OUKA
The University of Osaka
Institutional Knowledge Archive

Title	Isolation of free thiol-containing proteins/peptides from a complex biological sample
Author(s)	Donzeli Pereira, Caroline
Citation	大阪大学, 2014, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/50471
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

OSAKA UNIVERSITY

Isolation of free thiol-containing proteins/peptides from a complex biological sample

Caroline Donzeli Pereira
2014

Isolation of free thiol-containing proteins/peptides from a complex biological sample

By

Caroline Donzeli Pereira

Graduate School of Sciences
Department of Chemistry
Osaka University
August, 2014.

Table of Contents

Acknowledgments IV

Abbreviations VI

Abstract X

List of Figures XII

List of Tables XVIII

General Introduction **1**

References 13

Chapter I. A method for isolating cysteine containing peptides **17**

I.1. Introduction 17

Conventional methods for isolating peptides containing cysteine 19

I.2. Material and Methods 21

I.3. Results 24

I.4. Discussion 32

References 34

Chapter II. The method for isolating proteins containing free thiol groups **39**

II.1. Introduction 39

Papain 40
II.2. Material and Methods 43
II.3. Results 46
II.4. Discussion 50
References 54

Chapter III. Isolating proteins with free thiol groups in plasma 59
III.1. Introduction 59
Transthyretin 60
Amyloidosis process and other diseases 63
Oxidative Stress and related diseases 66
III.2. Material and Methods 70
III.3. Results 77
III.4. Discussion 90
References 97

General Conclusion 105

Acknowledgments

First of all, I would like to thank God, the Almighty. You made everything possible. In these four years, You taught me about faith and strength; thus, I could complete this research successfully.

I would like to thank my parents, Arlete and Aurito, and my sister Carine, for your love, prayers, encouragement and sacrifices for preparing me for my future.

I would like to express my deep and sincere gratitude to my research supervisor, Professor Toshifumi Takao, for your invaluable insight, support and hours of discussion. I like to acknowledge the assistance of the professors in the thesis committee, the members of the School of Sciences Faculty and the members of the laboratory of Protein Profile and Functional Proteomics.

Finally, I would thank to all my friends: Marina, Carlos, Ibrahim, Pedro, Olivia, Tiago, Akemi, Gilberto, Fátima, Vânia and Márcia. You were very important during my time here in Japan. I cannot even imagine how my life would have been without these amazing people around me. Last but not least, I would like to thank my friends in Brazil: Daniele, Kalló, Marcos, Thaise, Akemi, Rafael, Fernanda and Gustavo.

I owe my success to all of the wonderful people who have supported and assisted me during my studies.

Abbreviations

2DE	Bidimensional Electrophoresis
2-PDS	2,2'-Dithiodipyridine
ACTH	Adrenocorticotropic Hormone
ADH	Alcohol Dehydrogenase
ATC	N-{2-((2-Acryloyl)Amino)Ethyl}-1,3-Thiazolidine-4-Carboxamide
BCA	Bicinchoninic Acid Assay
Boc-ITC	Tert-Butoxycarbonyl-N-{2-((2-Iodoacetyl) Amino)-ethyl}-1,3-Thiazolidine-4-Carboxamide
BSA	Bovine Serum Albumin
CHCA	α-Cyano-4-Hydroxy Cinnamic acid
CRISPs	Cysteine-Rich Secretory Proteins
CysTRAQ	Cysteinyl Peptide Enrichment and Isobaric Tags for Relative and Absolute Quantitation
DbsA (B,C and G)	Bacterial Protein Disulfide Isomerase
dCMP	Deoxycytidylate Hydroxymethylases
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
DSIP	Delta Sleep-Inducing Peptide
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
ER	Endoplasmic Reticulum
ERO1	Oxidoreductin 1

FAC	Familial Amyloidotic Cardiomyopathy
FAD	Cofactor Flavin
FAP	Autosomal Dominant Systemic Familial Amyloidotic Polyneuropathy
GAPDH	Glyceraldehyde- 3-Phosphate Dehydrogenase
GSNO	S-Nitrosoglutathione
GSSG	Oxidized Glutathione
H_2O_2	Hydrogen Peroxide
Hcy	Homocysteine
HPPP	Human Plasma Proteome Project
HUPO	Human Proteome Organization
ICAT	Isotope-Coded Affinity Tag
IgG	Immunoglobulin G
iTRAQ	Isobaric Tags for Relative and Absolute Quantitation
LC	Liquid Chromatography
LC-SPDP	6-[3'-(2-Pyridyldithio)-Propionamido] Hexanoate
m/z	Mass to Charge Ratio
MALDI-TOF/TOF	Matrix-Assisted Laser Desorption/Ionization – Time of Flight
MDA	Malondialdehyde
MM	Molecular Weight Marker
mRNA	Messenger RNA
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MT	Metallothioneins
MW	Molecular Weight
NADPH/$NADP^+$	Nicotinamide Adenine Dinucleotide Phosphate

NEM	N-Ethylmaleimide
NF-κB	Factor-κB
NMR	Nuclear Magnetic Resonance
NO	Nitrogen Oxides
O_2	Oxygen
O_2^-	Superoxide Anion
·OH	Hydroxyl Radical
Ox-PTMs	Oxidative Posttranslational Modifications
PBD	Protein Data Bank
PBS	Phosphate Buffered Saline
PDI	Protein Disulfide Isomerase
PDO	Protein Disulfide Oxidoreductase
pNa	Pyr-Phe-Leu-pNa
Prx	Peroxiredoxin
PTM	Posttranslational Modifications
RBP	Retinol Binding Protein
RNA	Ribonucleic Acid
RNS	Reactive Nitrogen Species
ROS	Reactive Oxygen Species
S100	Low-Molecular-Weight Proteins Characterized by Two Calcium-Binding Sites
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
-SH	Sulfhydryl Group
-SOH	Sulfenic Acid

$-SO_2H$	Sulfinic Acid
$-SO_3H$	Sulfonic Acid
-SNO	S-Nitrosylation
SOD	Superoxide Dismutase
SPDP	N-Succinimidyl 3-(2-Pyridyldithio) Propionate
SSA	Systemic Amyloidosis
-SSG	S-Glutathionylation
TBS	Tris Buffered Saline
TIFF	Tagged Image File Format
TTR	Transthyretin
ZIP	ZRT, IRT-like Protein

Abstract

Cys has a reactive thiol group that can suffer from various modifications such as oxidation to sulfenic (-SOH), sulfinic (-SO_2H), sulfonic acid (-SO_3H), etc. This amino acid is the second least abundant amino acid, but it is highly conserved in amino acid sequences among species, and plays important roles for protein functions as active sites, regulatory sites, structure-stabilizing disulfides, cofactor bindings, etc. The importance of Cys has also been shown by the fact that its mutation into other amino acids can frequently trigger dysfunctions of proteins or genetic diseases. Since Cys is typically present in all classes of proteins and peptides and has the unique chemical property of containing a thiol group, a number of methods have been developed for isolating cysteine-containing peptides/proteins. However, only a few methods have been reported for the isolation of cysteine-containing proteins in the literature. In this sense, a aminopropyl resin coupled with a hetero-bifunctional cross-linker [*N*-succinimidyl 3-(2-pyridyldithio) propionate (SPDP) or [succinimidyl 6-[3'-(2-pyridylthio)- propionamido] hexanoate] (LC-SPDP) allows the proteins and peptides containing free thiols to be specifically captured on the resin and to be recovered by elution with a buffer containing a reducing agent. To evaluate the method, a tryptic digest of bovine serum albumin and an extract from *Carica papaya* latex, which contains the cysteine protease papain, were applied to the system. The method could successfully work with small amounts of a sample, down to low picomoles; in the case of the cysteine protease, the method allowed the separation of the active (unmodified) enzyme from the inactive one (modified form). Moreover, the method has been applied to human plasma to see whether or not the plasma contains free-thiol containing proteins. The eluted fraction predominantly contained transthyretin (TTR), which has a single Cys in its sequence. Despite that this free Cys has been known to display little physiological function, it has been reported to be modified with various low-molecular-weight compounds, such as glutathione, Cys, etc., and the products might be involved with the advent of TTR's amyloidogenic property. TTR has been deeply involved with some physical disorders and can be a useful biomarker for diagnosis of cancer and amyloidosis process. As such, data about the ratio of the free Cys form of TTR to the oxidized ones may correlate with one's physiological state, rationalized to show susceptibility to oxidative stress, or one's propensity to develop some diseases as a disease-specific signature.

List of Figures

Figure 1. (A) Representation of disulfide bond formation and isomerization by protein disulfide isomerase (PDI). (B) Regeneration of PDI by the endoplasmic reticulum oxidoreductin 1 (ERO1). After ERO1 transfers its disulfide bonds to PDI, the electrons are assimilated by FAD, and an oxygen molecule is disproportioned to hydrogen peroxide (H_2O_2). Figures based on the ones presented by Chakravarthi and coworkers (2006). …….7

Figure 2. Three-dimensional structure of the metallothionein (MTA) from sea urchin. (A) Alpha domain (PDB number 1QJK). (B) Beta domain (PDB number 1QJL). The spheres represent the metals ions, Cd^{2+} associated to each domain (Figure obtained using PyMOL software). …….9

Figure I.1. Scheme for Poros 20 NH SPDP system. Poros 20 NH is conjugated to the *N*-succinimidyl 3-(2-pyridyldithio) propionate, and peptides containing sulfhydryl groups can react with resin system. The peptides containing cysteine can be recovered using a reducing agent to cleave SPDP-peptide disulfide bonds. ……..…….24

Figure I.2. (A) MALDI-TOF mass spectrum of a peptide mixture. The molecular masses correspond to DSIP (849.9 Da), Bradykinin (1060.5 Da), Angiotensin I (1296.6), and cortistain (1721.0), respectively. (B) Mass spectrum of the eluted fraction from the Poros 20 NH SPDP resin. Only one peptide could be observed as the reduced form of cortistatin. ……...……...…25

Figure I.3. MALDI-TOFMS spectra from a mixture of synthetic peptides, angiotensin (1296.6 Da) and a synthetic cysteine-containing peptide (1338.7 Da). (A) a mixture 1:10 of the cysteine containing peptide and angiotensin I; (B) the eluted fraction obtained after incubation with the resin Poros 20 NH SPDP of 100 pmol of angiotensin I and 10 pmol of Cys-peptide; (C) the flow-through fraction after Poros 20 NH SPDP purification: only the angiotensin I was observed; (D) the eluted

fraction obtained after incubation with the resin Poros 20 NH SPDP of 1 nmol of angiotensin I and 10 pmol of Cys-peptide. ……..….27

Figure I-4. Nano-LC profile of the tryptic fragments of bovine serum albumin monitored at 220 nm. Line I indicates the flow-through fraction from the Poros 20 NH SPDP system. Line II represents the eluted fraction, which contains Cys-peptides. The dotted line represents the acetonitrile concentration used for peptide elution. The (*) indicates the fraction where the spiked peptide (GACL*LPKIETMR, L*= +7 Da) was observed. ....……...28

Figure I.5. Bovine serum albumin (BSA) amino acid sequence. In bold are the fragments observed in the eluted fraction from the Poros 20 NH SPDP resin. These fragments covered 70% of sequence and ≈60% of the tryptic peptides which contains cysteine in the sequence. …….31

Figure II.1. Cysteine protease papain three dimensional structure (PDB 9PAP). (A) Figure representing the endopeptidase surface. The Cys25 is highlighted in yellow inside the active pocket. (B) Cartoon of papain structure the Cys25 is presented in yellow. Figures obtained using PyMOL software. ……...……41

Figure II.2. SDS-PAGE 12.5% of *C. papaya* latex. 0.6 mg of dried latex was applied to the system. One third of each sample (total volume) was used on the SDS-PAGE. MM is the molecular mass standards. Line 1 and 2 represent the sample applied to SPDP system, flow-through fraction and eluted fraction, respectively. Line 3 and 4 represent the sample applied to the Poros 20 NH LC-SPDP system, flow-through and eluted fraction, respectively. The recovery of papain was higher when the LC-SPDP is used. ……..……46

Figure II.3. SDS-PAGE 12.5% of *C. papaya* latex. In this case, the concentration of the dried latex used was reduced; 0.3 mg was applied to the system and one third of each sample (total volume)

was used on the SDS-PAGE. MM is the molecular mass standards. Line 1 represents the flow-through fraction from Poros 20 NH LC-SPDP system. Line 2 represents the eluted fraction from Poros 20 NH LC-SPDP system where purified papain can be observed. ……47

Figure II.4. SDS-PAGE 12.5% of *C. papaya* latex sample used on enzymatic assay. MM is the molecular mass standards. Line 1 represents the flow-through fraction. Line 2 represents the eluted fraction from Poros 20 NH LC-SPDP system. ……48

Figure II.5. SDS-PAGE 12.5% of *C. papaya* latex. In this experiment, different times for linkage were evaluated. The sample (3 mg of dried latex) was exposed to the system for 2, 4 and 7 hours. At the seventh hour, the concentration of the active form of papain recovered was higher comparing to 2 and 4 hours. ……49

Figure III.1. Three dimensional structure of human wild type TTR dimer. The eight strands are indicated by the letters A-H. Figure obtained using PyMOL software (PBD number 3W3B). …...61

Figure III.2. Three dimensional structure of human wild type TTR tetramer. Each TTR monomer is highlighted by different colors. In the center, the channel is presented, where the thyroid hormones are transported (PDB 3W3B and 1DVQ). Figure obtained using PyMOL software. ….62

Figure III.3. Three dimensional structure of human TTR amyloid fibrils (10 amino acids; 105-115). The fragments adopt a β-strand conformation. Figure obtained using PyMOL software (PDB number 1RVS). ……64

Figure III.4. Effects of the oxidative stress in organism. Some diseases caused by the presence of free radicals or reactive oxygen species in the organs. This figure was modified, based on the scheme from the Online Resources for Disorders Caused by Oxidative Stress (**http://www.oxidativestressresource.org/**). ..……..68

Figure III.5. Identification of transthyretin in human plasma. (A) SDS-PAGE 12.5% of human plasma. MM is the molecular mass standards. Lane 1 represents the Flow-through fraction from the Poros 20 NH LC-SPDP system. Lane 2 represents the Eluted fraction from Poros-LC-SPDP system. The protein identified by mass spectrometry is indicated by the arrow. (B) Mascot data sheet of the protein identified as TTR. ……..……78

Figure III.6. SDS-PAGE 12.5% of human plasma. MM is de molecular mass standards. Lane 1 represents the Flow-through fraction from the Poros 20 NH LC-SPDP system. Lane 2: the Eluted fraction from the Poros 20 NH LC-SPDP system. The sample was not boiled at 90°C before applying it to the electrophoresis. The arrow indicates the TTR dimer. Lane 3: the Eluted fraction from the Poros 20 NH LC-SPDP system boiled at 90°C. …….79

Figure III.7. MALDI-TOFMS spectrum of the Eluted fraction obtained from Poros-LC-SPDP system. Each peak corresponds to TTR: monomer (13 761 Da), dimer (27 522 Da), trimer (41 283 Da), and tetramer (55 044 Da). .……80

Figure III.8. SDS-PAGE 12.5% of human plasma. In this experiment the effect of the reducing agent on the plasma proteins was evaluated using the Poros 20 NH LC-SPDP system. The sample has been treated with 0, 0.1, and 0.5 mM DTT before applying it to the Poros 20 NH system. …...81

Figure III.9. Western blot analysis of the plasma sample with anti-TTR antibody (rabbit polyclonal antibody). The samples were treated with (right panel) and without (left panel) 0.1 mM DTT before

applying the sample to the Poros 20 NH LC-SPDP system. The overall concentration of TTR increased after pretreatment with the reducing agent. Lane I and II showed the Flow-through and Eluted fraction, respectively. ..…….82

Figure III. 10. MALDI-TOFMS spectra of the Flow-through fraction obtained from the Poros 20 NH LC-SPDP system. (A) The Flow-through fraction, the modified forms of TTR, can be observed in the spectra. (B) Flow-through fraction treated with 5 mM DTT for 60 min at 60°C. After the process, only the free thiol form of TTR can be observed. ……..84

Figure III.11. SDS-PAGE 12.5% of human plasma. MM is the molecular mass standards. Lane 1: Flow-through fraction; lane 2: Eluted fraction; lane 3: Eluted fraction prepared from the sample treated with 0.1 mM DTT; lane 4: Eluted fraction prepared from the sample treated with 50 mM NEM at room temperature for 2 hours. ……85

Figure III.12. Optimization of incubation time for linking TTR to the Poros 20 NH LC-SPDP resin. The TTR prepared from the Eluted fraction was applied to the resin, and after 2, 4, 7 hours of incubation, the TTR was eluted with the reducing solvent and applied to SDS-PAGE. The TTR was visualized by Western blotting (see Figure. III.9). .……..86

Figure III.13. Effect of H_2O_2 treatment on the recovery of the free-Cys form of TTR. SDS-PAGE 12.5% of the Eluted fractions obtained from the human plasma. The human plasma was treated with 2mM, 5mM and 10mM H_2O_2, pre-purified, and applied to the Poros 20 NH LC-SPDP system resulting in the Flow-through and Eluted fractions (see Materials and Methods). ……..87

Figure III.14. Western blot analysis with anti-TTR antibody of the Flow-through (I) and Eluted (II) fractions obtained Fig. III.14. The free-Cys form of TTR completely disappeared after 10 mM H_2O_2 treatment. …….88

Figure III.15. Isolation of free thiol containing proteins from human plasma. (A) Comparison between the Poros 20 NH LC-SPDP system (lane 1 and 2) and Thiopropyl Sepharose 6B (lane 3 and 4). Lanes 1 and 3 were the Flow-through fractions; lanes 2 and 4 were the Eluted fractions. MM is the molecular mass standards. (B) Comparison between Poros 20 NH LC-SPDP (lane 1 and 2) and Toyo Pearl AF-Amino 650 LC-SPDP (lane 3 and 4). Lanes 1 and 3 were the Flow-through fractions; lanes 2 and 4 were the Eluted fractions. MM is the molecular mass standards. ……..89

Figure III.16. Three dimensional structures from papain (PDB 9PAP) (A) and TTR dimer (PDB 3W3B) (B). The free-Cys in both proteins (Cys25 and C10, respectively) are highlighted in yellow. Figures obtained using the PyMOL software. ...……..93

List of Tables

Table I.1. Peptides containing Cys observed in the eluted fraction in Figure I-4. The cysteine residues are depicted in bold. ……...30

Table II.1. Papain enzymatic assay using peptidyl-*p*NA. Three hundred seventy four micrograms of papain extract was applied to the Poros 20 NH LC-SPDP system and the same concentration was used as the total enzymatic activity for the assay. The values correspond to the amounts of *p*Na released in one minute. …….48

Table III.1. Protein concentrations, estimated by micro BCA protein assay kit, of the Flow-through fractions obtained in the experiment described in Figure III.14. ……..88

General Introduction

A hundred years before the DNA structure was elucidated by Watson and Crick, proteins had already been identified by scientists [1]. However, in the last fifteen years, many efforts have been made to identify and comprehend the genome of many organisms, mainly the human genome. The genome of more than seventy organisms has been sequenced [2, 3]. Although the gene sequences were identified, their function and interaction are not completely known yet [3]. Also, in genomic analysis, it has been observed that humans are not so "different" from other organisms, in the sense of the number of genes; therefore, the complexity of the organisms can be related not only to genes, but also to their products, i.e. the proteins. Splicing variation in mRNA, differential proteolysis and posttranslational modifications contribute to the wide variety of proteins' function *in vivo* [4]. In this sense, understanding the relationship between nucleotides and proteins and how they function and act in a complex system is one of the most promising fields in biological studies.

The proteins produced by the genome are understood as the focus of the proteomics studies. In the early 1900s, protein purification and sequencing were a laborious and time onerous job; protein sequencing was carried out manually and high amounts of the target protein were necessary for its identification [5]. Nowadays, the high throughput technologies, which are based on mass spectrometry, allow for a rapid and efficient analysis of proteins, which enable a comprehensive and system-wide analysis of proteins, using nanograms [6].The techniques essential for overall protein identification can be divided into two stages: separation and mass spectrometric techniques [1, 3]. The former includes chromatography, isoelectric focusing, electrophoresis, etc. The difference in chemical or physical proprieties of proteins is key to separate the desired protein. Especially, affinity chromatography using antibody or chemical tags designed for a specific structure has been successfully applied for separation of target proteins. The methods based on a chelating tag or metal ions that have high affinity for the phosphoryl group are often used for concentration or detection of phosphorylated peptides/proteins [7]. The mass spectrometry, in combination with sequence database search, has been a major method for protein identification.

Even with the cutting-edge technologies, there are still essential issues in covering whole proteins present in a sample. One of the intrinsic issues is the complexity of proteins in terms of a variety and abundance. The low-abundant proteins are hardly detected in the sample containing

high-abundant proteins. Nowadays, the proteomic analysis is often conducted for some specific character of proteins or pre-fractionated sample to obtain a more detailed protein profile [6]; e.g. since cytokines in the plasma are quite low abundant (one milliliter of plasma would provide 5 pg of interleukine 6, a marker for infections and inflammation processes), they could be profiled only after removing house-keeping proteins such as albumin, which constitutes 55% of the total amount of proteins present in the serum (35-50 mg/mL) [8]. To access the low abundant proteins, it is necessary to remove the abundant compounds by a serial pre-fractionation of the sample using different techniques, even though the multiple steps in the purification process can cause loss of proteins to some extent [6]. In this regard, improving proteomics methods and technologies could lead to an increase of data, which consistently provides meaningful information of biological specimen such as a cell, tissue, etc. [4].

The sulfur-amino acid, cysteine

Among the 20 amino acids produced by an organism, methionine and cysteine are the sulfur-containing amino acids; the other amino acids are composed of only carbon, oxygen, hydrogen and nitrogen. Oxygen and sulfur atoms belong to the same group on the periodic table, i.e. they can make the same number of covalent linkages; however, the difference in electronegativity makes the sulfur atom a versatile element: in the case of a protein, it plays an important role in displaying the protein's chemical proprieties; e.g. the replacement of the sulfur by oxygen atom in methionine would produce a less hydrophobic amino acid [9], the replacement of selenocystein (a selenium containing amino acid) by cysteine (Cys), which is located in the catalytic site, often diminishes the enzyme function. If the Sec is mutated with any other amino acid, the protein becomes inactive [10]. Moreover, reactive oxygen species (ROS), which react with various cellular elements and the sulfur amino acids, play important roles for maintaining the homeostasis in the environment. e.g. the phagocytes produce hypochlourus acid (HOCl) during the defense response and this compound reacts faster, in two orders of magnitude, with cysteine and methionine than the other amino acids [11].

Cys has a reactive thiol group that can suffer from various modifications such as oxidation to sulfenic (-SOH), sulfinic ($-SO_2H$), and sulfonic acid ($-SO_3H$) [12]. This amino acid may behave as a hydrophobic compound inside a protein structure, or a hydrophilic compound when it is exposed on the protein surface or in a polar microenvironment as the catalytic pocket [10]. The pKa of the cysteine is affected by the environment and the nearby amino acids — in the cellular milieu, it can vary up to 6 orders of magnitude. Under physiological conditions, free Cys has pKa of 8.5, oxidative modifications are unlikely to happen; the lower pKa of 7.2of the –SH group makes it thiolated or reactive to electrophiles [11, 13, 14]. The Cys residues exposed on protein surface, acting as nucleophile, tend to be modified and change their proprieties and the surrounding areas [10]. The oxidation of the thiol groups, occurring in protein synthesis in a cell, mainly produce disulfide bonds (S-S), which are important for protein folding and structural stability. S-glutathionylation (protein-SSG) is known to be a redox intermediate in disulfide formation, and S-nitrosylation (-SNO) is important for regulation of active NO [15]. This reactivity is useful for enzyme activity, controlling the regulatory process and linking the enzyme to coenzymes or prosthetic groups [13, 14]. A cluster of Cys residues in a protein, which are found in metallothionein, are located in a site for metal-binding: many metals and metalloids from outside can easily be trapped and detoxitised with the sulfur amino acids [10].

Cys is the second least abundant amino acid. However, they are often conserved in amino acid sequences among species, and play important roles for protein functions as active sites, regulatory sites, structure-stabilizing disulfides, cofactor binding, etc. The importance of Cys has also been shown by the fact that its mutations into other amino acids can frequently trigger dysfunctions of proteins or genetic diseases. Accordingly, Cys in a protein is strictly under evolutionary control [10].

Functional categories of Cys

1. Structural cysteine

The disulfide bonds are one of the most used strategies to stabilize protein structure. They are present in secreted proteins and in proteins resident in oxidizing environment such as mitochondria interspace and endoplasmic reticulum (ER). Meanwhile, transitory disulfide bonds can be formed in some cellular milieu, mainly in reducing conditions [10].

Intramolecular disulfide allows the linkage between distant sections – in the linear sequence – of the polypeptide chain; after the bonding there are 2.5 Å of distance between the residues [10]. The disulfide bond can be formed not only between two Cys in the same polypeptide chain, but also between Cys in two distinct peptides – an intermolecular disulfide link. The paring of two Cys, intra or intermolecular, can produce diverse disulfide networks and the complexity of the network is proportional to the number of Cys involved in the process [16]. An active protein has a specific and unique disulfide bond arrangement and there is a chemical and enzymatic control, in the cell, of the correct Cys parings. However, after the polypeptide synthesis, to reduce the entropy of the unfolded state, the environmental conditions can promote the incorrect bonding between two cysteines, which compromises protein activity [16, 17].

From bacteria to humans, a complex mechanism has been developed to control the protein folding via disulfide bonds. Archea and thermophilic bacteria encode a protein disulfide oxidoreductase (PDO) that catalyzes the disulfide bond formation in the cytoplasm. In other bacteria, the process occurs in the periplasm and involves four enzymes, DbsA, DsbB, DsbC, and DsbG. DsbA introduces the disulfide bond in proteins, meanwhile DsbB reoxidizes the DsbA. However, the disulfide bonds produced by DbsA are formed in a non-specific manner, and DsbC and DsbG are responsible for breaking and rearranging the bonds to correct ones [18].

In eukaryotic cells, cytosol, mitochondria and chloroplast have their own mechanisms for folding. However, it is in the ER that the folding process generally occurs, before the protein translocation in the cell or its secretion [18]. The ER has an ion concentration and oxidizing capabilities similar to the extracellular milieu; a redox state that allows the link between the Cys and also the reducing power to avoid incorrect bindings. This process is supported by some proteins that ensure the correct bonding formation and breaking of the erroneous disulfide [17, 19]. The protein

disulfide isomerase (PDI) and the endoplasmic reticulum oxidoreductin 1 (ERO1) are responsible for the correct bonding of the cysteines present in one protein. PDI is a 55 kDa protein and has five domains (a, a', b, b' and c). a and a' are homologous to thioredoxin and each domain has an active site where two cysteines can be found. The third and fourth domains, b and b', are inactive and have similar sequences. The last domain is the protein C-terminal and it has a positive charge [18, 19]. The sequence CGHC is present in a and a' domains and the two cysteine in the sequence can form an intramolecular bond or exist in the reduced form. When PDI is oxidized, it inserts disulfide bonds in a protein substrate; both active sites are involved in this process. On the other hand, to reduce a disulfide bond, only one site catalyzes the reaction. In the case of erroneous disulfide bond formation, the PDI acts on these links, breaking the substrate and rearranging the disulfide network (Figure 1). This process is called isomerization and can also occur when the correct disulfide bond is formed, if the pairing prevents the correct protein folding [19].

PDI activity is regulated by ERO1; this glycosylated flavoenzyme is associated with the ER membrane and oxidizes PDI, transferring its disulfide bonds to the disulfide isomerase, and the PDI can now oxidize the substrate again. This mechanism uses the oxygen and the cofactor flavin (FAD) to generate disulfide bonds. ERO1 has four cysteines involved in the disulfide exchange reaction; two in the active site and the other ones in a shuttle site. The first site is near the redox site of FAD and the shuttle site is present in a flexible loop – on the protein surface – and positioned near the active site. The disulfide bonds in the PDI are restructured by the activity of the cysteines present in the shuttle site; further, the cysteines in ERO1 site are re-oxidized by the Cys in the catalytic site. In the end, the sulfhydryl in the active site transfers its electrons to FAD and a molecule of hydrogen peroxide (H_2O_2) is formed (Figure 2) [20]. Furthermore, the ER environment controls the ERO1 activity, e.g. if the ER has an excess of oxidizing agents, the Cys will be linked by disulfide bonds and the active pocket of ERO1 will be blocked. In this case, the activity of ERO1 is reduced and the number of disulfide bonds formed by PDI likewise decreases [17, 18]. The process for activation and regulation of ERO1 is controlled by the amounts of glutathione (GSH) in the environment. Initially, GSH acts in ERO1 reducing the disulfide bond between Cys150-Cys295, which causes structural modifications on the other disulfide site, Cys90-Cys349, activating the enzyme. The concentration of GSH in the cytosol is enough to avoid disulfide bond formation. On the other hand, in the lumen the concentration of the oxidized glutathione (GSSG) is higher, which allows the activity of ERO1 and PDI [20, 21].

The final product of the disulfide bond formation, a molecule of H_2O_2, causes an oxidative stress in the cell; a high concentration of ROS can damage cellular components and induce apoptosis. It has been demonstrated that ERO1 is also controlled by the levels of H_2O_2 in the lumen. Decreasing ERO1 activity by RNA interference the amounts of ROS are reduced in the cell. In heavy secretory cells, 25% of the ROS present in the environment are generated from protein synthesis; in case of stress, the cells decrease mRNA translation and ERO1 activity to control ROS rates. Another mechanism to control ROS is associated to GSH. In case of stress, the synthesis of GSH can be amplified to eliminate these toxics compounds [21].

Due to Cys chemical proprieties, Cys residues are not normally present on protein surface. Using 15 000 protein sequences present in the protein data bank (PBD), Marino and Glaydshev (2012) observed that this sulfur amino acid is one of the least exposed residues; it can promote modifications in environmental conditions and physiological state. Usually, it is not present in mature proteins, unless this amino acid has a specific function in the molecule [10].

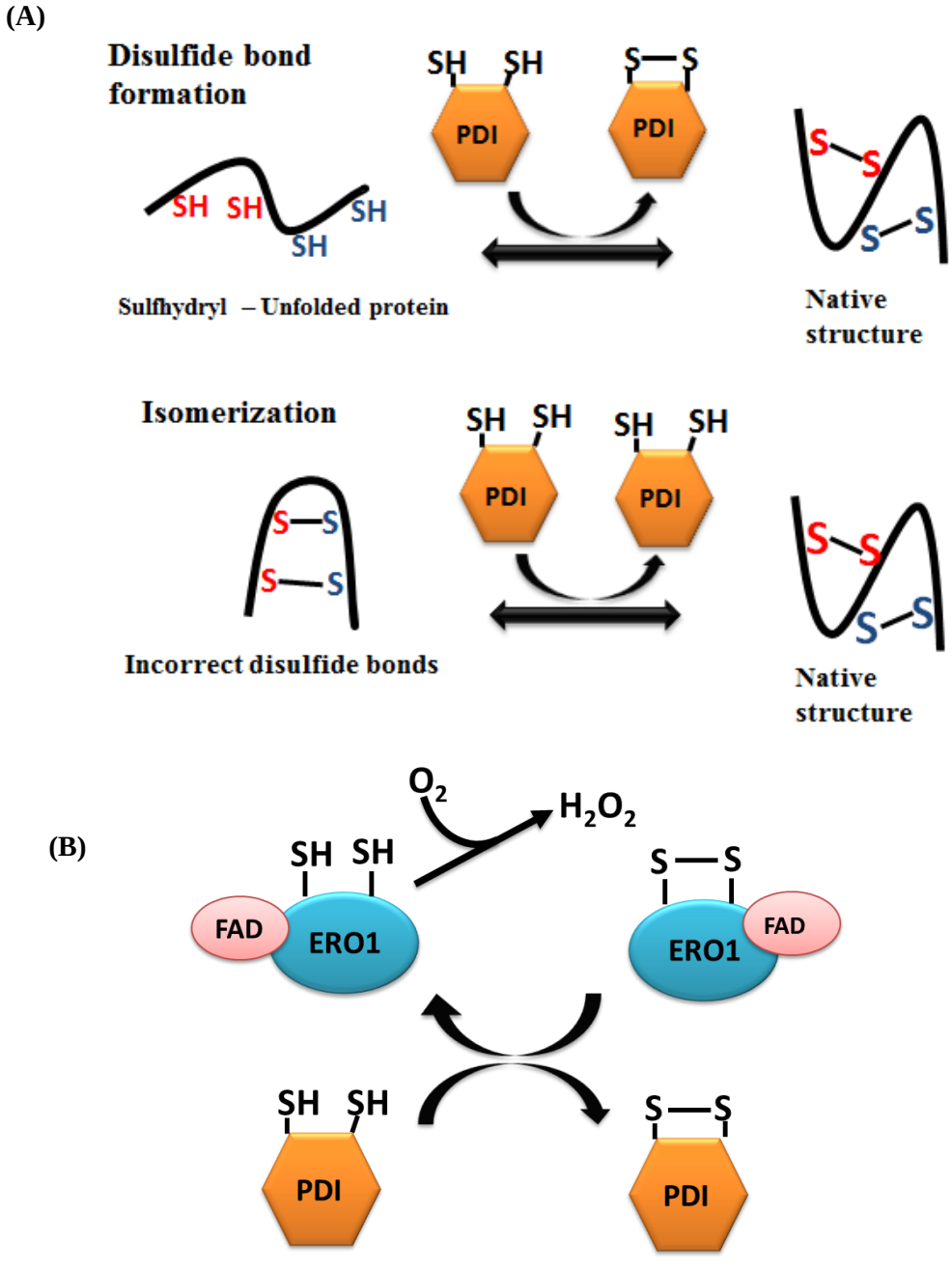


Figure 1. (A) Representation of disulfide bond formation and isomerization by protein disulfide isomerase (PDI). (B) Regeneration of PDI by the endoplasmic reticulum oxidoreductin 1 (ERO1). After ERO1 transfers its disulfide bonds to PDI, the electrons are assimilated by FAD, and an oxygen molecule is disproportioned to hydrogen peroxide (H_2O_2). Figures based on the ones presented by Chakravarthi and coworkers (2006) [21].

2. Metal-binding cysteine

Not only histidine (His), but also cysteine is used for metal coordination. Metals linked to proteins are useful to both structure stabilization and catalytic or regulatory sites [10]. Cys-rich proteins in bacteria, as in *Helicobacter pylori*, are responsible for Ni^{2+} homeostasis; the urease requires Ni^{2+} for its activity, converting urea in carbon dioxide and ammonia, which is essential for the infection process in the host [22].

Oxidized cysteine can form linkages to some metallic ions like iron and cobalt. Additionally, the thiolate can bind strongly to metal ions like Cu^{+}, Zn^{2+}, $Fe^{2+/3+}$, Cd^{2+}. This binding process is possible by virtue of changes in oxidation and molecule geometry. These characteristics allow this amino acid to link to different types of metal ions in the same cluster [23]. The high affinity of metals by the sulfur amino acid has been shown by Potocki and coworkers (2011); the N-termimal of ZIP (ZRT, IRT-like Protein) protein has a rich-Cys domain and three thiolates are coordinated to one metal ion in a stable pseudoplanar geometry [22]. Carbon monoxide dehydrogenase from *Carboxydothermus hydrogenoformans* has, in the active site, a [Ni-4Fe-5S] array, where each metal ion is linked to a separate Cys. In mammals, many metalloproteases and liver alcohol dehydrogenase (ADH) have zinc/sulfur complexes in the active site. If the metal ion is released from a protein, as a result of some oxidation process, the enzyme might be inactive; a strict redox control should be maintained to preserve enzyme activity [23]. ADH, in horses, represents a model for metal-binding to stabilize protein structure, in which Zn2+ is bound to Cys97, Cys100, Cys103, and Cys111 to form a tetrahedral arrangement [23].

Another class of metal-binding cysteines is the metallothioneins (MT). They are small Cys-rich peptides and the Cys present in the sequence do not form disulfide bonds (Figure 2). This group of peptides can bind to metal ions from group 11 and 12, reducing the pKa of the cysteines. The most studied MT is the mammalian MT, which has two domains, α and β. The first domain contains 9 cysteines and can link to 3 Cd^{2+} or Zn^{2+} or 6 Cu^{+} ions. The β-domain has 11 cysteine residues, which can coordinate 4 Cd^{2+} or Zn^{2+} or 6 Cu^{+} ions. These peptides were isolated in many organisms and due to the high contents of cysteine (~30%), they are involved in many biological processes such as metal ion homeostasis, detoxification of toxic metal ions and protection against oxidative stress. Toxic and nontoxic metal ion levels are controlled by the metallochaperones, transporting these ions in a controlled manner: e.g., many MTs are chaperones, transporting Zn^{2+} and Cu^{+} to apo-enzymes. About metal detoxification, aside from linking to metals from group 11 and 12, MTs

are capable of binding to Co^{2+}, Pb^{2+}, $Pt^{+2/4+}$, Fe^{2+}. In the cellular environment, these Cys-rich peptides preferentially coordinate many toxic metal ions. Additionally, the content of the sulfur amino acid allows MT to react and inhibit ROS. It has been shown that oxidative stress promotes expression of MT; in a myocardial injury, a cardiac MT is overexpressed to inhibit the perfusion of ischemia. Many structural studies have been conducted to understand these peptides and their interaction with metals; in rats, MT Cd_5Zn_2-MT-2 was analyzed by NMR and X-ray and it was observed that the cysteinyl-thiolates were tetrahedrally coordinated to ions. Metal ions of different sizes can be arranged in the cluster of the human Cd_7 and Cu_7-MT-2. Meanwhile, the $Cu_{(I)}$-MT has shown diagonal or trigonal geometries, depending on the $Cu_{(I)}$:MT ratio. The structural analysis of MT showed the effect caused by a metal ion in the peptide structure, such as the rearrangement of the peptide backbone and the coordination geometry of the metal [24].

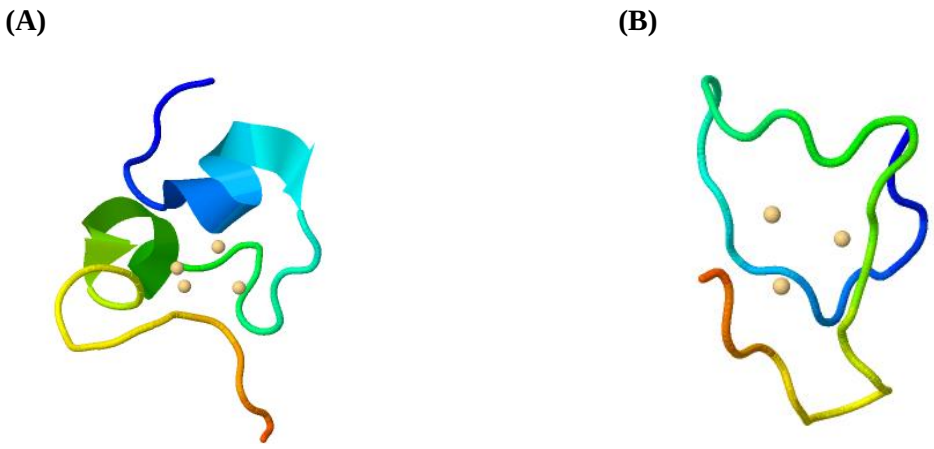


Figure 2. Three-dimensional structure of the metallothionein (MTA) from sea urchin. (A) Alpha domain (PDB number 1QJK). (B) Beta domain (PDB number 1QJL). The spheres represent the metals ions, Cd^{2+} associated to each domain (Figure obtained using PyMOL software).

The reactivity of metal ions, mainly zinc ions, to cysteine can also configure a problem in enzyme function. Hydrogenases and hydrolases have Cys in the active sites and when they are

linked to Zn or other divalent metals – such as Cd^{2+} and Cu^{+} – the enzyme activities are abolished. The nucleophile Cys represent potential ligands for theses metal ions [22, 23]. As the concentration of zinc ion in the cellular milieu is very low under normal conditions, those enzyme activities are not significantly affected. However, under oxidative stress, zinc ion is released into the cytosol, and zinc/sulfur bonds are formed at the active site of some proteins, such as caspase-3. This is a reversible and favorable reaction, since it rapidly protects the thiols from environmental oxidation that could modify the enzyme in an irreversible manner [23].

3. Regulatory cysteine

The cysteine residues are often involved in posttranslational modifications (PTM). Protein functions and its network interactions are modified by PTMs. Reaction with hydrogen sulfide (H_2S - sulfhydration), S-glutathionylation (RS-SG), sulfenic acid (Cys-SOH), *S*-nitrosylation (NO-Cys) and intra and intermolecular disulfide bonds are examples of Cys-PTMs [10, 25]. In general, these modifications are reversible and deeply involved with redox; the sulfur atom is electron rich and multiple oxidation states are allowed in its d-orbital [25, 26]. Moreover, the oxidative posttranslational modifications (Ox-PTMs) are usually induced by ROS, reactive nitrogen species (RNS) and small disulfide molecules; they are involved in different regulatory processes and diverse modifications can happen at the same Cys residue [25, 27].

Sulfhydration is reported in few proteins: a persulfide bond (R-SSH) is generated and the Cys becomes more reactive. It was reported that some enzymes become active upon this sufhydration with H_2S, such as the glyceraldehyde- 3-phosphate dehydrogenase (GAPDH) which has an increase of 700% in glycolytic activity after sulfhydration [28]. 25-50% of hepatic proteins are sulfhydrated. Also, when the protein p65 from the nuclear factor-κB (NF-κB) is modified by H_2S, the expression of the pro-survival genes is initiated and the apoptosis is inhibited in cells even in stress conditions [25, 28].

The S-glutathionylation is the reaction between a free thiol and oxidized glutathione (GSSG) or S-nitrosoglutathione (GSNO). Previously, this PTM was known to be related only to Cys residue protection against superoxidation, but it has been shown to directly affect some kinase and

phosphatase activities, presenting a strong correlation between phosphorylation and S-glutathionylation process [25, 29]. Protein kinase A, creatine kinase, protein kinase C, serine/threonine phosphatase, and protein tyrosine phosphatase are the candidates regulated by glutathionylation. In addition, some calcium binding proteins, members of S100 family, are targets for S-glutathionylation and they are up-regulated in tumor progression. S100A1 can be modified at Cys85 and its affinity for calcium increases in 10-folds. The S100B, a marker for malignant melanoma, has a susceptible glutathionylation site, at the Cys84 [29].

The S-nitrosylation protects important Cys residues from irreversible oxidative stress. Many authors have demonstrated the role played by NO-Cys in cardiovascular protection. In ischemic processes, this modification shields proteins from oxidative damages and the cells can restore their normal functions in a shorter time. It was observed during heart ischemia that the cellular mechanisms modify the Cys residues with NO and lower protein degradation level due to irreversible oxidation, and protein dysfunctional activity is reduced. More than one thousand proteins were described as being posttranslationally regulated and modified by NO. The Cys residues that suffer from the S-nitrosylation are specific and most of these proteins have only a single modification. Furthermore, NO-modified proteins participate in many cellular functions, such as energy production (mitochondria function and ion channel proteins) and protein localization [30, 31].

Sulfenylation (SOH) is an Ox-PTM. For a long time, it was considered a harmful modification in Cys residues; this PTM is not stable and can cause conformational changes in structure and modulation of protein activity, occurring mainly in cysteine proteases, such as deubiquitinating enzymes [25, 26]. It has also been reported that the oxidation of this amino acid in the active pocket inactivates the protein during cellular stress. Once the Cys is modified to be -SOH form, however, it can be readily oxidized to a more stable form ($-SO_2H$): it was reported that only 5% of the Cys was remained as -SOH. In many cases, when proteins are modified to be -SOH and $-SO_2H$ forms, their half-life is reduced to a few minutes [15, 26].

4. Catalytic cysteine

Many enzymes have free Cys residues in the active pockets. These enzymes have important roles in cell cycle regulation, apoptosis and signal transduction [32, 33]. Cys-enzymes, in general, are inhibited by mild oxidative conditions and could be reduced and recovered, *in vivo*, by glutathione or thioredoxin [10, 32]. Oxidoreductases, transferases, hydrolases and isomarases are examples of enzymes, all of which possess free Cys in their active sites. This residue is conserved and the catalytic cysteine is necessary to form a disulfide bond and/or a sulfenyl intermediate in oxidoreductases. The nucleophilic propriety of the sulfur atom is essential to phosphatases, protein disulfide isomarases and cysteine proteases [32]. The catalytic Cys can be subdivided into two groups: the catalytic non-redox cysteines and the catalytic redox cysteines [33]. In the first group, the Cys acts like a nucleophile in the catalysis process. During the reaction, the redox state of Cys is interchangeable. Those Cys residues are conserved and present in many protein families, such as the Cys peptidases, dCMP hydroxymethylases and protein-tyrosine phosphatases [10, 33]. In the second group, cysteine is involved in the oxidation, reduction and isomerization of disulfide bonds or in other kinds of reaction in which changes in the redox state of the residues occur. The best-known representative of this group is the oxidoreductase family [10].

References

[1] Soloviev, M., Finch, P., Peptidomics, current status. *J Chromatogr B Analyt Technol Biomed Life Sci* 2005, *815*, 11-24.

[2] Clynen, E., De Loof, A., Schoofs, L., The use of peptidomics in endocrine research. *Gen Comp Endocrinol* 2003, *132*, 1-9.

[3] Baggerman, G., Verleyen, P., Clynen, E., Huybrechts, J., *et al.*, Peptidomics. *J Chromatogr B Analyt Technol Biomed Life Sci* 2004, *803*, 3-16.

[4] Soloviev, M., Shaw, C., Andren, P., *Peptidomics: Methods and Applications*, John Wiley & Sons, Inc., Hoboken, New Jersey 2008.

[5] James, P., Protein identification in the post-genome era: the rapid rise of proteomics. *Q Rev Biophys* 1997, *30*, 279-331.

[6] Brewis, I. A., Brennan, P., Proteomics technologies for the global identification and quantification of proteins. *Adv Protein Chem Struct Biol* 2010, *80*, 1-44.

[7] Chen, X., Wu, D., Zhao, Y., Wong, B. H. C., Guo, L., Increasing phosphoproteome coverage and identification of phosphorylation motifs through combination of different HPLC fractionation methods. *Journal of Chromatography B* 2011, *879*, 25-34.

[8] Anderson, N. L., Anderson, N. G., The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 2002, *1*, 845-867.

[9] Brosnan, J. T., Brosnan, M. E., The sulfur-containing amino acids: an overview. *J Nutr* 2006, *136*, 1636S-1640S.

[10] Marino, S. M., Gladyshev, V. N., Analysis and functional prediction of reactive cysteine residues. *J Biol Chem* 2012, *287*, 4419-4425.

[11] Held, J. M., Gibson, B. W., Regulatory control or oxidative damage? Proteomic approaches to interrogate the role of cysteine oxidation status in biological processes. *Mol Cell Proteomics* 2012, *11*, R111.013037.

[12] Chang, Y. C., Huang, C. N., Lin, C. H., Chang, H. C., Wu, C. C., Mapping protein cysteine sulfonic acid modifications with specific enrichment and mass spectrometry: an integrated approach to explore the cysteine oxidation. *Proteomics* 2010, *10*, 2961-2971.

[13] Lipton, S. A., Choi, Y. B., Takahashi, H., Zhang, D., *et al.*, Cysteine regulation of protein function--as exemplified by NMDA-receptor modulation. *Trends Neurosci* 2002, *25*, 474-480.

[14] Fahey, R. C., Hunt, J. S., Windham, G. C., On the cysteine and cystine content of proteins. Differences between intracellular and extracellular proteins. *J Mol Evol* 1977, *10*, 155-160.

[15] Cai, Z., Yan, L. J., Protein Oxidative Modifications: Beneficial Roles in Disease and Health. *J Biochem Pharmacol Res* 2013, *1*, 15-26.

[16] Akaji, K., Kiso, Y., *Houben-Weyl, Methods of Organic Chemistry, Synthesis of Peptides and Peptidomimetics*, Thieme, Stuttgart 2002.

[17] Feige, M. J., Hendershot, L. M., Disulfide bonds in ER protein folding and homeostasis. *Curr Opin Cell Biol* 2011, *23*, 167-175.

[18] Heras, B., Kurz, M., Shouldice, S. R., Martin, J. L., The name's bond.....disulfide bond. *Curr Opin Struct Biol* 2007, *17*, 691-698.

[19] Wilkinson, B., Gilbert, H. F., Protein disulfide isomerase. *Biochim Biophys Acta* 2004, *1699*, 35-44.

[20] Sevier, C. S., Kaiser, C. A., Ero1 and redox homeostasis in the endoplasmic reticulum. *Biochim Biophys Acta* 2008, *1783*, 549-556.

[21] Chakravarthi, S., Jessop, C. E., Bulleid, N. J., The role of glutathione in disulphide bond formation and endoplasmic-reticulum-generated oxidative stress. *EMBO Rep* 2006, *7*, 271-275.

[22] Potocki, S., Rowinska-Zyrek, M., Valensin, D., Krzywoszynska, K., *et al.*, Metal binding ability of cysteine-rich peptide domain of ZIP13 Zn2+ ions transporter. *Inorg Chem* 2011, *50*, 6135-6145.

[23] Giles, N. M., Watts, A. B., Giles, G. I., Fry, F. H., *et al.*, Metal and redox modulation of cysteine protein function. *Chem Biol* 2003, *10*, 677-693.

[24] Sutherland, D. E., Stillman, M. J., The "magic numbers" of metallothionein. *Metallomics* 2011, *3*, 444-463.

[25] Chung, H. S., Wang, S. B., Venkatraman, V., Murray, C. I., Van Eyk, J. E., Cysteine oxidative posttranslational modifications: emerging regulation in the cardiovascular system. *Circ Res* 2013, *112*, 382-392.

[26] Lo Conte, M., Carroll, K. S., The redox biochemistry of protein sulfenylation and sulfinylation. *J Biol Chem* 2013, *288*, 26480-26488.

[27] Di Simplicio, P., Franconi, F., Frosalí, S., Di Giuseppe, D., Thiolation and nitrosation of cysteines in biological fluids and cells. *Amino Acids* 2003, *25*, 323-339.

[28] Paul, B. D., Snyder, S. H., H_2S signalling through protein sulfhydration and beyond. *Nat Rev Mol Cell Biol* 2012, *13*, 499-507.

[29] Pastore, A., Piemonte, F., S-Glutathionylation signaling in cell biology: progress and prospects. *Eur J Pharm Sci* 2012, *46*, 279-292.

[30] Wiseman, D. A., Thurmond, D. C., The good and bad effects of cysteine S-nitrosylation and tyrosine nitration upon insulin exocytosis: a balancing act. *Curr Diabetes Rev* 2012, *8*, 303-315.

[31] Murphy, E., Kohr, M., Menazza, S., Nguyen, T., *et al.*, Signaling by S-nitrosylation in the heart. *J Mol Cell Cardiol* 2014.

[32] Nagahara, N., Catalytic site cysteines of thiol enzyme: sulfurtransferases. *J Amino Acids* 2011, *2011*, 709404.

[33] Fomenko, D. E., Marino, S. M., Gladyshev, V. N., Functional diversity of cysteine residues in proteins and unique features of catalytic redox-active cysteines in thiol oxidoreductases. *Mol Cells* 2008, *26*, 228-235.

Chapter I. A method for isolating cysteine containing peptides

I.1. Introduction

Around 280 million years ago, the transition of a reducing to an oxidizing atmosphere had a direct effect upon the redox-state of the cysteine present in the protein sequence. Evolutional analysis presented that only after the emergence of aerobic microorganisms extracellular proteins could contain cysteine, and proteins with disulfide bonds were more expressed in aerobic microorganisms [1].Cysteine is characterized as a low mutability amino acid; it has rarely mutated during the course of evolution. Its frequency has increased over time: from 0.41% to 2.26% among the encoded amino acids [2, 3]. Furthermore, the toxicity of miscoding of this sulfur amino acid is low, which allows cysteine to be a suitable substitution at most positions in the protein sequence [1]. Some hypotheses state hat this amino acid was introduced recently in organism evolution; few cysteines are found in simple organisms (such as bacteria and fungus) and hence a limited number of protein domains containing this amino acid could be observed in these organisms. Studies about cysteine contents in all phyla report that mammals have four to five times more cysteine in the protein sequence than *Archea* [1]. Conversely, *Archea* as *Methanobacterium* and *Pyrococcus* have significant cysteine clustering in the proteome. This can be explained by environmental conditions – the optimal grown temperature for *Archea* is 103°C, and the CXXC cluster is conserved because it is involved either in disulfide bonds or in coordination with hetero atoms, presenting high levels of thermostability [4].

Prokaryotes and eukaryotes administrate a thiol-mediate switch mechanism to regulate many processes in the cell. The reversible redox switches in proteins and peptides can be associated with physiological and pathological conditions [5]. In the case of some parasites, as *Trypanosoma cruzi*, *Trypanosoma brucei* and *Leshimania spp*., the infectivity is directly related to cysteine proteases. They play an important role in parasite nutrition, cell invasion, host immune evasion and virulence. e.g., the cysteine proteases identified in *T. cruzi* are members of the papain family, and they have an intra and extracellular distribution [6]. In fungus, a group of small hydrophobic proteins (~150 amino acids length), called hydrophobins, has been described; they contain 8 cysteine residues and are associated with fruiting body formation and

pathogenicity. Among the hydrophobins group, a diverse amino acid sequence is observed, while the cysteine residues are strongly conserved [7].

In vertebrates, a group of proteins named cysteine-rich secretory proteins (CRISPs) have 16 conserved cysteines distributed in two domains – they are associated in an intradomain linkage [8, 9]. CRISPs are mainly expressed in the male reproductive tract, but can also be found in skeletal muscles, thymus and secretory glands. In the reproductive tract, they are related to gametes maturation in epididymis and the fertilization process. Furthermore, they were identified in the venom of some reptiles and in the saliva of lamprey fish; these Cys-rich proteins act upon the ion-channels. In snakes, CRISPs are the major venom component; they block the ion chain after the snake attacks its prey [9].

Not only proteins, but also cysteine containing peptides play an important role in several systems. Database analysis of non-redundant proteins present many peptides which have a high frequency of lysine, cysteine and arginine, comparing to larger proteins [10]. Produced or processed by many biological pathways, cysteine-rich peptides have many functions and a diverse sequence. At the same time, they have an analogous and rigorous system to control their production, regulation and signaling. Among these bioactive peptides, there are the defensins, granulins and chemokine groups [11, 12]. Defensins are cysteine-rich peptides which have less than 100 amino acids in the sequence and contain three disulfide bonds and a β-hairpin [10]. Defensins are found in animals and plants, playing an important role in the regulation of intestinal microbiome and defense mechanisms against microorganisms [11]. Granulins are putative growth factors which control mesenchymal and epithelial cells; they have a molecular mass of 6 kDa and a highly conserved motif of 12 cysteines, arranged in four cysteine pairs surrounded by two singletons towards in C- and N-terminal [13]. The chemokines are proinflammatory mediators attracting many lineages of lymphocytes and leukocytes to sites of microbial invasion; these peptides usually have 70-80 amino acids in length and four conserved cysteine residues [14]. Defensins, granulins and chemokines are produced by mammal leukocytes and are described as possessing cytotoxic activity, antimicrobial and cytokine proprieties. The intrachain disulfide bonds present in these peptides structures are essential for their biological activity [12].

In short, the occurrence of cysteine in the organism may be defined by the balance between the beneficial and harmful reactions of the thiol groups in proteins, which is controlled by intracellular mechanisms to protect the organism against adverse reaction [15]. The human genome encodes around 214,000 cysteines and ~89% of the proteins have, at least, one cysteine

residue [2, 16, 17] . In this regard, methods and technologies targeted at Cys-containing peptides and proteins can provide useful information to investigating proteomes.

Conventional methods for isolating peptides containing cysteine

To analyze the cysteine in the proteome, most techniques adopt an enzymatic treatment to cleave proteins into small peptides; then, signature peptides are utilized to identify the protein [18]. Usually, quantitative analysis of proteins and peptides involve techniques as protein separation, two-dimensional polyacrylamide gel electrophoresis and mass spectrometry [19].

Aiming for the purification and/or isolation of thiol-containing compounds, in 1973 the "covalent chromatography" method was developed by Brocklehurst and coworkers. It is based on the thiol-disulfide interchange: thiol-containing compounds are immobilized on a matrix by the formation of disulfide bonds. For this purpose a thiol activated matrix can be used: an agarose-bound glutathione with 2,2'-dithiodipyridine (2-PDS - activated thiol-sepharose 4B) and/or an agarose reacted with mercaptohydroxypropyl ether and 2-PDS (thiopropyl sepharose 6B). When the disulfide bond is formed, the pyridine 2-thione is liberated and the reaction can be monitored by spectrophotometry at 343 nm [20, 21].

A cation-exchange reverse phase chromatography may also be applied to identify disulfide-linked peptides. In acidic conditions, these peptides present variations in cationic proprieties that allow a strong retention of cysteine during the process. Another method commonly used is the diagonal paper electrophoresis; in which case, the proteins are cleaved into peptides by pepsin, and the disulfide linked peptides are separated at acidic pH on the one dimension of paper electrophoresis. Then, the electrophoretogram is treated with performic acid vapor and the second dimension can be performed [22].

Indeed, among all the techniques applied to isolate cysteine containing peptides, the most famous is the isotope-coded affinity tag (ICAT) developed in 1999. The Cys-peptides are linked to a probe, which is composed of three parts: a reactive group, an isotopic coded linker and a tag. The first portion is a thiol-specific reactive group, an iodoactyl group; the second one is a

deuterated linker – and the biotin is the tag applied for affinity isolation. ICAT uses light and heavy deuterated linkers that can react with the cysteinyl groups in different samples. The heavy and the light labeled samples are combined and enzymatically cleaved to obtain peptide fragments. Therefore, the cysteine tagged fragments – a reduced number of peptides – are isolated by an avidin affinity chromatography. This method provides a quantitative analysis of the proteins containing Cys residue(s) [19].

Another efficient technique has been applied to isolate Cys containing peptides: the covalent capture. The method relies on the formation of a covalent bond between a peptide and a stabilized structure. The most used method for covalent capture is the disulfide bond formation. However, the reaction between the C-terminal thioester and the N-terminal of the amino acid cysteine is also useful; in this case, an amide bond is formed in the process [23]. The covalent capture can also be applied together with the tag method. In this case, the reduced proteins are alkylated with a Boc-ITC group (tert-butoxycarbonyl-N-{2-((2-iodoacetyl) amino)-ethyl}-1,3-thiazolidine-4-carboxamide) or a ATC (N-{2-((2-acryloyl)amino)ethyl}-1,3-thiazolidine-4-carboxamide). After blocking the Cys residues, the sample suffers an enzymatic digestion and the tag is activated; using an aldehyde resin, the tagged peptides can be recovered after an elution process with O-methylhydroxylamine. This method is called cysteine-reactive tandem mass tags [24].

Recently, a new method to isolate and quantify the peptides containing cysteine was proposed by Tambor and coworkers (2012). The method combines a stable isotope labeling, isobaric tags for relative and absolute quantitation (iTRAQ) and cysteinyl peptide enrichment. The CysTRAQ is based on enzymatic digestion of the proteins in the sample, which were previously labeled with iTRAQ reagents. Subsequently, the cysteine peptides are isolated by an enrichment method – using a covalent chromatography process. Ultimately, the Cys peptides are quantified by the relative abundances of the iTRAQ chemicals in the sample [25].

Even though all the techniques mentioned above represent an evolution in the process for protein and peptide identification, they also have limitations and disadvantages. e.g. even the newest proteomic techniques are not capable of fully comprehending a complex protein digestion; 4 orders of concentration range is the limit observed on any method [25]. In this sense, to increase the amount of information about biological systems and how they can be affected by environmental conditions, it is necessary to develop new approaches to analyze complex peptide mixtures.

I.2.Material and Methods

Materials

Bovine serum albumin (BSA), trypsin from porcine pancreas, acetonitrile and α-cyano-4-hydroxy cinnamic acid (CHCA) were purchased from Sigma (St. Louis, MO). The synthetic peptides cortistatin, bradykinin, DSIP and angiotensin I were purchased from Peptides (Osaka, Japan), and a marked peptide derived from BSA was purchased from Bio-Synthesis (Lewisville, TX). *N*-succinimidyl 3-(2-pyridyldithio) propionate (SPDP) was purchased from Pierce (Rockford, IL). The reducing agent, dithiothreitol (DTT) was purchased from Wako (Osaka, Japan). HF Bond elut-C18 columns were purchased from Varian (Walnut Creeck, CA). The resin for perfusion chromatography, Poros 20 NH Amine, was produced by PerSeptive Biosystems (Framingham, MA).

Preparation of Pyridyldithiol-coupled Resin

Poros 20 NH resin (20 mg) was washed with acetonitrile (100% and 50%) and pure water for the removal of impurities. The process was repeated twice for each solvent. After the cleaning process, twenty millimolar of SPDP was dissolved in ethanol and applied to the resin, which was incubated with the heterobifunctional group for 4 hours at room temperature. The excess of SPDP was removed washing the resin with ethanol. The interaction between SPDP and the resin Poros 20 NH was evaluated based on the coefficient for pyridine 2-thoine extinction, according to the protocol proposed by the producer. A working solution was prepared containing 50 mM NH_4CO_3H pH 7.8, 25% ethanol, 100 mM NaCl and 1mM EDTA. Poros 20 NH resin and SPDP complex were washed three times using the working solution, before exposing the system to the sample.

Sample preparation

A standard mixture, containing 4 peptides, was prepared from a stock solution with 15 pmol of DSIP (WAGGDASGE), 30 pmol of bradykinin (RPPGFSPR), 25 pmol of angiotensin I (DRVYIHPFHL) and 20 pmol cortistain (PCKNFFWKTFSSCK). The disulfide bond present in the cortistatin was reduced using 10 mM of DTT in 50 mM NH_4CO_3H at 60° C for 1 hour. The reducing agent was removed using HF Bond-Elut C_{18}, and the peptide mixture was diluted in the working solution and then applied to the resin. The system was incubated for 3 hours at room temperature (batch incubation).

A tryptic digestion from BSA (25, 50 and 100 pmol) was prepared to evaluate the system against a complex peptide mixture. Previously, the protein structure was denatured and the disulfide bonds were reduced in 50 mM NH_4CO_3H containing 0.1% SDS and 10 mM of DTT at 60° C for 1 hour. The solution was exposed to the immobilized trypsin (20 pmol in 50mM NH_4CO_3H), aiming to produce peptide fragments, for 12 hours at 37° C. BSA fragments were subjected to HF Bond-Elut C_{18} to remove the DTT before applying this mixture to the resin. 20 pmol of the marked peptide (GACL*LPKIETMR; L*= +7 Da) was spiked in the sample. The tryptic fragments were suspended in working solution and incubated for 3 hours at room temperature (batch incubation).

Separation of Cysteine-containing peptides

After the linking process, the resin was centrifuged (2000xg for 10 seconds) and the supernatant recovered (flow-through fraction). The resin was extensively washed using working solution before recovering the peptides bound to the system. 50 mM DTT were suspended in 500 µL of working solution and used to recover the cysteine containing peptides (eluted fraction). The reducing solution was mixed with the resin for 1 hour at room temperature and supernatant was recovered. Both fractions were dried in a vacuum and desalted using HF Bond Elut-C18. The same process was applied to BSA peptides.

Nano-Flow Liquid Chromatography and Mass spectrometry

The flow-through and eluted fractions were analyzed in an Ultimate nano-LC system (Dionex, Idstein, Germany). A linear gradient of 0.1% trifluoroacetic acid in water (solvent A) and 0.1% of trifluoroacetic acid in acetonitrile (solvent B) were used to elute peptides separated from the chromatographic system (5% to 80% of solvent B in 60 minutes). First, peptides were concentrated with a C18 trapping column (0.3 mm x1 mm, Dionex) at a flow rate of 20 µL/min, and the separation occurred in a C18-Pepmap column (0.075 x 150 mm, Dionex) at a flow rate of 200 nL/min. Peptides obtained after the reverse phase chromatographic process were directly spoted at 1 minute intervals onto a MALDI sample plate over a 96 min period. After this step, a matrix solution (CHCA in 0.1% trifluoracetic acid/ aqueous 50% acetonitrile) was blotted manually onto each sample spot and dried. The MALDI-TOF/TOF instrument (4800 proteomics analyzer, Applied Biosystems, Framingham, MA) was operated to obtain peptide fragments. The instrument was calibrated using a mixture of four peptides: angiotensin I (*m/z* 1296.6), dynorphin (*m/z* 1604.0), ACTH 1-24 (*m/z* 2932.6), and β-endorphin (*m/z* 3463.8). Mascot version 2.3 (Matrix Science, Manchester, UK) was the database employed to identify the fragments.

I.3.Results

The combination of SPDP and Poros 20 NH makes this system efficient for affinity chromatography (Figure I-1). To evaluate the number of linkers coupled to the amino group of Poros 20 NH, a test using the coefficient of extinction for pyridine-2-thione (8.08 x 10^3 $M^{-1}cm^{-1}$) was made. It was observed that ≈5 mmol of the crosslinker reacts with 20 mg of resin.

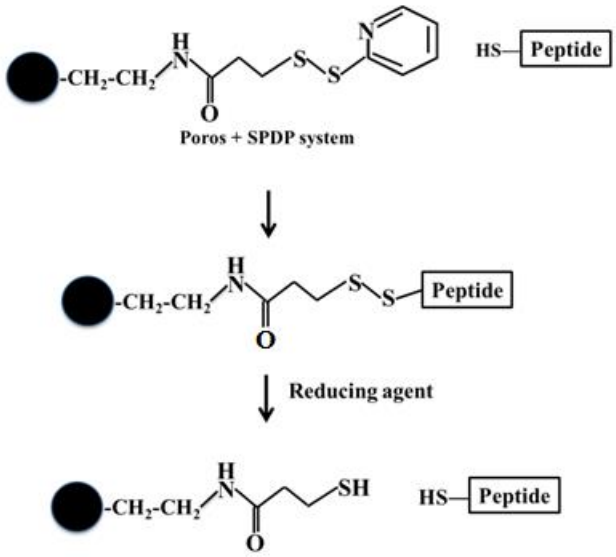


Figure I.1. Scheme for Poros 20 NH SPDP system. Poros 20 NH is conjugated to the *N*-succinimidyl 3-(2-pyridyldithio) propionate, and peptides containing sulfhydryl groups can react with resin system. The peptides containing cysteine can be recovered using a reducing agent to cleave SPDP-peptide disulfide bonds.

As a first test, a mixture of peptides was used to evaluate the Poros 20 NH SPDP system. Four peptides were used in different concentrations: 15 pmol DSPI (WAGGDASGE, *m/z* 849.9 Da), 30 pmol bradykinin (RPPGFSPR, *m/z* 1060.5 Da), 25 pmol angiotensin I (DRVYIHPFHL, *m/z* 1296.6) and 20 pmol cortistatin (PCKNFFWKTFSSCK, *m/z* 1721.0). Previously, the disulfide bond present in cortistatin was reduced using 10mM of DTT. Following the protocol

described in the material and methods section, the peptides were mixed and applied to the Poros 20 NH SPDP. Among the four standard peptides, only cortistatin, which has two cysteines in its sequence, interacted with the resin and was recovered after DTT elution (Figure I.2). Further, the flow-through fraction was analyzed by mass spectrometry and no trace of cortistatin was observed.

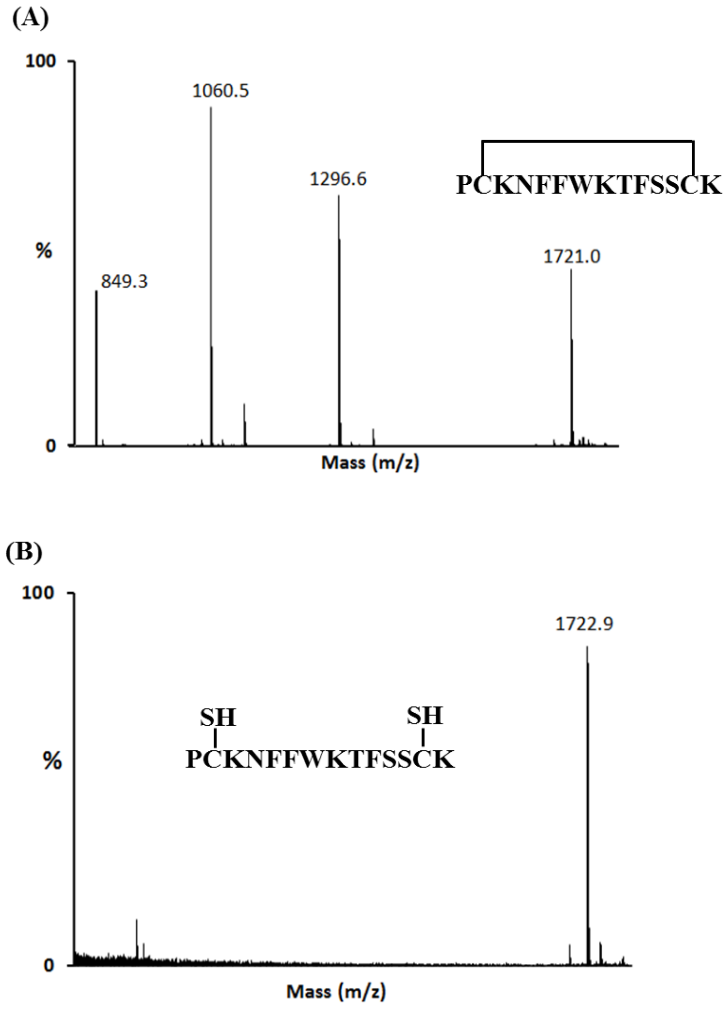


Figure I.2. (A) MALDI-TOF mass spectrum of a peptide mixture. The molecular masses correspond to DSIP (849.9 Da), Bradykinin (1060.5 Da), Angiotensin I (1296.6), and cortistain (1721.0), respectively. (B) Mass spectrum of the eluted fraction from the Poros 20 NH SPDP resin. Only one peptide could be observed as the reduced form of cortistatin.

Moreover, the system was evaluated regarding the sample charge capacity. The dynamic range of sample amounts applied for the system should be evaluated for the case of separation of low contents of target peptides in the sample: the following two peptides were used: angiotensin I (DRVYIHPFHL, *m/z* 1296.68) and a peptide containing one cysteine residue in its sequence (GACL*LPKIETMR, *m/z* 1338.67). The experiments were conducted using different concentrations of each peptide: (1) 100 pmol of angiotensin I and 10 pmol of the cysteine peptide and (2) 1 nmol of the angiotensin I and 1 pmol of the cysteine peptide. For the first experiment, a tenfold excess of angiotensin I was used (Figure I.3.A) and after the elution process, only the cysteine peptide was recovered (Figure I.3.B). The angiotensin I was observed only in flow-through fraction. In the second experiment, the concentration of angiotensin I was a hundredfold excess, against 10 pmol of the cysteine peptide. By mass spectrometry observations, in the original sample, the signal for the Cys-peptides is almost suppressed by the high concentration of angiotensin I (Figure I.3.C). However, after applying the sample to Poros 20 NH SPDP, the cysteine peptide could be recovered efficiently (Figure I.3. D).

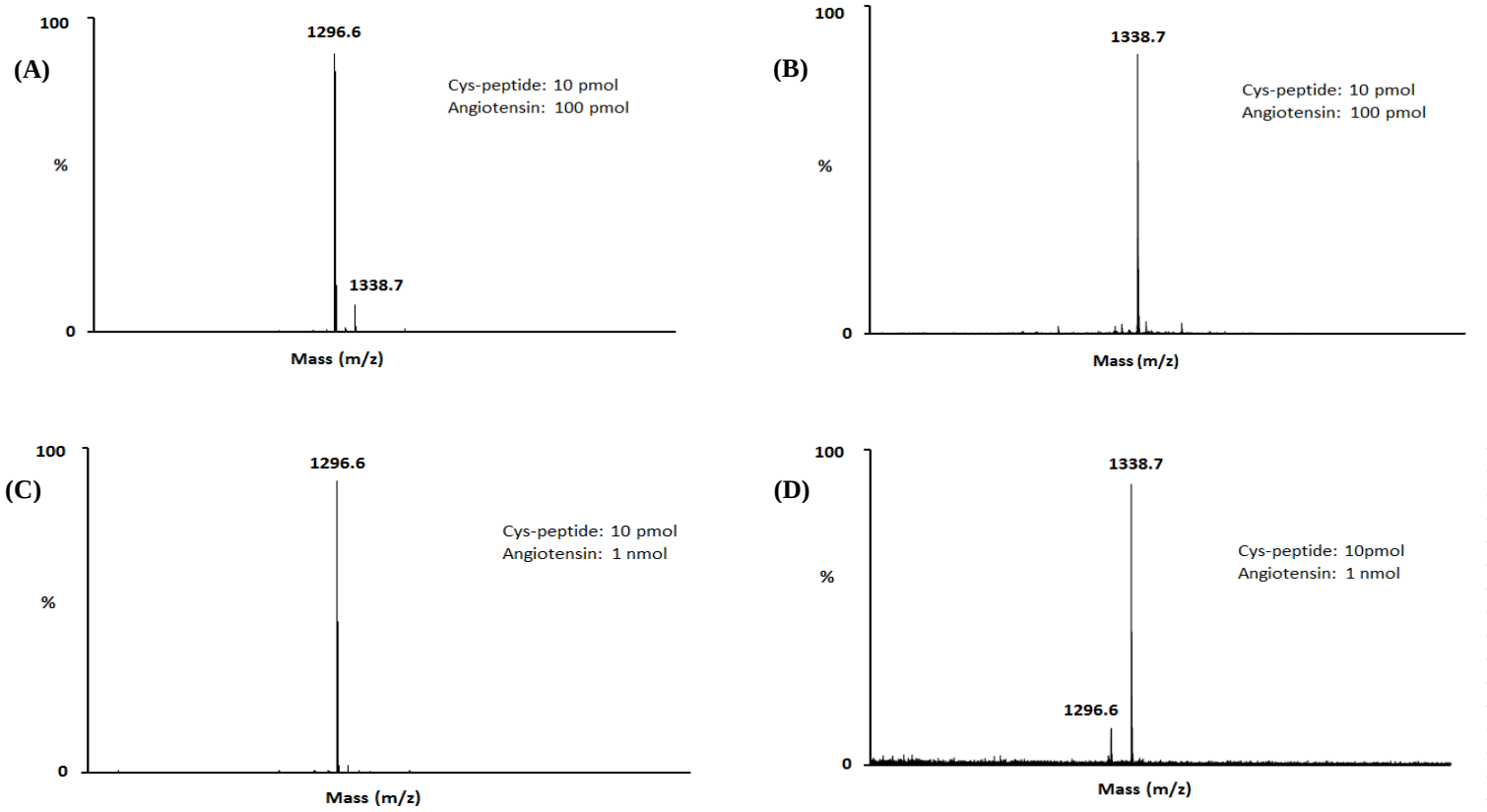


Figure I.3. MALDI-TOFMS spectra from a mixture of synthetic peptides, angiotensin (1296.6 Da) and a synthetic cysteine-containing peptide (1338.7 Da). (A) a mixture 1:10 of the cysteine containing peptide and angiotensin I; (B) the eluted fraction obtained after incubation with the resin Poros 20 NH SPDP of 100 pmol of angiotensin I and 10 pmol of Cys-peptide; (C) the flow-through fraction after Poros 20 NH SPDP purification: only the angiotensin I was observed; (D) the eluted fraction obtained after incubation with the resin Poros 20 NH SPDP of 1 nmol of angiotensin I and 10 pmol of Cys-peptide.

The next step was based on the experiments using a complex peptide mixture: the tryptic digestion from bovine serum albumin (BSA). 100 pmol of BSA were analyzed in each experiment. Likewise, 15 pmol of the marked peptide *m/z* 1338.7 (GACL*LPKIETMR, L*= +7 Da) was added to the peptide mixture to evaluate the recovery capacity of the system. The sample was incubated for 5 hours and the first fraction (flow-through), containing the peptides that were not linked to the resin, was recovered. After an extensively washing process, aiming to remove any unspecific binding, DTT was applied to the system. As a result, the second fraction, which contained the peptides that had reacted with the 2-pyridil moiety, was obtained. Afterwards, both fractions were analyzed on reverse-phase chromatography, Nano-LC C_{18}, and the chromatograms of these fractions have shown significant differences in concentration and peptide profile (Figure I.4).

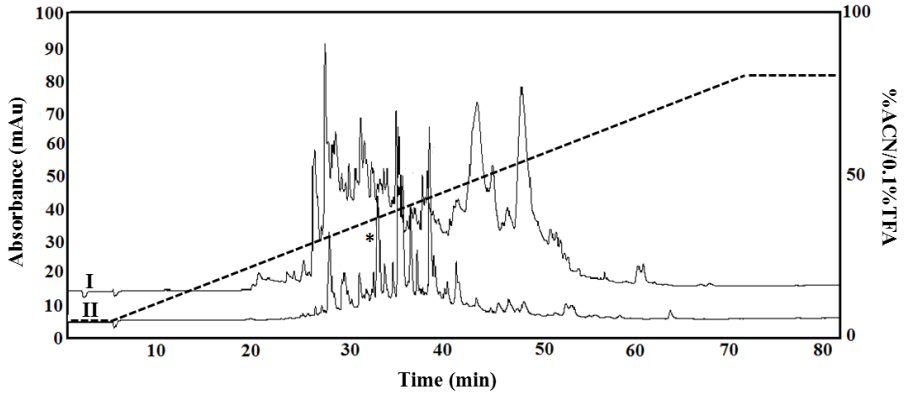


Figure I.4. Nano-LC profile of the tryptic fragments of bovine serum albumin monitored at 220 nm. Line I indicates the flow-through fraction from the Poros 20 NH SPDP system. Line II represents the eluted fraction, which contains Cys-peptides. The dotted line represents the acetonitrile concentration used for peptide elution. The (*) indicates the fraction where the spiked peptide (GACL*LPKIETMR, L*= +7 Da) was observed.

The process for tryptic digestion of BSA was repeated several times, using different concentrations of albumin and also different reaction times for enzyme digestion. In all cases, eluted fraction analysis by MALDI-TOF/TOF identified high concentration of BSA fragments, which contain, at least, one cysteine residue in their sequence (Table 1). The unbound peptides were also evaluated and 14% (6 peptides) of the identified peptides did not have cysteine in their sequence. The spiked peptide was only observed in the eluted fraction.

Table I.1. Peptides containing Cys observed in the eluted fraction in Figure I-4. The cysteine residues are depicted in bold.

$[M+H]^+$ (observed)	Sequence
841.4	L**C**VLHEK
918.5	LR**C**ASIQK
977.5	NE**C**FLSK
1011.4	QN**C**DQFEK
1015.4	SH**C**IAEVEK
1024.5	**CC**TESLVNR
1050.5	EA**C**FAVEGPK
1052.3	**CC**TKPESER
1177.4	E**CC**DKPLLEK
1197.9	DV**C**KNYQEAK
1209.5	**C**ASIQKFGER
1331.6	GA**C**LLPKIETMR
1138.4	**C**ASIQKFGER
1197.5	DV**C**KNYQEAK
1352.9	VTK**CC**TESLVNR
1362.5	SLHTLFGDEL**C**K
1364.5	ETYGDMAD**CC**EK
1386.6	YI**C**DNQDTISSK
1388.5	EYEATLEE**CC**AK
1407.6	LR**C**ASIQKFGER
1418.6	LKE**CC**DKPLLEK
1450.8	**C**ASIQKFGERALK
1456.8	VGTR**CC**TKPESER
1465.8	VGTR**CC**TKPESER
1497.6	DDPHA**C**YSTVFDK
1519.7	LKPDPNTL**C**DEFK
1578.6	E**CC**HGDLLE**C**ADDR
1588.8	GA**C**LLPKIETMREK
1616.8	QEPERNE**C**FLSHK
1633.7	YNGVFQE**CC**QAEDK
1667.9	MP**C**TEDYLSLILNR
1738.9	DDPHA**C**YSTVFDKLK
1756.7	**CC**AADDKEA**C**FAVEGPK
1811.2	L**C**VLHEKTPVSEKVTK
1824.1	RP**C**FSALTPDETYVPK
1851.2	LFTFHADI**C**TLPDTEK
1884.2	QEPERNE**C**FLSHKDDSPDLPK
1963.3	LKPDPNTL**C**DEFKADEK

2004.2	ETYGDMAD**CC**EKQEPER
2077.3	E**CC**HGDLLE**C**ADDRADLAK
2298.4	HLVDEPQNLIKQN**C**DQFEK
2316.7	YNGVFQE**CC**QAEDKGA**C**LLPK
2414.7	RP**C**FSALTPDETYVPKAFDEK
2435.8	GLVLIAFSQYLQQ**C**PFDEHVK
2441.8	VHKE**CC**HGDLLE**C**ADDRADLAK
2472.8	QN**C**DQFEKLGEYGFQNALIVR
2868.2	EYEATLEE**CC**AKDDPHA**C**YSTVFDK
2962.3	ETYGDMAD**CC**EKQEPERNE**C**FLSHK
3108.5	EYEATLEE**CC**AKDDPHA**C**YSTVFDKLK
3115.2	VGTR**CC**TKPESERMP**C**TEDYLSLILNR
3138.6	TVMENFVAFVDK**CC**AADDKEA**C**FAVEGPK
3398.9	SH**C**IAEVEKDAIPENLPPLTADFAEDKDV**C**K
3420.9	LAKEYEATLEE**CC**AKDDPHA**C**YSTVFDKLK
3759.3	HLVDEPQNLIKQN**C**DQFEKLGEYGFQNALIVR
3830.2	ETYGDMAD**CC**EKQEPERNE**C**FLSHKDDSPDLPK

1 MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EHFK**GLVLIA FSQYLQQCPF**
61 **DEHVKLVNEL TEFAK**TCVAD ESHAGCEK**SL HTLFGDELCK** VASLR**ETYGD MADCCEKQEP**
121 **ERNECFLSHK DDSPDLPKLK PDPNTLCDEF KADEK**KFWGK **YLYEIARRHP YFYAPELLYY**
181 **ANKYNGVFQE CCQAEDKGAC LLPK**IETMRE KVLASSARQR LR**CASIQKFG ER**ALKAWSVA
241 RLSQKFPKAE FVEVTKLVTD LTK**VHKECCH GDLLECADDR ADLAKYICDN QDTISSKLKE**
301 **CCDKPLLEKS HCIAEVEKDA IPENLPPLTA DFAEDKDVCK** NYQEAK**DAFL GSFLYEYSRR**
361 **HPEYAVSVLL RLAKEYEATL EECCAKDDPH ACYSTVFDKL KHLVDEPQNL IKQNCDQFEK**
421 **LGEYGFQNAL IVR**YTRKVPQ VSTPTLVEVS RSLGK**VGTRC CTKPESERMP CTEDYLSLIL**
481 **NRLCVLHEKT PVSEKVTKCC TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP**
541 **DTEK**QIKKQT ALVELLKHKP KATEEQLK**TV MENFVAFVDK CCAADDKEAC FAVEGPK**LVV
601 STQTALA

Figure I.5. Bovine serum albumin (BSA) amino acid sequence. In bold are the fragments observed in the eluted fraction from the Poros 20 NH SPDP resin. These fragments covered 70% of sequence and ≈60% of the tryptic peptides which contains cysteine in the sequence.

I.4. Discussion

Cysteine is present in all classes of proteins and peptides as hormones, growth factors, toxins and enzyme inhibitors [26]. This amino acid can suffer oxidative modifications and they are related to protein function; it has a role in diseases and health states, protecting proteins from oxidation and environment damages, and also participating in cascades in cellular defense system [27].

Many researchers have been trying to develop not only enrichment methods, but also methods to reduce sample complexity, in order to understand the role played by each protein in a sample. The analysis of Cys-peptides can be made by selective detection (fluorescent chemicals are used to detect the cysteine containing peptides) or selective reaction – which is the most used method. In the first case, fluorescent chemicals are used to detect the cysteine containing peptides [28]. The heterobifunctional linker SPDP can be used to isolate peptides containing cysteine residues because it is capable of forming disulfide bonds with any free thiol group present in the sample [29]. Using a simple peptide mixture, the method based on the Poros 20 NH resin and SPDP exhibits specificity for the peptides containing sulfhydryl groups. This could be corroborated by applying a hundredfold excess of angiotensin I against the Cys-peptide; this peptide could be recovered even in a low concentration. The reaction occurs through the interaction between the 2-pyridyl moiety and the sulfhydryl group, which promotes a selective modification of the thiol group of the crosslinker, causing thiol-disulfide interchange reaction. The thiol group (R_1SH) can act on any disulfide (R_2-S-S-R_3, where R_1, R_2 and R_3 are alkyl groups). The reaction takes place with low pKa and even in acidic media; this specificity allows a selective modification of distinct thiol groups [20].

BSA is a protein of 66.5 kDa and contains 35 cysteine residues in its sequence, which are arranged into 17 disulfide bonds [30]. As such, many articles have used BSA as a standard peptide mixture. Chang and coworkers (2010) used cysteinyl enrichment for isolating peptides containing cysteine derived from enzymatic digestion. For this purpose, the cysteine residues were oxidized using performic acid, and a nanoparticle affinity coupling with polyarginine was applied to recover the peptides [31]. They could identify 25 peptides using this method, a ratio of 40% of coverage of protein sequence. Additionally, Dai and collaborators (2005) applied the aforementioned method, but cited problems in identifying peptides using mass spectrometry in negative mode – the

equipment sensitivity is decreased by a factor of ~10 in this mode. They likewise observed that some peptides signals were suppressed due to sulfonic group, and that BSA peptide coverage, after reduction, alkylation and trypsin digestion, was less than 50% [18]. However, when BSA was applied to the current method, it was effective to isolate almost all fragments containing at least one cysteine residue in their sequence. 55 fragments containing cysteine were identified, including peptides obtained by miscleavage digestion of BSA, a ratio of 60% of sequence coverage. Peptides that had less than 700 Da weren't observed and peptides with mass value superior to 4000 Da did not have a significant abundance to be analyzed in this study. There weren't problems identifying these 55 fragments using positive mode on MALDI MS/MS.

Poros 20 NH SPDP is an advantageous method for proteomic analysis because it is simple and not time consuming. The sample only needs a step for disulfide bond oxidation, and a desalting process to remove reducing agents before applying the sample to the resin. Moreover, aiming to enhance the number of peptides identified in the analysis, one step of alkylation of disulfide bonds can be performed after recovering peptides from Poros 20 NH SPDP. This step can provide a high-quality mass spectrum for analysis.

Selection of the tryptic peptides which contain amino acids in low abundance, like Cys and His, reduces sample complexity and also provides enough information about protein profile, allowing proteome coverage [16, 24]. Poros 20 NH SPDP both provides a specific selection of peptides and allows the comprehension of proteins constitution in a complex sample.

Lastly, the benefits of developing new methods for covalent chromatography are applicable not only to purifying thiol-containing peptides, but also any molecule containing a free thiol group [21]. It has been described that N-acetyl-L-cysteine, a low-molecular weight thiol, and reduced glutathiones are related to the cleavage of disulfide bonds in plasmin (serine protease that functions in cell-migration and thrombolysis). This modification triggers the proteolysis of the N-terminal of the protein, producing angiostatin [32]. The stable –SOH observed in some proteins, like in aldose reductase, protects against perfusion and ischemic injuries on the heart [27]. Another case is the homocysteine (Hcy) – an amino acid derived from methionine that has a reactive sulfhydryl group; the high concentration of this sulfur amino acid in the plasma is associated to cardiovascular disease because it stimulates the inflammatory process in the organism, production of hydroxyl radical and lipid peroxidation. Further, in renal diseases, high levels of cysteine and Hyc are observed in the patients. In this case, data about both these amino acids can be used as a biomarker to cardiovascular and renal diseases [31, 33].

References

[1] Miseta, A., Csutora, P., Relationship between the occurrence of cysteine in proteins and the complexity of organisms. *Mol Biol Evol* 2000, *17*, 1232-1239.

[2] Jones, D. P., Go, Y. M., Mapping the cysteine proteome: analysis of redox-sensing thiols. *Curr Opin Chem Biol* 2011, *15*, 103-112.

[3] Brooks, D. J., Fresco, J. R., Lesk, A. M., Singh, M., Evolution of amino acid frequencies in proteins over deep time: inferred order of introduction of amino acids into the genetic code. *Mol Biol Evol* 2002, *19*, 1645-1655.

[4] Rosato, V., Pucello, N., Giuliano, G., Evidence for cysteine clustering in thermophilic proteomes. *Trends Genet* 2002, *18*, 278-281.

[5] Leonard, S. E., Carroll, K. S., Chemical 'omics' approaches for understanding protein cysteine oxidation in biology. *Curr Opin Chem Biol* 2011, *15*, 88-102.

[6] Nicoll-Griffith, D. A., Use of cysteine-reactive small molecules in drug discovery for trypanosomal disease. *Expert Opin Drug Discov* 2012, *7*, 353-366.

[7] Matthew D. Templeton, E. H. A. R. a. R. E. B., *Molecular Plant-Microbe Interactions* 1994, pp. 320-325.

[8] Koppers, A. J., Reddy, T., O'Bryan, M. K., The role of cysteine-rich secretory proteins in male fertility. *Asian J Androl* 2011, *13*, 111-117.

[9] Yamazaki, Y., Morita, T., Structure and function of snake venom cysteine-rich secretory proteins. *Toxicon* 2004, *44*, 227-231.

[10] White, S. H., Wimley, W. C., Selsted, M. E., Structure, function, and membrane integration of defensins. *Curr Opin Struct Biol* 1995, *5*, 521-527.

[11] Tinoco, A. D., Kim, Y. G., Tagore, D. M., Wiwczar, J., *et al.*, A peptidomics strategy to elucidate the proteolytic pathways that inactivate Peptide hormones. *Biochemistry* 2011, *50*, 2213-2222.

[12] Harwig, S. S., Chen, N. P., Park, A. S., Lehrer, R. I., Purification of cysteine-rich bioactive peptides from leukocytes by continuous acid-urea-polyacrylamide gel electrophoresis. *Anal Biochem* 1993, *208*, 382-386.

[13] Bateman, A., Bennett, H. P., Granulins: the structure and function of an emerging family of growth factors. *J Endocrinol* 1998, *158*, 145-151.

[14] Clark-Lewis, I., Kim, K. S., Rajarathnam, K., Gong, J. H., *et al.*, Structure-activity relationships of chemokines. *J Leukoc Biol* 1995, *57*, 703-711.

[15] Fahey, R. C., Hunt, J. S., Windham, G. C., On the cysteine and cystine content of proteins. Differences between intracellular and extracellular proteins. *J Mol Evol* 1977, *10*, 155-160.

[16] Wang, S., Zhang, X., Regnier, F. E., Quantitative proteomics strategy involving the selection of peptides containing both cysteine and histidine from tryptic digests of cell lysates. *J Chromatogr A* 2002, *949*, 153-162.

[17] Enjalbert, Q., Girod, M., Simon, R., Jeudy, J., *et al.*, Improved detection specificity for plasma proteins by targeting cysteine-containing peptides with photo-SRM. *Anal Bioanal Chem* 2013, *405*, 2321-2331.

[18] Dai, J., Wang, J., Zhang, Y., Lu, Z., *et al.*, Enrichment and identification of cysteine-containing peptides from tryptic digests of performic oxidized proteins by strong cation exchange LC and MALDI-TOF/TOF MS. *Anal Chem* 2005, *77*, 7594-7604.

[19] Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., *et al.*, Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 1999, *17*, 994-999.

[20] Brocklehurst, K., Specific covalent modification of thiols: applications in the study of enzymes and other biomolecules. *Int J Biochem* 1979, *10*, 259-274.

[21] Hillson, D. A., Freedman, R. B., Resolution of protein disulphide-isomerase and glutathione-insulin transhydrogenase activities by covalent chromatography. *Biochem J* 1980, *191*, 373-388.

[22] Gorman, J. J., Wallis, T. P., Pitt, J. J., Protein disulfide bond determination by mass spectrometry. *Mass Spectrom Rev* 2002, *21*, 183-216.

[23] Hartgerink, J. D., Covalent capture: a natural complement to self-assembly. *Curr Opin Chem Biol* 2004, *8*, 604-609.

[24] Giron, P., Dayon, L., Mihala, N., Sanchez, J. C., Rose, K., Cysteine-reactive covalent capture tags for enrichment of cysteine-containing peptides. *Rapid Commun Mass Spectrom* 2009, *23*, 3377-3386.

[25] Tambor, V., Hunter, C. L., Seymour, S. L., Kacerovsky, M., *et al.*, CysTRAQ - A combination of iTRAQ and enrichment of cysteinyl peptides for uncovering and quantifying hidden proteomes. *J Proteomics* 2012, *75*, 857-867.

[26] Akaji, K., Kiso, Y., *Houben-Weyl, Methods of Organic Chemistry, Synthesis of Peptides and Peptidomimetics*, Thieme, Stuttgart 2002.

[27] Cai, Z., Yan, L. J., Protein Oxidative Modifications: Beneficial Roles in Disease and Health. *J Biochem Pharmacol Res* 2013, *1*, 15-26.

[28] Seiwert, B., Hayen, H., Karst, U., Differential labeling of free and disulfide-bound thiol functions in proteins. *J Am Soc Mass Spectrom* 2008, *19*, 1-7.

[29] Carlsson, J., Drevin, H., Axén, R., Protein thiolation and reversible protein-protein conjugation. N-Succinimidyl 3-(2-pyridyldithio)propionate, a new heterobifunctional reagent. *Biochem J* 1978, *173*, 723-737.

[30] Feldhoff, R. C., Peters, T., Fragments of bovine serum albumin produced by limited proteolysis. Isolation and characterization of peptic fragments. *Biochemistry* 1975, *14*, 4508-4514.

[31] Chang, Y. C., Huang, C. N., Lin, C. H., Chang, H. C., Wu, C. C., Mapping protein cysteine sulfonic acid modifications with specific enrichment and mass spectrometry: an integrated approach to explore the cysteine oxidation. *Proteomics* 2010, *10*, 2961-2971.

[32] Hogg, P. J., Disulfide bonds as switches for protein function. *Trends Biochem Sci* 2003, *28*, 210-214.

[33] Miller, J. W., Beresford, S. A., Neuhouser, M. L., Cheng, T. Y., *et al.*, Homocysteine, cysteine, and risk of incident colorectal cancer in the Women's Health Initiative observational cohort. *Am J Clin Nutr* 2013, *97*, 827-834.

Chapter II. The method for isolating proteins containing free thiol groups

II.1. Introduction

In an organism, the global thiol-disulfide redox status is defined as the distribution of thiols and disulfides in diverse cellular pools. The activity and stability of many enzymes, chaperones and transcription factors are dependent on the maintenance of this status inside the different cellular compartments. Due to thiol's high reactivity, free thiol groups are mostly observed in proteins, not in peptides sequences. This reactivity is determined by structural environment and the cysteine pKa. Free cysteines have a lower pKa (from 3.5 to 5.6), which makes the thiol groups occur in the deprotonated form, the reactive thiolate anions (RS-) [1].

In eukaryotes, 2-3% of the proteins contain free thiol groups and these residues are involved in substrate biding or catalytic process [2, 3]. Moreover, these residues may be subject to posttranslational modifications acting as mediators for redox signaling and regulation. Over the past 15 years, many papers have described the role of free cysteines in redox-switches regulated proteins [4, 5].

Inside the cell, some free thiol groups are partially oxidized at normal cellular function, but when any stressor is present in the environment, these residues respond by attempting to maintain homeostasis [6]. Furthermore, reversible thiol modification can be caused by ROS or RNS and S-acylation (when fatty acids are attached through a thiol-ester group). Changes in the protein redox status are a physiological process, and the cysteine residues which are modified in physiological conditions represent the most sensitive thiol in the proteome [4, 6].

Most of the modifications are reversible, but in the case of extensive exposure of proteins to reactive species, an over-oxidation process may take place －the cysteines could be irreversibly modified to sulfonic and sulfinic acid [5]. One of most cited enzymes for thiol modification is the peroxiredoxin (Prx) family. These enzymes are described as protection to antioxidants, i.e. detoxifying by reducing ROS. The Prxs have one (1-Cys category) or two redox-active cysteines (2-Cys category) and these cysteine residues are directly involved in catalysis. The 1-Cys, the

peroxidatic cysteine, reacts to peroxide substrate and it is modified by sulfenic acid (Cys-SOH). Therefore, a thiol-electron donor is responsible for reducing the cysteine again and making the protein active once more. In the second group's case, the Cys-2, the resolving cysteine, oxidizes the Cys-SOH, forming a disulfide bond with the peroxidatic Cys. Subsequently, some enzymes can reduce the aforementioned bond and reestablish protein activity. The modification reset by Prx is used as an intracellular signal for many enzymes and transcriptional regulators [7]. Not only Prx, but also gene transcription, metabolism, apoptosis, signal transduction and protein trafficking are involved in the processes modulated by reversible modifications at cysteine residues [4].

There are several methods used for isolating Cys-containing peptides, and most of them utilize labeling of the thiol groups with various compounds. However, when these methods are applied for the analysis of free thiol groups in large protein molecules, they could not react with all possible thiols [8]. To analyze protein containing free-cysteines and the modifications sustained by this amino acid, the most applied technique is the tandem mass spectrometry: it allows the distinction between the different modifications [6]. In the case of cysteine proteases, some techniques using the covalent attachment to water-soluble supports have been applied, such as the resin thiopropyl sepharose 6B, as previously mentioned [9]. Differential labelling of thiols and disulfides, followed by electrophoretic separation or chromatographic techniques can also be used. However, determination of the relative abundance between proteins containing free thiol groups and proteins whose thiol groups are modified is yet a challenge [1].

Papain

Papain (EC 3.4.22.2) is a cysteine endopeptidase isolated for the first time from the latex of *Carica papaya* (papaya tree) in 1937 by Balls and coworkers. The latex production by plants is a defense mechanism against insects and herbivorous; more than 12,000 plant species are reported to contain latex. Papaya latex has a milky appearance and it is composed of 15% of dry matter and also of cardenolide, alkaloids, chitinases, limanarase and some enzymes [10-12].

Among the enzymes described in this fluid are glycyl endopeptidase, papain, chymopapain and caricain. These four endopeptidases represent 80% of the enzyme fraction present in the latex

and they have molar concentration of 1mM. Papain represents only 5-8% of this total, being the minor component among the endopeptidases [10, 13]. Inside the plant, the enzymes are stored in their inactive forms, the preforms, but they are converted into active enzymes when the latex is released from the plant. This is a mechanism to protect the plant from proteases' activity; after the plant is injured, less than two minutes are enough for all enzymes to turn active, and due to the intense proteolytic cleavages, the latex coagulates quickly [10].

The cysteine proteases in papaya latex contain 212 to 218 amino acids in their sequences and they have a strong degree of homology; the active site containing a cysteine, histidine and asparagine residues, and the amino acids close to the catalytic triad are highly conserved [10, 14]. Papain is a globular protein (Figure 3); it has 212 residues folded to form two domains with a cleft between them [15]. The active cysteine (Cys25) is embedded in the highly conserved sequence $CGSC^{25}WAFS$ and it is positioned on L1 α-helix at the surface on the left domain.

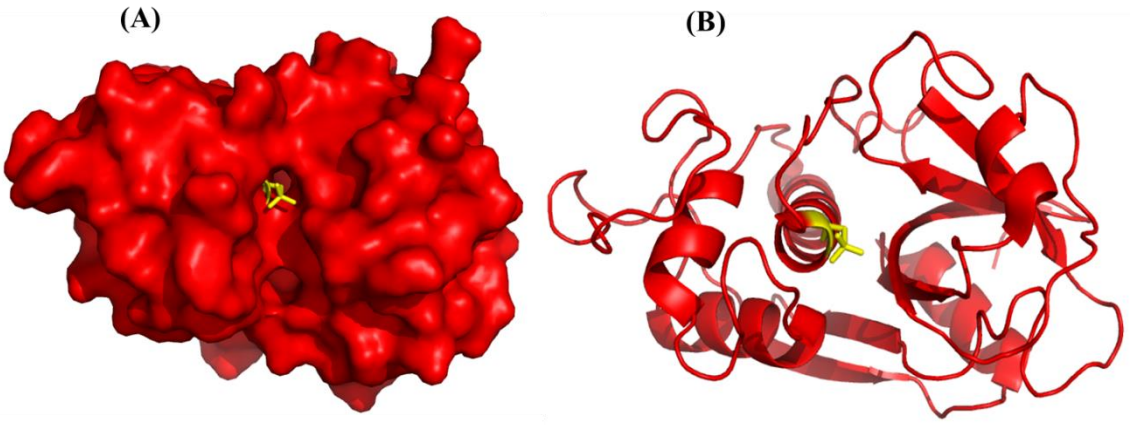


Figure II.1. Cysteine protease papain three dimensional structure (PDB 9PAP). (A) Figure representing the endopeptidase surface. The Cys25 is highlighted in yellow inside the active pocket. (B) Cartoon of papain structure the Cys25 is presented in yellow. Figures obtained using PyMOL software.

Papain-like cysteine proteases can be found in animals, fungi, RNA and DNA viruses, and bacteria [15]. For many years papain has been used by the food industry, in cosmetics, peptide production, synthesis of molecules and pharmaceutical industries. This enzyme has been applied to prevent wool shrinkage, chill hazes in beers and also as meat tenderizer [16]. Papain-like cysteine proteases were related to nonspecific protein degradation in lysosomes and other cellular compartments. Nevertheless, it has been shown that this protease plays an important role in antigen presentation and extracellular matrix turnover [14, 17].

Several techniques have been proposed to isolate papain from papaya latex. One of the first methods applied to isolate the active form of papain makes use of the p-aminophenylmercuric acetate attached to the sepharose resin, because this group is capable of reacting with any sulfhydyl group. In this case, all the inactive forms of the enzyme are washed out of the system. Another method applied was a sepharose column modified with the tetrapeptide glycyl-glycyl-(O-benzyl)-L-tyrosyl-L-arginine. It was described that this peptide has high affinity to the active form of papain and the enzyme can be recovered using a pH gradient. Conversely, one of the most used methods to isolate papain is the covalent chromatography with the 2-pyridyl group (thiopropyl sepharose 6B) [16]. Although all these methods have already been established, many authors emphasize problems with low purity. This may occur because most of the techniques utilized salt precipitation, which gave less than 50% of purity [17]. In this regard, as the method proposed in this work is efficient to isolate peptides containing cysteine, the technique has been tested for the crude extract from papaya latex, aiming to obtain a pure fraction of the active form of papain.

II.2. Material and Methods

The experimental and analytical procedures, reagents and instruments were the same as described in the section .1.3., except those described below.

Material

Papaya latex crude powder and extract was purchased from Sigma and Roche (St. Louis, MO and Basel, CH, respectively). Amicon® Ultra-0.5 centrifugal filter was obtained from Millipore (Darmstadt, German). The succinimidyl 6-[3'-(2-pyridyldithio)-propionamido] hexanoate (LC-SPDP) was purchased from Pierce (Rockford, IL). Dimethyl sulfoxide (DMSO) was purchased from Watanabe Chemical Industries LTD (Hiroshima, Japan). Pyr-Phe-Leu-pNa (papain substrate) was purchased from Peptides (Osaka, Japan). Other chemicals were of analytical grade and were purchased from Nacalai Tesque (Kyoto, Japan).

Resin preparation

A similar protocol was used to prepare the Poros 20 NH LC-SPDP. However, the concentration of the resin was increased. In all analysis, 40 mg of Poros 20 NH and 20 mM of LC-SPDP was used.

Sample preparation

Following the protocol proposed by Hillson, (1980, with some modifications), 0.1mM of dithiothreitol (DTT) was added to the working solution to activate the thiol groups present in the proteins. The sample containing papain was dissolved in 50 mM NH_4CO_3H containing 0.1mM DTT

and incubated in 30°C for 30 minutes. This concentration of the reducing agent is enough to activate the thiol groups in the protein, but the disulfide bonds in the protein were not broken. The DTT was removed using the ultra-0.5 centrifugal filter (cut off 10 kDa). The reduction solution was replaced by the working solution containing 50 mM NH_4CO_3H, 25% ethanol, 100 mM NaCl and 1 mM EDTA pH 7.8. Even using this weak reducing condition, it was possible to observe some protein aggregates; in this case, these were removed from the tests by centrifugation.

Recovery of the active fraction of the papain

The sample was applied to the Poros 20 NH LC-SPDP system and they were incubated for 10 hours at room temperature. After the linking process, the resin was centrifuged (2000xg for 10 seconds) and the supernatant recovered (flow-through fraction). The resin was extensively washed using working solution before recovering the papain linked to the system. Fifty millimolar of DTT were suspended in 500 µL of working solution and used to obtain the active papain (eluted fraction). The reducing solution was mixed with the resin for 1 hour at room temperature and the supernatant was recovered.

Polyacrylamide Gel electrophoresis (PAGE)

SDS-PAGE was carried out as described by Laemmli (1970). The sample was dissolved in 50 mM Tris-HCl buffer, pH 6.8, containing 0.1% bromophenol blue, 2% SDS and 10% sucrose, and the sample was applied to 12.5% acrylamide gel (9 cm x 6 cm x 0.1 cm). The running buffer contained 25 mM Tris- 192 mM glycine and 0.1% SDS, pH 8.4. PAGE was conducted at a constant current of 20 mA for 80 minutes. The SDS-PAGE was stained using the silver staining protocol proposed by Blum (1987).

Enzymatic activity assay

The assay for papain activity was done as described by Lottenberg and coworkers (1981). After removing the reducing agent, the flow-through fraction and the eluted fraction were diluted in phosphate buffer 100 mM pH 6.8 containing 100 mM NaCl and 1 mM of EDTA. Additionally, the crude extract of papain latex was treated under the same conditions. Ten milimolar of a substrate (Pyr-Phe-Leu-*p*Na) was diluted in DMSO. After this, thirty microliters of each sample was added to 2940 µL of buffer and 30 µL of the substrate was also added to the solution. The final solution was incubated for three minutes and the absorption at 405 nm was measured. The enzyme activity was calculated based on the absorption coefficient of *p*Na (ε_{405nm}= 9920). One unit of activity corresponds to the amount of protease that can hydrolyze one micromol of the substrate per minute under the aforementioned conditions.

II.3.Results

To evaluate the proposed method of recovering proteins containing free thiol groups, a cysteine protease, papain, was chosen as a model. After free thiol group activation using 0.1mM of DTT, the reducing agent was removed from the solution using the Amicon (cut off 10 kDa). The sample, dissolved in the working solution, was applied to the Poros 20 NH SPDP, and the system was incubated for 8 hours at 4°C (to reduce the autolysis process). The proteins which were not linked to the resin were removed by washing the system exhaustively with the working solution. To recover the papain attached to the resin, 50 mM of DTT was used. However, using the heterobifuncional SPDP, a low concentration of the enzyme in the eluted fraction was observed, while most of the papain was washed out of the resin. In this sense, the strategy to recover papain was modified. The heterobifunctional group SPDP was replaced by the succinimidyl 6-[3'-(2-pyridyldithio)-propionamido] hexanoate (LC-SPDP). After lengthening the linker, it was observed that the concentration of papain attached to the resin increased 5 times (Fig.II.2).

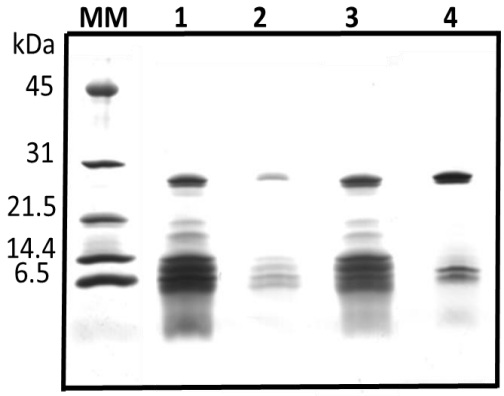


Figure II.2. SDS-PAGE 12.5% of *C. papaya* latex. 0.6 mg of dried latex was applied to the system. One third of each sample (total volume) was used on the SDS-PAGE. MM is the molecular mass standards. Line 1 and 2 represent the sample applied to SPDP system, flow-through fraction and eluted fraction, respectively. Line 3 and 4 represent the sample applied to the Poros 20 NH LC-SPDP system, flow-through and eluted fraction, respectively. The recovery of papain was higher when the LC-SPDP is used.

Although the recovery of the enzyme increased, papain was also observed in the flow-through fraction. Thus the enzyme concentration applied to the system was decreased to ensure that the protein did not react with the system because the capacity of the system was exceeded. Even though smaller amounts were used on the experiments, papain was still observed in the flow-through fraction (Figure II.3).

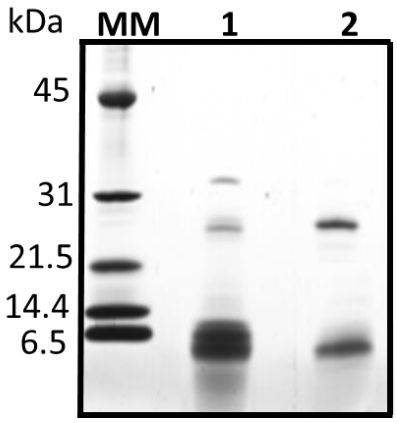


Figure II.3. SDS-PAGE 12.5% of *C. papaya* latex. In this case, the concentration of the dried latex used was reduced; 0.3 mg was applied to the system and one third of each sample (total volume) was used on the SDS-PAGE. MM is the molecular mass standards. Line 1 represents the flow-through fraction from Poros 20 NH LC-SPDP system. Line 2 represents the eluted fraction from Poros 20 NH LC-SPDP system where purified papain can be observed.

After obtaining the flow-through and eluted fraction, an activity assay was performed to evaluate if the enzyme present in the flow-through fraction was the active form of papain. Following the protocol proposed by Lottenberg (1981), the peptidyl-*p*Na substrate was used to quantify the papain enzymatic activity, and the amount of *p*Na released was measured by an UV spectrometer. The enzyme activity was predominantly recovered in the eluted fraction while only low enzyme activity was observed in the flow-through fraction (Table II.1).

Table II.1. Papain enzymatic assay using peptidyl-*p*NA. Three hundred seventy four micrograms of papain extract was applied to the Poros 20 NH LC-SPDP system and the same concentration was used as the total enzymatic activity for the assay. The values correspond to the amounts of *p*Na released in one minute.

Sample	Amounts of *p*NA released (µmol.min^{-1})
Crude extract	0.37
Flow-through fraction	0.03
Eluted fraction	0.21

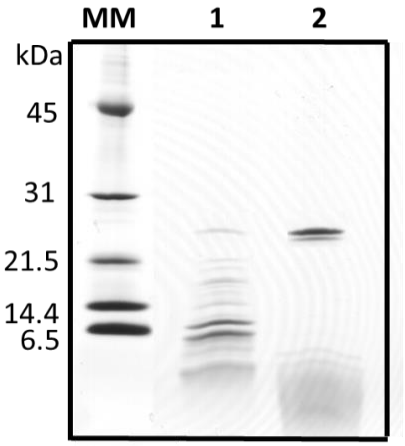


Figure II.4. SDS-PAGE 12.5% of *C. papaya* latex sample used on enzymatic assay. MM is the molecular mass standards. Line 1 represents the flow-through fraction. Line 2 represents the eluted fraction from Poros 20 NH LC-SPDP system.

In some cases, several low molecular protein bands were also observed in the SDS-PAGE from eluted fraction. Considering that the sample was previously centrifuged using a filter of 10 -kDa molecular cut off, these peptides could be derived from protein autolysis during the experimental procedures. To diminish the degradation process of the active form of papain, the

incubation time was reduced. However, reducing the time of exposure between the enzyme and the system could affect the recovery of papain. Therefore, the time reaction was evaluated in different stages. After preparing the system, the papain was incubated for 2, 4 and 7 hours. In each case, after the incubation time, the standard procedures were made and the cysteine protease was recovered. The samples obtained in this experiment were analyzed by SDS-PAGE and it was observed that in the case of papain, it is necessary to incubate the enzyme with the system for at least seven hours to get a higher concentration of enzyme (Figure II.5).

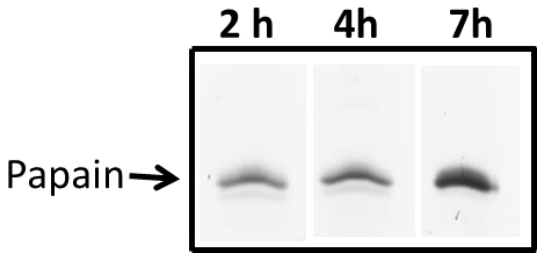


Figure II.5. SDS-PAGE 12.5% of *C. papaya* latex. In this experiment, different times for linkage were evaluated. The sample (3 mg of dried latex) was exposed to the system for 2, 4 and 7 hours. At the seventh hour, the concentration of the active form of papain recovered was higher comparing to 2 and 4 hours.

II.4.Discussion

Cysteine proteases are involved in many biological processes – degradation of muscular proteins, immunopharmacological responses and some diseases are caused by deregulation of these enzymes, and, therefore, they have been studied over the years [15]. From papaya latex, four cysteine proteases can be obtained: papain (that corresponds to 10% of the enzymes content), chymopapain (26%), glycyl endopeptidase (23-28%) and caricain (14-26%) [10, 18]. Papain cleaves peptide bonds in basic amino acids as lysine, arginine and residues following phenylalanine; it acts against proteins, peptides, amino acids esters and amide links [19].

Usually, papain has been isolated from the crude latex by cation-exchange chromatography, and the profiles presented using this method show that papaya latex is a complex mixture of enzymes. On the other hand, some hydrophobic resins can be used to purify the papain from the latex as the fractogel TSK butyl 650, fractogel EMD propyl and phenyl sepharose. However, the profile for elution of the protease is not as efficient as the one observed for cation-exchange chromatography [10].

Proteins can be combined with crosslinker reagents through a reaction with the side chain of a specific amino acid; this technique has been effective to analyze several types of proteins [20]. The bridges or links formed can be intra or intermolecular type: the intramolecular crosslinking has been applied to insert additional tertiary structure into proteins, e.g. to increase their conformational stability. The intermolecular crosslinking has been used to modify cell membranes or other molecular assemblies and bind proteins (protein conjugation) [21]. In this regard, the reactivity of the 2-pyridyl disulfide in the LC-SPDP allows the separation between the thiol-containing molecules from those without this group. The electrophilic propriety of the pyridyl disulfide has been used in kinetic studies due to the high reactivity with active-centre thiol groups in a wide range of pHs; conformational transitions and changes in ionization state can be detected using this reactive group [22]. Additionally, the thiol group capability of affecting the microenvironment in the catalytic pocket allows different kinds of enzymes to bind to a specific linker with distinct affinities [23]. e.g. actinidin is a cysteine protease produced by *Actinidia chinensis* (kiwifruit) and it has a similar polypeptide conformation to that of the papain; however, the activity of these two enzymes against the 2-pyridyl group is different [12, 24]. In the case of papain, the Cys25 has high

reactivity with the 2-pyridyl disulfide and it is also supported by the interaction between the 2-pyridyl disulfide and the imidazole ring from His159 [25].

Firstly, the SPDP was chosen as the crosslinker to isolate papain from papaya latex. After the standard procedures, it was observed that the concentration of enzyme recovered during the chromatographic process was low (Figure II.2). Several factors were considered to explain the low efficiency of the method. One hypothesis studied was the positioning of the Cys25 in the active site. The Cys25 is located in the cleft between the two subunits and, to recover the papain, a direct interaction between the 2-pyridyl group and the –SH group at Cys25 is necessary. However, as the SPDP is linked to the Poros 20 NH, the accessibility of the 2-pyridyl to the Cys25 may be limited. Aiming to increase the amount of enzyme recovered during the process, the SPDP was replaced by the LC-SPDP. This linker has a longer chain, with a space arm length of 15.6Å, as opposed to the 6.8 Å arm length of SPDP (Thermo Scientific manual). The longer arm can easily access the active site and react with Cys25; after modifying the linker, the amount of papain obtained in the process increased 5 times (Figure II.2).

Despite of higher protein recovery, some papain was still observed in the flow-through fraction. Taking into account the fact that a free thiol in a protein is sensitive to oxidation or modification, the Cys residue in the papain obtained in the flow-through fraction might have been modified or oxidized, while the papain in the eluted fraction might contain a free thiol group. Since the endopeptidases present in the latex of many plants can cause damage to the plants themselves, these enzymes are rendered inactive when they are stored in the latex; when the fresh papaya latex is collected, the activity among the endopeptidades present in this fluid is low [26].

It was described by Lowe and coworkers (1976) that in papaya latex three forms of papain can be observed: an irreversible inactivated papain, a reversible inactivated papain and an active form; in the case of modified papain, the sulfur atom from Cys25 might link to other compounds, *e. g.* cysteine (reversibly modified) or sulfinic acid (irreversible modification) [27, 28]. In this sense, before exposing the sample to the Poros 20 NH LC-SPDP, the papain crude extract was activated using 0.1mM DTT. This concentration of reducing agent is enough to activate the thiol groups that are reversibly modified, while the intramolecular disulfide bonds in the papain structure are not affected [23].

Even though the former gentle reduction was performed, small amounts of papain could still be observed in the flow-through fraction. In this case, the papain present in the flow-through may represent the percentage of inactive enzyme; according to Azarkan and coworkers (2003), 30 to

50% of the enzymes present in the papaya latex are irreversibly oxidized [10]. Moreover, once the activator (reducing agent) is removed from the sample, the papain may readily become inactive again as was demonstrated by Klein and Kirsch (1969) [27, 28].

In the case of the active form, the papain catalytic pocket may accommodate seven amino acid residues, but the residue P2 (the second position for substrate biding in the active site) is the most important residue for enzyme specificity; to hydrolyze a peptide and/or protein, papain has a preference for phenylalanine at the P2 position. The reaction can occur at pH range of 4-8, but the maximal velocity of hydrolysis occurs at pH 6.5; enzyme activity is proportional to the number of titratable thiol on the enzyme [27-29]. As such, in order to confirm that the papain present in the flow-through fraction is the inactive form, assay using the Pyr-Phe-Leu-*p*Na was performed. This substrate contains the chromogenic group p-nitroaniline (*p*Na), and when the peptide bond is cleaved, the amount of *p*Na released can be monitored at 405 nm [30]. The assay data demonstrated that little activity was observed for the flow-through fraction; the activity was mainly recovered in the eluted fraction with 56.8% yield. (Table II.1, Figure II.4).

In some cases, some proteins with low molecular masses on the SDS-PAGE could be observed (Figure II.2 and 3). This contamination might be attributed to partial autolysis of the enzyme. Papain is a very stable protein; even in extreme conditions such as high concentration of denaturating agents and low pHs, the protein remains active [31]. Taking into account that the incubation time during the experiments was ten hours, it was considered that a longer time could cause degradation of the enzyme. According to Sluyterman and Wijdenes (1980), once the enzymes were in the active form, they can attack the papain itself, producing some fragments containing free-cysteine that can react with the Poros 20 NH LC-SPDP system [31]. Hence, to diminish the degraded products derived from autolysis, the time necessary for the maximum recovery of papain was evaluated. Three incubation times were examined: 2, 4 and 7 hours (Figure II.5). As a result, seven-hour incubation gave the highest recovery of papain, and also the degradation was reduced.

The method described here, using the 2-pyridyl from LC-SPDP, was effective to isolate the cysteine protease, papain. Additionally, the system makes it possible to isolate only the active form of papain in one step from a biological resource, the latex. Hereof, the present technique that can separate the active forms of thiol-enzymes should be useful to obtain a highly purified enzyme. Of note, the ratio of the free Cys form (active form) to the oxidized or modified ones (inactive form) would reflect the developmental and/or physiological status of *C. papaya* (or any other organism

that produces cysteine proteases), which then could be rationalized to be how the enzymatic activity is controlled *in vivo*.

References

[1] Hansen, R. E., Roth, D., Winther, J. R., Quantifying the global cellular thiol-disulfide status. *Proc Natl Acad Sci U S A* 2009, *106*, 422-427.

[2] Seiwert, B., Hayen, H., Karst, U., Differential labeling of free and disulfide-bound thiol functions in proteins. *J Am Soc Mass Spectrom* 2008, *19*, 1-7.

[3] Hu, W., Tedesco, S., Faedda, R., Petrone, G., *et al.*, Covalent selection of the thiol proteome on activated thiol sepharose: a robust tool for redox proteomics. *Talanta* 2010, *80*, 1569-1575.

[4] Guo, J., Gaffrey, M. J., Su, D., Liu, T., *et al.*, Resin-assisted enrichment of thiols as a general strategy for proteomic profiling of cysteine-based reversible modifications. *Nat Protoc* 2014, *9*, 64-75.

[5] Brandes, N., Schmitt, S., Jakob, U., Thiol-based redox switches in eukaryotic proteins. *Antioxid Redox Signal* 2009, *11*, 997-1014.

[6] Jones, D. P., Go, Y. M., Mapping the cysteine proteome: analysis of redox-sensing thiols. *Curr Opin Chem Biol* 2011, *15*, 103-112.

[7] Wood, Z. A., Schröder, E., Robin Harris, J., Poole, L. B., Structure, mechanism and regulation of peroxiredoxins. *Trends Biochem Sci* 2003, *28*, 32-40.

[8] Egorov, T. A., in: Jörnvall, H., Höög, J.-O., Gustavsson, A.-M. (Eds.), *Methods in Protein Sequence Analysis*, Birkhäuser Basel 1991, pp. 177-185.

[9] Carlsson, J., Axén, R., Unge, T., Reversible, covalent immobilization of enzymes by thiol-disulphide interchange. *Eur J Biochem* 1975, *59*, 567-572.

[10] Azarkan, M., El Moussaoui, A., van Wuytswinkel, D., Dehon, G., Looze, Y., Fractionation and purification of the enzymes stored in the latex of Carica papaya. *J Chromatogr B Analyt Technol Biomed Life Sci* 2003, *790*, 229-238.

[11] Konno, K., Hirayama, C., Nakamura, M., Tateishi, K., *et al.*, Papain protects papaya trees from herbivorous insects: role of cysteine proteases in latex. *Plant J* 2004, *37*, 370-378.

[12] Brocklehurst, K., Baines, B. S., Malthouse, J. P., Differences in the interaction of the catalytic groups of the active centres of actinidin and papain. Rapid purification of fully active actinidin by covalent chromatography and characterization of its active centre by use of two-protonic-state reactivity probes. *Biochem J* 1981, *197*, 739-746.

[13] Nitsawang, S., Hatti-Kaul, R., Kanasawud, P., Purification of papain from Carica papaya latex: Aqueous two-phase extraction versus two-step salt precipitation. *Enzyme and Microbial Technology* 2006, *39*, 1103-1107.

[14] Lecaille, F., Kaleta, J., Brömme, D., Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. *Chem Rev* 2002, *102*, 4459-4488.

[15] Harrison, M. J., Burton, N. A., Hillier, I. H., Catalytic Mechanism of the Enzyme Papain: Predictions with a Hybrid Quantum Mechanical/Molecular Mechanical Potential. *Journal of the American Chemical Society* 1997, *119*, 12285-12291.

[16] Lowe, G., The Cysteine proteinases. *Tetrahedron* 1976, *32*, 291-302.

[17] He, Y., bin Tuan Chik, S. M. S., Chong, F. C., Purification of papain from unclarified papaya juice using reversed phase expanded bed adsorption chromatography (RP-EBAC). *Journal of Industrial and Engineering Chemistry*.

[18] Chaiwut, P., Pintathong, P., Rawdkuen, S., Extraction and three-phase partitioning behavior of proteases from papaya peels. *Process Biochemistry* 2010, *45*, 1172-1175.

[19] Amri, E., Mamboya, F., *American Journal of Biochemistry and Biotechnology* 2012, pp. 99-104.

[20] Lambert, W., Söderberg, C. A., Rutsdottir, G., Boelens, W. C., Emanuelsson, C., Thiol-exchange in DTSSP crosslinked peptides is proportional to cysteine content and precisely controlled in crosslink detection by two-step LC-MALDI MSMS. *Protein Sci* 2011, *20*, 1682-1691.

[21] Carlsson, J., Drevin, H., Axén, R., Protein thiolation and reversible protein-protein conjugation. N-Succinimidyl 3-(2-pyridyldithio)propionate, a new heterobifunctional reagent. *Biochem J* 1978, *173*, 723-737.

[22] Baines, B. S., Brocklehurst, K., Characterization of papaya peptidase A as a cysteine proteinase of Carica papaya L. with active-centre properties that differ from those of papain by using 2,2'-dipyridyl disulphide and 4-chloro-7-nitrobenzofurazan as reactivity probes. Use of two-protonic-state electrophiles in the identification of catalytic-site thiol groups. *Biochem J* 1982, *205*, 205-211.

[23] Hillson, D. A., Freedman, R. B., Resolution of protein disulphide-isomerase and glutathione-insulin transhydrogenase activities by covalent chromatography. *Biochem J* 1980, *191*, 373-388.

[24] Brocklehurst, K., O'Driscoll, M., Kowlessur, D., Phillips, I. R., *et al.*, The interplay of electrostatic and binding interactions determining active centre chemistry and catalytic activity in actinidin and papain. *Biochem J* 1989, *257*, 309-310.

[25] Baines, B. S., Brocklehurst, K., *Journal of Protein Chemistry* 1982, pp. 119-139.

[26] Azarkan, M., Dibiani, R., Baulard, C., Baeyens-Volant, D., Effects of mechanical wounding on Carica papaya cysteine endopeptidases accumulation and activity. *Int J Biol Macromol* 2006, *38*, 216-224.

[27] Klein, I. B., Kirsch, J. F., The activation of papain and the inhibition of the active enzyme by carbonyl reagents. *J Biol Chem* 1969, *244*, 5928-5935.

[28] Klein, I. B., Kirsch, J. F., The mechanism of the activation of papain. *Biochem Biophys Res Commun* 1969, *34*, 575-581.

[29] Filippova, I. Y., Lysogorskaya, E. N., Oksenoit, E. S., Rudenskaya, G. N., Stepanov, V. M., l-Pyroglutamyl-l-phenylalanyl-l-leucine-p-nitroanilide—A chromogenic substrate for thiol proteinase assay. *Analytical Biochemistry* 1984, *143*, 293-297.

[30] Lottenberg, R., Christensen, U., Jackson, C. M., Coleman, P. L., Assay of coagulation proteases using peptide chromogenic and fluorogenic substrates. *Methods Enzymol* 1981, *80 Pt C*, 341-361.

[31] Sluyterman, L. A., Wijdenes, J., The effect of thiol compounds on the autolysis of papain. *Eur J Biochem* 1980, *113*, 189-193.

Chapter III. Isolating proteins with free thiol groups in plasma

III.1. Introduction

Blood connects all cells and tissues in an organism and it is composed of cells (red cells, white cells and platelets), proteins, lipids, carbohydrates, amino acids, nucleic acids and hormones. Blood does not possess its own genome nor transcriptome; some researches propose that blood genome is the "collection" of the organism's foundational genome with all variation as mutations and gene duplications, and despite the lack of its own transcriptome, blood may contain transcripts of any cell in the body [1].

Plasma is the liquid portion of the blood, obtained after all cell types are removed. On the other hand, in the absence of anticoagulant substances, the prothrombin, present in blood, is cleaved by thrombin and a fibrin clot is formed; after a centrifugation process the serum is obtained. The serum has a lower protein concentration (in the order of 3-4%), compared to the plasma, due to the removal of the fibrinogen and some proteins that are involved in fibrin clot [1, 2].

Even if an impressive number of peptides and proteins have been identified in the plasma, only 22 proteins make up ≈ 95% of the mass of proteins in this fluid [3, 4]. The most abundant proteins in the blood are albumin (3.5 g/mL), IgG (1.2 g/mL), Apolipoprotein I (1.0g/mL), and transferrin (0.2g/mL) [2, 5]. Conversely, several proteins and peptides, as protein inhibitors and hormones, are present in low concentrations in the blood and have a potent physiologic activity [2, 5].

Plasma and serum are applied for medical diagnosis and prognosis, because they represent the health conditions at any given point in time and many health disorders manifest themselves by modifying the protein contents in the blood [6]. Proteins present in the plasma carry out their functions in the circulation, but many other proteins and peptides may also be present in this fluid for different reasons such as the result of tissue damages, metabolic turnover, etc [2]. Hippocrates

was the first physician who postulated that diseases could be related to the blood, and nowadays blood is utilized for clinical diagnosis, ranging from cancer to diabetes [1, 2].

At the same time, proteins in the plasma have been used as a model for studying metabolism, synthesis and degradation [7]. e. g. low levels of albumin, due to an increased catabolism process, are related to acute infections, myocardial infarction or traumas [8]. Moreover, many products obtained from plasma have been used in the pharmaceutical industry. e.g., in the last few years, twenty plasma proteins have been used for treating diseases or injuries [2, 5].

One of the first attempts to identify the plasma proteome was made in 1976, using the two dimensional electrophoresis, and only 40 proteins could be described on that occasion. Due to dynamic range limitation in electrophoresis techniques, merely another 20 proteins could be identified in the past 16 years [9]. However, in 2002, Anderson and Anderson reported 289 proteins that were isolated by the combination of different techniques as immunoassay and 2D electrophoresis [2]. Moreover, in the same year, the Human Proteome Organization (HUPO) initiated the Human Plasma Proteome Project (HPPP) wherein 35 laboratories concentrated their efforts to determine the protein contents of the plasma. Nowadays, 3020 proteins are identified [10].

Beyond the protein abundance in the plasma, greater than 10 orders of magnitude, the heterogeneity level of this fluid caused by inter- and intra-individual variations, correlated to age, gender and genetic factors, make the plasma one of the most complex samples [10]. Although plasma has been one of the most studied samples derived from human, several researchers believe that much important information is yet unknown [1].

Transtherytin

Transthyretin (TTR) is a protein produced by the liver and secreted in the plasma; it can also be found in cerebrospinal fluid. It was sequenced for the first time by Kanda and coworkers in 1974, and it was formerly called prealbumin because in electrophoresis analysis, this protein runs ahead of albumin among serum proteins [11]. The levels of TTR in the plasma, in adults, are 20-40 mg/dL,

but the expression of this protein decreases over the age of fifty. TTR can bind to thyroid hormones and retinol binding proteins. It is composed of 127 amino acid residues, which are organized in a β-sheet structure with a small α-helix, formed by seven amino acids. The β-sheets are organized in eight strands of A-H into two anti-parallels sheets and the α-helix is present between strands E and F (Figure III.1). This protein is observed in a homotetrameric structure with 55 kDa and each monomer is associated with another monomer, forming a αβ dimer which interacts with strong hydrogen bonds in the β-sheet region [12-14].

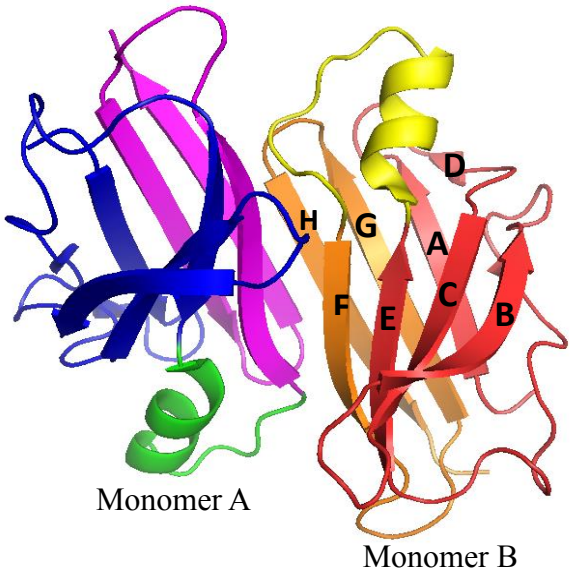


Figure III.1. Three dimensional structure of human wild type TTR dimer. The eight strands are indicated by the letters A-H. Figure obtained using PyMOL software (PBD number 3W3B).

On the other hand, the contact between the two dimers corresponds to a small area within the AB loop of one monomer and the H strand of the two primed monomer; hydrophilic and hydrophobic interactions maintain the quaternary structure, but they are weaker comparing to the

forces existing between the dimer structure, which makes the dimeric structure the basic unit for TTR [15].

The structure called “dimer of dimer” produces a twofold symmetry channel in which the hormones are bound (Figure III.2). This channel accommodates two molecules of thyroxine (50 Å and 8 Å in diameter); the center is a hydrophilic site containing the hydroxyl groups. As an asymmetric molecule, each thyroxine is positioned in a different orientation when it is bound to TTR. Furthermore, there is a negative cooperativity between the thyroxine molecules. When one thyroxine binds to the site, conformational changes modify the interaction among the TTR subunits, affecting the bidding affinity of the second molecule. In this case, after the first link, the channel becomes more rigid, making it more difficult for the second molecule of thyroxine to access the site [14, 15].

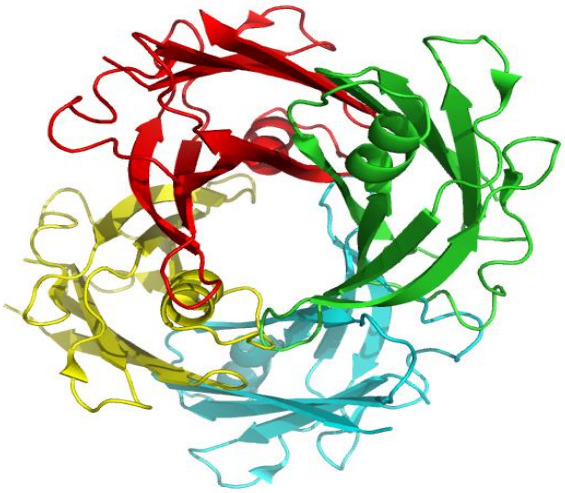


Figure III.2. Three dimensional structure of human wild type TTR tetramer. Each TTR monomer is highlighted by different colors. In the center, the channel is presented, where the thyroid hormones are transported (PDB 3W3B and 1DVQ). Figure obtained using PyMOL software.

In addition, TTR is responsible for transporting vitamin A (retinol) from the liver to the cells. The retinol is bound to the retinol binding protein (RBP - 21 kDa) and the complex binds to TTR on the outer surface; the main interaction occurs among the hydrophobic residues of TTR α-helix and RBP loop region between β-strands 7 and 8 [12, 15, 16]. The link among TTR-RBP-Retinol protects the retinol from degradation and prevents vitamin A from being released to non-target cells. Two RBP/retinol complexes interact with TTR tetramer; two RBP molecules bind to the same TTR dimer and this linkage blocks the other two sites from RBP interactions [12, 16]. However, the link between RPB/retinol and TTR does not affect the affinity of this protein for thyroxine [15].

Amyloidosis process and other diseases

TTR has a short half-life – around two days – in the blood stream; it is metabolized and excreted by the kidneys. Level changes of TTR in the serum can be correlated with nutritional status and health conditions: *e. g*. the level of transient TTR decreases during inflammation process and postoperative state; conversely, alcohol intoxication, prednisone therapy, and metabolites produced during pregnancy, increase the level of this protein [17]. Studies conducted by Gaudiani and coworkers (2014) reported that TTR can be used like a biomarker in case of severe illness and in malnourished patients [17]. Regarding hypoglycemia (low levels of sugar in the blood), the concentration of TTR in the blood stream is reduced.

The gene that codes TTR is located on the long arm of chromosome 18 and is 7kb long; it contains four exons of 200 base pairs each and three introns. The first exon codifies only three amino acid residues; the second exon the 4-47, the third exon the 48-92, and the last exon the 93-127. A sequence upstream to the 5'cap is related to gene expression and it is highly conserved among different species; humans, rats and mice present ≈85% of homology in TTR sequence. The small size of the exons enables easy identification of any mutation in the gene [18, 19]. Genetic mutation in TTR gene may lead to the production of an abnormal protein, which may form an incorrect complex or be deficient in the complex formation, resulting in the degradation and/or

precipitation as an insoluble amyloid fibril in organs and tissues [20]. The extracellular insoluble protein fibrils are responsible for amyloidosis diseases. This aggregation is observed in the ordered structure in which the peptide's chains fold back and forth on themselves to form antiparallel β-sheets (called cross-β); this strong intra-strand stabilization, due to exposure of the hydrophobic residues, makes the fibrils also resistant to proteolytic digestion (Figure III.3) [12, 14, 21]. These fibrils accumulate in some organs and the peripheral nerves, causing tissue damage and disruption of organ function [15].



Figure III.3. Three dimensional structure of human TTR amyloid fibrils (10 amino acids; 105-115). The fragments adopt a β-strand conformation. Figure obtained using PyMOL software (PDB number 1RVS).

The most common amyloidosis process caused by TTR is due to mutations in peptide sequence: some mutations cause a weaker or tighter hormone binding and the destabilization of the quaternary structure. However, not all mutations will destabilize the protein structure. There are over 100 points of mutations in TTR sequence; the most amyloidogenic points for mutations are in the edge of C strand, CD loop, D-strand and in some points on DE-loop (Figure. III-1) [13, 15].

The autosomal dominant systemic familial amyloidotic polyneuropathy (FAP) can be caused by mutations in several genes such as apoliprotein I, apoliprotein II, fibrinogen, lyzomyme and

transthyretin. Among these mutations, the one with the most prevalence to develop amyloidosis (around 90% of the cases) occurs in TTR [22]. FAP is expressed in all ages (≥20 year) and the fibrils are accumulated in peripheral nerves and muscular tissues, promoting digestive, urinary, sexual and cardiovascular dysfunctions [16]. It was described that even before the abnormal TTR forms aggregate into amyloidosis deposits, they can have a toxic property on tissues and the peripheral nerves [20].

The first mutation in TTR that was associated to FAP was the Val30Met, common among people from Sweden, Portugal, Japan, and Brazil. Additionally, other mutations promote amyloid deposits in cardiac tissue, causing heart failure; this type of TTR amyloidosis is called the familial amyloidotic cardiomyopathy (FAC). The TTR variant Val122Ile is the most prevalent mutation in FAC, observed in 3-4% of African-American population [15].

Moreover, TTR is the most abundant protein observed in elderly individuals with fibrils formation in case of systemic amyloidosis (SSA) [13, 14]. In SSA, the fibrils are composed of a wild type TTR and it affects ~25% of the population over the age of eighty [15]. According to Johnson and coworkers (2012) aging is a risk factor for the development of amyloidosis, because aging-associated deficiencies may affect protein homeostasis. In addition, unknown factor(s) and/or protein can contribute to the development of the amyloidosis process [21, 23].

In TTR sequence there is a single cysteine residue per monomer (Cys10); this residue is exposed in the beginning of the helical region, near strand A (see Figure.III.16) [14, 24]. The Cys10 can form a disulfide bond with various reactive thiol molecules, as adducts of sulfonate, cysteine, homocysteine (Hcy), etc. However, 10-15% of the Cys10 are known to be the reduced form. In healthy and amyloidosis patients, TTR tetramer is composed of a mixture of Cys10 modified species [25, 26]. The most stable link is observed between Cys10 and sulfite; this molecule provides an extra hydrogen bond between one of oxygen of sulfite and the main chain at Gly75 [16]. Several studies showed that this residue played an important role in protein stability; it has been reported that the ratio of TTR with the modified Cys to TTR with a free Cys is high in patients with FAP [25, 26]. According to Giustarini and coworkers (2006), the thiol composition in proteins correlates with age; in this case, the increase in thiol contents, induced by age, may contribute to SSA development, which is attributed to the destabilization of the TTR structure caused by the Cys modification [26-28].

Amyloidosis causes millions of deaths every year; many researchers are trying not only to understand the causes for fibril formation, but also to produce drugs to inhibit the amyloidosis

process [15]. As such, some non-invasive tests have been carried out by screening TTR molecular masses of the mutated ones or the modified species in the blood stream; a shift in mass and/or charge state in ESI-MS can be correlated to amino acid mutations or the Cys modification. A mass shift greater than 10 Da was observed in some TTR variants, comparing to the wild type [22]. Also, studies about TTR thermodynamics suggest whether or not a mutant protein can form fibrils; the rate of tetramer dissociation (kinetic stability) implies the degree of amyloid fibril formation or the age of onset [15].

Amyloidosis process can be related both to the modification of Cys10 in TTR and to diseases caused by hyperhomocysteinemia or homocystinuria, e.g. developments in chronic renal failure and cardiac disorders. In these cases, the level of cysteine and homocysteine in the plasma increases, and, as the result, the proteins with free thiol groups can easily be modified. Lim and colleagues (2003) observed that for the patients with end stage renal disease, the level of TTR-Cys-Cys-Hcys increases, where homocysteine can react with Cys10, producing thiolate TTR, which may increase TTR's propensity to develop pathophysiological conditions [24]. From several lines of evidence, the modifications at the Cys of TTR can be used as a marker for disorders in sulfur metabolism and aging process [27, 29].

Oxidative Stress and related diseases

The presence of oxygen inside the organism can be seen as a bittersweet existence – on one hand, this molecule is used by the electron transport chains and some enzymes in the respiration process; on the other hand, the system suffers from the toxic effect of O_2. In this sense, the oxidative stress is classified as an imbalance between the production of reactive oxygen species (ROS) and the activity of the antioxidant defenses [30]. Since 1950s many researchers have been interested in oxidative stress and how this phenomenon is related to the aging process and the development of

some diseases. Biological systems can produce reactive oxygen species and free radicals through normal metabolic processes, but, at the same time, the organisms are capable of protecting themselves from the toxicity of these species. In humans, a defense system against ROS-damages can be observed: enzymes represented by catalases and glutathione peroxidase remove these species from the organism. Catalase is placed in the peroxisomes and the glutathione peroxidase is in the cytosol and mitochondria, in which the superoxide dismutase (SOD) enzyme is placed [30]. Additionally, the uneven ratios between the redox couples, oxidized and reduced, such as glutathione (GSH/GSSG) or NADPH/$NADP^+$, can cause oxidative stress, which is unrelated to overproduction of free radicals [31].

Human bodies are perpetually exposed to free radicals, not only generated by electromagnetic radiation and other environment sources such as pollutants, but also by cosmic radiation and cellular metabolism. Damages to the biomolecules generated by free radicals lead to many chronic disorders (Figure III.4). The brain and the neuronal tissue are sensitive to free radicals; when neurons and glial cells are attacked by ROS, it causes neuronal damages. Moreover, it has been described that reactive oxygen species may stimulate the production of new antigens, which provoke immune responses [30]. Atherosclerosis, cancer, diabetics, rheumatoid arthritis, post-ischemic perfusion injury, myocardial infarction, cardiovascular diseases, chronic inflammation, stroke and septic shock, aging, and other degenerative diseases are some of the disorders associated with oxidative stress [32].



Figure III.4. Effects of the oxidative stress in organism. Some diseases caused by the presence of free radicals or reactive oxygen species in the organs. This figure was modified, based on the scheme from the Online Resources for Disorders Caused by Oxidative Stress (**http://www.oxidativestressresource.org/**).

Free radicals such as the superoxide anion ($O_2^{\cdot-}$) and, hydroxyl radical (·OH) can generate hydrogen peroxide (H_2O_2) and these reactive species lead to cell damage or cell death, due to DNA damage, protein and lipid oxidation and cellular dysfunction [33]. As proteins are the most abundant compounds in intra and extracellular environments, the likelihood of interaction between proteins and free-radicals is high. The oxidative stress alters the function of receptors, transport proteins, enzymes, etc. [30, 34].

During oxidative stress, free radicals attack proteins, at the side chains of amino acid residues, generating other free radicals, which may be transferred from other amino acids in the protein; the

reaction causes a modification in the electrical charge net, which may induce protein fragmentation. This process occurs more frequently at the α-carbon. Modifications in the primary structure cause protein denaturation and increase the hydrophobicity degree due the structural changes. At the same time, new covalent bonds could also be formed among the modified compounds, which promote protein aggregation. Meanwhile, when the proteins are oxidatively modified *in vivo*, intracellular proteolytic mechanisms could recognize damaged proteins and degrade them [35, 36]. In the case of enzymes, when they have free thiol groups in the catalytic pocket, they could be modified, inactivating the enzymes [33].

Furthermore, the oxidative products of lipids or carbohydrates induce protein modification. In this case, carbonyl groups are produced and can react with proteins [30, 37]. Products from lipid peroxidation can be found far away from where they are initially produced, because these aldehydes easily diffuse across the membranes. Malondialdehyde (MDA) is the most abundant carbonyl-terminal molecule in the circulation; it is produced due to lipid oxidation of the cell membranes. MDA in the circulation may form cross linkers between proteins or protein and phospholipids; it forms adducts with lysine residues or amine head-groups of phospholipids. Both MDA and other aldehydes can react with the sulfur atom of cysteine and the imidazole nitrogen of histidine [38, 39].

It is well known that elevated levels of ROS directly modify various proteins in the blood, because blood is exposed to severe oxidative stress, comparing to any other biological fluids [37]. In this regard, proteins play an important role to control or depress the oxidative stress in this biological fluid; 10-50% of potential antioxidants in the plasma are proteins [34]. Once prokaryotes and eukaryotes are exposed to oxidative stress, the organisms can respond to it, using a conserved response mechanism such as expression of oxidant scavengers, proteases, chaperones, and DNA repairing enzymes [40].

Free radicals and ROS have short lifetime and, are thus difficult to detect *in vivo*, but could be detected in indirect manners, in which modified proteins and lipids may be the surrogate markers of oxidative stress [37, 38]. TTR and serum albumin have a free Cys residue (position 34 of TTR), which can react with chloramines to form sulfonic acid. The sulfhydryl group at Cys is likely to act like an antioxidant in the plasma. However, the oxidation occurs in an irreversible manner and makes the protein unstable; in fact, high levels of modified TTR have been observed in nephrotic patients. The proteins modified by ROS retain their fingerprints in the modified residues, enabling this signature as a potential biomarker of oxidative stress; however, few markers for protein oxidation *in vivo* are available [37, 41].

III.2. Material and Methods

The experimental and analytical procedures, reagents and instruments were the same as described in the section 2.3., except those described below.

Material

The human plasma and the proteins trypsin and ribonuclease A were purchased from Sigma (St. Louis, MO). The micro BCA protein assay kit was purchased from Termo Scientific (Rockford, IL). The primary antibody against TTR was from Santa Cruz Biotechnology (Santa Cruz, CA) and the second antibody from BioRad (California, US). The western lightening ECL reagent was purchased from Perkin Elmer (Waltham, MA). The IgG and albumin depletion kit and the chromatographic resin Thiopropyl Sepharose 6B were purchased from GE Healthcare (Buckinghamshire, UK). The resin Toyopearl AF-Amino 650 was purchased from Tosoh Bioscience (Tokyo, Japan). The other chemicals were of analytical grade and were purchased from Nacalai Tesque (Kyoto, Japan).

Sample preparation

The dried human plasma was resuspended in phosphate buffered saline (PBS – 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 and 2 mM KH_2PO_4 pH 7.4) and then it was fractionated by molecular mass using Amicon® Ultra-0.5 centifugal filter (cut off 10 kDa). The peptide fraction (≤10 kDa) was recovered and stored at -80°C. The protein fraction was depleted using the albumim and IgG kit according to the manufacturer's instructions. Forty microliters were used on the depletion step, and half of the depleted sample was used for the Poros 20 NH LC-SPDP system. The PBS buffer was replaced by the working solution (50 mM NH_4CO_3H, 25% ethanol, 100 mM

NaCl and 1 mM EDTA pH 7.8) after the depletion step. Following the protocol proposed by Hillson, (1980, with some modifications) [42], 0.1mM of dithiothreitol (DTT) was added to the working solution to activate the thiol groups present in the proteins. The sample was diluted in the working solution containing DTT and incubated in 30°C for 30 minutes. This concentration of the reducing agent is enough to activate the thiol groups in the protein, but the disulfide bonds were not broken. The DTT was removed using the ultra-0.5 centrifugal filter (cut off 10 kDa). The reduction solution was replaced by the working solution. Even using this weak reducing condition, it was possible to observe some protein aggregates – which were removed from the tests by centrifugation.

Isolation of Proteins containing free thiol using LC-SPDP

The solution containing the proteins was exposed to the Poros 20 NH LC-SPDP resin for 12 hours, in batch incubation, at 4 °C to avoid proteolysis in the sample. After the incubation time, the resin was centrifuged at 2000g for 30 seconds. The supernatant was removed (Flow-through) and the resin was exhaustively washed with the working solution to remove all the proteins that were not attached to it. Bound proteins were recovered incubating the chromatography system for 1 hour at room temperature with a reducing solution, containing 50 mM of DTT. For posterior analysis, the samples – Flow-through and Eluted fraction – were desalted using the ultra-0.5 centrifugal filter (cut off 10 kDa).

Polyacrylamide Gel electrophoresis (PAGE)

SDS- PAGE was carried out as described by Laemmli (1970) [43]. Sample were dissolved in 50 mM Tris-HCl buffer, pH 6.8, containing 0.1% bromophenol blue, 2% SDS and 10% sucrose, and they were applied to 12.5% acrylamide gel (9 cm x 6 cm x 0.1 cm). The running buffer contained 25 mM Tris- 192 mM glycine and 0.1% SDS, pH 8.4. PAGE was conducted at a constant

current of 20 mA for 80 minutes. The SDS-PAGE was stained using the silver staining protocol proposed by Blum (1987) and also Coomassie Blue G250 [44].

In-gel digestion of proteins

To identify the protein bands present in the SDS-PAGE, the gel bands were minced in 1-2 mm blocks. A solution of 50% methanol and 50mM NH_4CO_3H was added and the blocks were incubated at room temperature for 1 hour. The washing step was repeated twice. The solution was removed and a reducing solution containing 10mM of DTT in 50 mM NH_4CO_3H was applied to each block. The gel was incubated for 1 hour at 60°C. The solution was removed and 50 mM acrylamide in 100mM NH_4CO_3H was added to alkylate the cysteine residues. The sample was incubated for 30 minute at room temperature. The acrylamide solution was replaced by deionized water. The procedure was repeated twice. The blocks were washed with acetonitrile and the peptides fragments were produced using 20 pmol of trypsin. The digestion was carried out at 37°C overnight. The supernatant was collected. After this, a solution of 50% acetonitrile/0.1% TFA was added to quench the enzyme. The samples were sonicated for 10 minutes and the supernatant was recovered. The process was repeated twice. The supernatants were dried.

Protein identification

The tryptic fragments obtained after in gel digestion were analyzed via nano-flow LC linear ion-trap time-of-flight (TOF) mass spectrometer (NanoFrontier LD, Hitachi, Tokyo, Japan). The liquid chromatography was performed on a PicoFrit C18 column (75 µm x 100 mm, particle size of 3µm; New Objective, Wobum, Massachusetts). The peptides were eluted at 200 nl/min using a 2%–49% acetonitrile gradient over 45 min in the presence of 0.3% formic acid. The mass spectrometer

was operated in the data-dependent MS/MS mode. Precursor ions were measured by the TOF mass analyzer and subjected to fragmentation in the linear ion trap. Identification was performed in the Swiss-Prot database (SwissProt 2012 07) with *Homo sapiens* taxonomy (20,350 sequences). Mascot (version 2.3, Matrix Science, Manchester, UK) was used as the search engine. The variable amino acid modification was oxidized methionine, and carbamidomethylation of cysteines was set as the fixed modification.

Mass spectrometry

All mass spectrums were acquired on 4800 MALDI-TOFMS (Applied Biosystems, Framingham, MA, USA). A matrix, sinapinic acid, was prepared at the concentration of 10 $mg.ml^{-1}$ in 30% acetonitrile/ 0.1% TFA. The sample (0.5µl) was spotted onto the MALDI plate and mixed with the matrix (0.5µl) and allowed to dry at room temperature. The instrument was operated in the linear mode and the mass spectra were randomly acquired over the entire sample surface using an average of 300 laser shots. Mass calibration was accomplished using the $[M+H]^+$ ions of ribonuclease A (MW 13683.3, Sigma), trypsin (MW 23 449.7, Sigma) and bovine serum albumin (MW 66 430.1, Sigma). The calibration molecules were mixed with the matrix and deposited on the membranes either on or near the blotted areas.

Modification of a free thiol by the N-ethylmaleimide (NEM)

The free cysteine residues present in plasma proteins were alkylated using the N-Ethylmaleimide (NEM). Fifty milimolar of NEM in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 and 2 mM KH_2PO_4 pH 7.4) was added to the plasma and the sample was incubated at 37°C for 2 hours. The reaction was quenched by adding 25 mM cysteine to the solution; the sample

was incubated for another 30 minutes at room temperature. NEM excess and cysteine were removed from the solution by centrifugation using the ultra-0.5 centrifugal filter (cut off 10 kDa).

Protein detection by immunoblotting

Using the protocol proposed by Harlow and Lane (1988) [45], after the standard procedures for electrophoresis process, the proteins from the plasma were transferred from the gel onto the nitrocellulose membrane using a transfer buffer containing Tween-20 in tris buffered saline (TBS – 50 mM Tris-HCl, 150 mM NaCl pH 7.4). The process was carried out at a constant current of 100 mA and 20 V for 60 minutes. Then, the membrane was washed with the transfer buffer and, to avoid nonspecific binding, the proteins in the nitrocellulose membrane were blocked using 1% BSA in Tween-20 in TBS. The membrane was incubated with the solution at room temperature with agitation for two hours. After the blocking step, the nitrocellulose paper was washed twice for 5 minutes using the Tween-20 in TBS, and the primary antibody against TTR (rabbit polyclonal anti-prealbumin – 1:200 dilution) was applied on the paper. The blot was incubated with the primary antibody for at least one hour; then, the paper was washed to remove the primary antibody excess. The second antibody, a rabbit anti-IgG (1:200), was exposed to the membrane for at least 1 hour at room temperature. The blot was washed with Tween-20 in TBS and the target protein that was specifically recognized by the antibody was visualized using enhanced chemiluminescence reagents in a dark chamber. The blot images were processed as 16bit in TIFF format. The images were analyzed using the software ImageJ (National Institute of Health, Maryland – US). Using the software the area, pixel value statistics and intensity of the protein bands in the membrane were calculated.

Effect of the oxidative stress on TTR

To observe the effect of the ROS on TTR present in the plasma and also how this protein reacts with the Poros 20 NH LC-SPDP system after the oxidative stress, the plasma protein fraction was pretreated with hydrogen peroxide (H_2O_2). Based on the protocol proposed by Fligiel and coworkers (1984) [46], the sample was diluted in phosphate buffer 20mM 100mM NaCl pH 7.4 and 1-10 mM of H_2O_2 was added. The sample was incubated for 1 hour at 37°C. Before applying the sample to the Poros 20 NH LC-SPDP system, the sample was concentrated and the H_2O_2 removed by centrifugation, using the ultra-0.5 centrifugal filter (cut off 10 kDa). Further, the phosphate buffer was replaced by the working solution during the process. After the chromatographic process, the protein concentration in the Flow-through fraction was measured using the micro BCA protein assay kit.

Thiopropyl Sepharose 6B and Toyopearl AF-Amino 650

The commercial available resins Thiopropyl Sepharose 6B and the Toyopearl AF-Amino 650 were handled as recommended by the specification sheet. One gram of the resin Thiopropyl 6B was used in each experiment. The resin was suspended in distilled water and washed using the same solvent for 30 minutes. After the washing process, the thiopropyl 6B was washed using the working solution applied to the Poros 20 NH LC-SPDP system (50 mM NH_4CO_3H, 25% ethanol, 100 mM NaCl and 1 mM EDTA pH 7.8). Then, the solution containing the proteins was exposed to the resin for 12 hours, in batch incubation, at 4°C to avoid proteolysis in the sample. After, the unbound proteins, Flow-through fraction, were recovered and the resin was exhaustively washed with the working solution to remove all the proteins that were not attached to it. Bound proteins were recovered incubating the chromatography system for 1 hour at room temperature with a reducing solution, containing 50 mM of DTT.

One milliliter of the resin Toyopearl AF-Amino 650 was washed using distilled water for 30 minutes. After this, the resin was washed using 10 mL of ethanol. Twenty millimolar of LC-SPDP was dissolved in ethanol and applied to the resin, which was incubated with the heterobifunctional group for 4 hours at room temperature. The excess of LC-SPDP was removed washing the resin

with 20 mL of ethanol. The system Toyopearl AF-Amino 650 LC-SPDP was washed with 20 mL of the working solution, before exposing the system to the sample; the standard procedures used for the Poros 20 NH LC-SPDP were also applied to the Toyopearl system. For posterior analysis, the samples – Flow-through and Eluted fraction from Thiopropyl Sepharose 6B and Toyopearl AF-Amino 650 LC-SPDP – were desalted using the ultra-0.5 centrifugal filter (cut off 10 kDa).

III.3. Results

To analyze the proteins containing free thiol groups in the plasma, a commercial source of plasma, obtained from pooled human blood, was used. First, peptides were removed by centrifugation using molecular-weight 10 kDa-cutoff filter, and then, abundant proteins (albumin and IgG) were eliminated by the depletion kit (see Materials and Methods). Then, the protein fraction was applied to the POROS 20 NH LC-SPDP system. After 12 hours of incubation, the proteins that were not linked to the resin were removed by exhaustive washing of the resin with the working solution. The Flow-through fraction was monitored at 280 nm. The resin was washed until the effluent showed little absorbance at 280 nm. Then, the working solution containing a reducing agent was applied to the resin, rendering bound proteins to be liberated from the resin. The Flow-through and the Eluted fractions were analyzed by SDS-PAGE; in the Flow-through fraction, many proteins could be seen, meanwhile, a few protein bands were present in the Eluted fraction. Among these proteins, one with molecular mass ≈14 kDa was the most abundant in the sample (Figure III.5). In order to identify the compound, the band was extracted from the SDS-PAGE and subjected to in-gel digestion (see Material and Methods). The resultant digested peptides were analyzed by nano-flow LC linear ion-trap time-of-flight (TOF) mass spectrometer. The peaks in the spectra were identified using the MASCOT software and the protein observed in the SDS-PAGE was identified to be human transthyretin (TTR).

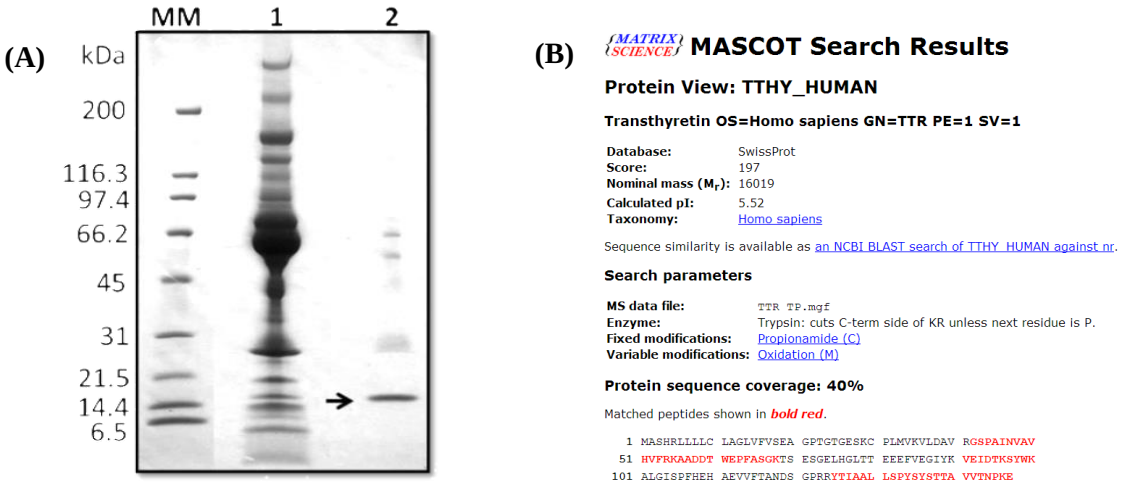


Figure III.5. Identification of transthyretin in human plasma. (A) SDS-PAGE 12.5% of human plasma. MM is the molecular mass standards. Lane 1 represents the Flow-through fraction from the Poros 20 NH LC-SPDP system. Lane 2 represents the Eluted fraction from Poros-LC-SPDP system. The protein identified by mass spectrometry is indicated by the arrow. (B) Mascot data sheet of the protein identified as TTR.

TTR has been described in many studies as a homotetramer; however, it is well established that the most stable form is the dimer. Nonetheless, on the SDS-PAGE only the monomer was observed. In order to observe the dimeric form of TTR, the Eluted fraction from the plasma was not heated at 90°C before applying the sample to the electrophoresis; when the sample was not denatured by high temperatures the dimer could be observed in the SDS-PAGE (Figure III.6).

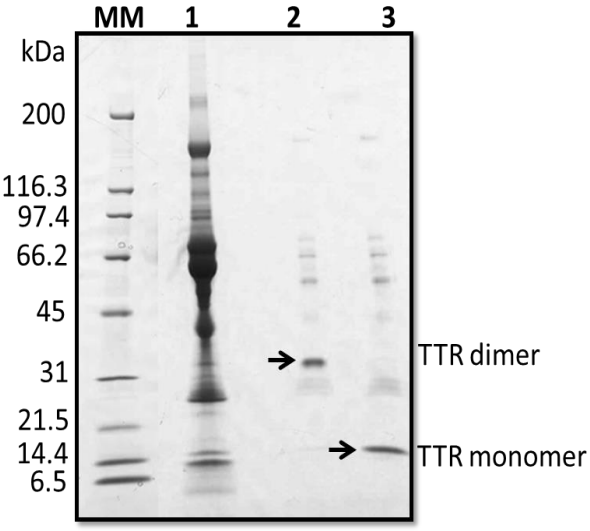


Figure III.6. SDS-PAGE 12.5% of human plasma. MM is de molecular mass standards. Lane 1 represents the Flow-through fraction from the Poros 20 NH LC-SPDP system. Lane 2: the Eluted fraction from the Poros 20 NH LC-SPDP system. The sample was not boiled at 90°C before applying it to the electrophoresis. The arrow indicates the TTR dimer. Lane 3: the Eluted fraction from the Poros 20 NH LC-SPDP system boiled at 90°C.

In addition, the Eluted fraction was analyzed by MALDI-TOFMS, which predominantly gave the molecular mass (*m/z* 13,761) of the monomer. Of note, the dimer, trimmer and tetramer observed in the same spectrum were ascribed to the non-physiological complex formed during MALDI process (Figure III.7).

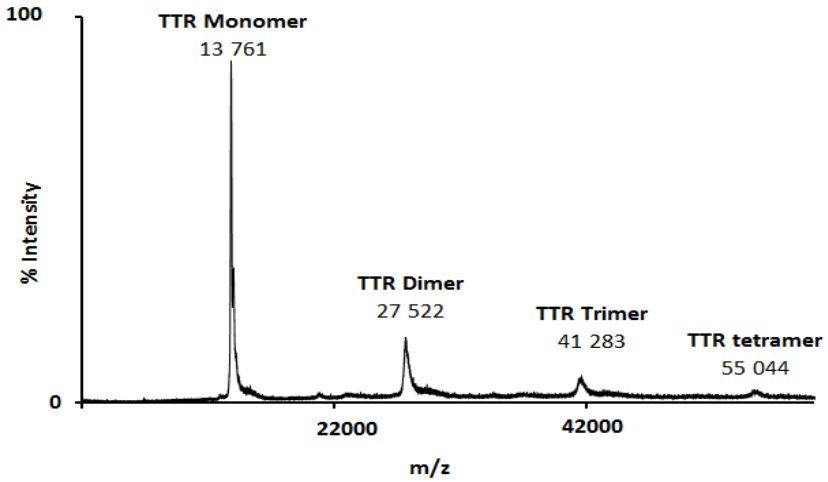


Figure III.7. MALDI-TOFMS spectrum of the Eluted fraction obtained from Poros-LC-SPDP system. Each peak corresponds to TTR: monomer (13 761 Da), dimer (27 522 Da), trimer (41 283 Da), and tetramer (55 044 Da).

TTR has a sole cysteine residue at position 10, which could be reversibly modified. As described in Chapter II, the usage of a low concentration of reducing agent could enable reversibly modified Cys to revert to the free thiol Cys. As such, 0.1 mM and 0.5 mM DTT were used to reduce the reversible disulfides, if present, between Cys10 and some small molecules. Subsequently, the sample was applied to the Poros 20 NH LC-SPDP system. SDS-PAGE of the Eluted fractions obviously showed increased recoveries of TTR while other several proteins, which probably contain disulfide bonds, were simultaneously obtained (Figure III.8).

The Eluted fraction obtained from the sample treated with 0.5 mM DTT was reduced, alkylated, and digested with trypsin, and then, the resultant tryptic peptides were subjected to nano-flow LC-MS/MS or MALDI-TOFMS (Materials and Methods). As a result, TTR was mainly identified as well as alpha-1-acid glycoprotein 2 (23.7 kDa), coagulation factor X (54.7 kDa), serum albumin (66.4 kDa), complement factor I (65.7 kDa), C2 domain-containing protein 2 (75.5 kDa),

and complement C3 (187.1 kDa). All these proteins possess cysteine residues that are involved in inter- or intra-disulfide linkages or free thiols.

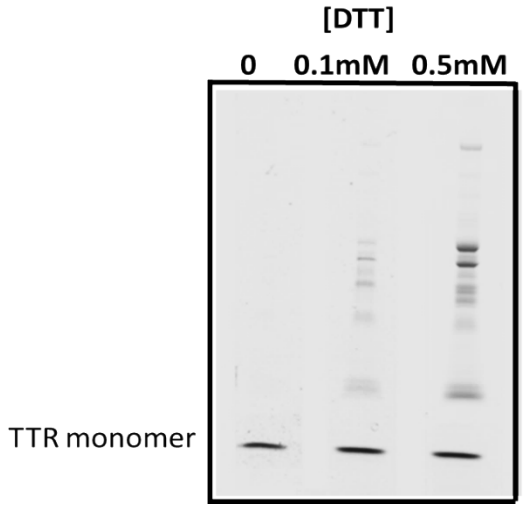


Figure III.8. SDS-PAGE 12.5% of human plasma. In this experiment the effect of the reducing agent on the plasma proteins was evaluated using the Poros 20 NH LC-SPDP system. The sample has been treated with 0, 0.1, and 0.5 mM DTT before applying it to the Poros 20 NH system.

A western blot analysis was also performed using the samples derived from the Poros 20 NH LC-SPDP system. In this experiment, the samples treated with and without 0.1 mM DTT were subjected to the western blot analysis. Following the electrophoresis process, the proteins were transferred to the nitrocellulose membrane and treated with the anti-TTR antibody (rabbit polyclonal antibody – Figure III.9). The amounts of TTR in the Flow-through fraction (probably TTR modified at Cys) and in the Eluted fraction (free thiol form) were estimated. Using the software ImageJ, the intensity of the fluorescence liberated by chemicals, due to the recognition of TTR by the antibody, was calculated for each fraction, which implies the concentration of the proteins in the samples. As a result, the ratio of the free-form of TTR (Eluted fraction) to the modified form (Flow-through fraction) in the commercially available plasma was estimated to be

1:4 (for the fraction without DTT) and 1:3 (for the fraction treated with 0.1 mM DTT). At the same time, the analysis of the sample that has been treated with 0.1 mM DTT showed that the amount of free thiol form of TTR significantly increased about 1.5 folds (Lane II in the right panel of Figure III.9).

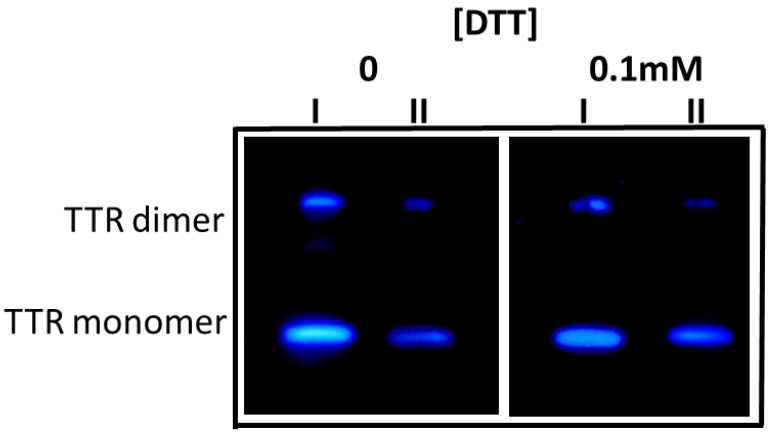


Figure III.9. Western blot analysis of the plasma sample with anti-TTR antibody (rabbit polyclonal antibody). The samples were treated with (right panel) and without (left panel) 0.1 mM DTT before applying the sample to the Poros 20 NH LC-SPDP system. The overall concentration of TTR increased after pretreatment with the reducing agent. Lane I and II showed the Flow-through and Eluted fraction, respectively.

In order to identify the modified form of TTR, the Flow-through fraction, after desalting, was subjected to MALDI-MS in *m/z* range from 10,000 to 70,000. As a result, several ion peaks were observed, among which the ion peak at *m/z* 66,567.8 (MH^+) and 33,410.8 ($[M+2H]^{2+}$) could be, most probably, assigned to albumin, based on the fact that it is the most abundant protein in plasma (Figure III. 10). In addition, there was the ion peak at *m/z* 13,875.6, which was larger in mass by

119 Da than the intact TTR, which could most probably correspond to a cysteine adduct. To further confirm the linkage of the Cys via a disulfide, the Flow-through fraction was treated with 5 mM of DTT for 60 min at 60°C, then analyzed by MALDI-MS, giving the molecular mass at *m/z* 13,761.2, which agreed well with the mass of the intact TTR. Moreover, the ion peak observed at *m/z* 13930.6 could be ascribed to S-cysteinylglycine-linked TTR. In fact, S-glutathionated (γ-Glu-Cys-Gly) TTR as well as the S-cysteinylated one have been reported previously as the major modified species in the plasma [47].

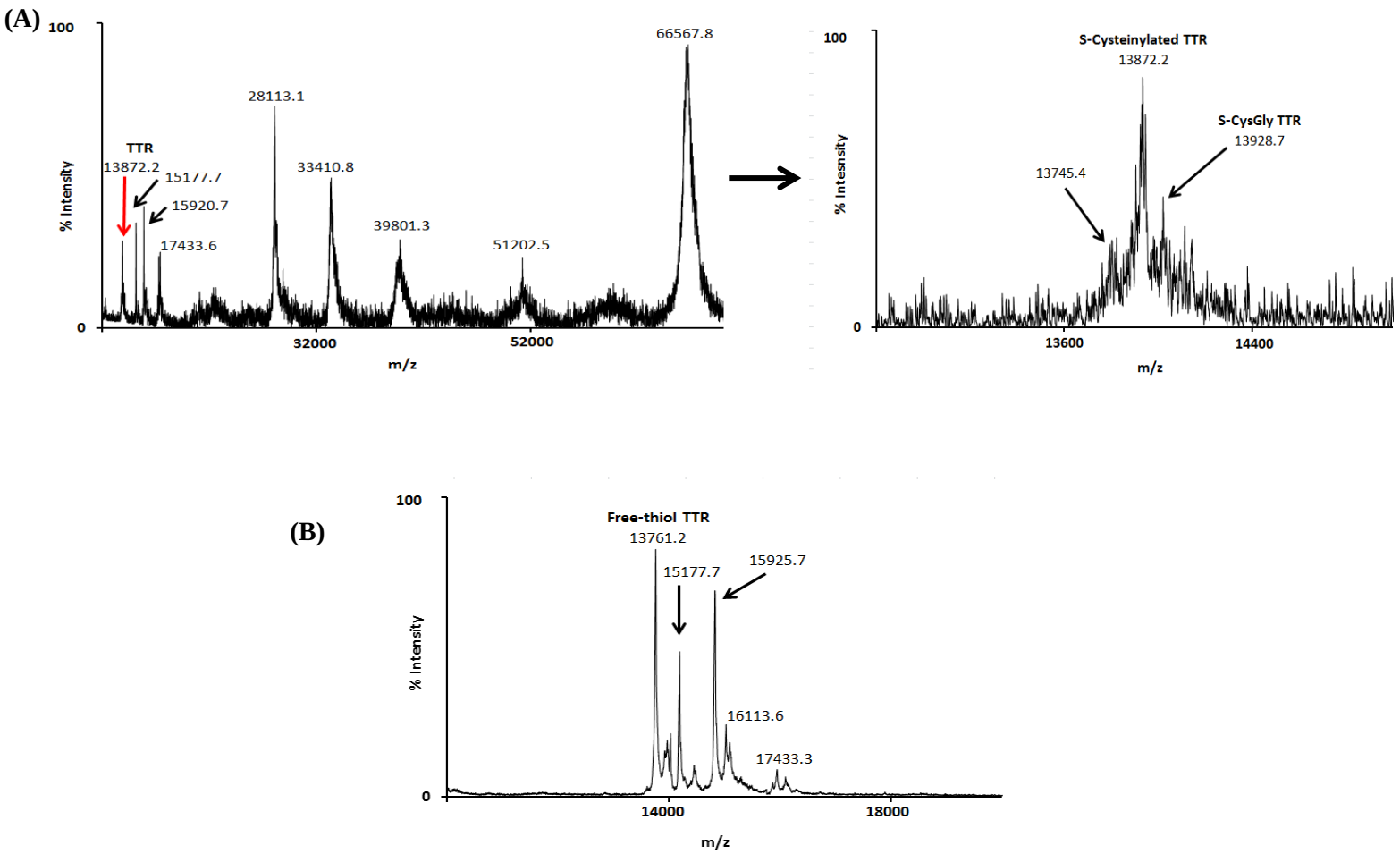


Figure III. 10. MALDI-TOFMS spectra of the Flow-through fraction obtained from the Poros 20 NH LC-SPDP system. (A) The Flow-through fraction, the modified forms of TTR, can be observed in the spectra. (B) Flow-through fraction treated with 5 mM DTT for 60 min at 60°C. After the process, only the free thiol form of TTR can be observed.

To demonstrate that TTR is linked to LC-SPDP via the thiol group of Cys10, the plasma sample treated with N-ethylmaleimide (NEM), a reagent which promptly reacts with a thiol, was applied to the Poros 20 NH LC-SPDP system. In contrast to both of the intact sample and that treated with 0.1 mM DTT, the NEM-treated sample did not show any TTR band in the Eluted fraction (Figure III.11), indicating that the TTR recovered in the Eluted fraction possesses a free thiol at the sole Cys10.

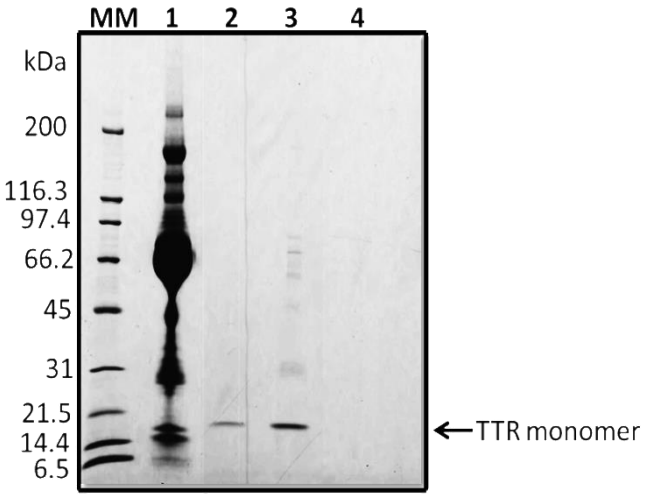


Figure III.11. SDS-PAGE 12.5% of human plasma. MM is the molecular mass standards. Lane 1: Flow-through fraction; lane 2: Eluted fraction; lane 3: Eluted fraction prepared from the sample treated with 0.1 mM DTT; lane 4: Eluted fraction prepared from the sample treated with 50 mM NEM at room temperature for 2 hours.

The standard procedure of the aforementioned method utilized a 12 hour incubation to accomplish the disulfide exchange of LC-SPDP with thiols in the sample. However, since free thiols are not stable and tend to be oxidized during incubation, the time for the incubation should be

optimized and shortened as much as possible. In the case of TTR in the plasma, 4-hours incubation turned out to be enough (Figure III.12).

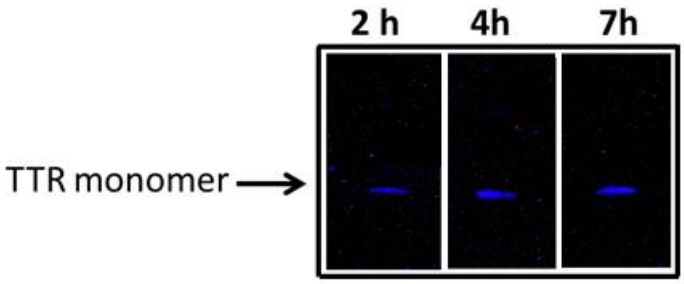


Figure III.12. Optimization of incubation time for linking TTR to the Poros 20 NH LC-SPDP resin. The TTR prepared from the Eluted fraction was applied to the resin, and after 2, 4, 7 hours of incubation, the TTR was eluted with the reducing solvent and applied to SDS-PAGE. The TTR was visualized by Western blotting (see Figure. III.9).

The presence of free thiol groups in proteins has been described as the mechanism for maintaining the homeostasis in case of oxidative stress, because this group is sensitive to free radicals or reactive oxygen and nitrogen species. In this regard, an experiment was performed to examine the effect of hydrogen peroxide (H_2O_2) on TTR and how the modification caused by this oxidative species could be monitored by using the Poros 20 NH LC-SPDP system that enables distinct separation of TTR with a free thiol at Cys10 from that with modified Cys (Figure. III.7, III.10). The protein fraction (>10 kDa and albumim and IgG-depleted, see Material and Methods) from the plasma was treated with different concentrations of H_2O_2 (0, 2, 5, 10 mM) at room temperature for 60 min. Following evaporation of H_2O_2 under vacuum, the residues were dissolved and subjected to the Poros 20 NH LC-SPDP system. Each resultant Eluted fraction was applied to

SDS-PAGE (Figure III.13). The TTR with a free thiol at Cys, which was recovered in the Eluted fraction (0 mM H_2O_2 in Figure III.13), significantly reduced after treatment with 2 mM H_2O_2, and completely disappeared with 10 mM H_2O_2 (Figure III.14). The result strongly suggested that the free thiol at Cys10 was oxidized with relatively low concentration of H_2O_2 (5 mM), rendering TTR modified ones that were recovered in the Flow-through fraction. In fact, the protein concentration of the Flow-through fraction increased upon treatment with H_2O_2 (Table III.1), suggesting that TTR and other proteins as well, seen in the Eluted fraction (Figure III.13), were then accumulated in the Flow-through fraction, most probably, owing to the free thiols of those proteins being modified with H_2O_2. The amounts of TTR in the Flow-through fractions increased 3- and 3.8-times, comparing with that of the intact one, after treatment with 5 and 10 mM H_2O_2 respectively (Figure III.14).

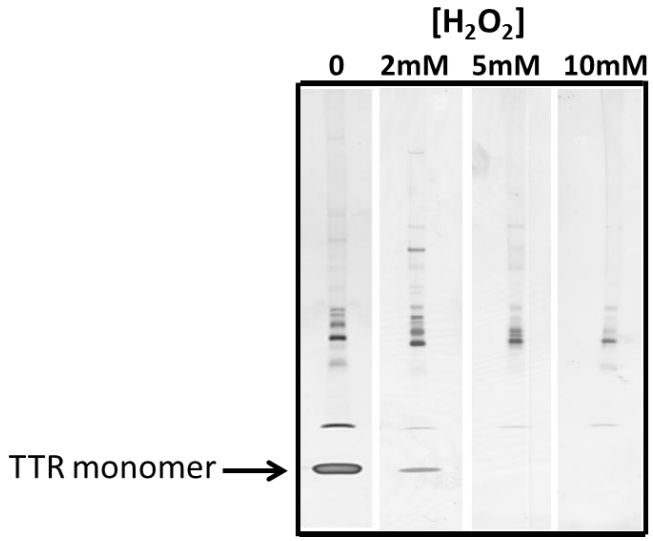


Figure III.13. Effect of H_2O_2 treatment on the recovery of the free-Cys form of TTR. SDS-PAGE 12.5% of the Eluted fractions obtained from the human plasma. The human plasma was treated with 2mM, 5mM and 10mM H_2O_2, pre-purified, and applied to the Poros 20 NH LC-SPDP system resulting in the Flow-through and Eluted fractions (see Materials and Methods).

Table III.1. Protein concentrations, estimated by micro BCA protein assay kit, of the Flow-through fractions obtained in the experiment described in Figure III.14.

Concentration of H_2O_2 (mM)	Protein Concentration ($\mu g.ml^{-1}$)
0	305.5
5	385.7
10	410.2

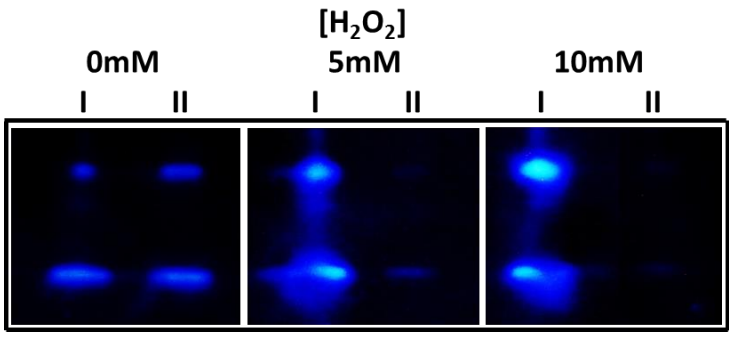


Figure III.14. Western blot analysis with anti-TTR antibody of the Flow-through (I) and Eluted (II) fractions obtained Fig. III.14. The free-Cys form of TTR completely disappeared after 10 mM H_2O_2 treatment.

As discussed previously in Chapter II, there are several methods to isolate proteins containing free thiol groups, among which the most used technique is the covalent chromatography using the Thiopropyl Sepharose 6B. In regard to specificity and recovery, the present method was assessed in comparison with that using Thiopropyl Sepharose 6B resin. Moreover, the resin Poros 20 NH was

replaced by the Toyo Pearl AF-Amino 650 resin and the heterobifunctional groups LC-SPDP was linked to it. Both resins were handled as described in the Material and Methods section; the same solvents and conditions for the elution and the same amounts of the plasma sample were used for all the separation systems. The recovery of free-Cys form of TTR in the Eluted fraction using the Poros 20 NH LC-SPDP system was higher than that obtained with the Thiopropyl 6B (Figure III.15.A) and Toyo Pearl AF-Amino 650 LC-SPDP (Figure III.15.B); several protein bands in the higher molecular mass were observed with the Thiopropyl 6B and Toyo Pearl AF-Amino 650 LC-SPDP, which could probably be derived from non-specific bindings, whereas the TTR could be solely obtained with the present method.

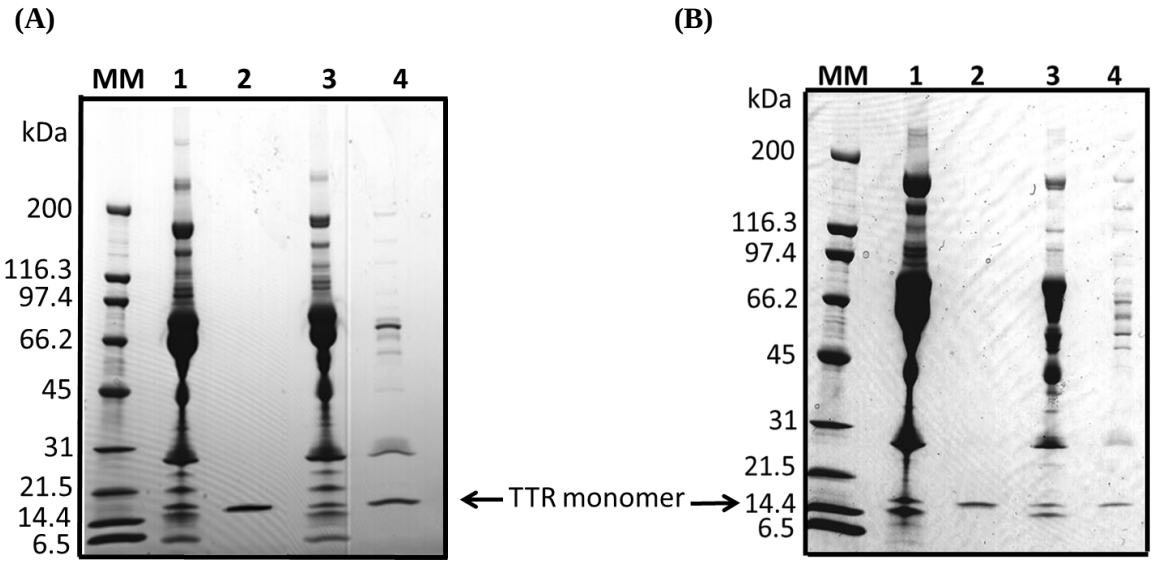


Figure III.15. Isolation of free thiol containing proteins from human plasma. (A) Comparison between the Poros 20 NH LC-SPDP system (lane 1 and 2) and Thiopropyl Sepharose 6B (lane 3 and 4). Lanes 1 and 3 were the Flow-through fractions; lanes 2 and 4 were the Eluted fractions. MM is the molecular mass standards. (B) Comparison between Poros 20 NH LC-SPDP (lane 1 and 2) and Toyo Pearl AF-Amino 650 LC-SPDP (lane 3 and 4). Lanes 1 and 3 were the Flow-through fractions; lanes 2 and 4 were the Eluted fractions. MM is the molecular mass standards.

III.4. Discussion

In the plasma most proteins produced by the organism can be found, as well as proteins derived from microorganisms such as bacteria and fungi and also from virus. Some of these proteins are present in low concentration ($pg.ml^{-1}$), while others are present in large amounts ($mg.ml^{-1}$) [9]. In this sense, the measurements of low- and high-abundant proteins have been essential for medical diagnosis, because this information could be used to classify and identify the stage of the disease, its progression, and the effectiveness of medical treatment [6].

In general, the plasma proteins that contain cysteine residues form disulfide bonds, which play an important role in protein structure [48]. Conversely, only 2-3% of the proteins contain free thiols, and such reactive group is involved in protein function and regulation [49, 50]. In this regard, the present method could specifically isolate these free thiol containing proteins, and this property is useful for both reduction of the complexity of biological samples and study on the role of those proteins in the environment.

At first, when the protein fraction derived from the plasma was analyzed by the present method, a protein band at ≈14kDa that was identified as transthyretin (Figure III.5) was observed. Albumin has been described as the main plasma protein, which contains a free thiol group at Cys34, and makes up 80% of the free thiol content in the plasma [51]. However, after depletion of albumin, it was observed that the most concentrated protein containing free-Cys is TTR. This protein has a sole cysteine residue at the position 10; it is synthetized by the liver, metabolized, and excreted by the kidneys. For many years, TTR has been studied not only for being related to amyloidosis process, but also because it can be used as a biomarker for nutritional status and severe illnesses [17].

It is described that TTR exists as a tetrameric form, but the dimer is the most stable form. The dimer is slightly rotated along the y axis in relation to each other – this arrangement promotes resistance against most alkali and acid solutions, denaturing agents such as urea, sodium-dodecyl-sulfate (SDS) and guanidine-HCl [19]. The dimer could be observed at ≈31 kDa in the SDS-PAGE, while the monomer was simultaneously observed at 14 kDa, when the sample was not heated at 90°C (Figure III.6).

Several studies present the relationship between TTR and amyloidosis process. In most cases, this disorder can be related to mutations or post-translation modifications (PTM) in the protein [22, 25]. Mass spectrometry is one of the most used tools to show primary sequence variation and PTM in TTR sequence, such as the modifications at the Cys10. Using the data obtained by SDS-PAGE and western blot analysis, in the commercial available plasma, 18% of the TTR exists as the free thiol form (Figure III.7 and 9): this content is in accordance with the data described in other studies, which present 80% to 95% of the total TTR in the circulation is posttranslationally modified at the Cys10; the values are similar to the level of modified albumin [25, 52, 53].

As mentioned, most of the sulfhydryl groups in proteins are modified, but some of these modifications are reversible. In this regard, a gentle reduction of the plasma proteins was carried out using 0.1 mM of DTT. An increment in TTR amounts recovered was expected as observed for the cysteine protease, papain (Chapter II); in fact, there was a gain of 1.5 times in the TTR concentration cause by the treatment with 0.1 mM DTT and this increment was also observed in the western blot analysis (Figure III.8 and 9). However, when the reducing agent was added to the sample, other proteins could also be seen in the SDS-PAGE (Figure III.8). According to Hillson (1980), even though low concentrations of the reducing agent are applied to the sample, some proteins could be reduced to be ready to react with the 2-pyridyl group [42]. Chang and coworkers (1997) likewise showed that the disulfide bond can be reduced using low concentrations of DTT (e.g. 0.5 mM) at room temperature [54]. In addition, the reduction of disulfide bonds between inter-chain polypeptides could more easily take place. The proteins identified in the SDS-PAGE (Figure III.8) were complement factor I and coagulation factor X, both of which possess disulfides in the inter-chain linkages. The inter-chain linkages are more susceptible to reduction than the intra-chain linkages; the disulfides present in the surface of a protein can be reduced without loss of the native structure [55, 56]. In this case, the DTT breaks the linkage between the chains, despite the intra-chain disulfide bonds being intact.

Meanwhile, some analyses were conducted using the Flow-through fraction of the Poros 20 NH LC-SPDP system, aiming to directly observe the PTMs at the Cys10; a shift of 180 Da in the molecular mass can be related to an addition of a sulphite adduct, 119 Da for cysteine, 176 Da for cysteinylglycine, and 305 Da for glutathione [22]. The mass spectrum of the Flow-through fraction gave two modified forms of TTR (Figure III.10). The most abundant one had a shift in mass by +119 Da that corresponds to a cysteine; a minor peak observed at 13,390 Da had a mass increment of 176 Da that corresponds to a cysteinyl-glycine. These results are in accordance with those

observed by Poulsen and coworkers (2012) who presented that the most prevalent modifications at Cys10 of TTR were S-Cys, S-Cys-Gly and S-glutathione, which were called the three S-thiolation [25, 47].

Next, the TTR obtained from the Flow-through fraction was exposed to 5 mM DTT, followed by molecular mass measurement, which clearly showed the mass shift from those of the above modified forms (+119 and +176 Da) to the unmodified form. The ratio between the free and the modified form of TTR may be used to estimate the propensity of the protein to be destabilized and form the amyloid fibrils, since the tetramer stability is related to the modification at Cys10 because the adducts can alter the native contact between the subunits [26, 57]. Hanyu (2009) and Kingsbury (2007) and coworkers presented the effect of the modification on the Cys10 by a cysteine; in FAP patients, Cys10 of the TTR was linked to a cysteine and this linkage destabilized the TTR structure. They also referred to when the level of Cys present in the plasma increases due to some disorders, the cysteine can react more with TTR, and the modified TTR triggers the formation of the amyloidogenic structure [25, 26].

In order to characterize the nature of the linkage formed between TTR and the resin, NEM, a reagent which reacts rapidly with a sulfhydryl group was added to the plasma sample. The sample treated with NEM did not contain any proteins in the Eluted fraction (Figure III.11), implying that the linkage of TTR to the Poros 20 NH LC-SPDP resin was formed by the interchange reaction of the sulfhydryl group (TTR) and the dithio-pyridine moiety (LC-SPDP).

In Chapter II it was demonstrated that for papain, the required incubation time for enzyme recovery was, at least, seven hours, whereas, a four hour incubation was optimal for recovering TTR. The shorter incubation time for TTR could be related to the location of the Cys10 in the surface of the molecule (the edge of β-strand in Figure III.16). Meanwhile, the Cys25 of papain is located in the inner part of the molecule (the cleft between the two subunits in Figure. II.1 and Figure III.16). The sulfhydryl of an exposed Cys has the pKa close to the physiological pH and small environmental variations of the Cys can alter the pKa of the sulfhydryl group; the Cys may act as nucleophilic agent, which, as a result, cause some modifications in the proximal portion of protein surface [58]. When Cys10 is not modified, it may interact with Gly57 in the β-strand D (see Figure III.1); the main chain of the Gly57 form a sulfur-hydrogen bond with the Cys10 [52].

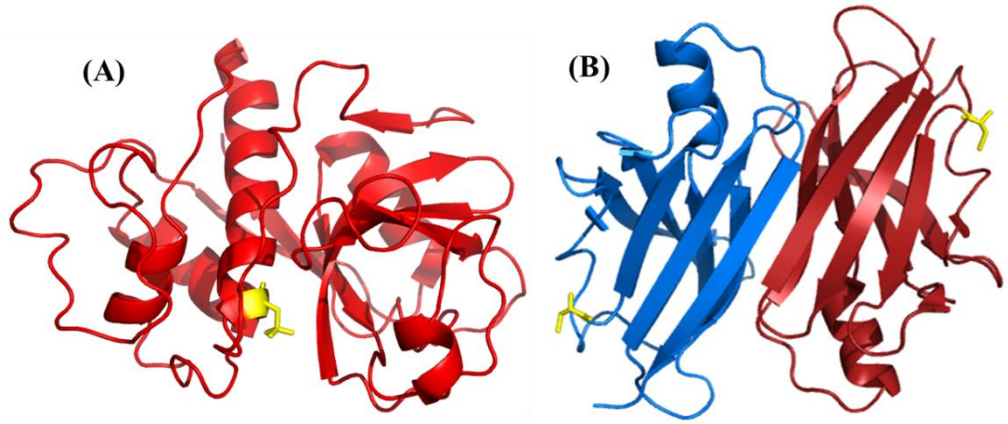


Figure III.16. Three dimensional structures from papain (PDB 9PAP) (A) and TTR dimer (PDB 3W3B) (B). The free-Cys in both proteins (Cys25 and C10, respectively) are highlighted in yellow. Figures obtained using the PyMOL software.

When the Cys10 is modified, the hydrogen bond between Cys10 and Gly57 cannot be formed and the disruption of this bonding might alter the protein structure, causing a shift in TTR conformation. Nonetheless, not all modifications will destabilize the protein structure. Each Cys10 in the tetramer can have different modifications and this creates a microheterogeneity in the protein structure. According to Lim and coworkers (2003), a tetramer with different PTMs per subunit could promote differences in protein stability [24, 52]. In the case of TTR Val30Met variants, when the Cys 10 is S-cysteinylated, the degree of fibril formation increases, compared to the unmodified Val30Met. The S-sulfonated form of TTR is cited as one of the most amyloidogenic forms because this adduct adds a negative charge to each monomer, and the net charge and changes in conformational state can induce the amyloid fibril formation [52, 59].

After presenting the efficacy of the method to isolate the proteins containing free thiol groups from the plasma, the experiment to simulate the effect of oxidative stress on free thiol containing proteins were conducted. Proteins containing free thiol groups are responsible for buffering the redox state in the plasma. This group is capable of deoxidizing the compounds in the environment and acts as a non-enzymatic defense against ROS. Proteins could absorb around 70% of the reactive species present in the environment, which cause covalent modifications in the amino acid side chain [51, 60]. Using hydrogen peroxide and the Poros 20 NH LC-SPDP system, it was possible to observe the response of the oxidative stress on a number of proteins recovered in the Eluted fraction (Figure III. 13 and 14).

It was described that tryptophan, tyrosine, histidine and cysteine are the first amino acids that are modified in the presence of $^{\bullet}OH$. However, Cys has the highest constant rate with $^{\bullet}OH$; it is three times higher than Trp, Tyr and His [36, 61]. The proteins containing free thiol groups are susceptible to two-electron oxidation not only by H_2O_2, but also by lipid peroxides, aldehydes, quinones, and disulfides [62]. Among the reactive oxygen species, the effect of the H_2O_2 upon the cysteine residues has been studied by many researches. The oxidation of the cysteine residues by ROS/RNS and free-radicals is one of the main mechanisms for the transduction of biological signal, transcriptional regulation, cell proliferation, angiogenesis and apoptosis [62, 63]. The –SH group works as a redox sensor, controlling protein function. Not all thiol groups in proteins promptly react with oxidants, but some of them readily ionize to thiolate ($Cys\text{-}S^-$), and this reactive thiolate could provide the specificity of the redox signaling [63, 64].

The susceptibility to peroxidation may be related to amino acid sequence, but it is difficult to predict if a cysteine residue is a redox-active residue only by the primary structure; a protein structure influences the susceptibility of each amino acid to react with free radicals in the environment, favoring the access of the oxidants to some residues [34, 63]. In this sense, the position of Cys10 in the surface of TTR may make this residue more susceptible to H_2O_2 attack. The "loss" of the free thiol group was observed when the plasma sample was previously exposed to 2 -mM of H_2O_2; the concentration of TTR in the Eluted fraction was significantly reduced. The treatment of the plasma with 5 and 10 mM of H_2O_2 abolished the free-form TTR (Figure III. 13). Furthermore, Table III.1 presents the protein concentration of the Flow-through fraction; in the sample treated with 5 and 10 mM H_2O_2, the total protein concentration increased, compared to the control (sample without H_2O_2 treatment). These values may indirectly represent the proteins with

free thiol groups, but oxidized by H_2O_2, being accumulated and recovered in the Flow-through fraction.

The western blot analysis of the *in vitro* oxidative stress also represents the scavenger propriety of TTR. On the samples exposed to H_2O_2, an accumulation of TTR in the Flow-through fraction (Figure III.14) was observed. Image analysis of the blot indicates that the concentration of TTR in the Flow-through fraction increased in 3 and 3.8 times after the treatment with 5 and 10 mM H_2O_2, respectively. The attack of ROS on proteins could make or break disulfide bonds, triggering protein unfolding. Protein aggregates can be formed due to the new disulfide bonds and covalent and non-covalent bonds between the unfolded proteins [36, 65]. In Figure III.14, the high-molecular-mass protein band at the dimer position obviously increased in proportion to the H_2O_2 concentration.

In diseases related to oxidative stress, such as chronicle kidney disease, the redox status of the proteins containing free thiol groups are changed and this rate can be used as a marker for the progression of the disease. In 2001, Mimic-Oka used the protein-SH levels in the plasma as a marker for oxidative stress in patients with chronic renal disorder, because the level of free thiol groups in the plasma was significantly reduced in end stage patients [37]. Moreover, according to the degree of renal dysfunction, irreversible modification in the thiol group in albumin can directly induce the carbonyl formation in plasma proteins and these modifications aggravate the uremia [51]. Renal disorders, inflammatory diseases and cancer can cause oxidative modifications on albumin and other proteins containing free thiol groups [66]. Estimating the ratio between modified and unmodified thiol groups may allow the development of a standard in changes in the redox state related to diseases [67]. In this case, the comparison between control samples (heathy volunteers) and those from diseases related to oxidative stress, which may be applied to redox proteomics, provides important information on molecular response in vivo to the oxidative stress [51, 60].

The present method was compared to a commercially available method in regards to isolation of proteins containing free thiol groups. The concentration of TTR in the plasma is relatively low (150-200 $\mu g.ml^{-1}$); and to study this protein, diverse methods have been used: affinity chromatography – the mediation of RBP or thyroxine coupled with a sepharose resin – and immunological methods, such as immunoprecipitation and immunoaffinity chromatography [19, 68, 69]. However, some of these methods, which employ multiple steps, are time consuming and yield low recovery. The present method has also been compared with other similar resin-based methods, i.e. Thiopropyl Sepharose 6B, one of the most utilized chromatography systems, and a polymer-

based resin, Toyo Pearl AF-Amino 650 (see Materials and Methods) (Figure III.15). Although the TTR could be recovered in the Eluted fraction with these resins, the Poros 20 NH LC-SPDP resin resulted in TTR with higher purity and yield. Meanwhile, other two methods gave some high-molecular-weight proteins other than TTR, which could be attributed to non-specific bindings to those resins.

References

[1] Issaq, H. J., Xiao, Z., Veenstra, T. D., Serum and plasma proteomics. *Chem Rev* 2007, *107*, 3601-3620.

[2] Anderson, N. L., Anderson, N. G., The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 2002, *1*, 845-867.

[3] Deutsch, E. W., Eng, J. K., Zhang, H., King, N. L., *et al.*, Human Plasma PeptideAtlas. *Proteomics* 2005, *5*, 3497-3500.

[4] Hortin, G. L., The MALDI-TOF mass spectrometric view of the plasma proteome and peptidome. *Clin Chem* 2006, *52*, 1223-1237.

[5] Burnouf, T., Modern plasma fractionation. *Transfus Med Rev* 2007, *21*, 101-117.

[6] Merrel, K., *Department of Chemistry and Biochemistry*, Brigham Young, Utah, United States of America 2009, p. 155.

[7] ANKER, H. S., The biosynthesis of plasma proteins. *Ann N Y Acad Sci* 1961, *94*, 14-22.

[8] Jarnum, S., Turnover of plasma proteins. *J Clin Pathol Suppl (Assoc Clin Pathol)* 1975, *6*, 13-21.

[9] Anderson, N. L., Polanski, M., Pieper, R., Gatlin, T., *et al.*, The human plasma proteome: a nonredundant list developed by combination of four separate sources. *Mol Cell Proteomics* 2004, *3*, 311-326.

[10] Nanjappa, V., Thomas, J. K., Marimuthu, A., Muthusamy, B., *et al.*, Plasma Proteome Database as a resource for proteomics research: 2014 update. *Nucleic Acids Res* 2014, *42*, D959-965.

[11] Kanda, Y., Goodman, D. S., Canfield, R. E., Morgan, F. J., The amino acid sequence of human plasma prealbumin. *J Biol Chem* 1974, *249*, 6796-6805.

[12] Hamilton, J. A., Benson, M. D., Transthyretin: a review from a structural perspective. *Cell Mol Life Sci* 2001, *58*, 1491-1521.

[13] Cody, V., Wojtczak, A., in: Richardson, S., Cody, V. (Eds.), *Recent Advances in Transthyretin Evolution, Structure and Biological Functions*, Springer Berlin Heidelberg 2009, pp. 1-21.

[14] Gales, L., Saraiva, M. J., Damas, A. M., Structural basis for the protective role of sulfite against transthyretin amyloid formation. *Biochim Biophys Acta* 2007, *1774*, 59-64.

[15] Hörnberg, A., *Transthyretin from a Structural Perspective*, Univ. 2004.

[16] Altland, K., Winter, P., Potential treatment of transthyretin-type amyloidoses by sulfite. *Neurogenetics* 1999, *2*, 183-188.

[17] Gaudiani, J. L., Sabel, A. L., Mehler, P. S., Low prealbumin is a significant predictor of medical complications in severe anorexia nervosa. *Int J Eat Disord* 2014, *47*, 148-156.

[18] Saraiva, M. J., Transthyretin mutations in health and disease. *Hum Mutat* 1995, *5*, 191-196.

[19] Ingenbleek, Y., Young, V., Transthyretin (prealbumin) in health and disease: nutritional implications. *Annu Rev Nutr* 1994, *14*, 495-533.

[20] Saraiva, M. J., Birken, S., Costa, P. P., Goodman, D. S., Amyloid fibril protein in familial amyloidotic polyneuropathy, Portuguese type. Definition of molecular abnormality in transthyretin (prealbumin). *J Clin Invest* 1984, *74*, 104-119.

[21] Johnson, S. M., Connelly, S., Fearns, C., Powers, E. T., Kelly, J. W., The transthyretin amyloidoses: from delineating the molecular mechanism of aggregation linked to pathology to a regulatory-agency-approved drug. *J Mol Biol* 2012, *421*, 185-203.

[22] Reddi, H. V., Thomas, B. C., Willkomm, K. S., Ferber, M. J., *et al.*, Discordant results between biochemical and molecular transthyretin assays: lessons learned from a unique testing algorithm at the Mayo Clinic. *J Genet* 2013, *92*, 599-604.

[23] Ando, Y., Nyhlin, N., Suhr, O., Holmgren, G., *et al.*, Oxidative stress is found in amyloid deposits in systemic amyloidosis. *Biochem Biophys Res Commun* 1997, *232*, 497-502.

[24] Lim, A., Sengupta, S., McComb, M. E., Théberge, R., *et al.*, In vitro and in vivo interactions of homocysteine with human plasma transthyretin. *J Biol Chem* 2003, *278*, 49707-49713.

[25] Hanyu, N., Shimizu, T., Yamauchi, K., Okumura, N., Hidaka, H., Characterization of cysteine and homocysteine bound to human serum transthyretin. *Clin Chim Acta* 2009, *403*, 70-75.

[26] Kingsbury, J. S., Klimtchuk, E. S., Théberge, R., Costello, C. E., Connors, L. H., Expression, purification, and in vitro cysteine-10 modification of native sequence recombinant human transthyretin. *Protein Expr Purif* 2007, *53*, 370-377.

[27] Giustarini, D., Dalle-Donne, I., Lorenzini, S., Milzani, A., Rossi, R., Age-related influence on thiol, disulfide, and protein-mixed disulfide levels in human plasma. *J Gerontol A Biol Sci Med Sci* 2006, *61*, 1030-1038.

[28] Suhr, O. B., Ando, Y., Ohlsson, P. I., Olofsson, A., *et al.*, Investigation into thiol conjugation of transthyretin in hereditary transthyretin amyloidosis. *Eur J Clin Invest* 1998, *28*, 687-692.

[29] Sass, J. O., Nakanishi, T., Sato, T., Sperl, W., Shimizu, A., S-homocysteinylation of transthyretin is detected in plasma and serum of humans with different types of hyperhomocysteinemia. *Biochem Biophys Res Commun* 2003, *310*, 242-246.

[30] Aruoma, O., Free radicals, oxidative stress, and antioxidants in human health and disease. *Journal of the American Oil Chemists' Society* 1998, *75*, 199-212.

[31] Hybertson, B. M., Gao, B., Bose, S. K., McCord, J. M., Oxidative stress in health and disease: the therapeutic potential of Nrf2 activation. *Mol Aspects Med* 2011, *32*, 234-246.

[32] Uttara, B., Singh, A. V., Zamboni, P., Mahajan, R. T., Oxidative stress and neurodegenerative diseases: a review of upstream and downstream antioxidant therapeutic options. *Curr Neuropharmacol* 2009, *7*, 65-74.

[33] Ghezzi, P., Oxidoreduction of protein thiols in redox regulation. *Biochem Soc Trans* 2005, *33*, 1378-1381.

[34] Gebicki, S., Gebicki, J. M., Formation of peroxides in amino acids and proteins exposed to oxygen free radicals. *Biochem J* 1993, *289 (Pt 3)*, 743-749.

[35] Davies, K. J., Protein damage and degradation by oxygen radicals. I. general aspects. *J Biol Chem* 1987, *262*, 9895-9901.

[36] Davies, K. J., Delsignore, M. E., Lin, S. W., Protein damage and degradation by oxygen radicals. II. Modification of amino acids. *J Biol Chem* 1987, *262*, 9902-9907.

[37] Mimić-Oka, J., Simić, T., Plješa, M., Stupar, N., Turković, S., Oxidative modifications of plasma proteins in different stages of chronic renal failure. *Facta universitatis-series: Medicine and Biology* 2001, *8*, 1-5.

[38] Cighetti, G., Paroni, R., Marzorati, S., Borotto, E., *et al.*, Evaluation of oxidative stress in serum of critically ill patients by a commercial assay and gas chromatography-mass spectrometry. *Clin Chem* 2005, *51*, 1515-1517.

[39] Grimsrud, P. A., Xie, H., Griffin, T. J., Bernlohr, D. A., Oxidative stress and covalent modification of protein with bioactive aldehydes. *J Biol Chem* 2008, *283*, 21837-21841.

[40] Brandes, N., Schmitt, S., Jakob, U., Thiol-based redox switches in eukaryotic proteins. *Antioxid Redox Signal* 2009, *11*, 997-1014.

[41] Bruschi, M., Candiano, G., Santucci, L., Ghiggeri, G. M., Oxidized albumin. The long way of a protein of uncertain function. *Biochim Biophys Acta* 2013, *1830*, 5473-5479.

[42] Hillson, D. A., Freedman, R. B., Resolution of protein disulphide-isomerase and glutathione-insulin transhydrogenase activities by covalent chromatography. *Biochem J* 1980, *191*, 373-388.

[43] Laemmli, U. K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970, *227*, 680-685.

[44] Blum, H., Beier, H., Gross, H. J., Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *ELECTROPHORESIS* 1987, *8*, 93-99.

[45] Harlow, E., Lane, D., *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory 1988.

[46] Fligiel, S. E., Lee, E. C., McCoy, J. P., Johnson, K. J., Varani, J., Protein degradation following treatment with hydrogen peroxide. *Am J Pathol* 1984, *115*, 418-425.

[47] Poulsen, K., Bahl, J. M., Tanassi, J. T., Simonsen, A. H., Heegaard, N. H., Characterization and stability of transthyretin isoforms in cerebrospinal fluid examined by immunoprecipitation and high-resolution mass spectrometry of intact protein. *Methods* 2012, *56*, 284-292.

[48] Fahey, R. C., Hunt, J. S., Windham, G. C., On the cysteine and cystine content of proteins. Differences between intracellular and extracellular proteins. *J Mol Evol* 1977, *10*, 155-160.

[49] Brosnan, J. T., Brosnan, M. E., The sulfur-containing amino acids: an overview. *J Nutr* 2006, *136*, 1636S-1640S.

[50] Seiwert, B., Hayen, H., Karst, U., Differential labeling of free and disulfide-bound thiol functions in proteins. *J Am Soc Mass Spectrom* 2008, *19*, 1-7.

[51] Matsuyama, Y., Terawaki, H., Terada, T., Era, S., Albumin thiol oxidation and serum protein carbonyl formation are progressively enhanced with advancing stages of chronic kidney disease. *Clin Exp Nephrol* 2009, *13*, 308-315.

[52] Lim, A., Prokaeva, T., McComb, M. E., Connors, L. H., *et al.*, Identification of S-sulfonation and S-thiolation of a novel transthyretin Phe33Cys variant from a patient diagnosed with familial transthyretin amyloidosis. *Protein Sci* 2003, *12*, 1775-1785.

[53] Heegaard, N. H., Hansen, M. Z., Sen, J. W., Christiansen, M., Westermark, P., Immunoaffinity chromatographic and immunoprecipitation methods combined with mass spectrometry for characterization of circulating transthyretin. *J Sep Sci* 2006, *29*, 371-377.

[54] Chang, J. Y., A two-stage mechanism for the reductive unfolding of disulfide-containing proteins. *J Biol Chem* 1997, *272*, 69-75.

[55] Sears, D. W., Mohrer, J., Beychok, S., Relative susceptibilities of the interchain disulfides of an immunoglobulin G molecule to reduction by dithiothreitol. *Biochemistry* 1977, *16*, 2031-2035.

[56] Kuwajima, K., Ikeguchi, M., Sugawara, T., Hiraoka, Y., Sugai, S., Kinetics of disulfide bond reduction in alpha-lactalbumin by dithiothreitol and molecular basis of superreactivity of the Cys6-Cys120 disulfide bond. *Biochemistry* 1990, *29*, 8240-8249.

[57] Liu, L., Hou, J., Du, J., Chumanov, R. S., *et al.*, Differential modification of Cys10 alters transthyretin's effect on beta-amyloid aggregation and toxicity. *Protein Eng Des Sel* 2009, *22*, 479-488.

[58] Marino, S. M., Gladyshev, V. N., Analysis and functional prediction of reactive cysteine residues. *J Biol Chem* 2012, *287*, 4419-4425.

[59] Nakanishi, T., Sato, T., Sakoda, S., Yoshioka, M., Shimizu, A., Modification of cysteine residue in transthyretin and a synthetic peptide: analyses by electrospray ionization mass spectrometry. *Biochim Biophys Acta* 2004, *1698*, 45-53.

[60] Hu, W., Tedesco, S., Faedda, R., Petrone, G., *et al.*, Covalent selection of the thiol proteome on activated thiol sepharose: a robust tool for redox proteomics. *Talanta* 2010, *80*, 1569-1575.

[61] Held, J. M., Gibson, B. W., Regulatory control or oxidative damage? Proteomic approaches to interrogate the role of cysteine oxidation status in biological processes. *Mol Cell Proteomics* 2012, *11*, R111.013037.

[62] Jones, D. P., Radical-free biology of oxidative stress. *Am J Physiol Cell Physiol* 2008, *295*, C849-868.

[63] Eaton, P., Protein thiol oxidation in health and disease: techniques for measuring disulfides and related modifications in complex protein mixtures. *Free Radic Biol Med* 2006, *40*, 1889-1899.

[64] Kim, J. R., Yoon, H. W., Kwon, K. S., Lee, S. R., Rhee, S. G., Identification of proteins containing cysteine residues that are sensitive to oxidation by hydrogen peroxide at neutral pH. *Anal Biochem* 2000, *283*, 214-221.

[65] Grune, T., Reinheckel, T., Davies, K. J., Degradation of oxidized proteins in mammalian cells. *FASEB J* 1997, *11*, 526-534.

[66] Ellidag, H. Y., Bulbuller, N., Eren, E., Abusoglu, S., *et al.*, Ischemia-modified albumin: could it be a new oxidative stress biomarker for colorectal carcinoma? *Gut Liver* 2013, *7*, 675-680.

[67] Hill, B. G., Reily, C., Oh, J. Y., Johnson, M. S., Landar, A., Methods for the determination and quantification of the reactive thiol proteome. *Free Radic Biol Med* 2009, *47*, 675-683.

[68] Raghu, P., Ravinder, P., Sivakumar, B., A new method for purification of human plasma retinol-binding protein and transthyretin. *Biotechnol Appl Biochem* 2003, *38*, 19-24.

[69] Mahn, A., Lienqueo, M. E., Quilodrán, C., Olivera-Nappa, A., Purification of transthyretin as nutritional biomarker of selenium status. *J Sep Sci* 2012, *35*, 3184-3189.

General Conclusions

The hetero-bifunctional cross-linkers [N-succinimidyl 3-(2-pyridyldithio) propionate] (SPDP) and [succinimidyl 6-[3'-(2-pyridylthio)- propionamido] hexanoate] (LC-SPDP) bound to "Poros" (a polymer-based resin) was found to be effective for isolating free Cys-containing peptides/proteins with a few picomole from a complex mixture. The method has shown high recovery, selectivity and reproducibility in all experiments performed, and these features are important for making the method a useful procedure in proteomic analysis.

Furthermore, the application of the method to plasma proteins revealed that transthyretin (TTR) is the major free Cys-containing protein in this biological fluid. Protein profiling of complex biological samples (derived from cells, tissues, organs, etc.) can be modified in various physiological conditions, such as diseases. In this sense, the results obtained in this study can be applied for estimating a ratio of the free-Cys form of TTR to modified or total TTR in plasma. TTR has been reported to form a disulfide bond with some other reactive thiol compounds, such as sulfonate, cysteine, homocysteine, etc. However, 10-20% of the Cys10 remain in the reduced form. This protein has been deeply involved with some disorders and can be useful for diagnoses of cancer and amyloidosis process. In this sense, data regarding the ratio of the free Cys form to the oxidized ones may correlate with the physiological status of the organism or its propensity to develop some diseases as a disease-specific signature; it could be a beneficial marker, at the proteomic level, for early diagnosis or prediction of disease progression, which can be closely related to protein expression and post-translational modifications.

Proteins in the plasma have the important role of maintaining the redox-state in equilibrium. Protein containing free thiol groups, such as albumin, can react 100 times faster with ROS/RNS than RNA and DNA. This well-known mechanism strongly suggests that free cysteines in proteins are capable of sinking the ROS produced during respiration and physiological disturbances [61]. Taking into account the sensitivity and reactivity of the free Cys residues, the ratio of free-Cys form(s) to the oxidized one(s) could reflect a physiological state rationalized to show susceptibility to oxidative stress. In this regard, the Poros 20 NH LC-SPDP system will be a useful technique to estimate the ratios of the unmodified TTR (with a free Cys10) to the Cys-modified ones in the plasma between healthy and disease states.