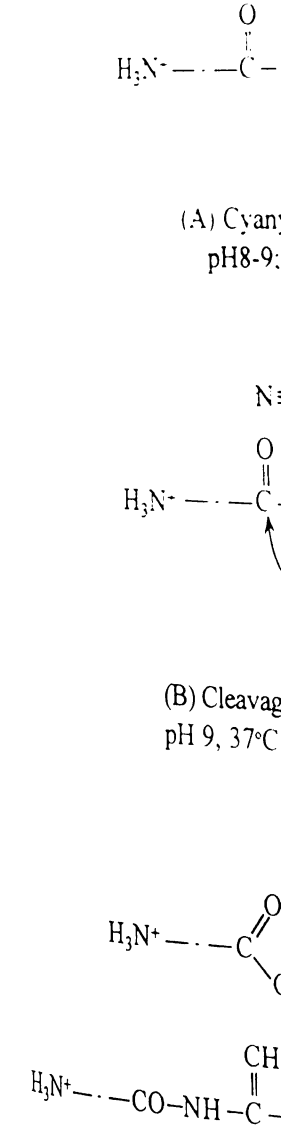


Figure 2.1. Reaction betwee
(NTCB), (A) Cyanylation ar
cleavage reaction under alkal

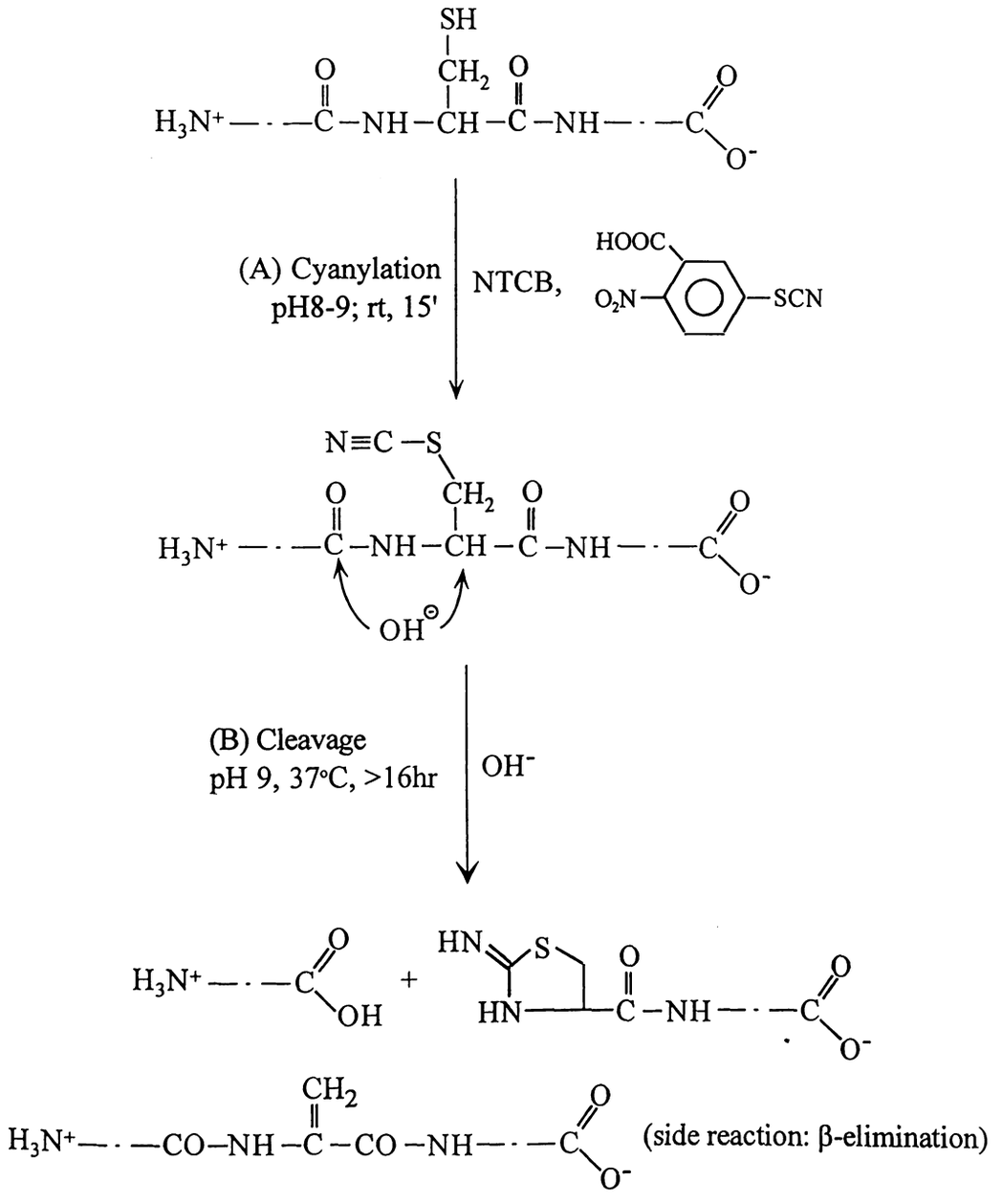


Figure 2.1. Reaction between cysteine residue and 2-nitro-5-thiocyanobenzoic acid (NTCB), (A) Cyanylation and (B) Cleavage reaction. β-Elimination competes with cleavage reaction under alkaline conditions.

formation of **n+1** peptide fragme

and location of cysteine residues

reaction can come to completion

proteins tested, Degani and Pat

alkaline conditions, competes wi

peptides showed marked variati

depending on the structural prop

excess of NTCB and a low tota

the side reaction of displacemen

thiol groups (8).

Other reagents for cyany

(9) and by Brocklehurst *et al* (1

NTCB, but their utility needs to

Jacobson *et al* (7) pro

mechanism implicates that hy

influence of rising pH on the rat

of OH^- is shown in Figure 2.2.

out the nucleophilic attack (th

reaction probably proceeds only

This generates a much more

cyclization and cleavage, withou

formation of **n+1** peptide fragments, mass analysis of the fragments indicates the number and location of cysteine residues. Although the original paper claimed that the cleavage reaction can come to completion with little side reactions for most of the peptides and proteins tested, Degani and Patchornik (8) found β-elimination, occurring also under alkaline conditions, competes with the cleavage as an adverse reaction. The cyanylated peptides showed marked variations in their tendency to undergo β-elimination reaction depending on the structural properties of the respective peptides. Experimentally, a large excess of NTCB and a low total concentration of thiol groups must be applied to avoid the side reaction of displacement of CN- from S-cyanocysteine residues by the unreacted thiol groups (8).

Other reagents for cyanylation of SH groups have been prepared by Wakselman (9) and by Brocklehurst *et al* (10, 11). These reagents may offer some advantages over NTCB, but their utility needs to be further examined.

Jacobson *et al* (7) proposed a mechanism for the cleavage reaction. This mechanism implicates that hydroxide ion catalyzes the cleavage, as shown by the influence of rising pH on the rate of cleavage reaction. A possible explanation of the role of OH^- is shown in Figure 2.2. The amide nitrogen is apparently too weak a base to carry out the nucleophilic attack (the pKa of a protonated amide is -4.0 or less), and the reaction probably proceeds only after attack of OH^- on the carbonyl carbon of the amide. This generates a much more basic nitrogen, which then can participate in concerted cyclization and cleavage, without formation of an acyliminothiazolidine intermediate.

Figure 2.2. A me

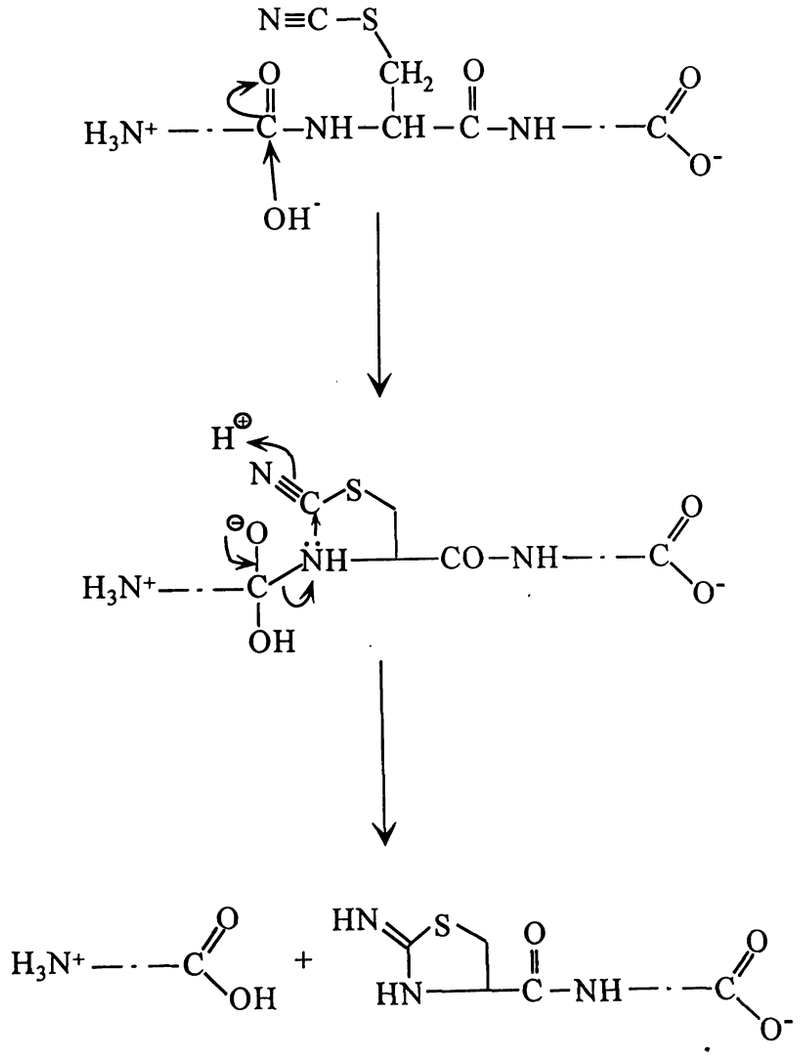


Figure 2.2. A mechanism for base catalyzed cleavage reaction.

One feature of the reacti

related to the easy displacement

nitrothiophenolate. The cyanyla

are unreactive to NTCB. The

sulfhydryls can thus be achieve

reaction without prior reduction

While potentially a very

used for sequence determinatio

by the iminothiazolidinecarboxy

to remove the ITC group fro

cleavage by NTCB has been u

number of side reactions that c

the reversibility of the cyanolys

Because cysteines are

usually produces large fragmen

polyacrylamide gel electrophor

using this approach are often co

Recently, Papayannop

spectrometry to sequence the

isolated from *Sarcophaga bul*

could be used to sequence pe

One feature of the reaction is that the reactivity of NTCB toward a thiol group is related to the easy displacement by the surfur nucleophile of the good leaving group *p*-nitrothiophenolate. The cyanylation is selective to free sulfhydryl groups, disulfide bonds are unreactive to NTCB. The selective cyanylation and subsequent cleavage of free sulfhydryls can thus be achieved in the presence of disulfide bonds by carrying out the reaction without prior reduction of the protein (7).

While potentially a very useful method, this cleavage reaction has seldom been used for sequence determination because all but the N-terminal peptide becomes blocked by the iminothiazolidinecarboxylyl (ITC) group. Thus far, there is no convenient method to remove the ITC group from the cleavage products (12). Another reason that Cys cleavage by NTCB has been used infrequently is that the reaction conditions cause a number of side reactions that compete with the desired cleavage reaction. These include the reversibility of the cyanolysis reaction and the β-elimination of thiocyanylate (8).

Because cysteines are relatively scarce in proteins, cleavage at these residues usually produces large fragments which can be mass mapped by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (13-32). However, peptide assignments using this approach are often complicated by its poor mass accuracy (error > 5%).

Recently, Papayannopoulos and Biemann (34) used CID tandem mass spectrometry to sequence the NTCB cleavage reaction products of a protease inhibitor isolated from *Sarcophaga bullata.* Their work demonstrated that mass spectrometry could be used to sequence peptides from the NTCB cleavage reaction in spite of the

blocked N-terminus. However,

and hence cysteine-rich proteins,

We developed a simple

location of both cystine and fre

chemical cleavage reaction desc

peptides by MALDI-TOF MS (

or denatured protein is first allo

chain at free sulfhydryl sites.

disulfide bonds by tris(2-carbo

reagent for this purpose since, u

thiol groups that may react wit

the resulting peptide fragments

deduce the number and locatio

the same cleavage reaction aft

on the number and location of

of the protein. The selectiv

sensitivity, speed, and mass a

approach. We also demonstra

to remove excess reagents an

spectra. Experimental cond

sulfhydryl cyanylation, and

Zetabind membranes.

blocked N-terminus. However, tandem MS is practically limited to low-mass peptides, and hence cysteine-rich proteins, because of the effective mass limit of CID (< 3000 Da).

We developed a simple and sensitive methodology to recognize the number and location of both cystine and free sulfhydryls in peptides and proteins using the specific chemical cleavage reaction described above, followed by mass mapping of the resulting peptides by MALDI-TOF MS (35). In these analyses, as shown in Figure 2.3, a peptide or denatured protein is first allowed to react with NTCB to selectively cleave the peptide chain at free sulfhydryl sites. The reaction mixture is then subjected to reduction of disulfide bonds by tris(2-carboxyethyl)phosphine hydrochloride (TCEP) (36), a useful reagent for this purpose since, unlike other common reducing reagents, it does not contain thiol groups that may react with NTCB under the reaction conditions. Mass mapping of the resulting peptide fragments by MALDI-MS provides information from which one can deduce the number and location of the free sulfhydryls. Parallel experiments involving the same cleavage reaction after reduction of disulfide bonds provide mass spectral data on the number and location of total cysteines, which can be used to confirm the sequence of the protein. The selectivity of the NTCB cleavage reaction combined with the sensitivity, speed, and mass accuracy of MALDI-TOF MS are attractive features of this approach. We also demonstrated the reactions performed on a Zetabind membrane used to remove excess reagents and other salts and buffers in order to improve the MALDI spectra. Experimental conditions are described for protein disulfide bond reduction, sulfhydryl cyanylation, and cleavage reactions performed either in solution or on Zetabind membranes.

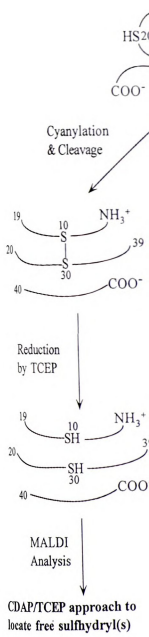


Figure 2.3. Chemical modif

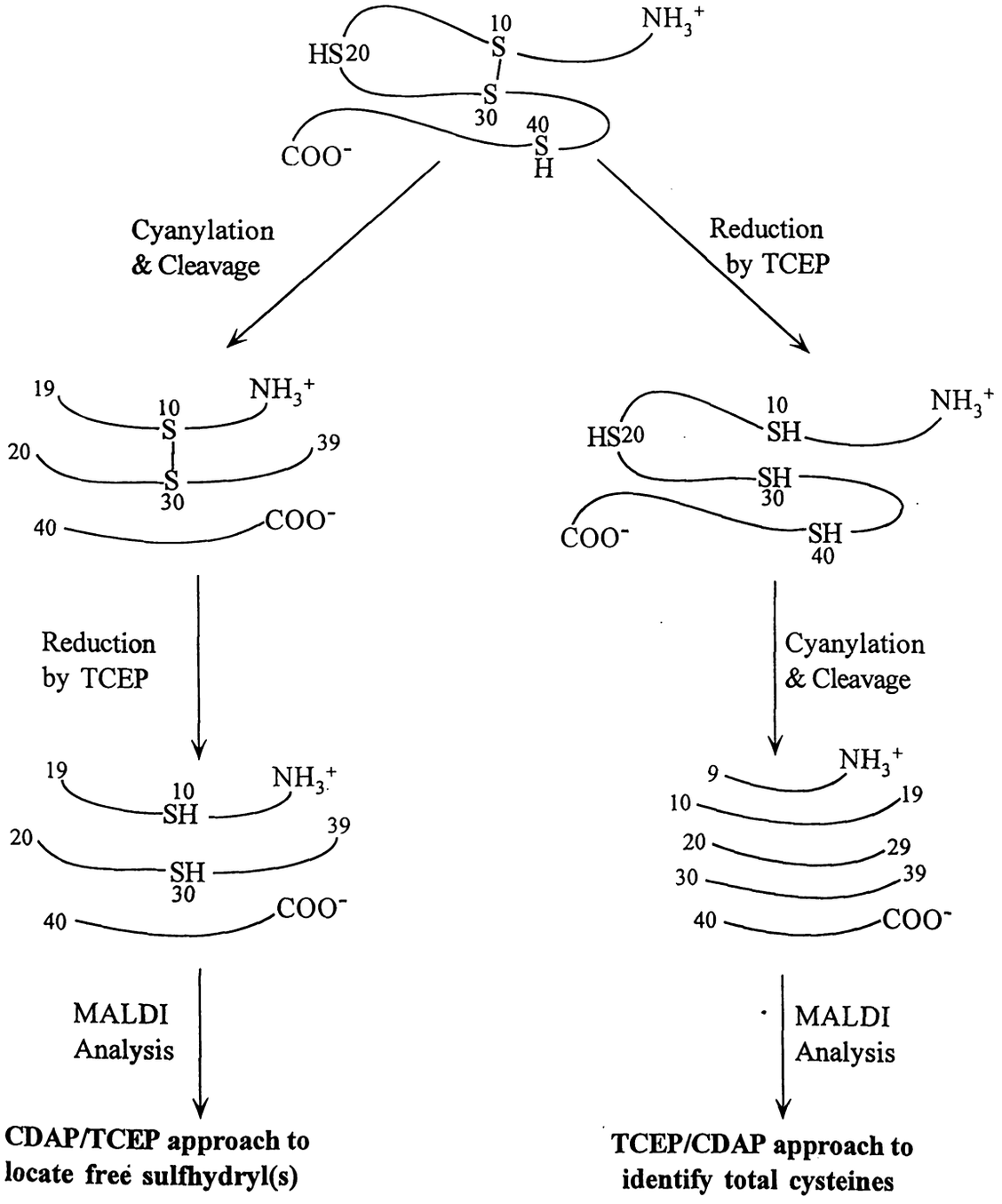


Figure 2.3. Chemical modification to identify free sulfhydryls and total cysteines.

MALDI-TOF MS

MALDI mass spectra w
Vestec Corp., Houston, TX) t
nitrogen laser (337 nm, 3-ns pul
26 kV. Data were acquired in
recorder with 2-ns resolution.
or internal calibration using star
(1 pmol of each standard) obt
experiments were performed u
Co., Milwaukee, WI) as the ma
(v/v) solution of acetonitrile/ac
peptide or protein samples and
allowed to air dry before beir
immobilized on an inert Zetab
mounted membrane containing
matrix/protein solution was a
acetonitrile/0.1% TFA solutio
introduction into the mass spec

III. Experimental Section

MALDI-TOF MS

MALDI mass spectra were obtained on a Vestec LaserTec Research (VT 2000, Vestec Corp., Houston, TX) time-of-flight (TOF) mass spectrometer equipped with a nitrogen laser (337 nm, 3-ns pulse). The accelerating voltage in the ion source was set to 26 kV. Data were acquired in the linear mode of operation, using a 500-MHz transient recorder with 2-ns resolution. Time-to-mass conversion was achieved by either external or internal calibration using standards of bradykinin (m/z 1061.2) and insulin (m/z 5734.5) (1 pmol of each standard) obtained from Sigma Chemical Co. (St. Louis, MO). All experiments were performed using α-cyano-4-hydroxycinnamic acid (Aldrich Chemical Co., Milwaukee, WI) as the matrix. Saturated matrix solutions were prepared in a 50% (v/v) solution of acetonitrile/aqueous 0.1% TFA and mixed in equal volumes with the peptide or protein samples and applied to a stainless-steel probe tip. The mixture was allowed to air dry before being introduced to the mass spectrometer. For a protein immobilized on an inert Zetabind membrane, matrix solution was applied to a probe-mounted membrane containing the adsorbed protein or reaction products (37). The matrix/protein solution was allowed to dry under the saturated vapor of the 50% acetonitrile/0.1% TFA solution and cocrystallized on the membrane surface prior to introduction into the mass spectrometer.

Chemicals

Tris(2-carboxyethyl)phos

Rockford, IL. Guanidine hy

Biochemicals (Indianapolis,

MHRQEAVDCLKKFNARRKI

peptides and proteins, 2-nitro-5-

purchased from Sigma and use

NTCB was prepared in eith

guanidine-HCl solution and th

aqueous solution was prepared

base, and stored under N_2 in -

suitable concentration before us

Analysis of Free sulfhydryls f

Peptide or protein samp

M tris-HCl buffer (pH 8.0)

Cyanylation of the free sulfhy

excess of NTCB (over free sul

then raised to 9.0 by adding

protein was accomplished afte

bonds was performed at 37°C

Chemicals

Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) was purchased from Pierce, Rockford, IL. Guanidine hydrochloride was a product of Boehringer-Mannheim Biochemicals (Indianapolis, IN). Peptides DRVYIHPCHLLYYS and MHRQEAVDCLKKFNARRKLKGA were purchased from Bachem, California. Other peptides and proteins, 2-nitro-5-thiocyanobenzoic acid (NTCB), and Trizma buffer, were purchased from Sigma and used without further purification. The 0.01 M solution of NTCB was prepared in either 0.1 M tris-HCl buffer or the buffer made in 6 M guanidine-HCl solution and the pH of the solution was adjusted to pH 8.0. TCEP aqueous solution was prepared as 0.1 M stock solution and adjusted to pH 8.0 with tris base, and stored under N_2 in -20°C freezer. The stock solution was further diluted to a suitable concentration before use.

Analysis of Free sulfhydryls for Proteins in Solution (NTCB/TCEP procedure)

Peptide or protein samples (1 nmol to 100 pmol) were solubilized in 2~5 µl of 0.1 M tris-HCl buffer (pH 8.0) containing 6 N guanidine-HCl as denaturing agent. Cyanylation of the free sulfhydryl groups was achieved by adding a 5-10 fold molar excess of NTCB (over free sulfhydryls) and reacting at 37°C for 30 minutes. The pH was then raised to 9.0 by adding 3.0 M tris base. Cleavage of the cyanylated peptide or protein was accomplished after incubation at 37°C for 16 hours. Reduction of disulfide bonds was performed at 37°C for 30 minutes by providing a 5-10 fold molar excess of

TCEP reducing reagent (over d

volume of a 50% (v/v) acetonitri

Confirmation of Total Cysteine

The above protein sampl

TCEP (over disulfide bonds) at

protein. Then a 5-10 fold mol

modify the total cysteines (nati

with 3.0 M tris base and the mix

diluted 50-fold with a 50% (v/v

MS.

Microscale Analysis of Protei

Figure 2.4 illustrates a

treatment and processing of

membrane (0.45-μm pore siz

Products Inc., Denver, CO) wa

A sample containing 10

tris-HCl buffered (pH 8.0) 6 N

molar excess of NTCB solutio

The membrane-attached probe

water that served to prevent

TCEP reducing reagent (over disulfide bonds). Samples were diluted with a 50-fold volume of a 50% (v/v) acetonitrile/0.1% TFA solution prior to analysis by MALDI-MS.

Confirmation of Total Cysteines (TCEP/NTCB procedure)

The above protein samples were allowed to react with a 5-10 fold molar excess of TCEP (over disulfide bonds) at 37°C for 30 minutes to reduce the disulfide bonds in the protein. Then a 5-10 fold molar excess of NTCB (over total cysteines) was applied to modify the total cysteines (native and nascent sulfhydryls). The pH was adjusted to 9.0 with 3.0 M tris base and the mixture was incubated at 37°C for 16 hours. The sample was diluted 50-fold with a 50% (v/v) acetonitrile/0.1% TFA solution for analysis by MALDI-MS.

Microscale Analysis of Proteins Immobilized on an Inert Membrane

Figure 2.4 illustrates a schematic flowchart for various protocols of chemical treatment and processing of the sample prior to analysis by MALDI. A Zetabind membrane (0.45-μm pore size and 50-μm thickness, purchased from Life Science Products Inc., Denver, CO) was fixed to a probe tip as described previously (37).

A sample containing 10 to 15 pmol of protein (0.5-μl volume, dissolved in 0.05 M tris-HCl buffered (pH 8.0) 6 N guanidine-HCl) was applied to the membrane. A 50-fold molar excess of NTCB solution was then added to the protein on the membrane surface. The membrane-attached probe tip was put in a closed vessel containing a few drops of water that served to prevent drying of the reaction mixture droplet on the membrane

surface. The vessel was incuba
NTCB. The reaction mixture w
immersion in 0.01% NH_4OH fo
other salts.

After air drying, 1.0 µl o
HCl solution was applied to the
for 16 hours to promote cleava
50-fold molar excess (~1 µl,
applied to the membrane to re
37°C for another 30 minutes.
in 0.1% TFA aqueous solution

A 1.5-µl aliquot of a
acid in a 50% acetonitrile/0.1%
to resolubilize the peptides.
period) in a vessel saturated
solution so that the peptides
effectively resolubilized fro
subsequently cocrystallized
analyzed directly by MALDI-

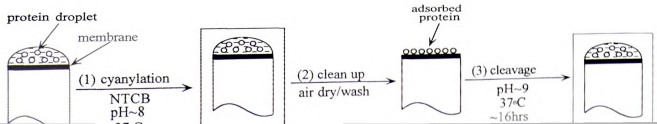
Parallel experiments w
membrane to confirm the tota

surface. The vessel was incubated at 37°C for 30 minutes to promote cyanylation with NTCB. The reaction mixture was allowed to dry on the membrane and then washed by immersion in 0.01% NH_4OH for 30 seconds to remove the excess NTCB reagents and other salts.

After air drying, 1.0 µl of 0.05 M sodium borate-buffered (pH 9.0) 6 N guanidine-HCl solution was applied to the membrane. The probe tip was incubated again at 37°C for 16 hours to promote cleavage in a closed vessel containing a few drops of water. A 50-fold molar excess (~1 µl, in this case) of TCEP (over disulfide bonds) was then applied to the membrane to reduce the disulfide bonds. The probe tip was incubated at 37°C for another 30 minutes. After air drying, the membrane was washed by immersion in 0.1% TFA aqueous solution for 30 seconds to remove the reagents and buffer salts.

A 1.5-µl aliquot of a saturated matrix solution of α-cyano-4-hydroxycinnamic acid in a 50% acetonitrile/0.1% TFA aqueous solution was applied to the dried membrane to resolubilize the peptides. The solution was allowed to dry slowly (over one hour period) in a vessel saturated by the vapor of the 50% acetonitrile/0.1% TFA aqueous solution so that the peptides and proteins immobilized on the membrane could be effectively resolubilized from the membrane surface by the matrix solution and subsequently cocrystallized with the matrix. The resulting peptide mixtures were analyzed directly by MALDI-MS (37).

Parallel experiments with the TCEP/NTCB procedure also were performed on the membrane to confirm the total number of cysteines in the protein.


protein droplet
membrane
(1) cyanylation
NTCB
pH~8
(2) clean up
air dry/wash
adsorbed
protein
(3) cleavage
pH~9
37°C
~16hrs

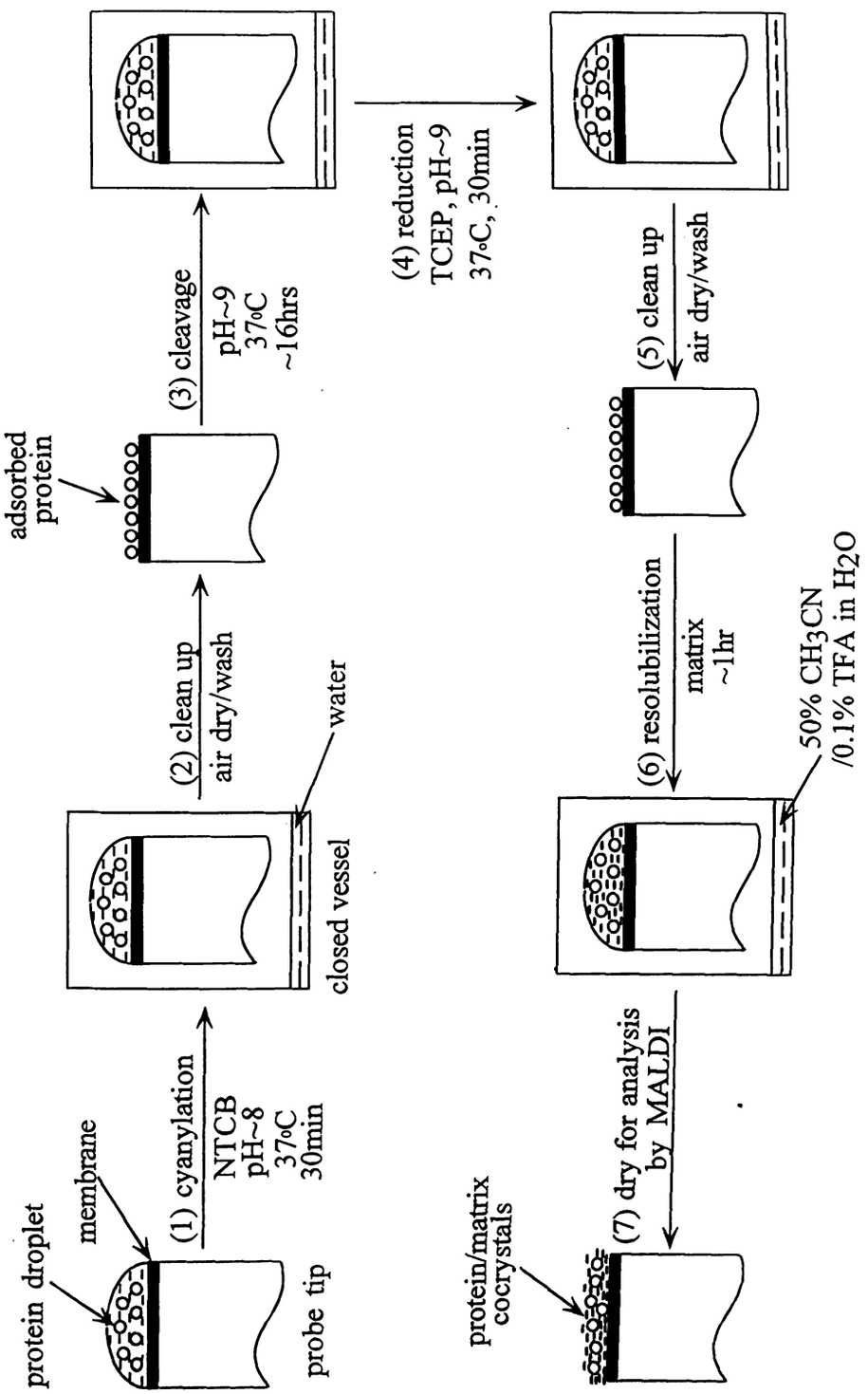


Figure 2.4. Conceptional scheme for microscale analysis of proteins immobilized on an inert membrane.

IV. Results and Discussion

A. Cyanylation and Cleavage

Preliminary experiments by MALDI-MS were conducte amino acid sequences and cy expected masses of their expect

Typical MALDI mass cleavage products, obtained by under the described experim cyanylation, a cyano group rep the original peptide (Figure 2.5 the observed value exactly ma at 37°C for 16 hours, the cyany 2.5B), two of them are cleava respectively. The peak at m/z with the cleavage reaction. T the mass shift from the cyany

Likewise, Figure 2.6 s remaining three peptides. T peptide mixtures. Most of th such as the cyanylated/uncl

IV. Results and Discussion

A. Cyanylation and Cleavage of Peptides by NTCB

Preliminary experiments involving NTCB cleavage of analytes prior to analysis by MALDI-MS were conducted with sulfhydryl-containing peptides having different amino acid sequences and cysteine locations. Table 2.1 lists the peptides and the expected masses of their expected products after cyanylation and cleavage.

Typical MALDI mass spectra of the cyanylated peptide (MW 2427.9) and its cleavage products, obtained by reacting 100 pmol of peptide with 500 pmol of NTCB under the described experimental conditions, are shown in Figure 2.5. After the cyanylation, a cyano group replaces hydrogen, resulting in a mass increase of 25 Da from the original peptide (Figure 2.5A). The expected cyanylated peptide has a m/z of 2453.9; the observed value exactly matches the expected data. After incubation in pH 9.0 buffer at 37°C for 16 hours, the cyanylated peptide is replaced by three new components (Figure 2.5B), two of them are cleavage products with the expected m/z of 1099.5 and 1373.5, respectively. The peak at m/z 2394.3 corresponds to the β-elimination product competing with the cleavage reaction. The β-elimination product can be easily identified because the mass shift from the cyanylated peptide is -59 Da.

Likewise, Figure 2.6 shows the MALDI spectra of the cleavage products from the remaining three peptides. These spectra reveal general features of MALDI spectra of peptide mixtures. Most of the major peaks can be assigned to expected specific species, such as the cyanylated/uncleaved peptide, the β-elimination products, and the specific

... sulfhydryl-containing peptides after the NTCB reaction

Table 2.1. Calculated m/z values of fragments for sulfhydryl-containing peptides after the NTCB reaction

amino acid sequence	$[M+H]^+$				
	intact peptide	cyanylated peptide	β-elimination product	fragment I	fragment II
TCVEWLRRYLKN	1581.9	1606.9	1547.9	120.1	1506.8
DRVYIHPCHLLYYS	1780.4	1805.4	1746.4	900.1	924.3
RYVVLPRPVCFEKGMNYTVR	2428.9	2453.9	2394.9	1099.5	1373.5
MHRQEAVDCLKKFNARRKLKGA	2601.4	2626.4	2567.4	986.1	1659.3

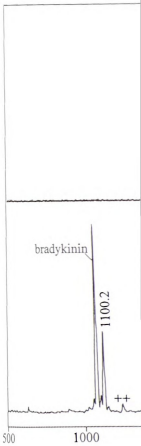


Figure 2.5. MALDI mass
reaction with a 5-fold mola
products. ++ indicates pea
for the correlation of calcul

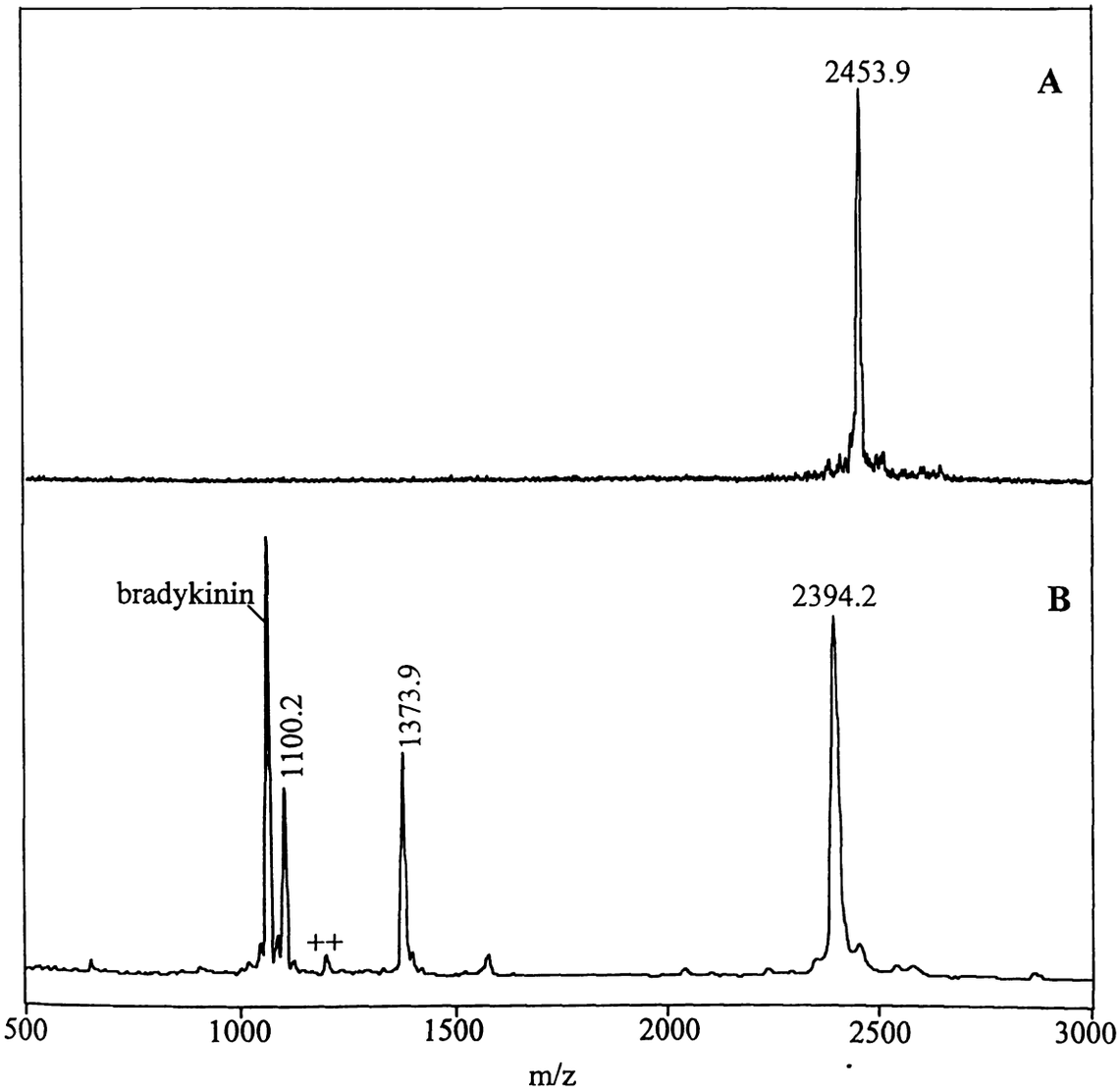


Figure 2.5. MALDI mass spectra of peptide RYVVLPRPVCFEKGMNYTVR after reaction with a 5-fold molar excess of NTCB, (A) cyanylated peptide and (B) cleavage products. ++ indicates peaks corresponding to doubly charged species. See Table 2.1 for the correlation of calculated and observed m/z values.

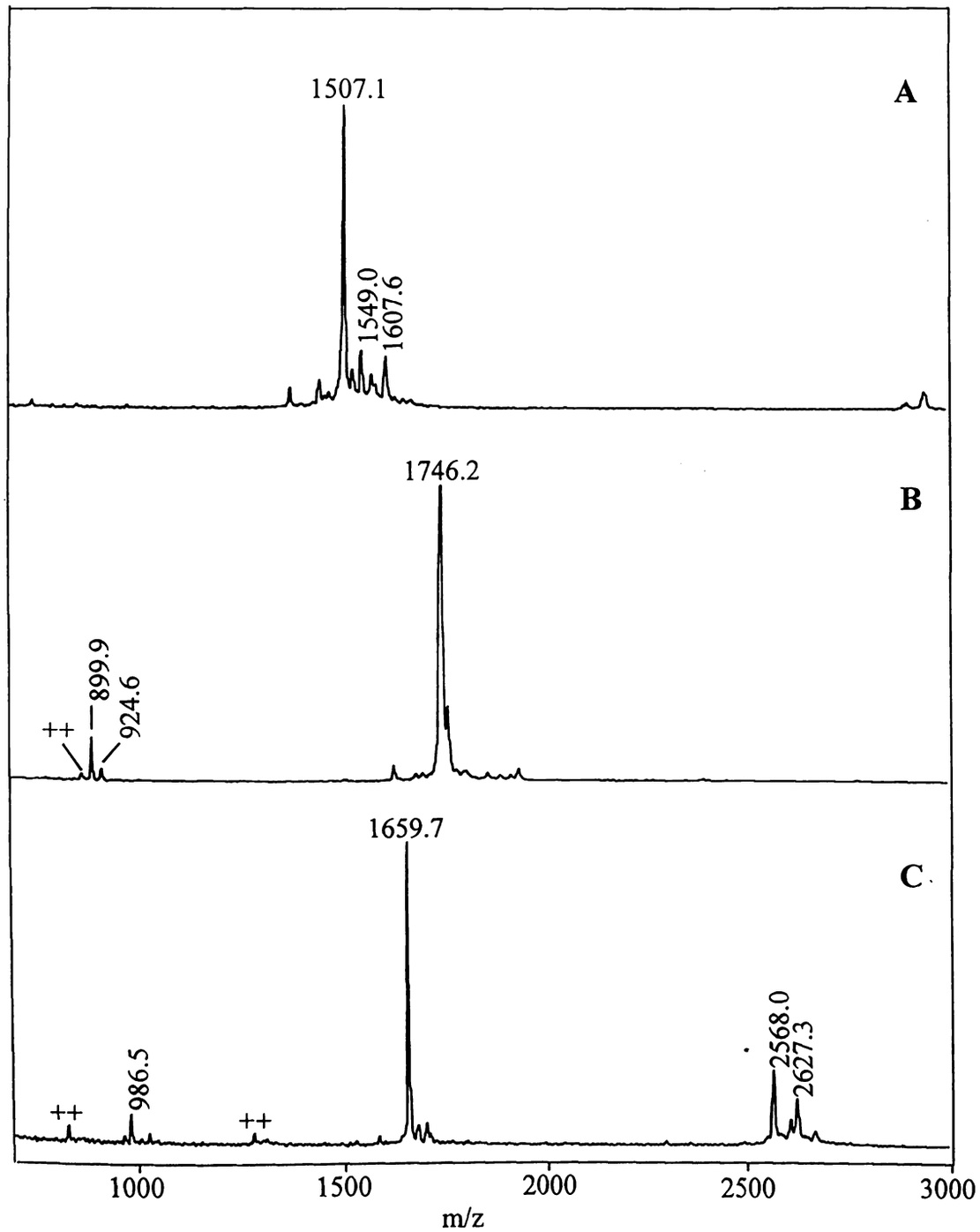


Figure 2.6. MALDI spectra of cleavage products of peptides (A) TCVEWLRRYLKN, (B) DRVYIHPCHLLYYS, and (C) MHRQEAVDCLKKFNARRKLKGA, after reaction with 5-fold molar excess of NTCB. See Table 2-1 for the identities of the peaks.

cleavage products, with a mas

used. Mass mapping by MAL

the peaks. However, the respo

greatly, depending on the natu

cleavage products, m/z 986.5

amount, however, the respon

different. Another observatio

elimination of cyanylated pep

of β-elimination products de

cyanylated cysteine residue

amino acid residues, such as

products, perhaps though s

underwent little β-eliminatio

DRVYIHPCHLLYYS. Exp

changing reaction conditions

NTCB/analyte were not succ

have the same pH dependenc

elimination products can als

reactions affect the yield of

information on the sulfhydry

structures on the cleavage rea

cleavage products, with a mass accuracy of better than 0.05% if internal standards are used. Mass mapping by MALDI, therefore, unambiguously recognizes the identities of the peaks. However, the responses of peptides on MALDI, i. e., the peak intensities, vary greatly, depending on the nature of the peptides. For example, in Figure 2.6C, the two cleavage products, m/z 986.5 and m/z 1659.7 respectively, should be of equal molar amount, however, the responses of the two fragments on MALDI are considerably different. Another observation is that, in addition to the expected cleavage products, β-elimination of cyanylated peptides always competes as a main side reaction. The yields of β-elimination products depend on the structures of amino acids adjacent to the cyanylated cysteine residue (19). Our experiments indicate that certain neighboring amino acid residues, such as Pro, His, or Phe, may increase the yield of β-elimination products, perhaps though steric hindrance. Therefore, while TCVEWLRRYLKN underwent little β-elimination, the β-elimination product predominated for the peptide DRVYIHPCHLLYYS. Experiments carried out to minimize this side reaction by changing reaction conditions such as pH, temperature, reaction time, and molar ratio of NTCB/analyte were not successful, because both cleavage and β-elimination reactions have the same pH dependency. Fortunately, both cyanylated/uncleaved peptides and β-elimination products can also be detected by MALDI-MS. Therefore, although these reactions affect the yield of the NTCB cleavage reaction, they provide complementary information on the sulfhydryl location and cleavage pattern. The effects of amino acid structures on the cleavage reaction will further be discussed in detail in Chapter 3.

Price (14) reported tha
during the cyanylation of sulfhy
between protein sulfhydryl grou
also observed this side reaction
experimental conditions. No
mixed disulfides has been obse

B. The Effects of pH and Te

Although the cleavage
systematic study and optimiza
reaction. The experimental co
protocol proposed by Jacobson
9.5 buffer for 16-80 hours alt
to promote the cleavage (38).
by conducting the cleavage re
and 9.5 buffer solutions,
cyanylation and cleavage can
using tris-HCl or sodium bor
be completed in a 5-fold m
minutes at 37°C, the complet
16 hours even at pH 9.0.
promote both cleavage and β-

Price (14) reported that another concurrent reaction may occur less frequently during the cyanylation of sulfhydryl groups and lead to the formation of a mixed disulfide between protein sulfhydryl group and NTCB reagent. Denslow and Nguyen (38) recently also observed this side reaction. But such a side reaction seems to be minor under our experimental conditions. No trace of the products corresponding to the formation of mixed disulfides has been observed on the analysis by MALDI-MS.

B. The Effects of pH and Temperature on the Kinetics of the Cleavage Reaction

Although the cleavage reaction by NTCB was proposed two decades ago (7), little systematic study and optimization have been carried out regarding to the kinetics of the reaction. The experimental conditions used today are exactly the same as that used in the protocol proposed by Jacobson *et al* (7). Typically, the cleavage is performed in pH 8.0 ~ 9.5 buffer for 16-80 hours although higher temperature was also occasionally employed to promote the cleavage (38). We have studied the effects of pH on the cleavage reaction by conducting the cleavage reaction on the above model peptides in pH 7.5, 8.0, 8.5, 9.0, and 9.5 buffer solutions, respectively. Time-course studies indicated that both cyanylation and cleavage can be carried out under mildly alkaline conditions (pH 8-9) using tris-HCl or sodium borate as buffers. While cyanylation of sulfhydryls can easily be completed in a 5-fold molar excess of NTCB solution even at pH 7.5 within 30 minutes at 37°C, the complete cleavage of the cyanylated peptide chains takes as long as 16 hours even at pH 9.0. Experimental results showed that higher pH can greatly promote both cleavage and β-elimination, while the relative yields of the two competitive

reactions do not vary significa

resulted in incomplete cleavag

resulting from long hours of in

It should be pointed out that th

from peptide to peptide, depen

to cysteines. Efficient cleavag

The cleavage reactio

temperature (e. g., 50°C) has

were usually conducted in ma

reaction system by dialysis o

analysis on microscale (pico

although higher temperature

after long hours of incubatio

of solvents. In addition, o

remaining reagents (e. g.,

temperatures have to be caref

C. **Identification of Free S**

Proteins may contain

forms. The NTCB/TCEP an

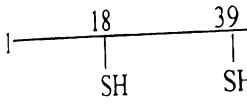
to determine the free sulfhy

reactions do not vary significantly over a wide pH range. The lower pH (< 7.5) often resulted in incomplete cleavage (cyanylated/uncleaved product) and other side reactions resulting from long hours of incubation, as will be discussed in greater detail in chapter 3. It should be pointed out that the kinetics of both cleavage and β-elimination vary greatly from peptide to peptide, depending considerably on the structure of amino acids adjacent to cysteines. Efficient cleavage requires careful monitoring of the reaction.

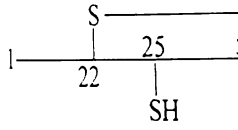
The cleavage reaction was previously performed at 37°C, although higher temperature (e. g., 50°C) has also been used. However, previous cleavage experiments were usually conducted in macroscale (mg) after removing the excess reagents from the reaction system by dialysis or other purification procedures (7, 13-20). Our results from analysis on microscale (picomole of samples in microliters of solution) indicated that although higher temperature can accelerate the cleavage, it is not recommended because after long hours of incubation, the reaction vial can easily get dry due to the evaporation of solvents. In addition, our cleavage reaction is performed without removal of the remaining reagents (e. g., NTCB), the side reactions caused by NTCB at higher temperatures have to be carefully considered.

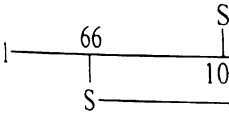
C. Identification of Free Sulfhydryls and Total Cysteines in Proteins

Proteins may contain free sulfhydryls, disulfide bonds, or a combination of both forms. The NTCB/TCEP and TCEP/NTCB procedures described in Figure 2.3 were used to determine the free sulfhydryls and total cysteines, respectively. Figure 2.7 lists the

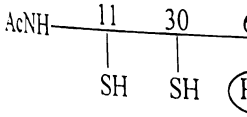


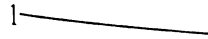
Sp



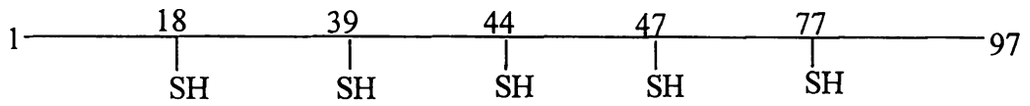


β

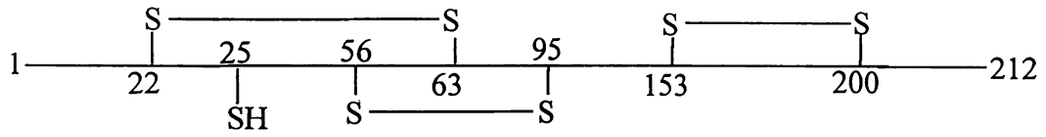




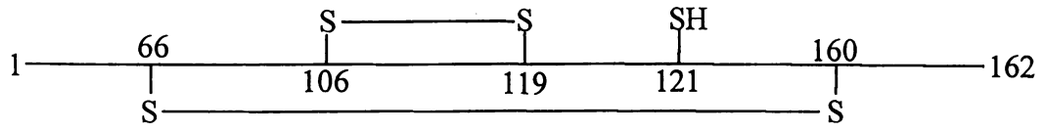
Figure



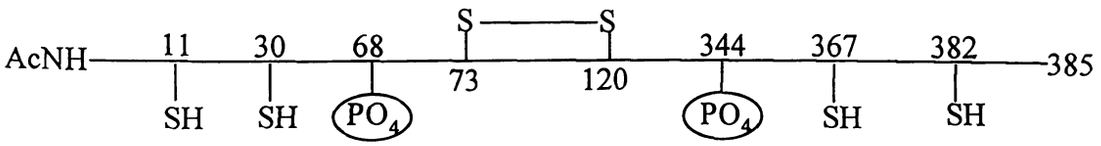
Spinach Ferredoxin (MW. 10483)



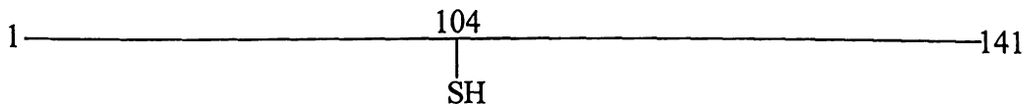
Papain (MW. 23426)



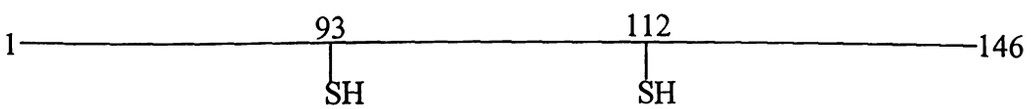
β-Lactoglobulin A (MW. 18368)



Ovalbumin (MW. 42699)



α-Hemoglobin A (MW. 15053)



β-Hemoglobin A (MW. 15954)

Figure 2.7. Structural features of model proteins.

Table 2.2. Mass assignm

Protein	NTCI Fragment
Ferredoxin[a]	1-17
	18-38
	39-43
	44-46
	47-76
	77-97
	18-43
	44-76
Papain	1-24
	25-212
Ovalbumin	1-10
	11-29
	30-366
	367-381
	382-385
	1-29
	1-29
	367-385
	367-385
β-Lactog-lobulin A	1-120
	121-162
α-Hemoglobin[a]	1-103
β-Hemoglobin	104-141
	1-92
	93-111
	112-146

[a] Only NTCB treatment was a
only contain free sulfhydryl g
[b] Not detected by MALDI-MS
[c] β-elimination products.
[d] Cyanylated/uncleaved peptic

Table 2.2. Mass assignment of peptide fragments after different treatments

Protein	NTCB/TCEP Treatment			TCEP/NTCB Treatment		
	Fragment	$[M+H]^+_{calc}$	$[M+H]^+_{obs}$	Fragment	$[M+H]^+_{calc}$	$[M+H]^+_{obs}$
Ferredoxin[a]	1-17	1839.1	1838.8			
	18-38	2352.2	2347.7			
	39-43	517.5	nd[b]			
	44-46	320.3	320.0			
	47-76	3337.6	3337.6			
	77-97	2332.5	2332.0			
	18-43	2792.0[c]	2792.8			
	44-76	3580.8[c]	3581.5			
Papain	1-24	2621.5	2622.1	1-21	2372.7	2371.9
	25-212	20848	nd	22-24	290.3	288.4
				25-55	3547.0	3546.7
				56-62	880.9	880.9
				63-94	3794.2	3792.4
				95-152	6309.2	6308.9
				153-199	5039.7	5038.9
				200-212	1503.7	1503.5
Ovalbumin	1-10	1011.1	nd	1-10	1011.1	nd
	11-29	2378.7	2378.5	11-29	2378.7	2377.4
	30-366	37684.7	nd	30-72	4762.5	4761.4
	367-381	1715.1	1715.2	73-119	5427.1	5426.1
	382-385	429.5	429.4	120-366	27581.1	nd
	1-29	3312.8[c]	3312.8	367-381	1715.1	1714.8
	1-29	3371.8[d]	3371.5	382-385	429.5	nd
	367-385	2067.6[c]	2067.7	1-29	3312.8[c]	3311.2
	367-385	2126.6[d]	2126.6	1-29	3371.8[d]	3370.3
				367-385	2067.6[c]	2068.1
β-Lactog-lobulin A	1-120	4905.7	4901.1	1-65	7247.4	7238.2
	121-162	13506.7	13510	66-105	4651.6	4653.5
				106-118	1461.6	1462.1
				119-120	274.3	nd
				121-159	4551.3	4548.1
				160-162	396.5	nd
α-Hemoglobin[a]	1-103	11073.6	11065			
β-Hemoglobin	104-141	4097.8	4098.0			
	1-92	9919.0	9915.0			
	93-111	2207.6	2207.0			
	112-146	3828.4	3831.6			

[a] Only NTCB treatment was applied to spinach ferredoxin and hemoglobin samples since the proteins only contain free sulfhydryl groups.

[b] Not detected by MALDI-MS.

[c] β-elimination products.

[d] Cyanylated/uncleaved peptides.

structural features of the prote

observed masses of the fragmen

1. Spinach Ferredoxin

Spinach ferredoxin cont

biological systems, some of th

with Fe and/or Mo (41). The

applied to the protein solution

complexes. The MALDI mass

of ferredoxin reacted with 2.5

the MALDI probe tip), sh

corresponding to 39-43. The

good agreement, with a relativ

the TOF instrument, the peak

with the peak at m/z 2332.5

m/z 2347.7 deviated from th

The peak at m/z 1820.1 is at

1839.1) during the cleavage

m/z 2792.8 and 3581.5 corr

44-76 in which β-elimination

structural features of the proteins studied by the methodologies. The calculated and observed masses of the fragments from these proteins are listed in Table 2.2.

1. Spinach Ferredoxin

Spinach ferredoxin contains 5 free sulfhydryls and no disulfide bonds (39, 40). In biological systems, some of the free sulfhydryl groups in ferredoxin can form complexes with Fe and/or Mo (41). Therefore, prior to treatment with NTCB, 1 mM EDTA was applied to the protein solution to liberate the free sulfhydryls from possible S-Fe or S-Mo complexes. The MALDI mass spectrum in Figure 2.8, recorded from a 100-pmol sample of ferredoxin reacted with 2.5 nmol of NTCB (~1.5 pmol of the analyte was applied to the MALDI probe tip), shows all the protonated peptide fragments except that corresponding to 39-43. The calculated and observed $[M + H]^+$ values (Table 2.2) are in good agreement, with a relative mass deviation of < 0.1%. But, due to poor resolution of the TOF instrument, the peak at m/z 2352.5 (corresponding to fragment 18-38) overlaps with the peak at m/z 2332.5 (corresponding to fragment 77-97). The observed peak at m/z 2347.7 deviated from the calculated value by -4.8 Da, indicating much higher error. The peak at m/z 1820.1 is attributable to the dehydration product of fragment 1-17 (m/z 1839.1) during the cleavage reaction, as observed by Papov and Biemann (42). Peaks at m/z 2792.8 and 3581.5 correspond to the incomplete cleavage of fragments 18-43 and 44-76 in which β-elimination occurs at Cys39 and Cys47, respectively.

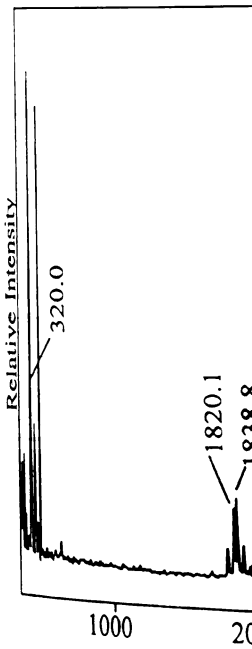


Figure 2.8. MALDI mass s
500 pmol of NTCB. Appro
Asterisks (*) indicates peak
the identities of marked pea

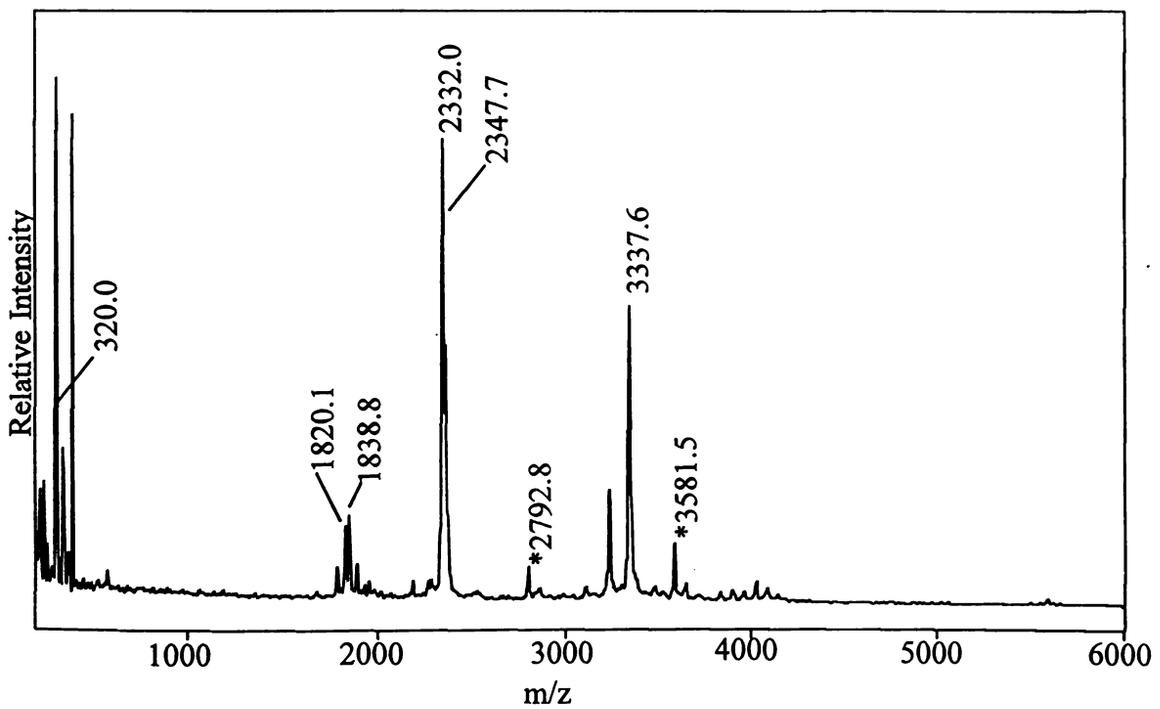


Figure 2.8. MALDI mass spectrum of 100 pmol of spinach ferredoxin after reaction with 500 pmol of NTCB. Approximately 1.5 pmol of analyte was applied to the probe tip. Asterisks (*) indicates peaks corresponding to β-elimination products. See Table 2.2 for the identities of marked peaks and the correlation of calculated and observed m/z values.

2. Papain

Papain, a member of a

study because it has one cystei

45). The NTCB/TCEP proc

sulfhydryl (Figure 2.9A). Bec

Cys25 should give two peptic

bond (Cys22-Cys66). The rec

and produces peptide fragmen

m/z 2621.5 and 20848.0, res

was observed corresponding

corresponding to peptide ch

probably due to suppression b

Another interesting ob

reaction mixture from papai

fragment 1-24 at m/z 2622

iminothiazolidinyl residue at

are still linked by an *inter*-ch

at Cys25. Thus, observatio

disulfide bond might have be

on β-lactoglobulin A (MW

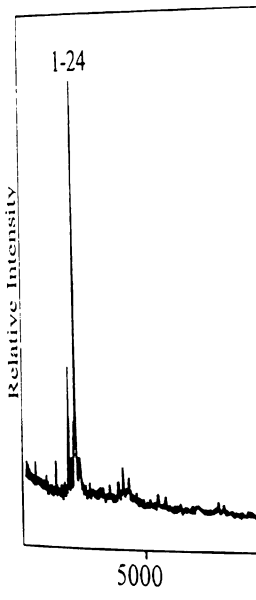
at Cys121, β-lactoglobulin

inter-chain disulfide bond (

2. Papain

Papain, a member of a family of over 40 thiol proteases (43), was chosen for this study because it has one cysteine (Cys25) at the active site and three disulfide bonds (44, 45). The NTCB/TCEP procedure was utilized to recognize the locus of the free sulfhydryl (Figure 2.9A). Because papain is a single chain polypeptide (44), cleavage at Cys25 should give two peptide chains which remain linked by an *inter*-chain disulfide bond (Cys22-Cys66). The reduction of disulfide bonds by TCEP cleaves the two chains and produces peptide fragments 1-24 and 25-212 having calculated values for $[M + H]^+$ at m/z 2621.5 and 20848.0, respectively. Experimentally, an intense peak at m/z 2622.1 was observed corresponding to the peptide chain 1-24. The expected high-mass peak corresponding to peptide chain 25-212 was not seen in the MALDI mass spectrum, probably due to suppression by the low-mass peptide fragment present in the sample.

Another interesting observation is that the MALDI analysis of the NTCB cleavage reaction mixture from papain prior to disulfide bond reduction also showed the peptide fragment 1-24 at m/z 2622.1. In principle, only an intact protein derivatized by an iminothiazolidinyl residue at Cys25 should be detected because the two peptide chains are still linked by an *inter*-chain disulfide bond (Cys22-Cys66) after the NTCB cleavage at Cys25. Thus, observation of the peak at m/z 2622.1 indicates that the *inter*-chain disulfide bond might have been cleaved during the MALDI process. Similar experiments on β-lactoglobulin A (MW 18368) (46) gave a similar result. After cleavage with NTCB at Cys121, β-lactoglobulin A should produce two peptide chains that are linked by an *inter*-chain disulfide bond (Cys66-Cys160). However, the MALDI mass spectrum of β-



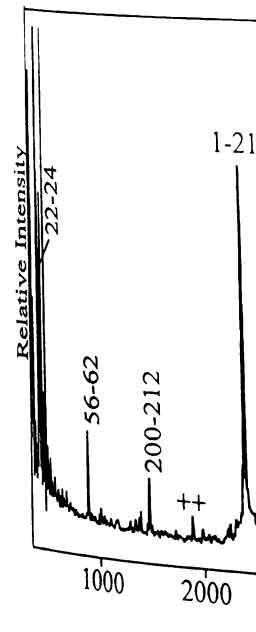


Figure 2.9. MALDI mass
TCEP/NTCB treatment.
correlation of calculated a

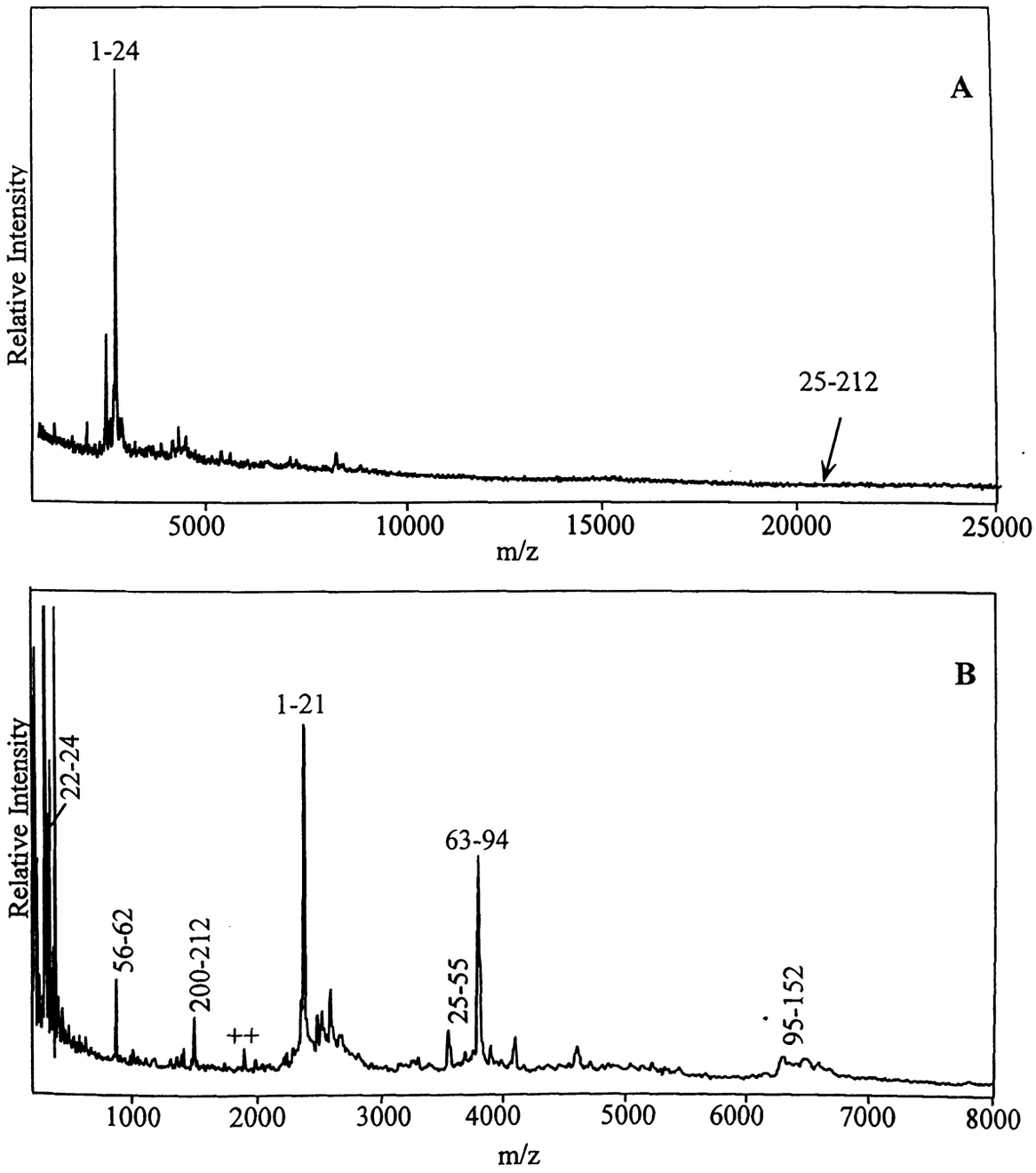


Figure 2.9. MALDI mass spectra of papain after (A) NTCB/TCEP treatment and (B) TCEP/NTCB treatment. See Table 2.2 for the identities of other marked peaks and the correlation of calculated and observed m/z values.

lactoglobulin A after reaction
resulting from the cleavage of
the recent report by Patterson
undergo "prompt fragmentatio
(47).

Total cysteine assessm
procedure. Most of the expect
seven cysteine residues were
peaks unable to be assigned to
in the commercial sample. Ta
their observed masses. In th
elimination were found, whic
polypeptide chain at all the c

3. Ovalbumin

Ovalbumin (48-50),
studied by the proposed pro
NTCB/TCEP procedure (Fi
gave two more peaks (at m/
and cleavage of the disulfid
agreement with the previou
Furthermore, two clusters

lactoglobulin A after reaction with NTCB showed peaks corresponding to fragments resulting from the cleavage of the *inter*-chain disulfide bond. Our observations support the recent report by Patterson *et al* that the *inter*-chain disulfide bond in peptides can undergo "prompt fragmentation" or "in-source fragmentation" in the MALDI ion source (47).

Total cysteine assessment of papain was performed according to the TCEP/NTCB procedure. Most of the expected fragments corresponding to the cleavage at each of the seven cysteine residues were found in the MALDI mass spectrum (Figure 2.9B). The peaks unable to be assigned to any of the fragments were most likely due to an impurity in the commercial sample. Table 2.2 also lists the calculated masses of the fragments and their observed masses. In this case, only a few fragments of incomplete cleavage or β-elimination were found, which indicates that the NTCB reagent can effectively cleave the polypeptide chain at all the cysteine residues in papain.

3. Ovalbumin

Ovalbumin (48-50), containing four free cysteines and one disulfide bond, was studied by the proposed procedures. Compared to the mass spectrum obtained from the NTCB/TCEP procedure (Figure 2.10A), that obtained with the TCEP/NTCB treatment gave two more peaks (at m/z 4761.4 and 5426.1) that were due to reduction, cyanylation, and cleavage of the disulfide bond Cys73-Cys120 (Figure 2.10B). This conclusion is in agreement with the previous assignment of the disulfide bond linkage in ovalbumin (50). Furthermore, two clusters of peaks were observed in the MALDI mass spectra which

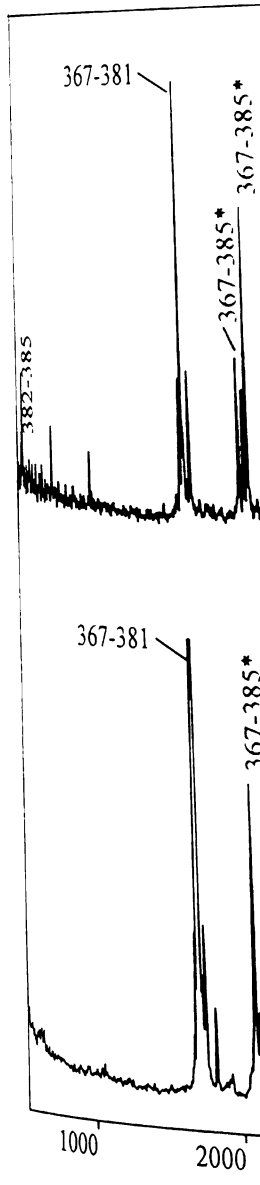


Figure 2.10. MALDI mas
TCEP/NTCB treatment.
peptides and/or their β-eli
peaks and the correlation

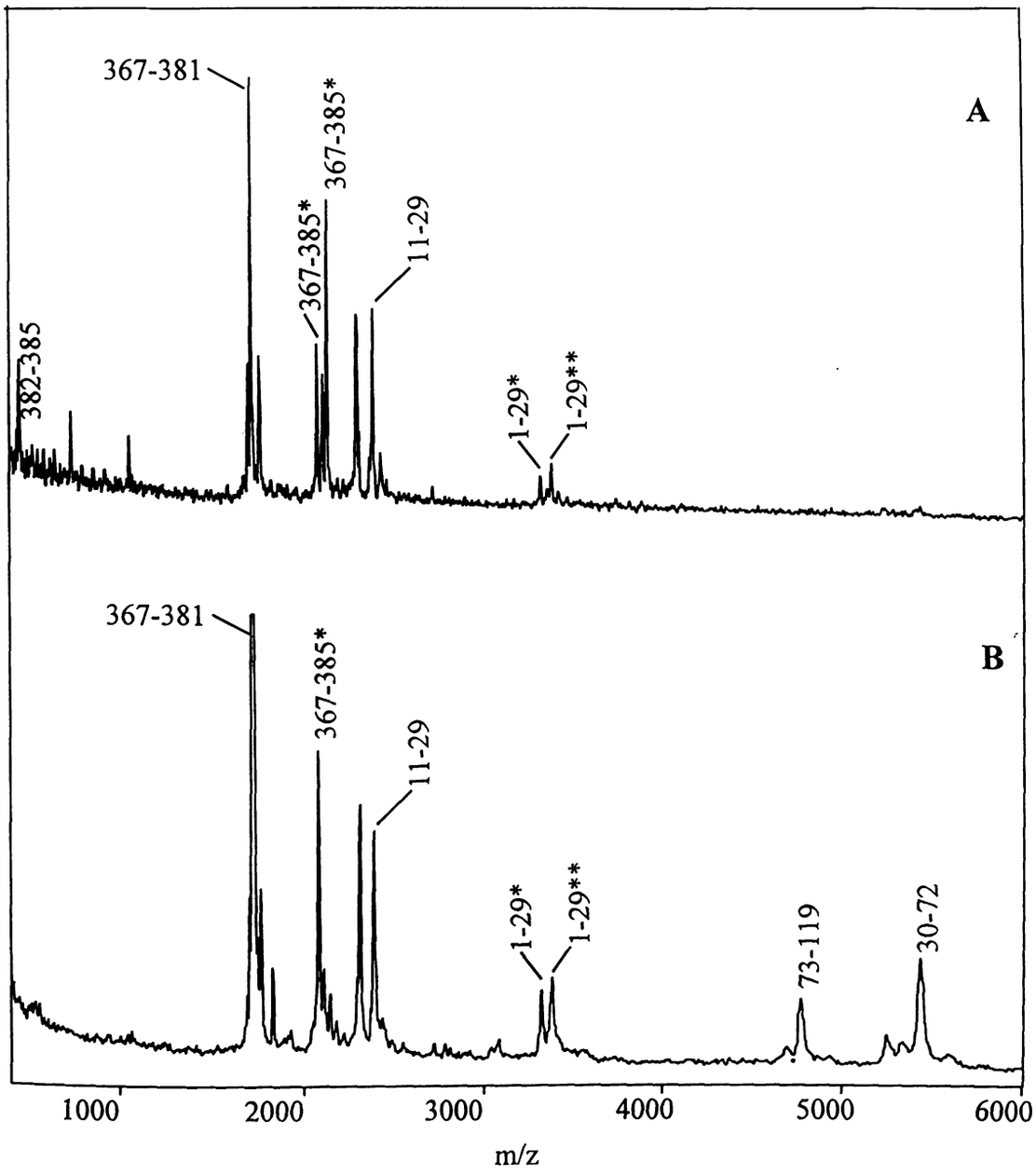


Figure 2.10. MALDI mass spectra of ovalbumin after (A) NTCB/TCEP and (B) TCEP/NTCB treatment. * indicates peaks corresponding to cyanylated/uncleaved peptides and/or their β-elimination products. See Table 2.2 for the identities of marked peaks and the correlation of calculated and observed m/z values.

correspond to the incomplete

peak at m/z 3371.5 is due to

Cys11 (expected m/z 3371.8).

at Cys11 (expected m/z 3312.

peptide and its β-elimination

and 2067.7 are likely due

uncleaved Cys382 and its

ovalbumin contains two ph

observed $[M+H]^-$ ion of the

phosphate group is still att

indicating that phosphoryla

experimental/ reaction condi

can be used to detect the p

segment, because it is princ

shift. This example demons

be used to confirm a protein

4. β-Lactoglobulin A

β-Lactoglobulin A

(46). After NTCB/TCEP t

121 are expected, with a

Experimentally, the fragme

correspond to the incomplete cleavage and/or β-elimination of sulfhydryl groups. The peak at m/z 3371.5 is due to the peptide chain 1-29 containing a cyanylated/uncleaved Cys11 (expected m/z 3371.8), while m/z 3311.1 corresponds to the β-elimination product at Cys11 (expected m/z 3312.8). The mass difference between the cyanylated/uncleaved peptide and its β-elimination product is 59 Da. For the same reason, peaks at m/z 2126.6 and 2067.7 are likely due to the peptide chain 367-385, containing a cyanylated/ uncleaved Cys382 and its β-elimination product, respectively. It is known that ovalbumin contains two phosphorylated residues at serines 68 and 344 (50). The observed $[M+H]^+$ ion of the fragment 30-72 (expected at m/z 4761.5 Da) shows that the phosphate group is still attached to Ser68 after cyanylation and cleavage reactions, indicating that phosphorylated proteins should be amenable to analysis under these experimental/ reaction conditions. On the other hand, the mass mapping of the fragments can be used to detect the possible posttranslational modification in a specific protein segment, because it is principally possible to recognize the modification from the mass shift. This example demonstrated that total cysteine determination by this approach can be used to confirm a protein's structure.

4. β-Lactoglobulin A

β-Lactoglobulin A contains five cysteines, of which Cys121 is a free cysteine (46). After NTCB/TCEP treatment, two fragments corresponding to the cleavage at Cys 121 are expected, with a m/z of 13506.7 (1-120) and 4905.7 (121-162), respectively. Experimentally, the fragment 1-120 presents as a small peak (Figure 2-11A), probably

due to the suppression from c

two expected fragments, 1-12

the other two fragments 121-1

cyanylation and cleavage at C

160 is insufficiently stable ir

thiolate ions that subsequen

Given the fact that β-lactogl

observed nonspecific cleav

exchange and subsequent c

mixture against 20% acetic

molecular weight salts and o

detected by MALDI (Fig. 2-

The peak at m/z 18390 is lik

containing fragment 1-120

doubly charged ion. The M

shows fragments 1-65, 66-

and 160-162, which fall i

fragments are essentially

MALDI responses of indi

similar MALDI spectra of

the disulfide bonds and sub

due to the suppression from other components (peptides and salts). In additional to the two expected fragments, 1-120 and 121-162, MALDI can also identify trace amounts of the other two fragments 121-159 (m/z 4551.2) and 1-65 (m/z 7247.0), attributable to both cyanylation and cleavage at Cys66 and Cys160, implying that the disulfide bond pair 66-160 is insufficiently stable in the presence of NTCB reagent and is hydrolyzed to form thiolate ions that subsequently participate in the cyanylation and cleavage reactions. Given the fact that β-lactoglobulin A is particularly liable to SH/S-S exchange (51), the observed nonspecific cleavage at Cys66 and Cys160 could also result from such exchange and subsequent cyanylation and cleavage. After dialysis of the cleavage mixture against 20% acetic acid for 12 hours (MW cut-off: 3000 Da) to remove low molecular weight salts and other contaminants, the fragment 1-120 can be unambiguously detected by MALDI (Fig. 2-11B), confirming that Cys121 is a cyanylation/cleavage site. The peak at m/z 18390 is likely due to the reoxidization during the dialysis of sulfhydryl-containing fragment 1-120 and 121-162, while the peak at m/z 9191 is obviously its doubly charged ion. The MALDI spectrum after TCEP/NTCB treatment (Figure 2-11C) shows fragments 1-65, 66-105, 106-118, 121-159, and two smallest fragments, 119-120 and 160-162, which fall in the region reserved for matrix peaks. Therefore, all the fragments are essentially detected, despite of the greatly variable differences in the MALDI responses of individual fragments. Denslow and Nguyen (38) have reported similar MALDI spectra of cleavage products of β-lactoglobulin obtained by reduction of the disulfide bonds and subsequent cyanylation and cleavage of the sulfhydryl groups.

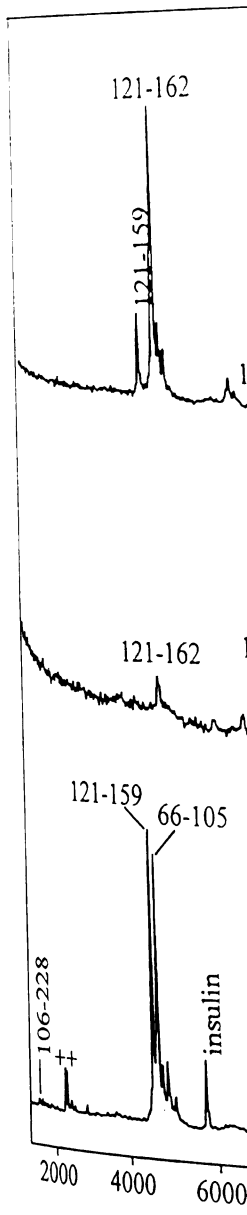


Figure 2.11. MALDI m
NTCB/Dialysis/TCEP,
correlation of calculate

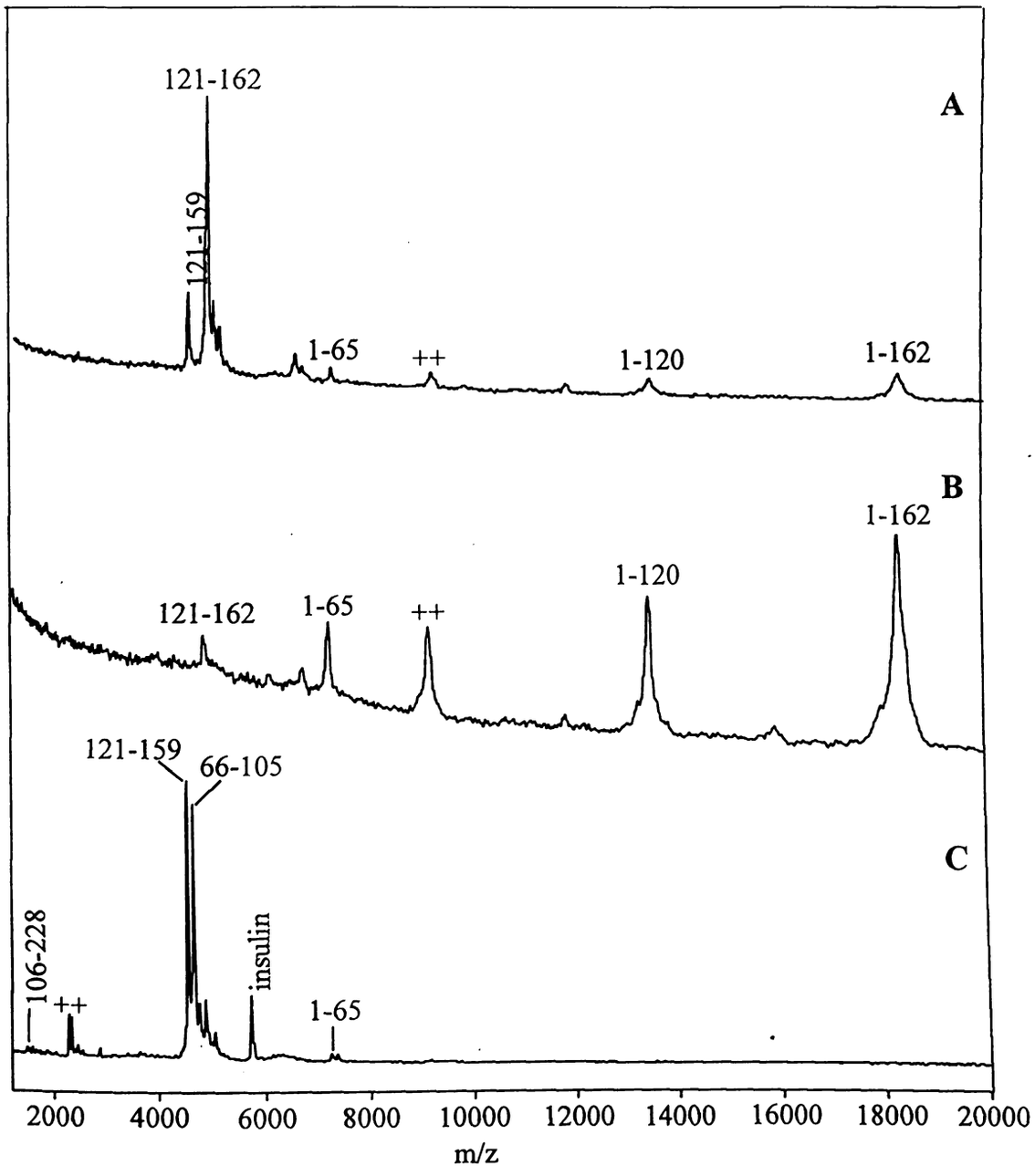


Figure 2.11. MALDI mass spectra of b-lactoglobulin after (A) NTCB/TCEP, (B) NTCB/Dialysis/TCEP, and (C) TCEP/NTCB treatment. See Table 2.2 for the correlation of calculated and observed m/z values.

4. Hemoglobin

Human hemoglobin A

one free cysteine at 104, wh

112. Since hemoglobin A ca

commercially available. The

mixture of α- and β-hemoglo

the expected fragments can

observation is that each of t

chain, is accompanied by a p

The identity of these peaks

optimization of the determin

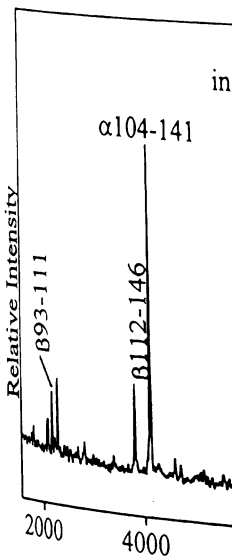


Figure 2.12. MALDI m
with NTCB. See table

4. Hemoglobin

Human hemoglobin A contains α- and β- chains (52). α-Hemoglobin A contains one free cysteine at 104, while β-hemoglobin A contains two free cysteines at 93 and 112. Since hemoglobin A can be easily oxidized in the air. The native hemoglobin is not commercially available. The hemoglobin we acquired was a service sample containing a mixture of α- and β-hemoglobin A variants. After NTCB cyanylation and cleavage, all the expected fragments can be unambiguously identified (Figure 2.12). An interesting observation is that each of the fragments, 1-92 and 93-111 of B chain, and 1-103 of A chain, is accompanied by a peak which shows a 102 Da mass increase from that fragment. The identity of these peaks are unclear. Because of the limited sample available, the optimization of the determination was not pursued further.

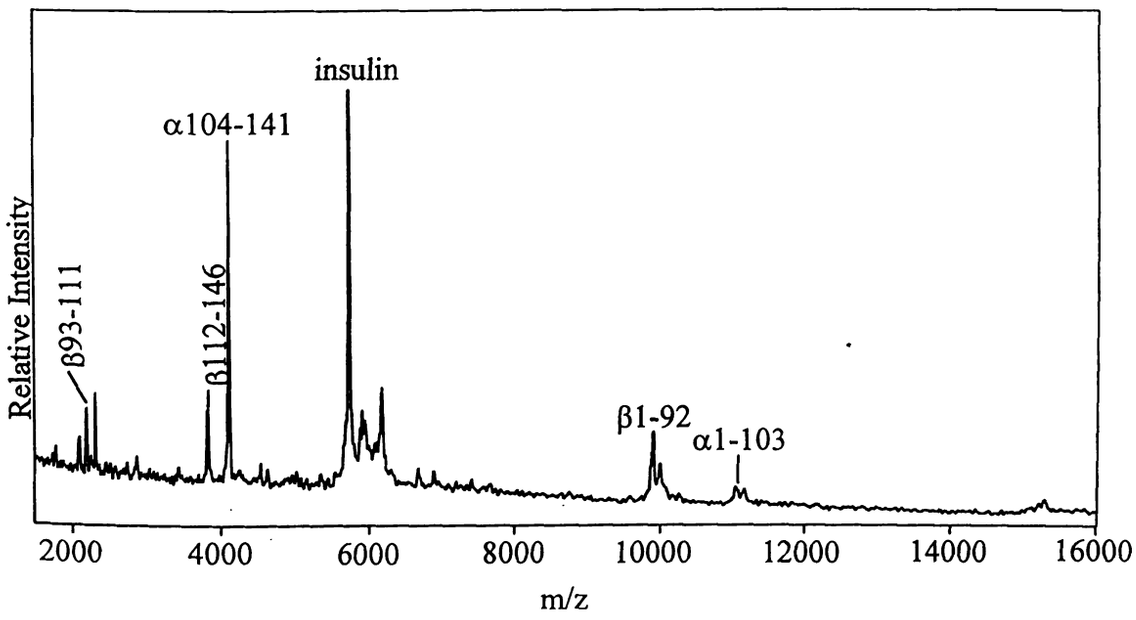


Figure 2.12. MALDI mass spectrum of α- and β-hemoglobin mixtures after reaction with NTCB. See table 2.2 for the correlation of calculated and observed m/z values.

E. Choice of Reducing Rea

Several reducing rea
reduction of protein disu
dithiothreitol (DTT) (54). ar
However, thiol-containing r
be avoided because excess
interferes with analyses
carboxyethyl)phosphine (T
disulfide bonds (36, 55, 56)
used in the presence of an o
is stable not only in acidic
reducing activity than DT
reaction solution is therefor
sufficient for complete red
larger amounts of TCEP d
MALDI-MS.

F. Stability of Disulfide B

Stoichiometric expe
for complete cyanylation
performed under the same

E. Choice of Reducing Reagent

Several reducing reagents have been developed for the efficient and selective reduction of protein disulfide bonds. Among these, β-mercaptoethanol (53), dithiothreitol (DTT) (54), and sodium borohydride ($NaBH_4$) (55) are traditionally used. However, thiol-containing reducing reagents such as β-mercaptoethanol and DTT must be avoided because excess reagents will react with NTCB. Sodium borohydride interferes with analyses by MALDI-MS (55). Tributylphosphine and tris(2-carboxyethyl)phosphine (TCEP) have proven to be selective and reactive towards disulfide bonds (36, 55, 56). However, tributylphosphine is water-insoluble and must be used in the presence of an organic solvent. TCEP is a water-soluble reducing reagent and is stable not only in acidic solution, but in basic solution as well. TCEP shows higher reducing activity than DTT over a wide pH range (57). The pH adjustment of the reaction solution is therefore unnecessary. Usually, a 5-fold molar excess of TCEP is sufficient for complete reduction of disulfide bonds in denaturing conditions although larger amounts of TCEP do not appear to have any adverse effect on the analysis by MALDI-MS.

F. Stability of Disulfide Bonds during the NTCB Reaction

Stoichiometric experiments indicated that the NTCB/thiol ratio of 5:1 is sufficient for complete cyanylation of thiol groups. Because both cyanylation and cleavage were performed under the same condition without removal of the NTCB reagent, the excess of

NTCB reagent can always ha

a long period of incubation,

The mechanism of th

acting as a nucleophile to at

cyanylated product, which c

amino peptide bond (7). In

groups. Selective cyanylati

presence of cystines becaus

specificity is not without lir

containing peptides indica

incubation time result in cle

the stability study of disul

excess of NTCB reagent u

disulfide bond. The same r

excess of NTCB reagent i

hours. MALDI spectra, as

generally stable in the pre

incubation at 37°C, cor

cyanylation/cleavage of the

NTCB exceeds 50-fold unc

correspond to the cleavage

disulfide bonds undergo

NTCB reagent can always have a chance to attack the unreacted sulfhydryl groups during a long period of incubation, which ensures the cyanylation goes completion.

The mechanism of the NTCB reaction involves the dissociated thiolate anion ($-S^-$) acting as a nucleophile to attack the cyanate moiety under alkaline conditions to form a cyanylated product, which can then undergo cleavage catalyzed by hydroxide ions at the amino peptide bond (7). In principle, the NTCB reagent is specific for free sulfhydryl groups. Selective cyanylation/cleavage at free sulfhydryl groups can be performed in the presence of cystines because the disulfide bonds do not react with NTCB. However, this specificity is not without limits under the reaction conditions. Experiments on disulfide-containing peptides indicate that a large excesses of NTCB reagent and excessive incubation time result in cleavage at disulfide bond residues. This has been confirmed by the stability study of disulfide-containing peptide(s) in the presence of different molar excess of NTCB reagent under the same buffer conditions. Somatostatin contains one disulfide bond. The same molar amount of somatostatin was mixed with 2, 5, 10, 50-fold excess of NTCB reagent in pH 9.0 buffer, respectively, and incubated at 37°C for 16 hours. MALDI spectra, as shown in Figure 2.13, show that while the disulfide bond is generally stable in the presence of a 10-fold molar excess of NTCB after 16 hours of incubation at 37°C, complete splitting of the disulfide bond and subsequent cyanylation/cleavage of the nascent cysteine residues take place when the molar excess of NTCB exceeds 50-fold under the same conditions. The masses of the resulting fragments correspond to the cleavage at one or both cysteine sites. One possible explanation is that disulfide bonds undergo hydrolysis under mildly alkaline conditions (58, 59). The

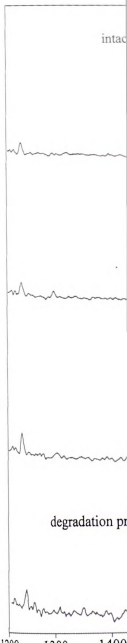


Figure 2.13. Stability stu
of NTCB reagent. After
at 37°C for 16 hours, sor

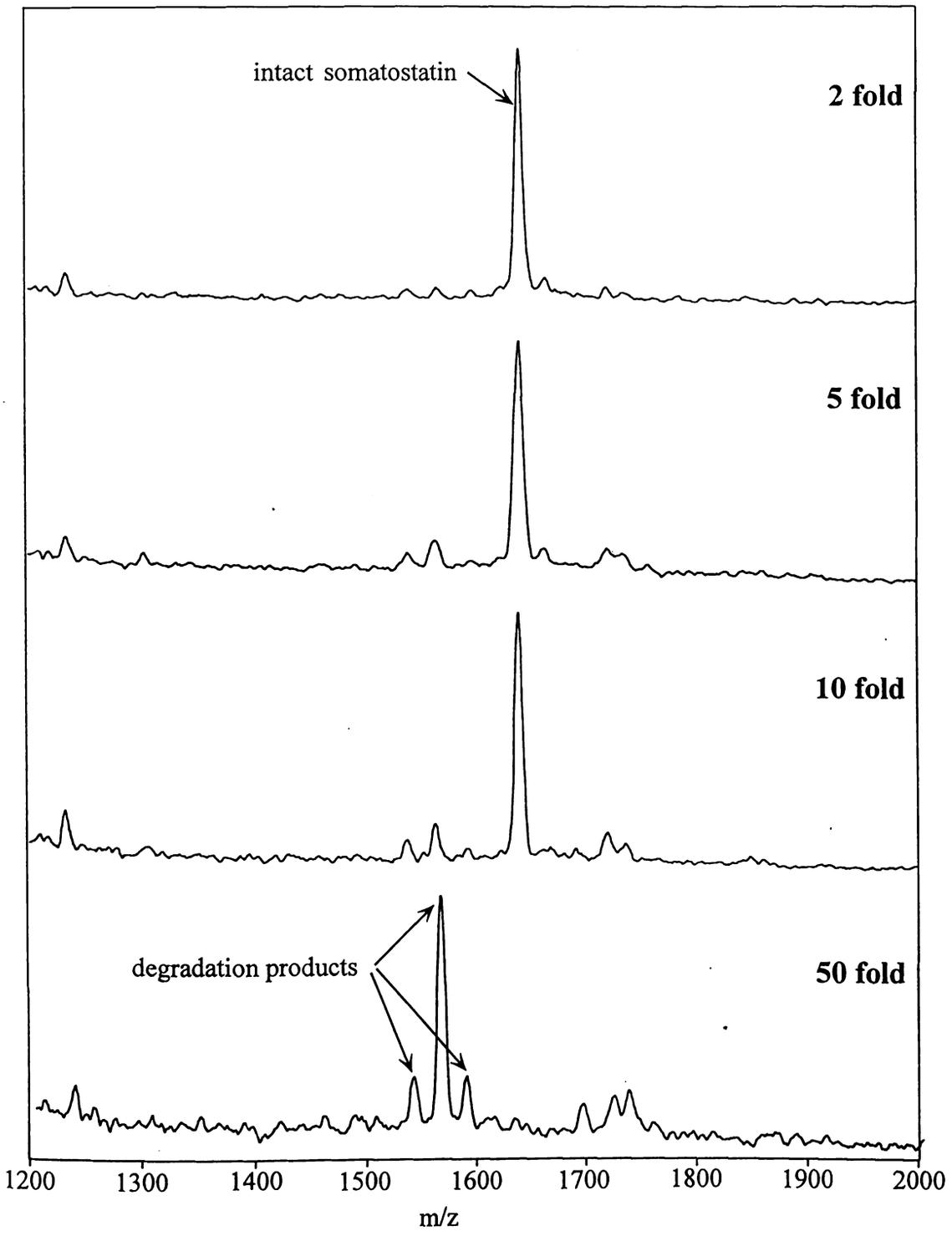


Figure 2.13. Stability study of somatostatin in the presence of different molar excesses of NTCB reagent. After exposed to 50-fold excess of NTCB reagent in pH 9.0 buffer at 37°C for 16 hours, somatostatin decomposes completely.

resulting thiolate anions attac

cleave the peptide chain

NTCB/TCEP procedure is u

NTCB must be applied to en

purified by chromatography

subsequent cleavage reactio

total cysteines (TCEP/NTC

prior to the NTCB reaction.

Another commonly

bond exchange that usually

9.0, 16 hours incubation) di

not affect the assignment

sulfhydryls become blocke

hand, the described metho

bond pairing patterns.

G. Analysis by MALDI-M

Because of the wid

obtained from NTCB cleav

are common. Previous

employed sodium dodecy

Unfortunately, its accurac

resulting thiolate anions attack the NTCB reagent to form cyanylated products that further cleave the peptide chain at the original disulfide bond sites. Therefore, if the NTCB/TCEP procedure is used for sulfhydryl-rich proteins where a high molar excess of NTCB must be applied to ensure complete cyanylation, the cyanylated proteins are better purified by chromatography or dialysis to remove the excess NTCB reagent prior to the subsequent cleavage reaction. This limitation is not a problem for the determination of total cysteines (TCEP/NTCB procedure) since all the disulfide bonds must be reduced prior to the NTCB reaction.

Another commonly encountered problem in cysteine characterization is disulfide bond exchange that usually occurs at pH > 8.0. Under our experimental conditions (pH 9.0, 16 hours incubation) disulfide bond exchange cannot be avoided. However, it should not affect the assignment of free sulfhydryl groups and/or total cysteines since free sulfhydryls become blocked immediately after cyanylation with NTCB. On the other hand, the described method cannot directly be employed for the assignment of disulfide bond pairing patterns.

G. Analysis by MALDI-MS

Because of the widely variable occurrence of cysteine in proteins, the fragments obtained from NTCB cleavage vary greatly in size. Fragments with a mass over 5000 Da are common. Previous reports on the mass mapping of NTCB cleavage products employed sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Unfortunately, its accuracy of molecular weight determination is usually limited to ± 5-

10%. In many cases, pept
masses of two or more pept
Incomplete cleavage of prot
and the SDS-PAGE technic
products corresponding t
cyanylated/uncleaved peptic

MALDI-TOF MS
determine molecular weigh
0.01% and thereby provide
comparison of calculated v
assignment is better than ±
for calibration. This accura

MALDI spectra ar
though peaks for a sod
Occasionally, matrix addu
that the MALDI signal int
because the reaction produ
peptides may not give
discrimination effects are
of salts, or due to differe
peptides are more suppres

10%. In many cases, peptide assignments are inconclusive or are ambiguous if the masses of two or more peptides obtained from the NTCB digestion are similar (19, 21). Incomplete cleavage of protein chains is often observed with the NTCB reaction (60-62), and the SDS-PAGE technique has difficulty distinguishing among incomplete cleavage products corresponding to the peptide with an unreacted sulfhydryl group, a cyanylated/uncleaved peptide, or a β-elimination product.

MALDI-TOF MS has been a valuable complement to SDS-PAGE. It can determine molecular weights as high as 300,000 Da with the mass accuracy of ±0.1-0.01% and thereby provide an unambiguous answer to questions mentioned above. A comparison of calculated values and our observed data indicates that the error of mass assignment is better than ± 0.05% for most of the peptides if internal standards are used for calibration. This accuracy is sufficient for the mass mapping of cleavage products.

MALDI spectra are dominated by peaks for singly protonated species $[M+H]^+$, though peaks for a sodium adduct $[M+Na]^+$ often appear at a lower intensity. Occasionally, matrix adduct peaks are observed. A limitation of our present method is that the MALDI signal intensities vary considerably from peptide to peptide in a mixture because the reaction products are directly analyzed without purification by HPLC. Some peptides may not give a detectable signal during analysis of a mixture. These discrimination effects are either due to MALDI contaminants such as high concentrations of salts, or due to differential responses of the peptides, or both. Usually high-mass peptides are more suppressed and yield weaker signals.

H. Reactions of Proteins I

One of the great techniques, such as FAB a However, the high conce interferes with a successful form a thick cake or a res Samples containing contam of the sample to minimiz analyte often can be obser where only trace quantities

Recently, various t supports for direct analysis and proteins on an inert t soluble MALDI contamina while the analyte is retaine matrix solution. Moreov proteins are possible on th

The *in situ* disulfi reaction for protein sampl Zetabind membrane and a of which have been emplo 64). Experiments with th

H. Reactions of Proteins Immobilized on A Zetabind Membrane

One of the great advantages of MALDI-MS over other mass spectrometric techniques, such as FAB and ESI mass spectrometry, is its tolerance of contaminants. However, the high concentration of denaturing agents used in our procedure still interferes with a successful analysis by MALDI-MS. The protein-matrix mixtures often form a thick cake or a residue that does not completely cocrystallize on the probe tip. Samples containing contaminants frequently give no or very weak signals. Upon dilution of the sample to minimize the effect of the contaminants, a MALDI signal from the analyte often can be observed. Diluting the sample is, however, self-defeating in cases where only trace quantities of the analyte are available.

Recently, various transfer membranes (37, 63, 64) have been studied as sample supports for direct analysis of proteins by MALDI-MS. The immobilization of peptides and proteins on an inert transfer membrane can be used to facilitate removal of water-soluble MALDI contaminants, such as high concentrations of denaturing agents and salts, while the analyte is retained when the membrane is immersed in water prior to adding the matrix solution. Moreover, *in situ* chemical and enzymatic modifications of adsorbed proteins are possible on the membrane.

The *in situ* disulfide reduction by TCEP, cyanylation by NTCB, and cleavage reaction for protein samples as small as 10 pmol have been performed on a nylon-based Zetabind membrane and a porous polyethylene membrane (purchased from Fisher), both of which have been employed for direct MALDI-MS analysis of high-mass proteins (37, 64). Experiments with the NTCB reaction on the two membranes showed comparable

results. However, the pol
operationally difficult to ha

The MALDI mass
reactions with 10 pmol of
membrane, are shown in F
NTCB reagents was app
cyanylation of sulfhydryl
membrane by washing, w
the high concentration of
subsequent analysis by
disulfide bond sites, in the
the excess NTCB reagent
performed in solution and
spectra for papain. The c
complete for ovalbumin
observed in the MALDI n
Cys11. However, the res
poorer for the analytes im
adequately controlled.

One problem with
maintaining an aqueous s
This problem can be mini
containing a few drops c

results. However, the polyethylene membrane is a porous thin film on which it was operationally difficult to handle the chemical reactions.

The MALDI mass spectra of the reaction products, obtained respectively from reactions with 10 pmol of papain and 15 pmol of ovalbumin immobilized on a Zetabind membrane, are shown in Figures 2.14 and 2.15. A 50-fold molar excess of TCEP and NTCB reagents was applied to ensure complete reduction of disulfide bonds and cyanylation of sulfhydryls. Because all reagents can readily be removed from the membrane by washing, while the protein and/or peptide reaction products are retained, the high concentration of TCEP and NTCB do not present any large problem during subsequent analysis by MALDI-MS. Moreover, the possible cleavage reaction at disulfide bond sites, in the presence of a high molar excess of NTCB, was minimized as the excess NTCB reagent was washed out after the cyanylation reaction. The reactions performed in solution and on the surface of the Zetabind membrane gave almost the same spectra for papain. The cleavage reaction performed on the membrane was even more complete for ovalbumin relative to that performed in the solution; no peaks were observed in the MALDI mass spectrum that corresponded to the incomplete cleavages at Cys11. However, the resolution and accuracy of mass assignment can be substantially poorer for the analytes immobilized on the membrane if the power level of the laser is not adequately controlled.

One problem with performing microscale chemical reactions on a membrane is in maintaining an aqueous solution of the reactants on the membrane for an extended period. This problem can be minimized by performing reactions in a closed vessel (see Fig. 2-3) containing a few drops of water to saturate the air space with water vapor so that the

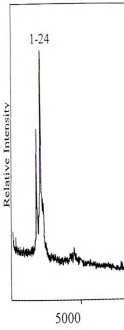


Figure 2.14. MALDI r
treatment on a Zetabinc

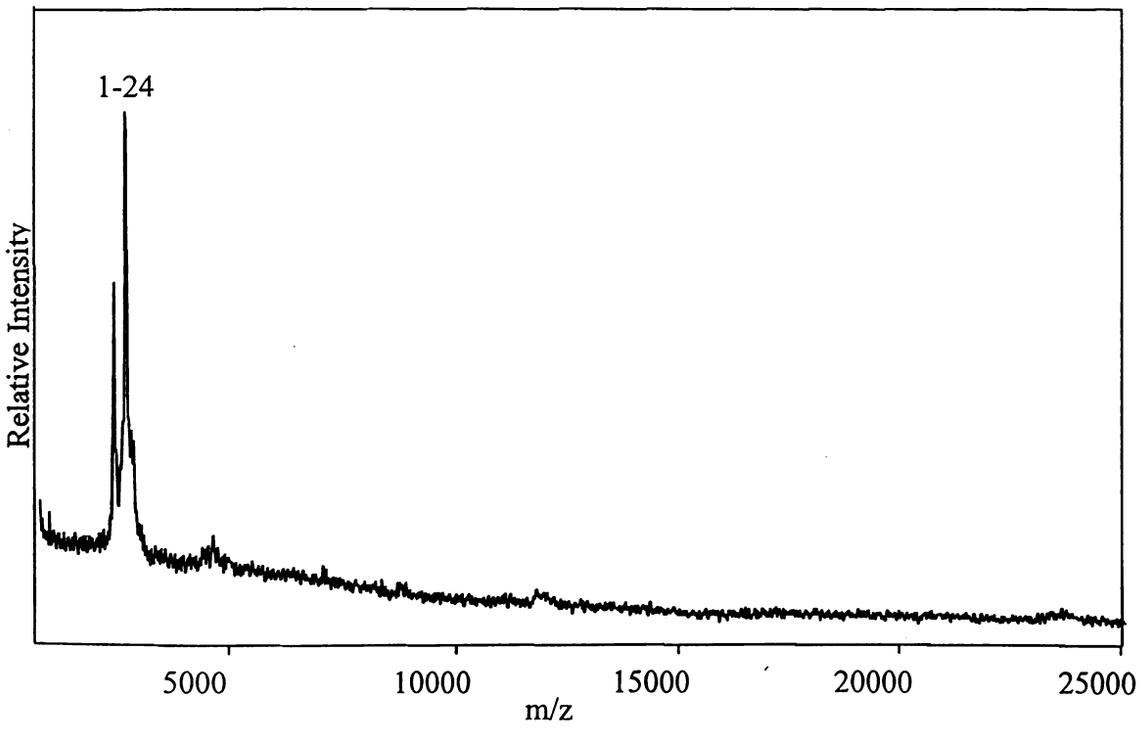


Figure 2.14. MALDI mass spectrum of 10 pmol of papain after *in situ* NTCB/TCEP treatment on a Zetabind membrane.

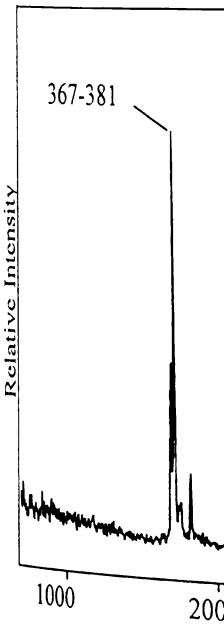


Figure 2.15. MALDI m
treatment on a Zetabind

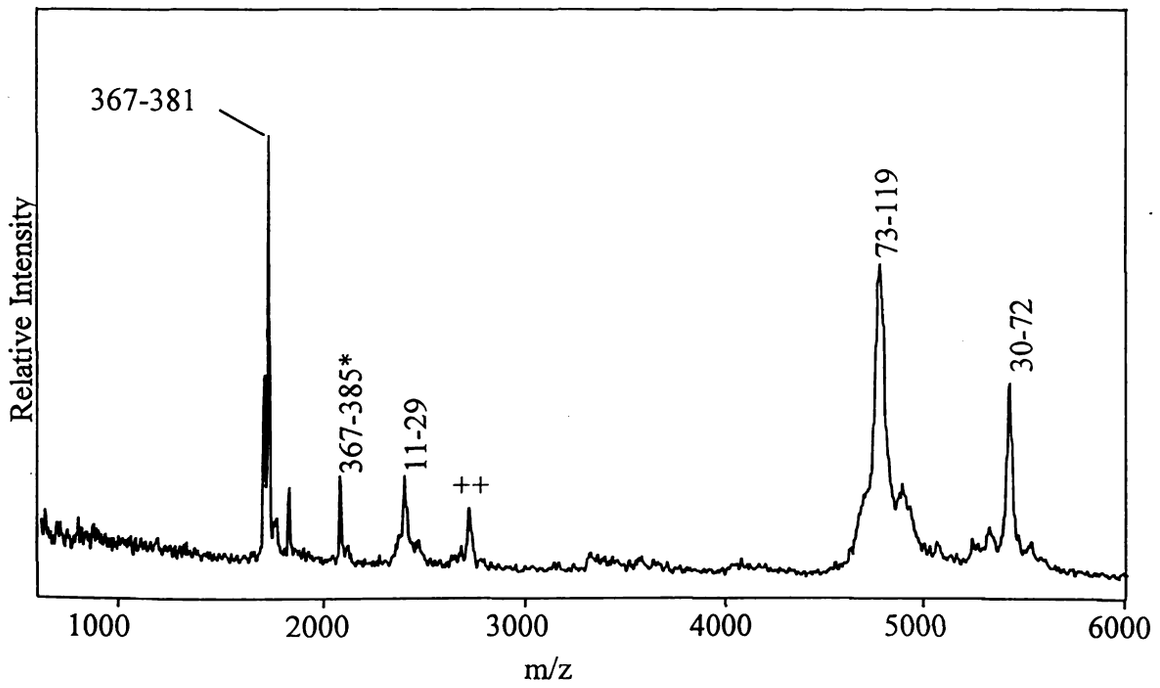


Figure 2.15. MALDI mass spectrum of 15 pmol of ovalbumin after *in situ* TCEP/NTCB treatment on a Zetabind membrane.

droplet of aqueous reaction

reach completion.

The resolubilizatio

subsequent cocrystallizati

successful analysis by MA

an aqueous solution of th

membrane and the analy

saturated with the vapor o

drying/cocrystallizing pro

improves the resolution in

After some of the

Nguyen (38) reported a s

after cyanylation on PV

MALDI-MS. In their pr

ribonuclease, bovine β-la

membrane. After rinsing

by placing the membrane

MALDI solvent by sonica

the efficiency of cleavage

partial cleavage fragmen

mixed disulfide bonds (

droplet of aqueous reaction mixture will be sustained long enough to allow the reaction to reach completion.

The resolubilization of reaction products from the membrane surface and subsequent cocrystallization of the analyte/matrix mixture are the key to eventual successful analysis by MALDI-MS. To assure the complete resolubilization of samples, an aqueous solution of the matrix in 50% acetonitrile/0.1% TFA was applied to the membrane and the analyte-matrix mixture was allowed to dry slowly in a container saturated with the vapor of the 50% acetonitrile/0.1% TFA aqueous solvent. This long drying/cocrystallizing process also helps produce fine analyte-matrix cocrystals and improves the resolution in the MALDI mass spectrum.

After some of the results presented in this chapter were published, Denslow and Nguyen (38) reported a similar approach to cleave blotted proteins at cysteine residues after cyanylation on PVDF membranes. The resulting fragments were analyzed by MALDI-MS. In their protocol, the cyanylation of 50 pmol each of bovine pancreatic ribonuclease, bovine β-lactoglobulin, and bovine insulin was carried out on a PVDF membrane. After rinsing the membrane to remove buffers, cleavage was then carried out by placing the membrane in alkaline buffer and the cleavage products were extracted into MALDI solvent by sonication. Although most of the expected fragments were detected, the efficiency of cleavage is somewhat less than that performed in free solution. Several partial cleavage fragments were found, some of them correspond to β-elimination or mixed disulfide bonds (38). These results support our conclusion that the cleavage

reaction can be achieved o

that the MALDI analysis o

V. Conclusions

We have demonstr

treatment of a protein an

simple and sensitive me

sulfhydryl residues in the

following TCEP and NTC

posttranslational modifica

reaction products. We als

remove excess reagents a

NTCB reagent is unique

protocol provides the adv

high sensitivity.

reaction can be achieved on membranes. Moreover, our experiments further demonstrate that the MALDI analysis of cleavage products can directly be performed on membranes.

V. Conclusions

We have demonstrated here that the combination of NTCB and TCEP chemical treatment of a protein and subsequent mass-mapping by MALDI-TOF MS provides a simple and sensitive methodology for determining the number and location of free sulfhydryl residues in the presence of disulfide bonds. On the other hand, mass-mapping following TCEP and NTCB procedures can be used to confirm the primary sequence and posttranslational modification of a protein and to identify the cleavage products and side reaction products. We also demonstrated that an inert Zetabind membrane can be used to remove excess reagents and other salts to improve the MALDI spectra. The use of the NTCB reagent is unique in that it specifically targets the sites being analyzed. This protocol provides the advantages of fast analysis, easy operation, high mass accuracy, and high sensitivity.

VI. References

1. Darbre, A. *Practical*

   149.

2. Chin, C. C. Q., and W

3. Hirayama, K., and Ak

   (Matsuo, T., Caprioli,

   Ltd., (1994). p. 299-3

4. Robertson, J. G., Ad

   Villafranca, J. J., *Bioc*

5. Glocker, M. O., Arbo

   Acad. Sci. USA., **91**, 5

6. Catsimpoolas, N., and

7. Jacobson, G. R., Sch

   248, 6583-6591 (197

8. Degani, Y., and Patch

9. Wakselman, M., Gui

   Chem. Commun., **1**, 2

10. Brocklehurst, K., Ma

    A., Mushiri, M. S., an

11. Brocklehurst, K., Mal

12. Schaffer, M. H., and

    (1976).

VI. References

1. Darbre, A. *Practical Protein Chemistry: A Handbook*, John wiley & Sons, (1985) p. 149.
2. Chin, C. C. Q., and Wold, F., *Anal. Biochem.*, **214**, 128-134 (1993).
3. Hirayama, K., and Akashi, S., in *Biological Mass Spectrometry: Present and Future* (Matsuo, T., Caprioli, R. M., Gross, M. L., and Seyama, Y, Eds.), John Wiley & Sons Ltd., (1994). p. 299-312.
4. Robertson, J. G., Adams, G. W., Medzihradszky, K. F., Burlingame, A. L., and Villafranca, J. J., *Biochemistry*, **33**, 11563-11575 (1994).
5. Glocker, M. O., Arbogast, B., Milley, R., Cowgill, C., and Deinzer, M. L., *Proc. Natl. Acad. Sci. USA.*, **91**, 5868-5872 (1994).
6. Catsimpoolas, N., and Wood, J. L., *J. Biol. Chem.*, **241**, 1790-1796 (1966).
7. Jacobson, G. R., Schaffer, M. H., Stark, G. R., and Vanaman, T. C., *J. Biol. Chem.* **248**, 6583-6591 (1973).
8. Degani, Y., and Patchornik, A., *Biochemistry,* **13**, 1-11 (1974).
9. Wakselman, M., Guibe-Jampel, E., Raoult, A., and Busse, W. D., *J. Chem. Soc. Chem. Commun.*, **1**, 21(1976).
10. Brocklehurst, K., Malthouse, J. P. G., Baines, B. S., Blenkinsop, R. D., Churcher, J. A., Mushiri, M. S., and Ormerod, F., *Biochem. Soc. Transactions*, **6**, 261 (1978).
11. Brocklehurst, K., Malthouse, J. P. G., *Biochem, J.*, **175**, 761 (1978).
12. Schaffer, M. H., and Stark, G. R., *Biochem. Biophys. Res. Comm.,* **71**(4), 1040-1047 (1976).

13. Casey, R., and Lang, A

14. Price N. C., *Biochem.*

15. Schaffer, M. H., and

(1976).

16. Lu, H. S., and Gracy,

17. Thomas, M. L., Janat

USA, **79**, 1054-1058 (

18. Roberts, D. D., and G

19. Matsudaira, P., Jakes,

82, 6788-6792 (1985)

20. Mahboub, S., Richard

(1986).

21. Bahler, M., Benfenati

108, 1841-1849 (198

22. Minard, P., Desmad

Biochem., **185**, 419-4

23. Nefsky, B., and Brets

24. Eshdat, Y., Chapot, M

25. Katayama, E., *J. Bioc*

26. Sutherland, C., and W

27. Tang, C., Yuksel, K.

283(1), 12-19 (1990).

13. Casey, R., and Lang, A., *Biochem. J.*, **145**, 251-261 (1975).

14. Price N. C., *Biochem. J.*, **159**, 177-180 (1976).

15. Schaffer, M. H., and Stark, G. R. *Biochem. Biophys. Res. Comm.* **71**, 1040-1047 (1976).

16. Lu, H. S., and Gracy, R. W., *Arch. Biochem. Biophys.*, **212**(2), 347-359 (1981).

17. Thomas, M. L., Janatova, J., Gray, W. R., and Tack, B. F., *Proc. Natl. Acad. Sci. USA*, **79**, 1054-1058 (1982).

18. Roberts, D. D., and Goldstein, I. J., *J. Biol. Chem.*, **259**, 909-914 (1984).

19. Matsudaira, P., Jakes, R., Cameron, L., and Atherton, E., *Proc. Natl. Acad. Sci. USA.* **82**, 6788-6792 (1985).

20. Mahboub, S., Richard, C., Delacourte, A., and Han, K., *Anal. Biochem.*, **154**, 171-182 (1986).

21. Bahler, M., Benfenati, F., Valtorta, F., Czernik, A. J., and Greengard, P., *J. Cell Biol.*, **108**, 1841-1849 (1989).

22. Minard, P., Desmadril, M., Ballery, N., Perahia, D., and Mouawad, L., *Eur. J. Biochem.*, **185**, 419-423 (1989).

23. Nefsky, B., and Bretscher, A., *Proc. Natl. Acad. Sci. USA*, **86**, 3549-3553 (1989).

24. Eshdat, Y., Chapot, M. P., and Strosberg, A. D., *FEBS Letters*, **1-2**, 166-170 (1989).

25. Katayama, E., *J. Biochem.*, **106**, 988-993 (1989).

26. Sutherland, C., and Walsh, M. P., *J. Biol. Chem.*, **264**, 578-583(1989).

27. Tang, C., Yuksel, K. U., Jacobson, T. M., and Gracy, R. W., *Arch. Biochem. Phys.*, **283**(1), 12-19 (1990).

28. May, J. M., Buchs, A.

29. Wines, B. D., and Eas
(1991)

30. Som, S., and Friedma

31. Altamirano, M. M., P
(1992).

32. Goold, R., and Baines

33. Stark, G. R. in *Meth*
(1977), Academic Pre

34. Papayannopoulos, I.A

35. Wu, J., Gage, D. A., a

36. Burns, J. A., Butler,
(1991).

37. Zaluzec, E. J., Gag
Spectrom., **5**, 230-23

38. Denslow, N. D., and
p.241.

39. Matsubara, H., and S

40. Matsubara, H., and S

41. Chakrabarti, P., *Bioc*

42. Papov, V. V., and B
Spectrometry and Al

28. May, J. M., Buchs, A. and Carter-Su, C., *Biochem.*, **29**, 10393-10398 (1990).

29. Wines, B. D., and Easterbrook-Smith, S. B., *Molecular Immunology*, **28**(8), 855-863 (1991)

30. Som, S., and Friedman, S., *J. Biol. Chem.*, **266**, 2937-2945 (1991).

31. Altamirano, M. M., Plumbridge, J. A., and Calcagno, M. L., *Biochem.*, **31**, 1153-1158 (1992).

32. Goold, R., and Baines, A. J., *Eur. J. Biochem.*, **224**, 229-240 (1994).

33. Stark, G. R. in *Methods Enzymol.* (Jakoby, W. B., and Wilchek, M., Eds.), vol. 47, (1977), Academic Press, San Diego, p. 129.

34. Papayannopoulos, I.A., and Biemann, K., *Protein Sci.*, **1**, 278-288 (1992).

35. Wu, J., Gage, D. A., and Watson, J. T., *Anal. Biochem.*, **235**, 161-174(1996).

36. Burns, J. A., Butler, J. C., and Whitesides, G. M., *J. Org. Chem.*, **56**, 2648-2650 (1991).

37. Zaluzec, E. J., Gage, D. A., Allison, J., and Watson, J. T., *J. Am. Soc. Mass Spectrom.*, **5**, 230-237 (1994).

38. Denslow, N. D., and Nguyen, H. P., in *Techniques in Protein Chemistry VII*, (1996), p.241.

39. Matsubara, H., and Sasaki, R.M., *Proc. Natl. Acad. Sci. USA.*, **57**, 439-445 (1967).

40. Matsubara, H., and Sasaki, R.M., *J. Biol. Chem.* **243**, 1732-1757 (1968).

41. Chakrabarti, P., *Biochem.*, **28**, 6081-6085(1989).

42. Papov, V. V., and Biemann, K. Proceedings of the 42nd ASMS Conference on Mass Spectrometry and Allied Topics, Chicago, IL, May29-June3, (1994), pp.688.

43. Thibault, P., Pleasance

Mass Spectrom. Ion Pr

44. Arnon, R., in *Methods*

Press, San Diego, vol.

45. Cohen, L. W., Coghlar

46. Mckenzie, H. A., Ra

(1972).

47. Patterson, S., and Katt

48. Tompson, E.O.P., and

49. Takatera, K., and Wat

50. Nisbet, A.D., Saundry

J. Biochem., **115**, 335

51. Phelan, P., and Malth

52. Voet, D., and Voet, J.

53. Kawamura, M., and N

54. Konigsberg, K., in *M*

Academic Press, San

55. Zaluzec, E. J., Gage,

(1994).

56. Fischer, W. H., Rivie

225-228 (1993).

57. Han, J. C., and Han,

43. Thibault, P., Pleasance, S., Laycock, M.V., MacKay, R.M., and Boyd, R.K., *Int. J. Mass Spectrom. Ion Proc.,* **111**, 317-353 (1991).

44. Arnon, R., in *Methods Enzymol.* (Perlmann, G. E., and Lorand, L., Eds.), Academic Press, San Diego, vol. 19, (1970), pp. 226-244.

45. Cohen, L. W., Coghlan, V. W., and Dihel, L. C., *Gene*, **48**, 219-227(1986).

46. Mckenzie, H. A., Ralston, G. B., and Shaw, D. C., *Biochem.*, **11**(24), 4539-4547 (1972).

47. Patterson, S., and Katta, V., *Anal. Chem.*, **66**, 3727-3732(1994).

48. Tompson, E.O.P., and Fisher, W.K., *Aust. J. Biol. Sci.*, **31**, 433-442 (1978).

49. Takatera, K., and Watanabe, T., *Anal. Chem.*, **65**, 3644-3646(1993).

50. Nisbet, A.D., Saundry, R.H., Moir, A.J.G., Fothergill, L.A., and Fothergill, J.E., *Eur. J. Biochem.,* **115**, 335-345 (1981).

51. Phelan, P., and Malthouse, J. P. G., *Biochem. J.*, **302**, 511-516 (1994).

52. Voet, D., and Voet, J. G., *Biochemistry*, John Wilet & Sons, (1990), p. 218-219.

53. Kawamura, M., and Nagano, K., *Biochim. Biophys. Acta,* **774**, 188 (1984).

54. Konigsberg, K., in *Methods in Enzymol.,* (Hirs, C. H. W., and Timasheff, S. N., Eds.), Academic Press, San Diego,vol. 25, pp. 185.

55. Zaluzec, E. J., Gage, D. A., Watson, J. T., *J. Am. Soc. Mass Spectrom.* **5**, 359-366 (1994).

56. Fischer, W. H., Rivier, J.E., and Craig, A.G., *Rapid Commun. in Mass Spectrom.* **7**, 225-228 (1993).

57. Han, J. C., and Han, G. Y., *Anal. Biochem.,* **220**, 5-10 (1994).

58. Allen, G., in *Labora*

(Burdon, R. H., and Kn

59. Torchinsky, Y. M., in

Oxford, (1981).

60. May, J. M., Buchs, A.,

61. Tang, C. Y., Yuksel,

Biophys., **283**, 12-19 (1

62. Goold, R., and Baines,

63. Vestling, M.M., and Fe

64. Blackledge, J.A., and A

58. Allen, G., in *Laboratory Techniques in Biochemistry and Molecular Biology,* (Burdon, R. H., and Knippenberg, P. H., Eds), Elsevier, Vol.9, (1989), pp. 318.

59. Torchinsky, Y. M., in *Surfur in Protein;* (Metzler, D., Translation Ed.), Pergamon, Oxford, (1981).

60. May, J. M., Buchs, A., and Carter-Su, C., *Biochemistry*, **29**, 10393-10398 (1990).

61. Tang, C. Y., Yuksel, U., Jacobson, T. M., and Gracy, R. W., *Arch. Biochem. Biophys.*, **283**, 12-19 (1990).

62. Goold, R., and Baines, A. J., *Eur. J. Biochem.*, **224**, 229-240 (1994).

63. Vestling, M.M., and Fenselau, C., *Anal. Chem.*, **66**, 471-477 (1994).

64. Blackledge, J.A., and Alexander, A.J., *Anal. Chem.*, **67**, 843-848 (1995).

FURTHER STUD

CYSTEINE RESIDU

MS: OP

I. Introduction

It has been known f

the N-terminus of cysteine

amino-terminal peptide and

(1, 2). The cleavage reactio

37°C for a prolonged pe

polypeptide chain at the cy

bonds are unreactive to th

frequency of the occurrenc

in size and can usually be

electrophoresis (SDS-PAG

Cyanylated cysteine can a

thiocyanate and dehydroal

mildly alkaline conditions,

a moderate degree of cleav

nearly to completion for a

Patchornik (9) found that d

CHAPTER 3

FURTHER STUDY ON THE LOCALIZATION OF PROTEIN CYSTEINE RESIDUES BY CHEMICAL CLEAVAGE AND MALDI-MS: OPTIMIZATION AND APPLICATION

I. Introduction

It has been known for some time that peptide chains can selectively be cleaved at the N-terminus of cysteine residues after cyanylation of sulfhydryl groups to form an amino-terminal peptide and a series of 2-iminothiazolidine-4-carboxylyl (ITC) peptides (1, 2). The cleavage reaction can be carried out in mildly alkaline conditions (pH8-10) at 37°C for a prolonged period of incubation (12-80h). Selective cleavage of the polypeptide chain at the cyanylated cysteine residues may be achieved because disulfide bonds are unreactive to the cyanylation and cleavage. Because of the widely variable frequency of the occurrence of cysteine in proteins, the resulting fragments vary greatly in size and can usually be mass-mapped by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (3-7) or more recently by mass spectrometry (2, 8). Cyanylated cysteine can also undergo a base-catalyzed β-elimination reaction to form thiocyanate and dehydroalanine. Since both cleavage and β-elimination occur under mildly alkaline conditions, the two reactions are competitive, sometimes resulting in only a moderate degree of cleavage of a peptide bond. Although the cleavage seemed to go nearly to completion for a variety of proteins in Jacobson's experiments, Degani and Patchornik (9) found that dehydroalanyl residues were formed in variable amounts from

the β-elimination of cyany

β-elimination over cleav

peptides. Higher pH was

Nakagawa *et al* (10) rece

hormone, hPTH [1-84], ir

biologically active α-amid

to produce with an 80% y

recombinant techniques.

been extensively used wit

In chapter 2, we d

in proteins using this spe

TOF MS. This methodo

simple, fast, and sensitive

proteins. One of the inhe

reaction takes too long to

exposure to alkaline sol

results.

In this chapter, v

function of pH, solvents

amino acids adjacent to

show that a higher pH

nevertheless, in contras

the β-elimination of cyanylated cysteine residues at pH 8-10. Increasing pH favored the β-elimination over cleavage, depending on the respective structural properties of peptides. Higher pH was therefore avoided in order to suppress the β-elimination. But Nakagawa *et al* (10) recently reported the cleavage of cyanylated human parathyroid hormone, hPTH [1-84], in a high concentration of ammonia in an attempt to produce a biologically active α-amidated peptide. By the specific cleavage reaction they were able to produce with an 80% yield the α-amidated peptide which could not be produced by recombinant techniques. In all, the protocol (1, 9) developed by Jacobson *et al* (1) has been extensively used with little modification over the past two decades.

In chapter 2, we described a new methodology to localize free sulfhydryl groups in proteins using this specific cleavage reaction followed by mass-mapping by MALDI-TOF MS. This methodology is advantageous over conventional approaches in that it is simple, fast, and sensitive, and is especially useful for locating sulfhydryl groups in large proteins. One of the inherent problems with this procedure, however, is that the cleavage reaction takes too long to accomplish. Various side reactions may rise after long hours of exposure to alkaline solution, which brings about some uncertainty into the analytical results.

In this chapter, we present our systematic studies on the cleavage reaction as a function of pH, solvents, and amino acid structures for polypeptides containing different amino acids adjacent to N-terminus of cyanylated cysteines. Our experimental results show that a higher pH can greatly promote both cleavage and β-elimination reactions, nevertheless, in contrast to the previous reports, for most of the peptides studied, the

extent of β-elimination d

Ammonia, a stronger nucl

of peptide chains and min

most of the peptides stud

within an hour at room ter

effects of peptide structur

The optimized conditions

in a variety of proteins wi

II. Experimental Sectio

Mass Spectrometry

MALDI mass spe

spectrometer (PerSeptive

VSL-337ND nitrogen la

the ion source was set t

Time-to-mass conversio

standards of bradykinin

skeletal myoglobin (m/z

All experiments were

Chemical Co., Milwauk

in a 50% (v/v) solution

extent of β-elimination does not increase significantly even at a pH as high as 12. Ammonia, a stronger nucleophile than hydroxyl anions, greatly accelerates the cleavage of peptide chains and minimizes side reactions related to the prolonged incubation. For most of the peptides studied, the cleavage in 1 M NH_4OH solution can be complete within an hour at room temperature. Based on the results from over a dozen peptides, the effects of peptide structure on the rate and yield of the cleavage reaction were evaluated. The optimized conditions have been applied to the recognition of free sulfhydryl groups in a variety of proteins with both known and unknown cysteine structures.

II. Experimental Section

Mass Spectrometry

MALDI mass spectra were obtained on a Voyager Elite time-of-flight (TOF) mass spectrometer (PerSeptive Biosystems Inc., Framingham, MA) equipped with a model VSL-337ND nitrogen laser (Laser Science, Newton, MA). The accelerating voltage in the ion source was set to 20 kV. Data were acquired in the linear mode of operation. Time-to-mass conversion was achieved by external and/or internal calibration using standards of bradykinin (m/z 1061.2), bovine pancreatic insulin (m/z 5734.5), and horse skeletal myoglobin (m/z 16,952) obtained from Sigma Chemical Co. (St. Louis, MO). All experiments were performed using α-cyano-4-hydroxycinnamic acid (Aldrich Chemical Co., Milwaukee, WI) as the matrix. Saturated matrix solutions were prepared in a 50% (v/v) solution of acetonitrile/aqueous 0.1% TFA, and mixed in equal volumes

with peptide or protein s

mixture was allowed to ai

Chemicals

Guanidine hydroc

(Indianapolis, IN). 1-C

peptides TCVEWLRRYI

ovalbumin, and rabbit m

used without further puri

DRVYIHPCHLLYYS, A

American Peptide Comp

California Inc. The 0.10

HCl was freshly prepare

prepared in 0.1 M citrate

Cyanylation of SH Gro

To the peptide a

cyano-4-dimethylamino-

sulfhydryl content. Cya

room temperature for 1

reversed-phase HPLC

manually, the masses of

with peptide or protein samples, and applied to a stainless-steel sample plate. The mixture was allowed to air dry before being introduced into the mass spectrometer.

Chemicals

Guanidine hydrochloride was a product of Boehringer-Mannheim Biochemicals (Indianapolis, IN). 1-Cyano-4-dimethylamino-pyridinium tetrafluoroborate (CDAP), peptides TCVEWLRRYLKN and RYVVLPRPVCFEKGMNYTVR, spinach ferredoxin, ovalbumin, and rabbit muscle creatine phosphokinase were purchased from Sigma and used without further purification. Acetonitrile and TFA were of HPLC grade. Peptides DRVYIHPCHLLYYS, ALLETYCATPAKSE, and SLRRSSCFGGR were products of American Peptide Company. The rest of the peptides were purchased from Bachem California Inc. The 0.10 M CDAP solution in pH 3.0, 0.1 M citrate buffer-4M guanidine-HCl was freshly prepared prior to use. The 1 mM protein and peptide solutions were prepared in 0.1 M citrate buffer, pH 3.0, containing 4 M guanidine-HCl.

Cyanylation of SH Groups

To the peptide and protein solutions was added a 10-fold molar excess of 1-cyano-4-dimethylamino-pyridinium tetrafluoroborate (CDAP) solution over the sulfhydryl content. Cyanylation of sulfhydryl groups was accomplished by incubation at room temperature for 10-15 min. The modified proteins or peptides were purified by reversed-phase HPLC under gradient elution. The HPLC fractions were collected manually, the masses of which were determined by MALDI-MS. Those corresponding to

the cyanylated proteins ar

further use.

Cleavage of the Cyanyla

The cyanylated pe

4 M guanidine-HCl solut

equal volume of 4 M gua

0.25 M tris-HCl buffer, (

The pH 8.0 and pH 9.0 v

the cleavage reaction, w

ambient temperature for

were taken and diluted to

MS. The rest of the so

HPLC fractions were co

of the cleavage and β-eli

under the peaks identifie

The cleavage c

phosphokinase at a conc

tris-HCl buffer, pH 9.0,

NH_4OH solution-4 M g

Aliquots of 1 µl of th

CH_3CN/0.1%TFA for

the cyanylated proteins and peptides were dried in a speed vac and kept in a freezer for further use.

Cleavage of the Cyanylated Peptides and Proteins

The cyanylated peptide samples were reconstructed to a concentration of 1 mM in 4 M guanidine-HCl solution. A portion of the above solutions (~5 μl) was mixed with an equal volume of 4 M guanidine-HCl solution in pH 8.0, 0.25 M tris-HCl buffer, pH 9.0, 0.25 M tris-HCl buffer, 0.02 M NaOH (pH 12), and 2 M NH_4OH solution, respectively. The pH 8.0 and pH 9.0 vials were incubated in 37°C water bath for 18 hours to promote the cleavage reaction, while the vials with NaOH and NH_4OH solutions were placed at ambient temperature for 1 hour. After the reaction, 1-μl aliquots of the above solutions were taken and diluted to 100 μl with 1:1 CH_3CN/0.1%TFA for analysis by MALDI-TOF MS. The rest of the solutions was acidified and injected into an HPLC column. The HPLC fractions were collected and identified by MALDI-TOF MS. The relative yields of the cleavage and β-elimination products were calculated by integrating the HPLC areas under the peaks identified by MALDI.

The cleavage of cyanylated proteins, ferredoxin, ovalbumin, and creatine phosphokinase at a concentration of 0.5 mM, was carried out by incubation both in 0.1 M tris-HCl buffer, pH 9.0, containing 4 M guanidine-HCl at 37°C for 18 hours and in 1 M NH_4OH solution-4 M guanidine-HCl at ambient temperature for one hour, respectively. Aliquots of 1 μl of the above solutions were taken and diluted to 100 μl with 1:1 CH_3CN/0.1%TFA for analysis by MALDI-TOF MS. The rest of the solution was

acidified and then injecte

identified by MALDI-TO

HPLC Separation

The purification

cleavage products were

using Waters model 600

nm. The column used th

size, 300-Å pore, 4.6×2

solution and CH_3CN co

from 5-50% B in 50 r

conditions for individua

major HPLC fractions

determined by MALDI-

III. Results and Discus

A. Cyanylation of SH

The cyanylation

thiocyanobenzoic acid (

was believed to be spe

formation of mixed disu

acidified and then injected to the HPLC column. The HPLC fractions were collected and identified by MALDI-TOF MS.

HPLC Separation

The purification of the cyanylated proteins and peptides and the separation of cleavage products were achieved by reversed-phase HPLC with linear gradient elution using Waters model 6000 pumps controlled by a PC computer. UV detection was at 215 nm. The column used throughout the study was a Vydac C18 (#218TP54, 10 μm particle size, 300-Å pore, 4.6×250 mm). Mobile phases A and B contain 0.1% TFA aqueous solution and CH_3CN containing 0.1% TFA, respectively. Typically, a gradient elution from 5-50% B in 50 minutes was applied in preliminary experiments. The HPLC conditions for individual proteins and peptides were then modified and optimized. The major HPLC fractions were collected manually, and the masses of the fractions were determined by MALDI-MS.

III. Results and Discussion

A. Cyanylation of SH Groups by CDAP

The cyanylation of sulfhydryl groups was traditionally accomplished by 2-nitro-5-thiocyanobenzoic acid (NTCB) under mildly alkaline conditions (pH 8-10). The reagent was believed to be specific to sulfhydryl groups, although side reactions, such as the formation of mixed disulfide bonds between the NTCB and protein SH groups, were also

reported (9, 11, 12). Sinc

under alkaline conditions

be difficult even though t

1-cyano-4-dimethylamin

cyanylation of SH group

reagent has not been use

The CDAP is adv

groups in the presence c

minimize sulfhydryl/dis

peptides are stable und

subsequent cleavage rea

showed that complete cy

molar excess of the rea

temperature. Our study

the published results. F

excess of CDAP (~50-f

hours) could result in t

product can be separ

implicating the notable

Mass analysis of the sic

corresponding cyanylate

by CDAP needs to be

reaction. After extensiv

reported (9, 11, 12). Since both the cyanylation and the consequent cleavage are achieved under alkaline conditions, the separate study of cyanylation and cleavage reactions might be difficult even though the reaction conditions are carefully controlled. Another reagent, 1-cyano-4-dimethylamino-pyridinium tetrafluoroborate (CDAP), was proposed for the cyanylation of SH groups under slightly acidic conditions (pH 3-7) (14). However, this reagent has not been used as extensively as the NTCB.

The CDAP is advantageous over the NTCB for specific cyanylation of protein SH groups in the presence of disulfide bonds, because the acidic conditions can effectively minimize sulfhydryl/disulfide exchange. Furthermore, the cyanylated proteins or peptides are stable under acidic conditions, the kinetics of the cyanylation and the subsequent cleavage reaction can be studied independently. Previous papers (10, 14, 15) showed that complete cyanylation by the CDAP can be carried out using three to five fold molar excess of the reagent over free SH groups under very mild conditions at room temperature. Our study with sulfhydryl-containing peptides listed in Table 3.1 confirmed the published results. Further experiments with the model peptides showed that a large excess of CDAP (~50-fold over peptide SH groups) and excessive incubation time (> 2 hours) could result in the formation of an unidentified side reaction product. The side product can be separated from the corresponding cyanylated peptide on HPLC, implicating the notable differences between the hydrophobicity of the two products. Mass analysis of the side product by MALDI shows a mass increase of 43 Da from the corresponding cyanylated peptide. Therefore, like cyanylation by NTCB, the cyanylation by CDAP needs to be performed under controlled conditions to minimize the side reaction. After extensive examination with different peptides, we optimized the reaction

Table 3.1. Calculated m
cleavage reaction

amino acid sequence
PHCKRM
RGPCRAFI
SRNRCNDQ
SLRRSSCFGGR
MSRPACPNDKYE
YKTTICGKGLSATV
EKPLQNFTLCFR
ALLETYCATPALSE
TCVEWLRRYLKN
DRVYIHPCHLLYYS
KRNPGSNKRFPSNCG
RYVVLPRPVCFEKGM
MHRQEAVDCLKKFN

Table 3.1. Calculated m/z values of fragments for sulfhydryl-containing peptides after the cleavage reaction

amino acid sequence	M.W.	$[M+H]^+$		
		fragment I	fragment II	β-elimination
PHCKRM	771.0	253.3	562.7	738.0
RGPCRAFI	919.3	329.3	636.0	887.3
SRNRCNDQ	991.2	532.6	503.6	958.2
SLRRSSCFGGR	1225.7	705.9	564.8	1192.7
MSRPACPNDKYE	1410.0	561.7	892.3	1377.0
YKTTICGKGLSATV	1441.9	625.8	861.1	1407.9
EKPLQNFTLCFR	1493.9	1088.4	450.5	1460.9
ALLETYCATPALSE	1496.6	709.7	831.9	1463.6
TCVEWLRRYLKN	1580.9	120.1	1506.8	1547.9
DRVYIHPCHLLYYS	1779.4	900.1	924.3	1746.4
KRNPGSNKRFPSNCGRD	1947.3	1517.8	476.5	1914.3
RYVVLPRPVCFEKGMNYTVR	2427.9	1099.5	1373.5	2394.9
MHRQEAVDCLKKFNARRKLKGA	2600.4	986.1	1659.3	2567.4

conditions: the cyanylatio
CDAP (over sulfhydryl
Under these conditions t
reaction was observed.

B. Optimization of Cle

Although the sele
of cyanylated cysteine re
today are almost the sam
(1). Degani and Patchor
reactions and the effects
elimination are both bas
while β-elimination see
conditions (pH 8-9) w
incubation time differs
hours, depending on the
also affect the rate of cl
cleavage, no appreciab
conclusion made by D
dipeptides and tripeptid
model peptides represe

conditions: the cyanylation was carried out by treatment with a 10-fold molar excess of CDAP (over sulfhydryl groups) in pH 3.0 buffer for 15 min at ambient temperature. Under these conditions the cyanylation reaction went to completion and very little side reaction was observed.

B. Optimization of Cleavage Reactions

Although the selective chemical cleavage of peptide chains at the N-peptide bonds of cyanylated cysteine residues was proposed long ago, the experimental conditions used today are almost the same as those described in the protocol developed by Jacobson *et al* (1). Degani and Patchornik (9) studied the pH dependence of cleavage and β-elimination reactions and the effects of amino acid structures on both reactions. The cleavage and β-elimination are both base catalyzed and therefore competitive under alkaline conditions, while β-elimination seems to be favored over cleavage at higher pH. Mildly alkaline conditions (pH 8-9) were suggested for the cleavage reaction in early study. The incubation time differs significantly under the above conditions, ranging from 16 to 80 hours, depending on the structures of proteins under study. Structural features of proteins also affect the rate of cleavage reaction (16). Proteins have to be denatured prior to the cleavage, no appreciable cleavage was observed for native proteins. However, the conclusion made by Degani and Patchornik was based on the study on a few simple dipeptides and tripeptides with special structural factors. The behavior of cysteines in the model peptides represents to a less extent general structural features of cysteine residues

in peptides and proteins.

better understanding and

We have systema

function of pH and reac

amino acids adjacent to

were first cyanylated and

cyanylated peptides, afte

12.0, and 1 M NH_4OH

The cleavage mixtures w

fractions identified by M

side reactions were eval

Results of the k

cleavage can be accom

within one hour at amb

37°C in pH 8-9 buffers.

the greatest concerns e

favor β-elimination ov

However, our study on

that it is not the case.

differ significantly fo

conditions. Some of t

in peptides and proteins. Further examination of the cleavage reaction is needed for the better understanding and optimization of the cleavage reaction.

We have systematically studied the rate and yield of the cleavage reaction as a function of pH and reaction medium for a variety of polypeptides containing different amino acids adjacent to the cyanylated cysteines. The peptides under study (Table 3.1) were first cyanylated and the completion of cyanylation was monitored by MALDI. The cyanylated peptides, after purified by HPLC, were subjected to cleavage at pH 8.0, 9.0, 12.0, and 1 M NH_4OH (pH~11) solutions containing 4 M guanidine-HCl, respectively. The cleavage mixtures were separated by HPLC, the fractions collected, the masses of the fractions identified by MALDI. The relative yields of the cleavage products and other side reactions were evaluated by measuring HPLC peak areas.

Results of the kinetics study showed that, for most of the peptides studied, the cleavage can be accomplished in both ammonia solution and 0.01 M NaOH solution within one hour at ambient temperature, whereas it requires 18 hours of incubation at 37°C in pH 8-9 buffers. Although elevated pH can accelerate cleavage reactions, one of the greatest concerns expressed in previous papers was that a higher pH would largely favor β-elimination over cleavage, causing a lower yield of the cleavage reaction. However, our study on the relative yields under various pH and reaction media indicated that it is not the case. As shown in Table 3.2, the relative yields of β-elimination do not differ significantly for most of the peptides studied under the four experimental conditions. Some of the peptides show even higher cleavage yields in 1 M NH_4OH or

Table 3.2. Relative y

amino acid sequenc	
PHCKRM[d]	
RGPCRAFI	
SLRRSSCFGGR	
MSRPACPNDKYE	
YKTTICGKGLSATV	
EKPLQNFTLCFR	
ALLETYCATPALSE	
TCVEWLRRYLKN	
DRVYIHPCHLLYYS	
KRNPGSNKRFPSNC	
RYVVLPRPVCFEKG	
MHRQEAVDCLKKFN	

[a] in each case, β-elimin
[b] side reactions other th
[c] peptide undergoes ap
[d] incomplete conversio

Table 3.2. Relative yields (%) of β-elimination under different cleavage conditions

amino acid sequence	pH 8.0, 37°C,18h	pH 9.0, 37°C,18h	pH 12 , rt, 1h	1M NH_4OH, rt, 1h
PHCKRM[d]	9.9[a]	15.9	25.1	14.8
RGPCRAFI	31.6[b]	87.5	82.5	70.1
SLRRSSCFGGR	2.8	2.9	3.9	4.6
MSRPACPNDKYE	6.1	5.8	11.3	3.9
YKTTICGKGLSATV	16.8	18.2	22.2	20.3
EKPLQNFTLCFR	15.2[c]	29.9	49.8	20.9
ALLETYCATPALSE	~100	~100	~100	~100
TCVEWLRRYLKN	20.9	13.8	13.1	18.3
DRVYIHPCHLLYYS	56.2	90.7	98.5	93.8
KRNPGSNKRFPSNCGRD	5.0	5.0	5.0	5.0
RYVVLPRPVCFEKGMNYTVR	34.3	41.2	52.4	51.5
MHRQEAVDCLKKFNARRKLKGA	12.8	11.1	18.4	13.2

[a] in each case, β-elimination is a main side reaction

[b] side reactions other than β-elimination and dimerization are observed

[c] peptide undergoes appreciable dimerization

[d] incomplete conversion

Figure 3.1
(B) peptide

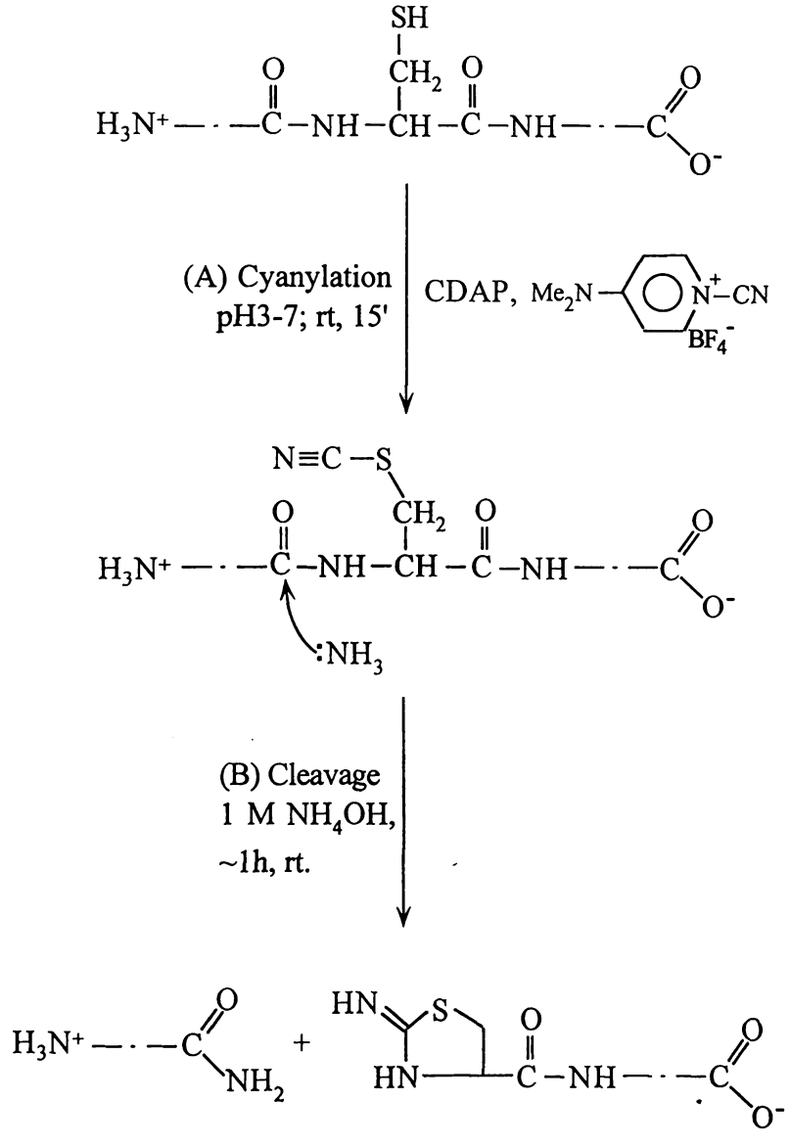


Figure 3.1. (A) Cyanylation of sulfhydryl group by CDAP and (B) peptide bond cleavage catalyzed by ammonia

NaOH solutions. These

reaction under stronger al

The mechanism c

2) indicates that the hyd

rising pH on the rate of

nucleophilic attack of C

mechanism, a stronger

condition that the size

nucleophilic group to t

requirement. Ammonia

the hydroxyl anion, the

carbonyl carbon would

as shown in Figure 3.1.

hydroxyl anion catalyze

structure replaces carbo

Although compa

and 1 M NH_4OH soluti

1 M NH_4OH solution p

0.01 M OH^- is available

agent will greatly accel

is a better nucleophile t

anion in terms of the c

NaOH solutions. These results demonstrate the promise to carry out the cleavage reaction under stronger alkaline conditions without sacrificing the cleavage yields.

The mechanism of cleavage reaction proposed by Jacobson *et al* (1) (see chapter 2) indicates that the hydroxide ion catalyzes the cleavage, as shown by the influence of rising pH on the rate of cleavage reaction. The reaction probably proceeds only after nucleophilic attack of OH^- on the carbonyl carbon of the amide. According to the mechanism, a stronger nucleophile should promote the cleavage reaction, given the condition that the size of the nucleophilic group is small so that the attack of the nucleophilic group to the carbonyl atom is sterically feasible. Ammonia meets the requirement. Ammonia has a stronger nucleophilic property than hydroxyl anion. Like the hydroxyl anion, the size of ammonia is small. Nucleophilic attack of ammonia to the carbonyl carbon would facilitate the cleavage giving an α-amidated N-terminal peptide, as shown in Figure 3.1. Therefore, the N-terminal fragments obtained by ammonia and hydroxyl anion catalyzed cleavage have one mass difference in mass spectrometry (amide structure replaces carboxyl structure).

Although comparable results were obtained for peptide cleavage in 0.01 M NaOH and 1 M NH_4OH solutions, the latter is preferred as it provides several advantages. First, 1 M NH_4OH solution provides much more nucleophilic species ($:NH_3$ ~1 M), while only 0.01 M OH^- is available in pH 12 NaOH solution. The high concentration of nucleophilic agent will greatly accelerate the concerted cleavage reaction. Second, although ammonia is a better nucleophile than hydroxyl anion, it has a smaller proton affinity than hydroxyl anion in terms of the capability of catalyzing β-elimination reaction. Therefore, the β-

elimination is anticipated

from our experimental

NaOH solution are high

solution containing 4 M

chain is more stable ar

Finally, the excess vol

cleavage.

The representativ

in 1 M NH_4OH solution

gave comparable result

elimination reactions ar

typical MALDI spectra

NH_4OH solution. The r

a main competitive side

observed. One such mi

free SH group again,

reported by others (9, 1

which probably arises

groups. However, th

negligible.

elimination is anticipated to be minor in 1 M NH_4OH. This tendency has been supported from our experimental results (Table 3.2) that the yields of β-elimination in 0.01 M NaOH solution are higher than that in 1 M NH_4OH solution. Third, the 1 M NH_4OH solution containing 4 M guanidine-HCl has a moderate pH (~11), at which the peptide chain is more stable and alkali-induced damage of peptide chains can be prevented. Finally, the excess volatile ammonia can easily be removed from the mixture after cleavage.

The representative chromatograms of peptides, after cleavage in pH 9.0 buffer and in 1 M NH_4OH solution, are presented in Figure 3.2 to Figure 3.5. The two conditions gave comparable results in terms of relative yields of competitive cleavage and β-elimination reactions and other possible side reactions. Figure 3.6 to Figure 3.9 show typical MALDI spectra of these peptides, after cleavage in pH 9.0 buffer and in 1 M NH_4OH solution. The results from both HPLC and MALDI confirm that β-elimination is a main competitive side reaction with cleavage reaction. Other side reactions were also observed. One such minor reaction was the conversion of cyanylated peptides to form a free SH group again, indicating the reversibility of the cyanylation reaction as also reported by others (9, 10). Another minor side reaction was the dimerization of peptides, which probably arises from nucleophilic attack of nascent thioate anions on thiocyano groups. However, these two side reactions were always minor and in most cases negligible.

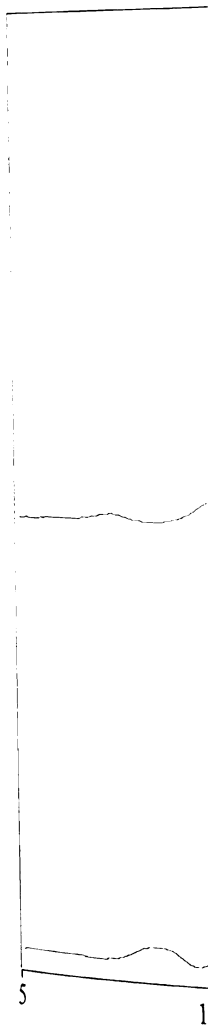


Figure 3.2. HPLC
in (A) pH 9.0, 37°C
I, II, β, C/U, and D
uncleaved species,

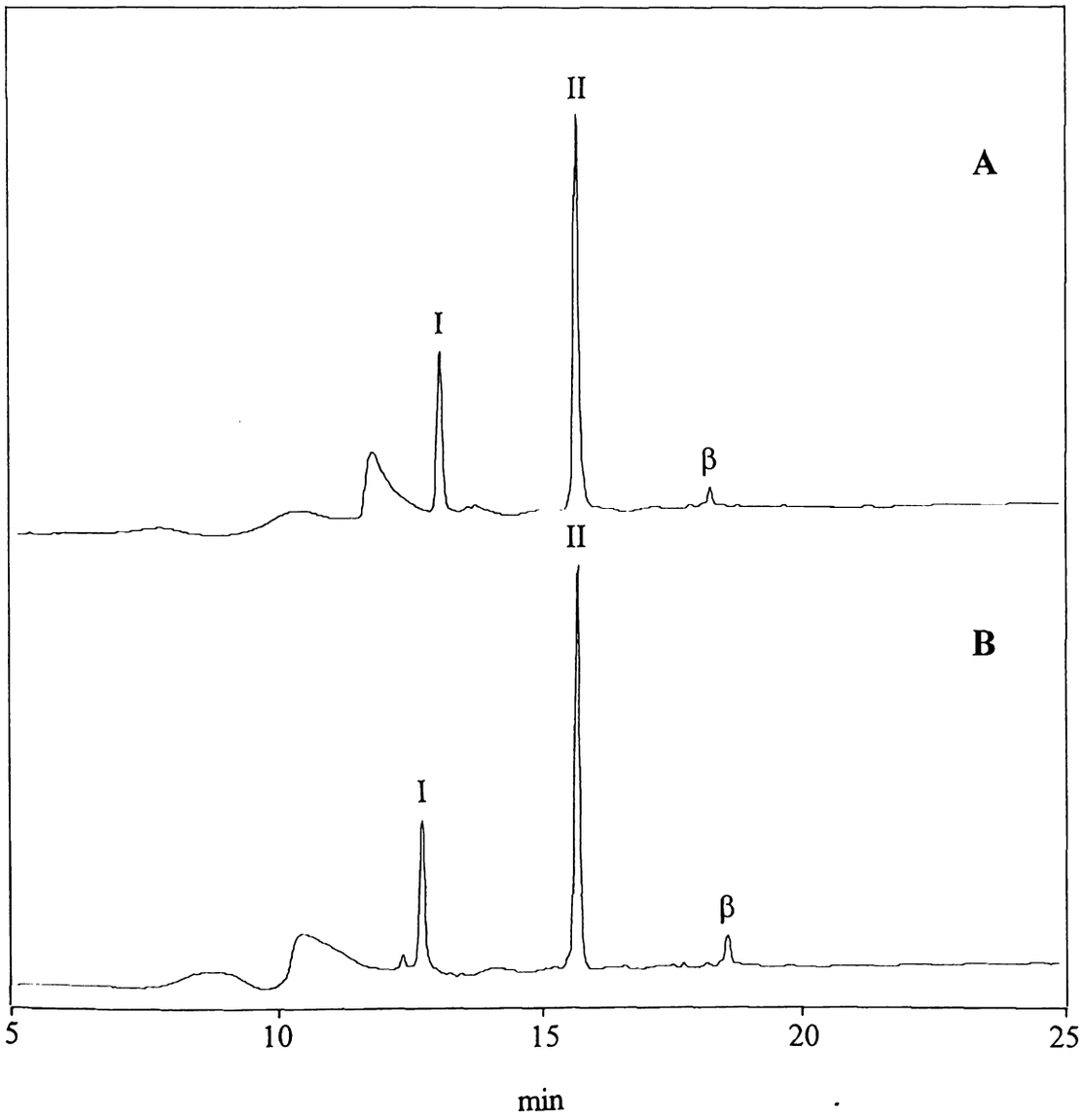


Figure 3.2. HPLC chromatograms of cleavage products of SLRRSSCFGGR in (A) pH 9.0, 37°C, 18h, and (B) 1 M NH_4OH, rt, 1h. The peaks marked with I, II, β, C/U, and D are fragment I, fragment II, β-elimination product, cyanylated/ uncleaved species, and peptide dimer, respectively.

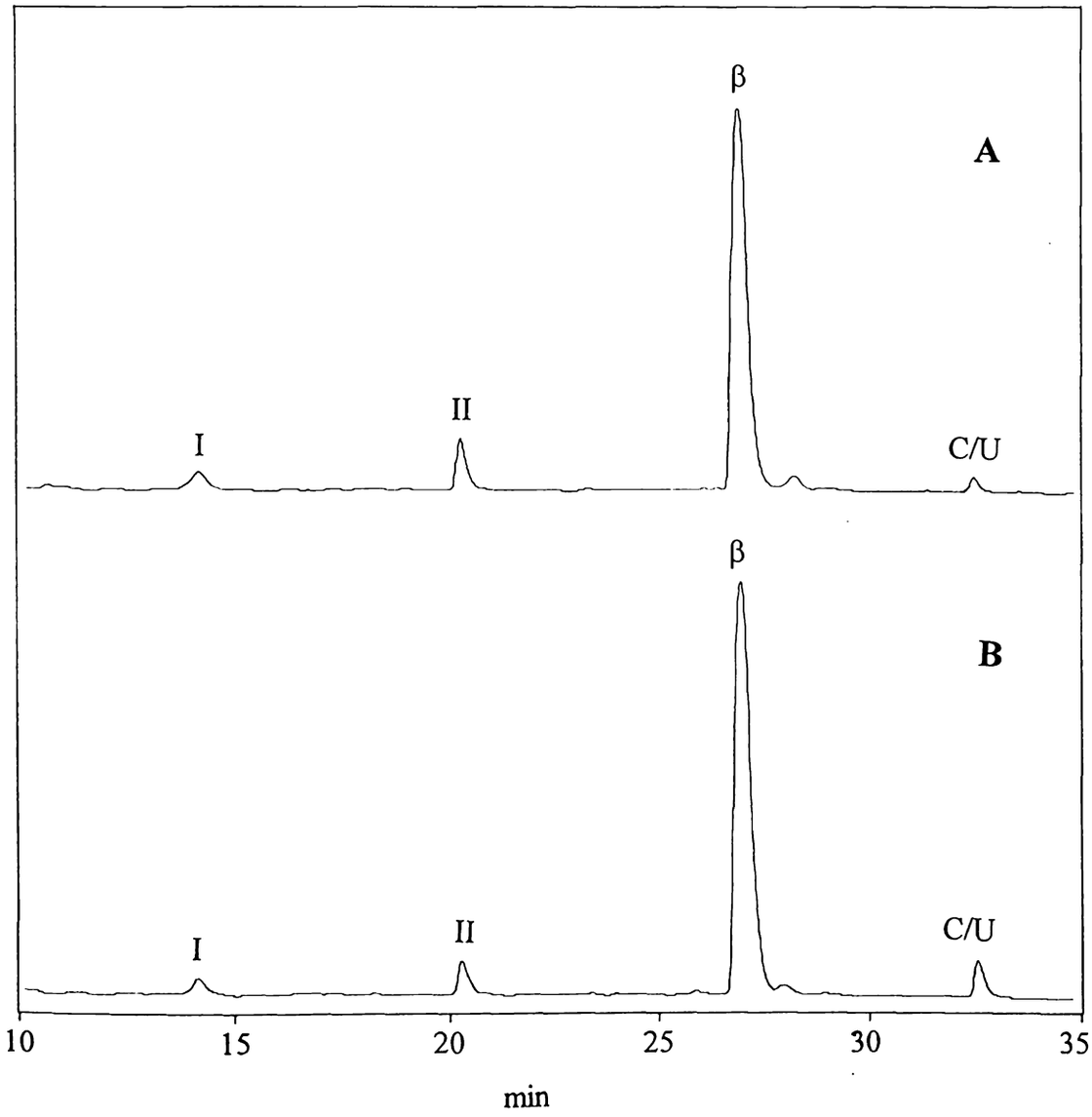


Figure 3.3. HPLC chromatograms of cleavage products of DRVYIHPCHLLYYS in (A) pH 9.0, 37°C, 18h, and (B) 1 M NH_4OH, rt, 1h, respectively.

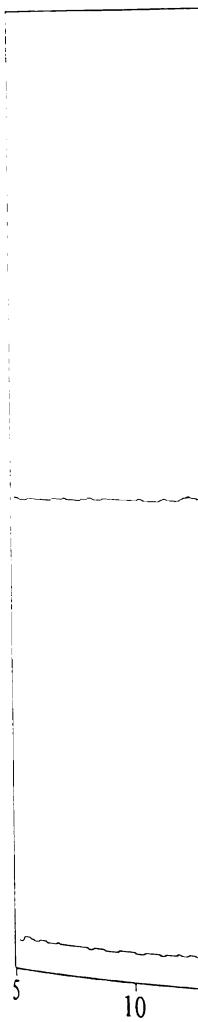


Figure 3.4. HPL
in (A) pH 9.0, 37

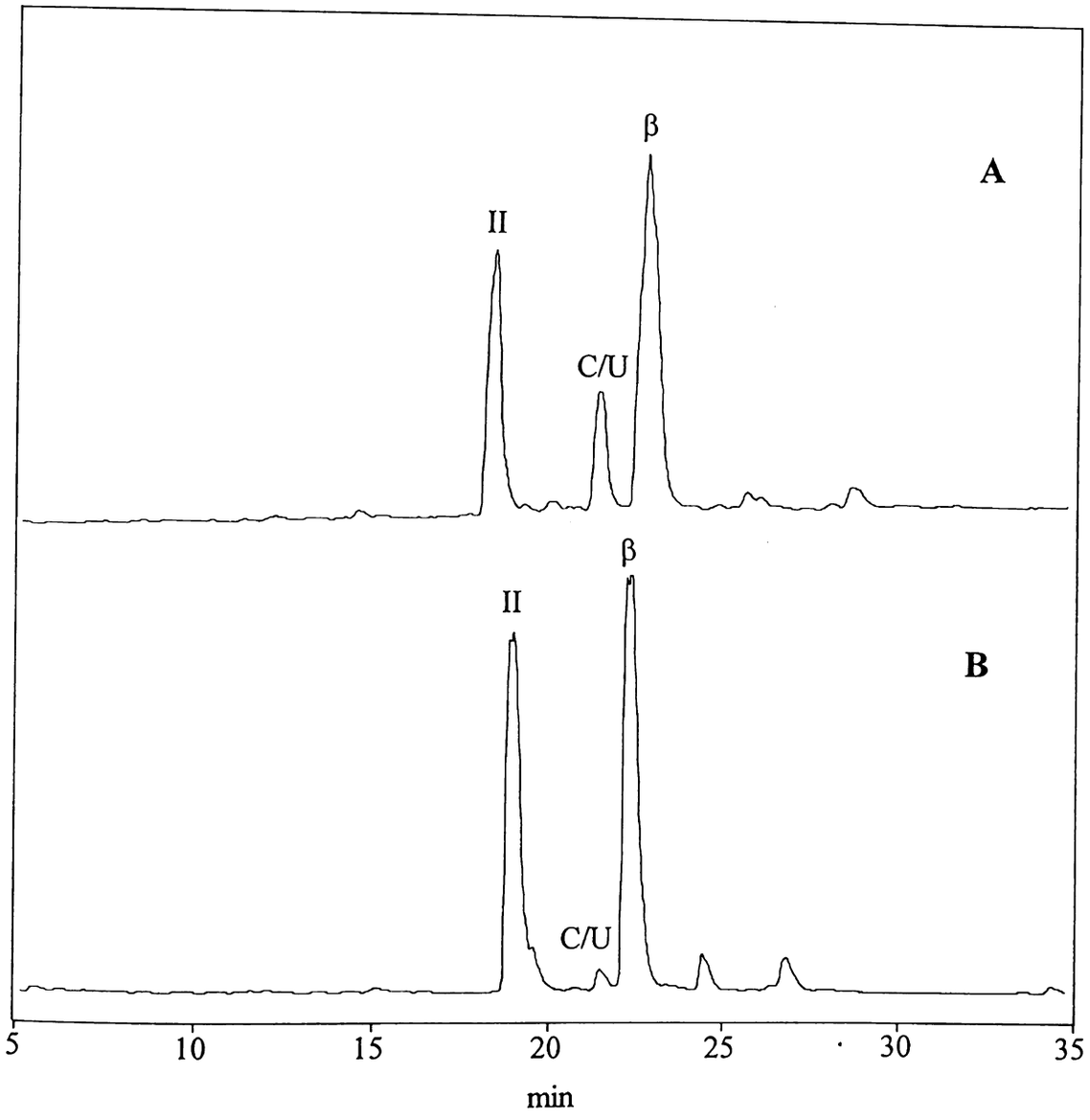


Figure 3.4. HPLC chromatograms of cleavage products of EKPLQNFTLCFR in (A) pH 9.0, 37°C, 18h, and (B) 1 M NH_4OH, rt, 1h, respectively.

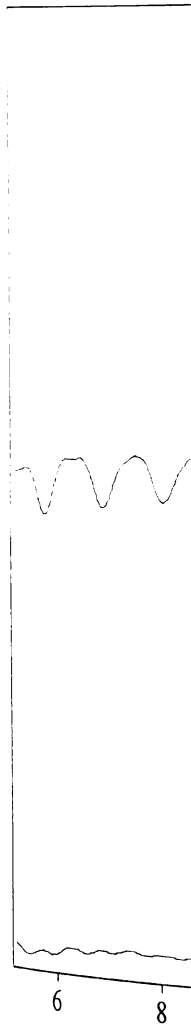


Figure 3.5. HPL
MHRQEAVDCL
NH_4OH, rt, 1h, r

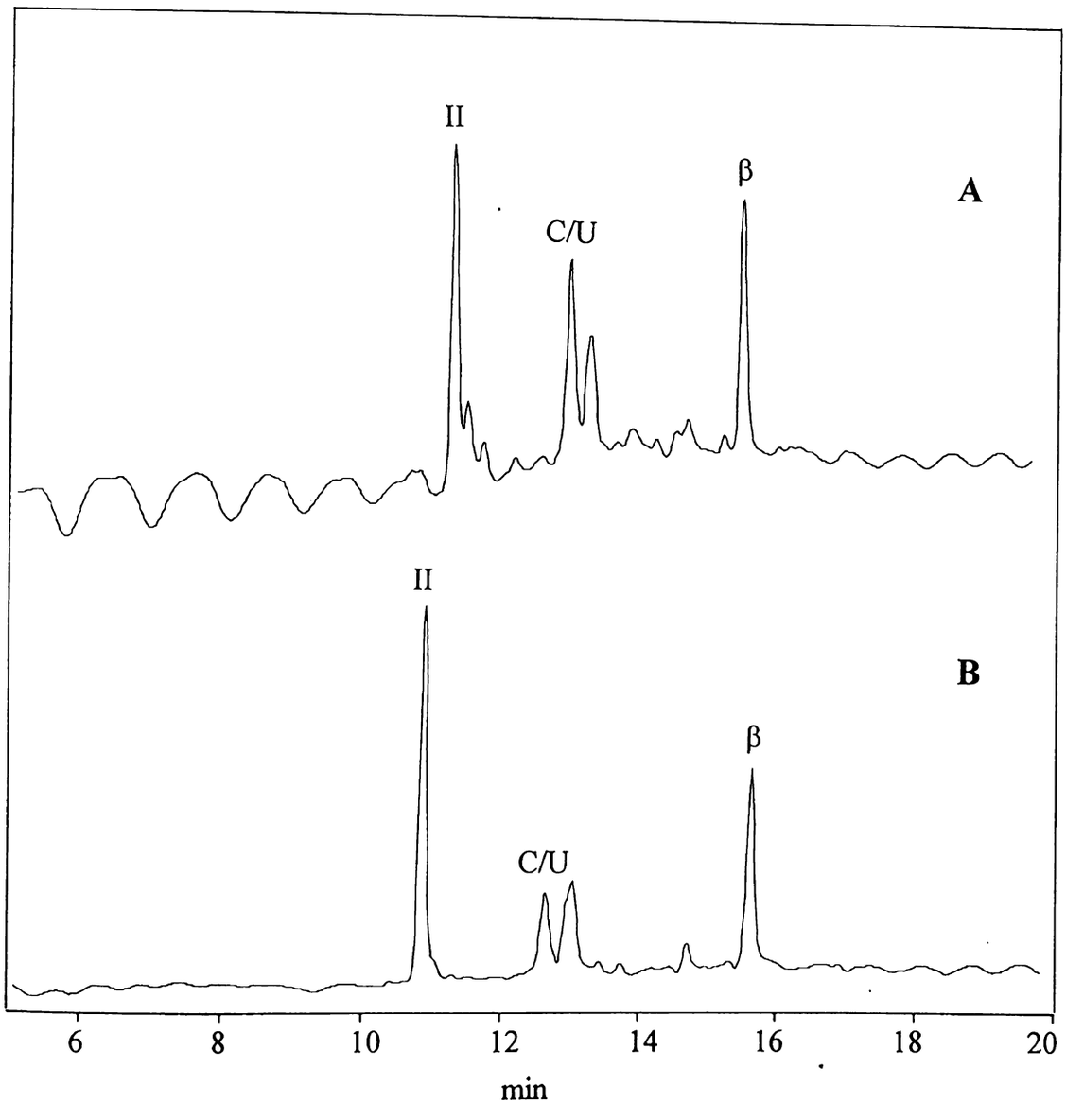


Figure 3.5. HPLC chromatograms of cleavage products of peptide MHRQEAVDCLKKFNARRKLKGA in (A) pH 9.0, 37°C, 18h, and (B) 1 M NH_4OH, rt, 1h, respectively.

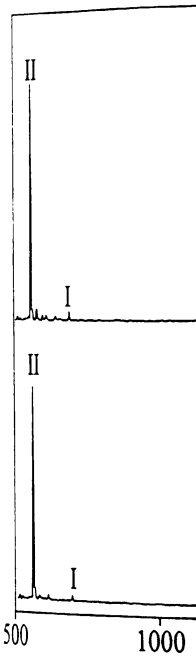


Figure 3.6. MALDI
in (A) pH 9.0, 37°C,
I, II, D, C/U, and β re
uncleaved peptide, ar

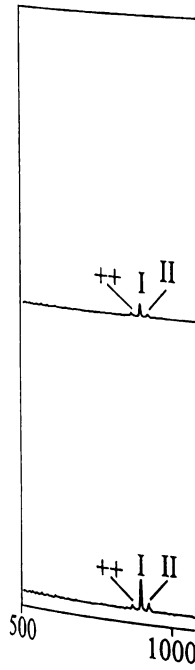


Figure 3.7. MALD
cleavage in (A) pH
Symbol ++ represe

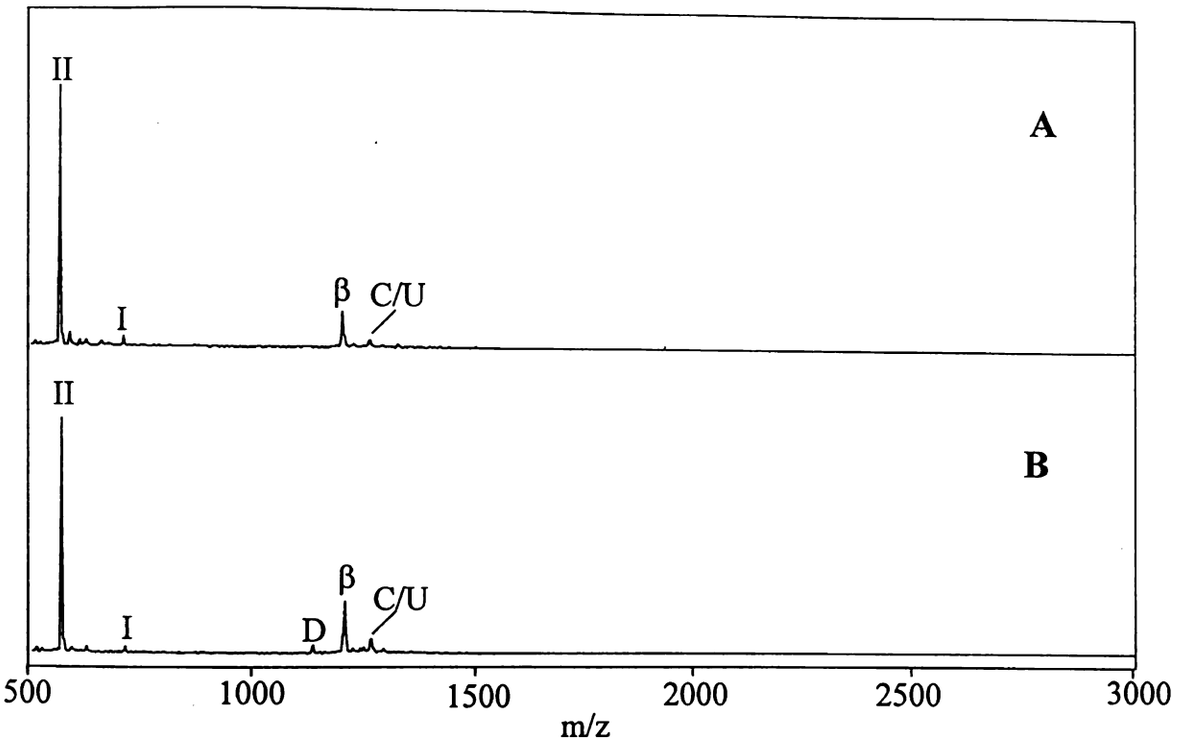


Figure 3.6. MALDI spectra of the cyanylated peptide SLRRSSCFGGR, after cleavage in (A) pH 9.0, 37°C, 18h and (B) 1 M NH_4OH solution, rt, 1h, respectively. Symbols I, II, D, C/U, and β represent fragment I, fragment II, peptide dimer, cyanylated/ uncleaved peptide, and its β-elimination product, respectively.

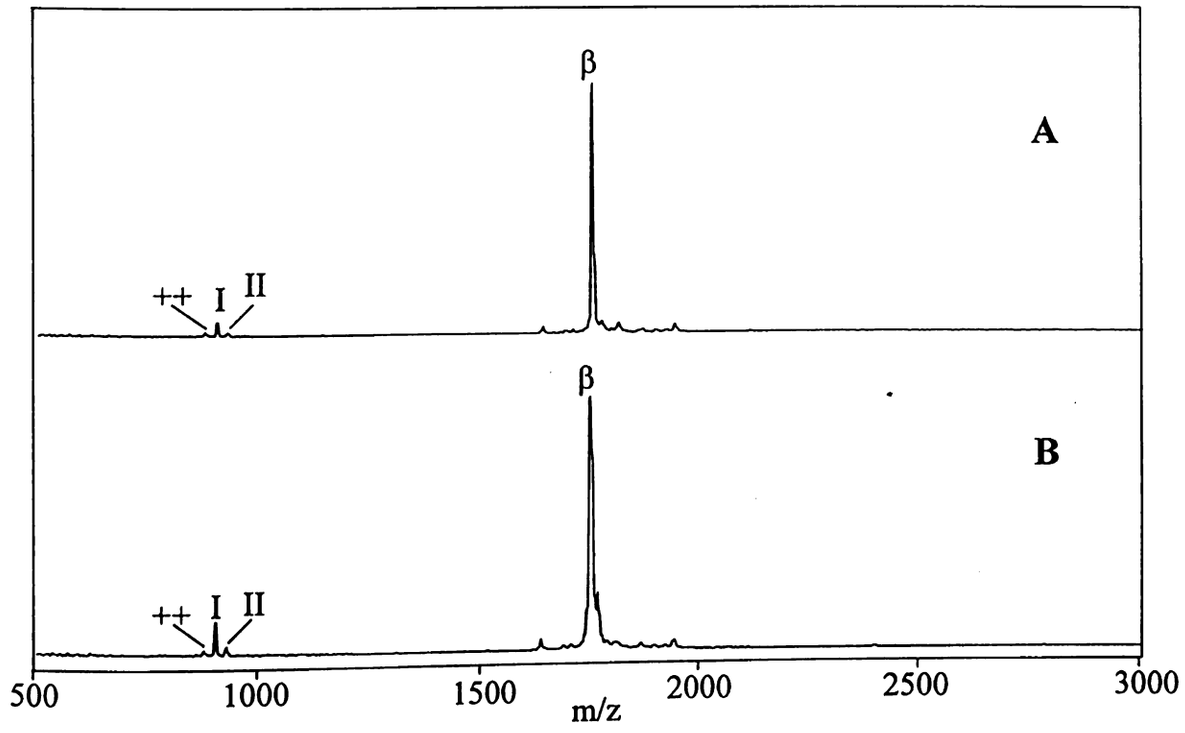


Figure 3.7. MALDI spectra of the cyanylated peptide DRVYIHPCHLLYYS, after cleavage in (A) pH 9.0, 37°C, 18h and (B) 1 M NH_4OH solution, rt, 1h, respectively. Symbol ++ represents a doubly charged species.

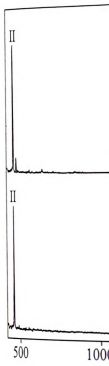


Figure 3.8. MALDI sp
in (A) pH 9.0, 37°C, 1
IP represents the intac

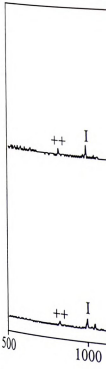


Figure 3.9. MALDI
after cleavage in (A)

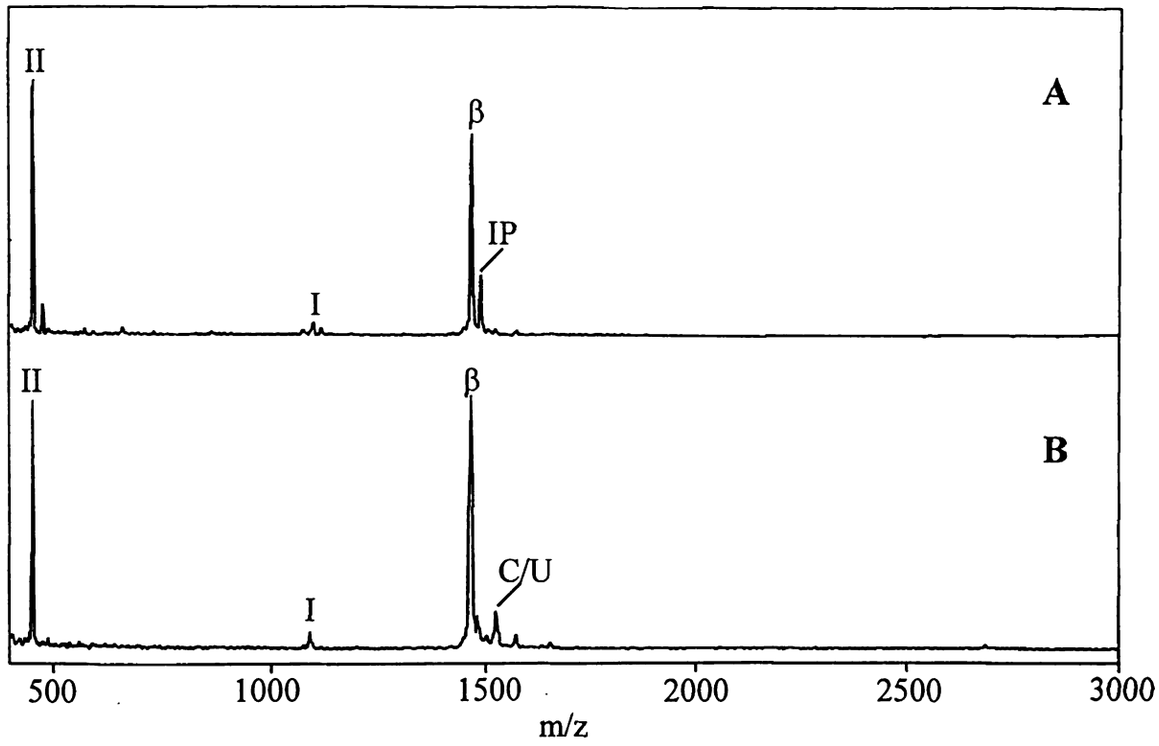


Figure 3.8. MALDI spectra of the cyanylated peptide EKPLQNFTLCFR, after cleavage in (A) pH 9.0, 37°C, 18h and (B) 1 M NH_4OH solution, rt, 1h, respectively. Symbol IP represents the intact peptide.

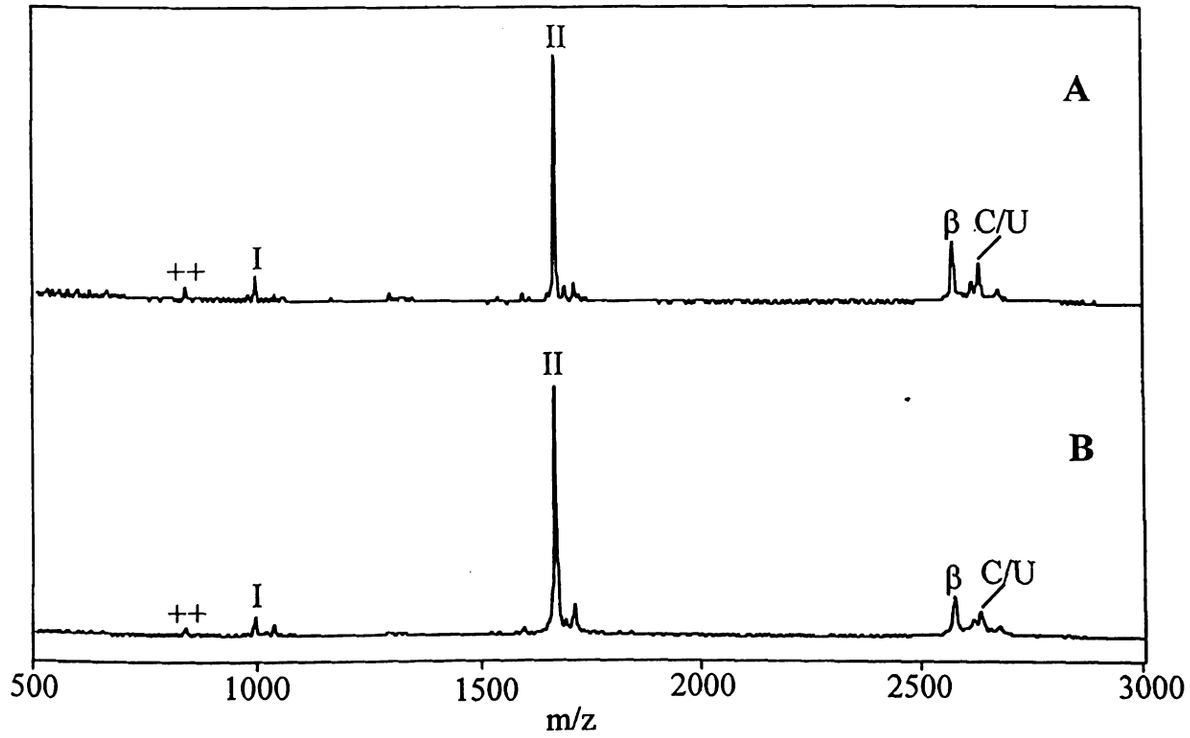


Figure 3.9. MALDI spectra of the cyanylated MHRQEAVDCLKKFNARRKLKGA, after cleavage in (A) pH 9.0, 37°C, 18h and (B) 1 M NH_4OH solution, rt, 1h, respectively.

C. **Effects of Amino A**

Degani and Patc

rates of β-elimination

dinitrophenylcysteine

Negatively charged carb

modified cysteine resid

take place, retard the

esterified amino acid a

rate, presumably becau

the functional groups

encountered if the cyan

C-terminal side. In mo

the cyanylated cystein

structural diversity on

acids. This conclusio

adjacent to the cyany

undergo β-elimination

residues may have the

The structural

cysteine significantly

denatured peptides ex

with rigid or bulky s

C. Effects of Amino Acid Structures on Cleavage and β-Elimination Reactions

Degani and Patchornik (9) showed that the effects of amino acid structures on the rates of β-elimination closely resemble those found in the β-elimination of S-2,4-dinitrophenylcysteine compounds whose β-elimination is also base-catalyzed. Negatively charged carboxylate groups located in the vicinity of the C-α hydrogen of the modified cysteine residue, where the attack of the catalyzing hydroxyl ion is supposed to take place, retard the rate of the elimination. On the other hand, the presence of an esterified amino acid adjacent to the cysteine residue greatly enhances the elimination rate, presumably because of the electron-withdrawing effect of the ester group. However, the functional groups mentioned in Degani and Patchornik's (9) examples are rarely encountered if the cyanylated cysteine residue in peptides does not locate in either N- or C-terminal side. In most cases, the group located in the vicinity of the C-α hydrogen of the cyanylated cysteine is an amide back bone of peptide chain, which imposes little structural diversity on the C-α hydrogen, regardless of structural features of the amino acids. This conclusion was supported by the observation that the same amino acid adjacent to the cyanylated cysteine presents considerable variation in the tendency to undergo β-elimination, whereas different amino acids adjacent to cyanylated cysteine residues may have the same relative yield (Table 3.2).

The structural features of the amino acid on the N-terminal side of cyanylated cysteine significantly affects both rate and yield of the cleavage reaction. For most of the denatured peptides examined, the cleavage products are predominant. But amino acids with rigid or bulky side chains, such as Pro and Tyr, are more resistant to cleavage,

giving β-elimination as

Som and Friedman's r

cysteine residues after c

the reaction, in which th

atom to provide the req

SCN group must be ori

nitrogen atom of the c

steric hindrance of the

residue, resulting in

neighboring amino ac

reaction, which was de

with the identical amin

other hand, is less depe

cysteine. However,

tendency and the rela

variation in pH has

products, whereas di

tendency to form cle

dominates the relative

It should be n

22 amino acid residue

role in the cleavage, e

giving β-elimination as a main product (Table 3.2). This observation is consistent with Som and Friedman's report (17) that the -Pro-Cys- was protected from cleavage at cysteine residues after cyanylation. This is understandable in terms of the mechanism of the reaction, in which the hydroxyl anion (or ammonia) must attack the scissile carbonyl atom to provide the required anionic tetrahedral intermediate, and the carbon atom of the SCN group must be oriented appropriately for nucleophilic attack by the activated amide nitrogen atom of the cyanylated cysteine. Both steps are significantly retarded by the steric hindrance of the side chain of the amino acid adjacent to the cyanylated cysteine residue, resulting in the decrease of the cleavage reaction rate. In additional to neighboring amino acids, the peptide's solution structure also affects the cleavage reaction, which was demonstrated by the slightly different yields obtained from peptides with the identical amino acid on the N-terminal side of cysteine. β-Elimination, on the other hand, is less dependent on the structure of the amino acid adjacent to the cyanylated cysteine. However, both cleavage and β-elimination reactions show the same pH tendency and the relative rate can be partially compensated. This explains why the variation in pH has little effect on the relative ratio of cleavage and β-elimination products, whereas different neighboring amino acids give a large variation in the tendency to form cleavage products. It is the structure of the peptide, not pH, that dominates the relative rate and yield of cleavage and β-elimination reactions.

It should be noted that the data presented here are deduced from peptides with 6-22 amino acid residues. For large proteins, solution structures may play a more important role in the cleavage, even if the protein is denatured. Many exceptions may be observed.

D. Application of Opt

The optimized
peptide chains at cya
methodology was exam
in proteins with diff
phosphokinase were ev

Figure 3.10 sho
NH_4OH solution and ir
"experimental Section"
47, and 77, respective
Experimentally, we co
(m/z 320.3 and 517.5
cleavage products und
of incubation at 37°C i
an overlapped peptide
not observed if the cl
1820.1 in pH 9.0 buff
1-17 (m/z 1839.1) du
dehydration did not oc

One of the gre
and subsequent cleav

D. Application of Optimized Methodology

The optimized conditions described above greatly facilitated the cleavage of peptide chains at cyanylated cysteine residues. The feasibility of the improved methodology was examined by experiments carried out to localize free sulfhydryl groups in proteins with different cysteine status. Ferredoxin, ovalbumin, and creatine phosphokinase were evaluated under different cleavage conditions to compare the results.

Figure 3.10 shows the MALDI spectra of spinach ferredoxin after cleavage in 1 M NH_4OH solution and in pH 9.0 buffer under the experimental conditions described in the "experimental Section". Ferredoxin contains five free sulfhydryls at positions 18, 39, 44, 47, and 77, respectively (18). Six fragments are anticipated after complete cleavage. Experimentally, we could detect by MALDI four out of six, two low-mass fragments (m/z 320.3 and 517.5, respectively) were missing. The mass spectra obtained from cleavage products under both conditions show a similar pattern. However, after 18 hours of incubation at 37°C in pH 9.0 buffer, a peak at m/z 3584.0 was observed, attributable to an overlapped peptide 44-76 with β-elimination at Cys47. This overlapped peptide was not observed if the cleavage was performed in 1 M NH_4OH solution. The peak at m/z 1820.1 in pH 9.0 buffer is probably attributable to the dehydration products of fragment 1-17 (m/z 1839.1) during prolonged exposure to the alkaline conditions (2), while this dehydration did not occur in 1 M NH_4OH solution.

One of the great advantages of the proposed methodology is that the cyanylation and subsequent cleavage are specific to free cysteine residues, selective cleavage of

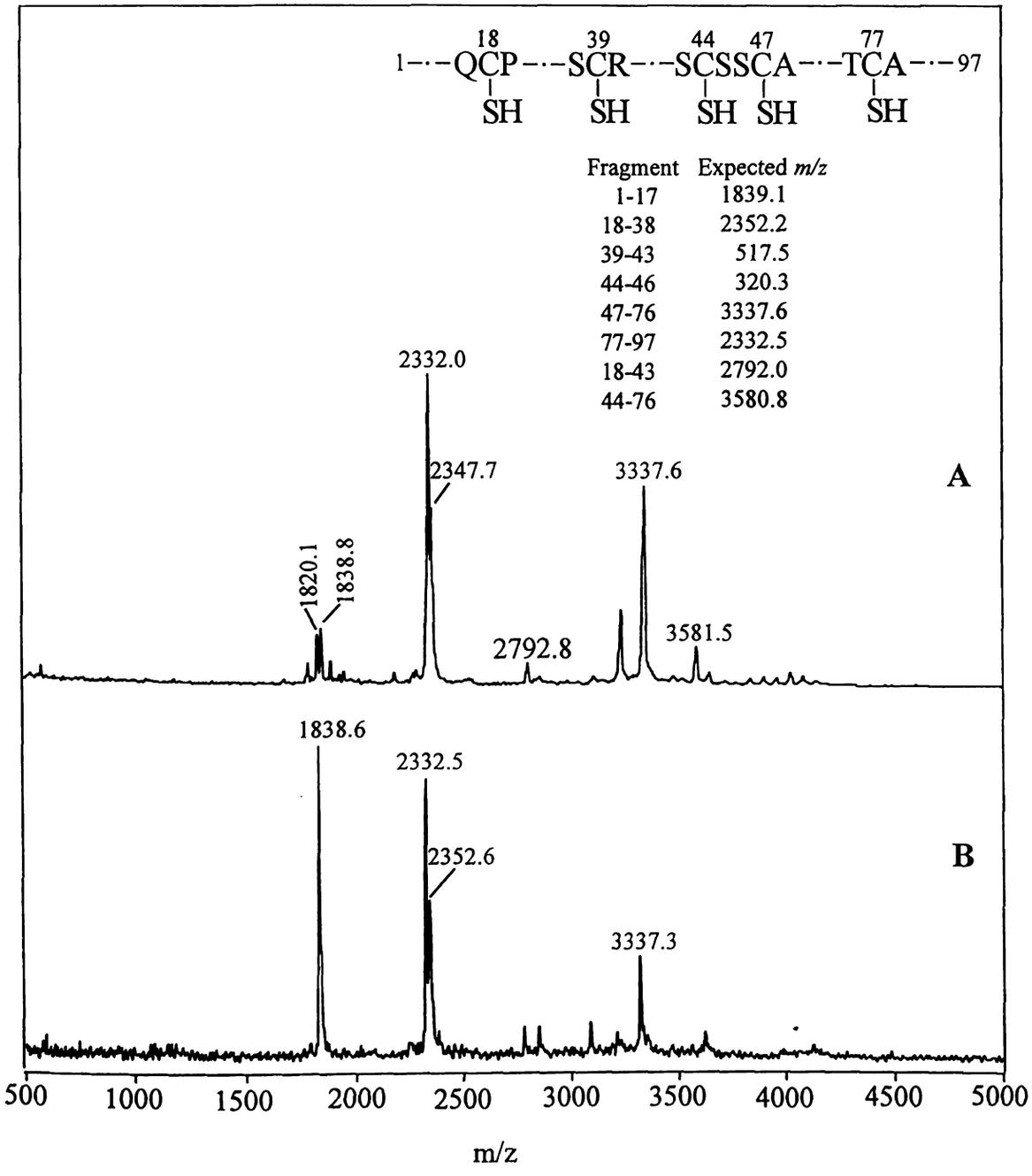


Figure 3.10. MALDI spectra of spinach ferredoxin (MW. 10483) after cyanylation by the CDAP and subsequent cleavage in (A) pH 9.0 buffer for 18 hours at 37°C and (B) 1 M NH_4OH solution for one hour at room temperature, respectively.

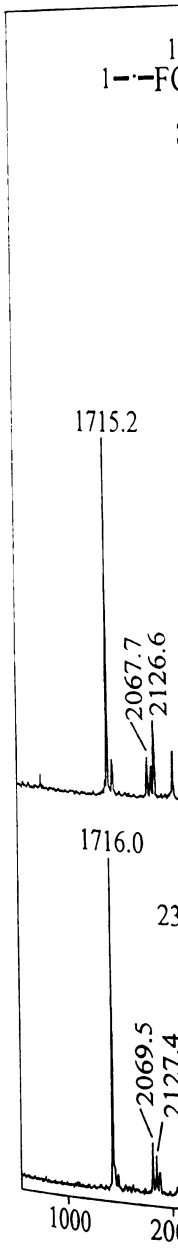


Figure 3.11. MA
CDAP and subse
NH_4OH solution
cyanylated/uncle

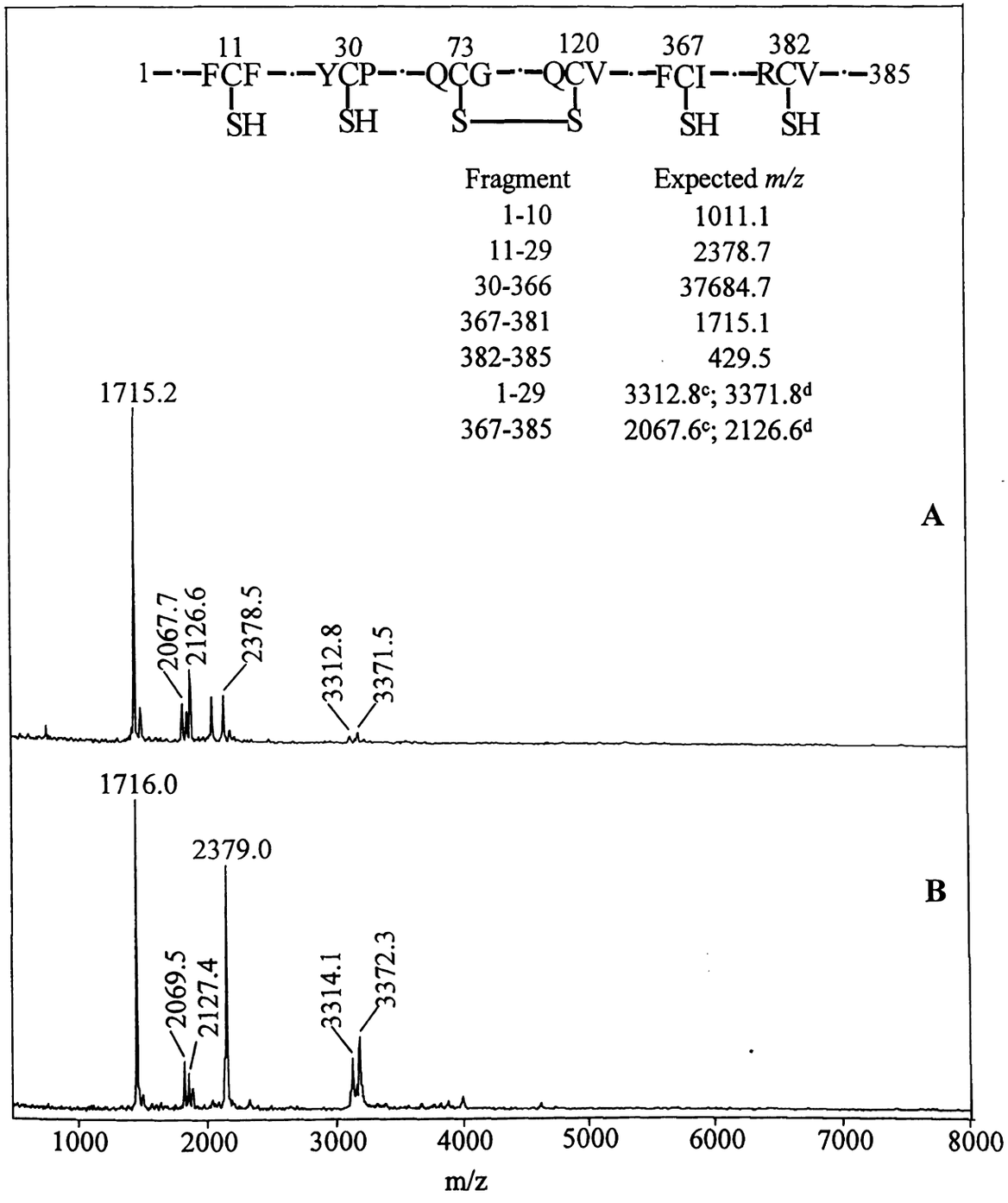


Figure 3.11. MALDI mass spectra of ovalbumin (MW. 42699) after cyanylation by CDAP and subsequent cleavage in (A) pH 9.0 buffer at 37°C for 18 hours and (B) 1M NH_4OH solution at room temperature for one hour, respectively. *c* and *d* represent cyanylated/uncleaved peptides and its β-elimination products, respectively.

protein chains at N-pep

cystinyl residues. This

both types of residue

demonstrated by the

ovalbumin, a protein c

cyanylation by CDAP,

HCl buffer for 18 ho

temperature, respectiv

complete cleavage at

detected by MALDI

informative because

cyanylation and cleav

385, corresponding to

Cys10 and Cys382, re

cleavage sites. In co

cleavage occurs at th

almost identical. An

occurs at -Tyr-Cys-

fragments 11-29 and

Our methodc

sulfhydryl groups in

amino acids. The pr

protein chains at N-peptide bonds of cysteinyl residues can be achieved in the presence of cystinyl residues. This option is particularly useful in cases of native proteins containing both types of residues, or in partially reduced proteins (15). This feasibility was demonstrated by the selective cyanylation and cleavage of the free SH groups in ovalbumin, a protein containing four free sulfhydryls and one disulfide bond (19). After cyanylation by CDAP, the modified protein was subjected to cleavage both in pH 9.0 tris-HCl buffer for 18 hours at 37°C and in 1 M NH_4OH solution for one hour at room temperature, respectively. Among five fragments expected corresponding to the complete cleavage at the four free SH residues, only two, 367-381 and 11-29, were detected by MALDI (Figure 3.11). These two fragments were, however, very informative because they indicated that Cys10, 30, 367, and 382 had undergone the cyanylation and cleavage. In addition, two sets of overlapped fragments, 1-29 and 367-385, corresponding to the cyanylated/uncleaved species and the β-elimination products at Cys10 and Cys382, respectively, can be detected, which implies that Cys30 and 367 are cleavage sites. In contrary, Cys 73 and Cys 120 are linked by a disulfide bond, and no cleavage occurs at these residues. The results obtained by the cleavage at pH 9.0 are almost identical. An exception to our previous discussion, however, is that the cleavage occurs at -Tyr-Cys- (Cys30) structure in this example, resulting in the formation of fragments 11-29 and 1-29, whereas no such cleavage was found for the model peptide.

Our methodology was further applied to recognizing the location of free sulfhydryl groups in rabbit muscle creatine phosphokinase, a protein containing 380 amino acids. The primary structure of this protein was established from cDNA clones

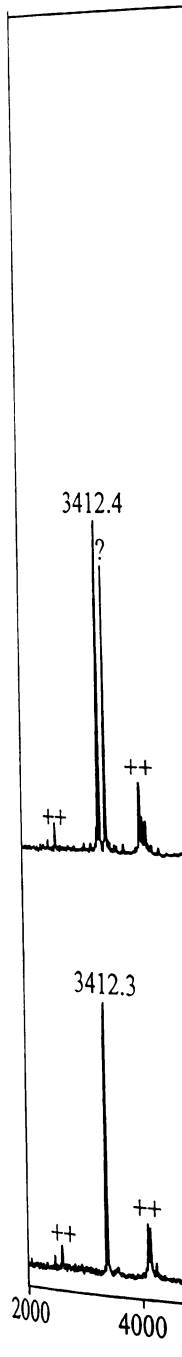


Figure 3.12. MA
42977) after cyan
for 18 hours and
respectively. The
in the sample. T

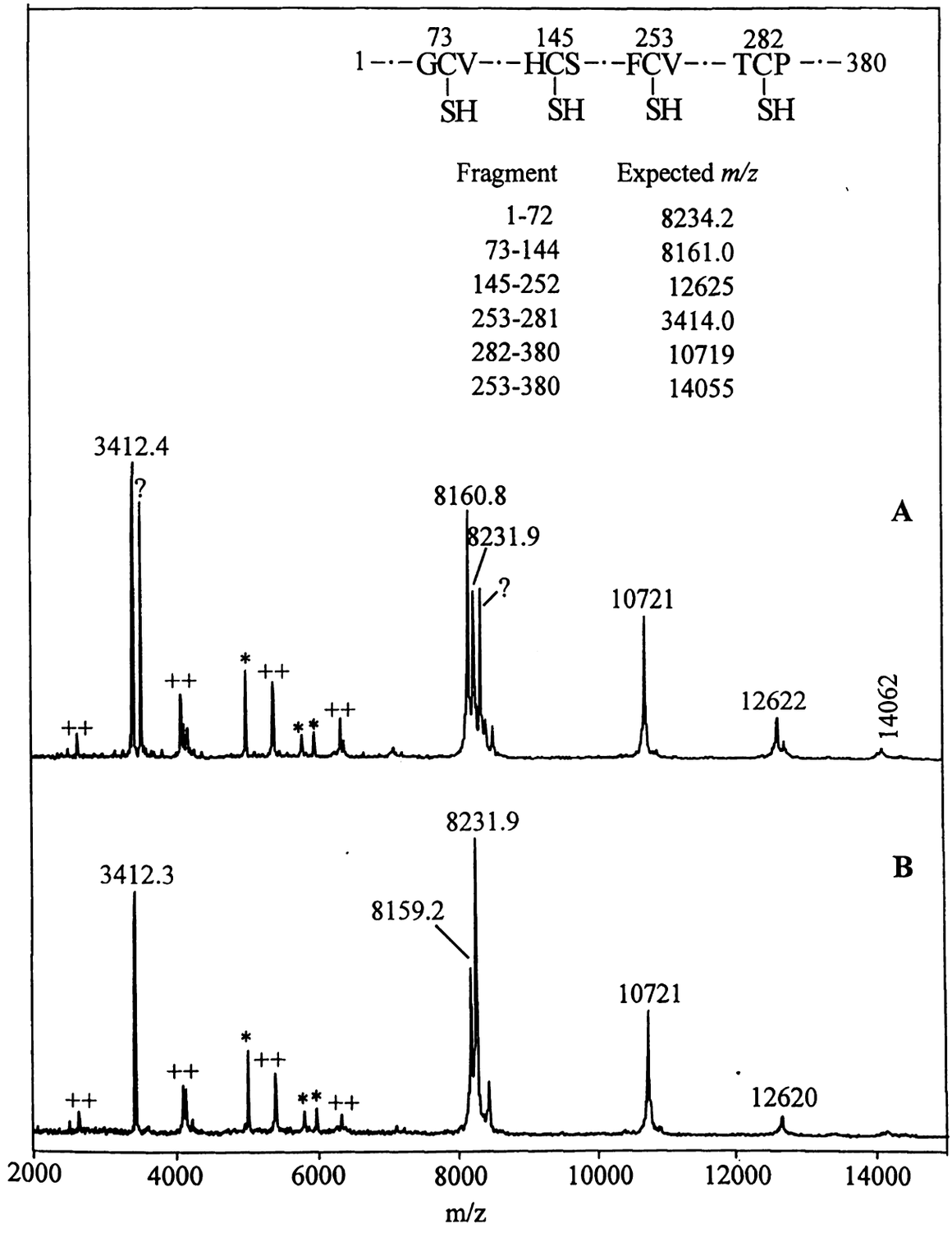


Figure 3.12. MALDI mass spectra of rabbit muscle creatine phosphokinase (MW. 42977) after cyanylation by CDAP and subsequent cleavage in (A) pH 9.0 at 37°C for 18 hours and (B) 1 M NH_4OH solution at room temperature for 1 hour, respectively. The peaks marked with "*" are due to "carry-over" from an impurity in the sample. The question marks "?" in (A) are related to two unidentified species.

and confirmed by othe

cysteines located at po

is a reactive site. But i

the cysteines or if all

available creatine pho

modified protein was

other coexisting protei

9.0 buffer and 1 M NH

products under both co

each of the four cystei

to "carry-over" from

creatine phosphokinas

cleavage products fro

fragment, m/z 14062

Cys253, but β-elimin

are in reduced cystei

creatine phosphokinas

It should be r

shows two unidentifi

peaks are likely due t

long exposure (18 ho

was reported by Naka

and confirmed by other physical methods (20). Creatine phosphokinase contains four cysteines located at positions 73, 145, 253, and 282, respectively, among which Cys282 is a reactive site. But it is unclear yet if there is a disulfide bond linkage between two of the cysteines or if all the four cysteines are in the reduced form. The commercially available creatine phosphokinase was subjected to cyanylation by the CDAP and the modified protein was purified by reversed-phase HPLC to remove excess reagents and other coexisting proteins. The purified creatine phosphokinase was then cleaved in pH 9.0 buffer and 1 M NH_4OH solution, respectively. MALDI mass spectra of the cleavage products under both conditions show all five fragments, corresponding to the cleavage at each of the four cysteine residues (Figure 3.12). The peaks marked with asterisk are due to "carry-over" from cleavage products of a protein impurity poorly resolved from creatine phosphokinase, as demonstrated by the analysis of MALDI spectrum of the cleavage products from the corresponding protein. In addition, an overlapped peptide fragment, m/z 14062, was observed in pH 9.0 buffer, corresponding to cleavage at Cys253, but β-elimination at Cys282. It is concluded that all the four cysteine residues are in reduced cysteine form and there is no disulfide-linked cystine in rabbit muscle creatine phosphokinase.

It should be noted that analysis of the cleavage products from the pH 9 buffer shows two unidentified peaks (with question mark) on the MALDI spectrum. These peaks are likely due to protein chain decomposition or other side reactions because of the long exposure (18 hours) to alkaline conditions during incubation. A similar observation was reported by Nakagawa et al (21) who noticed that prolonged exposure of cyanylated

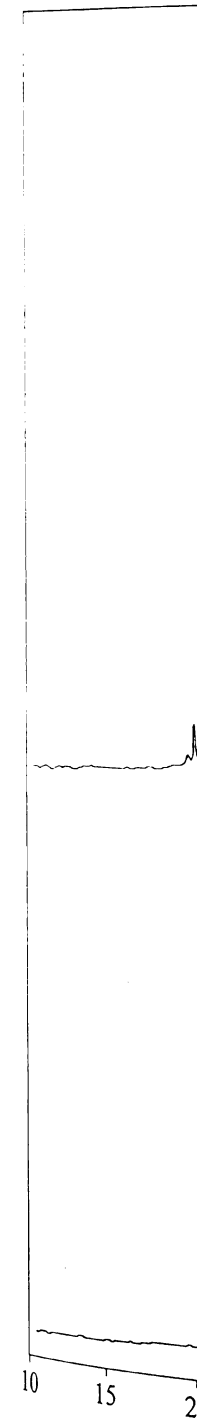


Figure 3.13. HPLC
in (A) pH 9.0, 37°C
to the specific cleav

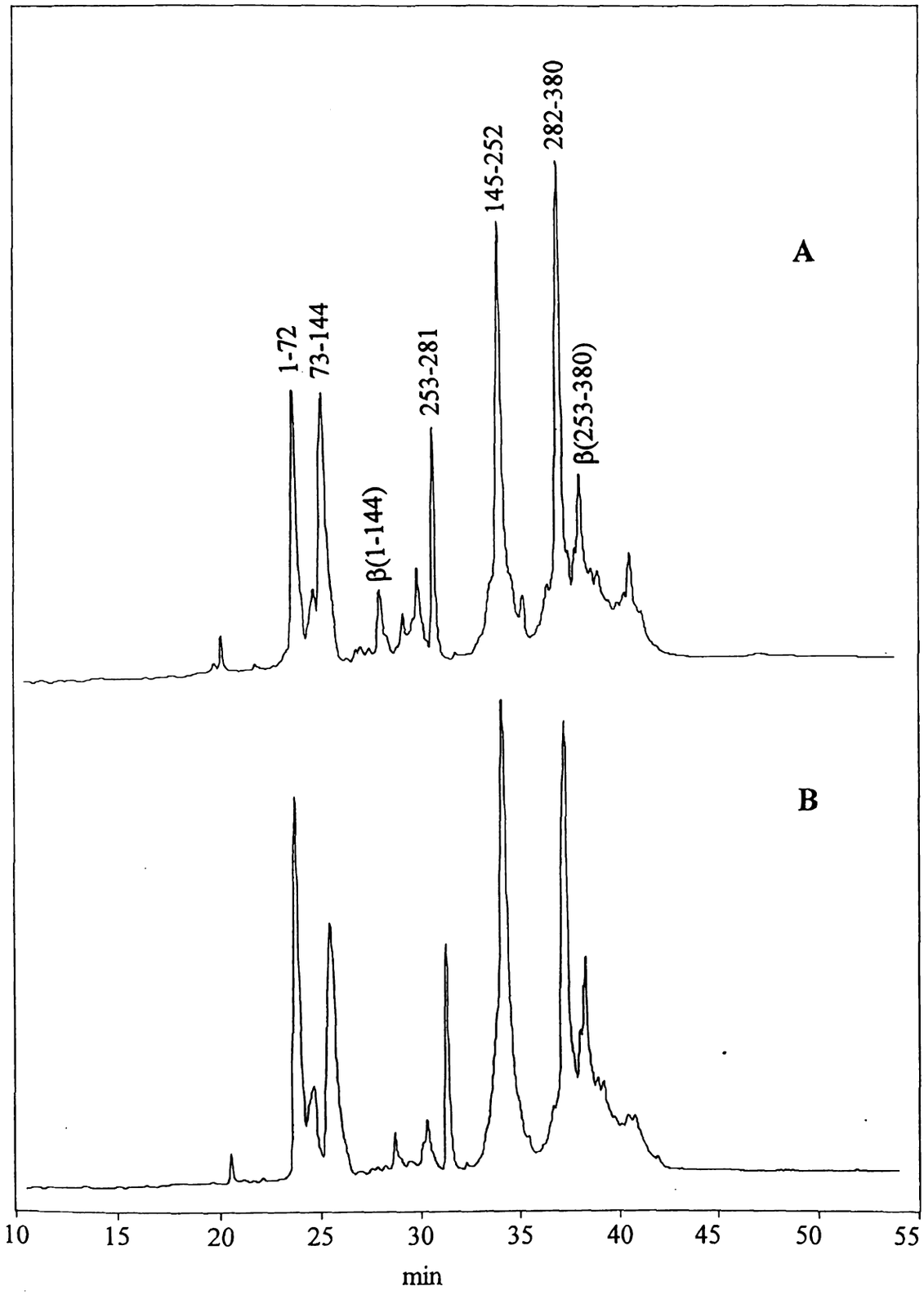


Figure 3.13. HPLC chromatograms of cleavage products of creatine phosphokinase in (A) pH 9.0, 37°C, 18h, and (B) 1 M NH_4OH, rt, 1h. The peaks corresponding to the specific cleavage at cysteine residues are marked.

peptide under mildly a

succinimide formation

incubation in a modera

The MALDI s

analytes. For instance

digests, whereas a very

results obtained from

followed by mass-ma

obtained under both c

fragments correspond

Furthermore, the β-el

result confirms the co

form; there is no disul

IV. Conclusions

In this chapte

structural factors on

elimination. The C

sulfhydryl groups be

oxidation of SH grou

might occur during c

two reactions are co

peptide under mildly alkaline conditions resulted in the oxidative modification of Met, succinimide formation of Asp, and other side reactions. It seems that the short time of incubation in a moderate concentration of ammonia minimizes such side reactions.

The MALDI signal is not readily related to the relative concentration of the analytes. For instance, the less intense peak might represent a main component in the digests, whereas a very intense peak might represent a minor component. To confirm the results obtained from MALDI, reversed-phase HPLC fractionation of the fragments followed by mass-mapping of the fragments by MALDI was carried out for digests obtained under both conditions. All the main HPLC peaks can be assigned to individual fragments corresponding to specific cleavage at the cysteine residues (Figure 3.13). Furthermore, the β-elimination product, 353-380, presents as a minor product. This result confirms the conclusion that the cysteine in creatine phosphokinase is in reduced form; there is no disulfide bond between cysteine residues.

IV. Conclusions

In this chapter, we have studied the effects of pH conditions and amino acid structural factors on the kinetics and relative yields of the cleavage reaction and β-elimination. The CDAP was found superior to the NTCB for the cyanylation of sulfhydryl groups because it reacts with SH groups under acidic conditions, where the oxidation of SH groups is slow, the protein is kept relatively stable, and the cleavage that might occur during cyanylation is minimal. In contrast to previous reports, although the two reactions are competitive, higher pH does not elevate the relative yields of β-

elimination significant

because it can greatly

nucleophile than hydro

The cleavage in 1 M

prolonged incubation

terminal side of cyar

reaction. But for most

The optimized

number of proteins w

fast, simple, sensitive

presence of disulfide

V. Future Work

The work des

cysteine residues afte

provide insight into

modification of prot

cleavage may not b

MALDI, either beca

because of the suppr

MALDI. In the cha

sample support for d

of peptides and prote

elimination significantly. 1 M ammonia solution is superior to other buffer solutions because it can greatly accelerate the cleavage reaction because it provides a stronger nucleophile than hydroxyl anions, and it can easily be removed from the reaction mixture. The cleavage in 1 M NH_4OH solution also minimizes the side reactions related to the prolonged incubation of proteins. The structures of amino acids adjacent to the N-terminal side of cyanylated cysteine residues are critical to the rate of the cleavage reaction. But for most of the peptides studied, the cleavage products are dominant.

The optimized procedure has been used to locate free sulfhydryl groups in a number of proteins with both known and unknown cysteine structures. This approach is fast, simple, sensitive, and can be applied to the localization of cysteine residues in the presence of disulfide bonds.

V. Future Work

The work described here and in chapter 2 demonstrates that chemical cleavage at cysteine residues after cyanylation and subsequent mass-mapping by MALDI-TOF MS provide insight into the location of free sulfhydryl groups and posttranslational modification of protein structures. However, some of the segments from the specific cleavage may not be detectable by the direct analysis of the digestion mixture by MALDI, either because of the interference of salts, buffers, and denature agents, or because of the suppression of the responses from coexisting peptides during analysis by MALDI. In the chapter 2, we proposed to use an immobilized inert membrane as a sample support for direct analysis of proteins. We demonstrated that the immobilization of peptides and proteins on an inert Zetabind membrane can be used to facilitate removal

of water-soluble MAL

and salts, while the ana

adding the matrix solu

possible on the memb

MALDI analysis is su

As a matter of fact, thi

possible reason is tha

between sample holde

The electric field the i

positions.

To overcome t

be tested. Instead of a

a membrane material

MALDI probe tips as

mixtures. This thin

expected to improve t

So far, two r

examined in this lab

cyanylate SH groups

quantitative cyanylat

cyanylation is slow d

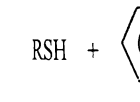
reagent, thiocyanopy

NTCB. The TCP, li

of water-soluble MALDI contaminants, such as high concentrations of denaturing agents and salts, while the analyte is retained when the membrane is immersed in water prior to adding the matrix solution. Moreover, *in situ* chemical cleavage of adsorbed proteins is possible on the membrane. However, the resolution, accuracy, and reproducibility of MALDI analysis is substantially poorer for the analytes immobilized on the membrane. As a matter of fact, this becomes a common drawback of using various membranes. One possible reason is that the membrane is too thick, it behaves as an insulating layer between sample holder (high voltage) and the protein/matrix sample on the membrane. The electric field the ions experience is not identical for those desorbed from the different positions.

To overcome the limitation of the current membranes, a new sample support can be tested. Instead of a thick membrane layer, a polymer film can be formed by dissolving a membrane material, such as nitrocellulose (22), in an organic solvent, coat it onto MALDI probe tips as a thin film, and allow it to air dry prior to deposition of peptide mixtures. This thin film would act as an inert membrane mentioned above, but is expected to improve the MALDI signal.

So far, two reagents for specific cyanylation of sulfhydryl groups have been examined in this laboratory. The CDAP is superior to the NTCB in its capacity to cyanylate SH groups under acidic conditions. However, CDAP is unstable at pH>4.5, quantitative cyanylation of SH groups thus may not be achieved in cases where cyanylation is slow due to structural factors of proteins. Another unexplored cyanylation reagent, thiocyanopyridine (TCP) (23), may provide advantages over the CDAP and NTCB. The TCP, like NTCB, can cyanylate sulfhydryl groups though the exchange of



(2,2'-dipy

RS

Figure 3.14. C
thiocyanopyrid

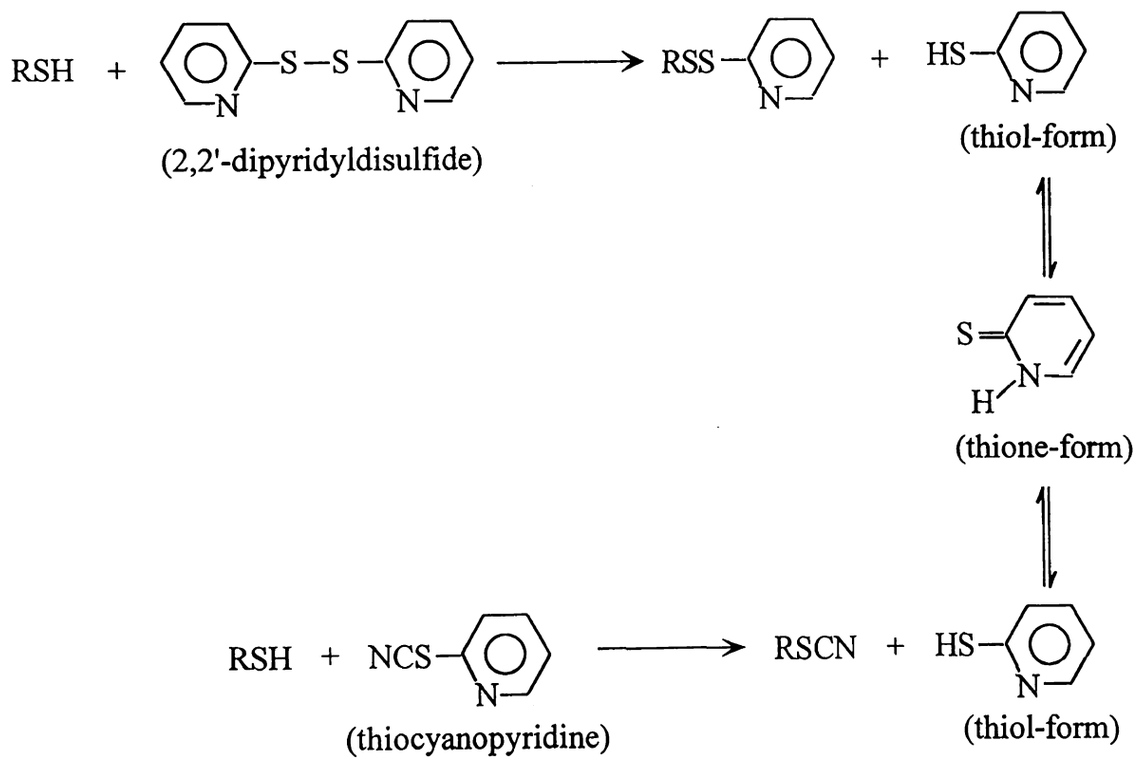


Figure 3.14. Chemical modification of SH groups by 2,2'-dipyridyldisulfide and thiocyanopyridine, respectively.

the cyano group. Bu
reactive to SH group
expectation is deduc
dipyridyldisulfide (2-
with thiol groups in a
resonance structure of
compensated for by
protonation of its ring
shares the same struct
(Figure 3.14). The
potentially powerful
this reagent is retarde

In all, the use
reaction conditions
simpler and more pla

the cyano group. But we expect the TCP, due to its structural features, will be more reactive to SH groups and can be used in lower pH than the NTCB reaction. This expectation is deduced from the comparison of structural similarity between 2,2'-dipyridyldisulfide (2-PDS) and TCP (Figure 3.16). The former can quantitatively react with thiol groups in a wide pH range (pH 1-9), as the releasing group is stabilized by resonance structure of the thione. The low nucleophilicity of thiol groups at acidic pH is compensated for by an increased electrophilicity of the 2-PDS as a result of the protonation of its ring nitrogen (pKa ~3). The releasing group of TCP after cyanylation shares the same structure as the product of 2-DPS, both having the stable thione structure (Figure 3.14). The structural similarity between 2-PDS and TCP makes TCP a potentially powerful reagent for cyanylation. Unfortunately, the further exploration of this reagent is retarded at this moment, because the TCP is not commercially available.

In all, the use of a new cyanylation reagent, together with the optimized cleavage reaction conditions and new membrane techniques, is promising to provide an even simpler and more plausible application of our cyanylation methodology.

VI. References

1. Jacobson, G. R., S

   248, 6583-6591(19

2. Wu, J., Gage, D. A

3. Nefsky, B., and Br

4. Thomas, M. L., J

   USA, **79**, 1054-10

5. Roberts, D. D., an

6. Matsudaira, P., Ja

   82, 6788-6792(19

7. Mahboub, S., Ric

   182(1986).

8. Papayannopoulos

9. Degani, Y., and P

10. Nakagawa, S., Ta

    Nishimura, O., F

11. Denslow, N. D.

    Academic Press,

12. Price, N. C., *Bioc*

13. Liao, T-H., and W

14. Wakselman, M.,

15. Wu, J., and Wats

VI. References

1. Jacobson, G. R., Schaffer, M. H., Stark, G. R., and Vanaman, T. C., *J. Biol. Chem.* **248**, 6583-6591(1973).
2. Wu, J., Gage, D. A., and Watson, J. T., *Anal. Biochem.*, **235**, 161-174(1996).
3. Nefsky, B., and Bretscher, A., *Proc. Natl. Acad. Sci. USA*, **86**, 3549-3553(1989).
4. Thomas, M. L., Janatova, J., Gray, W. R., and Tack, B. F., *Proc. Natl. Acad. Sci. USA*, **79**, 1054-1058(1982).
5. Roberts, D. D., and Goldstein, I. J., *J. Biol. Chem.*, **259**, 909-914(1984).
6. Matsudaira, P., Jakes, R., Cameron, L., and Atherton, E., *Proc. Natl. Acad. Sci. USA.* **82**, 6788-6792(1985).
7. Mahboub, S., Richard, C., Delacourte, A., and Han, K., *Anal. Biochem.*, **154**, 171-182(1986).
8. Papayannopoulos, I.A., and Biemann, K., *Protein Sci.*, **1**, 278-288(1992).
9. Degani, Y., and Patchornik, A., *Biochemistry,* **13**, 1-11(1974).
10. Nakagawa, S., Tamakashi, Y., Hamana, T., Kawase, M., Taketomi, S., Ishibashi, Y., Nishimura, O., Fukuda, T., *J. Am. Chem. Soc.,* **116**, 5513-5514(1994).
11. Denslow, N. D., and Nguyen, H. P., in *Techniques in Protein Chemistry VII*, Academic Press, Inc., (1996), p.241-248.
12. Price, N. C., *Biochem. J.*, **159**, 177-180(1976).
13. Liao, T-H., and Wadano, A., *J. Biol. Chem.*, **254**(19), 9602-9607(1979).
14. Wakselman, M., and Guibe-Jampel, E., *J. C. S. Chem. Comm.*, 21-22(1976).
15. Wu, J., and Watson, J. T., *Protein Sci.,* **6**, 391-398(1997)

16. Stark, G, R., in *Me*

17. Som, S., and Fried

18. Matsubara, H., and

19. Nisbet, A.D., Saun

J. Biochem., **115**,

20. Putney, S., Herli

Belagaje, R., Bier

259(23), 14317-14

21. Nakagawa, S., Ta

O., Fukuda, T., *Bi*

22. Preston, L. M., N

550(1993).

23. Brocklehurst, K.,

A., Mushiri, M. S

16. Stark, G, R., in *Methods Enzymol.*, **47**, 129-132(1977).

17. Som, S., and Friedman, S., *J. Biol. Chem.*, **266**(5), 2937-2945(1991).

18. Matsubara, H., and Sasaki, R.M., *J. Biol. Chem.*, **243**, 1732-1757(1968).

19. Nisbet, A.D., Saundry, R.H., Moir, A.J.G., Fothergill, L.A., and Fothergill, J.E., *Eur. J. Biochem.*, **115**, 335-345(1981).

20. Putney, S., Herlihy, W., Royal, N., Pang, H., Aposhian, H. V., Pickering, L., Belagaje, R., Biemann, K., Page, D., Kuby, S., and Schimmel, P., *J. Biol. Chem.*, **259**(23), 14317-14320(1984)

21. Nakagawa, S., Tamakashi, Y., Ishibashi, Y., Kawase, M., Taketomi, S., Nishimura, O., Fukuda, T., *Biochem. Biophys. Res. Commun.*, **200,** 1735-1741(1994).

22. Preston, L. M., Murray, K. K., and Russell, D. H., *Biol. Mass Spectrom.*, **22**, 544-550(1993).

23. Brocklehurst, K., Malthouse, J. P. G., Baines, B. S., Blenkinsop, R. D., Churcher, J. A., Mushiri, M. S., and Ormerod, F., *Biochem. Soc. Trans.*, Vol. 6, 261-263(1978).

A NOVEL METH

I. Introduction

Proteins are syn
added, one by one, fro
contain SH and/or S-S
However, of the two,
chain. Cysteine is carr
the codons UGC and
dehydrogenation of S
proteins. Inside the
glutathione. Next to
observation suggests t
disulfide bonds in a
generally lost complet

Many biologi
disulfide bonds. Beca
it is necessary to know
any unique contributi
With the advent of r

CHAPTER 4

A NOVEL METHODOLOGY FOR ASSIGNMENT OF DISULFIDE BOND PAIRINGS IN PROTEINS

I. Introduction

Proteins are synthesized in cells by a stepwise process in which amino acids are added, one by one, from the N-termini of the chains. Most of the proteins are found to contain SH and/or S-S groups, which are associated with cysteine and cystine residues. However, of the two, only cysteine is incorporated directly into the initial polypeptide chain. Cysteine is carried to the ribosome by its corresponding tRNA, which recognizes the codons UGC and UGU. Protein S-S groups arise by subsequent pairing and dehydrogenation of SH groups. Disulfide bonds occur frequently in extracellular proteins. Inside the cell the sulfhydryl group is maintained in a reduced state by glutathione. Next to tryptophan, cyst(e)ine is the most conserved amino acid. This observation suggests that cyst(e)ine is very important for structure and function. If all of disulfide bonds in a protein are reduced, both the tertiary structure and function are generally lost completely.

Many biologically and therapeutically important proteins (peptides) contain disulfide bonds. Because the disulfide bond is an important element of protein structure, it is necessary to know the locations of disulfide bonds in order to understand more fully any unique contributions they may make to protein (peptide) structure and function. With the advent of recombinant DNA techniques for mutagenesis and expression of

cloned genes, it has be
disulfide bonds into p
that the products of
proteins, or characteri
an important aspect in
methods for quantify
determination of the
chemists.

Current meth
cleavage of protein c
such as cyanogen br
bond. The resulting
sequences or molecu
mass spectrometry
protein leads to the r

Although thi
it is limited in many
their amino acid con
requirement may no
protein is very sma
cleaved between ev
more than one disu
residue by an enzym

cloned genes, it has become possible to make altered structures, such as insertion of new disulfide bonds into proteins to increase stability. In these cases, one must either verify that the products of genetic manipulations are identical to those found in the native proteins, or characterize their differences. The assignment of disulfide bonds is therefore an important aspect in the structural characterization of proteins. Although there are good methods for quantifying the number of disulfide bonds in proteins, the unambiguous determination of the location or pairing of disulfide bonds continues to challenge protein chemists.

Current methodology (1, 2) for assignment of disulfide bonds in proteins involves cleavage of protein chains between half-cystinyl residues with specific cleavage reagents, such as cyanogen bromide and trypsin, to obtain peptides that contain only one disulfide bond. The resulting mixture of peptides is separated, and the amino acid compositions, sequences or molecular masses of the peptides are determined by Edman degradation or mass spectrometry or both. Assignment of these peptides to specific segments of the protein leads to the recognition of disulfide crosslinkages.

Although this approach is well established and has been used with much success, it is limited in many respects. First, disulfide-containing peptides can be identified by their amino acid composition or sequence only if they are purified to homogeneity. This requirement may not be achieved in the case of large proteins, or when the quantity of protein is very small. Second, the above methodology requires the protein chain be cleaved between every half-cystinyl residues so that the resulting peptides contain no more than one disulfide pair. If the protein is not cleaved between every half-cystinyl residue by an enzyme or chemical reagent, the peptide obtained by the cleavage must be

cleaved again using an
enzymatic or chemical
proteins are often res
difficult or impossible
spaced or adjacent cy
are frequently used to
disulfide scrambling (
approach may not be
The problem of disu
proteins, irrespective
methodologies such
been proposed to av
obtained by these nor

Gray (5, 6) ha
highly bridged sma
experiments, peptide
the partially reduced
positions of alkylate
related to the disulfi
obviously tedious, i
protein using this ap

cleaved again using another enzyme or chemical reagent. However, even multiple-step enzymatic or chemical digestions rarely yield a full set of diagnostic fragments because proteins are often resistant to some specific cleavage reagents. It will be even more difficult or impossible to find an enzyme which is likely to cleave between two closely spaced or adjacent cysteines (1). It is likewise important to note that the conditions that are frequently used to cleave a protein chain are also the conditions that frequently lead to disulfide scrambling (3). As a result, the location of disulfide bonds determined by this approach may not be the same as the location of disulfide bonds in the native protein. The problem of disulfide exchange plagues all attempts to locate disulfide bonds in proteins, irrespective of the general method currently used. Although supplementary methodologies such as non-specific fragmentation by partial acid hydrolysis (4) have been proposed to avoid disulfide bond scrambling, it is difficult to deal with the data obtained by these non-specific techniques.

Gray (5, 6) has described an approach for analyzing disulfide linkage patterns in highly bridged small peptides with close or adjacent cysteine residues. In his experiments, peptides were partially reduced under controlled conditions, the isomers of the partially reduced protein separated by HPLC, the nascent free thiols alkylated, and the positions of alkylated cysteines recognized from the results of sequence analysis and related to the disulfide bond pair that had been reduced and cyanylated. However, it is obviously tedious, if not impractical, to sequence an alkylated high-mass peptide or protein using this approach.

II. Our Novel Strate

In chapter 2, w

free cysteine groups

chemical reaction bet

selectively cyanylate

cysteinyl residue can

peptide and a series

mapped to the sequen

react with NTCB an

groups. In chapter

conditions, and studi

elimination. The o

makes this methodo

interpretation of the

In this chap

assignment of disul

cleavage of partially

resulting peptides by

for a simple protein

reduced by tris(2-c

produce a mixture

Conditions can be o

II. Our Novel Strategy for Disulfide Assignment

In chapter 2, we developed a simple methodology for recognizing the location of free cysteine groups in peptides and proteins (7). This approach employs a specific chemical reaction between sulfhydryls and 2-nitro-5-thiocyanobenzoic acid (NTCB) to selectively cyanylate cysteine thiols (8). The N-terminal peptide bond of the modified cysteinyl residue can then be cleaved under alkaline conditions to form an amino-terminal peptide and a series of 2-iminothiazolidine-4-carboxylyl peptides which can be mass mapped to the sequence of the original molecule by MALDI-MS. Disulfide bonds do not react with NTCB and therefore do not interfere with the determination of sulfhydryl groups. In chapter 3, we further improved and optimized the cleavage reaction conditions, and studied the effects of pH and amino acid structure on cleavage and β-elimination. The optimized cleavage conditions greatly facilitated the cleavage and makes this methodology much more attractive in terms of speed of the analysis and interpretation of the data.

In this chapter, we report a novel strategy (outlined in Figure 4.1) for the assignment of disulfide bond pairings in proteins using the above specific chemical cleavage of partially reduced and cyanylated protein isomers with mass mapping of the resulting peptides by MALDI-TOF MS (9). In this methodology, as shown in Figure 4.2 for a simple protein containing only two disulfide bonds, the denatured protein is partially reduced by tris(2-carboxyethyl)phosphine (TCEP) in a buffer solution at pH 3.0 to produce a mixture of residual intact protein and isomers of partially reduced species. Conditions can be optimized so that the predominant products are isoforms in which only

Figure 4.1. D
of disulfide b

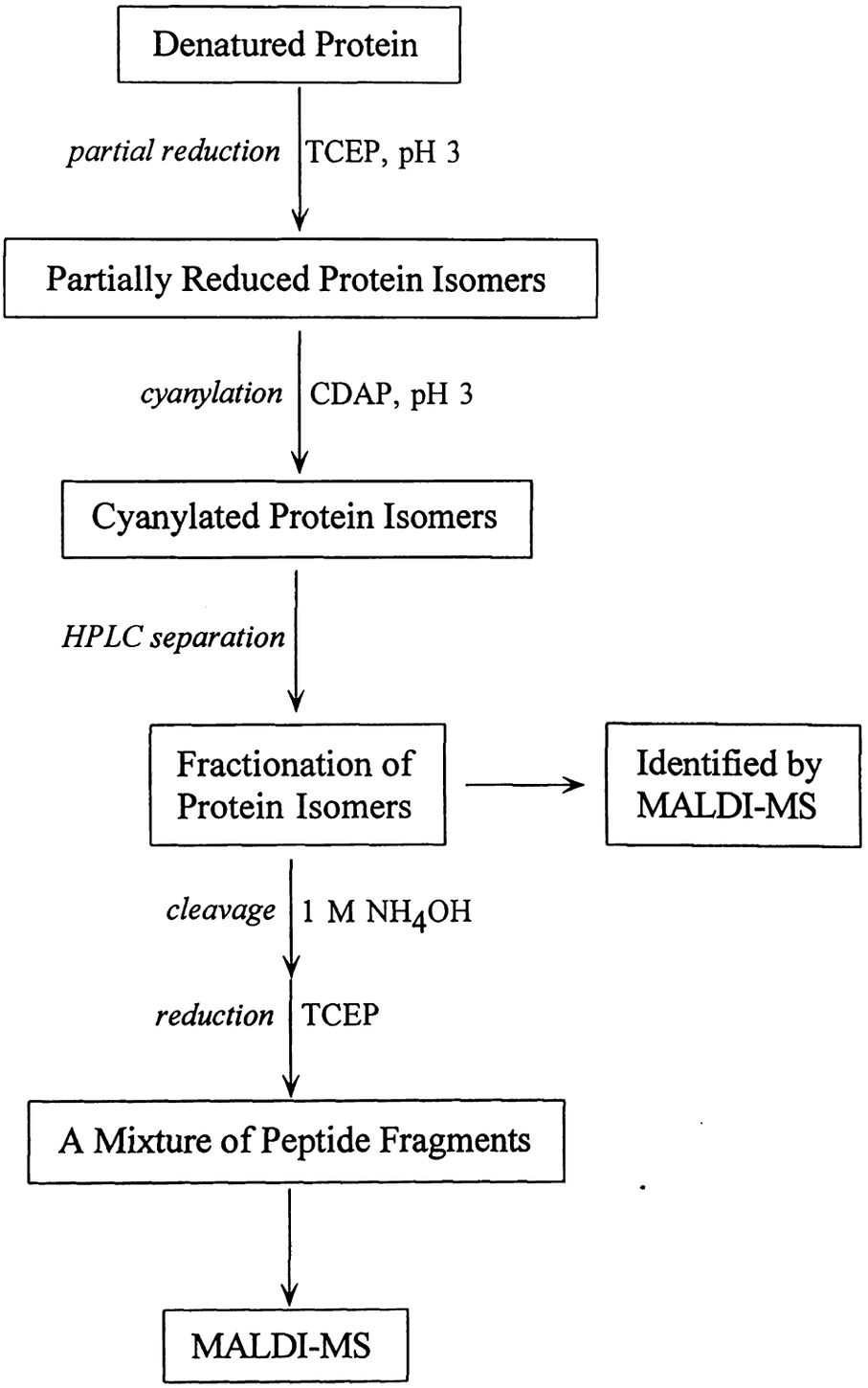


Figure 4.1. Descriptive overview of our proposed methodology for assignment of disulfide bond pairings in proteins.

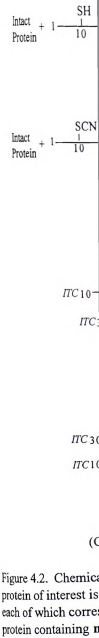


Figure 4.2. Chemica
protein of interest is
each of which corres
protein containing n
ITC stands for imin

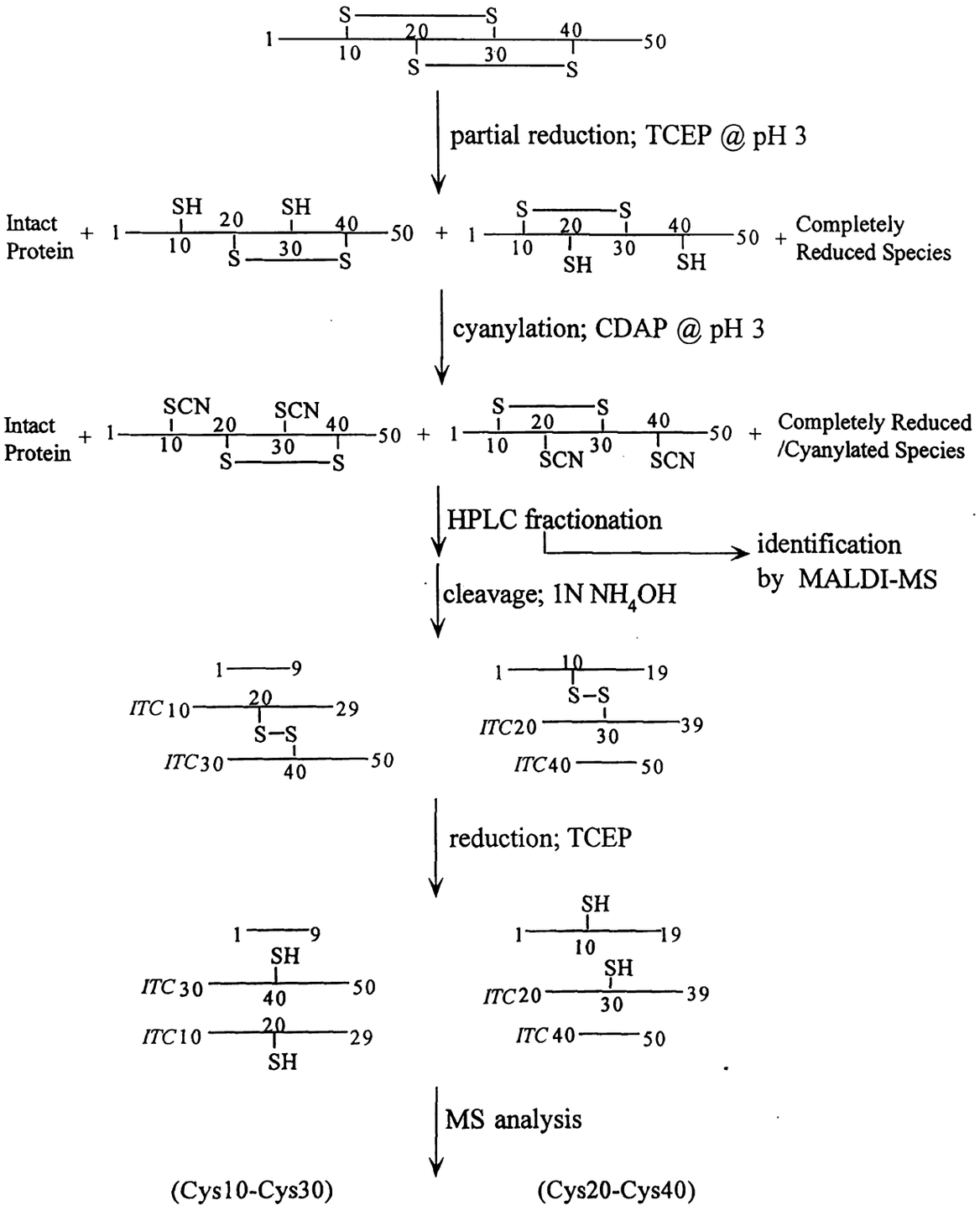


Figure 4.2. Chemical overview of our methodology. Partial reduction means that the protein of interest is reduced under controlled conditions into a mixture of isoforms, each of which corresponds to reduction of only one of its disulfide bonds; for a protein containing **n** cystines, **n** isomers of the singly-reduced protein will result. *ITC* stands for iminothiazolidinyl carboxyl residue at the amino terminus.

a single disulfide bo
cyanylated by 1-cyano-
the same buffer condi
then separated by rev
MALDI-MS to determ
+52 Da correspond
corresponds to doubly
shift from the mass
cleavage in aqueous
disulfide bonds, are
mass-mapped by MA
to the location of th
cleavage. A primary
obtained from our ap
protein isomers are p
needed to define a
confirm an assignm
related to disulfid
performed in an aci
adjacent or close
Therefore, two pr
methodology. Th
number of mod

a single disulfide bond has been reduced. Nascent sulfhydryls are immediately cyanylated by 1-cyano-4-dimethylamino-pyridinium tetrafluoroborate (CDAP) (10) under the same buffer conditions. The partially reduced and cyanylated protein isomers are then separated by reversed-phase HPLC, followed by analysis of HPLC fractions by MALDI-MS to determine which isomers are singly reduced/cyanylated. Those shifted by +52 Da correspond to a singly reduced/cyanylated species; a shift of +104 Da corresponds to doubly reduced/cyanylated species, etc. Those isomers with a 52-Da mass shift from the mass of the intact protein are dried and subjected to specific chemical cleavage in aqueous ammonia. The cleaved peptides, which may be linked by residual disulfide bonds, are then completely reduced to give a mixture of peptides that can be mass-mapped by MALDI-MS. The masses of the resulting peptide fragments are related to the location of the paired cysteines that had undergone reduction, cyanylation, and cleavage. A primary advantage of this approach is its underlying simplicity. The data obtained from our approach are straightforward and easily interpreted because only a few protein isomers are produced, and each is relevant. Typically, only **n-1** intermediates are needed to define an **n**-bridge system. Redundant information is always obtained to confirm an assignment. Secondly, we describe a practical way to circumvent problems related to disulfide bond scrambling because both reduction and cyanylation are performed in an acidic medium. Finally, our approach can be used for the assignment of adjacent or closely spaced cysteines for which conventional methodology fails. Therefore, two problems associated with current methodologies are solved using our methodology. The feasibility of the new approach is demonstrated by analyzing a number of model proteins with various disulfide bond linkages. The reported

experimental condition
cyanylation, HPLC sep

III. Experimental Se

Mass Spectrometry

MALDI mass
spectrometer (PerSep
VSL-337ND nitroger
the ion source was s
mode of operation.
calibration using sta
5734.5), and horse s
(St. Louis, MO). A
acid (Aldrich Chem
were prepared in a
equal volumes with
plate. The mixtu
spectrometer.

experimental conditions are optimized for partial reduction of the protein, sulfhydryl cyanylation, HPLC separation, and cleavage.

III. Experimental Section

Mass Spectrometry

MALDI mass spectra were obtained on a Voyager Elite time-of-flight (TOF) mass spectrometer (PerSeptive Biosystems Inc., Framingham, MA) equipped with a model VSL-337ND nitrogen laser (Laser Science, Newton, MA). The accelerating voltage in the ion source was set to 25 kV. Data were acquired in the positive or negative linear mode of operation. Time-to-mass conversion was achieved by external and/or internal calibration using standards of bradykinin (m/z 1061.2), bovine pancreatic insulin (m/z 5734.5), and horse skeletal myoglobin (m/z 16,952) obtained from Sigma Chemical Co. (St. Louis, MO). All experiments were performed using α-cyano-4-hydroxycinnamic acid (Aldrich Chemical Co., Milwaukee, WI) as the matrix. Saturated matrix solutions were prepared in a 50% (v/v) solution of acetonitrile/aqueous 0.1% TFA, and mixed in equal volumes with peptide or protein samples, and applied to a stainless-steel sample plate. The mixture was allowed to air dry before being introduced into the mass spectrometer.

Chemicals

Tris(2-carboxy
Chemical Co. (Rockf
Mannheim Biochemic
A, bovine milk α-lact
inhibitor (BPTI), citr
tetrafluoroborate (Cl
purification. Recom
factor (hEGF, contai
also purchased fro
Acetonitrile and TFA
pH 3.0 was prepared
with little deteriorati
freshly prepared pric

Partial Reduction

Ten-nmol p
citrate buffer (pH
disulfide bonds wa
in the protein (e.g.
mol of ribonuclea
temperature for 1

Chemicals

Tris(2-carboxyethyl)phosphine (TCEP) hydrochloride was purchased from Pierce Chemical Co. (Rockford, IL). Guanidine hydrochloride was a product of Boehringer-Mannheim Biochemicals (Indianapolis, IN). Bovine pancreatic ribonuclease A type III-A, bovine milk α-lactalbumin, soybean trypsin inhibitor (STI), bovine pancreatic trypsin inhibitor (BPTI), citric acid, sodium citrate, and 1-cyano-4-dimethylamino-pyridinium tetrafluoroborate (CDAP) were purchased from Sigma and used without further purification. Recombinant LR^3IGF-I (I1271), recombinant human epidermal growth factor (hEGF, containing a Met at the N-terminal as the initiator of biosynthesis) were also purchased from Sigma and purified by reversed-phase HPLC before use. Acetonitrile and TFA were of HPLC grade. The TCEP solution in 0.1 M citrate buffer at pH 3.0 was prepared as 0.10 M stock solution and stored under N_2 at -20°C for weeks with little deterioration. The 0.10 M CDAP solution in 0.1 M citrate buffer at pH 3.0 was freshly prepared prior to use.

Partial Reduction of Proteins

Ten-nmol protein samples were solubilized and denatured in 10 μl of 0.1 M citrate buffer (pH 3.0) containing 6M guanidine-HCl. Partial reduction of protein disulfide bonds was carried out by adding an equivalent of TCEP for the cystine content in the protein (e.g., 40 nmol of TCEP was reacted with 10 nmol of ribonuclease A as 1 mol of ribonuclease A contains 4 moles of cystine), followed by incubation at room temperature for 10-15 minutes. Depending on the protein understudy, the amount of

reducing agent may b
reduction products.

Cyanylation of Nasc

To the partia
CDAP solution over
groups was accompli

HPLC Separation o

Partially redu
with linear gradien
computer. UV det
(#218TP54, 10μm p
10μm particle size,
proteins were sligh
fractions were colle
determined by MA
sample size of 10-μ
columns; the use
sizes.

reducing agent may be adjusted to ensure the singly reduced isoforms are predominant reduction products.

Cyanylation of Nascent Sulfhydryls

To the partially reduced protein mixture was added a 20-fold molar excess of CDAP solution over the total cysteine content. Cyanylation of the nascent sulfhydryl groups was accomplished by incubation at room temperature for another 10-15 minutes.

HPLC Separation of Partially Reduced and Cyanylated Protein Isomers

Partially reduced and cyanylated species were separated by reversed-phase HPLC with linear gradient elution using Waters model 6000 pumps controlled by a PC computer. UV detection was at 215 nm. The columns were either a Vydac C18 (#218TP54, 10μm particle size, 300-Å pore, 4.6×250 mm) or a Vydac C4 (#214TP54, 10μm particle size, 300-Å pore, 4.6×250 mm). The HPLC conditions for individual proteins were slightly different (see chromatograms shown below). The major HPLC fractions were collected manually and the masses of the collected protein isomers were determined by MALDI-MS. Appropriate fractions were then dried for further use. The sample size of 10-nanomoles was used for convenient detection from conventional HPLC columns; the use of microbore columns should allow the use of much smaller sample sizes.

Cleavage of Singly R

To the dried H
aqueous solution to d
of the peptide chain w
removed in a vacuum

Complete Reduction

Truncated pe
reduced by reacting
minimize the possibi
acetonitrile/0.1% TF

IV. Results and Di

A. Partial Reducti

There are tw
reduction and parti
bonds in a protein
denaturing conditio
selectively split by
three disulfide bon
(11). Sequential

Cleavage of Singly Reduced and Cyanylated Protein Isomers

To the dried HPLC fraction was added 2μl of 6 M guanidine-HCl in 1M NH_4OH aqueous solution to dissolve the protein residue and then 5μl of 1 M NH_4OH. Cleavage of the peptide chain was performed at room temperature for 1 hour. Excess ammonia was removed in a vacuum system.

Complete Reduction of Remaining Disulfide Bonds

Truncated peptides, still linked by residual disulfide bonds, were completely reduced by reacting with 2μl of 0.1 M TCEP solution at 37°C for 30 minutes at pH 3-5 to minimize the possibility of reoxidation. Samples were diluted with 100μl of a 50% (v/v) acetonitrile/0.1% TFA solution prior to analysis by MALDI-MS.

IV. Results and Discussion

A. Partial Reduction of Proteins

There are two terminologies frequently used in disulfide reduction: sequential reduction and partial reduction. The former refers to the situation in which disulfide bonds in a protein are opened one by one by using different reducing agents and/or denaturing conditions. For example, only the most labile disulfide bond in papain can be selectively split by the addition of 2-mercaptoethanol in the presence of 8M urea, but all three disulfide bonds can be reduced in 6M guanidine-HCl by the same reducing agent (11). Sequential reduction is practically limited in use because it is difficult, if not

impossible, to choose
protein can be cleave
latter term, partial re
small portion of each
desired products are
bridges. As illustra
controlled so that ea
resulting in a mixture
is understandable as
same protein molecu
of each disulfide bo
which are negligi
Consequently, the
minor singly reduce
can easily be ach
stoichiometry of re

Although e
reduced to a simila
of various disulfid
different, dependi
conformational ene
(12-15). Our m

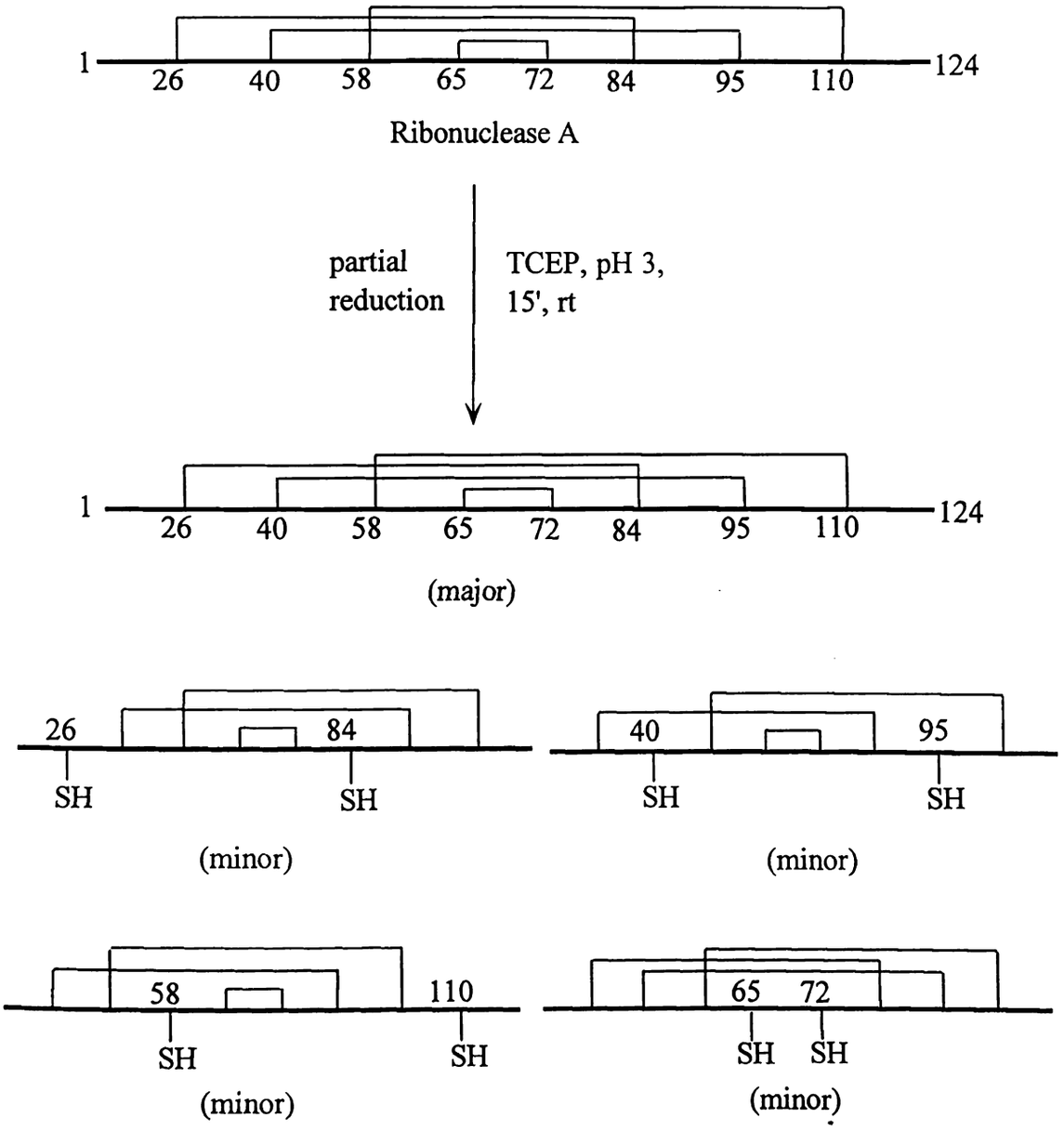
impossible, to choose appropriate conditions so that only one of the disulfide bonds in a protein can be cleaved at a time without the reduction of the other disulfide bonds. The latter term, partial reduction, refers to limited, controlled conditions under which only a small portion of each of the disulfide bonds, e. g., 10% of each, is reduced. That is, the desired products are those in which reduction has opened some, but not all disulfide bridges. As illustrated in Figure 4.3 for ribonuclease A, reduction conditions can be controlled so that each of the four disulfide bonds is reduced to a minor extent ($\leq$ 10%), resulting in a mixture of four singly reduced isomers as a main reduced form. This result is understandable as the probabilities of reducing two and three disulfide bonds in the same protein molecule (under the above conditions where we observed a 10% reduction of each disulfide bond) are 10%×10% = 1% and 10%×10%×10% = 0.1%, respectively, which are negligible in comparison to obtaining the singly reduced isoforms. Consequently, the mixture of partially reduced protein contains major intact proteins, minor singly reduced isoform, trace doubly reduced isoform, and so on. Practically, this can easily be achieved by controlling incubation time, reaction temperature, and stoichiometry of reducing agents.

Although every disulfide bond has the same redox potential and should be reduced to a similar extent from the standpoint of thermodynamics, the reduction kinetics of various disulfide bonds in solutions of native or partially denatured protein are quite different, depending on accessibility of the disulfide bonds to reducing agents, the conformational energy of the disulfide bonds, and the redox potential of the environment (12-15). Our methodology requires that individual disulfide bonds be broken at

(m

(m

Figure



+ doubly reduced isoforms (trace)
+ triply reduced isoforms (very trace)
.....

Figure 4.3. Conceptional scheme to illustrate "Partial Reduction".

comparable rates. T
study were denature
accessibility of reduc

Water-soluble
bonds (5-7, 9). Red
bond scrambling. Fu
controlled which ma

The mixture
reduction states suc
conducted with ribo
and LR^{3}IGF-I, show
cystine content in th
15 min at room tem
20-fold excess of T
From the limited n
control is not critic
more readily be co
content may not b
protein of unknow

comparable rates. To minimize structural diversity of disulfide bonds, proteins under study were denatured by dissolution in 6 M guanidine so that differences in the accessibility of reducing agent to each disulfide bond was minimized.

Water-soluble TCEP has proved to be an excellent reducing agent for disulfide bonds (5-7, 9). Reduction by TCEP can be carried out at pH 3.0 to suppress disulfide bond scrambling. Furthermore, at pH 3.0, the reduction of disulfide bonds is kinetically controlled which makes partial reduction possible (5).

The mixture of partially reduced proteins contained intermediates with different reduction states such as singly reduced and doubly reduced isoforms. Experiments conducted with ribonuclease A, insulin, soybean trypsin inhibitor, α-lactalbumin, hEGF, and LR^3IGF-I, showed that by using approximately an equivalent of TCEP for the total cystine content in the proteins, about 5~10% of each disulfide bond was reduced within 15 min at room temperature. However, for bovine pancreatic trypsin inhibitor, at least a 20-fold excess of TCEP was required to achieve 10% reduction of each disulfide bond. From the limited number of proteins we have examined it appears that stoichiometry control is not critical, although BPTI was an exception. The extent of the reduction can more readily be controlled by the reaction time and temperature. However, as the cystine content may not be known *a priori*, the extent of reduction should be monitored for a protein of unknown cystine content.

B. Cyanylation of N

Jacobson *et a*

specifically cyanylate

cyano-4-dimethylami

same purpose (10).

chemistry. We empl

free and total cysteir

selectivity of the N

conditions which ma

has been reported

conditions (9, 10, 1

in peptide and pro

cyanylation by CDA

e.g., pH 3-5 at roo

unnecessary to rem

CDAP. However, a

over peptide sulfhy

could result in mod

control the stoichi

hand, CDAP react

reaction system wi

cyanylation is poss

unknown protein.

B. Cyanylation of Nascent Sulfhydryls

Jacobson *et al.* (8) first showed that 2-nitro-5-thiocyanobenzoic acid (NTCB) specifically cyanylates cysteine thiols in mildly alkaline solutions. Another reagent, 1-cyano-4-dimethylamino-pyridinium tetrafluoroborate (CDAP) was later proposed for the same purpose (10). However, only NTCB has been extensively applied in protein chemistry. We employed the NTCB reaction to characterize the number and location of free and total cysteine groups in peptides and proteins (7). In spite of the reactivity and selectivity of the NTCB reagent, the cyanylation must be performed under alkaline conditions which may permit some disulfide bond scrambling. CDAP, on the other hand, has been reported to be a selective and reactive cyanylation reagent under acidic conditions (9, 10, 16, 17). The utility of the CDAP for cyanylation of sulfhydryl groups in peptide and proteins has been extensively investigated in chapter 3. Complete cyanylation by CDAP was possible by using a 5-fold molar excess under mild conditions, e.g., pH 3-5 at room temperature, which is compatible with TCEP. It is therefore unnecessary to remove excess TCEP, change buffer, or readjust the pH prior to using CDAP. However, as shown in chapter 3, a large excess of CDAP (~50-fold molar excess over peptide sulfhydryls) and excessive incubation time (> 2 hours at room temperature) could result in modification of other amino acid side chains. Therefore, it is necessary to control the stoichiometry of CDAP in order to minimize side reactions. On the other hand, CDAP reacts instantly with TCEP even at pH 3.0, the excess of TCEP in the reaction system will be killed upon addition of CDAP, and no further reduction during cyanylation is possible. This is definitely an advantage for the control of reduction in an unknown protein. However, it should be noted that a larger amount of CDAP has to be

applied in case more
structure, such as BPT

C. HPLC Separatio

The mixture
intact protein and a
reversed-phase HPL
bond(s), i. e., partial
4.4 is a typical
ribonuclease A. Bes
are four other peaks
greater than that
reduced/cyanylated
the same, suggestin
disulfide bonds in
other proteins will

The separa
hydrophobicity of
structure, exposes
protein's hydroph
reduced disulfide
thus, should elute
support this spec

applied in case more TCEP is required for partial reduction of proteins with a tight structure, such as BPTI.

C. HPLC Separation of Partially Reduced and Cyanylated Protein Isomers

The mixture obtained from the above reactions contains a majority of residual intact protein and a minority of partially reduced/cyanylated isoforms. The capacity of reversed-phase HPLC to separate these protein isomers with various residual disulfide bond(s), i. e., partially reduced isoforms, is demonstrated in following examples. Figure 4.4 is a typical chromatogram of the partially reduced/cyanylated isoforms of ribonuclease A. Besides the main peak that corresponds to residual ribonuclease A, there are four other peaks (marked 1-4), each corresponding to a species having a mass ~52 Da greater than that of the original protein. These peaks are attributable to singly reduced/cyanylated protein isomers. The intensities of the four peaks are approximately the same, suggesting that under our reaction conditions, the reducing rates of the four disulfide bonds in denatured ribonuclease A are comparable. The HPLC separation of other proteins will be discussed in individual sections below.

The separation mechanism in HPLC is based on the difference in the hydrophobicity of analytes (5, 18). Opening a given disulfide bond disrupts protein structure, exposes the protein's interior hydrophobic amino acids, and increases the protein's hydrophobicity to different extents. Therefore, isomers with one or more reduced disulfide bond(s) should tend to interact more strongly with HPLC columns and, thus, should elute at later retention time than the original protein. The experimental data support this speculation. Almost all partially reduced isomers show longer retention

Figure 4.4. HPL
reduced/cyanylat
phase HPLC on
gradient 20-40%
in CH_3CN. Peak
MALDI-TOF an

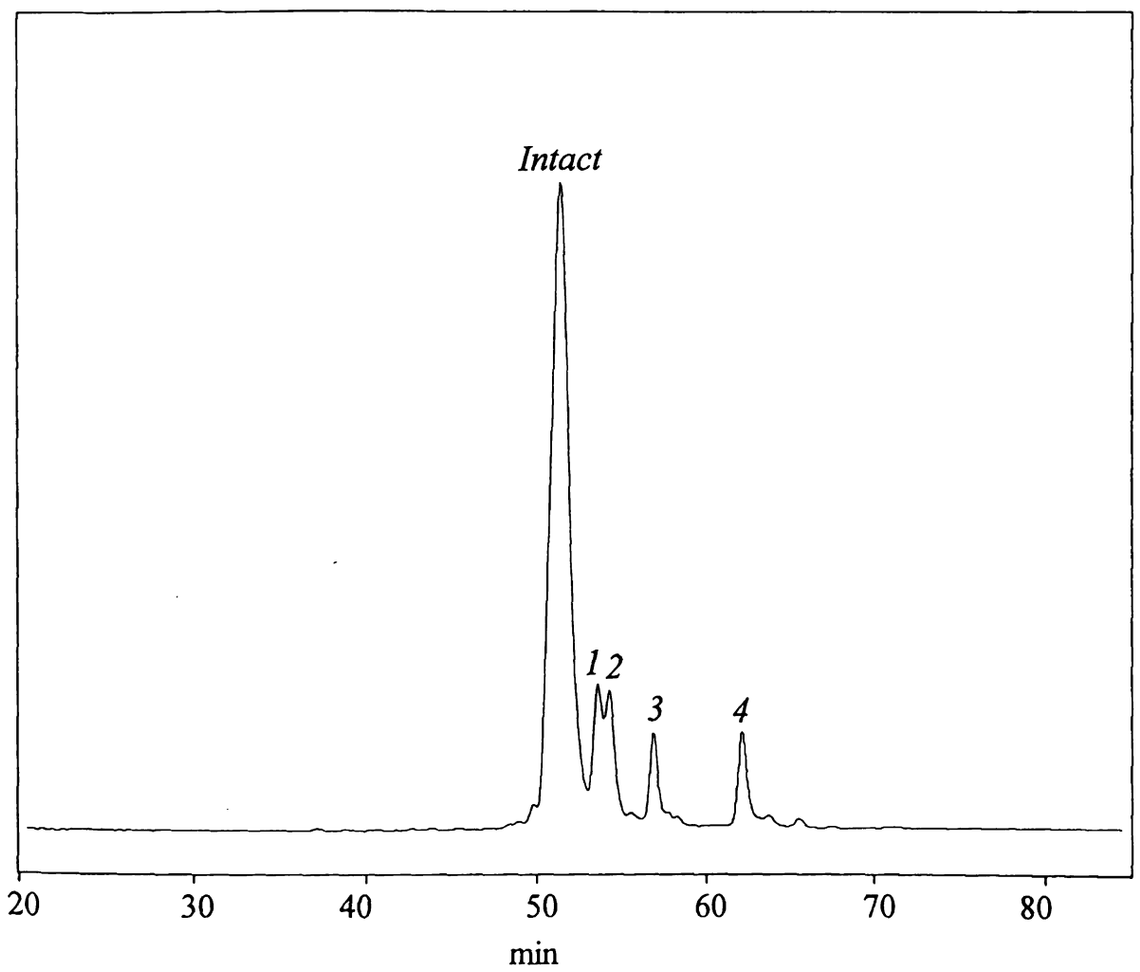


Figure 4.4. HPLC separation of denatured ribonuclease A and its partially reduced/cyanylated isomers. Ten-nmol of the protein were separated by reversed-phase HPLC on a Vydac C18 column at a flow rate of 1.5 ml/min with a linear gradient 20-40% B in 90 minutes, where A = 0.1% TFA in water and B = 0.1% TFA in CH_3CN. Peaks 1-4 represent singly reduced/cyanylated species, as determined by MALDI-TOF analysis.

times than their intact

from the HPLC data,

dramatically change t

the unusually broad

most likely due to di

suggested by Gray (5

and cyanylation, exh

column in the water-

this case for faster

propanol within a re

It should be

isomers, it is unnec

below for ribonucle

two respective prot

unambiguous assig

described methodol

D. Cleavage of Pe

One of the

bond of cyanylate

usually employed

hours at 37°C) in a

times than their intact original proteins. The differences in hydrophobicity, as reflected from the HPLC data, indirectly suggests that reduction of individual disulfide bonds can dramatically change the three-dimensional structure of a protein. Another observation is the unusually broad HPLC peak for some proteins (e.g., α-lactalbumin, STI), which is most likely due to different conformations that show slightly different retention times, as suggested by Gray (5). Some hydrophobic proteins (e.g., STI), especially after reduction and cyanylation, exhibited a long retention time or even irreversible retention on the column in the water-acetonitrile mobile phase. 1-Propanol was substituted for CH_3CN in this case for faster elution. Most of the proteins can be eluted by 40% aqueous 1-propanol within a reasonable time (19).

It should be pointed out that although HPLC is able to separate the protein isomers, it is unnecessary for all components to be baseline separated. As described below for ribonuclease A, even with poorly resolved chromatographic peaks representing two respective protein isomers, interpretation of MALDI data is still possible for the unambiguous assignment of disulfide bonds. This provides a valuable advantage of the described methodology for analyzing complicated mixtures.

D. Cleavage of Peptide Chains

One of the key features of our methodology is cleavage at the N-terminal peptide bond of cyanylated cysteinyl residues. Previously described methods (7, 8, 20-23) usually employed mildly alkaline conditions (pH~9) and long hours of incubation (>16 hours at 37°C) in attempts to control the β-elimination reaction, a side reaction that results

in a lower yield of t

elimination over the

cleavage kinetics as

amino acids adjacent

show that higher pH

for most of the p

significantly even at

within an hour at roo

1 M NH_4OH (pH~1

our experiments, 1 N

from the reaction sy

using higher pH is

reversed. Modifica

conditions describe

After cleava

linked by the remai

the peptide mixtu

subsequent analysi

E. Interpretation

Unlike the

products, MALD

in a lower yield of the cleavage reaction. An elevated pH was reported to favor β-elimination over the cleavage reaction (20, 21). We have systematically studied the cleavage kinetics as a function of pH for cyanylated polypeptides containing different amino acids adjacent to the N-terminus of cyanylated cysteines. The results in chapter 3 show that higher pH can greatly accelerate the cleavage and β-elimination reactions, but for most of the peptides studied, the extent of β-elimination does not increase significantly even at a pH as high as 12, a condition at which cleavage can be complete within an hour at room temperature. Cleavage can be accomplished in 0.01 M NaOH or 1 M NH_4OH (pH~11) solutions containing denaturing agent; both give similar results. In our experiments, 1 M NH_4OH is preferred over 0.01 M NaOH because it can be removed from the reaction system after completion of the cleavage reaction. One side reaction in using higher pH is the neutralization of -COOH groups; however, this can be easily reversed. Modification or cleavage of other side chains was rarely observed under the conditions described above.

After cleavage at cyanylated cysteinyl residues, the truncated peptide chains, still linked by the remaining disulfide bonds, can be easily reduced by excess TCEP. Finally, the peptide mixture is diluted to minimize the adverse effect of guanidine on the subsequent analysis by MALDI.

E. Interpretation of MALDI Data

Unlike the MALDI spectra of a complicated mixture of enzymatic digestion products, MALDI data derived from analysis of the cleavage reaction mixture of a

cyanylated protein are

cyanylated cysteinyl

reduced/cyanylated

corresponding to the

related to the positio

to deduce the disul

cleavage, provides

otherwise would hav

assignment.

Ribonuclease

four disulfide bon

Cys72. Figure 4.5

F

Table 4.1 lists the

peptide chains at

cyanylated. Figur

cleavage of isome

peaks 1-4, respect

cyanylated protein are easy to interpret. Cleavage of the peptide chain takes place only at cyanylated cysteinyl sites, which in principle yields three fragments for each singly reduced/cyanylated protein isomer (and sometimes two overlapped fragments corresponding to the β-elimination at one cysteine site). The mass of each fragment is related to the position of the two cyanylated cysteinyl residues which in turn can be used to deduce the disulfide bond linkage. β-Elimination, an alternative to peptide chain cleavage, provides mass spectral data corresponding to overlapped peptides (that otherwise would have cleaved) and serves as a confirmation for the disulfide bond pairing assignment.

Ribonuclease A (Mr = 13,683) (24) contains 124 amino acids that are linked by four disulfide bonds: Cys26-Cys84, Cys40-Cys95, Cys58-Cys110, and Cys65-Cys72. Figure 4.5 shows the disulfide structure of ribonuclease A.

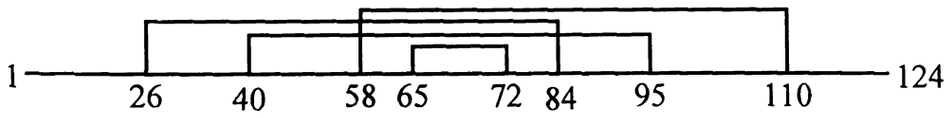


Figure 4.5. Disulfide structure of ribonuclease A.

Table 4.1 lists the calculated m/z values for possible fragments due to cleavage of the peptide chains at different sites depending on which disulfide bond was reduced and cyanylated. Figure 4.6A-D are four MALDI spectra of peptide mixtures resulting from cleavage of isomers of singly reduced/cyanylated ribonuclease A corresponding to HPLC peaks 1-4, respectively.

Table 4.1.
resulting fr
designated

Reductio of Disulf
Cys26-C
Cys65-C
Cys58-C
Cys40

Table 4.1. Calculated and observed m/z values for possible fragments resulting from the cleavage reaction of ribonuclease A chains at sites of designated cysteine pairs

Reduction of Disulfide	Fragment	Calculated m/z	Observed m/z
	1-25	2702.8	2705.3
	26-83	6547.3	6548.5
Cys26-Cys84	84-124	4526.0	4527.4
	1-83	9176.2	9176.7
	26-124	10995.3	10998.6
	1-64	7083.9	7083.8
	65-71	789.8	nd
Cys65-Cys72	72-124	5906.5	5907.7
	1-71	7795.7	7790.0
	65-124	6618.3	6617.9
	1-57	6353.1	6351.1
	58-109	5767.4	5766.8
Cys58-Cys110	110-124	1659.8	1659.8
	1-109	12042.4	12036.7
	58-124	7349.1	nd
	1-39	4413.9	4414.4
	40-94	6063.6	6061.2
Cys40-Cys95	95-124	3302.7	3303.7
	1-94	10399.5	10430.5
	40-124	9288.3	9293.4

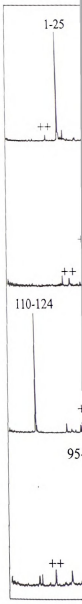


Figure 4.6. Th
of the four sing
HPLC peaks 1
doubly-charge
peaks 1-124 in
occurred at on
information. S

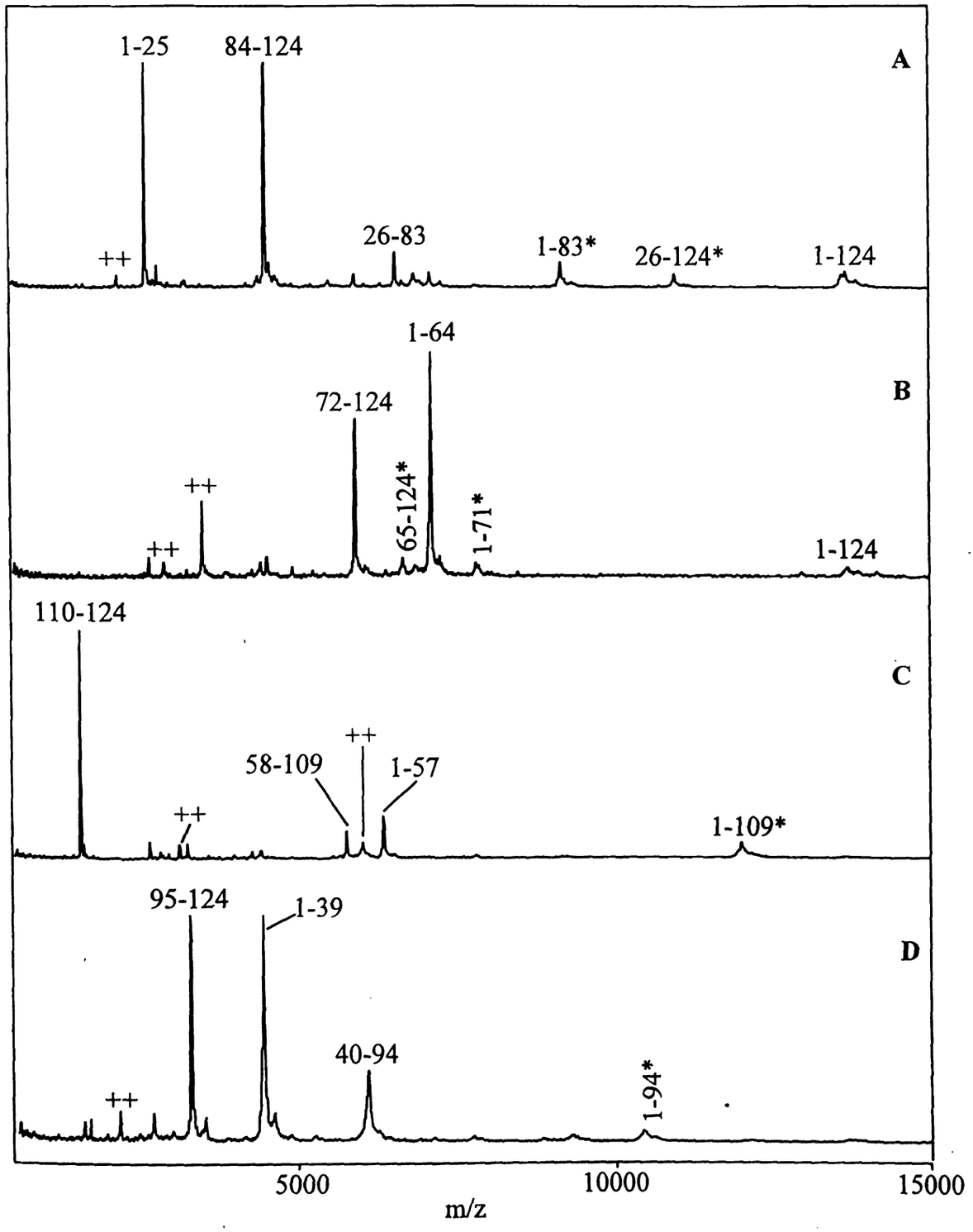


Figure 4.6. The MALDI mass spectra of peptide mixtures resulting from the cleavage of the four singly reduced/cyanylated ribonuclease A isomers, corresponding to the HPLC peaks 1-4 in Figure 4.4, respectively. The symbols ++ and * represent the doubly-charged species and protonated β-elimination products, respectively. The peaks 1-124 in (A) and (B) represent intact proteins in which β-elimination only occurred at one cysteinyl residue. These products do not provide any specific information. See also Table 4.1 for calculated and observed m/z values.

The mass spectrum
HPLC peak 1 in Fig
fragments 1-25, 26
4526.0) with a relat
peptide chain cleav
m/z 9176.7 corresp
cleavage at Cys84.
10998.6 is another
at Cys84. Overall,
deduced.

With simil
Cys58-Cys110, als
The MALDI spe
represented by H
7083.8, correspon
fragment 65-71 i
these two fragme
elimination proc
confirmation of t
65-124, while tha

The HPL
carefully collect

The mass spectrum in Figure 4.6A corresponds to the cleavage products represented by HPLC peak 1 in Figure 4.4. Three peaks at m/z 2705.3, 6548.5, and 4527.4 are due to fragments 1-25, 26-83, and 84-124, respectively (expected m/z 2706.8, 6547.3, and 4526.0) with a relative mass deviation of <0.05%. From these data, one can deduce that peptide chain cleavages occur at Cys26 and Cys84. Additionally, the MALDI peak at m/z 9176.7 corresponds to an overlapped peptide, 1-83, resulting from peptide chain cleavage at Cys84, but β-elimination at Cys26. Likewise, the MALDI peak at m/z 10998.6 is another overlapped peptide, 26-124, with cleavage at Cys26, but β-elimination at Cys84. Overall, a disulfide bond linkage between Cys26-Cys84 can be unambiguously deduced.

With similar strategy, two other disulfide bond linkages, Cys40-Cys95 and Cys58-Cys110, also can be recognized from Figure 4.6C and Figure 4.6D, respectively. The MALDI spectrum in Figure 4.6B, corresponding to the cleavage products represented by HPLC peak 2 (Figure 4.4), shows two main peaks at m/z 5907.7 and 7083.8, corresponding to fragment 72-124 and 1-64, respectively. Another expected fragment 65-71 is missing. However, it is still possible to deduce Cys65-Cys72 from these two fragments because no other combination gives such masses. Two minor β-elimination products, m/z 6617.9 and 7790.0, are particularly informative for confirmation of the assignment in this case. The peak at m/z 6617.9 represents residues 65-124, while that at m/z 7795.7 covers residues 1-71.

The HPLC peaks 1 and 2 in Figure 4.4 were not resolved completely. Even carefully collected HPLC fractions from one component still contained a small amount of

the other component

the cleavage produc

corresponding to t

assignment of the re

possible because on

In the earli

optimized, the com

collected as a broad

fraction is given in

spectrum A and B

pairs simultaneous

pairs in a mixture.

and doubly reduce

bond pairs from

gives more compl

that the assignmer

more complicated

responses from th

isoform might be

correct assignme

the other component. This is reflected in the MALDI spectra (Figures. 4.6A and B) of the cleavage products of these two fractions, each of which contains small fragments corresponding to the cleavage products of the other fraction. The unambiguous assignment of the respective disulfide bond pairs in the presence of another isomer is still possible because only a few fragments are produced.

In the earlier stage of this research, the HPLC conditions were not fully optimized, the compounds represented by peaks 1 and 2 in Figure 4.4 coeluted and were collected as a broad peak. The MALDI spectrum of the cleavage products of that broad fraction is given in Figure 4.7. It is apparent that this spectrum is a perfect overlap of the spectrum A and B in Figure 4.6. From this spectrum, one can assign two disulfide bond pairs simultaneously, demonstrating the capability of the methodology to assign disulfide pairs in a mixture. On the other hand, if the HPLC peak contains a mixture of both singly and doubly reduced/cyanylated isoforms, Caution should be used in assigning disulfide bond pairs from the fragments. Because a doubly reduced/cyanylated protein isomer gives more complicated fragments than a singly reduced/cyanylated protein, it is expected that the assignment from such a mixture will be challenging. The situation could be even more complicated, because, due to the discrimination of MALDI responses, the MALDI responses from the cleavage products of even trace amount of doubly reduced/cyanylated isoform might be more intense than those of the singly reduced/cyanylated isoform. The correct assignment may be retarded in such a case.

Figure 4.7. Th
of a mixture o
charged specie

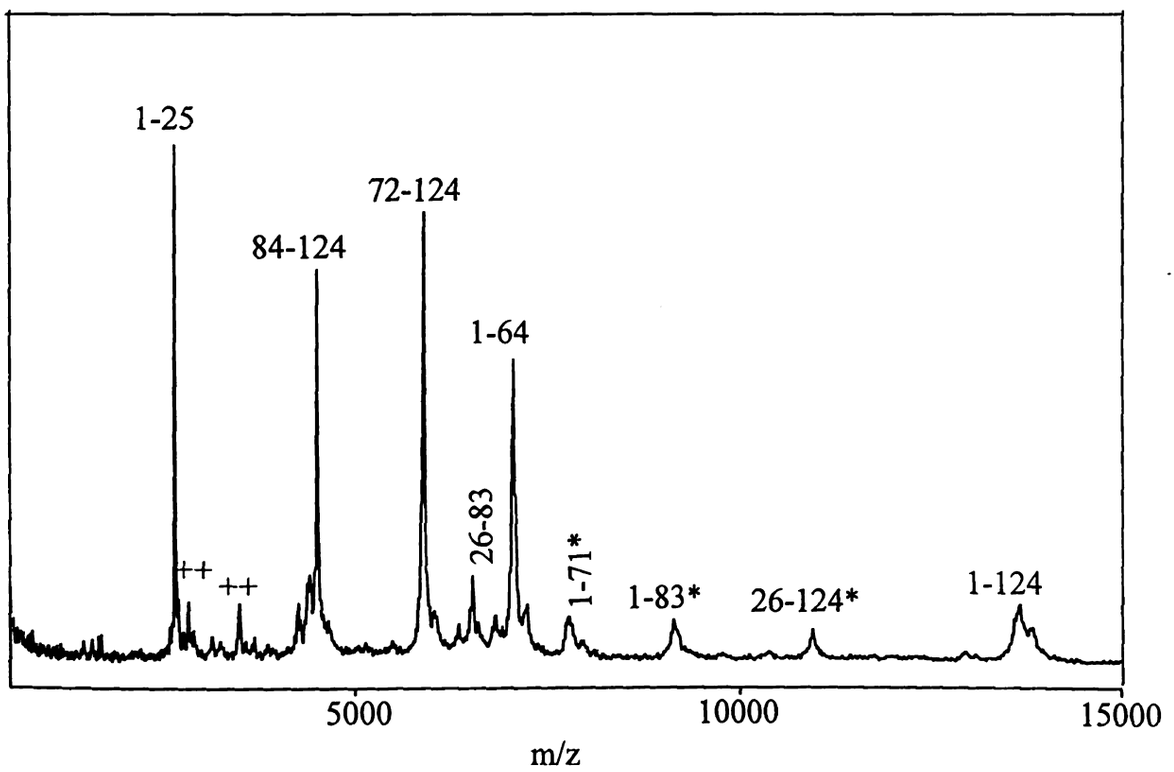


Figure 4.7. The MALDI mass spectrum of peptide mixtures resulting from the cleavage of a mixture of HPLC peaks 1 and 2. The symbols ++ and * represent the doubly-charged species and protonated β-elimination products, respectively.

We have obse

chain usually is less

5767.4, and m/z 606

for the middle fragm

reduces the yield of

α-Lactalbum

linked as 4 disulfide

and isomers of i

reduced/cyanylated

marked 1-4) are ob

isomers have sim

different disulfide

peaks 1-4 in Figur

corresponding MA

calculated m/z va

lactalbumin chain

Although

mass increase of

protein, MALDI

517.6) which wa

not detected, pos

it impossible to

We have observed that the middle fragment from cleavage of a cyanylated peptide chain usually is less abundant in the MALDI spectrum (e.g., m/z 6547.3, m/z 789.8, m/z 5767.4, and m/z 6063.6 in Table 4.1). A possible explanation for this observation is that for the middle fragment, β-elimination is able to occur at either side, which significantly reduces the yield of the expected cleavage product.

α-Lactalbumin (123 amino acids, Mr = 14,175) also contains 8 cysteines that are linked as 4 disulfide bonds (25). Figure 4.8 shows the HPLC separation of α-lactalbumin and isomers of its partially reduced/cyanylated species. Peaks for four singly reduced/cyanylated α-lactalbumin isomers (~52-Da mass shift from original molecule, marked 1-4) are observed in addition to that for residual α-lactalbumin. Again, the four isomers have similar abundances which suggest comparable reduction rates for the different disulfide bonds in denatured α-lactalbumin. The four HPLC fractions (HPLC peaks 1-4 in Figure 4.8) were subjected to cleavage under the described conditions. The corresponding MALDI mass spectra are shown in Figure 4.9A-D. Table 4.2 lists the calculated m/z values for possible fragments resulting from the cleavage reaction of α-lactalbumin chains at sites corresponding to different cysteine pairs.

Although the compound represented by HPLC peak 1 in Figure 4.8 showed a mass increase of ~52 Da, corresponding to an isomer of a singly reduced/cyanylated protein, MALDI-MS (Figure 4.9A) only detected one expected cleavage product (m/z 517.6) which was due to fragment 120-123. Two other fragments, 1-5 and 6-119, were not detected, possibly because of signal suppression. The insufficient information makes it impossible to deduce a disulfide bond linkage from the MALDI data.

Figure 4.8. HI
cyanylated iso
of 1.5 ml/min
in water and F
reduced/cyany

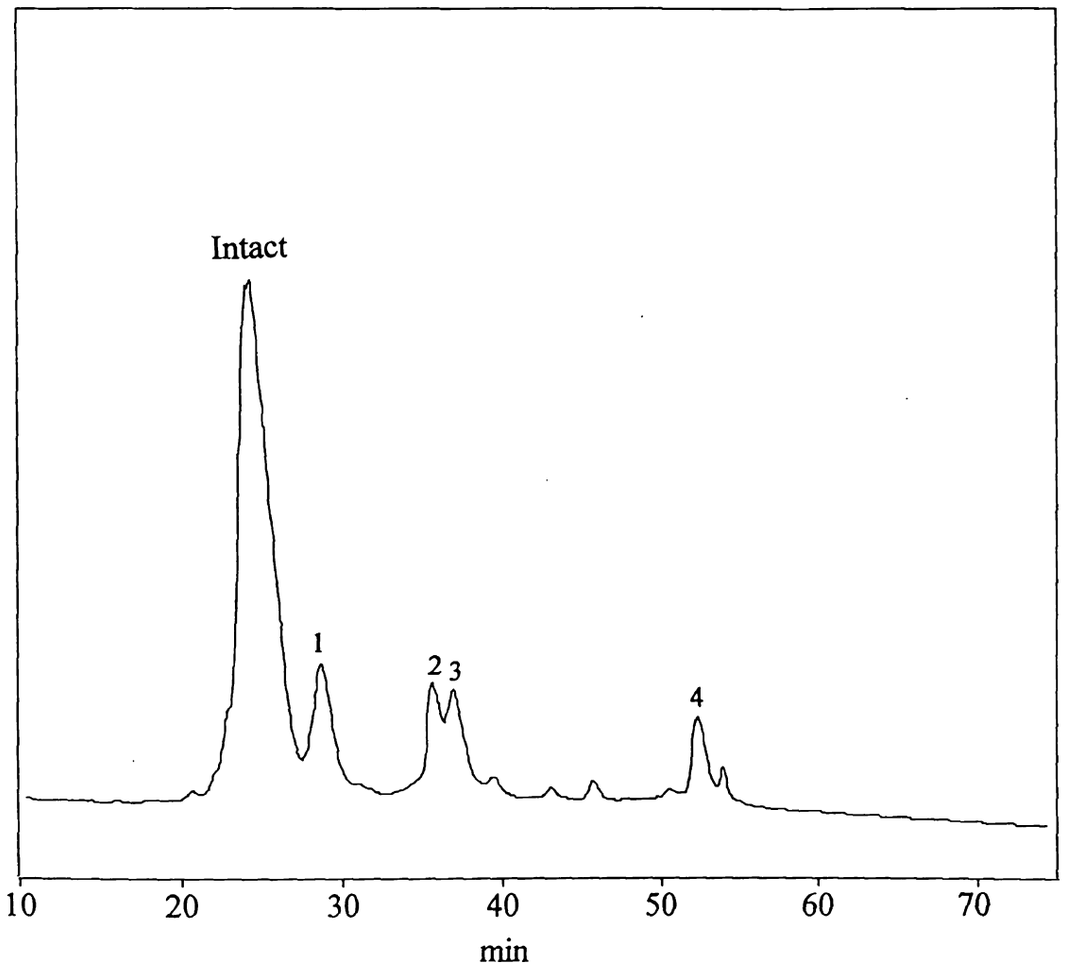


Figure 4.8. HPLC separation of denatured α-lactalbumin and its partially reduced/ cyanylated isomers. Separation was carried out on a Vydac C4 column at a flow rate of 1.5 ml/min with a linear gradient 40-50% B in 90 minutes, where A = 0.1% TFA in water and B = 0.1% TFA in CH_3CN. Peaks 1-4 represent singly reduced/cyanylated species, as determined by MALDI-TOF analysis.

Table 4.2. C
resulting fro
designated

| Reductio
of Disulf |
| --- |
| Cys6-Cys |
| Cys28-C |
| Cys61-C |
| Cys73 |

Table 4.2. Calculated and observed m/z values for possible fragments resulting from the cleavage reaction ofα-lactalbumin chains at sites of designated cysteine pairs

Reduction of Disulfide	Fragment	Calculated m/z	Observed m/z
	1-5	618.7	nd
	6-119	13135.7	nd
Cys6-Cys120	120-123	517.6	517.6
	1-118	13676	nd
	6-123	13575	nd
	1-27	3125.6	3122.5
	28-110	9525.6	9517.1
Cys28-Cys111	111-123	1620.8	1620.1
	1-110	12573	12560
	28-123	11068	nd
	1-60	6918.9	6914.4
	61-76	1800.8	1798.9
Cys61-Cys77	77-123	5552.5	5557.5
	1-76	8641.5	8648.6
	61-123	7275.3	7280.6
	1-72	8258.1	8264.5
	73-90	2098.3	2099.0
Cys73-Cys91	91-123	3915.6	3912.8
	1-90	10278	10303
	73-123	5936.0	nd

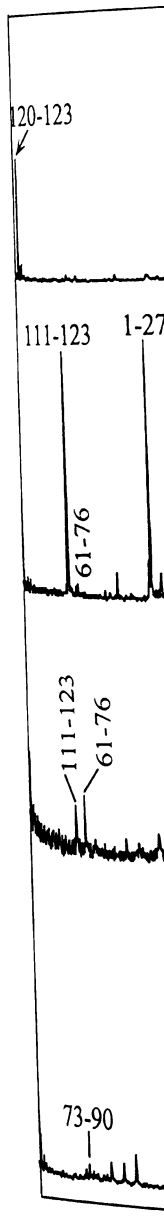


Figure 4.9. The
the four singly
peaks 1-4 in Fi
charged specie
for the calculat

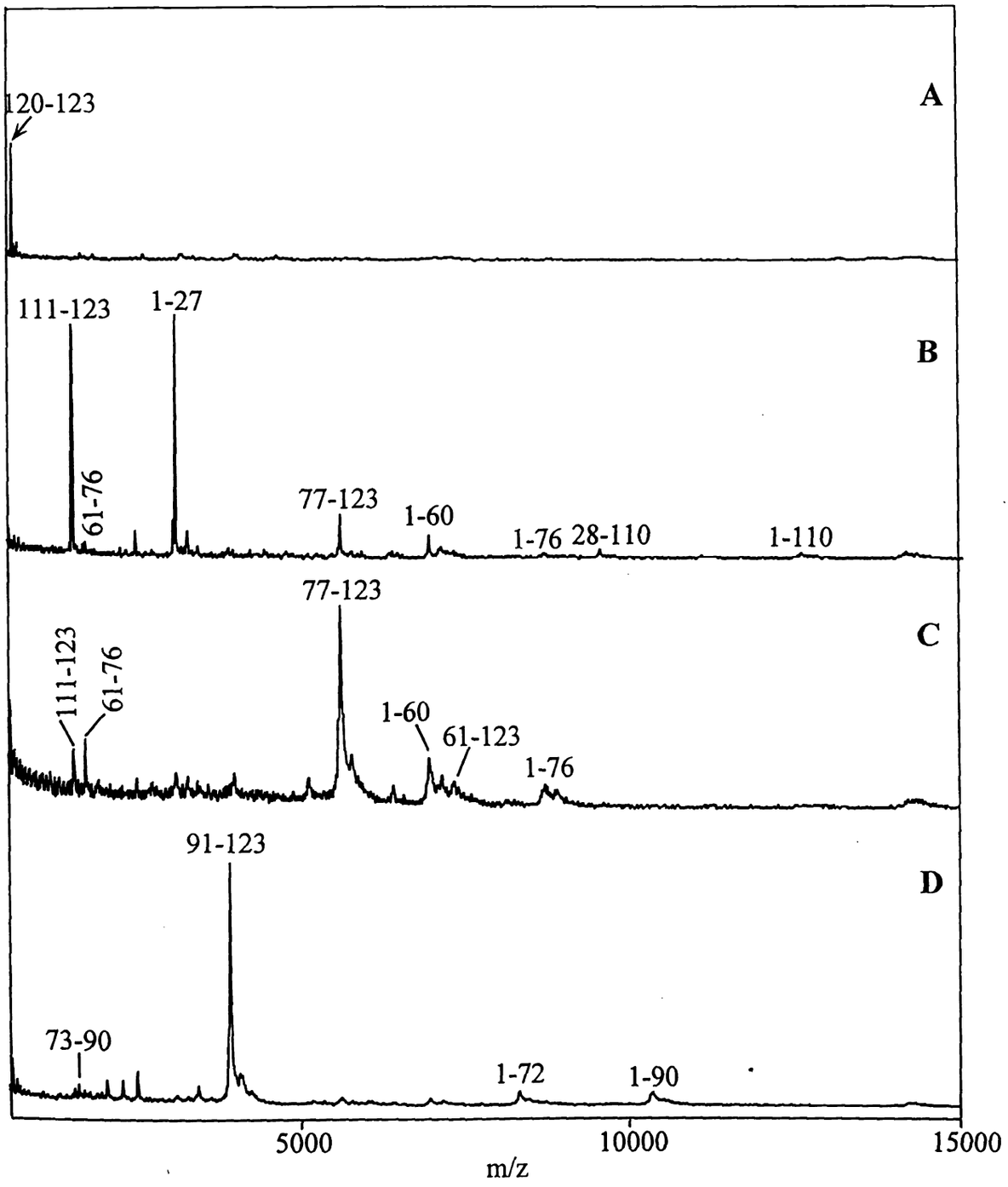


Figure 4.9. The MALDI mass spectra of peptide mixtures resulting from the cleavage of the four singly reduced/cyanylated α-lactalbumin isomers, corresponding to the HPLC peaks 1-4 in Figure 4.8., respectively. The symbols ++ and * represent the doubly-charged species and protonated b-elimination products, respectively. See also Table 4.2 for the calculated and observed m/z values.

The HPLC p

peak broadening. H

all the fragments re

singly reduced/cyan

1620.1 and 3122.5

peak at m/z 9517.

usually shows low

elimination. Furth

product 1-110. Th

peaks at m/z 179

respectively, resul

76 (m/z 8640.3) f

information is co

deduced from the

The Cys6

arising from anal

3 in Figure 4.8).

the fragments 6

products, 61-122

is linked to Cys

fraction 2. Fig

products in HP

The HPLC peaks 2 and 3 in Figure 4.8 were not baseline separated due to severe peak broadening. However, the MALDI-MS data interpretation permitted recognition of all the fragments resulting from the cleavage of the mixture of the two isomers of the singly reduced/cyanylated protein. As seen in Figure 4.9B, two intense peaks at m/z 1620.1 and 3122.5 can be assigned to fragments 111-123 and 1-27, respectively. A third peak at m/z 9517.1 matches the middle fragment 28-110 which, as mentioned above, usually shows lower abundance due to its opportunity for double participation in β-elimination. Furthermore, the peak at m/z 12560 is an indication of the β-elimination product 1-110. Thus, a disulfide bond, Cys28-Cys111, can be deduced. The other three peaks at m/z 1798.9, 5552.1, and 6909.9 match fragments 61-76, 77-123, and 1-60, respectively, resulting from cleavage at Cys61 and Cys77. The β-elimination product 1-76 (m/z 8640.3) further confirms that Cys77 is a cleavage site. The combination of this information is consistent with the assignment. Thus, two disulfide bond pairs can be deduced from the mass spectrum of an impure HPLC fraction.

The Cys61-Cys77 pair is further confirmed by the mass spectrum in Fig. 4-9C, arising from analysis by MALDI-MS of the cleavage products in HPLC fraction 3 (peak 3 in Figure 4.8). The three main peaks at m/z 1798.9, 5557.5, and 6914.4 correspond to the fragments 61-76, 77-123, and 1-60, respectively. In addition, two β-elimination products, 61-123 at m/z 7280.6 and 1-76 at m/z 8648.6, provide confirmation that Cys61 is linked to Cys77. The peak at m/z 1620.0 is likely due to some "carry-over" from fraction 2. Figure 4.9D was obtained from analysis by MALDI-MS of the cleavage products in HPLC fraction 4 (Figure 4.8). The three MALDI peaks at m/z 2099.0,

3912.8, and 8264.5
elimination product
Cys73 is linked to C

The other e
STI (Mr. 19977) co
and Cys136-Cys14
MALDI analysis
indicating the samp
had a very poor H
between STI and
analysis of STI w
but several impur
samples were sub
4.10, the elevated
difficult. Table
Preliminary MA
given in Figure
4.11A) correspo
indicating Cys1
they are not rea
detected for frag
is insufficient

3912.8, and 8264.5 represent fragments 73-90, 91-123, and 1-72, respectively. The β-elimination product (m/z 10300.2) is due to fragment 1-90. These data indicate that Cys73 is linked to Cys91.

The other example, soybean trypsin inhibitor (STI), is also demonstrated here. STI (Mr. 19977) contains 180 amino acids linked by two disulfide bonds, Cys39-Cys86 and Cys136-Cys145. (another variant of STI contains 181 amino acids with Mr. 20095. MALDI analysis of commercial STI gave a m/z 19988 Da as a predominant peak, indicating the sample contains 180 amino acids). Being a more hydrophobic protein, STI had a very poor HPLC behavior in CH_3CN/H_2O mobile phase. The strong interaction between STI and C18 or C4 column makes the HPLC separation impractical. The HPLC analysis of STI was improved greatly by using *n*-propanol as a mobile phase modifier, but several impurities were still poorly resolved from the STI peak. The purified STI samples were subjected to partial reduction and cyanylation. As illustrated in Figure 4.10, the elevated base line makes the separation of intact STI and its reduced isoforms difficult. Table 4.3 lists the m/z values of the expected and observed fragments. Preliminary MALDI analysis of the cleavage products from HPLC peaks 1 and 2 are given in Figure 4.11A and B. The MALDI peaks at m/z 4036.3 and 5017.2 (Figure 4.11A) correspond to the fragment 145-180 and an overlapped peptide 136-180, indicating Cys136 and Cys145 have been cleaved. Other peaks are also observed, but they are not readily related to the cleavage at any cysteine sites. Since no signals were detected for fragments 1-136 and 136-144, the information obtained by MALDI analysis is insufficient to draw any conclusion. Figure 4.11B was obtained by the MALDI

Figure 4.10. I
partially reduc
column at a fl
where A = 0.
TFA. Peaks
MALDI-TOF

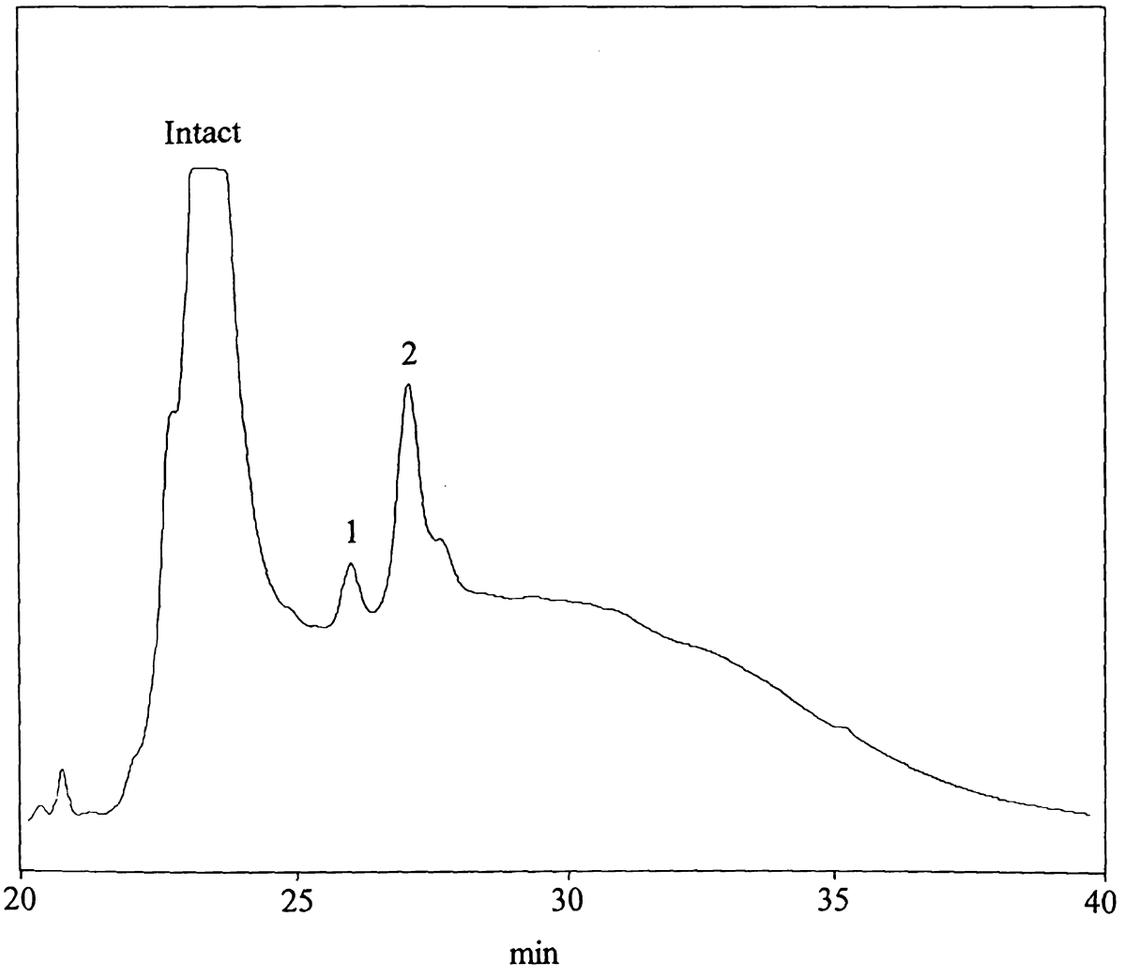


Figure 4.10. HPLC separation of denatured soybean trypsin inhibitor and its partially reduced/ cyanylated isomers. Separation was carried out on a Vydac C4 column at a flow rate of 1.2 ml/min with a linear gradient 15-50% B in 30 minutes, where A = 0.1% TFA in water and B = 80% 1-propanol/20% H_2O containing 0.1% TFA. Peaks 1 and 2 represent singly reduced/cyanylated species, as determined by MALDI-TOF analysis.

Table 4.3.
resulting f
at sites of

Reducti of Disu
Cys39-C
Cys136
Cys39 Cys13

Table 4.3. Calculated and observed m/z values for possible fragments resulting from the cleavage reaction of soybean trypsin inhibitor chains at sites of designated cysteine pairs

Reduction of Disulfide	Fragment	Calculated m/z	Observed m/z
Cys39-Cys86	1-38	4058.4	4059.1
	39-85	5258.1	5259.9
	86-180	10747	10752
Cys136-Cys145	1-135	14971	nd
	136-144	1057.1	nd
	145-180	4035.5	4036.3
	β(1-144)	15951	nd
	β(136-180)	5015.6	5017.2
Cys39-Cys86 Cys136-Cys145	1-38	4058.4	4057.1
	39-85	5258.1	5257.8
	86-135	5740.4	5738.8
	β(136-144)	1057.1	nd
	β(145-180)	4035.5	4037.9

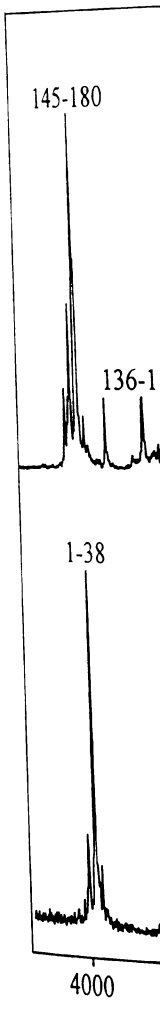


Figure 4.11. Th
of the two sing
corresponding
peaks are due t

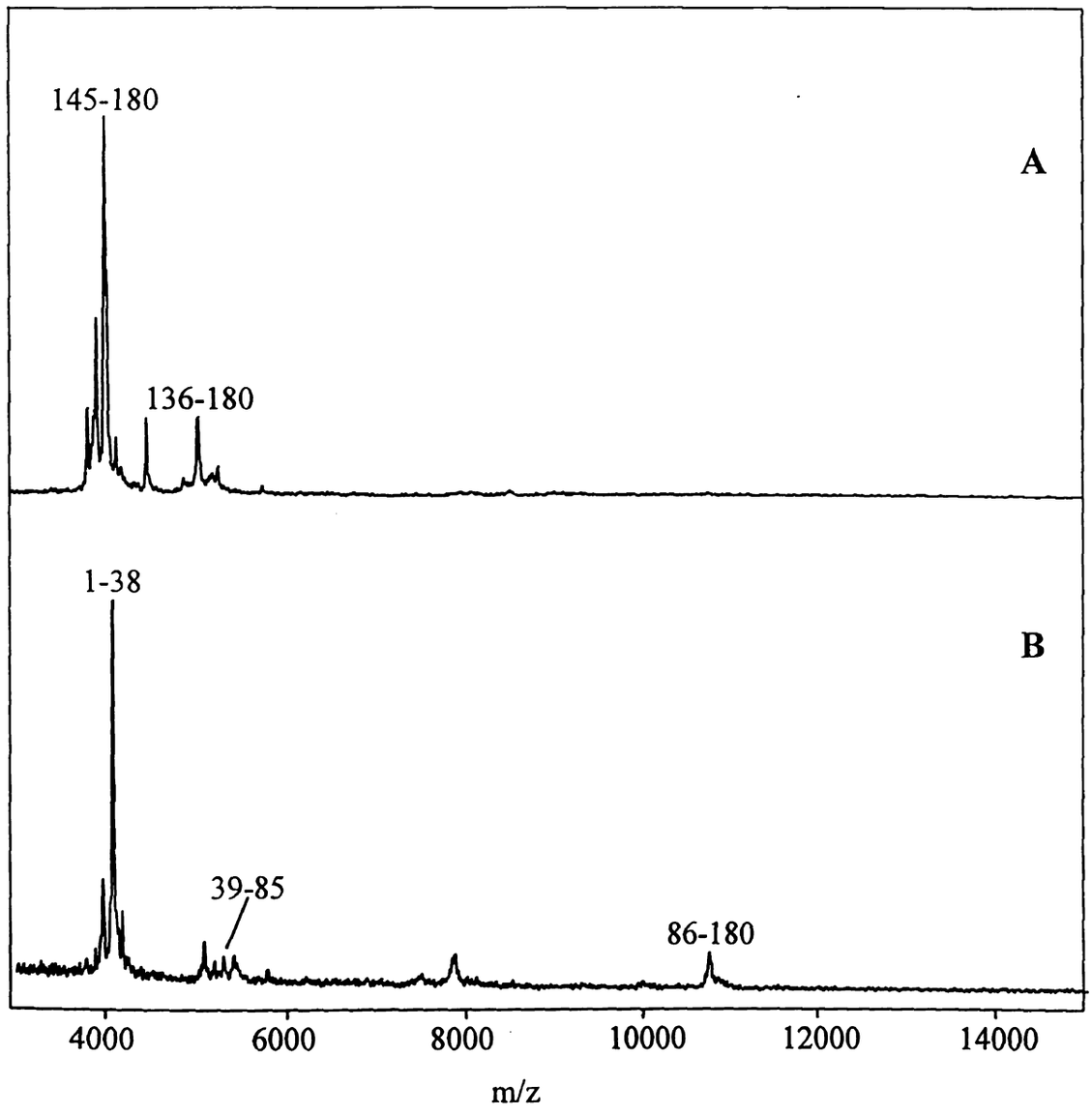


Figure 4.11. The MALDI mass spectra of peptide mixtures resulting from the cleavage of the two singly reduced/cyanylated soybean trypsin inhibitor (STI) isomers, corresponding to the HPLC peaks 1 and 2 in Figure 4.10, respectively. The unidentified peaks are due to impurities in the sample.

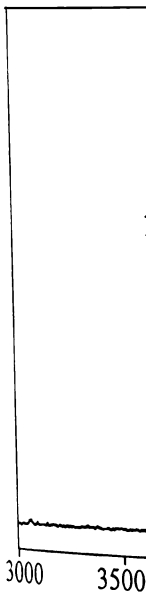


Figure 4.12. T
cleavage of un
isomers. Befo
remove excess

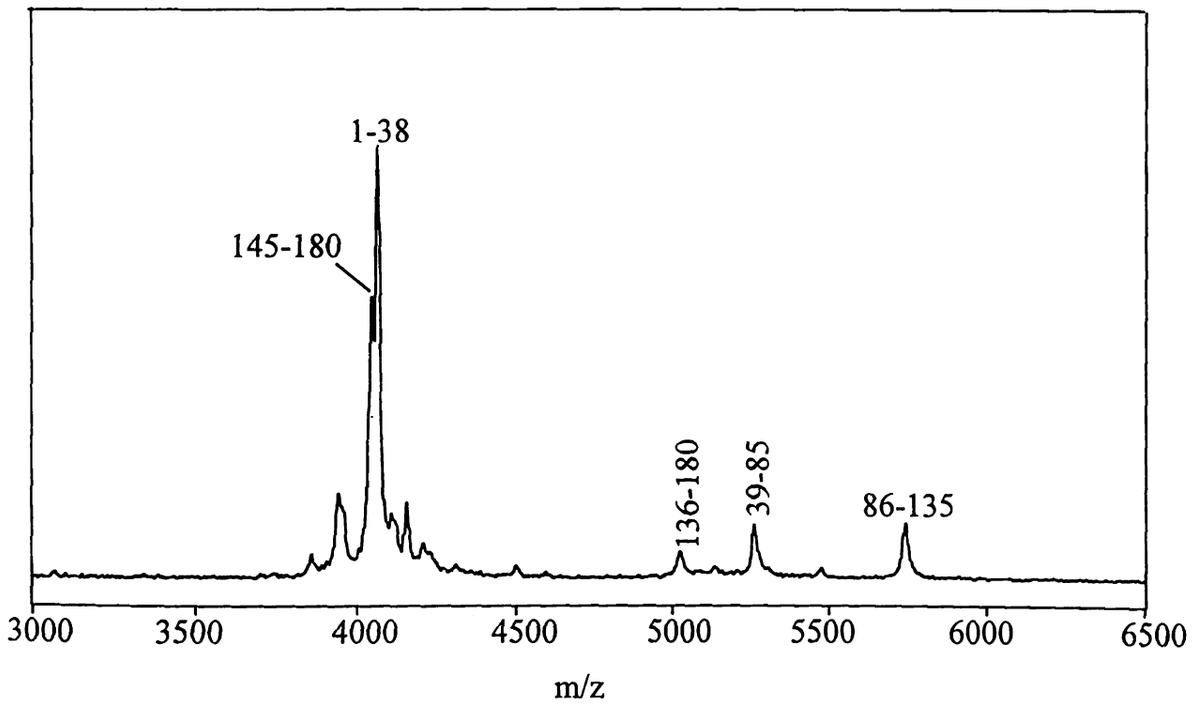


Figure 4.12. The MALDI mass spectrum of peptide mixtures resulting from the cleavage of unseparated, partially reduced/cyanylated soybean trypsin inhibitor (STI) isomers. Before the cleavage, the mixture was dialysed against 1% HAc solution to remove excess reagents. See table 4-3 for possible fragments and their m/z values.

analysis of the clea

5259.9, 10752 are

unidentified peaks

data, a Cys39-86 li

In the pres

remove excess rea

poor chromatogra

separation, instead

two disulfide bo

reduced/cyanylate

spectrum of the

fragments corresp

at both of the disu

attributable to

reduced/cyanylat

at each of the fo

would be difficul

F. Application

As we re

bond linkage by

under alkaline c

closely spaced

analysis of the cleavage products from HPLC peak 2. The fragments at m/z at 4059.1, 5259.9, 10752 are due to the fragments 1-38, 39-85 and 86-180, respectively. Other unidentified peaks are apparently due to impurities in the HPLC fraction. From these data, a Cys39-86 linkage is deduced.

In the present procedure, the cyanylated proteins are separated by HPLC to remove excess reagents and to separate the resulting isomers. For simple proteins with poor chromatographic performances, such as STI, it could be better to avoid HPLC separation, instead, use dialysis to get rid of excess reagents. Since STI only contains two disulfide bonds, the cleavage products from the mixture of the partially reduced/cyanylated isoforms were expected to be reasonably simple. The MALDI spectrum of the cleavage products from the unseparated isomers (Figure 4.12) shows fragments corresponding to the cleavage at each of the two disulfide bond sites and also at both of the disulfide bonds (see Table 4.3). For example, the fragment at m/z 5738.8 is attributable to fragment 86-135. Apparently, the mixture contains the doubly reduced/cyanylated STI, which gives rise to the fragments corresponding to the cleavage at each of the four cysteine residues. The unambiguous assignment of disulfide pairs would be difficult in such cases.

F. Application to Proteins Containing Closely Spaced or Adjacent Cysteines

As we reviewed in chapter 1, two problems remain in the recognition of disulfide bond linkage by conventional methods; one is the disulfide bond scrambling occurring under alkaline conditions and another is the failure to cleave protein chains between two closely spaced or adjacent cysteines. Our methodology can circumvent problems with

disulfide bond exch

conditions. Equall

containing closely s

the proteolytic clea

of cystine and cyan

The feasibi

containing closely

human recombina

(LR^3IGF-I), and i

Recombina

resembled its nat

initiator of protei

the disulfide bon

assignment of dis

by an Asn residu

containing only c

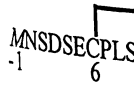


Figure 4.13.
epidermal gr

disulfide bond exchange, because the chemical reactions were performed under acidic conditions. Equally important is that our methodology is also applicable to proteins containing closely spaced or adjacent cysteines, because our approach does not depend on the proteolytic cleavage between cysteine residues, but depends on the partial reduction of cystine and cyanylation of the corresponding nascent cysteine residues.

The feasibility of our methodology to recognize disulfide linkage in proteins containing closely spaced and adjacent cysteine residues is demonstrated here using human recombinant epidermal growth factor (hEGF), LONG R^3 insulin growth factor-I (LR^3 IGF-I), and insulin as examples.

Recombinant hEGF used in this research was purchased from Sigma. It resembled its native form (53 amino acids), but with a Met at the N-terminus as an initiator of protein biosynthesis (26, 27). Figure 4.13 shows the amino acid sequence and the disulfide bond linkage in recombinant hEGF. Conventional methodologies for the assignment of disulfide pairs in hEGF is tedious because Cys31 and Cys33 are separated by an Asn residue, which is usually resistant to proteolytic digestion to give peptides containing only one disulfide bond.

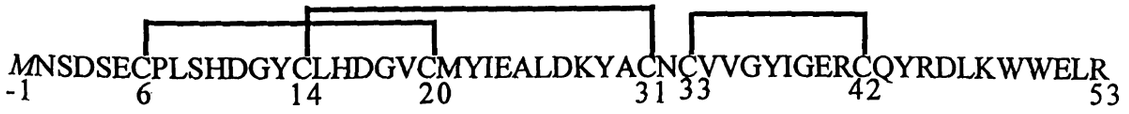


Figure 4.13. Primary structure and disulfide bond linkage of recombinant human epidermal growth factor (hEGF) (MW. 6347.2)

Table 4.4.
resulting fi
designated

Reductio of Disulfi
Cys6-Cys
Cys14-Cy
Cys33-Cy
Cys6-Cys Cys14-C
Cys6-Cy Cys33-C
Cys14- Cys33-
Cys6- Cys14 Cys33

Table 4.4. Calculated and observed m/z values for possible fragments resulting from the cleavage of recombinant hEGF chains at sites of designated cysteine pairs

Reduction of Disulfide	Fragment	Calculated m/z	Observed m/z
	-1-5	682.7	nd
	6-19	1541.6	1540.1
Cys6-Cys20	20-53	4217.9	4218.3
	6-53	5676.5(5735.5)	5734.4
	-1-13	1555.6	1554.6
Cys14-Cys31	14-30	1970.3	1969.2
	31-53	2916.3	2917.1
	-1-32	3694.1	3692.9
	33-41	1020.1	1020.2
Cys33-Cys42	42-53	1721.0	1722.0
	-1-41	4641.2	4643.1
	33-53	2664.1(2723.1)	2666.3(2725.9)
	-1-5	682.7	nd
	6-13	915.9	915.9
Cys6-Cys20	14-19	667.7	nd
Cys14-Cys31	20-30	1344.6	1345.0
	31-53	2916.3	2916.0
	14-30	1936.3	1934.7
	-1-5	682.7	nd
	6-19	1541.6	1541.1
Cys6-Cys20	20-32	1560.8	nd
Cys33-Cys42	33-41	1020.1	1020.0
	42-53	1721.0	1721.7
	33-53	2664.1(2723.1)	2664.9(2725.1)
	-1-13	1555.6	1555.2
	14-30	1970.3	1970.1
Cys14-Cys31	31-32	260.2	nd
Cys33-Cys42	33-41	1020.1	1020.6
	42-53	1721.0	1722.5
	33-53	2664.1(2723.1)	2666.9(2726.4)
	-1-5	682.7	nd
	6-13	915.9	915.8
	14-19	667.7	nd
Cys6-Cys20	20-30	1344.6	1344.3
Cys14-Cys31	31-32	260.2	nd
Cys33-Cys42	33-41	1020.1	1020.1
	42-53	1721.0	1722.0
	14-30	1936.3	1935.8
	33-53	2664.1(2723.1)	2665.9(2727.3)

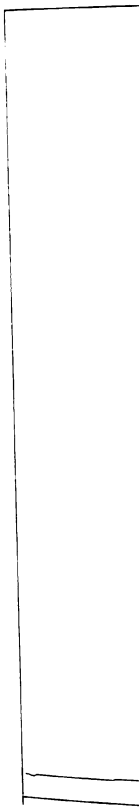


Figure 4.14.
cyanylated is
rate of 1.0 m
15-50 min, v
represent sig
analysis.

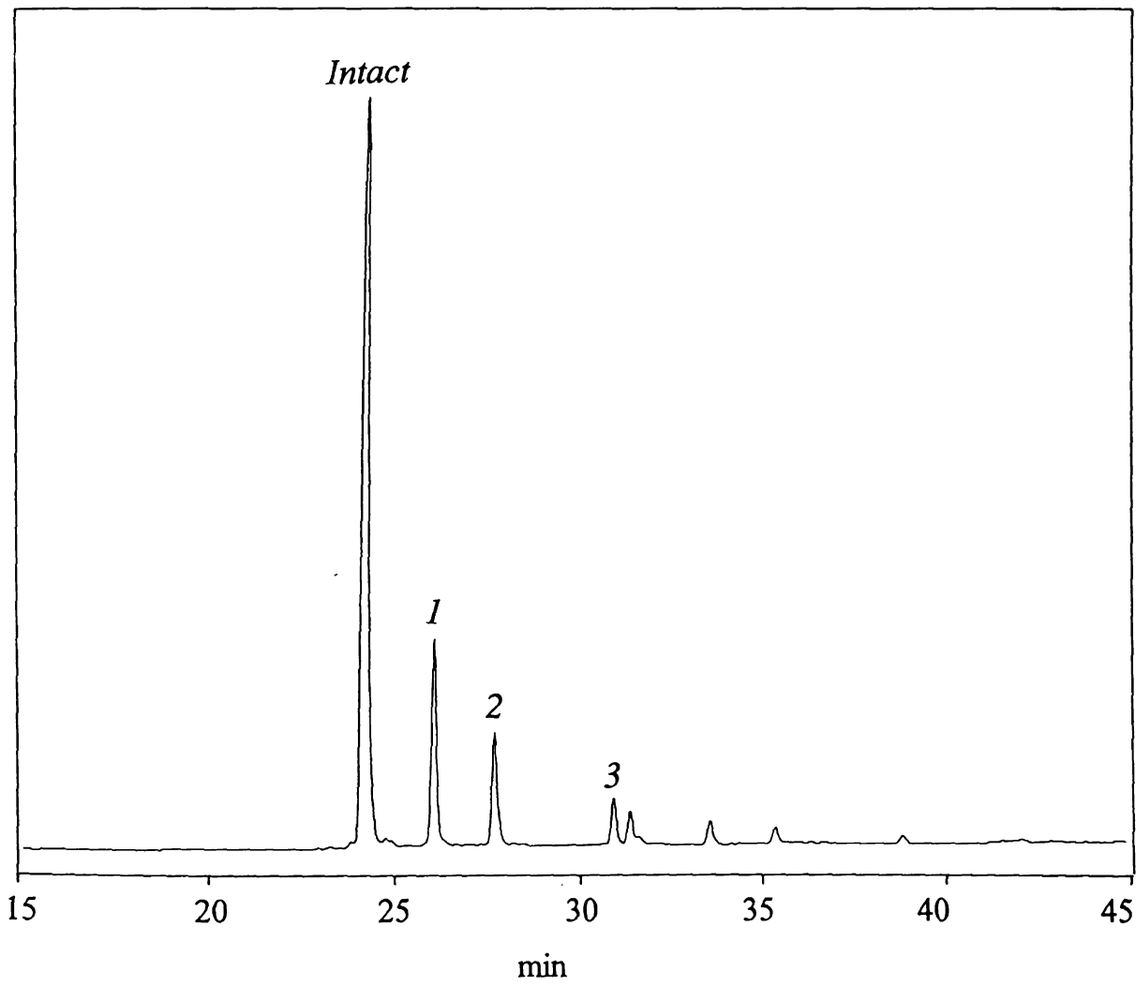


Figure 4.14. HPLC separation of denatured hEGF and its partially reduced/ cyanylated isomers. Separation was carried out on a Vydac C18 çolumn at a flow rate of 1.0 ml/min with a linear gradient 15-35% B in 15 min, and 35-55% B from 15-50 min, where A = 0.1% TFA in water and B = 0.1% TFA in CH_3CN. Peaks 1-3 represent singly reduced/cyanylated species, as determined by MALDI-TOF analysis.

After partial

obtained by HPLC

the first peak is int

doubly reduced/cy

Table 4.4 lists the

peptide chains at

reduced and cya

lactalbumin, the r

experimental co

characterized. Fi

cleavage of singl

respectively. Fig

isoforms and the

The mas

represented by H

to fragments 6

cyanylated/uncl

is not detectabl

chain cleavage

cleavage produ

41, and 42-53

4.15B). Additi

After partial reduction and cyanylation of hEGF, eight well-resolved peaks were obtained by HPLC (Figure 4.14). Mass analysis of the peaks by MALDI indicated that the first peak is intact protein, while peaks 1-3, 4-6, and 7 are singly reduced/cyanylated, doubly reduced/cyanylated, and completely reduced/cyanylated isoforms, respectively. Table 4.4 lists the expected and observed m/z values for fragments due to cleavage of the peptide chains at different cysteine sites depending on which disulfide bond(s) were reduced and cyanylated. Unlike the chromatograms of ribonuclease A and α-lactalbumin, the reduction of two disulfide bonds in hEGF was also observed under the experimental conditions and, therefore, the corresponding fractions were also characterized. Figure 4.15A-C are MALDI spectra of peptide mixtures resulting from cleavage of singly reduced/cyanylated hEGF isomers corresponding to HPLC peaks 1-3, respectively. Figure 4.16A-D are the MALDI mass spectra of doubly reduced/cyanylated isoforms and the completely reduced/cyanylated isoform, respectively.

The mass spectrum in Figure 4.15A corresponds to the cleavage products represented by HPLC peak 1 in Figure 4.14. The peaks at m/z 1540.1 and 4218.3 are due to fragments 6-19 and 20-53, respectively. The peak at m/z 5734.4 is due to a cyanylated/uncleaved fragment 6-53 with an expected m/z of 5738.5. The fragment -1-5 is not detectable by MALDI in any case. From these data, one can deduce that peptide chain cleavage occurs at Cys6 and Cys20. Therefore, Cys6 is linked to Cys20. The cleavage products of HPLC peak 2 shows m/z values attributable to fragments -1-32, 33-41, and 42-53, respectively, suggesting Cys33 and Cys42 are cleavage sites (Figure 4.15B). Additionally, two overlapped peptides, 33-53 and -1-41 confirm the conclusion.

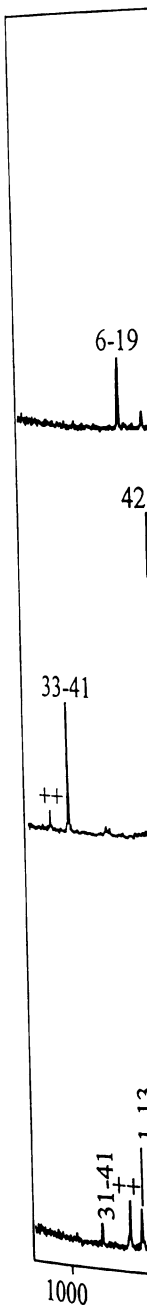


Figure 4.15.
of the three si
1-3 in Fig. 4.
species, β-eli
also Table 4.

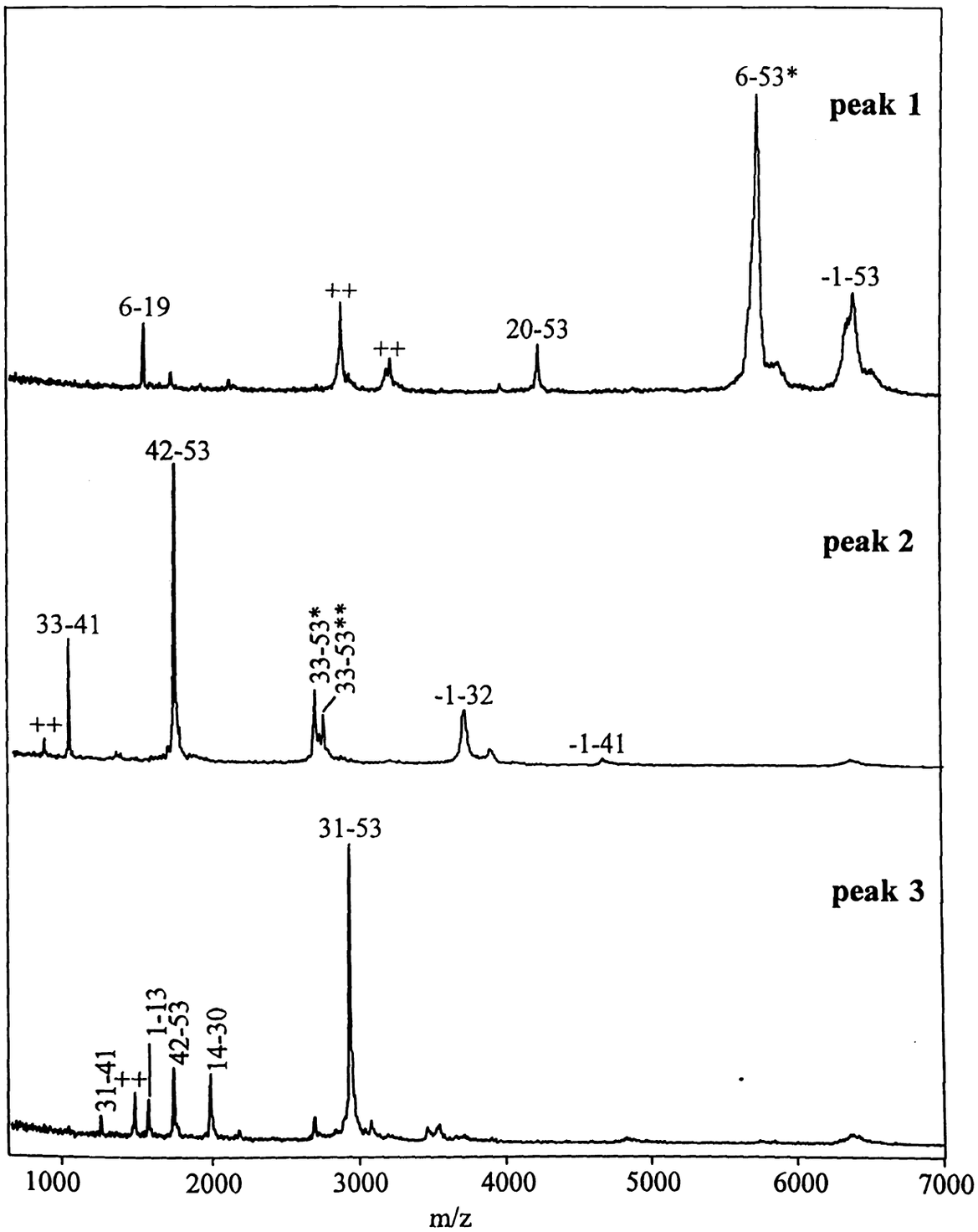


Figure 4.15. The MALDI mass spectra of peptide mixtures resulting from the cleavage of the three singly reduced/cyanylated hEGF isomers, corresponding to the HPLC peaks 1-3 in Fig. 4.14, respectively. The symbols ++ , *, and ** represent the doubly-charged species, β-elimination products, and cyanylated/uncleaved products, respectively. See also Table 4.4 for the calculated and observed m/z values.

Overall, a disulfide

The cleavage produ

4.15C). In additi

disulfide bond link

detected, correspo

5) confirm that th

sites. However, l

bring about confu

The inter

reduced/cyanylat

reduced/cyanylat

strategy is help

corresponding to

sites", that is, th

(Figure 4.16A)

fragments 6-19,

Cys42, respecti

observed. Obv

strategy, the M

the disulfide li

MALDI spectr

that cleavage c

g., fragment -1

Overall, a disulfide bond linkage between Cys33-Cys42 can be unambiguously deduced. The cleavage products from HPLC peak 3 present a little more complicated case (Figure 4.15C). In addition to the desired fragment -1-13, 14-30, and 31-53 implicating the disulfide bond linkage between Cys14-Cys31, two fragments, 31-41 and 42-53, were also detected, corresponding to the undesired cleavage at Cys42. Further experiments (chapter 5) confirm that the cleavage at Cys42 is minor in comparison with the cleavage at other sites. However, because of the high responses on MALDI, these two fragments could bring about confusion in data interpretation.

The interpretation of MALDI data from the cleavage products of doubly reduced/cyanylated isoforms is never as straightforward as that from singly reduced/cyanylated isomers, because more fragments are observed. A slightly different strategy is helpful for the assignment. Instead of looking at the MALDI peaks corresponding to the cleavage sites, the absence of such peaks represents the "uncleaved sites", that is, those still linked by a disulfide bond. For example, the MALDI spectrum (Figure 4.16A) of the cleavage products from HPLC peak 4 shows peaks due to fragments 6-19, 33-41, and 42-53, indicating the cleavage at Cys6, Cys20, Cys33, and Cys42, respectively. No fragments due to the cleavage at Cys14 and Cys31 can be observed. Obviously, these two cysteines are linked as a disulfide pair. With the same strategy, the MALDI spectra of the cleavage products from HPLC peaks 5 and 6 reveal the disulfide linkage Cys33-Cys42 and Cys6-Cys20, respectively. The analysis of the MALDI spectrum of the cleavage products from the completely cyanylated hEGF shows that cleavage occurs at each of the cysteine positions, although some of the fragments (e. g., fragment -1-5) might be missing on MALDI.

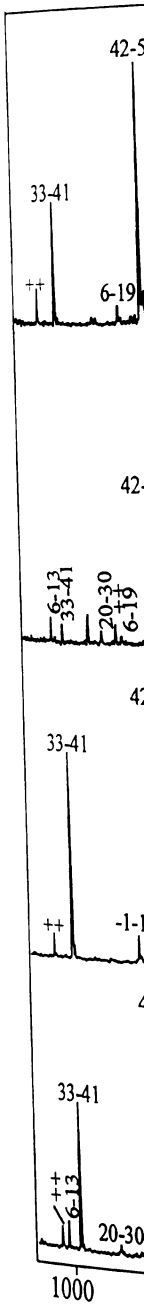


Figure 4.16.
of the three
cyanylated h
The symbols
and cyanylat
and observe

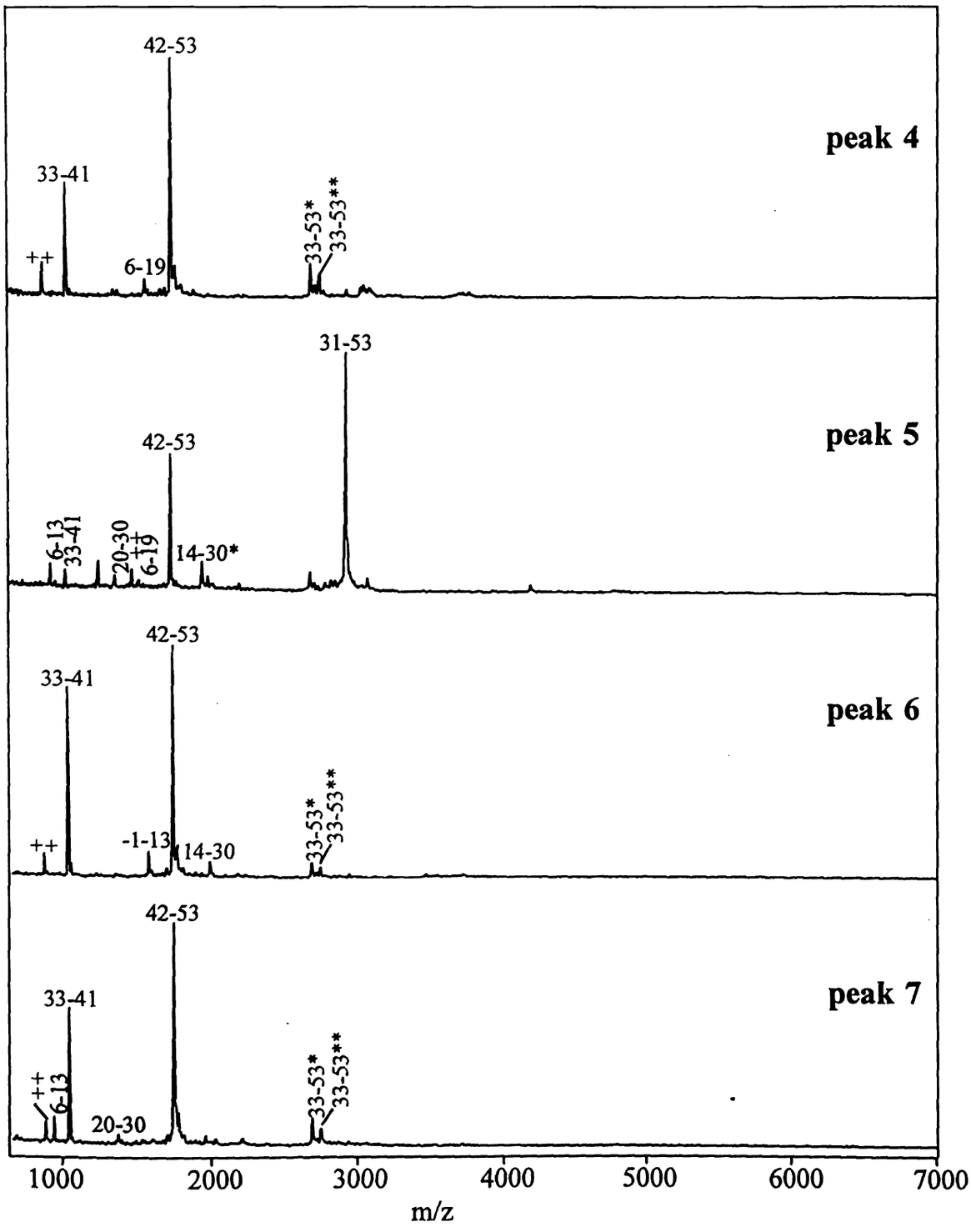


Figure 4.16. The MALDI mass spectra of peptide mixtures resulting from the cleavage of the three doubly reduced/cyanylated hEGF isomers and a completely reduced/ cyanylated hEGF,corresponding to the HPLC peaks 4-7 in Fig. 4-14, respectively. The symbols ++ , *, and ** represent doubly-charged species, β-elimination products, and cyanylated/uncleaved products, respectively. See also Table 4.4 for the calculated and observed m/z values.

Recombina

replacing glutamat

LR^{3}IGF-I shows hi

contains adjacent

scheme, assigned

below, has never

recombinant sour

as for IGF-I isola

have a major infl

bonded isomers c

1

Figure 4

reduced/cyanyla

with the typica

First, the major

the singly redu

become smalle

most of the p

Second, only t

2 in Figure 4.1

Recombinant LR^{3}IGF-I is a variant of human IGF-I that contains arginine replacing glutamate-3 as well as an amino terminal extension of 13 amino acids (28). LR^{3}IGF-I shows higher biological activities than its analog, IGF-I. Like IGF-I, LR^{3}IGF-I contains adjacent cysteine residues at 60 and 61 positions. The disulfide bonding scheme, assigned on the basis of homology to the insulin (or IGF-I) sequence and shown below, has never been verified experimentally. In preparing authentic LR^{3}IGF-I from recombinant sources, it is important to confirm that the disulfide bond linkage is the same as for IGF-I isolated from natural sources, since a mismatching of disulfide bonds could have a major influence on any biological activity, as has been observed with disulfide-bonded isomers of insulin (29).

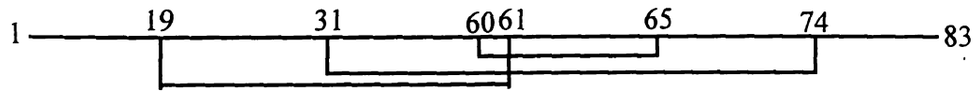


Figure 4.17 shows the HPLC separation of LR^{3}IGF-I and isomers of its partially reduced/cyanylated species. The chromatogram shows two exceptions in comparison with the typical chromatographic pattern of the partially reduced/cyanylated isoforms. First, the major peak representing the intact protein shows longer retention than one of the singly reduced isoforms. This result indicates that the hydrophobicity of the protein become smaller than that of the intact protein after opening one disulfide bond, whereas most of the proteins show larger hydrophobicity after splitting the disulfide bond(s). Second, only two singly reduced/cyanylated protein isomers were observed (peaks 1 and 2 in Figure 4.17), even though the protein contains three disulfide bonds. Stronger

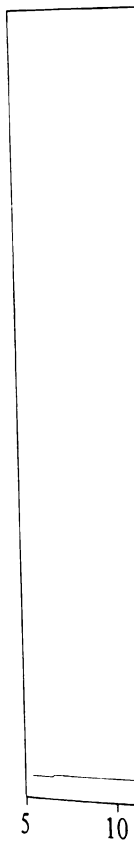


Figure 4.17
cyanylated
rate of 1.0
TFA in wa
reduced/cy

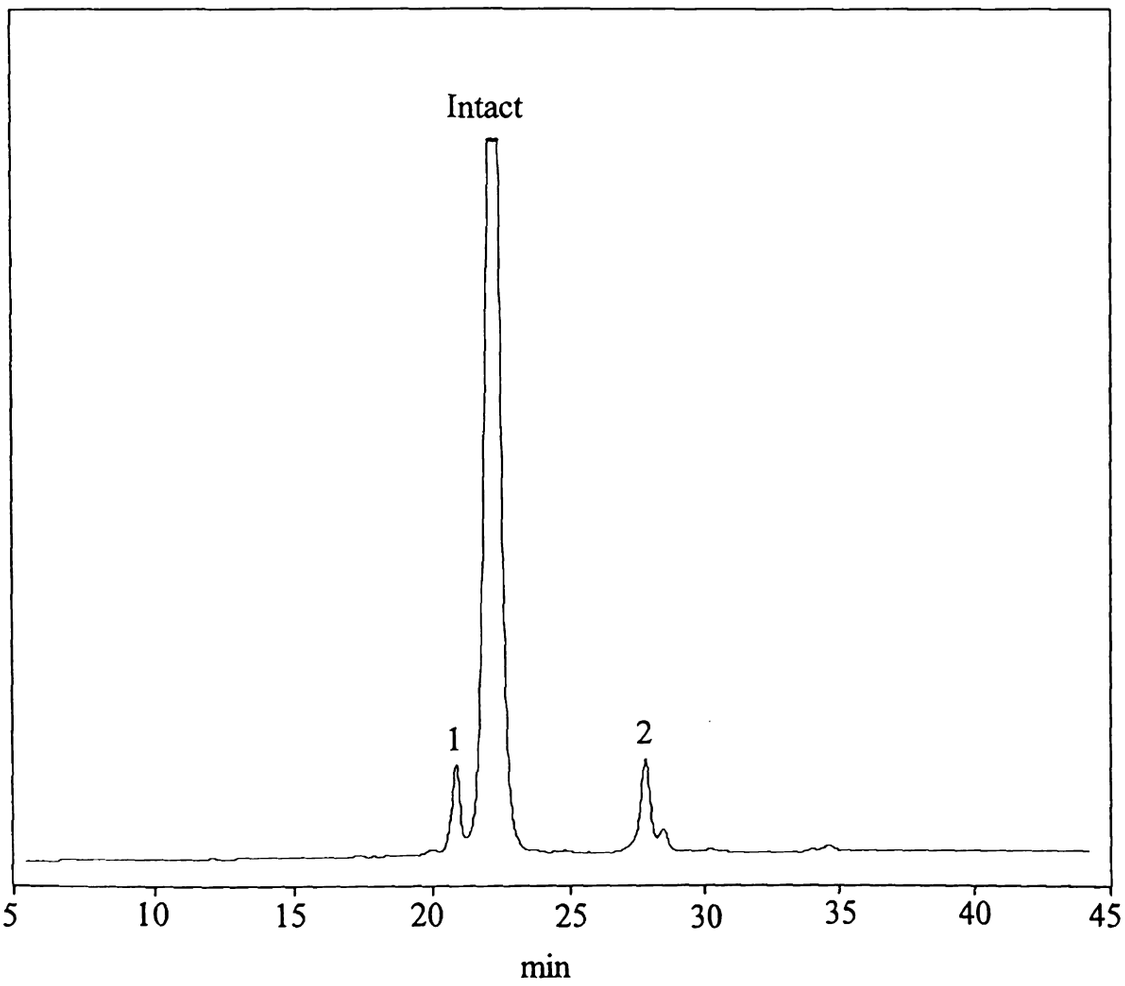


Figure 4.17. HPLC separation of denatured LR^{3}IGF-I and its partially reduced/ cyanylated isomers. Separation was carried out on a Vydac C18 column at a flow rate of 1.0 ml/min with a linear gradient 30-50% B in 45 minutes, where A = 0.1% TFA in water and B = 0.1% TFA in CH_3CN. Peaks 1 and 2 represent singly reduced/cyanylated species, as determined by MALDI-TOF analysis.

reducing condition

prolong the reduct

isomers, but the th

isomer from the

retention time as

MALDI-MS of th

disulfide bond (C

under the reducti

understandable f

stability of IGF-I

in LR^{3}IGF-I) is

might explain th

Table 4.

cleavage reactio

The two HPLC

under the descr

Figure 4.18A a

fragments 65-8

the cleavage re

m/z 2752.9 an

where β-elimi

cleavage sites.

reducing conditions (increase the ratio of reducing agent, apply higher temperature, and prolong the reduction time) result in the formation of doubly reduced/cyanylated protein isomers, but the third disulfide bond still refuses to reduce. At first we assume that the isomer from the reduction and cyanylation of the third disulfide bond had the same retention time as its intact protein and, therefore, coelutes with the intact one. But MALDI-MS of the cleavage products indicates that it is not the case. Apparently one disulfide bond (Cys31-Cys74, as concluded from the next paragraph) is inert to reduction under the reduction conditions, even if the protein has been denatured. This conclusion is understandable from the thermodynamic point of view. A recent study on the disulfide stability of IGF-I showed that the disulfide bond 18-61 (corresponding to disulfide 31-74 in LR^3IGF-I) is superstable and preferentially formed in the folding process (30), which might explain that this disulfide pair is structurally stable enough to resist the reduction.

Table 4.5 lists the calculated m/z values for possible fragments resulting from the cleavage reaction of LR^3IGF-I chains at sites corresponding to different cysteine pairs. The two HPLC fractions (HPLC peaks 1 and 2 in Figure 4.17) were subjected to cleavage under the described conditions. The corresponding MALDI mass spectra are shown in Figure 4.18A and B. The peaks at m/z 2188.7 and 6372.9 in Figure 4.18A are due to fragments 65-83 and 1-59, respectively. The fragment 60-64, a middle fragment during the cleavage reactions, was missing from the MALDI spectrum. However, the peaks at m/z 2752.9 and 6936.5, representing two overlapped peptide fragments 60-83 and 1-64, where β-elimination occurs at Cys65 and Cys60, respectively, confirm that 60 and 65 are cleavage sites. From these data, one can deduce that Cys60 is linked to Cys65. The

Table 4.5
resulting
designat

| Reduc
of Dis |
| --- |
| Cys60 |
| Cys1 |
| Cys |

Table 4.5. Calculated and observed m/z values for possible fragments resulting from the cleavage reaction of LR^3IGF-I chains at sites of designated cysteine pairs

Reduction of Disulfide	Fragment	Calculated m/z	Observed m/z
Cys60-Cys65	1-59	6371.2	6372.9
	60-64	637.7	nd
	65-83	2188.6	2188.7
	β(1-64)	6931.9	6936.5
	β(60-83)	2749.3	2752.9
Cys19-Cys61	1-18	1978.4	1976.9
	19-60	4538.0	4538.6
	61-83	2681.1	2682.5
	β(1-60)	6439.4	6441.9
	β(19-83)	7142.1	7139.1
Cys31-Cys74	1-30	3223.8	
	31-73	4964.5	
	74-83	1009.2	no reduction
	β(1-73)	8111.3	
	β(31-83)	5897.7	

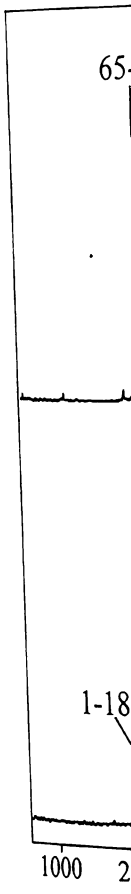


Figure 4.18.
of the two si
peaks 1 and
β-eliminatio
also Table 4

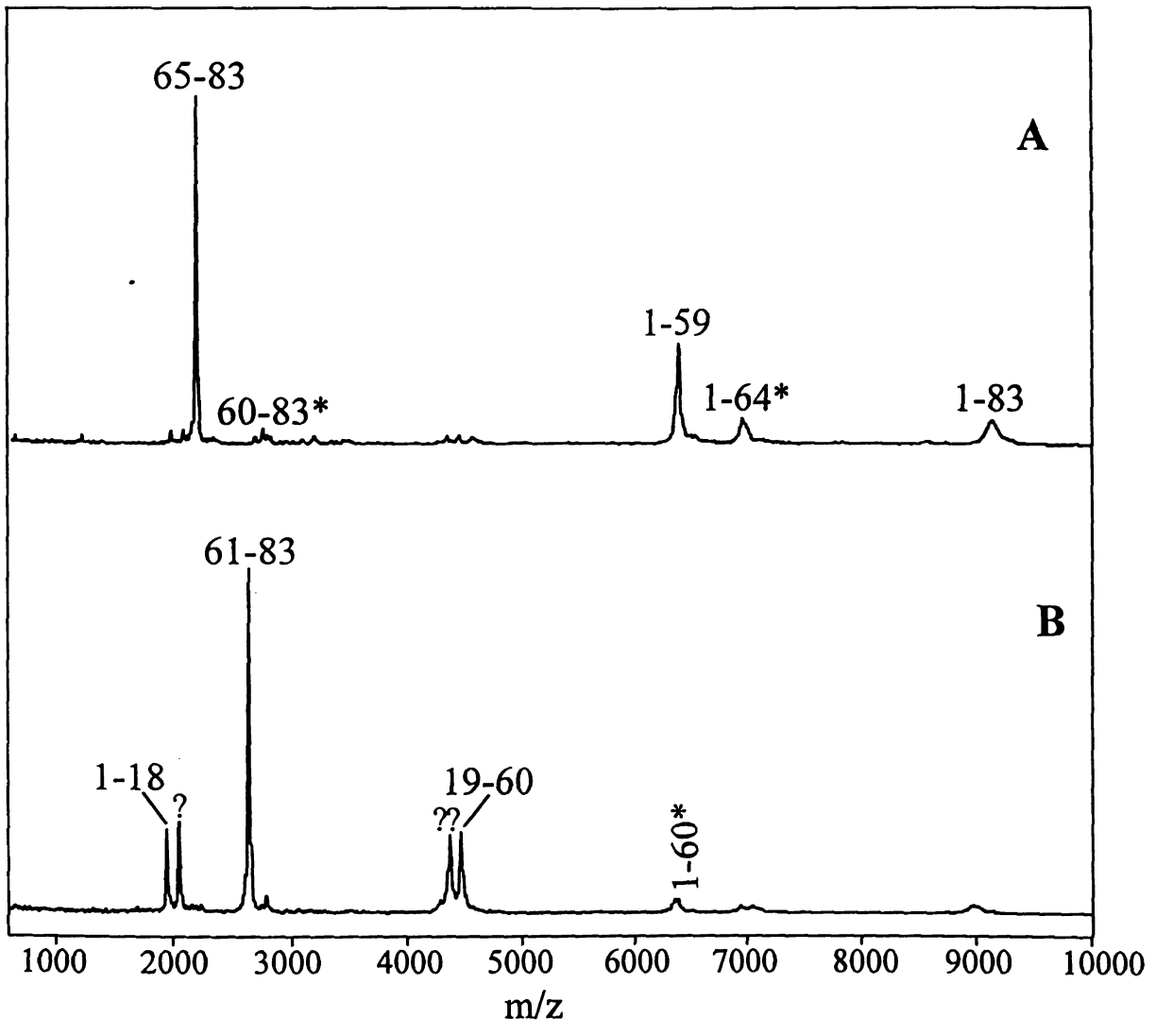


Figure 4.18. The MALDI mass spectra of peptide mixtures resulting from the cleavage of the two singly reduced/cyanylated LR^3IGF-I isomers, corresponding to the HPLC peaks 1 and 2 in Figure 4-17, respectively. The symbol * represent the protonated β-elimination products. The peaks with question marks are discussed in the text. See also Table 4.5 for the calculated and observed m/z values.

cleavage products
60, and 61-83, re
4.18B). Addition
linked to Cys61.
can be easily assi
digestion. As in
assume adjacent
disulfide structur
IGF-I and IGF-
shown). It is als
to study the refo

It should
confusing in na
103 Da from th
the peak with
fragment 19-60
caused by the
exchange betw
unexplainable,
Cys-Gly- struc
than disulfide
other hand, u

cleavage products from HPLC peak 2 show m/z values attributable to fragments 1-18, 19-60, and 61-83, respectively, suggesting Cys19 and Cys61 are cleavage sites (Figure 4.18B). Additionally, an overlapped peptide, 1-60, confirm the conclusion that Cys19 is linked to Cys61. Overall, the disulfide bond structure containing two adjacent cysteines can be easily assigned without multiple enzymatic digestions and/or multiple step Edman digestion. As insulin growth factor contains a number of protein analogs which all assume adjacent cysteine structures, the data presented here are of great importance to the disulfide structure study of such a protease. As evidence, preliminary experiments on IGF-I and IGF-II, both containing adjacent cysteines, show similar results (data not shown). It is also expected that the approach described here will be sufficiently powerful to study the refolding intermediates of the IGF analogs (30).

It should be pointed out that the question-marked peaks in Figure 4.18B are very confusing in nature. The peak with a single question mark showed a mass increase of 103 Da from the fragment 1-18, suggesting a cysteine attached to the fragment, whereas the peak with double question marks showed a mass decrease of 103 Da from the fragment 19-60, suggesting a cysteine removed from the fragment. The latter might be caused by the cleavage at Cys59, rather than Cys60, assuming that the disulfide bond exchange between two adjacent cysteine residues occurs. The former one is unexplainable, because such an exchange can never occur at Cys19, which has the -Lys-Cys-Gly- structure. Obviously structural heterogeneity of recombinant proteins, other than disulfide exchange, would be critical to wield occurrence of such fragments. On the other hand, unambiguous assignment of the disulfide structure is evident even if such

undesired fragme

native disulfide b

Insulin co

intrachain (Cys A

residues (Cys A

cyanylated insul

A ch

B ch

The ch

disulfide bond

preferred. Ou

partially reduc

reduced/cyany

disulfide bond

A prol

gives a good

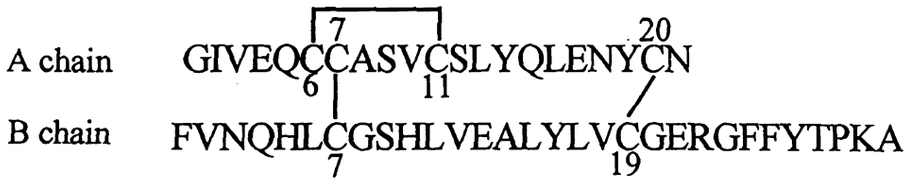
mode, while

expected and

for fragments

undesired fragments are present, because the fragments from specific cleavage at the native disulfide bond are sufficient to make a positive conclusion.

Insulin contains two interchain (Cys A7-Cys B7 and Cys A20-Cys B19) and one intrachain (Cys A6-Cys A11) disulfide bridges, and includes a pair of adjacent cysteine residues (Cys A6, Cys A7). The HPLC chromatogram of the partially reduced/ cyanylated insulin isoforms is shown in Figure 4.19.



The chromatogram shown by Gray (5) indicates that the reduction rates of the disulfide bonds in native insulin vary widely and that one route of reduction was preferred. Our chromatogram (Figure 4.19) of denatured insulin and isomers of its partially reduced/cyanylated form shows a peak for each of the three possible singly reduced/cyanylated species. Our data indicate that under denaturing conditions, the three disulfide bonds undergo reduction at comparable rates.

A problem with the analysis of insulin fragments by MALDI is that its A-chain gives a good signal in the negative ion mode, but a very weak signal in the positive ion mode, while the B-chain gives a good response in both modes (31). Table 4.6 shows the expected and observed m/z values (deprotonated molecules detected in negative mode) for fragments resulting from the cleavage reaction of insulin chains at sites corresponding

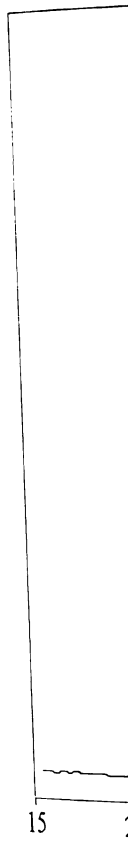


Figure 4.19
cyanylated
rate of 1.0
TFA in wa
reduced/cy
marked A

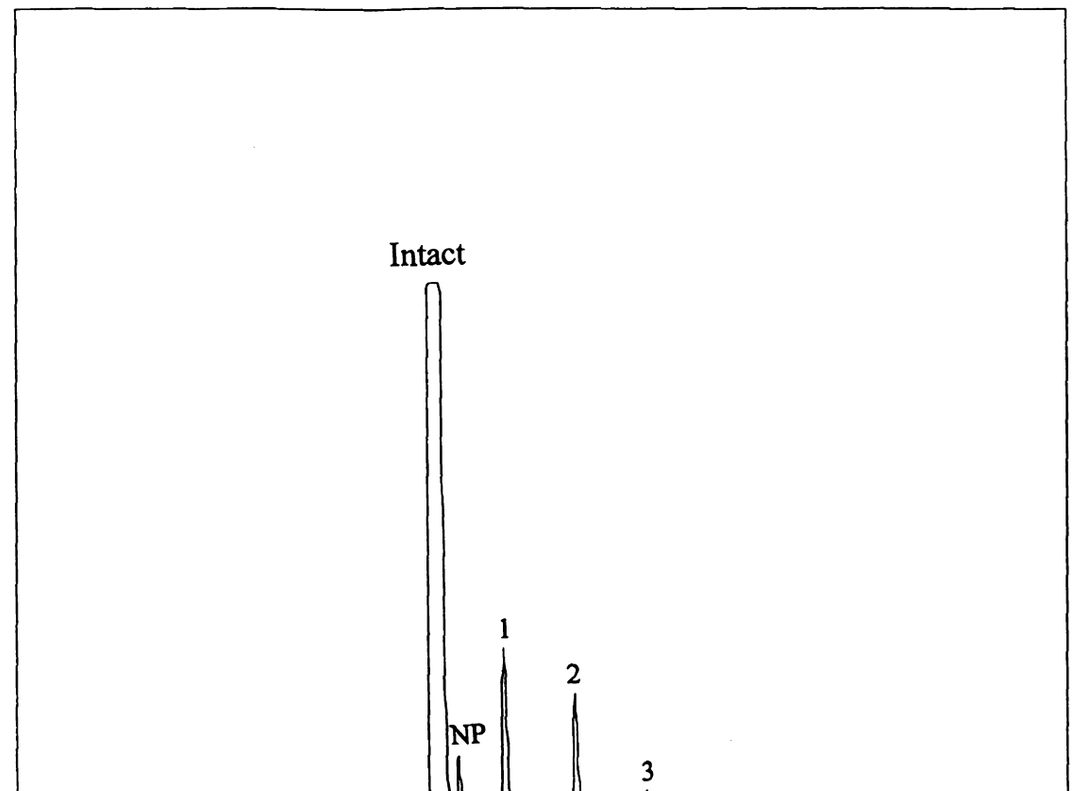


Figure 4.19. HPLC separation of denatured insulin and its partially reduced/ cyanylated isomers. Separation was carried out on a Vydac C18 column at a flow rate of 1.0 ml/min with a linear gradient 20-50% B in 40 minutes, where A = 0.1% TFA in water and B = 0.1% TFA in CH_3CN. Peaks 1-3 represent singly reduced/cyanylated species, as determined by MALDI-TOF analysis. The peaks marked A and B are insulin A-chain and B-chain, respectively.

Table 4.6
fragment
of design

| Reduc
of Dis |
| --- |
| CysA6- |
| CysA7- |
| CysA2 |

Table 4.6. Calculated and observed m/z values $[M-H]^-$ for possible fragments resulting from the cleavage reaction of insulin chains at sites of designated cysteine pairs

Reduction of Disulfide	Fragment	Calculated m/z	Observed m/z
CysA6-CysA11	A1-5	543.6	nd
	A6-10	505.6	nd
	A11-21	1373.8	1374.2
	A6-21	1803.4	1804.8
	A1-21	2340.0	2339.6
	B1-30	3399.9	3400.3
CysA7-CysB7	A1-6	646.7	nd
	A7-21	1734.9	1735.7
	B1-6	755.9	nd
	B7-30	2685.1	2686.3
CysA20-CysB19	A1-19	2123.7	2122.6
	A20-21	261.3	nd
	B1-18	2043.4	2041.5
	B19-30	1401.5	1402.3
	B1-30	3400.0 (3366.0)	3400.1 (3366.1)

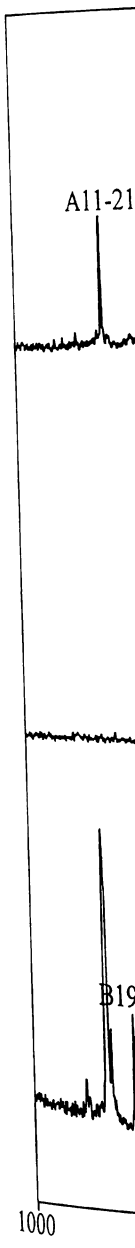


Figure 4.20.
the cleavage
HPLC peak
β-eliminatio
Table 4.6 fo

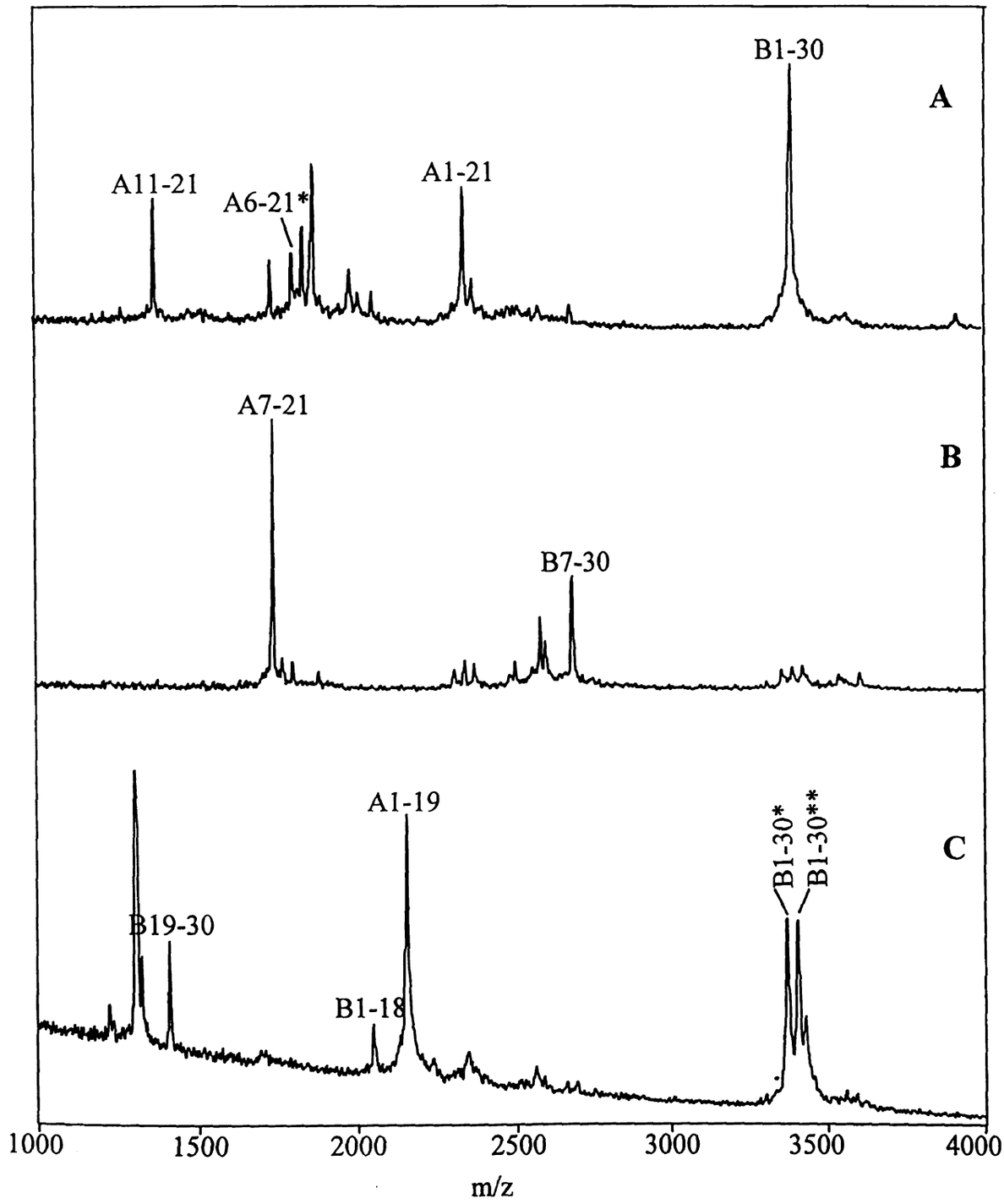


Figure 4.20. The negative ion MALDI mass spectra of peptide mixtures resulting from the cleavage of the three singly reduced/cyanylated insulin isomers, corresponding to the HPLC peaks 1-3 in Figure 4.19, respectively. The symbols * and ** represent the β-elimination products and cyanylated/uncleaved peptides, respectively. See also Table 4.6 for the calculated and observed m/z values.

to different cystei

mass spectra. Fig

in Figure 4.19) sl

resulting in the fi

A1-5 and A6-10

plagued by the

cleavage. Howe

of the adjacent

4.19 show simil

corresponding t

1735.7 is attrib

of the insulin

specific fragme

described here

proteins contai

A1-19, B 1-18

linked to CysB

V. Picomole

So far,

by the detect

methodologie

disulfide bon

to different cysteine pairs. Figures 4.20A-C are the corresponding negative ion MALDI mass spectra. Figure 4.20A (corresponding to the cleavage products of the HPLC peak 1 in Figure 4.19) shows that chemical cleavage occurs at Cys6 and Cys11 of the A chain, resulting in the fragment A11-21 and A6-21. Two low-mass fragments corresponding to A1-5 and A6-10 are not readily detected by MALDI. Furthermore, the mass spectrum is plagued by the unidentified impurity peaks, probably resulting from the nonspecific cleavage. However, it is still convincing to conclude that cleavage indeed occurred at one of the adjacent cysteine sites, CysA6. The cleavage products from HPLC 2 in Figure 4.19 show similar results (Figure 4.20B). In addition to an expected peak at m/z 2686.3, corresponding to fragment B7-30 (calculated m/z 2685.1), another intense peak at m/z 1735.7 is attributed to fragment A7-21 (calculated m/z 1734.9), resulting from cleavage of the insulin A-chain at Cys A7 which is adjacently linked to Cys A6. These two specific fragments suggest that Cys A7 be linked to Cys B7. Therefore, the methodology described here shows the potential for the assignment of disulfide bond pairings in proteins containing close or adjacent cysteine residues. Figure 4.20C gives fragments A1-19, B 1-18, B19-30, plus an overlapped peptide B1-30. It is obvious that CysA20 is linked to CysB19.

V. Picomole Scale Reactions

So far, the sample size used in our protocol is ~10 nmol. The sensitivity is limited by the detection limit of our HPLC systems. Although it is better than conventional methodologies, our methodology described here should have the potential to assign disulfide bonds at the picomole level of proteins. One improvement in the methodology

Figure 4.
species b
same as t

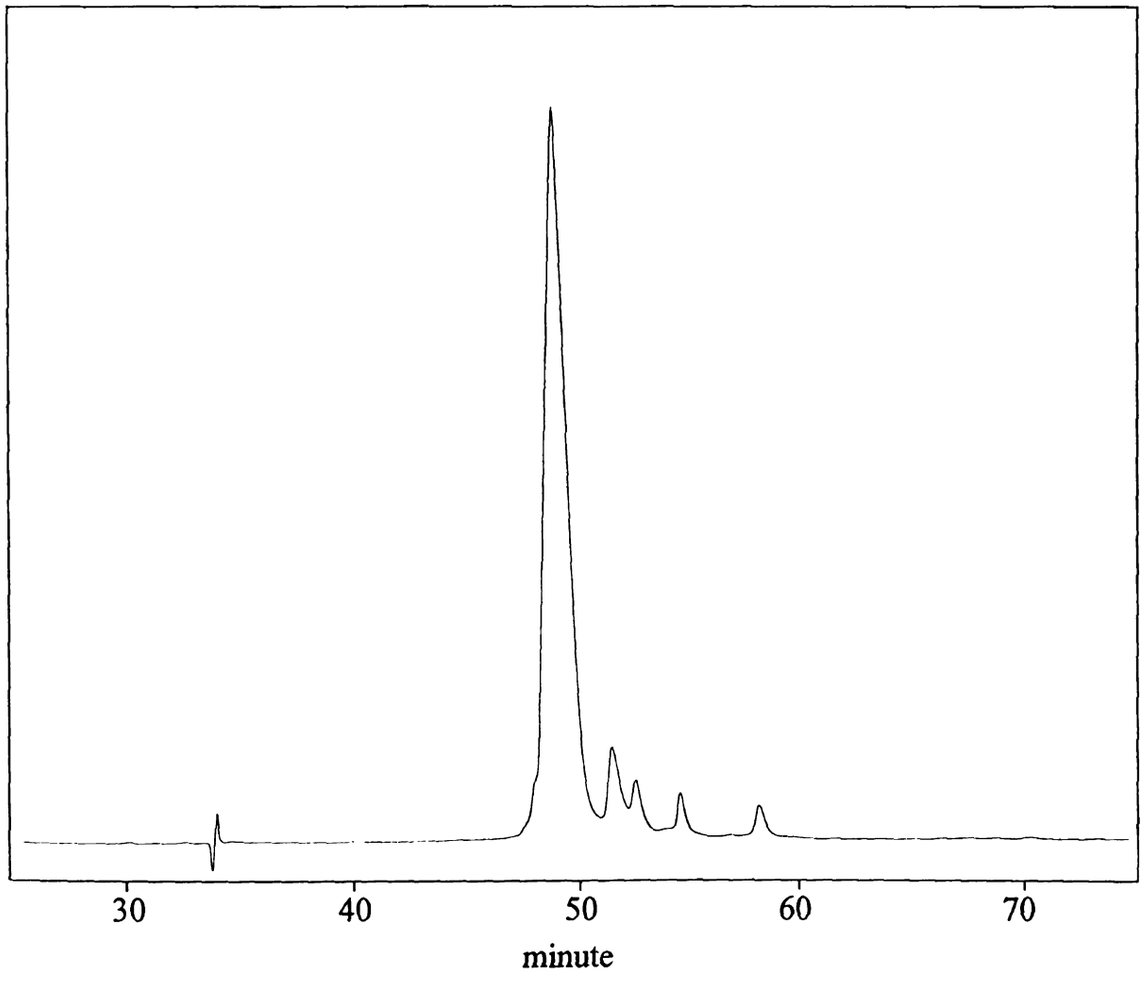


Figure 4.21. separation of ribonuclease A and its partially reduced/cyanylated species by microbore C18 HPLC column (1×100mm). HPLC conditions are the same as those in Figure 4.4, except the flow rate is 0.05 ml/min.

is to use a microbo

microscale is a ch

Twice as much th

The CDAP conc

used in protocol

the lower conce

reduced/cyanylat

mm C18 microb

was obtained o

preliminary resu

samples.

VI. Character

The ana

facets of simpl

the case of mul

only one of th

cysteines from

masses of the

reduced isofo

itself to comp

elimination p

is to use a microbore column HPLC to separate the protein isoforms. To run chemistry at microscale is a challenge to our protocol. Appropriate modifications have to be made. Twice as much the TCEP reducing agent had to be used to promote the partial reduction. The CDAP concentration for the subsequent cyanylation is five times as much as that used in protocol to prevent the autohydrolysis of CDAP reagent, which is accelerated in the lower concentration medium. Figure 4.21 shows the separation of the partially reduced/cyanylated isoforms of 500 picomoles of ribonuclease A obtained on a 1×100 mm C18 microbore column (thanks to the generosity of Dr. Gage). A better separation was obtained on the microbore column, while the retention time is shorter. These preliminary results indicate that our chemistry can be used at microscale level for protein samples.

VI. Characteristics of the Methodology

The analytical advantage of our proposed methodology accrues from several facets of simplicity. First, the optimized procedure for partial reduction assures that in the case of multiple disulfides there will be significant quantities of all isoforms in which only one of the cystines has been reduced. Second, cleavage of a cyanylated pair of cysteines from a newly reduced cystine yields only three fragments. Third, the sum of masses of the three cleavage fragments is equal to the mass of the cyanylated singly reduced isoform plus the mass of two molecules of water, a feature that readily lends itself to computerized assessment of the experimentally determined data. Fourth, β-elimination products are usually observed which confirms the assignment of cleavage

data by peaks tha

sites. This is an

consistency of d

likelihood that a

is quite low, eve

resolvable by M

reduced and cya

results reported

is the risk of dis

Because our e

scrambling are

evidence that

indeed observe

anticipated fro

Signal

MALDI is stil

does make ass

those shown i

electrospray

problem. Fu

which may b

batch analysi

data by peaks that represent sequences of residues that overlap one of the two cleavage sites. This is another feature that lends itself to computerized assessment of the internal consistency of data interpretation. Fifth, for relatively large proteins (> 10 kDa) the likelihood that alternative disulfide pairings would lead to the same mass spectral pattern is quite low, even if some mass spectral peaks for two alternative pairings may not be resolvable by MALDI. Sixth, complete HPLC separation of the isomers of the singly reduced and cyanylated species is unnecessary for identification as demonstrated in our results reported here. Seventh, one of the severe problems in conventional methodology is the risk of disulfide bond scrambling which usually occurs in an alkaline medium (1-3). Because our experiments are performed under acidic conditions, opportunities for scrambling are kinetically suppressed. Complete analysis of the MALDI data gives no evidence that scrambling occurred during our procedure. Some unexpected peaks are indeed observed in our spectra, but their masses do not appear to be related to fragments anticipated from scrambled disulfide bonds.

Signal suppression for some components of a complex mixture in the analysis by MALDI is still a problem (7). The suppression of signal from certain peptide fragments does make assignments risky when they are represented by minor peaks such as some of those shown in Figure 6. The complementary nature of analytical results by MALDI and electrospray (ESI) (32) would seem to provide one promising way to deal with this problem. Furthermore, ESI provides ready access for analysis of mixtures by LC-MS, which may be an effective alternative to problems of analyte suppression encountered in batch analysis.

VII. Conclusion

We have

chemical modifi

subsequent mas

alternative meth

known primary

minimizing disu

Furthermore, th

cysteines for w

sensitive. We

ionization to th

biological sour

VII. Conclusions

We have demonstrated that the combination of partial reduction of a protein, chemical modification of the sulfhydryl groups, cleavage of the peptide chain, and subsequent mass-mapping by MALDI-TOF MS provides a simple and effective alternative methodology for the assignment of disulfide bond pairings in proteins of known primary structure. Our novel approach offers an important advantage of minimizing disulfide bond scrambling, a concern in most conventional methodologies. Furthermore, this approach may be applied to proteins containing close or adjacent cysteines for which conventional approaches fail. The procedure is simple, fast, and sensitive. We are currently applying this methodology in combination with electrospray ionization to the characterization of disulfide bond pairings in proteins from a variety of biological sources.

VIII. Reference

1. Hirayama, K

(Matsuo T, C

p299-312.

2. Smith, D. L.

3. Ryle, A. P.,

4. Zhou, Z., an

5. Gray, W. R.

6. Gray, W. R

7. Wu ,J., Gag

8. Jacobson, C

248, 6583-

9. Wu, J., and

10. Wakselma

11. Shapira, E

12. Creighton

13. Kuwajima

29, 8240-

14. Reeve, J.

15. Takeda, F

679-684(

16. Nakagaw

Nishimu

VIII. References

1. Hirayama, K., and Akashi, S.,. In *Biological mass spectrometry: present and future*, (Matsuo T, Caprioli RM, Gross ML, Seyama Y, eds.), New Nork, Wiley, (1994), p299-312.
2. Smith, D. L., and Zhou, Z., *Methods Enzymol* **193**, 374-389(1990).
3. Ryle, A. P., and Sanger, F., *Biochem. J.*, **60**, 535-540(1955).
4. Zhou, Z., and Smith, D. L., *J. Protein Chem.* **9**, 523-532(1990).
5. Gray, W. R., *Protein Sci.*, **2**, 1732-1748(1993).
6. Gray, W. R., *Protein Sci.*, **2**, 1749-1755(1993).
7. Wu ,J., Gage, D. A., and Watson, J. T., *Anal. Biochem.*, **235**, 161-174(1996).
8. Jacobson, G. R., Schaffer, M.H., Stark, G. R., and Vanaman, T. C., *J. Biol. Chem.*, **248**, 6583-6591(1973).
9. Wu, J., and Watson, J. T., *Protein Sci.*, **6**, 391-398(1997).
10. Wakselman, M., and Guibe-Jampel, E., *J. C. S. Chem. Comm.*, 21-22(1976).
11. Shapira, E., and Arnon, R., *J. Biol. Chem.*, **244**(4), 1026-1032(1969).
12. Creighton, T. E., *Prog. Biophys. Mol. Biol.*, **33**, 231-297(1978).
13. Kuwajima, K., Ikeguchi, M., Sugawara, T., Hiraoka, Y., and Sugai, S., *Biochemistry*, **29**, 8240-8249(1990).
14. Reeve, J. R., and Pierce, J. G., *Int. J. Pept. Protein Res.*, **18**, 79(1981).
15. Takeda, K., Ogawa, K., Ohara, M., Hamada, S., Moriyama, Y., *J. Protein Chem.*, **14**, 679-684(1995).
16. Nakagawa, S., Tamakashi, Y., Hamana, T., Kawase, M., Taketomi, S., Ishibashi, Y., Nishimura, O., and Fukuda, T., *J. Am. Chem. Soc.*, **116**, 5513-5514(1994).

17. Nakagawa, S.
O., and Fukuc
18. Regnier, F. E
19. Rubinstein, N
20. Degani, Y., a
21. Stark, G. R.,
22. Lu, H. S., Gr
23. Papayannop
24. Smyth, D. C
25. Brew, K., C
4570-4582(
26. Heck, D. E
27. Chang, J-Y
9207-9216
28. Tomas, F.
Ballard, F.
29. Rinderkne
30. Hober, S.,
1749-175(
31. Zaluzec, I
5, 230-23
32. Siuzdak,

17. Nakagawa, S., Tamakashi, Y., Ishibashi, Y., Kawase, M., Taketomi, S., Nishimura, O., and Fukuda, T., *Biochem. Biophys. Res. Commun.*, **200**, 1735-1741(1994).

18. Regnier, F. E., and Gooding, K. M., *Anal. Biochem.*, **103**, 1-25(1980).

19. Rubinstein, M., *Anal. Biochem.*, **98**, 1-7(1979).

20. Degani, Y., and Patchornik, A., *Biochemistry* **13**, 1-11(1974).

21. Stark, G. R., *Methods Enzymol.*, **47**, 129-132(1977).

22. Lu, H. S., Gracy, R. W., *Arch. Biochem. Biophys.*, **212**, 347-359(1981).

23. Papayannopoulos, I. A., and Biemann, K., *Protein Sci.*, **1**, 278-288(1992).

24. Smyth, D. G., Stein, W. H., and Moore, S., *J. Biol. Chem.*, **238**, 227-234(1963).

25. Brew, K., Castellino, F. J., Vanaman, T. C., and Hill, R. L., *J. Biol. Chem.*, **245**, 4570-4582(1970).

26. Heck, D. E., *J. Biol. Chem.*, **267**, 21277(1992).

27. Chang, J-Y., Schindler, P., Ramseier, U., and Lai, P-H., *J. Biol. Chem.*, **270**(16), 9207-9216(1995).

28. Tomas, F. M., Knowles, S. E., Chandler, C. S., Francis, G. L., Owen, P. C., and Ballard, F. J., *J. Endocrinology*, **137**, 413-421(1993).

29. Rinderknecht, E., Humbel, R. E., *FEBS Lett.* **89**, 283-286(1978).

30. Hober, S., Forsberg, G., Palm, G., Hartmanis, M., and Nilsson, B., *Biochemistry*, **31**, 1749-1756(1992).

31. Zaluzec, E. J., Gage, D. A., Allison, J., and Watson, J.T., *J. Am. So. Mass Spectrom.*, **5**, 230-237(1994).

32. Siuzdak, G., *Proc. Natl. Acad. Sci. USA*, **91**, 11290-11297(1994).

DISULFID

RECOMB

(hEGF) BY

I. Introduction

Proteins a

proteins must a

information for a

acids in the po

process involvin

that the major e

perhaps even af

into their nativ

protein's prima

scientists have

is, the folding

attempts to sol

The stu

questions: (1)

biologically-ac

CHAPTER 5

DISULFIDE MAPPING OF FOLDING INTERMEDIATES OF RECOMBINANT HUMAN EPIDERMAL GROWTH FACTOR (hEGF) BY CDAP TRAPPING, CHEMICAL CLEAVAGE, AND MASS SPECTROMETRY

I. Introduction

Proteins are synthesized within the cell on ribosomes. To be biologically active, all proteins must adopt specific folded three-dimensional structures. Yet the genetic information for a protein specifies only the primary structure, the linear sequence of amino acids in the polypeptide backbone. Although folding in the cell is a highly complex process involving a cascade of helper proteins called the molecular chaperones, it is clear that the major events in the folding process occur after departure from the ribosome and perhaps even after release from the chaperones (1, 2). Proteins can spontaneously fold into their native conformation under physiological conditions. That implies that a protein's primary structure dictates its three-dimensional structure. For many years, scientists have endeavored to understand the process by which protein folding occurs, that is, the folding pathway. Although a wealth of information has been accumulated in attempts to solve this problem, the answer is far from complete.

The study of protein folding can be broken down into three different but related questions: (1) By what kinetic process or pathway does a protein adopt its native and biologically-active folded conformation? (2) What is the physical basis of the stability of

the folded confo

folding process a

Protein

practical point o

to exploit such

process of *in vi*

Many o

be overcome if

protein-stabiliz

its oxidation/re

the disulfide

folding pathw

requires the

refolding acc

been proven

6). Two us

developed te

amide group

be applied t

not amenabl

intermediate

bonds, is o

intermediat

the folded conformation? (3) Why does the amino acid sequence determine one particular folding process and resultant three-dimensional structure, instead some other?

Protein folding study is of great significance both from the theoretical and practical point of view. Recent advances in genetic engineering have provided the means to exploit such knowledge, thereby adding a new urgency for the need to understand the process of *in vitro* protein folding.

Many of the difficulties associated with determining protein folding pathways can be overcome if folding is coupled to disulfide bond formation. This is the only type of protein-stabilizing interaction that is susceptible to specific experimental control, due to its oxidation/reduction nature. Since 80% of proteins in nature contain disulfide bond(s), the disulfide formation study is of great significance for the understanding of protein folding pathways. Folding is coupled to disulfide formation if the folded conformation requires the correct disulfide bond linkage. The reduced protein is unfolded, and refolding accompanies disulfide formation. Characterization of folding intermediates has been proven to be a very powerful tool in the elucidation of protein-folding pathways (3-6). Two useful techniques are currently available to tackle this problem. A newly developed technique of pulsed-label NMR (7, 8) permits trapping and identification of amide groups that are engaged in the structured elements. This technique can in principle be applied to all types of proteins. However, intermediates trapped by this method are not amenable to chromatographic purification because of the transient nature of folding intermediates (9). An established method, which utilizes the formation of disulfide bonds, is often used. In this method, as described in Figure 5.1, the disulfide-bonded intermediates which form during the folding process are chemically trapped in a time

course manner.

kinetics of the

connected to a

containing prote

further purified

governed by the

the simplest ap

small

course manner. The trapped intermediates are isolated and structurally characterized, the kinetics of the interconversion are determined, and the different components are connected to a folding pathway. This technique is limited to the analysis of disulfide containing proteins, but a unique advantage is that species of trapped intermediates can be further purified for characterization. Because the formation of disulfide bonds is governed by the thiol/disulfide concentration of the redox couple used, it has represented the simplest approach to study the folding of proteins. By varying the concentrations of small

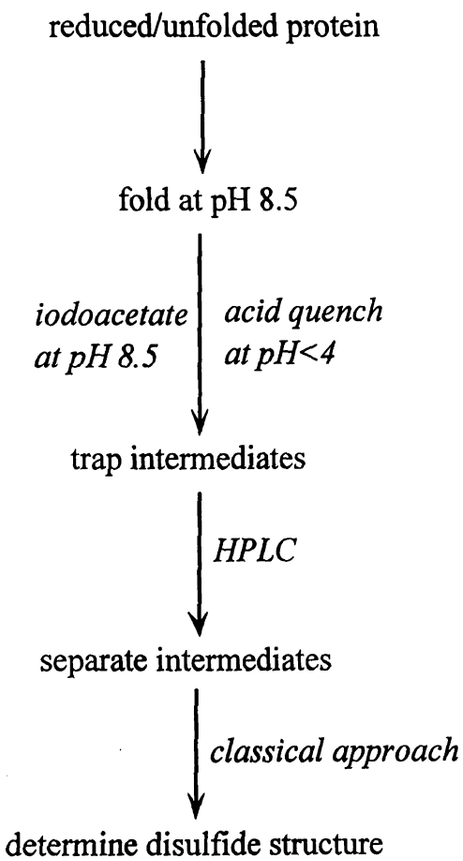


Figure 5.1. Conventional approach for disulfide mapping of protein folding intermediates

molecular weigl

intermediates as

predictable man

accumulated and

Even wi

have been char

have been stud

(BPTI) (15-18)

The firs

type protease

original mode

intermediates

and those tha

important int

modern separ

existed five s

species) and

the model

intermediate

stabilize the

dictating the

mainly of n

molecular weight thiol and disulfide components in solution, the distribution of intermediates as well as the rate of regeneration of native protein can be controlled in a predictable manner. This method thus allows construction of disulfide bond isomers accumulated and trapped along the process of folding.

Even with these advantages, the problem remains nontrivial, and few pathways have been characterized completely. The folding pathways of several proteins (10-14) have been studied in this manner. Among them, bovine pancreatic trypsin inhibitor (BPTI) (15-18), and ribonuclease A (9, 19-22) have been the most extensively studied.

The first case of a well-documented disulfide folding pathway is BPTI, a Kunitz-type protease inhibitor which comprises 58 amino acids and three disulfides. In the original model of BPTI folding (15, 23), eight well-populated 1- and 2-disulfide intermediates were identified. Five were shown to contain exclusively native disulfides, and those that adopted non-native disulfides have been suggested to be kinetically important intermediates. This original BPTI model was recently reexamined using modern separation and analytical methodologies (17, 18, 24), it was concluded that there existed five species of well-populated intermediates (two 1-disulfide and three 2-disulfide species) and all of them contain only native disulfide bonds. Despite the inconsistency, the model of BPTI folding maintains three principal conclusions: (a) the folding intermediates consist of only one- and two-disulfide species; (b) specific interactions that stabilize the native BPTI play a crucial role in guiding the folding early on and hence dictating the formation of limited numbers of well-populated intermediates that consist mainly of native disulfides. The rate-limiting step of BPTI folding is the intramolecular

rearrangements

rapidly forms the

However

nontrivial. The

number of cysti

three or fewer

disulfide protei

internal disulfi

separation of i

tasks. A clea

proposed only

Traditi

reagent that a

exchange. H

intermediates

ribonuclease

scale as alky

iodoacetate c

scale as the

problem for

of alkylatio

Because the

to quench f

rearrangements of two-disulfide intermediates to generate the immediate precursor that rapidly forms the third and final native disulfide.

However, although this approach has limited success so far, the problem remains nontrivial. The number of possible disulfide intermediates increases rapidly with the number of cystines in the protein, and success has been claimed with proteins having only three or fewer disulfide bonds. The determination of a folding pathway for a three disulfide protein is already a difficult undertaking, there being 74 possible different internal disulfide-bonded intermediates excluding the reduced and native states. The separation of intermediates and the determination of disulfide structures are formidable tasks. A clearly defined pathway for formation of a three-disulfide species has been proposed only in the case of BPTI.

Traditionally, folding intermediates were trapped by addition of iodoacetate, a reagent that alkylates free thiols and thereby prevents further oxidation or thiol-disulfide exchange. However, this method was recently questioned. The rearrangement of intermediates during trapping with iodoacetate was observed for both BPTI (17) and ribonuclease A (25) since thiol-disulfide exchange is expected to occur on the same time scale as alkylation by iodoacetate. For example, alkylation of thiols with 100 mM of iodoacetate occurs with a half-time of about 1 second at pH 8, which is on the same time scale as the kinetics of disulfide rearrangement (17). This could be a more severe problem for partially structured intermediates where steric hindrance would retard the rate of alkylation of some thiols (26). As an alternative, acid quenching was proposed. Because the thiolate anion is the reactive species in thiol/disulfide exchange, it is possible to quench folding extremely rapidly by lowering the pH. The subsequent reversed-phase

HPLC at pH 2

introducing sig

acid quenching

quenched inter

occur (17, 27).

before the disu

procedure.

Both o

(19), who insi

completely, a

2-aminoethyl

because it rea

thiol/disulfid

exchange rea

bonds, thus c

Rece

folding int

phenylarson

the presenc

allowing tr

yet formed

much on th

and, theref

HPLC at pH 2 is well suited for separation of acid-quenched intermediates without introducing significant rearrangements during the separation. A practical advantage of acid quenching is that it is reversible. As a result, it is possible to purify an acid-quenched intermediate and subsequently to allow further rearrangement or folding to occur (17, 27). On the other hand, the intermediates trapped by acid have to be modified before the disulfide bond structure can be defined. In this respective, it is a cumbersome procedure.

Both of the above trapping methods were attacked by Rothwarf and Scheraga (19), who insisted that the criteria for a good blocking agent are that it block quickly, completely, and without modifying the protein at sites other than thiols. They proposed 2-aminoethyl methanethiosulfonate [$(NH_2)C_2H_5SSO_2CH_3$] (AEMTS) as a trapping agent because it reacts specifically and much more quickly with thiol groups and thus prevents thiol/disulfide bond exchange. However, the trapping by AEMTS is based on the SH/S-S exchange reaction between an intermediate and the reagent to form mixed disulfide bonds, thus disulfide scrambling during analytical processing is still questionable.

Recently, Happersberger *et al* (28) proposed a novel method for trapping protein-folding intermediates by selective modification of bis-cysteine sulfhydryls with phenylarsonous acid derivatives (PAA), forming cyclic dithioesters in a wide pH range in the presence of two suitably spaced sulfhydryl groups of a peptide or protein, thereby allowing trapping and characterization of protein folding intermediates which have not yet formed covalent disulfide bonds. But, derivatization by this reagent depends too much on the structural accessibility of bis-cysteine SH groups which undergo the reaction and, therefore, is limited in practical use.

The isola

based on charge

of disulfide bon

efficiency of ic

separation of ir

better alternativ

In addi

for recognition

cumbersome.

To su

intermediates

large demand

In cha

pairings in pr

demonstrated

applicable to

chapter, we e

in folding int

Figur

intermediate

free sulfhyd

by reversed

the disulfide

The isolation of intermediates can be achieved by ion exchange chromatography, based on charge differences of species under separation. The information about the state of disulfide bonding is thus predictable from the elution order. However, the low column efficiency of ion chromatography prevents it from becoming a powerful tool for the separation of individual intermediates. Reversed-phase HPLC has been proven to be a better alternative for the separation, especially for the separation of small protein isomers.

In addition to the inconsistency of trapping methods, the conventional approach for recognition of disulfide bond structures of folding intermediates is very tedious and cumbersome. Problems mentioned in chapter 4 still plague the methodology.

To summarize, disulfide bond structure elucidation of protein folding intermediates is still a touchy and formidable problem in protein chemistry. There is a large demand for the development of new techniques.

In chapter 4, we have developed a novel methodology to assign disulfide bond pairings in proteins by partial reduction/cyanylation/cleavage/mass mapping. We have demonstrated that the methodology can minimize disulfide bond scrambling and be applicable to proteins containing closely spaced or adjacent cysteine residues. In this chapter, we exploit the feasibility of our novel methodology to study disulfide structures in folding intermediates and its potential advantages over conventional approaches.

Figure 5.2 illustrates our new procedure. After folding is initiated, the folding intermediates are trapped by CDAP acidic solution (pH 3) which specifically cyanylates free sulfhydryl groups and thus effectively quenches the folding process. After separation by reversed-phase HPLC, the folding intermediates are identified by MALDI-MS, and the disulfide bond structures characterized by our novel approach described in chapter 4.

Figur

Com

techniques i

mapping tec

the trappin

sulfhydryl/c

quantitative

quench, wh

determine

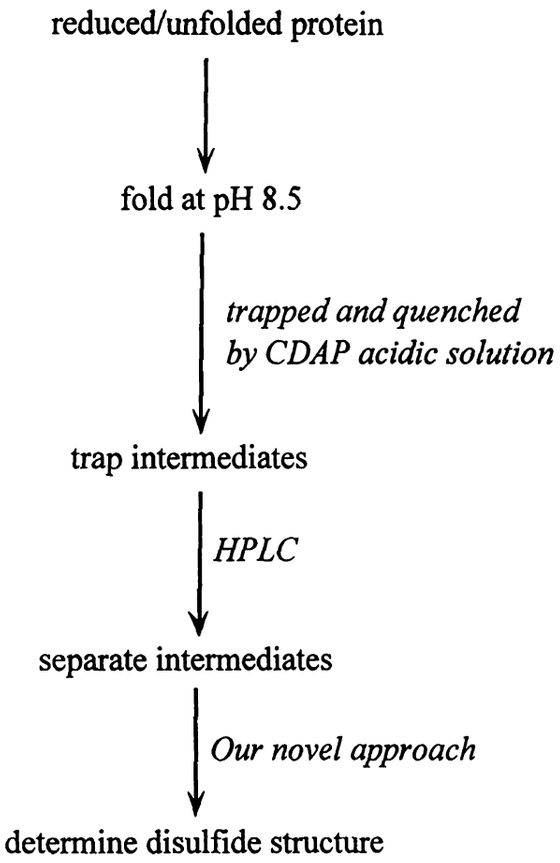


Figure 5.2. New approach for disulfide mapping of protein folding intermediates

Compared to the conventional approach, our new methodology introduces two techniques into the folding area: the disulfide trapping technique and the disulfide mapping technique. These improvements provide several important advantages. First, the trapping reaction, occurring in acidic solution, greatly reduces the risk of sulfhydryl/disulfide exchange that bothers iodoacetate trapping. Secondly, the CDAP quantitatively cyanylates free sulfhydryl groups to stop further folding. Unlike acid quench, which is totally reversible and requires further chemical modification in order to determine the disulfide structure of the folding intermediates, the cyanylated

intermediates ca

new approach.

advantages over

6
NSDSECPLSH

Figure

We us

protein contai

an example to

Prelin

were recently

intermediates

trap the inte

iodoacetate t

determined

disulfide str

We

CDAP tech

intermediates can be directly subjected to characterization of disulfide structure by our new approach. Thirdly, our disulfide bond mapping technique itself provides unique advantages over conventional techniques, as demonstrated in previous chapters.

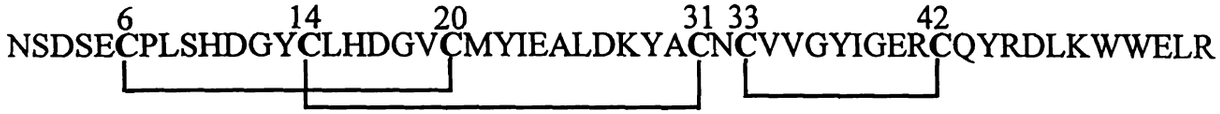


Figure 5.3. Disulfide linkage of human epidermal growth factor (hEGF).

We used recombinant human epidermal growth factor (hEGF), a small, compact protein containing 53 amino acid residues linked by three disulfide bonds (Figure 5.3), as an example to demonstrate the feasibility of our proposed methodology.

Preliminary results on the disulfide structures of folding intermediates of hEGF were recently reported by Chang *et al* (29). In this paper, the distribution of disulfide intermediates is different depending on whether iodoacetate or acidic quench was used to trap the intermediates, suggesting possible thiol/disulfide exchange in the process of iodoacetate trapping. The disulfide structures of a few well-populated intermediates were determined by classical methods, indicating that most of them contain non-native disulfide structure.

We isolated and identified over 18 disulfide folding intermediates trapped by the CDAP technique. The disulfide structures of seven well-populated were characterized by

our new approa

conventional or

is to elucidate t

methodology d

II. Experimer

Mass Spectro

MALI

spectrometer

extraction an

accelerating

voltages wer

acquired in

achieved by

1061.2), bo

16,952) obt

performed

WI) as the

acetonitrile

samples, ar

before beir

our new approach. Our methodology is simpler, faster, and more sensitive than the conventional one. Although the final goal of characterizing disulfide bond intermediates is to elucidate the folding pathway of the protein, this chapter will be only focused on the methodology development.

II. Experimental Section

Mass Spectrometry

MALDI mass spectra were obtained on a Voyager Elite time-of-flight (TOF) mass spectrometer (PerSeptive Biosystems Inc., Framingham, MA) equipped with delayed extraction and a model VSL-337ND nitrogen laser (Laser Science, Newton, MA). The accelerating voltage in the ion source was set to 20 kV. Grid voltage and guide wire voltages were 93.6% and 0.2% of the accelerating voltage, respectively. Data were acquired in the positive linear DE mode of operation. Time-to-mass conversion was achieved by external and/or internal calibration using standards of bradykinin (m/z 1061.2), bovine pancreatic insulin (m/z 5734.5),), and horse skeletal myoglobin (m/z 16,952) obtained from Sigma Chemical Co. (St. Louis, MO). All experiments were performed using α-cyano-4-hydroxycinnamic acid (Aldrich Chemical Co., Milwaukee, WI) as the matrix. Saturated matrix solutions were prepared in a 50% (v/v) solution of acetonitrile/aqueous 0.1% TFA, and mixed in equal volumes with peptide or protein samples, and applied to a stainless-steel sample plate. The mixture was allowed to air dry before being introduced into the mass spectrometer.

Materials

Recomb

Escherichia Co

purity of the p

active when

hydrochloride

hydrochloride

citric acid, soc

pyridinium te

further purifi

0.1 M citrate

at -20°C for

buffer at pH

Folding Exp

The

absence of r

0.1 M Citra

reducing ag

hours. The

the reduced

The

with 0.05

Materials

Recombinant human epidermal growth factor (hEGF) was derived from *Escherichia Coli.* Cells and was supplied by Protein Institute Inc., Broomall, PA. The purity of the protein was greater than 97%. The recombinant hEGF is fully biological active when compared with standard. Tris(2-carboxyethyl)phosphine (TCEP) hydrochloride was purchased from Pierce Chemical Co. (Rockford, IL). Guanidine hydrochloride was a product of Boehringer-Mannheim Biochemicals (Indianapolis, IN). citric acid, sodium citrate, hydrochloric acid, acetic acid, and 1-cyano-4-dimethylamino-pyridinium tetrafluoroborate (CDAP) were purchased from Sigma and used without further purification. Acetonitrile and TFA were of HPLC grade. The TCEP solution in 0.1 M citrate buffer at pH 3.0 was prepared as 0.10 M stock solution and stored under N_2 at -20°C for weeks with little deterioration. The 0.10 M CDAP solution in 0.1 M citrate buffer at pH 3.0 was freshly prepared prior to use.

Folding Experiments

The folding experiments described here were performed in Tris-HCl buffer in the absence of redox agents to control the folding. hEGF (1.0 mg) was dissolved in 0.2 ml of 0.1 M Citrate buffer, pH 3.0, containing 6 M of guanidine-HCl and 0.1 M of TCEP reducing agent. The denaturation and reduction of hEGF was carried out at 37°C for 2 hours. The reaction mixture was separated by HPLC and the fraction corresponding to the reduced hEGF was collected, recovered, dried and stored at -20°C.

The reduced and denatured proteins were refolded by diluting the protein sample with 0.05 M Tris-HCl buffer, pH 8.5 to a final protein concentration of 1 mg/ml. The

protein was sub

time course ma

solution contai

to 3 if necessa

temperature fo

fractions were

50 Da, 100 D

disulfide (III-

respectively.

Disulfide Ma

Deper

different strat

29). In addit

indicated the

from native

disulfide bo

procedure d

was estimat

exactly the

subjected to

protein was subjected to folding in the open air. Folding intermediates were trapped in a time course manner by mixing aliquots of the sample (0.1 ml) with 10 μl of 0.5 M HCl solution containing freshly prepared 0.2 M CDAP. The pH of the solution was adjusted to 3 if necessary. The cyanylation of free sulfhydryl groups was performed at room temperature for 15 minutes. The trapped intermediates were separated by HPLC. The fractions were collected manually and identified by MALDI-TOF MS. Those with 0 Da, 50 Da, 100 Da, 150 Da mass increases from the intact protein correspond to the 3-disulfide (III- or N), 2-disulfide (II-), 1-disulfide (I-), and 0-disulfide (R) species, respectively.

Disulfide Mapping of Purified, Well Populated Folding Intermediates

Depending on the number of disulfide bonds formed in the intermediates, different strategies were applied to the disulfide mapping of folding intermediates (15, 17, 29). In additional to the 2-disulfide and 1-disulfide bond intermediates, MALDI analysis indicated there are two 3-disulfide bond intermediates with quite different retention time from native hEGF, indicating they have scrambled disulfide bond structures. The disulfide bond assignment of these two intermediates was performed according to the procedure described in chapter 4. Again, the approximate amount of each intermediate was estimated by the relative peak area in the HPLC chromatogram. The procedure is exactly the same as we described in chapter 4. The quantity of the intermediates subjected to the treatment can be roughly estimated from the ratio of a given peak to the

total peak area

TCEP and CDA

For 2-

intermediate w

approach was

mapping of t

subjected to pa

4, to determi

protein, the r

species and

disulfide map

For t

purified and

adding 2µl c

the protein r

the peptide c

removed in

disulfide bo

37°C for 30

diluted with

MALDI-M

total peak areas representing all the protein intermediates. The concentration of the TCEP and CDAP can thus be calculated to ensure optimal results of the treatment.

For 2-disulfide intermediates, a portion of the purified and cyanylated intermediate was subjected to the same treatment as the 1-disulfide intermediates. This approach was applied to define the location of the two free sulfhydryl groups by mass mapping of the resulting peptides. Another portion of the same intermediate was subjected to partial reduction/cyanylation/cleavage/mass mapping, as described in chapter 4, to determine the disulfide bond structure. In additional to the original 2-disulfide protein, the mixture of partially reduced/cyanylated isomers contains two 1-disulfide species and one 0-disulfide species (completely reduced/cyanylated cysteine). The disulfide mapping of the products is described below.

For those containing only one disulfide bond (1-disulfide intermediates), the purified and cyanylated intermediates were directly subjected to chemical cleavage by adding 2μl of 1 M NH_4OH aqueous solution containing 6 M guanidine-HCl to dissolve the protein residue. Then an additional 5 μl of 1 M NH_4OH were added, and cleavage of the peptide chain was performed at room temperature for one hour. Excess ammonia was removed in a vacuum system. Truncated peptides, which may still be linked by a residual disulfide bond, were completely reduced by reacting with 2μl of 0.1 M TCEP solution at 37°C for 30 minutes at pH 3-5 to minimize the possibility of reoxidation. Samples were diluted with 100μl of a 50% (v/v) acetonitrile/0.1% TFA solution prior to analysis by MALDI-MS.

HPLC Separa

Both th

trapped interm

elution using V

at 215 nm. T

4.6×250 mm).

to facilitate th

acetonitrile/w

linear in 15 r

ml/min. The

protein isom

for further us

III. Results

A. Refoldi

The

the Chang'

However, i

achieved b

reduction

Complete

HPLC Separation

Both the purification of the reduced and denatured hEGF and the separation of trapped intermediates were performed by reversed-phase HPLC with linear gradient elution using Waters model 6000 pumps controlled by a PC computer. UV detection was at 215 nm. The column was a Vydac C18 (#218TP54, 10μm particle size, 300-Å pore, 4.6×250 mm). The chromatographic conditions resembled those used by Chang *et al* (29) to facilitate the comparison. Solvent A was water containing 0.1% TFA. Solvent B was acetonitrile/water (9:1, v/v) containing 0.1% TFA. The gradient was 14-34% solvent B linear in 15 min, 34-56% solvent B linear from 15 min to 50 min. The flow rate was 1 ml/min. The HPLC fractions were collected manually and the masses of the collected protein isomers were determined by MALDI-MS. Appropriate fractions were then dried for further use.

III. Results and Discussion

A. Refolding of hEGF under Controlled Conditions

The refolding conditions used in our experiments were the same as described in the Chang's paper (29). The hEGF samples were also obtained from the same source. However, in Chang's report, the reduction and denaturation of the native protein was achieved by adding dithiothreitol (DTT) in pH 8.5 buffer. In our experiments, the reduction and denaturation of the protein was performed in TCEP pH 3 solution. Complete reduction is possible under this condition and reoxidation of the reduced

species is minim

by passing thro

pH 8.5. The f

the same Tris-

with the same

typically low,

recovered. Fu

experiments,

the desalting

impurities.

oxidation for

The

concentratio

disulfide co

as the rate

Various re

(oxidized/re

to control t

alone the f

obtained fr

as standar

redox age

species is minimized. In Chang's experiments, the reduced/unfolded hEGF was desalted by passing though a PD-10 column (Pharmacia) equilibrated in 0.1 M Tris-HCl buffer, pH 8.5. The folding of the desalted protein was initiated immediately by diluting with the same Tris-HCl buffer to a final protein concentration of 1 mg/ml. Our experiments with the same procedure showed that the recovery of the protein from PD-10 column is typically low, although the procedure is simple. Only ~60% of the protein was recovered. Furthermore, the recovered protein has to be subjected to immediate refolding experiments, otherwise, the kinetics of the folding will be inaccurate. In our experiments, the desalting was done on an HPLC column which also purifies the protein from other impurities. Aliquots of reduced protein could be stored in dryness without detectable oxidation for months at -20°C, because the mobile phase (pH 2) stabilizes the protein.

The formation of disulfide bonds is governed by the thiol/disulfide bond concentration. By varying the concentration of small molecular weight thiol and disulfide components in solution, the distribution of intermediate disulfide forms as well as the rate of regeneration of native protein can be controlled in a predictable manner. Various redox agents, GSH/GSSG (reduced/oxidized glutathione), DTT^{ox}/DTT^{red} (oxidized/reduced dithiothreitol), and Cys/Cys-Cys (cysteine/cystine) were typically used to control the folding process, while the controlled folding was performed in Tris-HCl alone the folding is catalyzed by oxygen in air. The disulfide intermediates of hEGF obtained from controlled folding were extensively studied by Chang *et al* (29) and served as standards for measuring the efficiencies of EGF folding in the presence of various redox agents. In our experiments, only controlled folding was performed in order to

compare the re

structure mappi

B. Trapping

Conver

alkylation of s

alkaline condi

Recently Tor

iodoacetamid

advocated th

thiol/disulfid

common me

and the path

However, t

iodoacetate

either of the

Weissman a

and further

anion. Or

because th

conditions.

The

alkaline c

compare the results and to illustrate the feasibility of our methodology for disulfide structure mapping.

B. Trapping of Folding Intermediates

Conventional methodology to trap (or quench) the folding process utilizes the alkylation of sulfhydryl groups by iodoacetate (iodoacetic acid or iodoacetamide) under alkaline conditions, although the fidelity of iodoacetate trapping was recently questioned. Recently Torella et al (30) found that alkylation by iodoacetic acid, other than iodoacetamide, results in a side reaction at histidine residues. Scheraga's group (9, 19) advocated the use of AEMTS as a trapping agent, in spite of the potential risk of thiol/disulfide exchange, because of the fast reaction kinetics. Acid quenching is another common method to stop folding. The intermediates quenched by folding can be isolated and the pathway of the intermediate can be further studied by "stop-go" experiments. However, the kinetics and the distribution patterns of intermediates trapped by iodoacetate and acidic solution are frequently different, implying the possible error during either of the two trapping methods. The plausible kinetic analysis of acid quenching by Weissman and Kim (17) indicated that acid quenching can effectively slow down folding and further thiol/disulfide exchange by reducing the concentration of reactive thiolate anion. On the other hand, the results from iodoacetate trapping should be cautious because the thiol/disulfide exchange can hardly be eliminated under the reaction conditions.

The cyanylation of sulfhydryl groups by NTCB (chapter 2) was performed under alkaline conditions. However, CDAP was proved to be a useful alternative for the

cyanylation of

reagent indicat

alkylation, the

under which t

effectively tra

cyanylated pr

from folding

solution with

experimental

as the lower

would cataly

monitored b

C. Separat

Fold

were separ

structural i

MS. A to

represent 3

species.

reduced/ur

The result

cyanylation of SH groups under acidic conditions. Our experiments with the CDAP reagent indicated that, although the kinetics of the cyanylation is not as efficient as that of alkylation, the cyanylation can be finished at pH 3 within 15 min at room temperature, under which the thiol/disulfide exchange can be minimized. Not only can the CDAP effectively trap intermediates in low-pH to minimize the exchange reaction, but also the cyanylated products can be subjected to disulfide mapping after isolation. The pH change from folding (pH 8.5) to cyanylation (pH 3.0) was easily accomplished by acidifying the solution with a small volume of HCl solution containing CDAP, as described in the experimental section. The pH was important for a successful cyanylation by the CDAP, as the lower pH (<2) would slow down the reaction greatly, whereas the higher pH (>4.5) would catalyze the hydrolysis of the CDAP reagent. The final pH of the reaction was monitored by a pH meter.

C. Separation of Folding Intermediates by HPLC

Folding intermediates of hEGF trapped at 30min, 3h, 24h, and 48h, respectively, were separated by HPLC (Figure 5.4). In order to interpret these chromatograms, structural information of the fractionated intermediates was analyzed by MALDI-TOF MS. A total of 18 peaks were identified at different time points. Peaks 1, 3, and 4 represent 3-disulfide species. Peaks 2, 5, 6, 7, 10, and 12 represent 2-disulfide bond species. Peaks 13, 14, 16, and 17 are 1-disulfide bond species. Peak 18 is the reduced/unfolded hEGF. Peaks 8, 9, 11, and 15 contain both 2- and 1-disulfide species. The results revealed that the intermediates consist of eight 1-disulfide isomers and nine 2-

Figure
differer

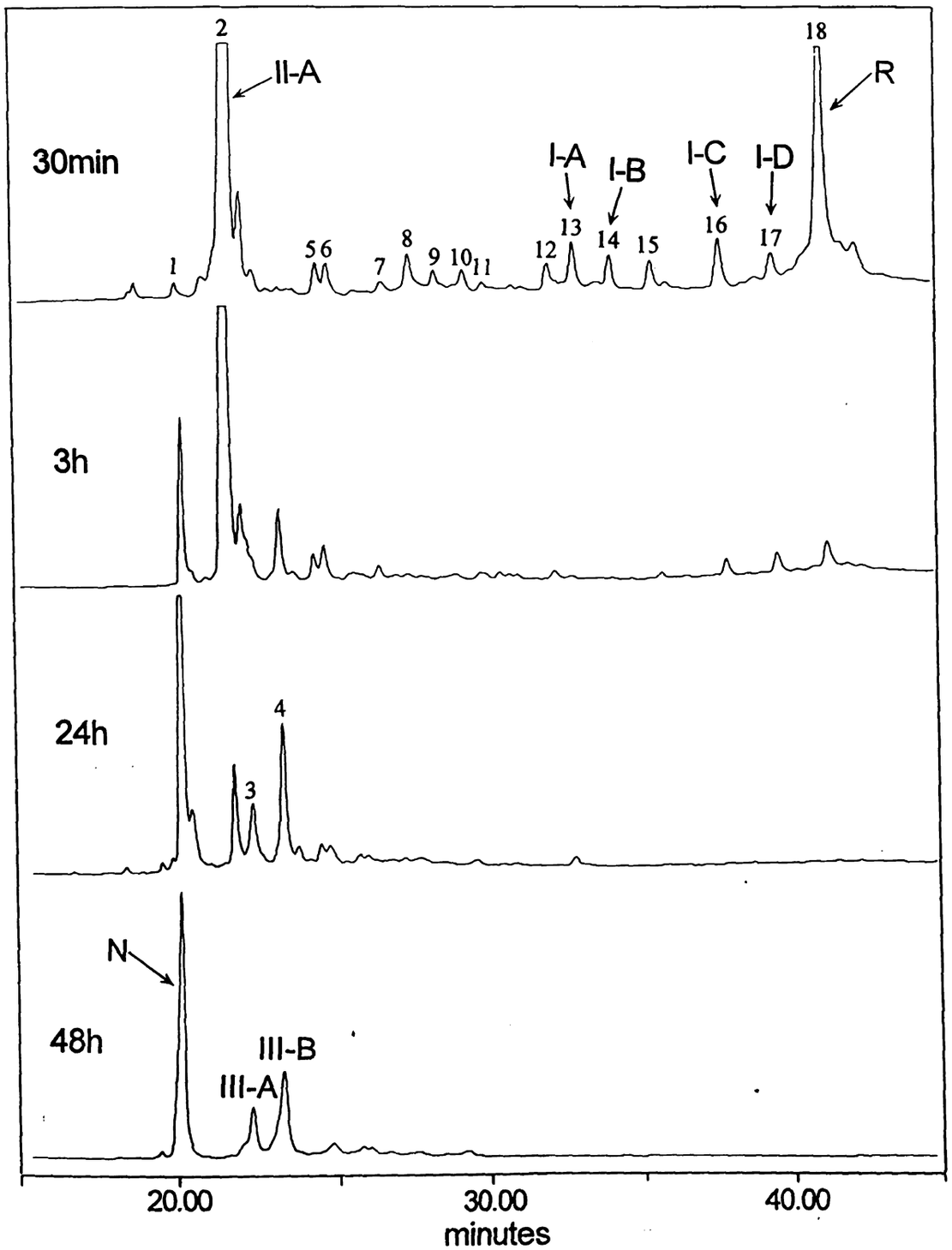


Figure 5.4. HPLC separation of CDAP-trapped folding intermediates of hEGF at different time courses.

disulfide isomer

are some overla

It was ap

of each disulfi

folding are ess

disulfide specie

of the species

after 3 hours o

the 2-disulfide

almost disappe

species were d

are similar to

Chang's pape

3- or 1-disulf

This is not su

than it should

non-native 3

MALDI-MS

chromatogra

As a

(iodoacetate,

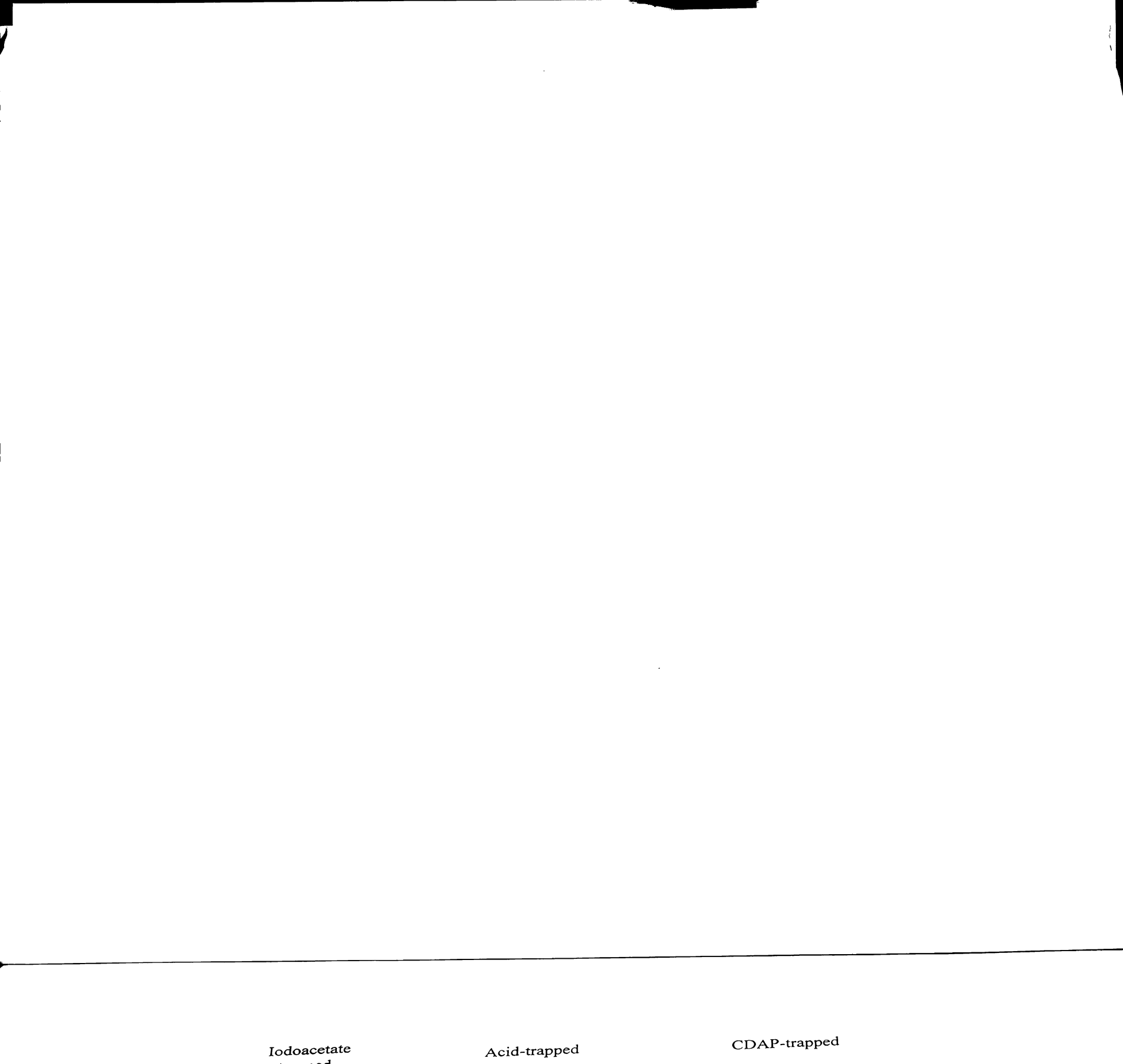
profiles sho

intermediate

disulfide isomers. Most 1-disulfide species eluted after 2-disulfide species, while there are some overlaps or reversed order.

It was apparent that along the folding process, equilibrium existed among isomers of each disulfide species. The HPLC profiles at 10- (data not shown) and 30-min of folding are essentially the same although the relative ratio of the most populated 2-disulfide species and the reduced/unfolded species is different. At 30 minutes, about 40% of the species presented as the reduced/unfolded isomer, whereas it almost disappeared after 3 hours of folding. The 3-disulfide species began to appear as a main peak, while the 2-disulfide species dominated. After 24 hours of folding, the 2-disulfide species almost disappeared and the native hEGF dominated. After 48 hours, only the 3-disulfide species were detected, representing complete folding. The HPLC profiles presented here are similar to some of Chang's results (29), however, there are some differences. In Chang's paper, thirteen 2-disulfide isomers were identified, some of them overlaped with 3- or 1-disulfide species. We were able to detect nine discernible 2-disulfide isomers. This is not surprising as the sensitivity of our detection system was substantially lower than it should be. Unlike Chang's chromatogram in which the two peaks representing the non-native 3-disulfide isomers contain a considerable amount of 2-disulfide species, MALDI-MS analysis showed that the corresponding peaks (3 and 4) in our chromatograms represented pure isomers.

As a comparison, the HPLC profiles obtained by different trapping techniques (iodoacetate, acid, or CDAP) at different time points are shown in Figure 5.5. The three profiles show almost the same pattern in terms of the appearance of disulfide intermediates. It should be noted however that, in Chang's experiments, the HPLC


Iodoacetate
Acid-trapped
CDAP-trapped

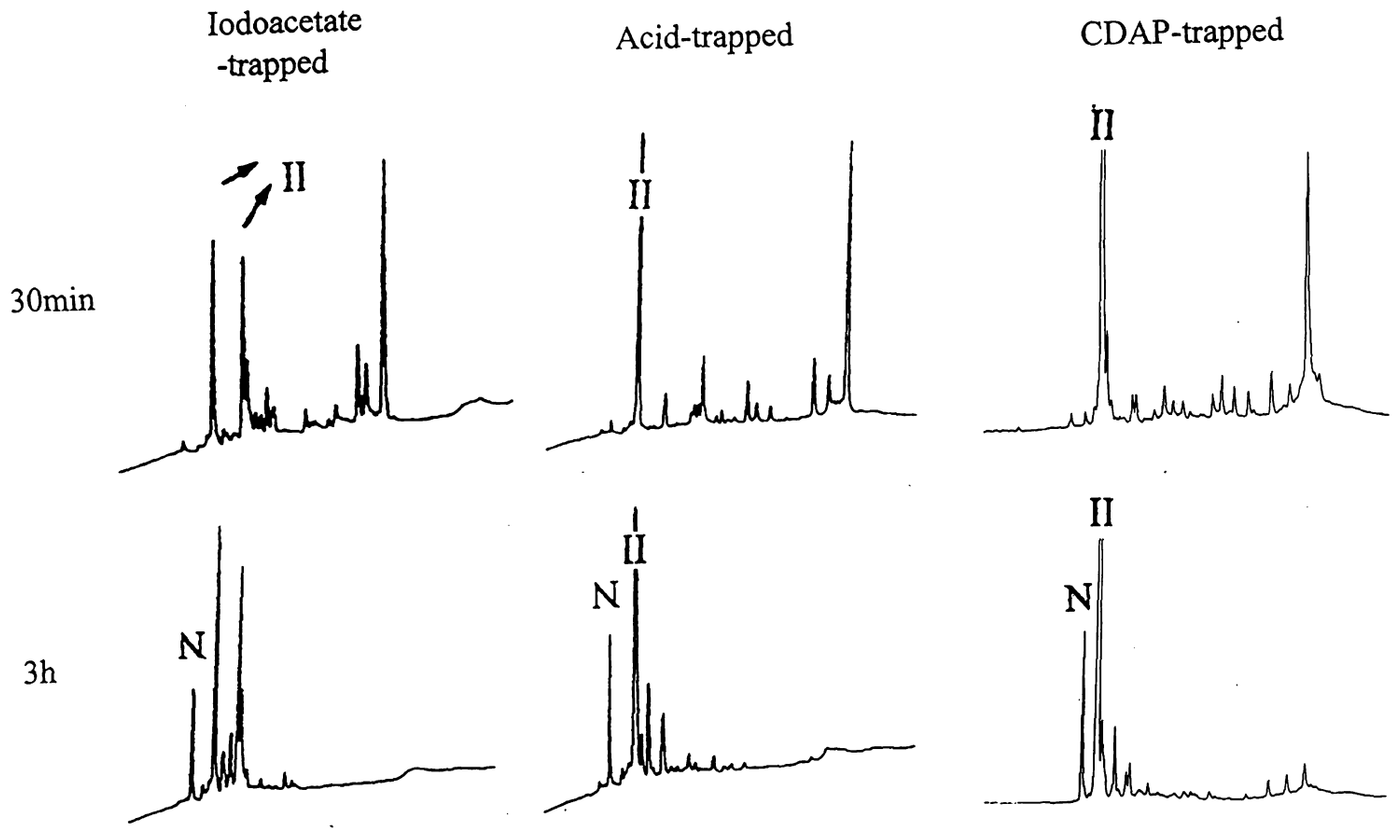


Figure 5.5. Comparison of HPLC profiles of iodoacetate-trapped, acid-trapped, and CDAP-trapped folding intermediates of hEGF.

profiles of acid-

counterparts. I

isomers trapped

from the under

or less. Howe

disulfide inter

disulfide intern

al (29) assum

means that all

the hydrophol

farfetched.

cyanylation d

SCN groups

acid and CDA

intermediates

intermediates

populated 2-

peak could c

characterizat

our results

intermediate

measuremen

profiles of acid-trapped intermediates did not fully resemble those of iodoacetate-trapped counterparts. It is understandable in terms of the difference in the hydrophobicity of isomers trapped by different manners. The hydrophobicity of alkylated protein differs from the underivatized isomers. Therefore, the relative retention time could change more or less. However, it is beyond understanding that the majority of iodoacetate-trapped 2-disulfide intermediates were eluted within three fractions, whereas acid-trapped 2-disulfide intermediate(s) were accumulated within one fraction (marked as II). Chang *et al* (29) assumed all the three fractions coeluted in acid-trapped HPLC profile, which means that all the three isomers had the same hydrophobicity before modification, while the hydrophobicities were substantially different after alkylation. This explanation is farfetched. Fortunately, the hydrophobicity of hEGF isomers before and after cyanylation does not change greatly (the polarity of SH and SCN is similar while SH and SCN groups are both small species), direct comparison of HPLC profiles obtained from acid and CDAP trappings is reasonable. It is noteworthy that the HPLC chromatogram of intermediates trapped by the CDAP technique is similar to that representing the intermediates trapped with acid. In both cases, only one peak corresponding to the most populated 2-disulfide intermediate(s) was obtained. To exclude the possibility that the peak could contain more than one 2-disulfide species, as suggested by Chang (29), the characterization of the disulfide structure of the species was carried out (see below) and our results indicated unambiguously that the peak contains ONLY one 2-disulfide intermediate. This result confirms that acid trapping provides a more accurate measurement of folding process. The two major isomers observed by Chang in the

chromatogram

caused by alk

It was

of each disu

intermediates

fractions for

behaved simi

previous rep

characterized

equilibrated

experiments,

(Figure 5.4,

convert the

alkaline me

Tyr, and a r

thiols to cat

D. Disulfi

Tra

one is "pep

Alternative

a color or

chromatogram representing iodoacetate trapping could be from thiol/disulfide exchange caused by alkylation under alkaline condition.

It was apparent that along the folding process, equilibrium existed among isomers of each disulfide species. For instance, the concentration of 1- and 2-disulfide intermediates ascended and then descended as folding progressed, but the relative ratio of fractions for each kind of isomers remain constant. Scrambled 3-disulfide species behaved similarly during the folding. These data supported the conclusion drawn by previous reports (29) that, like other proteins, the folding pathway of the hEGF was characterized by a sequential flow of unfolded EGF (R) through three groups of equilibrated intermediates, namely, 1-, 2-, and 3-disulfide (scrambled) isomers. In our experiments, about 40% of the folding intermediates remained as scrambled species (Figure 5.4, 48-h sample). The exposure to air for another 48 hours cannot automatically convert the scrambled species to the native form (furthermore, long hours of exposure to alkaline medium results in minor oxidation of some amino acids such as Met, His and Tyr, and a mass shift from the intact protein would be observed), due to the lack of free thiols to catalyze the disulfide reshuffling.

D. Disulfide Mapping of Well Populated Intermediates

Traditionally, this can be achieved by a number of strategies. The most common one is "peptide mapping" by isolating and analyzing every enzyme-fragmented peptide. Alternatively, it can be done by selective labeling of disulfide bonds after reduction with a color or fluorescent thiol-specific reagents (17, 31). Both methods need microgram

amounts of inte

and numerous a

As we c

of cyanylated

sulfhydryl grou

structure deter

defining disul

chemistry beh

intermediates

successfully cl

3-Disulfide Ir

After 4

Mass analysis

peak showed

confirmed th

were two no

collected an

strategy desc

as straightfo

isomers were

CDAP was t

Tabl

possible cya

amounts of intermediates, multiple enzymatic digestions, HPLC separation of peptides, and numerous attempts of sequence analysis.

As we demonstrated in chapters 2-4, the mass mapping of the cleavage products of cyanylated proteins provides specific information on the number and location of sulfhydryl groups and disulfide bond pairs. This strategy is also applied to the disulfide structure determination of the folding intermediates of hEGF. The procedures for defining disulfide structure in 1-, 2- and 3-disulfide species differ slightly, but the chemistry behind the methodology is exactly same. A total of seven well populated intermediates (marked with III-A, III-B, II-A, I-A, I-B, I-C, and I-D in Figure 5.4) were successfully characterized by our strategy.

3-Disulfide Intermediates

After 48-h folding (Figure 5.4, 48-h), three well populated species were separated. Mass analysis by MADLI indicated all three species are 3-disulfide isomers. The first peak showed the identical retention time as the native hEGF. Disulfide mapping confirmed that the isomer is the native hEGF (data not shown). The latter two peaks were two non-native isomers with scrambled disulfide structures. The two peaks were collected and subjected to the partial reduction/cyanylation/cleavage/mass mapping, a strategy described in chapter 4. The control of partial reduction is important but is never as straightforward as the protocols used in chapter 4. The molar amounts of protein isomers were estimated from the HPLC peak area and the concentration of the TCEP and CDAP was thus calculated.

Table 5.1 shows the expected m/z of fragments resulting from cleavage of 15 possible cyanylated disulfide isomers. Figure 5.6 shows the HPLC separation of partially

Table 5.1.
of 15 singly

Reductio of Disul
6-14
6-20
6-3
6-3
6-
14
1
1

Table 5.1. Expected m/z values for possible fragments resulting from the cleavage of 15 singly reduced/cyanylated hEGF isomers

Reduction of Disulfide	Fragment	m/z	Reduction of Disulfide	Fragment	m/z
	1-5	551.5		1-13	1423.4
6-14	6-13	915.9	14-42	14-41	3160.6
	14-53	4837.6		42-53	1721.0
	1-5	551.5		1-19	2047.1
6-20	6-19	1539.6	20-31	20-30	1344.6
	20-53	4214.0		31-53	2913.3
	1-5	551.5		1-19	2047.1
6-31	6-30	2840.2	20-33	20-32	1560.8
	31-53	2913.3		33-53	2697.2
	1-5	551.5		1-19	2047.1
6-33	6-32	3056.4	20-42	20-41	2536.9
	33-53	2697.2		42-53	1721.0
	1-5	551.5		1-30	3347.7
6-42	6-41	4031.5	31-33	31-32	261.3
	42-53	1721.0		33-53	2697.2
	1-13	1423.4		1-30	3347.7
14-20	14-19	667.7	31-42	31-41	1236.3
	20-53	2913.3		42-53	1721.0
	1-13	1423.4		1-32	3563.9
14-31	14-30	1968.3	33-42	33-41	1020.1
	31-53	2913.3		42-53	1721.0
	1-13	1423.4			
14-33	14-32	2184.5			
	33-53	2697.2			

Figure 5.6

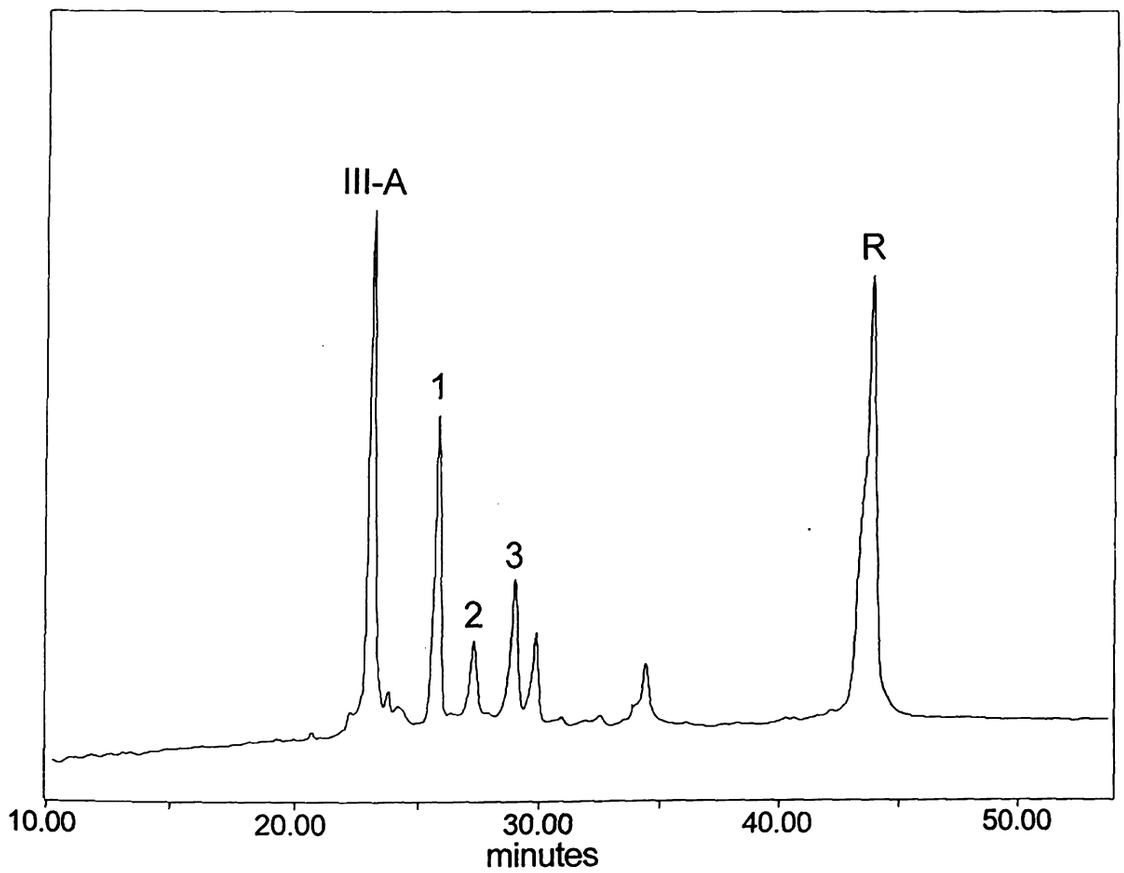


Figure 5.6. HPLC separation of III-A and its partially reduced/cyanylated species.

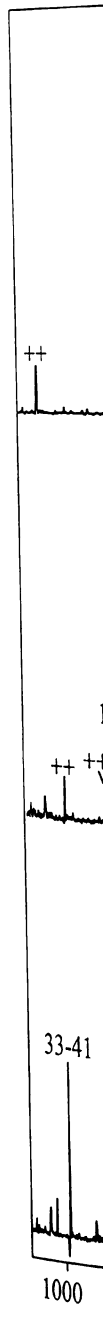


Figure 5.
of the thr
correspo
and * rep

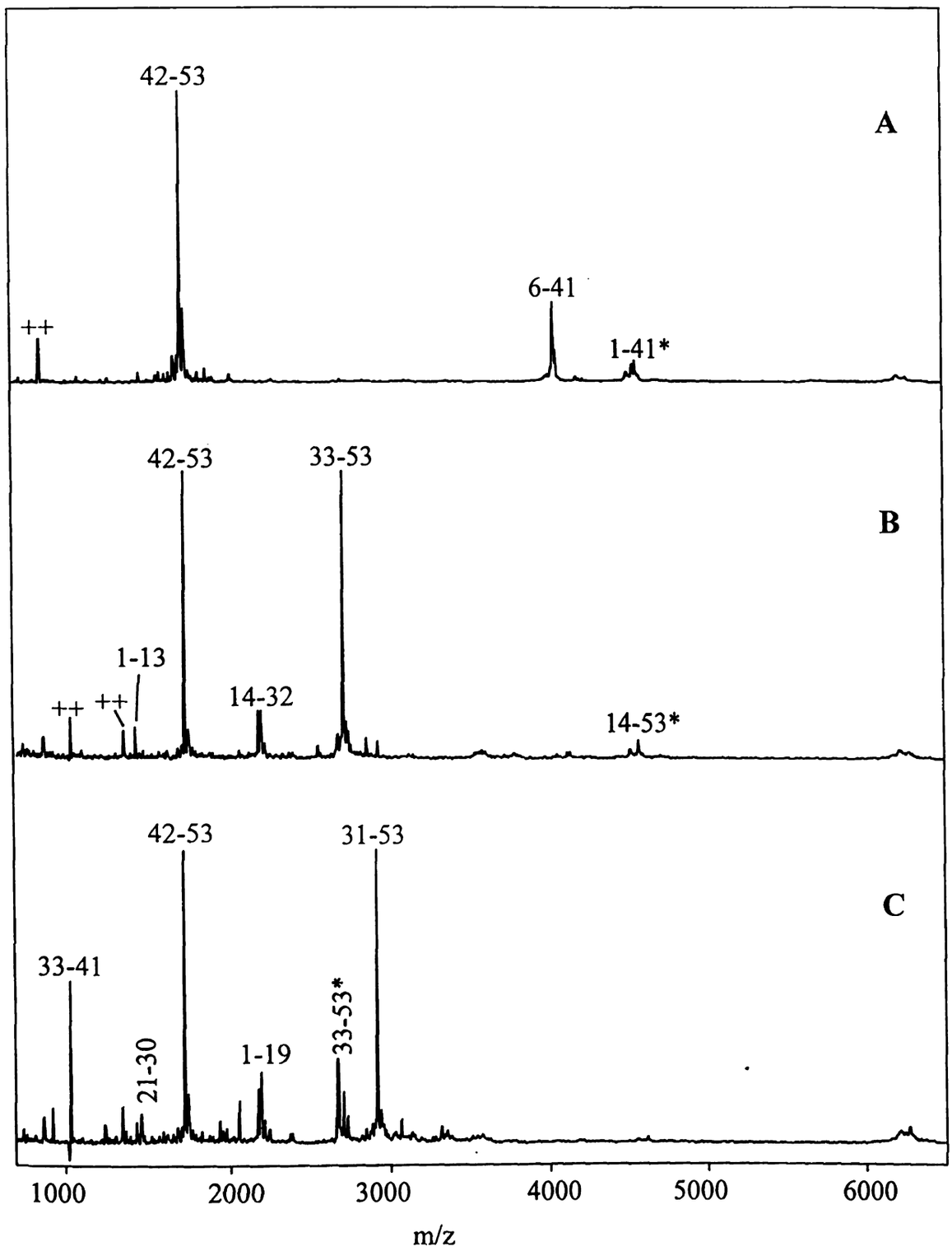


Figure 5.7. The MALDI mass spectra of peptide mixtures resulting from the cleavage of the three singly reduced/cyanylated species of non-native hEGF III-A, A-C corresponding to the HPLC peaks 1-3 in Figure 5-6, respectively. The symbols ++ and * represent the doubly-charged species and β-elimination products, respectively.

reduced/cyanyl

1-3 were singly

reduced EGF

distribution of

disulfide bond

completely red

products corre

spectrum (Fig

fragments 6-

fragment 1-5

Cys6 is conne

From

and 33-53 we

5.7B). An o

Therefore, C

fragment 42-

degree of c

cyanylation,

intensity of t

the assignme

Although fra

HPLC peak

reduced/cyanylated protein isomer III-A. Mass analysis of the peaks showed that peaks 1-3 were singly reduced/cyanylated species, the peak eluted at 42.5 min was a completely reduced EGF which should be typically present as a small peak. This atypical distribution of the partially reduced species suggests that the opening of the first two disulfide bonds facilitate the reduction of the last one, resulting in the accumulation of completely reduced species. Figure 5.7 shows the MALDI mass mapping of the cleavage products corresponding to HPLC peaks 1-3 in Figure 5.6, respectively. The MALDI spectrum (Figure 5.7A) of the cleavage products of HPLC peak 1 in Figure 5.6 shows fragments 6-41, 42-53, and an overlapped β-elimination product 1-41. The small fragment 1-5 was missing. It is clear that cleavage occurs at Cys6 and Cys42. Therefore, Cys6 is connected to Cys42 as a disulfide bond.

From the cleavage products of HPLC peak 2 in Figure 5.7, fragments 1-13, 14-32, and 33-53 were detected, indicating peptide chain cleavage at Cys14 and Cys33 (Figure 5.7B). An overlapped β-elimination product, 14-53, corroborates with the conclusion. Therefore, Cys14 is linked to Cys42. However, this assignment was questioned by the fragment 42-53, indicating that Cys42 also underwent cleavage. Apparently a minor degree of cleavage must have occured at Cys42, either by the disulfide split and cyanylation, or by some disulfide exchange during cleavage process, or both. The peak intensity of this fragment is so strong that further experiments were performed to confirm the assignment.

Although fragment 42-53 is indicated in the MALDI spectra of cleavage products from HPLC peaks 1-3, the absolute concentration of the fragment is quite different in the

Figur
reduc
2 in

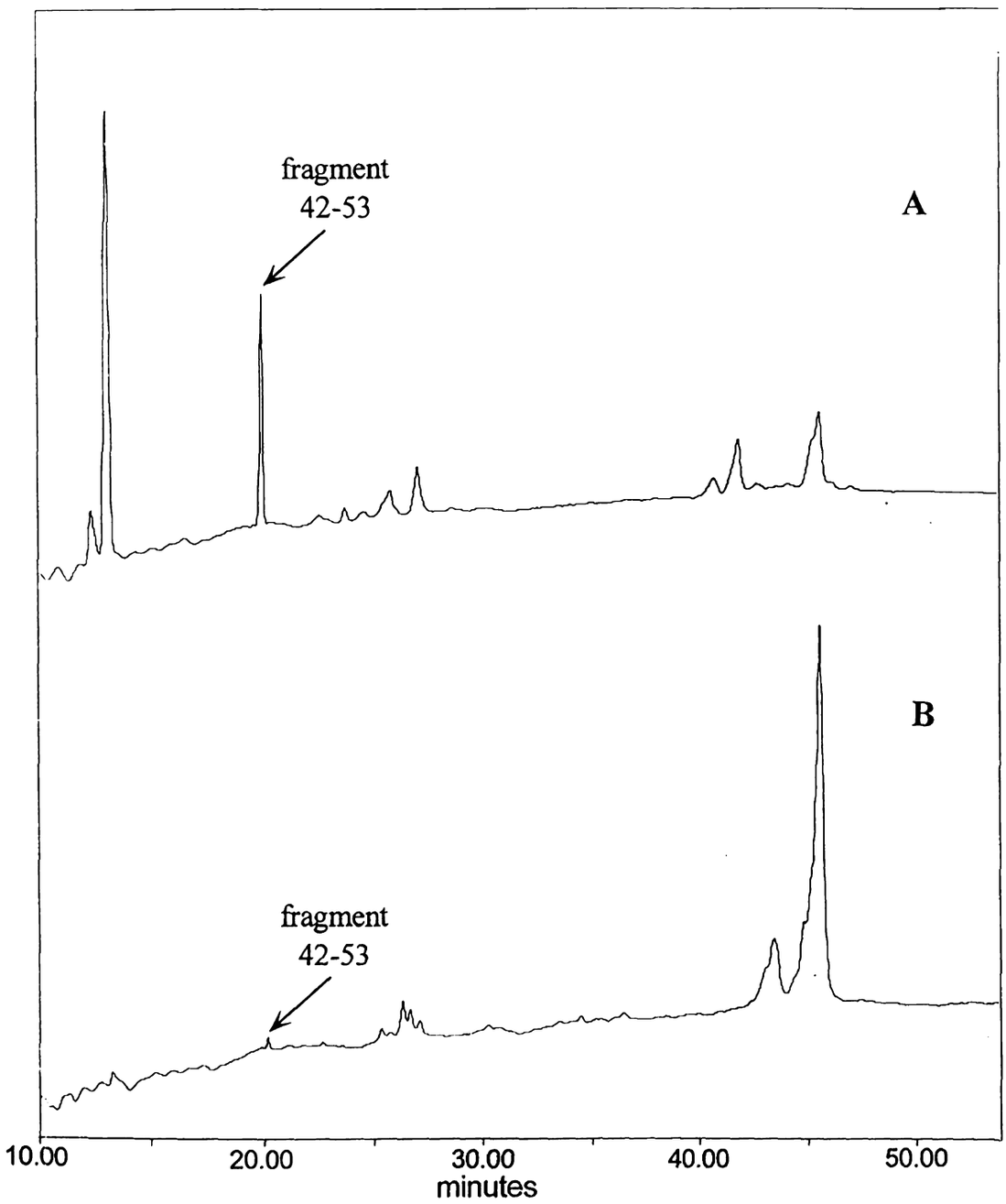


Figure 5.8. Comparison of HPLC profiles of cleavage products of two singly reduced/cyanylated III-A isomers. (A) and (B) are represented by peaks 1 and 2 in Figure 5-6, respectively.

three cases. Th

the same HPL

that the fractio

fragment repre

while the fragr

so small com

fragment inde

fact, the peak

mixture of cle

1% contamin

the data interp

The n

the fragments

pair (Cys20-

Again, the M

The

same strateg

species. Pe

analysis by

shown in F

elimination

three cases. The cleavage products from HPLC peaks 1 and 2 were re-fractionated under the same HPLC conditions (Figure 5.8). The MALDI analysis of the fractions indicated that the fraction eluting at 17.5 min corresponded to fragment 42-53. It was clear that this fragment represented one of the major cleavage products in HPLC peak 1 (Figure 5.7A), while the fragment 42-53 in the mixture of the cleavage products from HPLC peak 2 was so small compared to other fragment(s) that it could be neglected. However, this fragment indeed plagued every MALDI spectrum due to its high response. As a matter of fact, the peak corresponding to fragment 42-53 was still intense on MALDI even if the mixture of cleavage products from HPLC peak 1 was diluted by 100-fold. That means 1% contamination by fragment 42-53 could generate a false information and jeopardize the data interpretation.

The mixture of the cleavage products from HPLC peak 3 showed the presence of the fragments 1-19, 20-30, 31-53 (Figure 5.7C), indicating the linkage of last disulfide pair (Cys20-Cys31) which is already deducible after the first two pairs were defined. Again, the MALDI spectrum is also plagued by a trace of the undesired fragment 42-53.

The disulfide structure in hEGF isomer III-B was subjected to analysis by the same strategy. Figure 5.9 shows the HPLC separation of the partially reduced/cyanylated species. Peaks 1-3 are three singly reduced/cyanylated isomers, as confirmed by mass analysis by MALDI. The MALDI analysis of cleavage products from these peaks are shown in Figure 5.10A-C. Figure 5-10A shows the fragments 6-13, 14-53, and a β-elimination product 1-13. It is clear that Cys6 is hooked to Cys14. A small peak

Figure

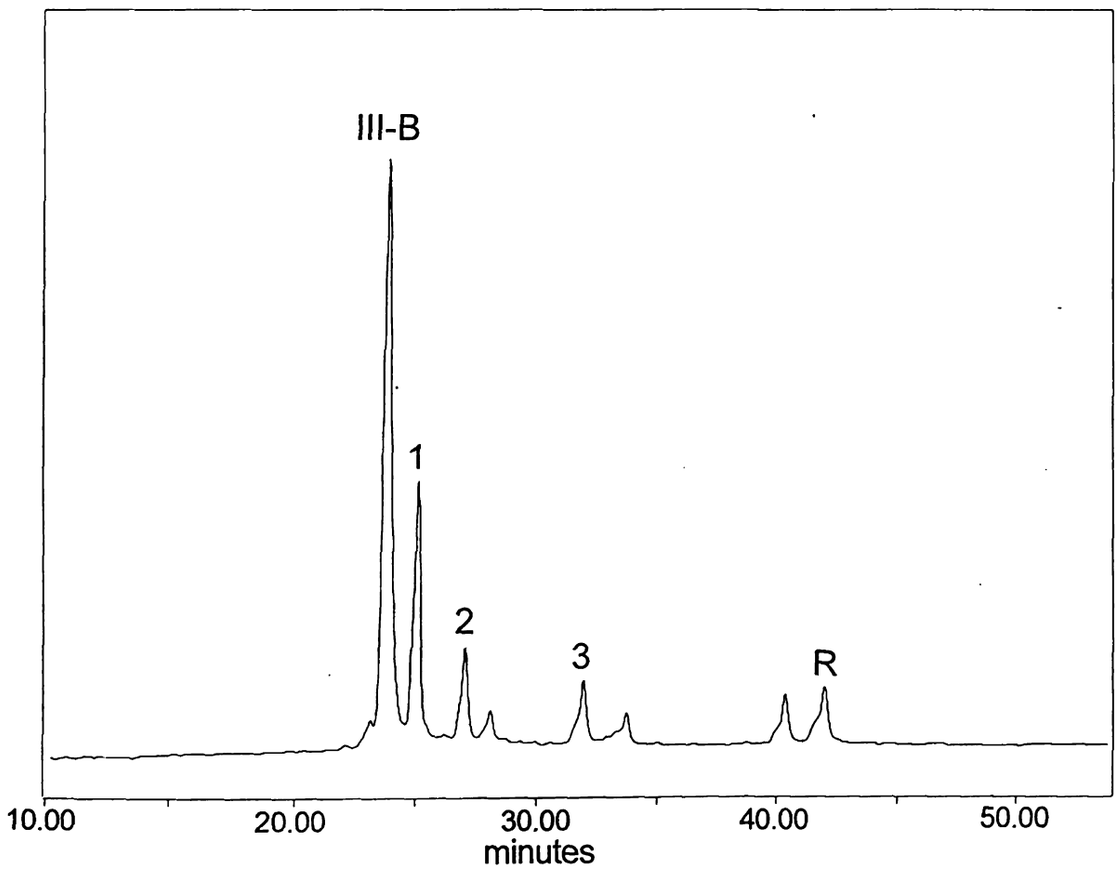


Figure 5.9. HPLC separation of III-B and its partially reduced/cyanylated species.

Figure 5.
3 singly
to the H

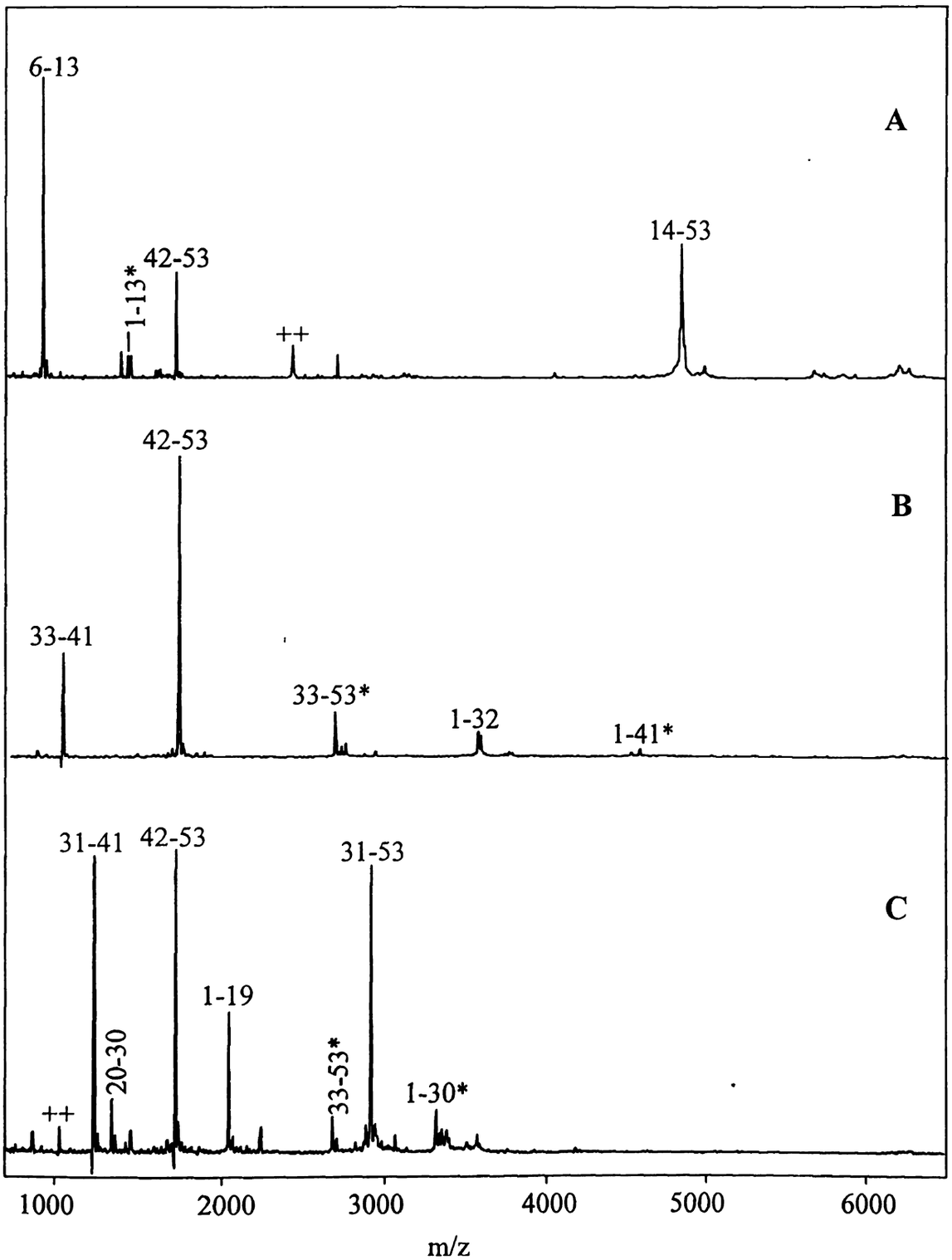


Figure 5.10. The MALDI mass spectra of peptide mixtures from the cleavage of the 3 singly reduced/cyanylated species of non-native hEGF III-B, A-C corresponding to the HPLC peaks 1-3 in Figure 5.9, respectively.

corresponding

5.10B) of the

fragments 1-3

elimination p

conclusively

therefore, car

pair, Cys20-C

fragments in

and a β-elimi

the assignme

the MALDI

31-53 can b

fractionation

HPLC/ESI e

2-Disulfide

In ou

the disulfid

55% and >

HPLC isola

11 shows th

the original

corresponding to the fragment 42-53 can be ignored. The MALDI spectrum (Figure 5.10B) of the cleavage products corresponding to HPLC peak 2 in Figure 5.9 shows fragments 1-32, 33, 41, 42-53, and two overlapped peptides corresponding to the β-elimination products 1-41 and 33-53, respectively. Disulfide bond linkage 33-42 can be conclusively assigned. In this spectrum, the fragment 42-53 presents as a huge peak and, therefore, cannot be ignored. From the disulfide structure of the first two pairs, the third pair, Cys20-Cys31, can be deduced, which is confirmed by the analysis by MALDI of the fragments in Figure 5.10C. In addition to the expected fragments, 1-19, 20-30, 31-53, and a β-elimination product 1-30, two undesired fragments, 31-41 and 42-53, can confuse the assignment, as these fragments indicate that Cys42 is also a cleavage site. Frankly, the MALDI is limited in this case for an unambiguous assignment, although the fragment 31-53 can be used as evidence that Cys42 is a minor cleavage site. The HPLC fractionation of the cleavage products combined with the MALDI identification or HPLC/ESI experiments should be informative to clarify the uncertainty.

2-Disulfide Intermediates

In our folding experiment, only one 2-S-S intermediate (II-A) was collected and the disulfide bond structure characterized. This most abundant intermediate counts for 55% and >70% of the total population at 30-min and 3-h folding, respectively. The HPLC isolated, cyanylated II-A was subjected to partial reduction/cyanylation. Figure 5-11 shows the chromatogram of the partially reduced/cyanylated products. In addition to the original 2-disulfide species, two 1-disulfide and one 0-disulfide species arise from the

Figure 5.1

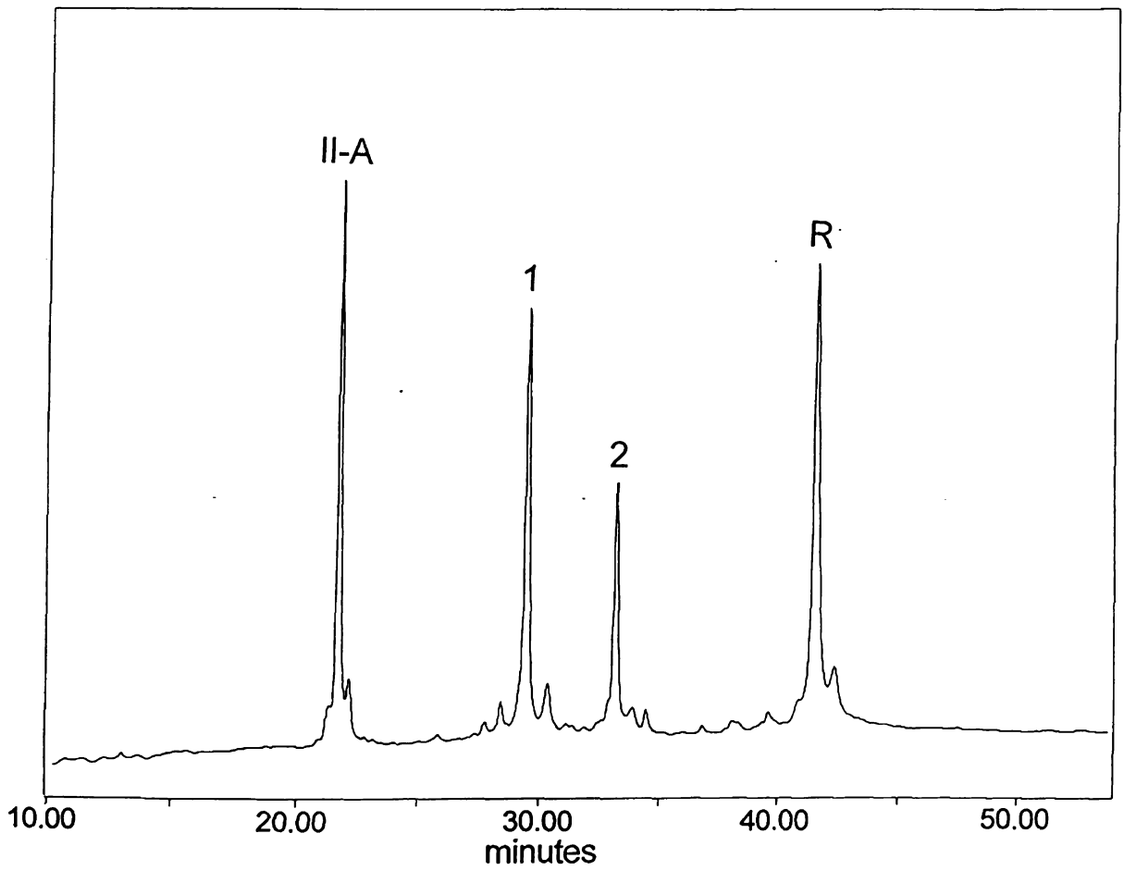


Figure 5.11. HPLC separation of II-A and its partially reduced/cyanylated species.

partial reducti

disulfide oper

disulfides ope

It shou

more than on

of the partiall

pure interme

one species,

complicated

Altho

products of

easier to loca

in II-A will

products rel

spectra. The

peak II-A sl

cleavage oc

residues. T

whatsoever

The

2 in Figure

indicating t

partial reduction and cyanylation. That is, the HPLC peak 1 (Figure 5.11) has one disulfide opened, peaks 2 and 3 have two disulfides opened, and peak 4 has all three disulfides opened, resulting in a familiar partial reduction pattern.

It should be pointed out that the HPLC fraction II-A in Figure 5.4 could contain more than one unresolved species, as suggested by Chang (29). From the HPLC pattern of the partially reduced protein isomers, it is clear that the original peak II-A represents a pure intermediate, other species, if any, are very minor. If the fraction contains more than one species, the partially reduced products of those species should be much more complicated than the present pattern. Obviously it is not the case.

Although the disulfide structure of the II-A can be directly mapped from cleavage products of the species represented by HPLC peaks 2 and 3 in Figure 5.11, it is much easier to locate first the cyanylated cysteine residues. The localization of the SH groups in II-A will greatly help the subsequent identification of the MALDI data of the cleavage products related to the reduced disulfide bonds, as the latter usually give more complex spectra. The MALDI spectrum (Figure 5.12A) of the cleavage products corresponding to peak II-A shows fragments 6-19, 20-53, and β-elimination product 6-53, indicating that cleavage occurred at cysteine residues 6 and 20, which represent two unoxidized SH residues. The peak 1-53 represents an intact II-A that did not undergo any cleavage whatsoever under our experimental conditions.

The MALDI spectrum (Figure 5.12B) of the cleavage products from HPLC peak 2 in Figure 5.11 shows fragments 6-19, 33-41, 42-53 and a β-elimination product 33-53, indicating that another disulfide pair, 33-42, must have been reduced, cyanylated and

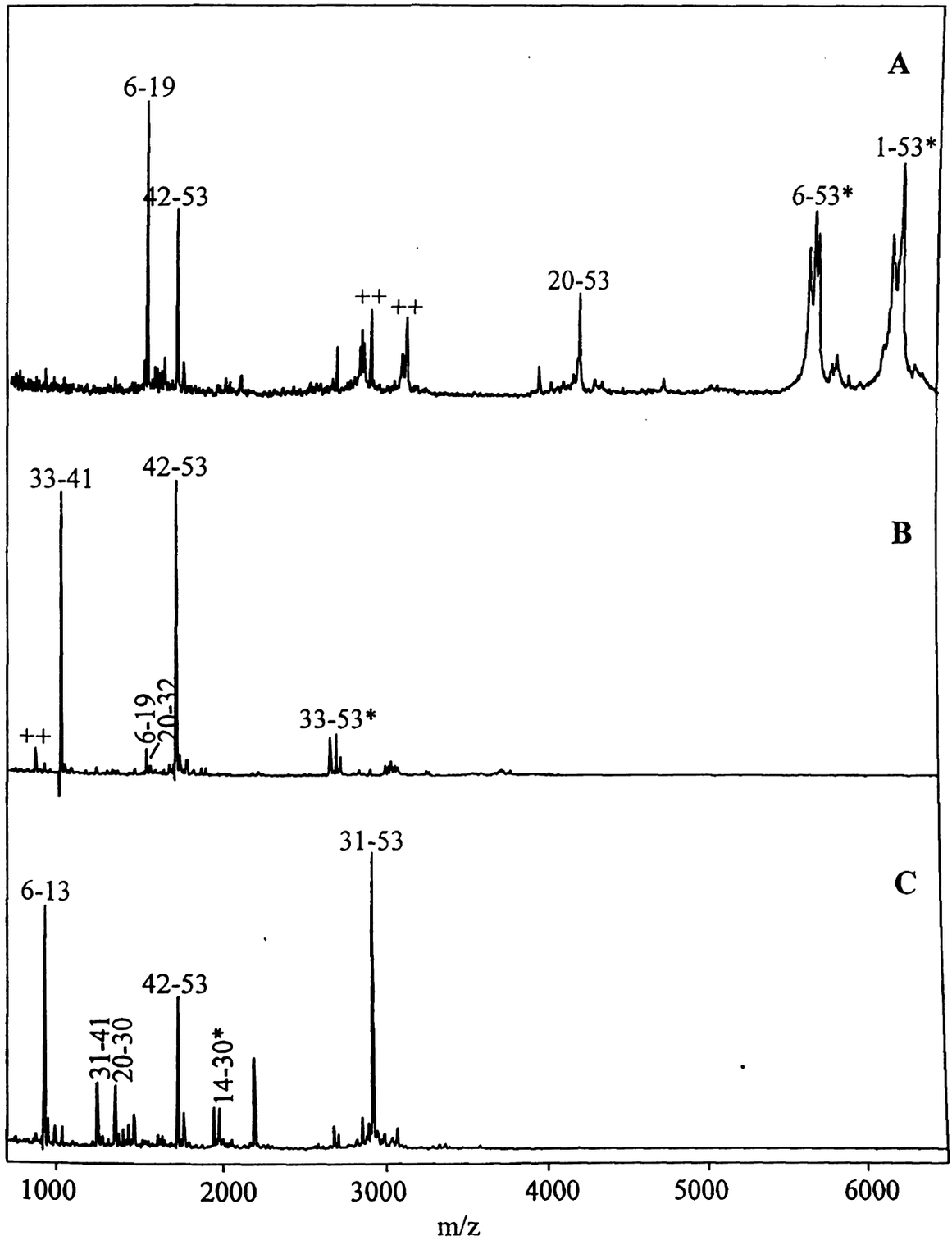


Figure 5.12. The MALDI mass spectra of peptide mixtures from the cleavage of (A) the cyanylated hEGF II-A, (B) and (C) the partially reduced/cyanylated species of II-A, corresponding to the HPLC peaks 1 and 2 in Figure 5.11, respectively.

cleaved. Therefore, Cys33 is connected to Cys42. Likewise, the MALDI spectrum (Figure 5.12C) of the cleavage products from HPLC peak 3 in Figure 5.11 shows fragments 6-13, 20-30, 14-30, and 31-53, indicating cleavage at cysteine residues 6, 14, 20, and 31, respectively. There is a minor degree of undesired cleavage at Cys42, resulting in the formation of fragments 31-41 and 42-53.

In summary, intermediate II-A, the most populated species during the early stage of folding, contains a native structure, Cys14-Cys31 and Cys33-Cys42.

1-Disulfide Intermediates

The recognition of disulfide structures in 1-disulfide intermediates is straightforward. The strategy is based on the mass mapping of cleavage products of the cyanylated 1-disulfide species. Instead of looking at the MALDI peaks corresponding to the cleavage sites, the absence of such peaks represents the "uncleaved sites", that is, those still linked by a disulfide bond. For a protein containing 6 cysteine residues, there are fifteen possibly 1-disulfide intermediates. The possible fragments corresponding to cleavage at cyanylated cysteine residues of the 15 isomers are listed in Table 5.2. Since some of the fragments may not be detected by MALDI, the mass mapping of the products from the completely reduced/cyanylated hEGF was used as a "figureprint" spectrum of the cleavage products. It is immediately clear that the fragment 1-5 cannot be detected by MALDI, while fragment 42-53 is very sensitive to MALDI detection. This profile helps the interpretation of MALDI data from other intermediates, especially if ambiguity arises.

The MALDI spectra of the cleavage products of the four well populated 1-disulfide intermediates are shown in Figure 5.13. Figure 5.13A is the MALDI spectrum

Table 5.2. Fragments of cleavage products corresponding to 15 cyanylated 1-disulfide intermediates of hEGF

Disulfide Linkage	Fragment	m/z	Disulfide Linkage	Fragment	m/z
	1-19	2047.1		1-5	551.5
	20-30	1344.6		6-19	1539.6
6-14	31-32	262.3	14-42	20-30	1344.6
	33-41	1020.1		31-32	262.3
	42-53	1721.0		33-53	2697.2
	1-13	1423.4		1-5	551.5
	14-30	1968.3		6-13	915.9
6-20	31-32	262.3	20-31	14-32	2184.5
	33-41	1020.1		33-41	1020.1
	42-53	1721.0		42-53	1721.0
	1-13	1423.4		1-5	551.5
	14-19	667.7		6-13	915.9
6-31	20-32	1560.8	20-33	14-30	1344.6
	33-41	1020.1		31-41	1236.3
	42-53	1721.0		42-53	1721.0
	1-13	1423.4		1-5	551.5
	14-19	667.7		6-13	915.9
6-33	20-30	1344.6	20-42	14-30	1968.3
	31-41	1236.3		31-32	262.3
	42-52	1721.0		33-53	2697.3
	1-13	1423.4		1-5	551.5
	14-19	667.7		6-13	915.9
6-42	20-30	1344.6	31-33	14-19	667.7
	31-32	262.3		20-41	2536.9
	33-53	2697.2		42-53	1721.0
	1-5	551.5		1-5	551.5
	6-30	2840.2		6-13	915.9
14-20	31-32	262.3	31-42	14-19	667.7
	33-41	1020.1		20-32	1560.8
	42-53	1721.0		33-53	2697.2
	1-5	551.5		1-5	551.5
	6-19	1539.6		6-13	915.9
14-31	20-32	1560.8	33-42	14-19	667.7
	33-41	1020.1		20-30	1344.6
	42-53	1721.0		31-53	2913.3
	1-5	551.5			
	6-19	1539.6			
14-33	20-30	1344.6			
	31-41	1236.3			
	42-53	1721.0			

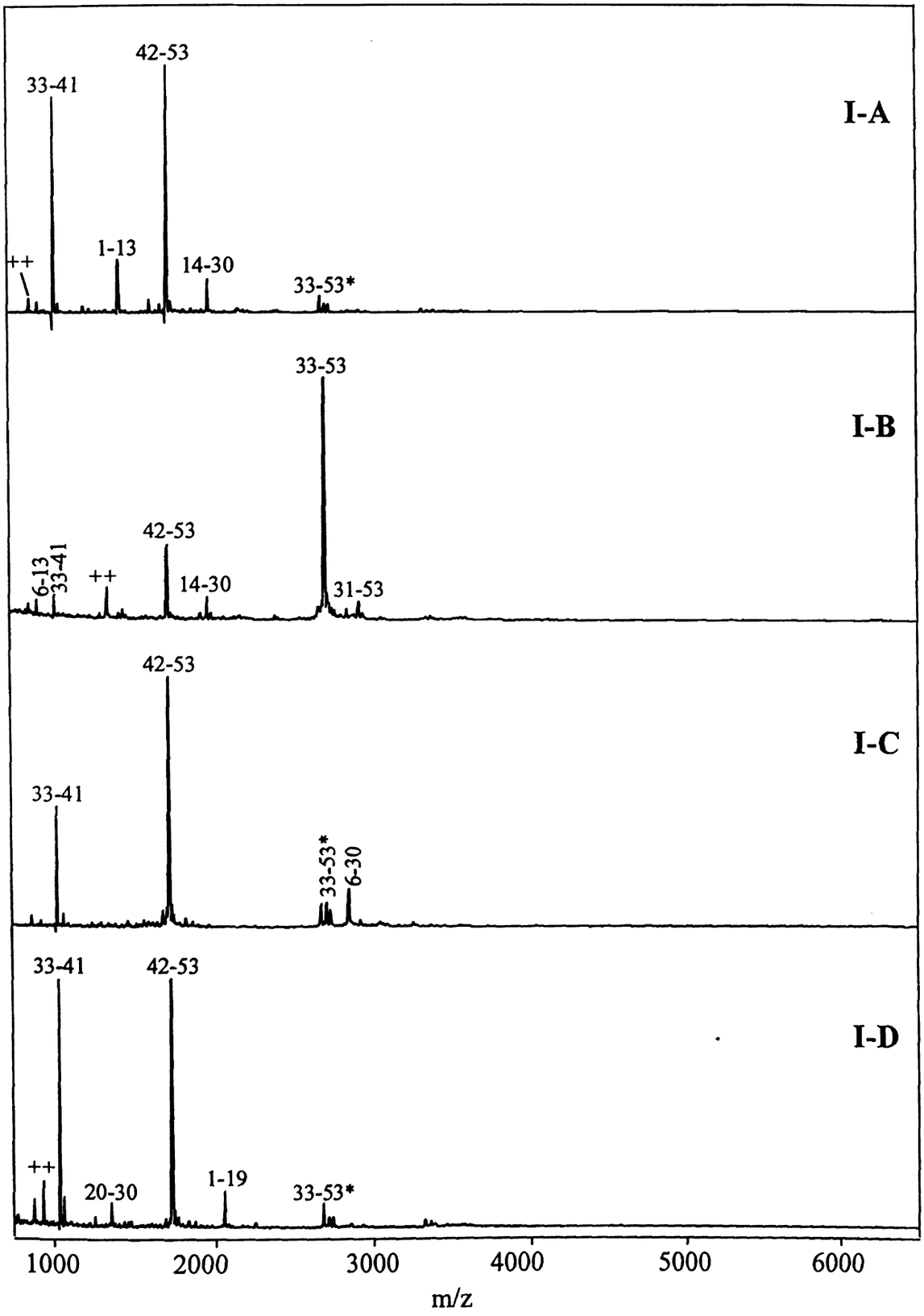


Figure 5.13. Disulfide mapping of 1-disulfide folding intermediates of hEGF.

of cleavage products corresponding to HPLC I-A in Figure 5.4. Fragments 1-13, 14-30, 33-41 and 42-53 were detected, implying chemical cleavage at a cyanylated cysteine at position 14, 31, 33, and 42, respectively. Obviously, no cleavage occurs at Cys6 and Cys20. Otherwise, fragments 1-5, 6-13, 14-19, and 20-30 would arise, giving m/z at 551.5, 915.9, 667.7, and 1344.6, respectively. Therefore, Cys6-Cys20 linkage is deduced. It is notable that this intermediate has a native disulfide structure, which was not detected in Chang's experiments (29). The second 1-disulfide intermediate (I-B) showed MALDI peaks corresponding to fragments 6-13, 14-30, 33-53, and 42-53, respectively. Cleavage occurs at Cys6, 14, 31, 33, and 42, but not Cys20. Extensive experiments indicated that the cleavage at Cys42 was very minor in comparison with the cleavage at other sites. Since fragment 33-53 is very intense, a reasonable conclusion can still be drawn that Cys42 does not present as a free sulfhydryl. Therefore, Cys20 is connected to Cys42. The MALDI spectrum for I-C (Figure 5.13C) is simple to interpret. Fragments 6-30, 33-41, and 42-53 show that Cys6, 31, 33, 42 are cleavage sites. Therefore, the disulfide pair 14-20 can be deduced. By the same strategy, the MALDI spectrum (Figure 5.13D) of I-D shows peaks for the fragments 1-19, 20-30, 33-42, and 42-53, suggesting the Cys6-Cys14 linkage. The disulfide structures of the last two intermediates are in agreement with the structures proposed by Chang *et al* (29). By our strategy, we were able to recognize the disulfide structure in I-A and I-B, neither of which was assigned by Chang *et al*. Among the four well populated 1-disulfide intermediates, only one has a native disulfide structure.

IV. Aspects for Further Improvement

One of the most important features of our methodology is the simplicity. Only a few fragments were produced, each of them is relevant to the assignment of disulfide structures. Some undesired fragments may be observed, but they rarely affect the assignment of disulfide bonds because our methodology only relies on the recognition of specific cleavage at cysteine residues. However, if the side products correspond to the cleavage at undesired cysteine site(s), the data interpretation would no longer be straightforward. This situation is more complicated by the fact that the MALDI response is not proportional to concentration or amount of analytes. Even a trace amount of an impurity may give a high MALDI response, as seen by the fragment 42-53 in the case of hEGF. Fortunately, HPLC can provide quantitative information on the relative concentration of analytes, if ambiguity arises. The most straightforward way is to perform HPLC/ESI analysis of cleavage products so that information on both quantities and masses of the analytes can be obtained.

Our CDAP trapping provides advantages over both iodoacetate trapping and acidic quenching. Effective trapping can be achieved in most cases as demonstrated by the MALDI analysis of the trapped intermediates. However, the CDAP concentration and the pH must be controlled because of the possible side reactions (see chapter 3) and the instability of CDAP at higher pH (hydrolysis). As the available SH groups vary in the folding course and the concentrations of various intermediates differ greatly, the optimal CDAP concentration is difficult to control. It should be a good investment to study more stable and specific novel reagents that can cyanylate SH groups in acidic solution. Thiocyanopyridine (TCP) (32) (see chapter 3) may act as such a reagent. Another

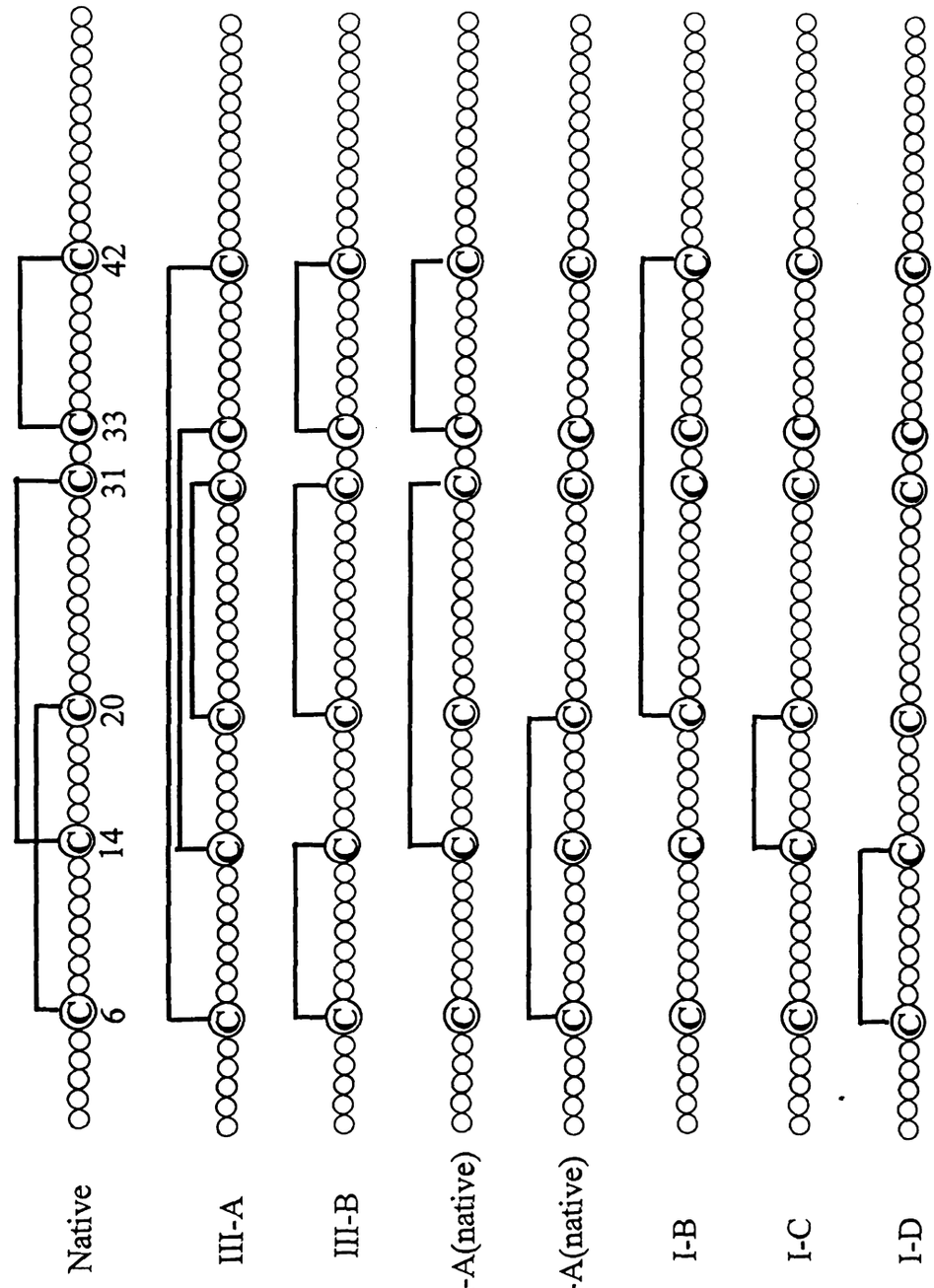


Figure 5.14. Disulfide structures of seven well populated folding intermediates of hEGF.

reagent, 4-thiocyanatoaniline (TCNA), was introduced in Biemann's group several years ago (33). This reagent was claimed to be a useful alternative to NTCB. But the capability of this reagent to modify SH groups in acidic solution remains to be examined.

V. Conclusions

To summarize, seven species of well populated intermediates in the refolding of hEGF have been isolated and characterized, which included four 1-disulfide, one 2-disulfide, and two 3-disulfide scrambled species (Figure 5.14), of which only two have native disulfide structure. Among the seven intermediates, I-A and I-B are newly identified; the disulfide structures of the other five intermediates are the same as these published previously (29). These results demonstrate the feasibility of our methodology. The significance of these intermediates to the folding pathway of hEGF needs to be further specified.

Our trapping technique, based on the cyanylation of free thiol groups under acidic conditions, showed similar results to those of acid trapping in terms of the distribution of intermediates during the folding process. This technique circumvents problems associated with the traditional iodoacetate trapping and greatly facilitates the subsequent characterization of disulfide structures of isolated folding intermediates. Our strategy for the determination of disulfide structures by chemical cleavage and mass mapping of the fragments is much simpler, faster, and conclusive, although the undesired cleavage at Cys42 indeed occurred in some cases.

This methodology opened a new door to the application of our chemistry. The preliminary experimental results presented in chapters 4 and 5 also show the potential to

characterize disulfide structures of folding intermediates of even more complicated proteins, such as those containing adjacent cysteine residues (e. g., IGFs).

VI. References

1. Ellis, R. J., *Curr. Opin. Struc. Biol.*, **4**, 117-122(1994).
2. Frydman, R. B., Nimmesgern, E., Ohtsuka, K., and Hartl, F. U., *Nature, Lond.* **370**, 111-117(1994).
3. Kim, P. S., and Baldwin, R. L., *Annu. Rev. Biochem.*, **59**, 631(1990).
4. Matthews, C. R., *Annu. Rev. Biochem.*, **62**, 653(1993).
5. Ptitsyn, O. B., *Curr. Opin. Struc. Biol.*, **5**, 74(1995).
6. Li, Y.-J., Rothwarf, D. M., and Scheraga, H. A., *Nat. Struct. Biol.*, **2**, 489(1995).
7. Jennings, P. A., and Wright, P. E., *Science*, **262**, 892-896(1993).
8. Balbach, J., Forge, V., van Nuland, N. A. J., Winder, S. L., Hore, P. J., and Dobson, C. M., *Nat. Struct. Biol.*, **2**, 865-870(1995).
9. Xu, X., Rothwarf, D. M., and Scheraga, H. A., *Biochemistry*, **35**, 6406-6417(1996).
10. Frech, C., and Schmid, F. X., *J. Mol. Biol.*, **251**, 135-149(1995).
11. Creighton, T. E., and Ewbank, J. J., *Biochemistry*, **33**, 1534-1538(1994).
12. Chang, J.-Y., *Biochemistry*, **35**, 11702-11709(1996).
13. Glocker, M. O., Arbogast, B., Milley, R., Cowgill, C., and Deinzer, M. L., *Proc. Natl. Acad. Sci., USA*, **91**, 5868-5872(1994).
14. Hober, S., Forsberg, G., Palm, G., Hartmanis, M., and Nilsson, B., *Biochemistry*, **31**, 1749-1756(1992).
15. Creighton, T. E., and Goldenberg, D. P., *J. Mol. Biol.*, **179**, 497(1984).
16. Darby, N. J., Norin, P. E., Talbo, G., and Creighton, T. E., *J. Mol. Biol.*, **249**, 463(1995).

17. Weissman, J. S., and Kim, P. S., *Science*, **253**, 1386(1991).

18. Weissman, J. S., and Kim, P. S., *Proc. Natl. Acad. Sci. USA*, **89**, 9900(1992).

19. Rothwarf, D. M., and Scheraga, H. A., *Biochemistry*, **32**, 2671(1993).

20. Rothwarf, D. M., and Scheraga, H. A., *Biochemistry*, **32**, 2680(1993).

21. Rothwarf, D. M., and Scheraga, H. A., *Biochemistry*, **32**, 2690(1993).

22. Rothwarf, D. M., and Scheraga, H. A., *Biochemistry*, **32**, 2698(1993).

23. Creighton, T. E., *Science*, **256**, 111-112(1992).

24. Weissman, J. S., and Kim, P. S., *Cell*, **71**, 841-851(1992).

25. Rothwarf, D. M., and Scheraga, H. A., *J. Am. Chem. Soc.*, **113**, 6293(1991).

26. Goto, Y., and Hamaguchi, *J. Mol. Biol.*, **146**, 321(1981).

27. Chang, J.Y., Biochemistry, 35, 11702-11709(1996).

28. Happersberger, P., Cowgill, C., Deinzer, M. L., and Glocker, M. O., *Proceedings of the 44th ASMS Conference on Mass Spectrometry and Allied Topics*, Portland, Oregon, May, 1996, p1050.

29. Chang, J.-Y., Schindler, P., Ramseier, U., and Lai, P.-H., *J. Biol. Chem.*, **270**(16), 9207-9216(1995).

30. Torella, C., Ruoppolo, M., Marino, G., and Pucci, P., *FEBS Letters*, **352**, 301-306(1994).

31. Chang, J.-Y., *J. Biol. Chem.*, **268**, 4043-4049(1993).

32. Brocklehurst, K., Malthouse, J. P. G., Baines, B. S., Blenkinsop, R. D., Churcher, J. A., Mushiri, M. S., and Ormerod, F., *Biochem. Soc. Trans.*, Vol. 6, 261-263(1978).

33. Zaia, J., Papayannopoulos, I. A., and Biemann, K., *Proceedings of the 39th ASMS Conference on Mass Spectrometry and Allied Topics*, Nashville, TN, May 1991.