NOTE TO USERS

This reproduction is the best copy available.

UMI®

University of Alberta

Microwave-assisted Acid/Base Hydrolysis and Mass Spectrometry for

Prion Structure Characterization

by

Bela Reiz   ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment

of the requirements for the degree of Master of Science

Department of Chemistry

Edmonton, Alberta

Spring 2009

Canada

Abstract

Proteinaceous infectious particles are endogenous proteins that by misfolding can cause a variety of fatal neurodegenerative diseases, which can affect both humans and animals. Characterization of the structural conformers, the dynamics and kinetics involved in the misfolding of this protein could be helpful in the design and development of diagnostic tools and therapeutic products.

In this work, we examined the use of microwave-assisted acid/base hydrolysis coupled with mass spectrometry for the characterization of the rSHPrP90-231 prion construct. We have identified an acid hydrolysis resistant core suggesting that the conformational conversion and aggregation of the protein involves part of α-helix 2, most part of α-helix 3 and the loop between these two helices when referred to the native PrP monomer. Circular dichroism, dynamic light scattering and electrophoresis data has also confirmed conversion and aggregation or oligomerization of the prion protein.

Acknowledgements

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACN	Acetonitrile
BSE	Bovine spongiform encephalopathy
CHCA	α-cyano-4-hydroxycinnamic acid
CID	Collision-induced dissociation
CJD	Creutzfeld Jacob Disease
CWD	Chronic Wasting Disease
DHB	2,5-dihydroxybenzoic acid
DLS	Dynamic light scattering
DTT	dithiothreitol
E. coli	*Escherichia coli*
ESI	Electrospray ionization
FFI	Fatal Familial Insomnia
FT-ICR	Fourier-transform ion cyclotron resonance
FT-ICR-MS	Fourier-transform ion cyclotron resonance mass spectrometry
FT-MS	Fourier-transform mass spectrometry
FWHM	Full width at half maximum
HPLC	High performance liquid chromatography
IR	Infrared
m/z	mass to charge
MAAH	Microwave-assisted acid hydrolysis

MABH	Microwave-assisted base hydrolysis
MALDI	Matrix-assisted laser desorption/ionization
MCP	Multichannel plate detector
MBM	Meat and bone meal
MS	Mass spectrometry
MW	Molecular weight
Nd:YAG	Neodymium dopped yttria-aluminium garnet
nm	nanometre
PrP	prion related protein
PrP^C	cellular PrP
PrP^{Sc}	scrapie PrP
PTM	Posttranslational modification
RP	Reversed-phase
rSHPrP	Recombinant Syrian hamster PrP
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate- polyacrylamide gel electrophoresis
S/N	Signal to noise ratio
SUS	Sodium undecyl sulphate
TFA	Trifluoroacetic acid
TIS	Timed ion selector
TOF	Time-of-flight

TSE	Transmissible spongiform encephalopathy
UV	Ultraviolet
m	milli- (10^{-3})
μ	micro- (10^{-6})
n	nano- (10^{-9})

Chapter 1: Introduction

Transmissible spongiform encephalopathy's (TSE's) are caused by the conformational changes of the normal body protein, called prion protein, resulting in a proteinase K resistant, amyloid fibril. These infectious diseases can affect both humans and animals. For examples, Kuru has decimated the population of Papua New Guinea more than fifty years ago, while Creutzfeldt-Jakob disease (CJD) caused the death of more than 250 individuals. Suggestions that transmission of bovine spongiform encephalopathy (BSE) to humans might be the cause of variant CJD caused a widespread health scare, causing devastating losses to the beef industry around the world. Up until now there are no treatments for these diseases. Therefore understanding the processes involved in the conformational changes of this protein from a cellular form into a proteinaceous infectious agent are of a great importance.

The motivation behind the work presented in this thesis was to gain a deeper understanding of the structure and dynamics of the prion protein which could help to the understanding of the etiology of TSE's and serve as basic information for the development of diagnostic tools and therapeutic products. In order to achieve this goal we have used a series of analytical techniques such as mass spectrometry, circular dichroism, dynamic light scattering and electrophoresis to better characterize the different conformational isoforms of this protein. Some of the key techniques used in my thesis work along with basic information on prion molecules and prion-related diseases are discussed in this chapter.

1.1. Principles and Instrumentation of Mass Spectrometry

The major components of an MS instrument are: an ion source, a mass analyzer, and the ion detector. The throughput, sensitivity, resolution and accuracy of a mass spectrometer are dependent on the efficiency of every single one of these components.

1.1.1. Ionization methods

In order to be able to determine the mass of a molecule using a mass spectrometer the molecules to be measured have to be charged. This can be done using several different ionization methods, such as, chemical ionization (CI), electron impact (EI), multiphoton ionization (MPI), desorption ionization, etc. Because biopolymers are polar, large, non-volatile and labile, soft ionization methods are needed to efficiently transfer the molecules into gas phase. There are two major soft ionization methods used for the detection of large, non-volatile and labile biomolecules. These are matrix-assisted laser desorption/ionization and electrospray ionization.

1.1.1.1. Matrix-Assisted Laser Desorption/Ionization Process

Matrix-assisted laser desorption/ionization or MALDI is one of the soft ionization methods used for the detection of large, non-volatile and labile molecules by mass spectrometry. The MALDI principle and acronym was first introduced by Franz Hillenkamp, Michael Karas and their colleagues in 1985[1] when they have shown that the amino acid alanine could be ionized easily if mixed with amino acid tryptophan and irradiated with a laser operated at 266 nm. The breakthrough came in 1987 when Tanaka and colleagues demonstrated that ionization of large biomolecules was possible with proper combination of laser wavelength and matrix[2]. They were able to ionize and detect carboxypeptidase-A, a 34.4 kDa protein. Tanaka was rewarded with the Nobel Prize in Chemistry in 2002. The method is based on the generation of vapour-phase ions of large non-volatile, fragile biomolecules directly from solid phase[3, 4]. For this purpose, a saturated or almost saturated solution of a small chemical matrix and analyte dissolved in a solvent or mixture of solvents are used. The two components are mixed either before deposition onto or directly on the MALDI plate. The typical deposition volumes of the analyte-matrix mixtures are 0.5 – 1.0 µL. Matrix to analyte ratio of 500:1 are usually used in order to isolate analyte molecules to prevent aggregation, minimize analyte-target interactions and avoid sample decomposition[5, 6]. After deposition the spots are usually slowly, air-dried during which crystallization of the matrix and

incorporation of the analyte occurs. The matrix incorporated analytes are then introduced into a mass spectrometer's ion source region. This region is usually kept under vacuum and desorption/ablation of the matrix and analyte molecules is achieved by the application of a pulsed laser beam. The mechanisms involved in the formation of charged matrix and analyte molecules in MALDI are poorly understood. To date, two different models have been proposed. The older model assumes that the analytes are incorporated in the matrix as neutral species. The first step in this model would be the photoionization of the matrix molecules resulting in a plume in which in the second step charge is transferred from the matrix to the analyte molecules[7]. The second model, also called the "lucky survivor" model[6] assumes that proteins are incorporated in the matrix crystals as charged species and after desorption most of them becomes re-neutralized within the desorbed clusters of matrix and analyte. Regardless of which of these two different proposed models is actually correct, after application of the laser pulse, desorption/ablation and ionization take place. In order to detect the analytes a short positive/negative voltage is applied to accelerate the positively/negatively charged species towards the mass analyzer. A simplified schematic of the MALDI process is shown in Figure 1.1.1
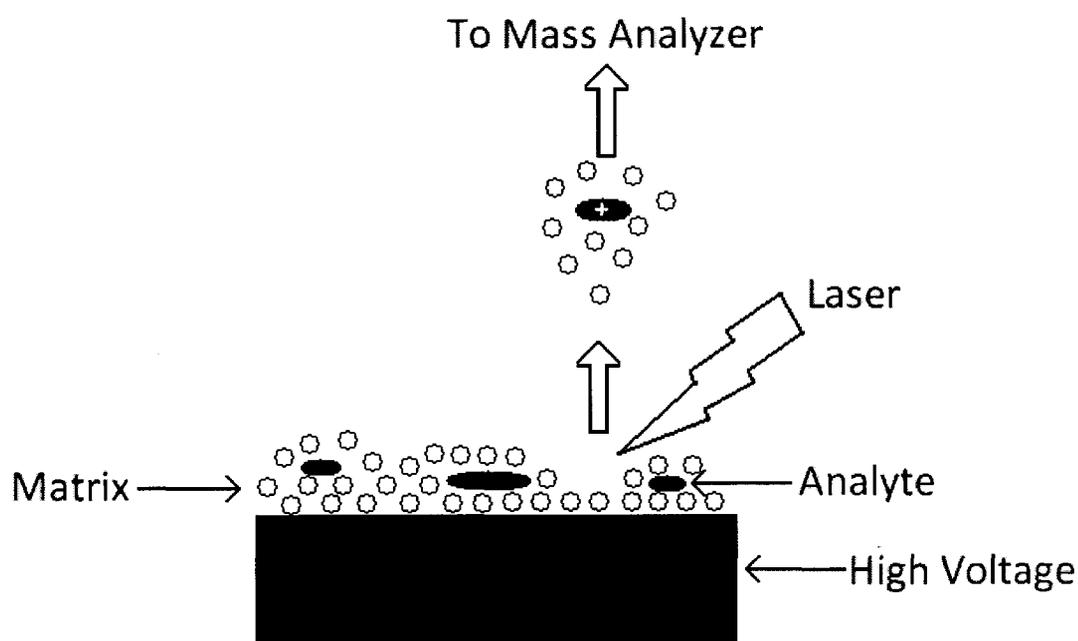
Figure 1.1.1 Simplified process of matrix-assisted laser desorption/ionization (MALDI)

Matrix molecules in MALDI are typically small, UV absorbing aromatic organic acids. There is no single MALDI matrix suitable for all types of analytes and applications. The most widely used matrices for peptides and proteins are CHCA (α-cyano-4-hydroxycinnamic acid) and DHB (2,5-Dihydroxy benzoic acid). CHCA is used in most proteomic applications, especially peptides for both MS and MS/MS analysis.

As might be expected, there is no single matrix or sample preparation method. Several different methods have been developed including: the dried-droplet[4], vacuum drying[8], crushed crystal[9], fast evaporation[10] and two-layer[11]. The most widely used method in our laboratory is the two-layer method using CHCA as a matrix. In this method a 1st layer consisting of matrix is deposited on a steel target plate. It involves the use of fast solvent evaporation to form a layer of small crystals which will serve as a seed for crystallization resulting in uniform and homogeneous crystals. The next step involves the deposition of a mixture of analyte and saturated matrix solution (2nd layer) on top of the first layer. After crystallization the spots can be washed with water to remove ionic contaminants or salts. The advantages of the two-layer sample preparation method over other methods are: high-resolution can be obtained easier, better signal reproducibility from shot to shot and better external calibration accuracy.

For UV-MALDI the most commonly used lasers are N_2 or Nd:YAG. Nitrogen lasers emit at a wavelength of 337 nm, close to the absorption maximum of commonly used matrices. The disadvantage of these lasers are their limited pulse repetition frequency (<100Hz) and the total number of emissions ($\leq 10^8$). Therefore, for high throughput analysis frequency-tripled Nd:YAG lasers are used. They emit at a wavelength of 355 nm. For IR-MALDI, CO_2 or Er:YAG lasers are used.

1.1.1.2. Electrospray Ionization

Electrospray ionization (ESI) is another one of the soft ionization methods used for the detection of large, non-volatile and labile molecules by mass spectrometry. The development of ESI-MS for the analysis of large biomolecules[12] was rewarded with the Nobel Prize in Chemistry to J.B. Fenn in 2002. ESI ionizes analytes that are dissolved in a

large amount of solvent that is pushed through a very small charged capillary. Four major processes have been described to be involved in the ESI ionization process: 1) production of charged droplets at the ESI capillary tip. The accumulation of charges at the surface of a liquid from a syringe pump, LC or some other source and the charges continuously build up at the liquid surface and cause the formation of a Taylor cone. If the electric field is high enough, the cone becomes unstable and a liquid filament is formed. Downstream the liquid filament becomes unstable and separate droplets are formed. 2) Shrinkage of charged ESI droplets. Shrinkage is caused by the evaporation of the solvent from the droplets. Evaporation can be achieved by using a N_2 gas flow at moderate temperatures to dry the droplet. 3) Repeated droplet disintegration. When the repulsion forces become bigger than the surface tension uneven fission of the droplet occurs. 4) Generation of gas phase ions. Consequent solvent evaporation and fission lead to extremely small droplets that contain only one ion and consequent evaporation of solvent will lead to conversion to a gas phase ion. The schematics of the ESI ionization process is shown in Figure 1.1.2.
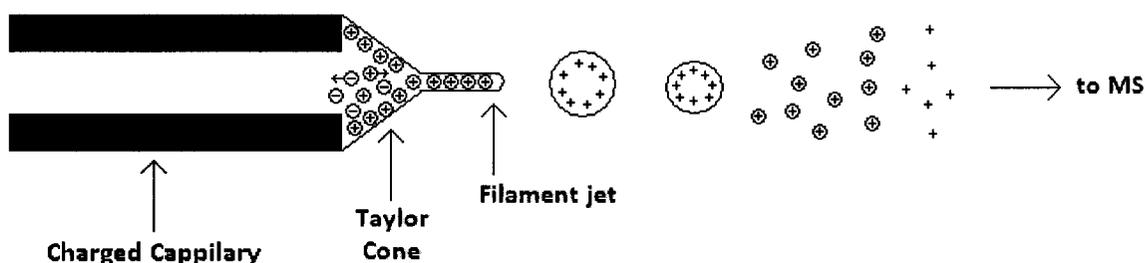
Figure 1.1.2 Simplified process of electrospray ionization (ESI)

1.1.2. Mass Analyzer

Mass analyzers are used to separate ion according to their mass-to-charge ratio (m/z). The most commonly used mass analyzers in proteomic research are the ion trap, quadrupole, time-of-flight, Fourier transform ion cyclotron resonance and Orbitrap. These are all very different in their design and are typically characterized by a number of important parameters, such as mass accuracy, mass resolution, sensitivity/limit of detection, signal-to-noise ratio (S/N) and scan speed. Sometimes, these mass analyzers

are used together in tandem mass spectrometers in order to take advantage of their different characteristics.

1.1.2.1. Time-of-Flight Mass Analyzer

Time-of-flight or TOF mass analyzers are one of the most widely used analyzers and by far the most widely used mass analyzer for MALDI applications. In TOF mass analyzers ions are accelerated with the same potential at a fixed point in time and are allowed to drift through a long field-free tube towards the detector. The ions travel with different velocities according to their mass-to-charge ratios, with lighter ions reaching the detector first. When a packet of ions with the mass m is accelerated by the application of the voltage V, the ions travel with the velocity v, expressed by:

$$Ve = \frac{1}{2}mv^2 \qquad \text{Eq. 1}$$

where e is the charge of the electron. This equation can be rearranged as:

$$v = \left(2Ve/m\right)^{1/2} \qquad \text{Eq. 2}$$

showing that velocity of the ions is inversely proportional with their mass.

Velocity can be defined as distance (length of flight tube in the TOF analyzer) over time:

$$v = D/t \qquad \text{Eq. 3}$$

Substituting Eq. 3 in Eq. 1 gives:

$$t = \left(mD^2/2Ve\right)^{1/2} \qquad \text{Eq. 4}$$

If the accelerating voltage and flight tube length are taken as constant the flight time will be given by:

$$t = km^{1/2} \qquad \text{Eq. 5}$$

which shows that the flight time will be directly proportional to the square root of the apparent mass. In case of a multiply charged ion the flight time will be proportional to the square root of the mass-to-charge ratio of the ion, as given by:

$$t = k\left(\frac{m}{z}\right)^{1/2}$$ Eq. 6

For a given experiment this equation can be inverted in order to get the calibration equation as m/z = kt^2. In order to account for an electronic delay and the effects from the ion generation, a second constant is usually added to this equation. One of the most commonly used calibration equations for TOF analyzers is:

$$^m/_z = kt^2 + c$$ Eq. 7

The mass resolution (R) in a TOF mass analyzer is determined by the peak width at a given mass in the spectrum:

$$R = {^m}/_{\Delta m} = \frac{t}{\Delta t}$$ Eq. 8

Ideally, in a TOF the ions would arrive at the detector in discrete packets at different times and all ions in one packet would arrive at the exact same time. However, in real applications the ions in a packet arrive at the detector spread out in a finite time window. There are two major sources of error in the arrival times in MALDI-TOF instruments. The first factor comes from the fact that the ions of the same mass are extracted from the plume having different initial velocities. The second factor degrading the resolution is given by the initial spatial distributions caused by irregularities in the matrix surface, although these are generally very small because of the very thin layer of matrix/analyte spotted on a flat metal surface.

To increase resolution, modern instruments are design to correct for the initial energy distribution of the ions. These instruments use time-lag-focusing also called delayed extraction[13-16] and an ion mirror called reflectron[17, 18].

In time-lag-focusing ions are extracted from the plume after a short period of delay after they were generated. The schematic of the time-lag-focusing is shown in Figure 1.1.3. Initially, the electric field between the repeller and the first electrode is kept very low or at zero and after some time this field is switched on. In this short period of time, usually tens of nanoseconds, ions with higher initial kinetic energies will travel farther from the repeller. When the field is switched on, because less energetic ions are closer to the repeller than more energetic ones, they will experience a higher potential difference and as a result they will catch up with the slower ions at some point in time. By properly selecting the delay time and pulse potential, with the detector spaced in the "space focus", peak broadening can be reduced and mass resolution can be greatly increased.
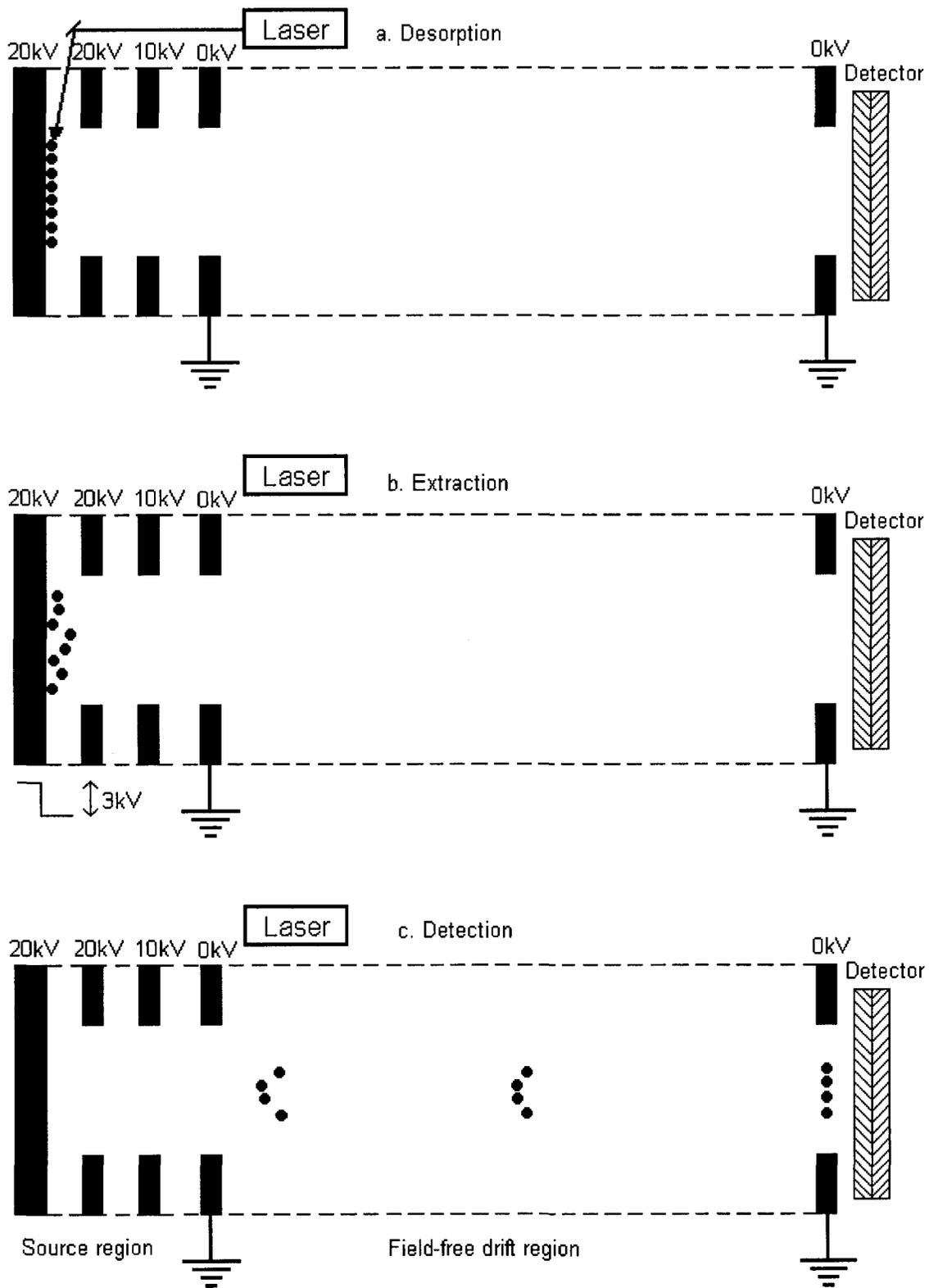
Figure 1.1.3 Principle of time-lag-focusing in a linear MALDI-TOF-MS

The second option of increasing the mass resolution is the use of a reflectron, Figure 1.1.4, which refocuses the ions in the direction of the detector. The reflector or ion mirror consists of a series of ring electrodes that ideally create a constant electric field through a linear voltage gradient. When the ions reach this electric field they will be slowed down and sent back towards the detector. When ions with the same m/z ratio have different initial energies, the ions with higher energies will move faster and arrive at the reflector before the ions with lower energies. These ions will be able to penetrate deeper in the reflectron than the ions with lower energies, thus increasing their flight time. By properly adjusting the voltages of the reflectron the ions can be refocused in order to hit the detector at the same time.
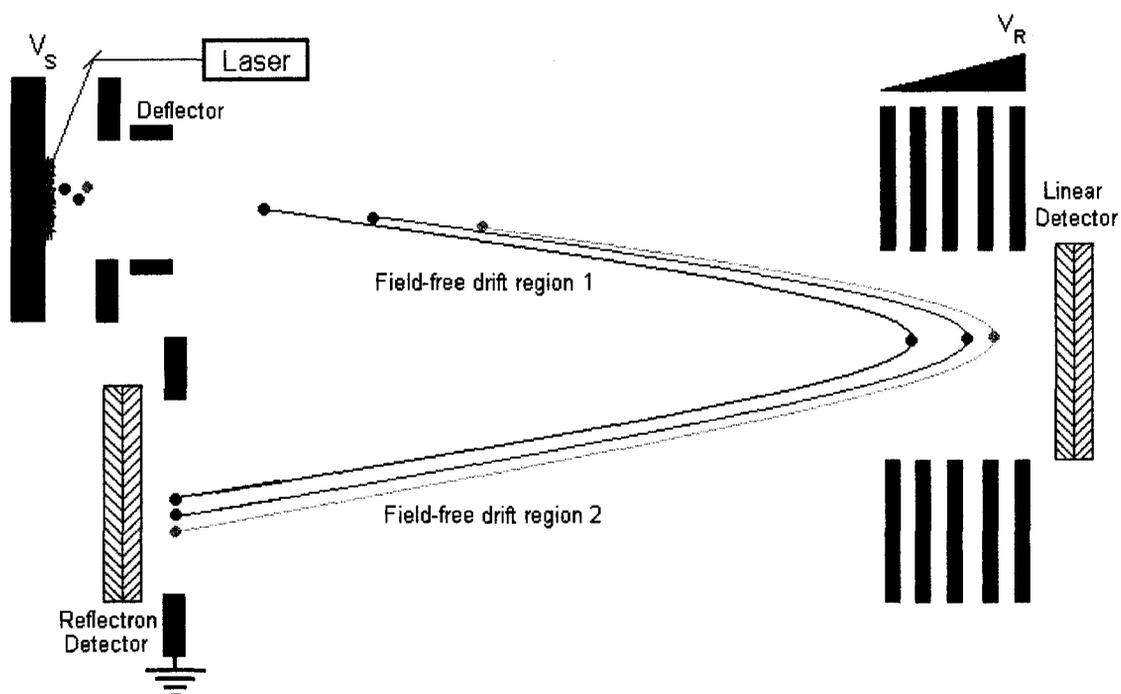
Figure 1.1.4 Simplified schematic of a reflectron equipped MALDI-TOF-MS

Modern MALDI-TOF instruments such as the Applied Biosystems/MDSSciex 4800 *Plus* MALDI-TOF/TOF™ Analyzer used for the work presented in this thesis incorporate not only delayed extraction and a reflectron but also a series of deflectors and lenses in order to increase both resolution and sensitivity. This instrument also includes a timed-ion-selector and a collision-induced dissociation (CID) cell which gives the instrument

MS/MS capability. The schematic diagram of this instrument is shown in Figure 1.1.5. The major hardware components of the 4800 MALDI-TOF/TOF™ Analyzer are: the TOF #1 region, timed-ion-selector (TIS), TOF #2 region and the detectors. The TOF #1 region contains the MALDI Source (Source 1), where the sample is ionized. After ionization a high voltage is applied to the stage to accelerate the ions, using a delayed extraction. The ions are focused by the Source 1 lens (Einzel lens) and then deflected by a set of deflectors past the laser mirror into the first field free region towards the TIS.

Einzel lenses usually consist of 3 or more cylindrical prisms to which an electric field is applied that will focus the ions towards the same end point when they pass through.

Linear detector

Mirror region

Source 2 region

QuanTIS™ Precursor Selector

Reflector detector

Sample stage and
sample plate

Deceleration Stack

Laser
Camera

OptiBeam™ on-axis laser

Source 1 Region
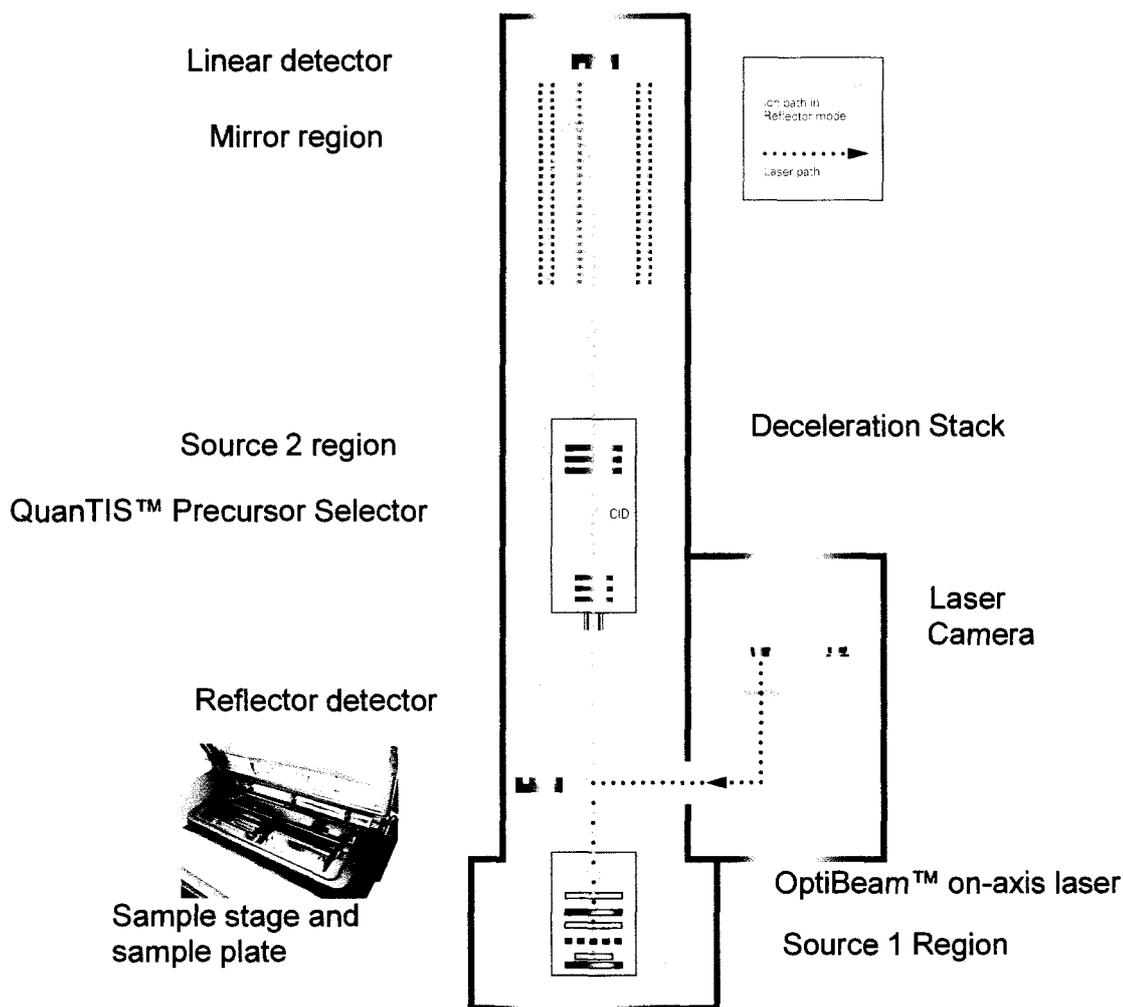
Figure 1.1.5 Schematic diagram of the Applied Biosystem 4800 MALDI TOFTOF™ Analyzer

The TIS is an ion gate that allows only selected ions with a specific m/z to pass through. It is composed by a set of 4 deflecting plates. Positive voltage is applied to 2 of the plates and negative to the other 2 plates. The TIS gate is either open or closed. When the gate is open all plates are at ground and all ions pass through freely. When the gate is closed, voltages are applied to the plates so that the ions are deflected from their path and will not reach the detector. Depending on the analysis type the TIS can have different functions. When the analyzer is used in an MS mode the TIS can be used as a low mass gate. In this case the gate will be closed until the arrival of the gated mass

12

when voltages will be turned off and ions can pass through freely. When the analyzer is used in the MS/MS mode the TIS is used to select the precursor ion. In this case before the arrival of the selected ion the gate is closed. When the precursor ion arrives the gate is opened and the precursor ion can pass through. In the MS/MS mode the ions are decelerated before reaching the TIS and entering the collision cell. After the precursor ion passes the gate is closed again. According to the manufacturer the resolution of the TIS gate on the 4800 MALDI Analyzer is m/z ± 2 Da.

The TOF #2 region contains the CID cell. This is used when the analyzer is in the MS/MS mode. The ions enter the CID cell and collide with the inert gas molecules (air, N_2, Ar, He etc) and fragment ions are produced. These are reaccelerated by the Source 2 using a delayed extraction and reach the field free path #2. In the MS mode the ions are refocused by Lens 3 and will reach the linear detector. In the MS/MS mode metastable ions are deflected by the metastable suppressor and the rest of the ions are focused by Lens 3 and then deflected by the mirror deflector into the 2-stage mirror. Ions are reflected by the mirror towards the reflectron detector.

1.1.2.2. Fourier Transform Ion Cyclotron Resonance Mass Analyzer

The best possible resolution and mass accuracy to date are achievable using Fourier transform ion cyclotron resonance mass spectrometers or FTICRMS (FTMS). This instrument uses an FTICR ion cell, Figure 1.1.6, in which the ions rotate around a fixed magnetic field at their cyclotron frequency[19]:

$$\omega_c = qB/m \qquad Eq. \ 9$$

where q is the charge of the ion, B is the magnetic field and m is the mass of the ion.

The ions are trapped in a Penning trap by applying a DC voltage between the two trapping plates. After trapping the ions, a frequency sweep pulse is applied which shifts through all the expected frequencies (and masses) so all the ions in the cell are excited in a spiralling orbit until the excitation pulsed is turned off or until the ion hits the wall. After the pulse is turned off the ions continue to oscillate at their final orbital radius and generate a small electric field on the detection plates. The resulting signal, called an

interferogram, consists of superimposed sine waves. Fast-Fourier transform is then used to generate a mass spectrum. This signal is amplified, digitized and stored.
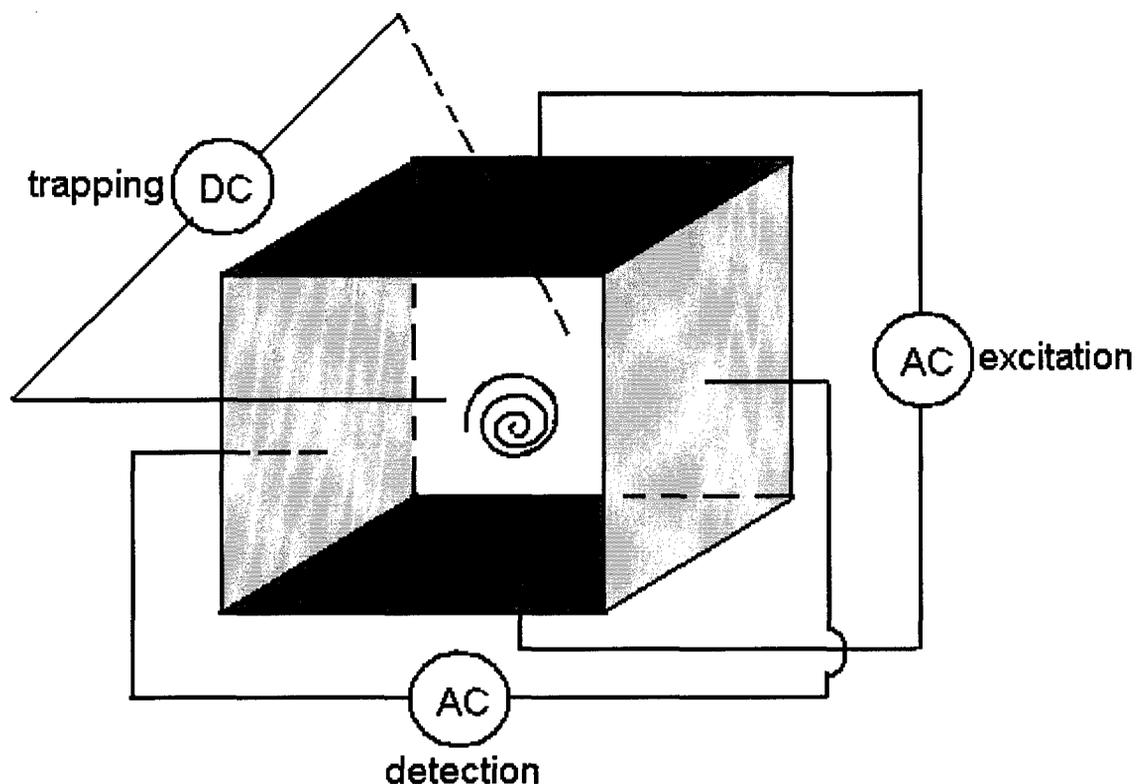
Figure 1.1.6 Simplified schematic of an FTICR ion cell

1.1.3. Detector

MALDI-TOF instruments are most commonly equipped with multi-channel plate detector or MCP[20]. Detection of the ions with an MCP, Figure 1.1.7, is by production of an amplified current due to direct impact of the ions on the surface of the detector. An MCP consists of a two-dimensional periodic array with millions of extremely small diameter glass capillaries fused together and sliced in a thin plate. Every single one of these capillaries in the array works as an independent electron multiplier. When a single ion enters one of these capillaries and hits the wall of the channel, an electron is emitted from the channel wall. The potential applied between the two ends of the capillaries will accelerate the emitted electron which by hitting the wall is going to produce more secondary electrons. Because this process is repeated many time along

the channel, this cascade process is generating several thousands of electrons. Two or more MCPs can be used in series so that a single ion is able to generate a very high output.
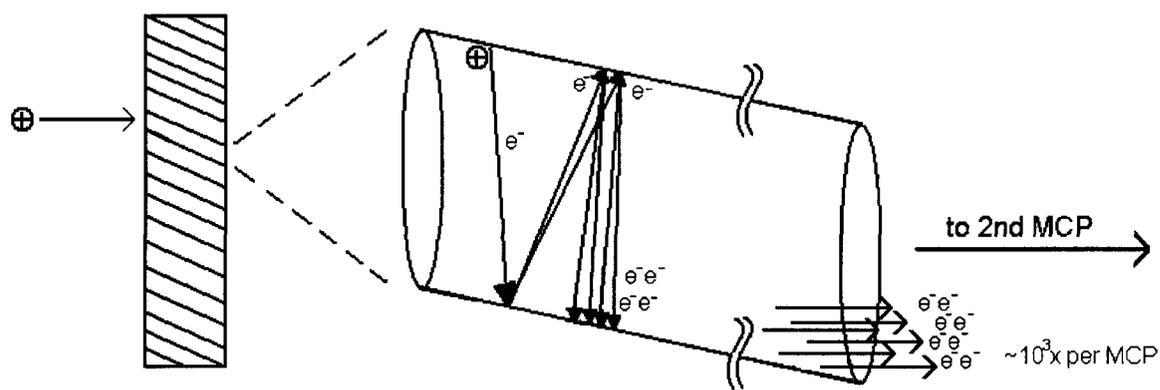
Figure 1.1.7 Schematic diagram of an MCP detector

1.1.4. Ion Fragmentation and Spectral Interpretation

MS/MS of peptide ions using a MALDI-TOF/TOF analyzer is achieved by high energy collision-induced dissociation (CID). In contrast to low energy CID in which a-, b- and y-fragments are the most abundant fragments, according to the Roepstorff/Fohlman nomenclature[21], shown in Figure 1.1.8, in high energy CID all ions shown below can be observed.
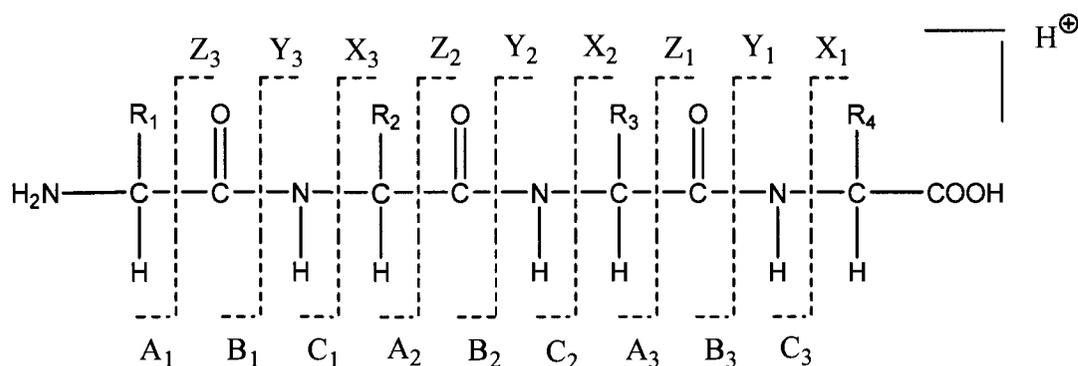
Figure 1.1.8 Nomenclature of peptide sequencing

In general, MS/MS spectral interpretation of a peptide is done by submitting the results into a search engine such as MASCOT[22]. The experimental mass values are compared by the program with calculated peptide masses or fragment ion mass values

and the closest match or matches is/are identified. If the sequence database does not contain the unknown protein, then the program will pull out those entries which have the closest homologies, often equivalent proteins from related species. The scoring algorithm is called Mowse scoring and it was described by Pappin *et al*[23].

1.2. Circular Dichroism

Circular dichroism or CD is a spectroscopic method based on the differentially absorbed right and left handed circularly polarized light. Far-UV CD in the spectral range 185-250 nm can be used to determine protein secondary structure. The chromophore in this region is the peptide bond and the CD response changes based on the folding of the protein. Different conformations of the protein (such as α-helix, β-sheet or random coil) give spectra that are characteristic in shape and signal intensity to each one of these conformations. Figure 1.1.9 shows the characteristic curves of poly-L-lysine in 3 different conformational states. Poly-L-lysine is in a disordered form at pH 7, but it changes into a α-helical structure if the pH is increased to 11.1. The β-sheet conformation can be achieved by heating of the sample to 65°C at pH 11.1. The characteristic features of an α-helix in the protein are negative bands at 222 nm and 208 nm and a strong positive band at around 190 nm. The characteristic features of a β-sheet are a negative band at 216 nm and a positive band at a positive band around 195 nm. For random coil no signal or very small bands are detected between 200 – 210 nm and a negative band is detected at around 196 nm. Using CD spectra, secondary structure of proteins can be estimated using a variety of computer algorithms.
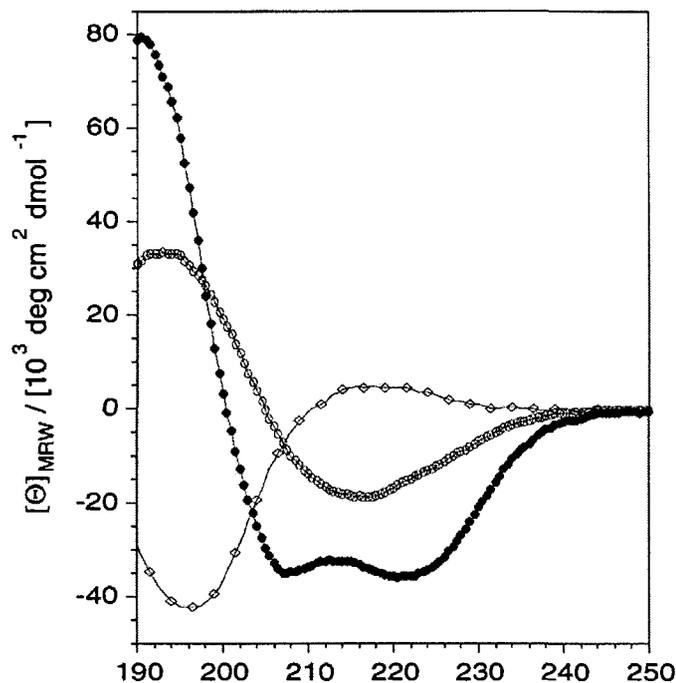
Figure 1.1.9 Far-UV CD spectra of different conformations of poly-L-lysine[24]; filled circles α-helix, open circles β-sheet and diamonds random coil.

1.3. Dynamic Light Scattering

Dynamic Light Scattering or DLS, also called Photon Correlation Spectroscopy or Quasi-Elastic Light Scattering, is a technique used to measure the size of particles in solution. The size of the particles is usually in the sub micron region. DLS measures the Brownian motion of the particles in solution and relates it to the size of the particle. The larger the particle, the slower the motion will be. Two important parameters that are necessary for accurate measurement of the Brownian motion (size) are the temperature and the viscosity of the medium. It is also necessary to keep the temperature constant throughout the measurement because viscosity is related to temperature. The hydrodynamic radius is calculated from the translational diffusion coefficient by using the Stokes-Einstein equation:

$$d(H) = \frac{kT}{3\pi\eta D} \qquad \text{Eq. 10}$$

where k is Boltzmann's constant, T is the temperature of the solution, η is the viscosity of the solution and D is the diffusion coefficient. Based on the hydrodynamic radius the molecular weight of a protein can be estimated using the following equation:

$$MW = 1.69R_h^{2.3394}$$ Eq. 11

where R_h is the measure hydrodynamic radius of the protein in a solvent.

1.4.   Polyacrylamide Gel Electrophoresis

Sodium dodecyl-sulphate polyacrylamide gel electrophoresis or SDS-PAGE[25] is probably the most widely used method for protein separation. The technique uses a porous gel as a sieve in which the proteins are separated based on their electrophoretic mobilities. The porous gel is formed by the copolymerization of acrylamide ($CH_2=CH-NH_2$) with the cross-linker bisacrylamide ($CH_2=CH-CO-NH-CH_2-NH-CO-CH=CH_2$). For SDS-PAGE, proteins are boiled with the strong anionic detergent SDS which will denature the proteins and applies the protein a negative charge that is proportional to its size. Depending on the procedure, a reducing agent can be added which will reduce the disulphide bonds. SDS binds to the protein in an approximately 1 SDS molecule to every 2 amino acids. SDS treated proteins will be negatively charged, denatured and will have approximately the same mass-to-charge ratios. During the electrophoresis process, SDS-treated proteins migrate towards the anode. Because of their approximately the same mass-to-charge ratios, the proteins will migrate with the same effective force, however because of their different sizes they will migrate with different velocities through the porous gel, the smaller proteins migrating faster than larger ones. There is an approximately linear relationship between migration distance and the log of MW. If proteins are not boiled in SDS before gel electrophoresis, different proteins will migrate differently in the gel having different sizes and mass-to-charge ratios. This is known as NATIVE-PAGE and is useful to detect protein complexes.

Proteins are visualized by staining with a dye such as Coomasie-blue or metal ions such as silver. The detection limit of the method is typically 20-100 ng for Coomasie staining and 1-10 ng for silver staining.

Gels have an optimal range of separation that is based on the percent of monomer present in the polymerization reaction. Gels containing different percentages of monomers such as 8%, 10% and 12% can effectively separate proteins between 24–205 kDa, 14-205 kDa, and 14-66 kDa proteins, respectively. Gels containing the same percentage of monomer are called fixed-percentage gels. The molecular weight range can be extended by using a gradient gel. Gradient gels also offer the advantage that proteins with similar molecular weight can be resolved as sharp bands. In gradient gels, proteins of a wide molecular weight enter the gel. Proteins with high molecular weight are going to be separated immediately while the proteins with lower MW are going to migrate freely until they reach the appropriate pore size. Proteins are going to migrate until they reach a limit of pore size, although migration will not stop completely[26].

1.5. Proteinaceous Infectious Particles

Proteinaceous infectious particles or prions are a type of infectious particles that were identified to be molecules of a normal body protein that have changed their three-dimensional structure. By misfolding these proteins cause a variety of fatal neurodegenerative diseases. As of today, no treatment of the disease other than palliation is possible. The available data for the past 20 years shows that more than 280,000 cattle have been reported to have been infected with Bovine Spongiform Encephalopathy (BSE), over 90% of them being reported in the UK[27]. Figure 1.1.1 shows the number of reported BSE cases worldwide and in the UK. Although the implementation of effective epidemiological screening systems has largely resolved the BSE crisis, the problems that prions cause to both human and animals are far from being resolved. Reports in the US have shown a significant rise in the number of CWD (chronic wasting disease) in elk and deer[28]. As in the case of humans, besides sporadic cases, there have been 4 reported cases of human variant Creutzfeld-Jakob (vCJD) disease

caused by blood transfusion[29-31]. This shows that the disease can be transmitted from human to human. In humans, prions can cause a variety of diseases such as Gerstmann-Straussler-Scheinker (GSS), Fatal familial insomnia (FFI) and Creutzfeldt-Jakob disease (CJD). A list of different prion diseases in both animals and humans along with the possible mechanisms of pathogenesis are shown in Table 1.1.1.
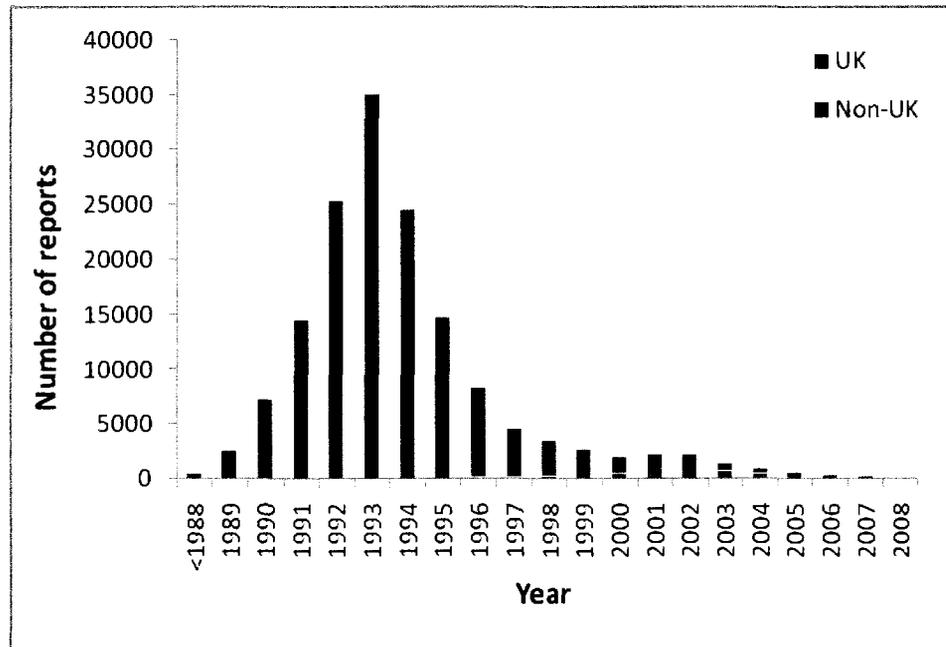
Figure 1.1.10 Incidence of reported BSE cases worldwide

Table 1.1.1 Prion related diseases and their possible mechanisms of pathogenesis

Disease	Mechanism of pathogenesis
Scrapie (sheep)	Infection in genetically susceptible sheep
Bovine Spongiform Encephalopathy (BSE, cattle)	Infection with prion-contaminated MBM (Meat and Bone Meal)
Transmissible Mink Encephalopathy (TME, mink)	Infection with prions from sheep and cattle
Chronic wasting disease (CWD, mule deer, elk)	Unknown
Feline spongiform encephalopathy (FSE, cats)	Infection with prion-contaminated MBM
Exotic ungulate encephalopathy (EUE, greater kudu, nyalal, oryx)	Infection with prion-contaminated MBM
Kuru (Fore people)	Infection through ritualistic cannibalism
Variant Creutzfeld-Jakob disease (vCJD)	Infection from prion-contaminated HGH, dura mater grafts etc.
Familial Creutzfeld-Jakob disease (fCJD)	Germline mutation in PrP gene
Gerstmann-Sträussler-Scheinker disease (GSS)	Germline mutation in PrP gene
Fatal familial insomnia (FFI)	Germline mutation in PrP gene (D178N and M129)
Sporadic Creutzfeld-Jakob disease (sCJD)	Somatic mutation of spontaneuous conversion of PrP^C into PrP^{Sc}?

1.5.1. Discovery of PRIONS

Bjorn Sigurdsson who studied scrapie in sheep suggested that the disease was caused by a "slow virus"[32]. A few years later others have suggested that other diseases such as kuru[33] and Creuzfeld-Jacob Disease[34] (CJD) are also caused by a slow virus. In all cases, scientists have found similarities in the microscopic images of the central nervous system (CNS). It was 1967 when the first paper was published by Alper[35] and coworkers in which the authors suggested that nucleic acids were not part of the scrapie infectious agent. They found that the scrapie agent was extremely resistant to inactivation by UV and ionizing irradiation. It was latter shown that not only the scrapie agent but also the kuru and CJD agents were also highly resistant to inactivation using ionizing irradiation[36]. It was not until 1982 when Prusiner[37] *et al.* published their finding in which they suggested that the scrapie agent lacked nucleic acid and was caused by a protein. In order to distinguish them from viruses plasmids and viroids the term prion has been proposed: *``Although the scrapie agent is irreversibly inactivated by alkali, five procedures with more specificity for modifying nucleic acids failed to cause inactivation. The agent shows heterogeneity with respect to size, apparently a result of its hydrophobicity; the smallest form may have a molecular weight of 50,000 or less. Because the novel properties of the scrapie agent distinguish it from viruses, plasmids, and viroids, a new term "prion" is proposed to denote a small proteinaceous infectious particle which is resistant to inactivation by most procedures that modify nucleic acids``.[37]* In 1997 Dr. Stanley B. Prusiner was awarded with the Nobel Prize in Physiology or Medicine for his discovery of prions.

1.5.2. Cellular PrP

The cellular form of the prion protein, also called PrP^C, is a naturally occurring protein encoded by the *Prnp* gene. It is a 254 amino acid glycoprotein present bound to the surface of the neurons by a glycosyl phosphatidyl inositol (GPI) anchor. Its secondary structure is dominated by three α-helices and 2 small β-strands, all situated in the C-

terminal end of the protein. The N-terminal part of the protein is unstructured and contains an octarepeat region. The protein presents 2 N-glycosilations one at residue N181 and one at residue N197 and a disulphide bond between 179C and 214C[38]. The GPI-anchor is found at position 231S. The structural features, and some biochemical properties of the protein are shown in Figure 1.1.11[39].

PK cleavage
site for PrP[Sc]

↓   ↓ ↓   CHO   CHO
              181N   179N   ⌐GPI anchor
1   18   51   90   134 | | 231   254

| Signal | | ▮▮▮▮▮ | CC | HC | PK-resistant fragment | | MA |

Octarepeat   111
region   └── S — S ──┘

Unstructured   Stuctured

β-sheet1   α-helix 1   β-sheet 2   α-helix 2   α-helix 3

CC - charged cluster
HC - hydrophonic core
GPI - glycosyl phosphatidyl inositol
CHO - glycosilation site
S-S disulphide bond
MA - membrane anchor region

	PrP[C]	PrP[Sc]
	➢Secondary structure: 3 α-helices, 2 β-sheet	➢Secondary structure: high percentage of β-sheet
	➢Soluble	➢Insoluble
	➢Easily digested by proteases	➢Resistant to digestion by proteases

Figure 1.1.11 Structural features and biochemical properties of the cellular prion protein[39]

The structure of the Syrian hamster prion protein has been determined in 1997 from NMR studies[40]. The protein is soluble and easily digested by proteases. The function of the protein however is still elusive. The functions attributed so far are: synaptic transmission[41, 42], copper binding[43], induction or protection against apoptosis [44, 45], immunoregulation and many others[46]. It has also been shown that postnatal

depletion of the PrP^C does not result in neurodegeneration[47]. The sequence homology between the Syrian hamster and human prion proteins from residue 90 to residue 231 is approximately 87%.

1.5.3. Scrapie PrP and prion conversion

The scrapie form of the prion protein, also called PrP^{Sc}, was discovered by enriching fractions from Syrian hamster infected brain extracts[48, 49]. It was found that the protein has the same primary structure as the cellular form[50]. Even after extensive research there is only limited information about the structure, function and epidemiology of prions. It is well documented that the scrapie form of the protein forms insoluble aggregates, it is partially resistant to proteinase K digestion and its secondary structure is dominated by β-sheet[51]. Because the PrP^{Sc} form of this protein forms large and insoluble aggregates, the use of conventional methods for structural determination to generate a high resolution structure, such as NMR and X-ray, have all failed.

There are 3 different proposed hypothesis concerning prion conversion and prion diseases:

    a) the protein-only hypothesis

    b) the virino hypothesis

    c) the hypothesis that for in vitro conversion the PrP^C into PrP^{Sc} requires an RNA molecule

The most widely accepted model is the protein-only hypothesis[52]. It has been shown that knockout mice have all failed to develop the disease when infected with infectious agent from brain homogenate by inoculation[53]. It has also been shown that reintroduction of the gene in the mice made them susceptible for the development of the disease[46]. Based on the protein only hypothesis, two different conversion mechanisms have been proposed. The template directed mechanism[54] proposed that upon infection of the host, PrP^{Sc} acts as a substrate for the conversion of PrP^C into PrP^{Sc}. This newly converted PrP^C then becomes the substrate for another PrP^C which will ultimately lead to the formation of the fibrils. The conversion of PrP^C into PrP^{Sc} is a

reversible process; with the equilibrium favouring the conformation of PrP^C. Converted PrP^{Sc} is able to maintain its conformation only when it adds onto a fibril-like seed or aggregate of PrP^{Sc}. The schematic of the mechanism is shown in Figure 1.1.12.
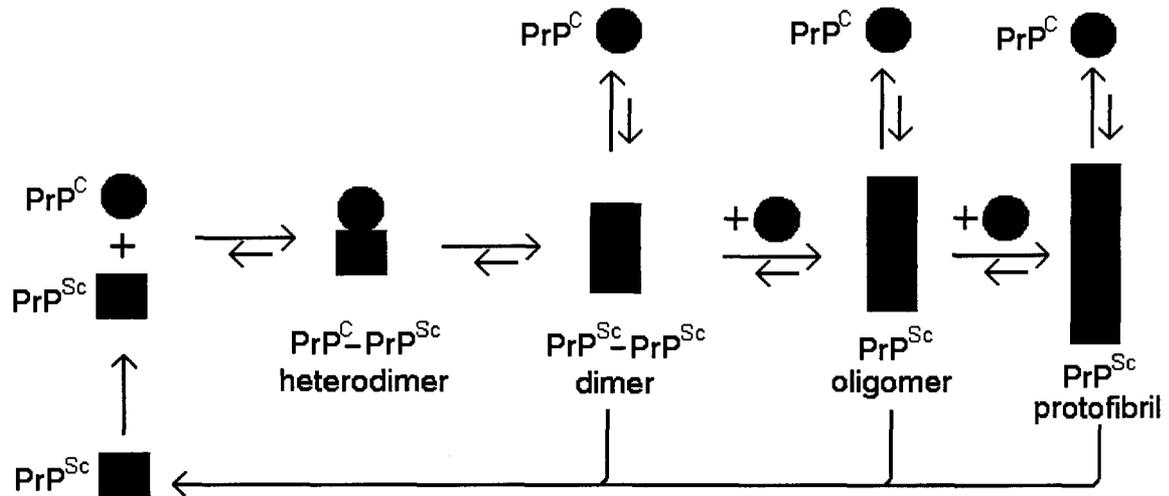
Figure 1.1.12 Template directed mechanism for prion conversion

It has been found that familial forms of the disease are related to mutations in the Prnp gene, therefore a second mechanism of spontaneous conversion has been proposed[55].

The protein only hypothesis is sustained by reports on in vitro conversion of the protein. Caughey *et al.*[56] have reported the conversion of PrP^c to protease-resistant forms similar to PrP^{Sc} in a cell-free system composed of purified constituents. The conversion required the presence of preexisting PrP^{Sc} in the conversion system, showing that PrP^{Sc} can derive from PrP^c-PrP^{Sc} interactions. Other in vitro conversion methods have been proposed by Baskakov *et al*[57]. They were able to convert the PrP^C into PrP^{Sc} without the presence of any seeding material. Although no seeding material was present, the presence of denaturing agents, a buffer system and shaking of the reaction mixture were needed for successful conversion. Although different conversion methods have been proposed over the years by different scientists, the condition for in vitro conversion of the prion protein are still unknown.

The virino hypothesis[58] proposed that the infectious agent in the case of scrapie is a nucleic acid that is associated or coated with a protein, for which PrP^c being the

most likely candidate. Current research has not been able to demonstrate the presence of any nucleic acids in prion infected samples.

The last hypothesis[59] was proposed when Deleault *et al.* who reported that presence of specific RNA molecules can stimulate prion conversion in mammals, but not in invertebrates. All these experiments were conducted in vitro and in vivo experiments are still needed to confirm this hypothesis.

1.5.4. Prion species barrier and protein X

One of the most disputed issues surrounding the prion protein and prion diseases is the existence of the species barrier. Early reports suggested that the transmission of the prion disease from one species to the other is not possible. Latest studies have revealed that transmission of the prion diseases is not necessarily species related. It has been found that inoculation with PrP^{Sc} from a different species is possible and can result in a very long incubation times for the first passage. The second passage is usually the same as infecting a homologous host. It has also been found that transmission between different strains of the same species is not possible or ineffective. Research performed on transgenic mice (Tg mice) has identified three factors involved in the species/transmission barrier: a) differences between the primary structure of the donor and acceptor; b) the strain of prion, and c) the species specificity of protein X. It has been reported based on genetic studies that there is a protein X[60, 61] which binds to the prion protein and makes the conversion from PrP^{C} to PrP^{Sc} much more favourable.

1.5.5. Future of the prion research and the objective of the thesis work

After extensive research involving prions for the last few decades there are still a lot of questions that prion researchers have to answer, such as:
- What is the exact physiological function of PrP^{C}?
- What is the precise physical structure of PrP^{Sc}?
- Is there a species or a transmission barrier?
- What factors determine the species barrier in prion infections?

- What are the host susceptibility factors that promote prion infection?
- What are the molecular mechanisms that could provide an efficacious therapy?

The purpose of my research described in this thesis is to develop new analytical methods, mostly based on mass spectrometry, in order to gain insight into the mechanisms involved in the folding and misfolding of the prion protein, more specifically to use the microwave-assisted acid and base hydrolysis of the prion protein combined with mass spectrometric analysis for the characterization of the prion protein.

1.6. References

(1)   Karas, M.; Bachmann, D.; Hillenkamp, F. *Anal. Chem.* **1985**, *57*, 2935-2939.

(2)   Koichi Tanaka, H. W., Yutaka Ido, Satoshi Akita, Yoshikazu Yoshida, Tamio Yoshida, T. Matsuo, *Rapid Communications in Mass Spectrometry* **1988**, *2*, 151-153.

(3)   Karas, M.; Bachmann, D.; Bahr, Y.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion processes* **1987**, *78*.

(4)   Karas, M.; Hillenkamp, F. *Anal Chem* **1988**, *60*, 2299-2301.

(5)   Hillenkamp, F.; Karas, M.; Beavis, R. C.; Chait, B. T. *Anal Chem* **1991**, *63*, 1193A-1203A.

(6)   Karas, M.; Gluckmann, M.; Schafer, J. *J Mass Spectrom* **2000**, *35*, 1-12.

(7)   Kaufmann, R.; Chaurand, P.; Kirsch, D.; Spengler, B. *Rapid Commun Mass Spectrom* **1996**, *10*, 1199-1208.

(8)   Weinberger, S. R.; Boernsen, K. O.; Finchy, J. W.; Robertson, V.; Musselman, B. D. *in Proceedings of the 41st ASMS Conference on Mass Spectrometry and Allied Topics; San Francisco, CA, May 31-June 4, 1993,8, 199-204.*

(9)   Xiang, F.; Beavis, R. C. *Rapid Commun Mass Spectrom* **1994**, *8*, 199-204.

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

(11)   Dai, Y.; Whittal, R. M.; Li, L. *Anal Chem* **1999**, *71*, 1087-1091.

(12)   Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.

(13)   Wiley, W. C.; McLaren, I. H. *Rev. Sci. Inst.* **1955**, *26*, 1150-1157.

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

(15)   Whittal, R. M.; Li, L. *American Laboratory* **1997**, *28*, 30-36.

(16)   Spengler, B.; Cotter, R. J. *Anal Chem* **1990**, *62*, 793-796.

(17)   Mamyrin, B. A.; Karataev, D. B.; Schmikk, B. A.; Zagulin, B. A. *Sov. Phys. - JETP* **1973**, *37*, 45-48.

(18)   Mamyrin, B. A. *Int. J. Mass Spectrom. Ion processes* **1994**, *131*, 1-19.

(19)   Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom Rev* **1998**, *17*, 1-35.

(20)   Wurz, P.; Gulber, L. *Rev. Sci. Inst.* **1996**, *67*, 1705-2031.

(21)   Roepstorff, P.; Fohlman, J. *Biomed Mass Spectrom* **1984**, *11*, 601.

(22)   *http://www.matrixscience.com/search_form_select.html* (last accessed Dec. 18th, 2008).

(23)   Pappin, D. J.; Hojrup, P.; Bleasby, A. J. *Curr Biol* **1993**, *3*, 327-332.

(24)   Fandrich, M.; Dobson, C. M. *EMBO J* **2002**, *21*, 5682-5690.

(25)   Laemmli, U. K. *Nature* **1970**, *227*, 680-685.

(26)   Margolis, J.; Kenrick, K. G. *Biochem Biophys Res Commun* **1967**, *27*, 68-73.

(27)   *http://www.oie.int/eng/info/en_esbmonde.htm* (last accessed Dec. 20th, 2008)

(28)   Sigurdson, C. J.; Manco, G.; Schwarz, P.; Liberski, P.; Hoover, E. A.; Hornemann, S.; Polymenidou, M.; Miller, M. W.; Glatzel, M.; Aguzzi, A. *J Virol* **2006**, *80*, 12303-12311.

(29)   Llewelyn, C. A.; Hewitt, P. E.; Knight, R. S.; Amar, K.; Cousens, S.; Mackenzie, J.; Will, R. G. *Lancet* **2004**, *363*, 417-421.

(30)   Peden, A. H.; Head, M. W.; Ritchie, D. L.; Bell, J. E.; Ironside, J. W. *Lancet* **2004**, *364*, 527-529.

(31)   Estebe, J. P. *Ann Fr Anesth Reanim* **1997**, *16*, 955-963.

(32)   Sigurdsson, B. *Br. Vet. J.* **1954**, *110*, 341-354.

(33)   Hadlow, W. J. *Lancet ii* **1959**, 289-290.

(34)   Klatzo, I.; Gajdusek, D. C.; Zigas, V. *Lab Invest* **1959**, *8*, 799-847.

(35)   Alper, T.; Cramp, W. A.; Haig, D. A.; Clarke, M. C. *Nature* **1967**, *214*, 764-766.

(36)   Gibbs, C. J., Jr.; Gajdusek, D. C.; Latarjet, R. *Proc Natl Acad Sci U S A* **1978**, *75*, 6268-6270.

(37)   Prusiner, S. B. *Science* **1982**, *216*, 136-144.

(38)   Turk, E.; Teplow, D. B.; Hood, L. E.; Prusiner, S. B. *Eur J Biochem* **1988**, *176*, 21-30.

(39)   Aguzzi, A.; Sigurdson, C.; Heikenwaelder, M. *Annu Rev Pathol* **2008**, *3*, 11-40.

(40)   James, T. L.; Liu, H.; Ulyanov, N. B.; Farr-Jones, S.; Zhang, H.; Donne, D. G.; Kaneko, K.; Groth, D.; Mehlhorn, I.; Prusiner, S. B.; Cohen, F. E. *Proc Natl Acad Sci U S A* **1997**, *94*, 10086-10091.

(41)   Collinge, J.; Whittington, M. A.; Sidle, K. C.; Smith, C. J.; Palmer, M. S.; Clarke, A. R.; Jefferys, J. G. *Nature* **1994**, *370*, 295-297.

(42)   Mallucci, G. R.; White, M. D.; Farmer, M.; Dickinson, A.; Khatun, H.; Powell, A. D.; Brandner, S.; Jefferys, J. G.; Collinge, J. *Neuron* **2007**, *53*, 325-335.

(43)   Millhauser, G. L. *Annu Rev Phys Chem* **2007**, *58*, 299-320.

(44)   Frigg, R.; Wenzel, A.; Samardzija, M.; Oesch, B.; Wariwoda, H.; Navarini, A. A.; Seeliger, M. W.; Tanimoto, N.; Reme, C.; Grimm, C. *Exp Eye Res* **2006**, *83*, 1350-1358.

(45)   Zhang, Y.; Qin, K.; Wang, J.; Hung, T.; Zhao, R. Y. *Biochem Biophys Res Commun* **2006**, *349*, 759-768.

(46)   Aguzzi, A.; Polymenidou, M. *Cell* **2004**, *116*, 313-327.

(47)   Mallucci, G. R.; Ratte, S.; Asante, E. A.; Linehan, J.; Gowland, I.; Jefferys, J. G.; Collinge, J. *EMBO J* **2002**, *21*, 202-210.

(48)   Bolton, D. C.; McKinley, M. P.; Prusiner, S. B. *Science* **1982**, *218*, 1309-1311.

(49)   Prusiner, S. B.; Bolton, D. C.; Groth, D. F.; Bowman, K. A.; Cochran, S. P.; McKinley, M. P. *Biochemistry* **1982**, *21*, 6942-6950.

(50)   Stahl, N.; Baldwin, M. A.; Teplow, D. B.; Hood, L.; Gibson, B. W.; Burlingame, A. L.; Prusiner, S. B. *Biochemistry* **1993**, *32*, 1991-2002.

(51)   Pan, K. M.; Baldwin, M.; Nguyen, J.; Gasset, M.; Serban, A.; Groth, D.; Mehlhorn, I.; Huang, Z.; Fletterick, R. J.; Cohen, F. E.; et al. *Proc Natl Acad Sci U S A* **1993**, *90*, 10962-10966.

(52)   Griffith, J. S. *Nature* **1967**, *215*, 1043-1044.

(53)   Bueler, H.; Aguzzi, A.; Sailer, A.; Greiner, R. A.; Autenried, P.; Aguet, M.; Weissmann, C. *Cell* **1993**, *73*, 1339-1347.

(54)   Gajdusek, D. C. *J Neuroimmunol* **1988**, *20*, 95-110.

(55)   Hsiao, K.; Prusiner, S. B. *Neurology* **1990**, *40*, 1820-1827.

(56)   Kocisko, D. A.; Come, J. H.; Priola, S. A.; Chesebro, B.; Raymond, G. J.; Lansbury, P. T.; Caughey, B. *Nature* **1994**, *370*, 471-474.

(57)   Breydo, L.; Makarava, N.; Baskakov, I. V. *Methods Mol Biol* **2008**, *459*, 105-115.

(58)   Kimberlin, R. H.; Wilesmith, J. W. *Ann N Y Acad Sci* **1994**, *724*, 210-220.

(59)   Deleault, N. R.; Lucassen, R. W.; Supattapone, S. *Nature* **2003**, *425*, 717-720.

(60)   Telling, G. C.; Scott, M.; Mastrianni, J.; Gabizon, R.; Torchia, M.; Cohen, F. E.; DeArmond, S. J.; Prusiner, S. B. *Cell* **1995**, *83*, 79-90.

(61)   Kaneko, K.; Zulianello, L.; Scott, M.; Cooper, C. M.; Wallace, A. C.; James, T. L.; Cohen, F. E.; Prusiner, S. B. *Proc Natl Acad Sci U S A* **1997**, *94*, 10069-10074.

Chapter 2: Microwave-assisted acid and base hydrolysis

2.1. Introduction

Simple and efficient digestion of proteins is a key factor in the characterization of proteins by mass spectrometry. To map the sequence of an entire protein and to detect any modifications along its sequence, most often the generation, detection and sequencing of short, sometimes partially overlapping sequences is needed. In some cases this can be very difficult. Sometimes the use of a number of different enzymes in combination with different types and amounts of detergents is needed which can make the process very time consuming and complex.

An alternative method to enzymatic digestion was developed by Zhong[1] et al. This method uses a high concentration of an acid coupled with microwave irradiation in order to hydrolyse proteins. This technique specifically cleaves at the amide bond of the intact proteins and produces sequence specific ladders containing N- and C- terminal amino acids. The mass spectrometric analysis of the hydrolysate results in a spectrum consisting of peaks of the polypeptide ladders from which the amino acid sequence along with any possible modifications can directly be read.

Protein stability is a major factor in terms of activity and is influenced by a number of factors. The contribution each residue of a specific protein makes to, or takes away from, the stability of a protein is small and therefore the stability of a protein is determined by a large number of small positive and negative interaction energies. Some of the factors that can influence the stability of the protein are hydrophobic effects, hydrogen bonds and disulphide bonds. The disulphide bond is an important post-translational modification. It can stabilize the folded form of a protein by holding two portions of the protein together, biasing the protein towards the folded conformation. Sometimes the disulfide bond may form the nucleus of a hydrophobic core of the folded protein. By linking together two different segments of the protein chain, the effective local concentration of protein residues is increased which results in a lower effective

local concentration of water molecules, thus water molecules will not be able to disrupt hydrogen bonds and break up the secondary structure of a protein. Although important for stability, the disulphide bond cannot fully stabilize a protein and it can be present in both the folded and unfolded form of a protein.

Since the prion molecules to be studied contain disulphide bonds, in this work we have investigated whether protein structure and the presence of disulphide bonds can influence the microwave-assisted acid hydrolysis process of standard proteins. We have also investigated the use of concentrated base and microwave irradiation as a means of degrading different proteins followed by MALDI-TOF-MS analysis.

2.2. Experimental

2.2.1. Materials and Reagents

Unless otherwise noted, all chemicals were purchased from Sigma (St. Louis, MO) and were of analytical grade. For HPLC separations, MS analysis, and preparation of digestions, Optima grade water, methanol and acetonitrile were used (Fisher Scientific, Mississauga, ON). 37% HCl (ACS Grade) was from Merck KGaA, Darmastadt, Germany. NaOH was from Fisher Scientific (Missisauga, ON). Horse heart Cytochrome C and Lysozyme were obtained from Sigma Aldrich (Oakville, ON).

2.2.2. Microwave-assisted acid hydrolysis

For the microwave-assisted acid hydrolysis (MAAH) sequencing experiments 1 mg mL^{-1} protein samples were prepared in 50 mM NaAc, pH= 5.5. An aliquot of 40 μL of this stock solution was mixed with an equal amount of 6M HCl and 1 μL of DTT in a 1.5 mL polypropylene vial. The vial was capped and placed inside a household microwave oven with 1000W power output at 2450MHz. A loosely covered water container containing 100 mL of water was placed beside the sample inside the microwave oven so that the extra energy would be absorbed by the water. Samples were subjected to microwave irradiation for various irradiation times ranging from 30 seconds to 4 minutes. The pressure inside the vial was released every 60 seconds by

opening the polypropylene vial's cap. Samples were allowed to cool at room temperature for about 5 minutes. The hydrolysates were dried in a vacuum centrifuge (SpeedVac) to remove all the acid. Samples were than reconstituted in 50% ACN/Water. The ACN was removed using a vacuum centrifuge. Samples were acidified and diluted using 0.1% TFA for desalting.

2.2.3. Microwave-assisted base hydrolysis

For the microwave-assisted base hydrolysis (MABH) experiments 1 mg mL^{-1} protein samples were prepared in 50 mM NaAc, pH= 5.5. An aliquot of 40 μL of this stock solution was mixed with 2 μL of 450 mM DTT and an equal amount of NaOH in a 1.5 mL polypropylene vial. The vial was capped and placed inside a household microwave oven with 1000W power output at 2450MHz. A loosely covered water container containing 100 mL of water was placed beside the sample inside the microwave oven so that the extra energy would be absorbed by the water. Samples were subjected to microwave irradiation for various irradiation times ranging from 30 seconds to 4 minutes. The pressure inside the vial was released every 30 seconds by opening the polypropylene vial's cap. Samples were allowed to cool at room temperature for about 5 minutes. Samples were than neutralized using concentrated HCl, to a final pH of 8.0. Dithiothreitol (DTT) was added in a final concentration of 15 mM to reduce the disulphide bonds. The mixture was incubated at 37°C for 60 minutes. After reduction the samples were acidified and diluted using 0.1 – 10% TFA for desalting.

2.2.4. Desalting and quantification

Desalting and peptide quantification were carried out in an Agilent 1100 HPLC system (Palo Alto, CA) using a newly developed method in our laboratory[2]. Briefly, desalting of the hydrolysates was performed on a 4.6 mm × 5 cm Polaris C18 A column with a particle size of 3 μm and 300 Å pore (Varian, USA). After loading of the peptide sample, the column was flushed with 97.5% mobile phase A (0.1% TFA in water) and 2.5% mobile phase B (0.1% TFA in ACN) and the salts were effectively removed. Subsequently, the concentration of mobile phase B in the mobile phase was step-wise

increased to 90% to ensure complete elution of the polypeptide fractions from the column. During the polypeptide elution process, a chromatographic peak was produced at 214 nm and based on the peak area the amount of peptides was determined. Four protein digests of various amounts were used as standards for the generation of a linear calibration curve between the peak area and the injected peptide amount. The calibration curve was generated as y=430.04x-269.56, where y refers to the peak area of the peptide sample, and x refers to the amount of peptide analyzed.

2.2.5. MALDI-TOF-MS of sequence ladders

The desalted samples were concentrated in a SpeedVac vacuum centrifuge. The samples were then mixed with matrix solution of α-cyano-4-hydrocynnamic acid (CHCA). Samples were then deposited on a MALDI plate using the two-layer sample preparation method[3] for matrix-assisted laser desorption ionization (MALDI) analysis. The MALDI-TOF mass spectra were obtained on an Applied Biosystems/MDSSciex 4800 Plus MALDI-TOF/TOF™ Analyzer. Ionization was performed with a diode-pumped Nd:YAG laser at 355 nm with a pulse frequency of 200 Hz with less than 500 picoseconds duration/pulse. The analyzer was used in a linear mode of operation with positive ion detection.

2.2.6. Circular Dichroism

CD spectra were recorded with an Olis DSM 17 spectrometer (Olis, Bogart, USA). Scans were performed with a bandwidth of 2 nm and data spacing of 1.0 nm. Each spectrum was recorded at 20°C using a 0.02 cm cuvette on protein samples in the concentration range from 0.5 to 1.2 µg µL^{-1}. Protein concentrations were estimated by measuring absorbance at 280 nm. All spectra represent the average of five individual scans after subtracting the background spectra. All buffers and solutions were filtered using a low-binding Millex®-GV 0.22µm filter (Millipore, Billerica, USA). To monitor the conformational changes under acidic conditions, protein samples were mixed 1:1 (v/v) with 6M HCl or 2 M NaOH followed by the recording of the CD spectra. Recording of the CD spectra was possible between 190 and 250 nm in 50mM NaAc, 200 and 250 nm in 3M HCl and 210 and 250nm in 1M NaOH. Deconvolution of the CD spectra was

performed using the CDPro software package. Secondary structure was determined using the programs included in the software package (CDSSTR, CONTINLL and SELCON3) and 3 different Reference Sets. The reference sets used were SP43, SDP48 and SMP56.

2.2.7. Dynamic Light Scattering

DLS spectra were recorded with a Malvern Zetasizer Nano-S (Malvern Instruments Ltd., UK). Scans were performed at an angle of 178° in a 1 mL cuvette. Samples were in 50 mM sodium acetate, pH 5.5, having a concentration of 1 mg mL^{-1}. For DLS experiments under acidic and basic conditions 2 mg mL^{-1} samples were prepared in 50mM sodium acetate and mixed 1 to 1 (v/v) with 6M HCl or 2 M NaOH before analysis. Buffer was filtered using a low-binding Millex®-GV 0.22µm pore size syringe filter (Millipore, Billerica, USA).

2.3. Results and Discussion

2.3.1. Principle of the MAAH

The MAAH was developed in 2004 by Zhong *et al.* using a series of different standard proteins. It was shown that proteins having a wide range of pI values and various post-translational modifications can be sequenced by mixing them with 6M HCl (3M final concentration), followed by a brief hydrolysis using microwave radiation and analyzed by MALDI-TOF-MS. The schematic of the technique is shown in Figure 2.3.1. As it is shown in the figure the technique is based on the analysis of a polypeptide ladder. Under optimal conditions the mass spectrum consists exclusively of N- and C-terminal peptides. For a shorter hydrolysis time than the optimum, peaks across the entire mass range can be detected. For optimum hydrolysis time the mass spectrum consists only of terminal peptides and no internal fragments. The presence of exclusively terminal peptides makes the reading of the protein sequence very easy. If the duration of the microwave radiation is increased internal peptide peaks are detected together with the terminal peptide peaks. If the irradiation is long enough the high mass peptide peaks and the molecular ion peak disappear and the mass spectrum will be dominated by low mass terminal and internal peaks.
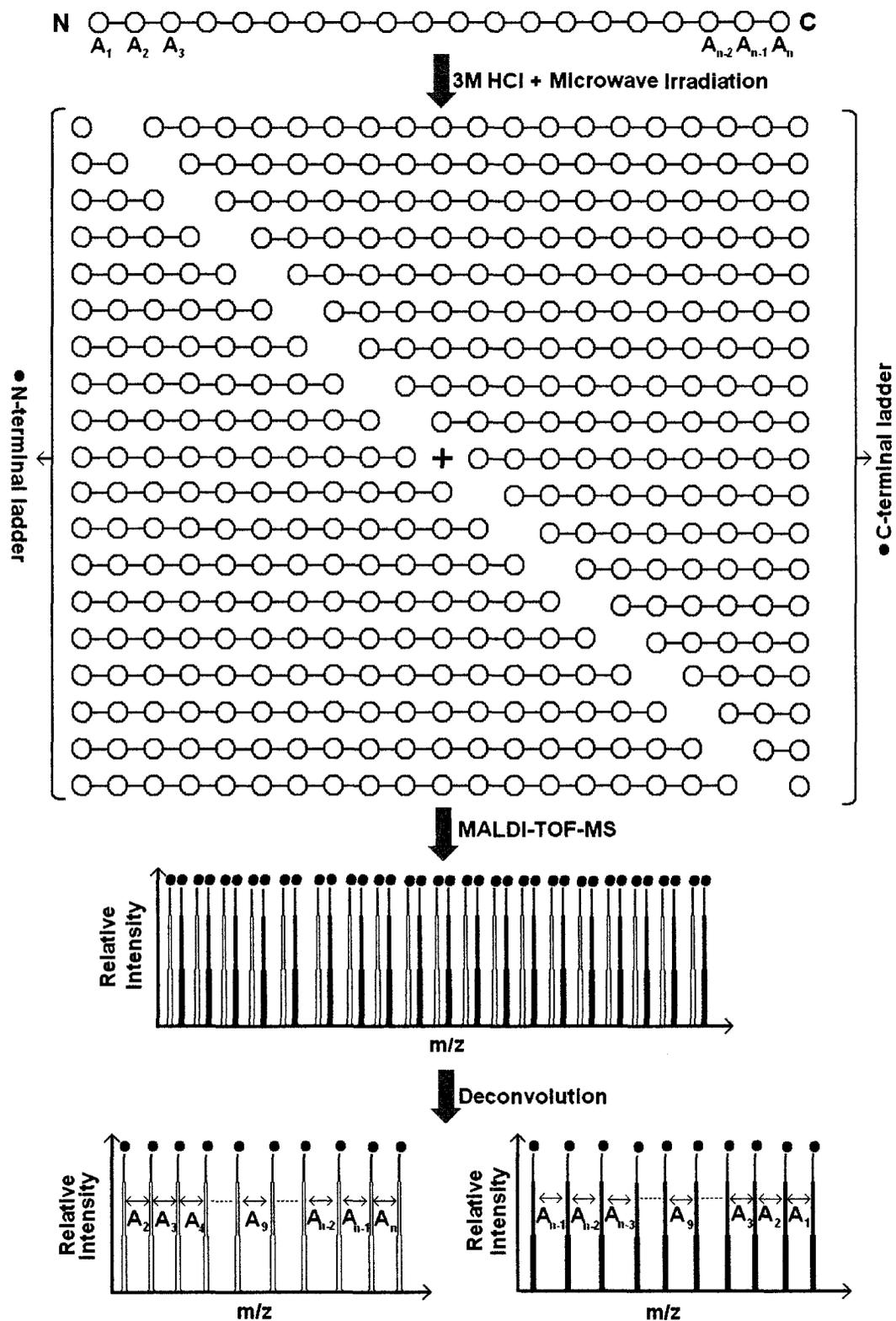
Figure 2.3.1 Schematic of the microwave assisted-acid hydrolysis technique. Under optimum conditions the mass spectrum consists exclusively form N- and C-terminal peptides. Amino acid sequences and modifications are directly read from the mass difference of adjacent peaks within the same series of the polypeptide ladder.

Because amides are less reactive and water is a weak nucleophile the hydrolysis of the amide bonds necessitates the presence of an acid or base for hydrolysis. Figure 2.3.2 shows a possible mechanism of the acid hydrolysis process[4]. Generally, the acid hydrolysis process involves the breakage of the acyl–nitrogen bond through a tetrahedral intermediate. Under acidic conditions the amidecarbonyl is protonated in the reaction between the amide and the hydronium ion which makes the amide bond more electrophilic. Now, water which is a weak nucleophile can attack the carbocation. The amino group will become a good leaving group. The electron pair from either oxygen can now be donated to the carbon atom with the protonated amino group becoming a leaving group. The proton of the carboxylic acid is extracted by the amide which now acts as a base.
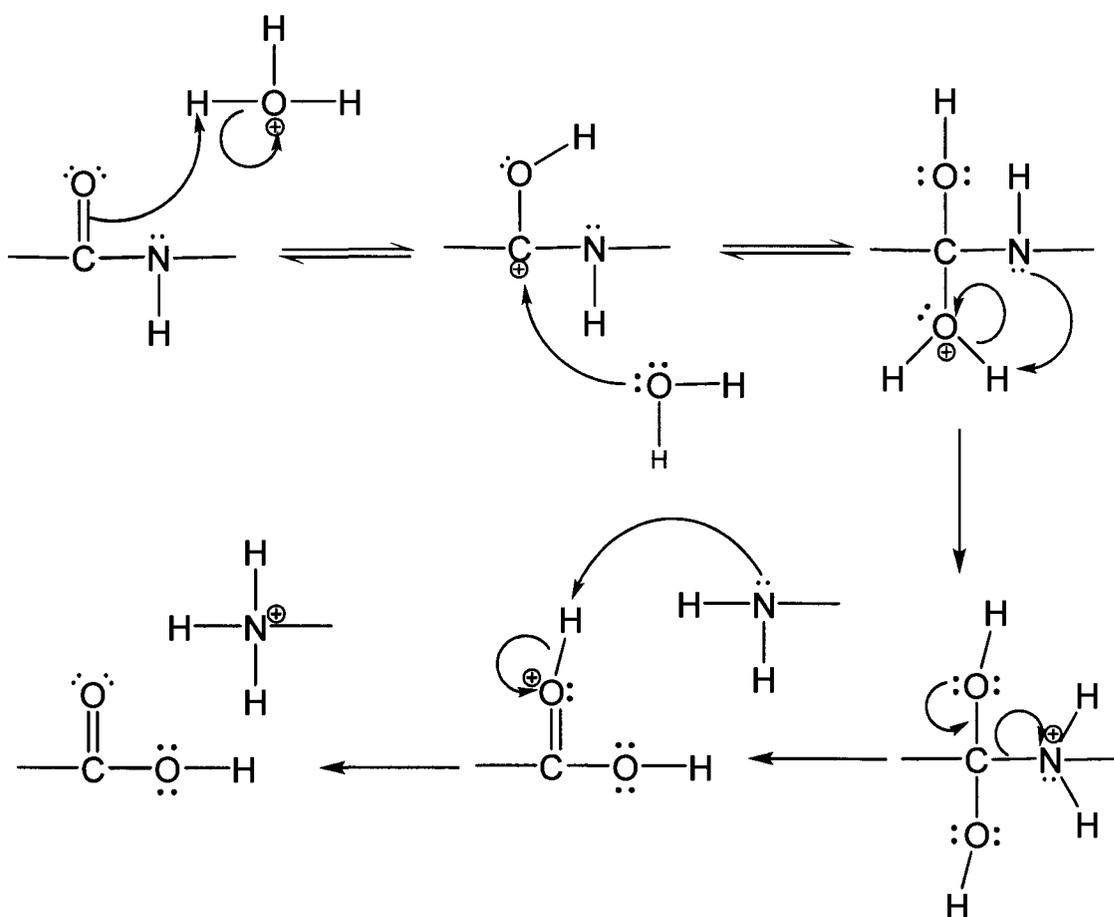
Figure 2.3.2 Possible reaction mechanism of the hydrolysis process of protein amide linkage in strong acid and microwave irradiation

2.3.2. Microwave-assisted acid hydrolysis of Horse Heart Cytochrome C

Horse Heart Cytochrome C, or CYC_HORSE, is a small, highly soluble heme protein (100 gL^{-1} in water), found loosely associated with the inner membrane of the mitochondrion. It is an essential component of the electron transport chain. The mature form of CYC_HORSE consists of 104 amino acids which, according to Qi et al[5] (PDB id: 2FRC), at pH 5.7 are organized in 4 helices containing 41 residues (39%) and 2 strands containing 2 residues (1%). The primary and secondary structure of the horse heart cytochrome C is shown in Figure 2.3.3.


```
        10            20            30            40            50
   ⌐ᴧᴧᴧᴧᴧᴧᴧᴧᴧᴧᴧᴧᴧᴧᴧ┐─────────────────────────────────►
    GDVEKGKKI  FVQKCAQCHT  VEKGGKHKTG  PNLHGLFGRK  TGQAPGFTYT

        60            70            80            90            100
   ᴧᴧᴧ────►───ᴧᴧᴧᴧᴧᴧ─────────────────────ᴧᴧᴧᴧᴧᴧᴧᴧᴧᴧᴧᴧ
    DANKNKGITW  KEETLMEYLE  NPKKYIPGTK  MIFAGIKKKT  EREDLIAYLK  KATNE

   ᴧᴧᴧ α-helix          ━━► β-sheet          ───── random coil
```
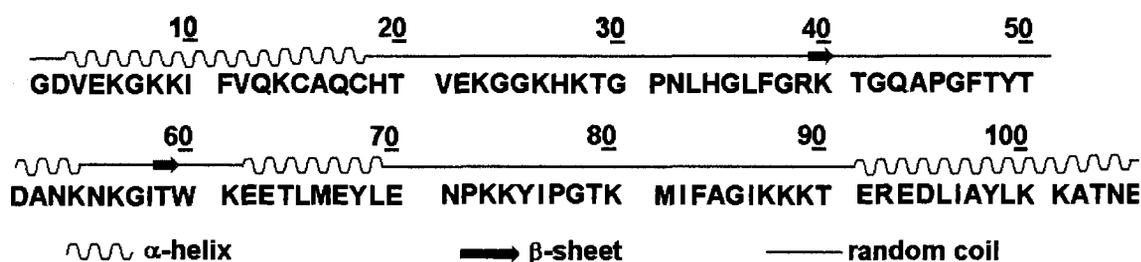
Figure 2.3.3 Primary and secondary structure of CYC_HORSE[5]

The MAAH of CYC_HORSE has been previously published[1]. It was shown that by a controlled hydrolysis the sequencing of the protein along with the detection of post-translational modifications such as removal of the first methionine, N-acetylation of the first glycine residue and the binding of the heme to the protein is possible. In this work, we generated similar results to those reported by Zhong et al [1] using MAAH combined with the newly introduced MALDI TOF/TOF system. It was found that this instrument provides better resolution and sensitivity than the one used by Zhong et al and, as a consequence, peptide ladders can be more easily detected using the ABI MALDI TOF/TOF MS.

Figure 2.3.4(a) shows the mass spectrum of CYC_HORSE in 3M HCl and 1 minute irradiation time. The mass spectrum consists of peaks corresponding to the N- and C-terminal ladders. From the mass spectrometric analysis of the polypeptide ladders the sequence of the protein can be easily determined (Figure 2.3.4(b)). A long stretch of sequence can be deduced from these two ladders which also overlap. The mass

difference between the molecular ion peak and the next peak is 99 Da which matches the mass of N-Acetylated glycine. This also shows the cleavage of the first methionine, a common post translational process[6]. Another modification of the CYC_HORSE is the binding of the heme. This can be detected in a mass shift for both the N- and C-terminal ladders the peaks containing Cys14 and Cys17, to which the heme group is covalently bonded. No cleavage sites were observed between Cys14 and Cys17 which shows a covalently bonded heme that is not affected by the hydrolysis process.

In the case of a protein with a known sequence from the mass difference between the molecular ion peak detected by MALDI-TOF-MS and the theoretical mass of the entire protein sequence modification can be detected. In this case the difference matches up with the mass of the heme, confirming the presence of the heme group. By performing MAAH and reading the sequence from the mass spectrum one can detect the exact location of the modification as explained above. Figure 2.3.5 shows the peptide bond hydrolysis map of the CYC_HORSE.
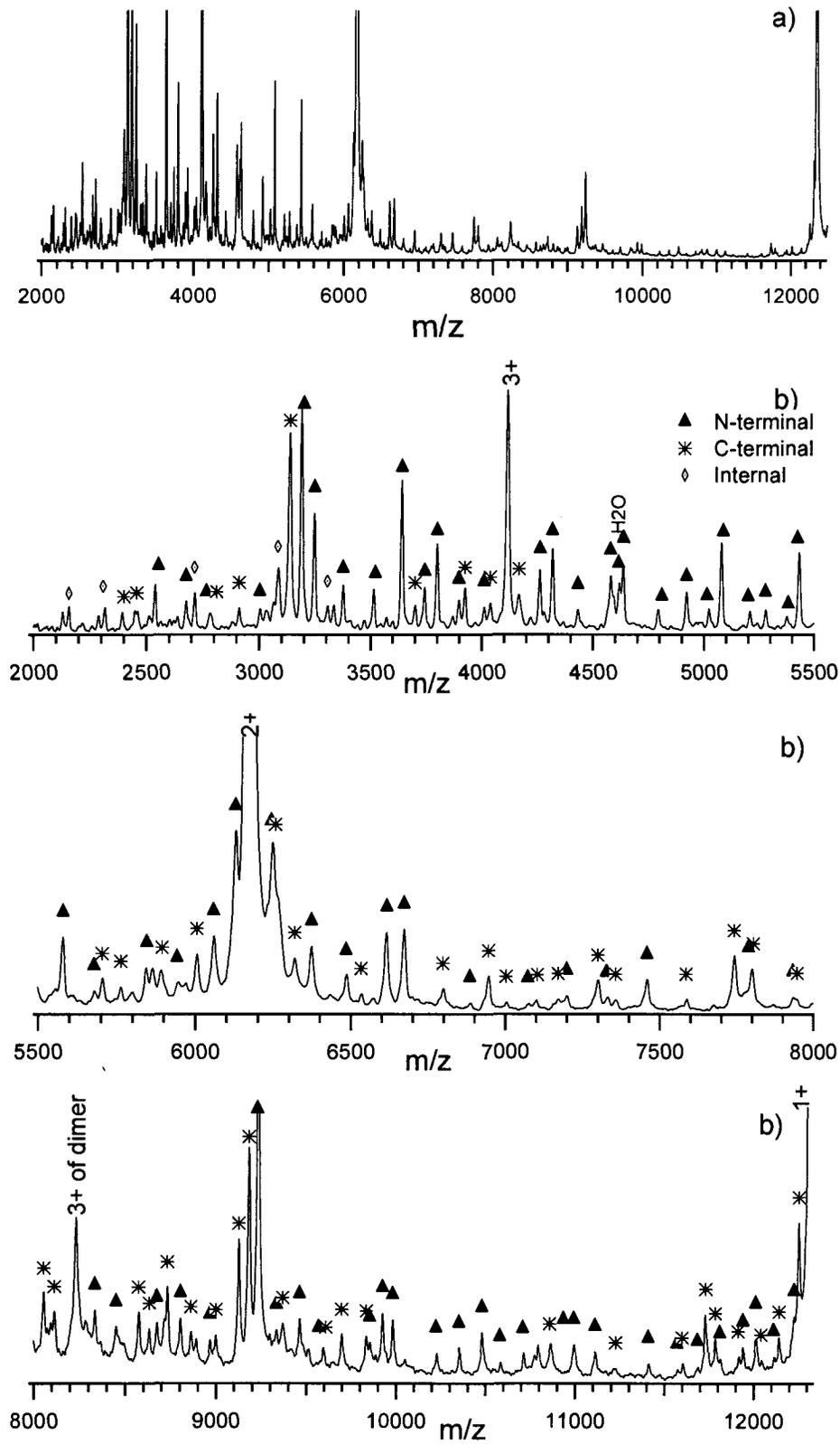
Figure 2.3.4 (a). MALDI spectra of Horse Heart Cytochrome C after MAAH in 3 M HCl and 1 minute irradiation time. (b) Expanded MALDI spectra of Heart Cytochrome C hydrolysate after 1 minute irradiation time in 3M HCl

```
         10         20              30        40           50
GDVEKGKKI  FVQKCAQCHT  VEKGGKHKTG  PNLHGLFGRK  TGQAPGFTYT

          60         70           80         90        100
DANKNKGITW  KEETLMEYLE  NPKKYIPGTK  MIFAGIKKKT  EREDLIAYLK  KATNE
```
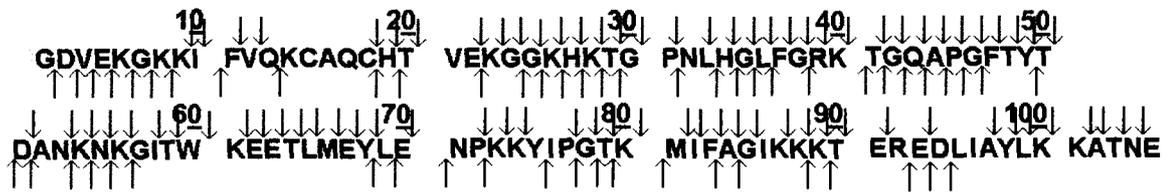
Figure 2.3.5 Peptide bond hydrolysis map of Horse Heart Cytochrome C in 3M HCl and 1 minute irradiation time. ↓ - hydrolysis site identified by the detection of the N-terminal peptide; ↑ - hydrolysis site identified by the detection of the C-terminal peptide. N-terminal purification tag shown in green.

Circular dichroism and dynamic light scattering were used to monitor possible changes in the secondary structure and to determine whether the protein is in a monomeric or oligomeric state under the hydrolysis condition. Figure 2.3.6 shows the obtained experimental CD spectra for CYC_HORSE in 50mM sodium acetate, pH 5.5, at 20°C and after mixing with 3M HCl. In 50 mM NaAc, the spectrum shows intense signals at both 222nm and 208 nm showing that the protein contains a high percentage of α-helix. The deconvolution of the CD spectra resulted in percentages that were significantly different from previously reported numbers, determined by NMR structural studies (PDB id: 2GIW[7]). The differences could be caused by the fact that the NMR structure has been determined at a pH 7.0 and our CD data has been recorded at pH 5.5. Another reason could be that, in general, to date, CD programs are not able to accurately estimate the β-sheet content. This problem becomes even more accentuated when the studied protein contains a low or no sheet content. In our case, it seems that the program is overestimating the β-sheet content and concomitantly underestimating the helical and random coil contents.

After mixing with 3M HCl at 20°C the spectra shows a decrease in the signal intensity at around 222nm but the increase at around 208nm suggests that the protein maintains some of its helical structure. Deconvolution results for CYC_HORSE in 50 mM NaAc and 3M HCl are shown in Table 2.3.1 and Table 2.3.2.
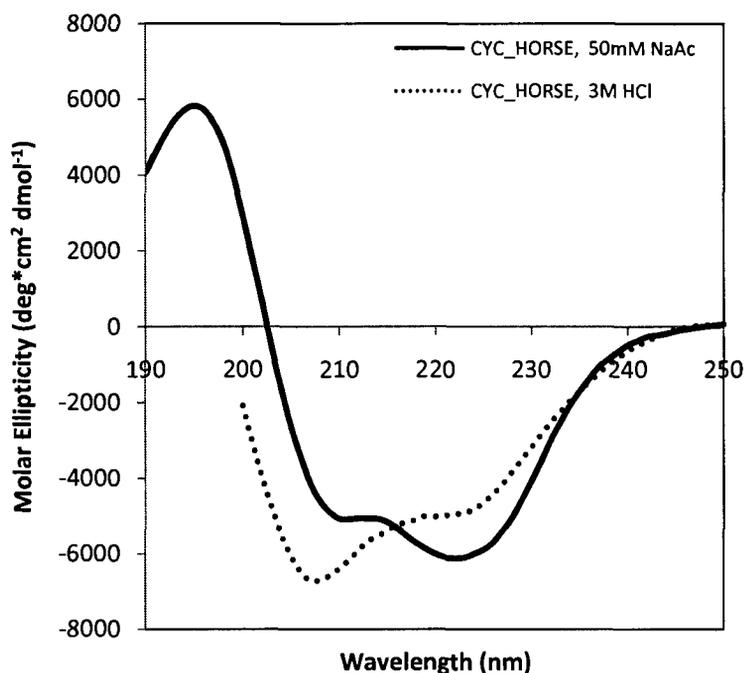
Figure 2.3.6 CD spectra of CYC_HORSE recorded at 20°C, pH 5.5 and in 3M HCl.

Table 2.3.1 Deconvolution results of the CD spectra for CYC_HORSE in 50 mM NaAc

NMR	pH = 7.0	36	0	15	49
Program	Ref. Set	Helix (%)	Sheet (%)	Turn (%)	Random (%)
	SP43	16	31	23	30
CDSSTR	SDP48	16	27	18	38
	SMP56	8	38	24	30
	SP43	20	27	20	33
CONTINLL	SDP48	18	26	17	39
	SMP56	20	27	21	33
	SP43	12	41	22	23
SELCON3	SDP48	12	41	22	23
	SMP56	11	40	23	24
Average±SD		15±4	33±7	21±2	30±6

Table 2.3.2 Deconvolution results of the CD spectra for CYC_HORSE in 3M HCl

Program	Ref. Set	Helix (%)	Sheet (%)	Turn (%)	Random (%)
	SP43	15	30	21	30
CDSSTR	SDP48	15	26	17	40
	SMP56	9	31	26	30
	SP43	18	29	21	33
CONTINLL	SDP48	16	24	16	38
	SMP56	18	28	22	33
	SP43	13	30	21	22
SELCON3	SDP48	12	29	20	22
	SMP56	13	30	21	24
Average		14±3	28±2	20±3	30±7

Figure 2.3.7 shows the light scattering intensities for CYC_HORSE in 50mM NaAc, pH 5.5 and after mixing with 6M HCl. At pH 5.5 a hydrodynamic radius of 1.7nm was determined which gives a predicted molecular weight of 12.1 kDa. This is very close to the expected molecular weight of 12.3 kDa and shows that the protein exists as a pure monomer. Even though the buffer was filtered using a 0.22 μm filter some contamination in the sample was detected, as shown in Figure 2.3.7 (c). Although these contamination peaks were observed in the size distribution by intensity, they were not present in the distribution by volume which shows that these particles are in very low concentration in the sample and they are contaminants from the buffer. Also after preparation of the stock solutions the samples were not filtered to make sure no large molecular weight species were filtered out.

After mixing with 6M HCl a mean hydrodynamic radius of 2.7 nm was determined. Considering that the protein maintains a spherical structure this gives a predicted molecular weight of 34.3 kDa, which could indicate that the sample is oligomerizing. It has to be noted that a relatively high polydispersity of 29% for this measurement has been observed. This indicates that the sample contains either oligomers of different sizes; it is a mixture of properly folded molecules and partially or fully unfolded conformers or it is a mixture of both oligomers and partially or fully folded/unfolded molecules. If a molecule unfolds, the measured hydrodynamic radius measured for the molecule will be higher than the radius in the folded state. Therefore using the formula shown in Eq. 11 (Chapter 1), which considers all particles as being spherical would result in an overestimation of the MW of the protein. The mean hydrodynamic radii of the DLS experiments are shown in Table 2.3.3.

Taken together, the results of the CD data which shows that the protein partially retains its helical structure with the results of the DLS experiments, it can be concluded that under highly acidic conditions CYC_HORSE exists as a mixture of partially or fully folded/unfolded conformers. It is also possible that in the presence of the microwave irradiation accompanied by the heating of the sample the protein becomes fully unfolded, making all the peptide bonds susceptible to the hydrolysis process. The data

also showed that only the peptides bonds were affected by the hydrolysis process and the covalently bound heme group remained attached to the backbone making possible the detection and localization of this post-translational modification.

Table 2.3.3 Hydrodynamic radii determined by dynamic light scattering for CYC_HORSE under different solvent conditions

Cytochrome C	R_H (nm)	Molecular Weight (kDa)	
		Expected	Estimated
50mM NaAc, pH 5.5	1.7	12.3	12.1
3M HCl	2.7	12.3	34.3
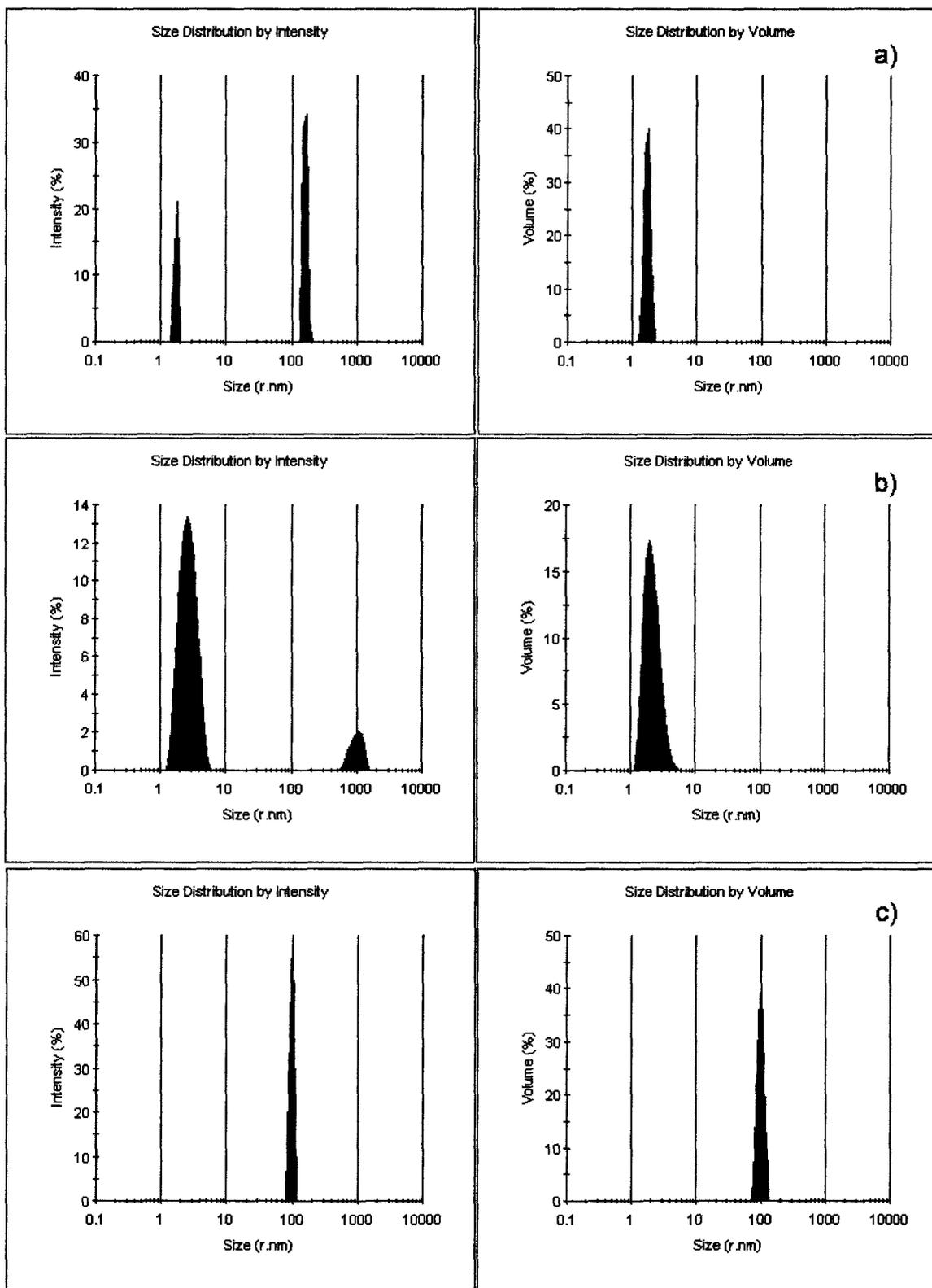
Figure 2.3.7 DLS spectra of: a) CYC_HORSE in 50mM NaAc, pH 5.5, b) CYC_HORSE in 3M HCl, c) 50mM NaAc, pH 5.5

2.3.3. Microwave-assisted acid hydrolysis of Hen Egg Whyte Lysozyme

Hen Egg White Lysozyme (HEWL) is a 14.3 kDa protein containing 129 amino acid residues. The protein contains 5 helices formed by 44 residues and 9 strands formed by 15 residues. The structure of the protein is stabilized by 4 different disulphide bonds linking all 8 cysteine residues of the protein. The primary and secondary structure of the protein with the 4 disulphide bonds can be seen in Figure 2.3.8.

```
           10              20              30              40              50
  ────────────────────  ──────────────────────────────────────────────────
  KVFGRCELAA   AMKRHGLDNY   RGYSLGNWVC   AAKFESNFNT   QATNRNTDGS

           60              70              80              90             100
  ──────────────────────────────────────────────────────────────────────
  TD YGILQINS   RWWCNDGRTP   GSRNLCNIPC   SALLSSDITA   SVNCAKKIVS

          110             120
  ──────────────────────────────────
  DGNGMNAWVA   WRNRCKGTDV   QAWIRGCRL

  ∿∿∿ α-helix          ➡ β-sheet          ─── random coil
```
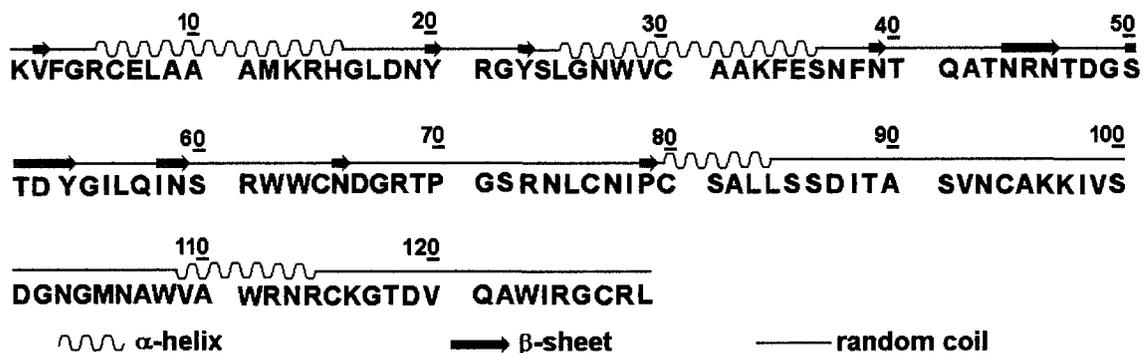
Figure 2.3.8 Primary and secondary structure of HEWL along with the 4 disulphide bonds[8]

The MAP sequencing of HEWL using MAAH followed by MALDI-TOF-MS has been previously reported. The protocol published by Zhong et al.[1] involved the reduction of the disulphide bonds prior to the microwave-assisted acid hydrolysis. The best signal-to-noise ratio and sequence coverage were obtained by a 15 hour reduction of the disulphide bond using DTT as a reduction agent followed by MAAH. To speed up this process and to determine how important is the reduction of the disulphide bond for the succes of the MAAH, we examined the effectivenes of the method without prior reduction of the disulphide bonds. Figure 2.3.9 shows the mass spectrum of HEWL after MAAH in 3M HCl and 1 minute irradiation time. The MAAH was performed without prior reduction of the disulphide bonds and also no reducing agent was present in the mixture during the MAAH process.
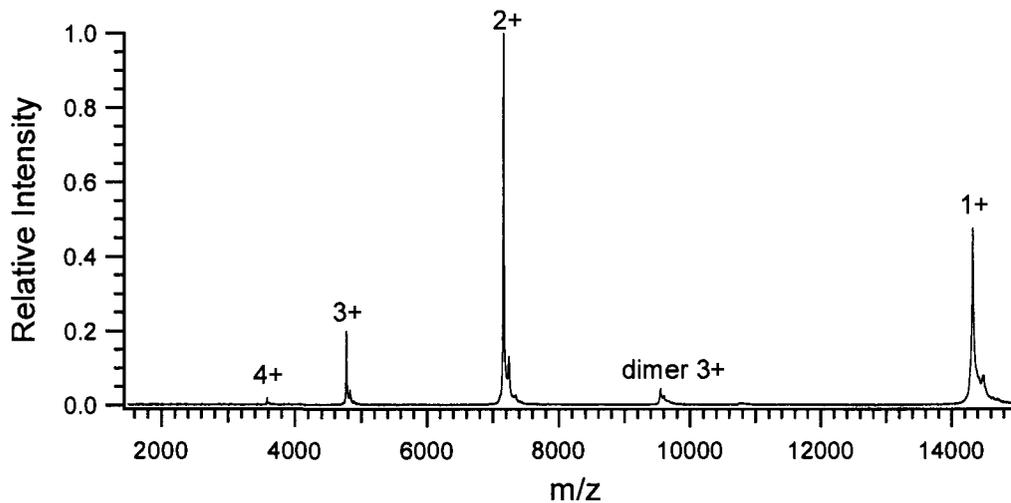
Figure 2.3.9 MALDI-TOF-MS spectra of HEWL after hydrolysis in 3M HCl and 1 minute of irradiation time.

The mass spectrum of the oxidized form of the HEWL shows only peaks corresponding to different charged states of the molecular ion peak and the triply charge dimer. No peaks corresponding to any hydrolysis product were observed. Previous reports on the structure of the HEWL suggest that the protein contains 4 different disulphide bonds, one of which connects Cys6 to Cys127. Even if hydrolysis of a peptide bond located between these two Cys residues occurs during the MAAH process the two peptide fragments corresponding to the N- and C-terminal fragments will be linked by this disulphide bonds and as a result the peaks of the hydrolyzed protein will show up as a single peak, with a molecular weight which would be 18 Da higher than the molecular weight ion. The other 3 disulphide bonds also link different parts of the protein which will further decrease the possibility of the detection of hydrolysate peaks when hydrolysis of peptide bonds occur between linked Cys residues.

If the hydrolysis of the peptide bond occurs between Arg5 and Cys6 the highest molecular weight fragment could be KVFGR with a MW of 606.37 Da. If the hydrolysis of the peptide bond occurs between Cys127 and Leu128 the resulted fragment will be LR having a mass 288.20 Da. These masses are hard to detect using MALDI-TOF-MS because of the presence of matrix peaks at low mass in MALDI-MS.

The absence of other peaks in the spectrum shows that when a controlled

hydrolysis is performed hydrolysis of the peptide bond occurs ones per molecule. The fact that no other peaks were detected also demonstrates that by performing a controlled hydrolysis of the peptide bond, such as is the case of the MAAH in 3M HCl and an optimum irradiation time of 1 min, no or very few internal fragments will be generated.

Figure 2.3.10 shows the mass spectrum of HEWL after MAAH in 3M HCl and 1 minute irradiation time, followed by the reduction of the disulphide bond. The mass spectrometric analysis shows peaks corresponding to the N- and C-terminal ladder. The first amino acid of the N-terminal polypeptide ladder is Lys19 which indicates that the first 18 amino acid residues, the so called signalling peptide, has been removed after the expression and transport of the protein to the organelles. This demonstrates the fact that MAAH can be used to detect and identify post-translational modification of proteins.
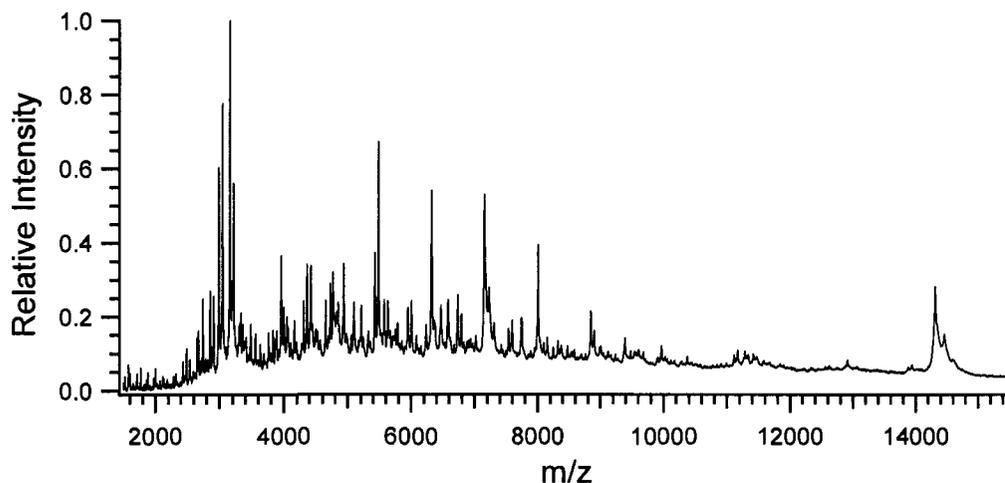
Figure 2.3.10 MALDI-TOF-MS spectra of HEWL after hydrolysis in 3M HCl and 1 minute of irradiation time followed by reduction with DTT.

Figure 2.3.11 represents the peptide bond hydrolysis map of the HEWL after MAAH in 3M HCl and 1 minute irradiation time followed by the reduction of the disulphide bond. The MAP shows high sequence coverage, having an overlap of the N- and C- terminal ladders.

Figure 2.3.11 Peptide bond hydrolysis map of HEWL in 3M HCl and 1 minute irradiation time followed by DTT reduction of the disulphide bonds. ↓ - hydrolysis site identified by the detection of the N-terminal peptide; ↑ - hydrolysis site identified by the detection of the C-terminal peptide. N-terminal purification tag shown in green.

In order to have a better understanding of the MAAH and the importance of disulphides in the MAAH process, we examined the structural changes of HEWL under similar conditions to those of the MAAH conditions. The secondary structure for the Hen Egg White Lysozyme determined by X-ray crystallography[8] (crystallization performed at pH 4.7, PDB 2VB1) was used as reference. CD experiments were performed on HEWL in 50mM acetate buffer, pH 5.5 at 20°C and after mixing with 6M HCl.

The resulting experimental CD data, shown in Figure 2.3.12, and the reference data are in a relatively good agreement. In 50 mM NaAc, pH 5.5, the CD spectra shows two peaks, one at 222 nm and on at 208 nm which shows that the protein contains a high percentage of α-helix organized in multiple helices. The results of the deconvolution of the spectrum are shown in Table 2.3.4.

After mixing with the 6M HCl an increase in signal at around 208 nm was observed which suggest an increase in the helical content of the protein. The results obtained by the deconvolution of the circular dichroism spectra, summarized in Table 2.3.5, confirm the slight increase in the helical content of the protein. The results of the CD experiments show that HEWL is stable under the studied conditions. No significant differences in the secondary structure can be observed between pH 5.5 and in 3M HCl.

Figure 2.3.12 CD spectra of HEWL recorded at 20°C, pH 5.5 and in 3M HCl

Table 2.3.4 Deconvolution results of the CD spectra for HEWL, 50 mM NaAc, pH 5.5

X-ray (pH 4.7)		40	10	24	26
Program	Ref. Set	Helix (%)	Sheet (%)	Turn (%)	Random (%)
CDSSTR	SP43	46	11	20	24
	SDP48	45	9	20	25
	SMP56	47	9	19	25
CONTINLL	SP43	44	11	19	26
	SDP48	44	10	18	28
	SMP56	44	10	20	27
SELCON3	SP43	42	13	20	26
	SDP48	42	12	20	26
	SMP56	42	13	20	26
Average±SD		44±2	11±2	20±1	26±1

Table 2.3.5 Deconvolution results of the CD spectra for HEWL in 3M HCl.

Program	Ref. Set	Helix (%)	Sheet (%)	Turn (%)	Random (%)
CDSSTR	SP43	56	10	15	19
	SDP48	56	10	14	19
	SMP56	56	9	14	22
CONTINLL	SP43	50	9	21	21
	SDP48	49	4	16	32
	SMP56	49	9	19	23
SELCON3	SP43	49	9	18	24
	SDP48	48	6	17	31
	SMP56	48	9	19	25
Average±SD		51±4	8±2	17±5	24±5

The DLS spectra of HEWL showed a peak of significant height. Figure 2.3.13 shows the light scattering intensity trace for HEWL. The DLS recorded in 50mM NaAc, pH 5.5, showed a hydrodynamic radius of 1.8nm corresponding to an estimated molecular weight of 13.3kDa which is very close to the theoretical value of 1.9nm hydrodynamic radius and 13.3kDa molecular weight, indicating that the peak is corresponding to a pure lysozyme monomer. Under acidic conditions, 3M HCl, the peak showed a hydrodynamic radius of 1.7nm, which is a clear indication that the lysozyme stays monomeric or stays compact under these conditions. The hydrodynamic radii of HEWL under the different solvent conditions are summarized in Table 2.3.6.

Table 2.3.6 Hydrodynamic radii determined by dynamic light scattering for HEWL under different solvent conditions

Lysozyme	R_H (nm)	Molecular Weight (kDa)	
		Expected	Estimated
50mM NaAc, pH 5.5	1.8	14.3	13.3
3M HCl	1.7	14.3	10.9
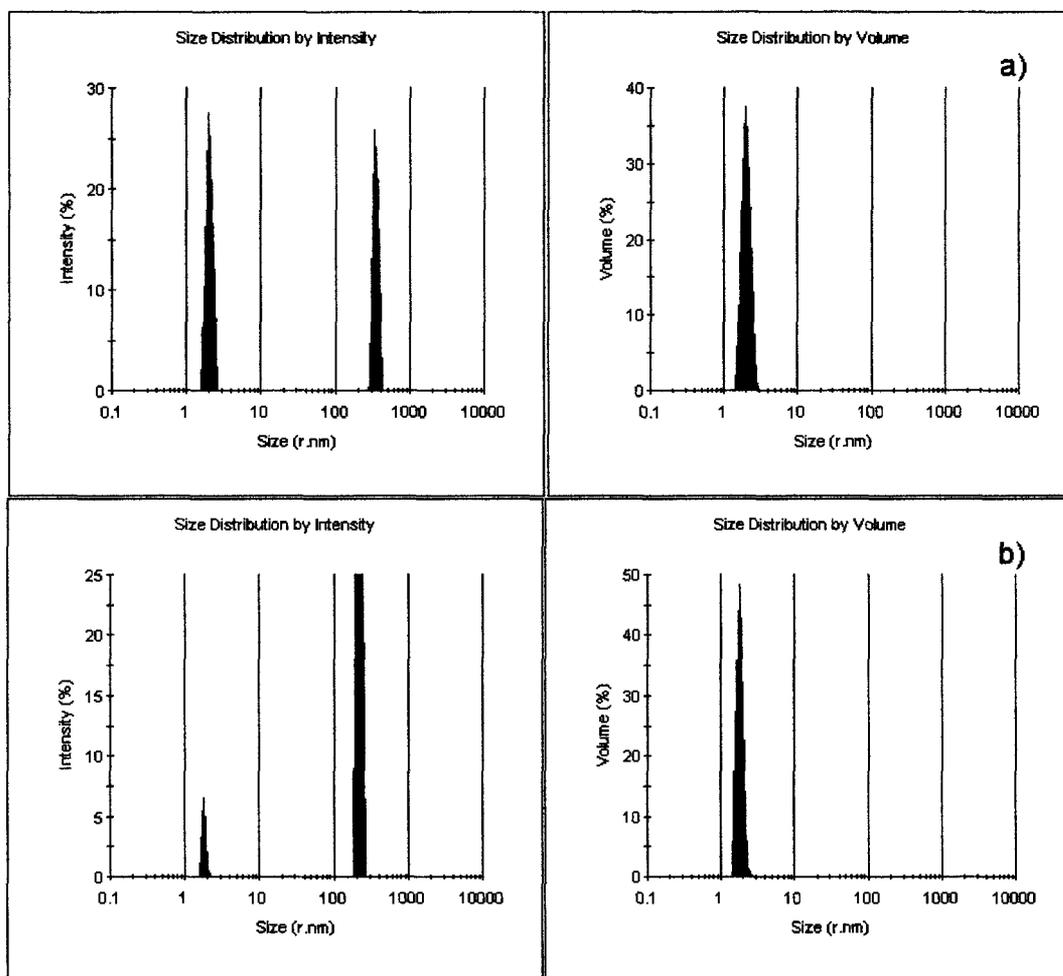
Figure 2.3.13 Light scattering trace of HEWL in (a) 50mM NaAc; (b) 3M HCl.

The CD data presented above for HEWL show that by mixing the protein dissolved in 50mM NaAC, pH 5.5, with 6M HCl does not alter the overall secondary structure of the protein. The DLS data showed that both at pH 5.5 and under highly acidic conditions the protein is mostly monomeric or compact. The mass spectrometric analysis of the hydrolysate resulted by MAAH under optimum conditions did not show a dependency between protein secondary structure and the hydrolysis process; peaks corresponding to hydrolyzed peptide bonds from α-helical, β-sheet, turn and random coil regions in the original lysozyme have all been detected, even in the presence of the stabilizing disulphide bonds. Unless severe changes in the secondary structure of the HEWL occurred because of the microwave irradiation, the data shows that hydrolysis of the peptide bond is possible as long as it is solvent accessible.

2.3.4. *Microwave-Assisted Base Hydrolysis*

The peptide bond is a less reactive bond and therefore the hydrolysis of the bond can be done under highly acidic or highly basic conditions. The acid hydrolysis of the peptide bond can be done using a concentrated acid such as HCl or TFA[1, 9]. It was already shown that the hydrolysis can be accelerated in the presence of microwave irradiation. As shown above, the MAAH method is capable of hydrolyzing different proteins having different structures and post translational modifications as long as the peptide bond is solvent accessible.

Here we examined the possibility of using microwave-assisted base hydrolysis as a mean of degrading proteins in order to be analyzed by MALDI-TOF-MS. Table 2.3.7 shows the list of standard proteins selected to examine the possibility of using MABH as a means of degrading proteins for mass spectrometric analysis.

Table 2.3.7. List of standard proteins examined for MABH.

Protein	ID (Swiss-Prot)	Average MW	pI	PTM's
Lysozyme – chicken	P00698	14313.1 Da	9.32	Signal peptide 1-18 cleaved ; 4 disulphide bonds
Myoglobin – horse	P02188	16951.5 Da	7.36	Initial M removed
Cytochrome C –horse	P00004	11701.5 Da	9.59	Acetylation at G2; heme group; Initial M removed

Figure 2.3.14 shows a possible mechanism of the microwave–assisted base hydrolysis process. Generally, the base hydrolysis process involves the breakage of the acyl–nitrogen bond through a tetrahedral intermediate. Hydrolysis starts with the rate determining nucleophilic attack[10] of a hydroxide ion on the carbon atom of the amide group, resulting in the formation of the hydroxide intermediate I_1. This intermediate is negatively charged and exhibits basic character, therefore it may induce dissociation of a solvent water molecule, which can occur through protonation of the nitrogen atom, yielding the intermediate I_2 and a solvated OH⁻ ion. In principle, breaking of the C–N bond in I_2 is possible with the formation of the final products I_4 and I_5. However, according to recent simulation studies[11] dissociation through a further sub-step, in which the OH⁻ group of I_2 is deprotonated by a hydroxide ion, was found to be strongly

favored. This reaction yields the intermediate I_3, which allows the opening of the C–N bond at much lower energetic costs. The final step of amide hydrolysis would be the dissociation of I_3 into the amide I_4 and the carboxylic acid I_5.
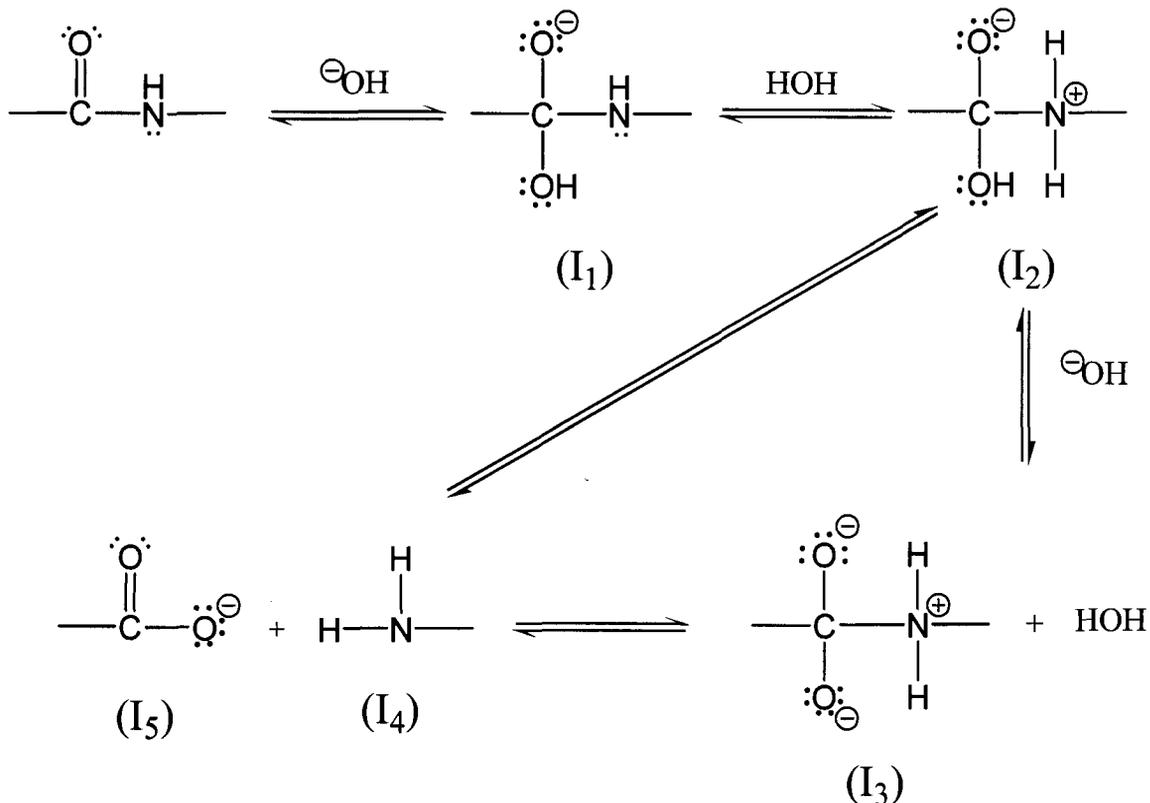
Figure 2.3.14 Possible reaction mechanism of the hydrolysis process of protein amide linkage in strong base and microwave irradiation

2.3.5. Microwave-assisted base hydrolysis of an amide bond.

The first protein tested was HEWL. We prepared a stock solution of 1 mg mL^{-1} and hydrolyzed it using NaOH. We have examined both the effect of concentration and irradiation time on the hydrolysis process. The tested concentrations of NaOH before mixing 1 to 1 (v/v) with the protein sample were 1, 2, 4 and 6M. Irradiation was performed for 30 seconds, 1, 2, and 3 minutes. It was found that microwave irradiation of the sample in 3M NaOH (6M initial concentration) was degrading the 1.5 mL Eppendorf vials. Therefore only initial concentrations of 1M, 2 M and 4 M NaOH were tested.

For 30 seconds irradiation time, we found MS spectra with good signal-to-noise ratios were obtained for all concentration ranges, as shown in Figure 2.3.15. The mass spectrometric analysis showed peaks corresponding to both N- and C-terminal ladders and the intact protein peak. Signal-to-noise ratios when mixing with 1M and 2 M NaOH are very similar, although a better sequence coverage has been obtained when mixing with 2 M NaOH. Increasing the concentration to 4 M did not improve the sequence coverage and resulted in peaks with similar or slightly lower signal-to-noise ratios compared to 2 M NaOH. As irradiation time was increased to 1 minute the best signal-to noise ratio was obtained when mixing with 1M NaOH. The spectrum consisted of N- and C-terminal peaks and low intensity peak of the intact protein. For 2 M and 4 M NaOH as initial concentration peaks with good signal-to-noise ratios were detected in the low mass region but very low intensity peaks or no peaks have been observed for the higher mass region, as shown in Figure 2.3.16. Even though good signal-to-noise have been obtained for 1 minute irradiation and 1M NaOH initial concentration the sequence coverage was lower than for 30 seconds irradiation and 2 M NaOH initial concentration. A further increase of the irradiation time to 2 minutes resulted in very low intensity or disappearance of high molecular weight and intact protein peaks and the appearance of internal fragment peaks. As all the intact protein was consumed after 2 and 3 minutes irradiation times, the internal peptide peaks become the dominant feature of the spectra. A long exposure to the microwave radiation will eventually result in a mixture of amino acids. The spectra for 2 and 3 minutes irradiation times are shown in Figure 2.3.17 and Figure 2.3.18, respectively. The optimal conditions for the MABH of HEWL were found to be 30 seconds irradiation time with 1 M NaOH as the final concentration. The expanded MALDI spectrum for this condition is showed in Figure 2.3.19. The corresponding hydrolysis map is shown in Figure 2.3.20.
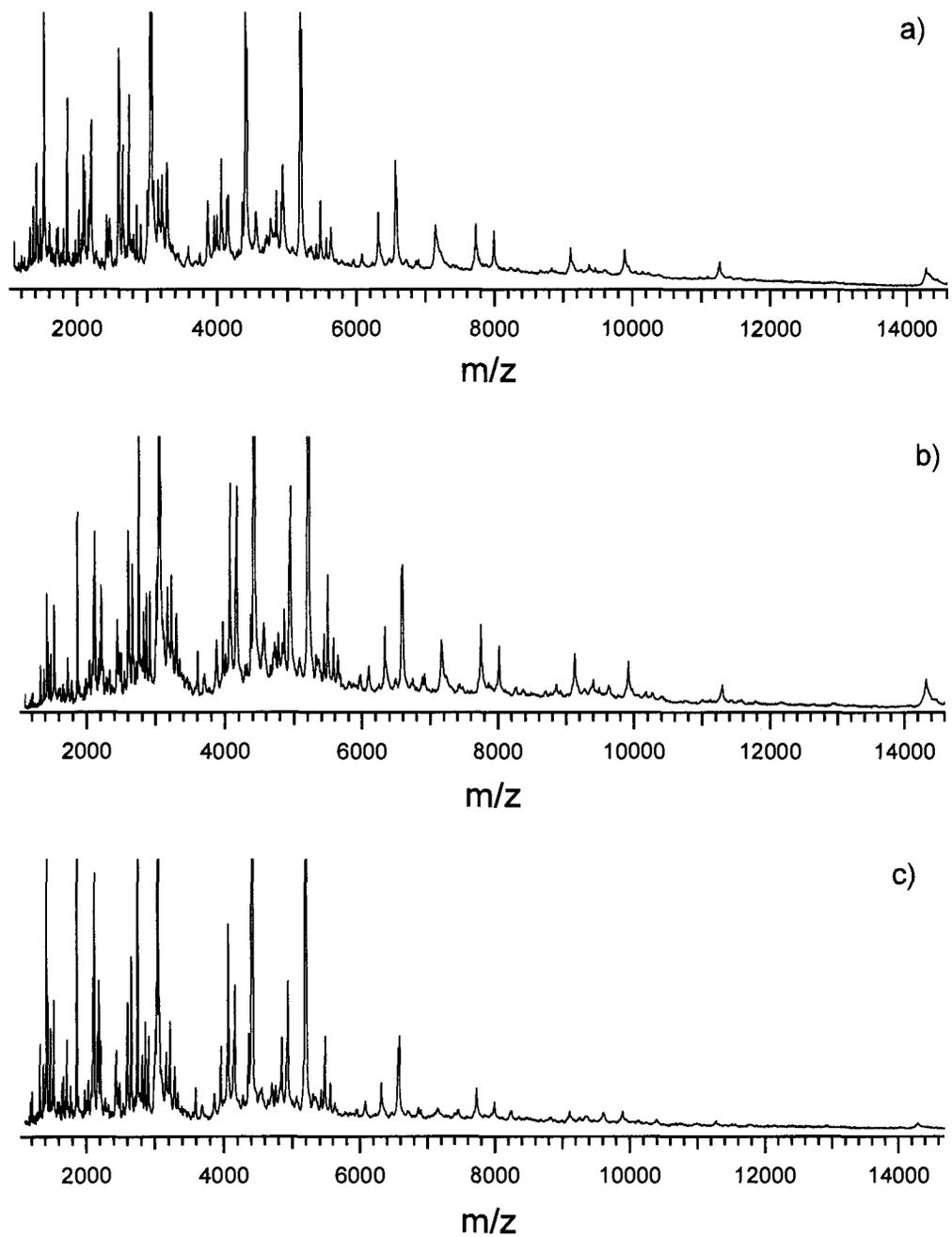
Figure 2.3.15 MALDI-TOF MS spectra of the hydrolysates of HEWL after 30 seconds irradiation time and mixing with: a) 1M NaOH, b) 2 M NaOH, c) 4 M NaOH
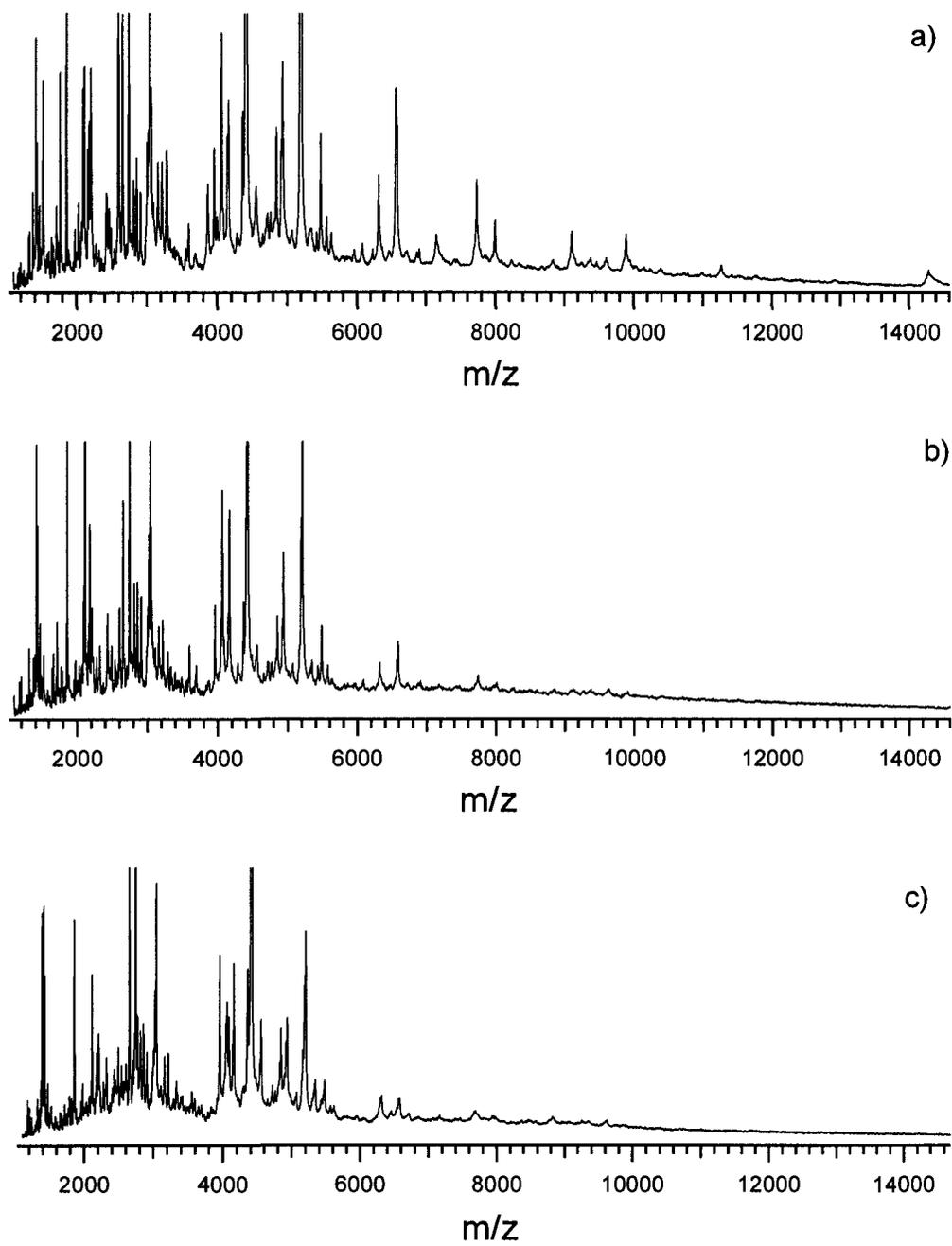
Figure 2.3.16 MALDI-TOF MS spectra of the hydrolysates after 1 minute irradiation
time and mixing with: a) 1M NaOH, b) 2 M NaOH, c) 4 M NaOH
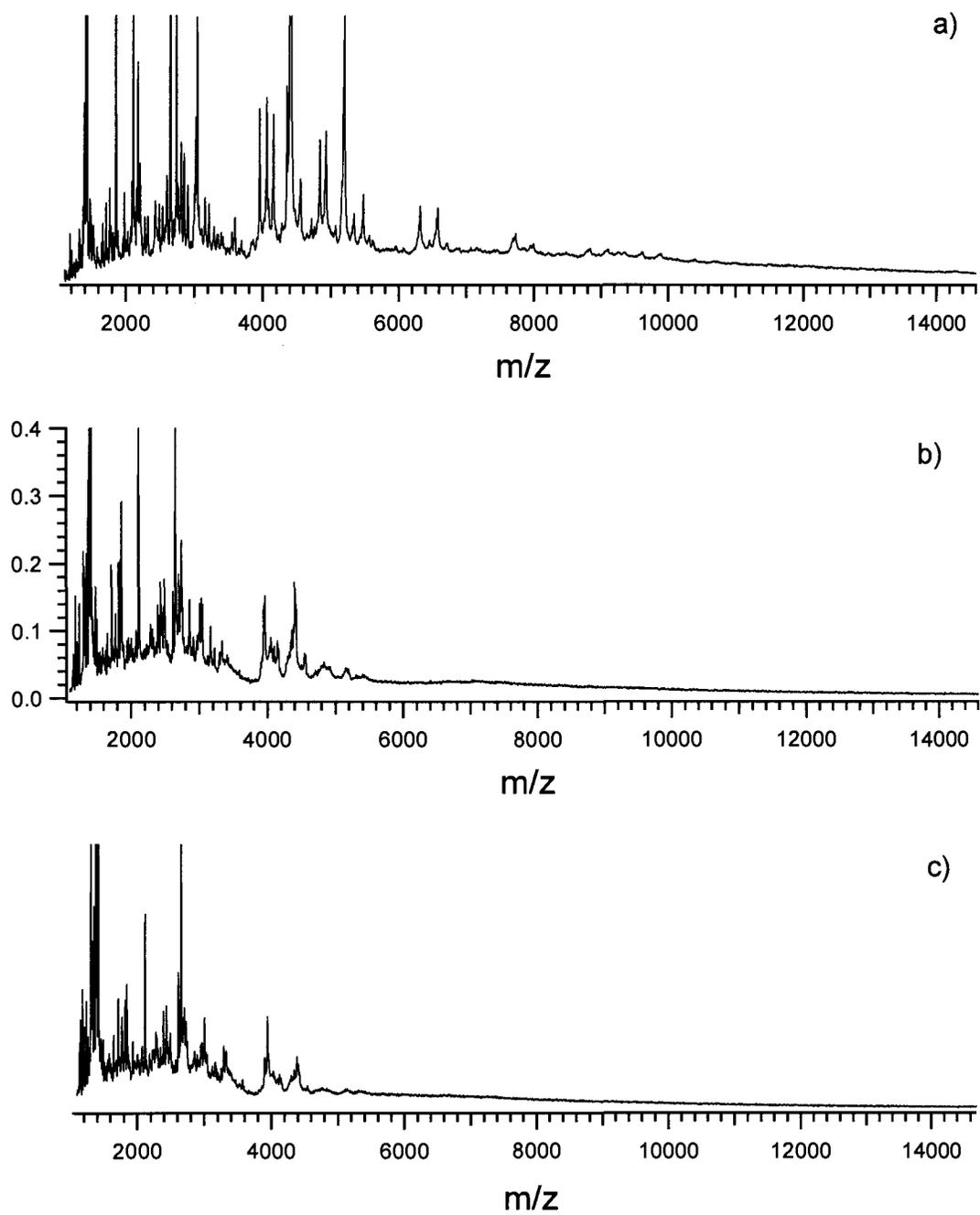
Figure 2.3.17 MALDI-TOF MS spectra of the hydrolysates after 2 minute irradiation
time and mixing with: a) 1M NaOH, b) 2 M NaOH, c) 4 M NaOH
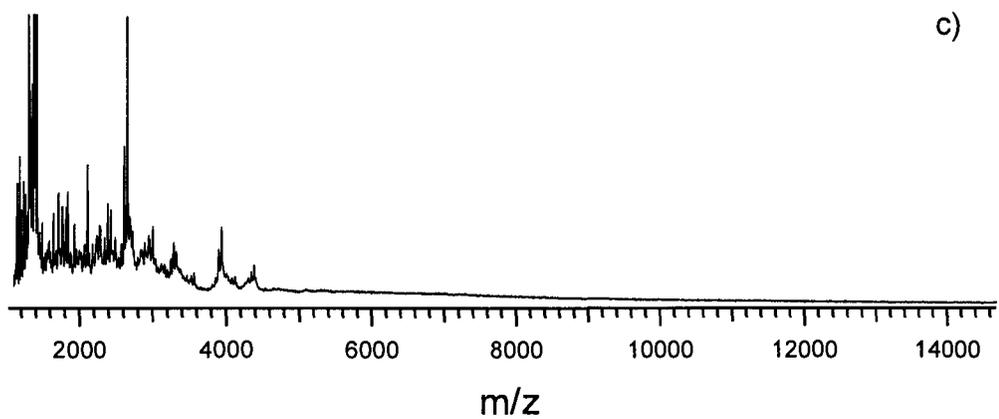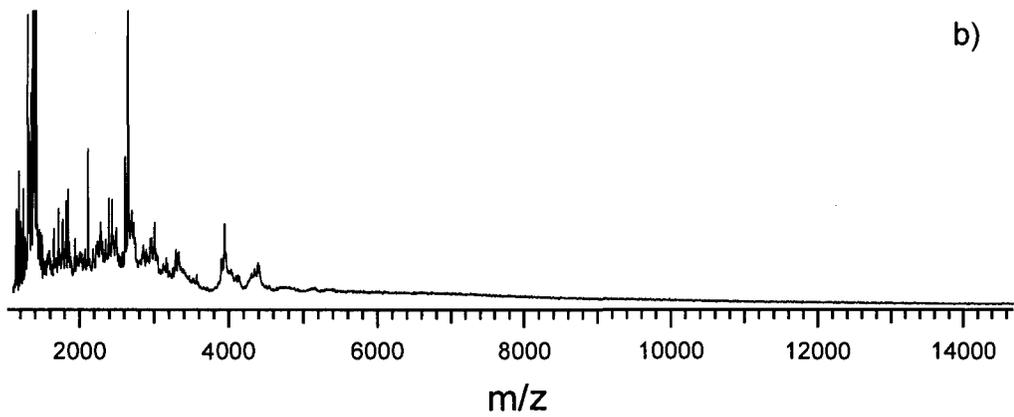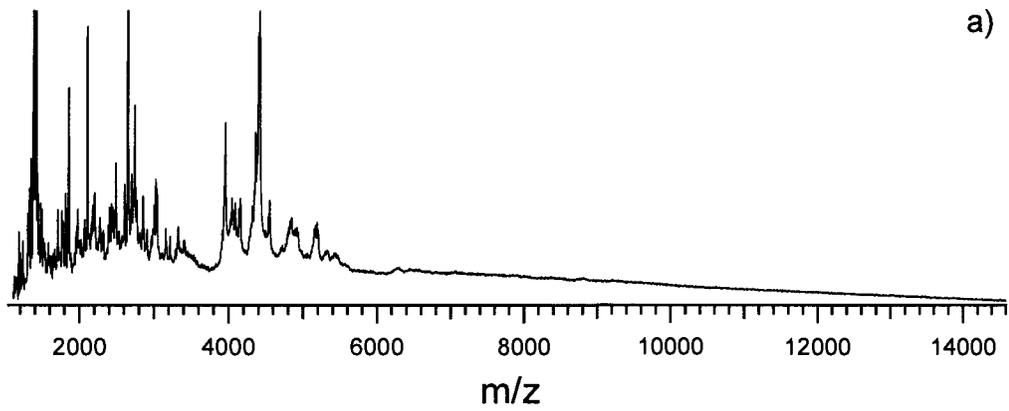
Figure 2.3.18 MALDI-TOF MS spectra of the hydrolysates after 3 minute irradiation
time and mixing with: a) 1M NaOH, b) 2 M NaOH, c) 4 M NaOH
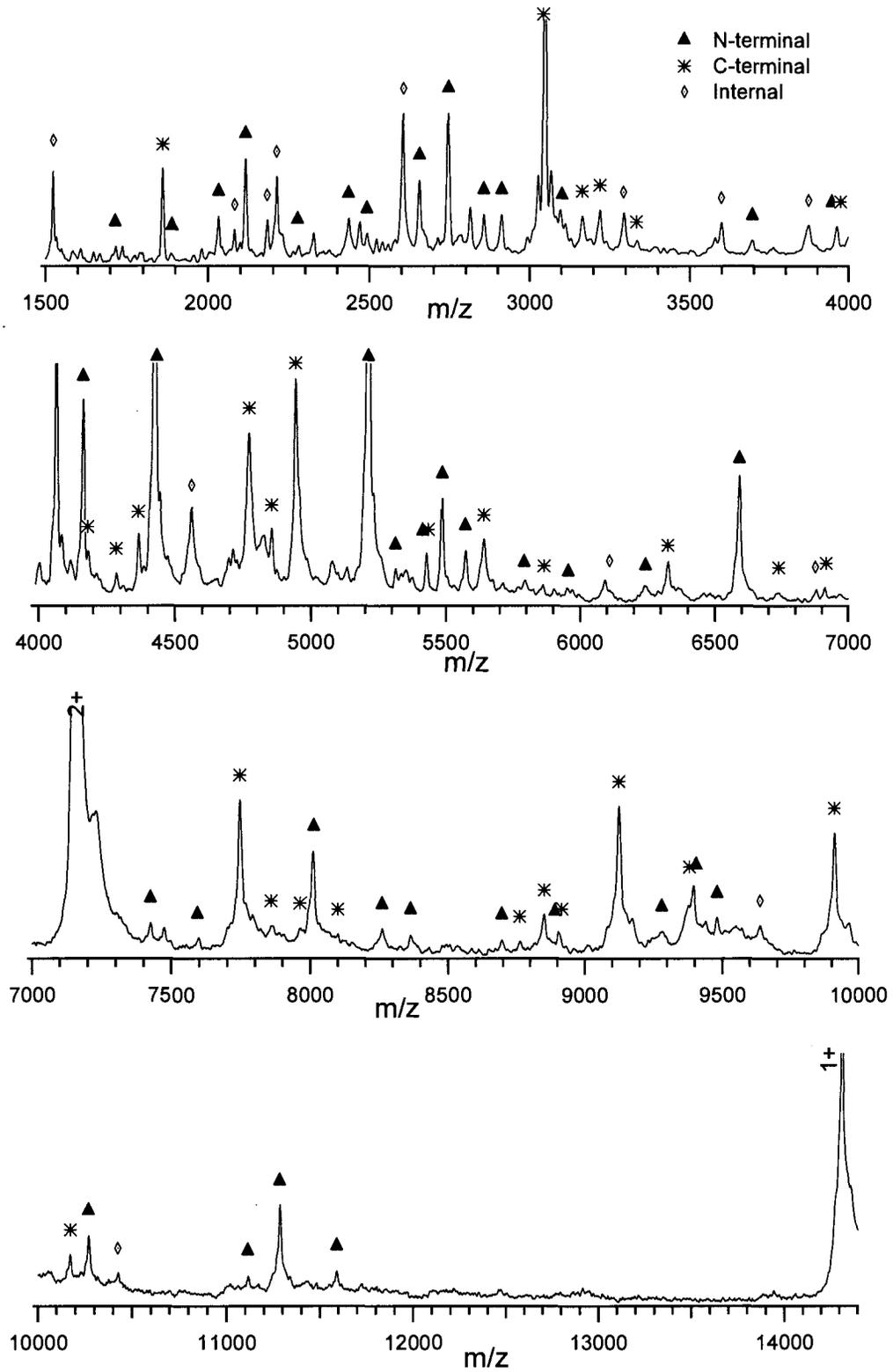
Figure 2.3.19 Expanded MALDI spectra of HEWL after mixing with 2 M NaOH and microwave irradiation for 30 seconds.

10↓ ↓↓ ↓↓ 20↓ ↓↓↓↓↓↓ 30 ↓↓ ↓↓↓ 40 ↓ ↓ ↓↓50↓
KVFGRCELAA AMKRHGLDNY RGYSLGNWVC AAKFESNFNT QATNRNTDGS
　　　　　　　　　　　　　　　　　　　　　　↑↑ ↑ ↑

↓↓↓60 ↓ ↓ 70 ↓ ↓↓↓ ↓ 80 ↓↓ 90 ↓ 100↓
TDYGILQINS RWWCNDGRTP GSRNLCNIPC SALLSSDITA SVNCAKKIVS
　↑ ↑ 　 ↑↑ 　↑↑ ↑↑ 　↑↑ 　↑ ↑↑ ↑

↓↓ 110 ↓ 120
DGNGMNAWVA WRNRCKGTDV QAWIRGCRL
↑↑↑ 　↑ ↑ ↑↑↑ ↑↑↑↑
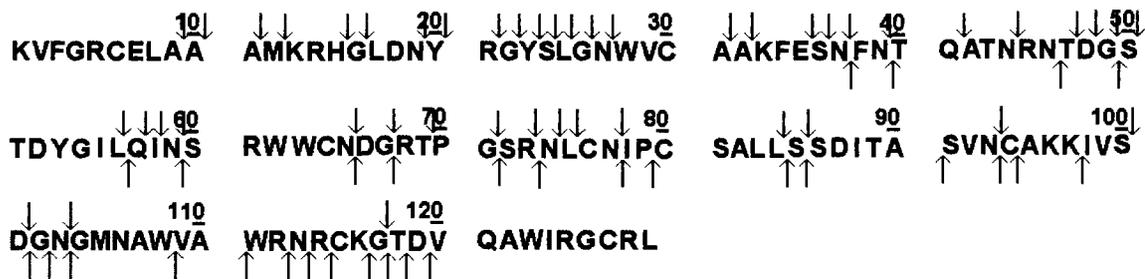
Figure 2.3.20 Peptide bond hydrolysis map of HEWL in 1M NaOH and 30 seconds irradiation time. ↓ - hydrolysis site identified by the detection of the N-terminal peptide; ↑ - hydrolysis site identified by the detection of the C-terminal peptide. N-terminal purification tag shown in green.

We also tested whether the presence of the base is affecting the overall structure of the protein and whether any oligomerization or aggregation of the sample occurs that could lead to solvent inaccessibility and subsequently a limited hydrolysis.

For CD analysis a stock solution of about 1.1 mg mL^{-1} of HEWL was prepared in 50 mM NaAC, pH 5.5. The sample was than mixed 1 to 1 (v/v) with 2 M NaOH. The CD spectra of HEWL at pH 5.5 and in 1M NaOH are shown in Figure 2.3.21. The spectrum shows a decrease in the signal at 222nm and 216 nm which suggests a decrease in the helical and sheet content of the protein. When using CDPro the results obtained, with the first and third reference sets did not reflect these decreases. Only when using reference set SDP48 was a decrease in values observed. This is probably caused by the fact that the spectrum was recorded in a very limited range, only 210 nm to 250 nm, which is missing the range between 190 nm to 210 nm that would give us a better estimation of the helical and random coil contents. Another reason could be that only SDP48 contains reference spectra for denatured proteins. The results of the deconvolution for HEWL in 1M NaOH is shown in Table 2.3.8.

At pH 5.5, in 50mM sodium acetate, the DLS spectra of HEWL showed a peak of significant height [Figure 2.3.22. (a)] with a hydrodynamic radius of 1.8 nm. This is very close to the expected 1.9nm, showing that the peak is corresponding to a pure lysozyme monomer. Figure 2.3.22 (b) shows the light scattering intensity trace for HEWL under basic conditions. This monodisperse peak showed a hydrodynamic radius of 2.1nm, which is a clear indication that the lysozyme stays monomeric or compact under these

conditions. The hydrodynamic radii of HEWL under the different solvent conditions are summarized in Table 2.3.9.

In conclusion, under the conditions of the MABH, lysozyme seems to partially maintain its structural features and stays in a monomeric state. It is possible that under the microwave irradiation and heating the protein is further denatured. This could also be favoured by the presence of DTT which reduces the disulphide bonds making the protein more prone to denaturation. Although the protein seems to have a less ordered structure than in the case of highly acidic conditions, the MABH was not able to produce an almost full N- and C-terminal ladder sequence as the MAAH is capable of doing.
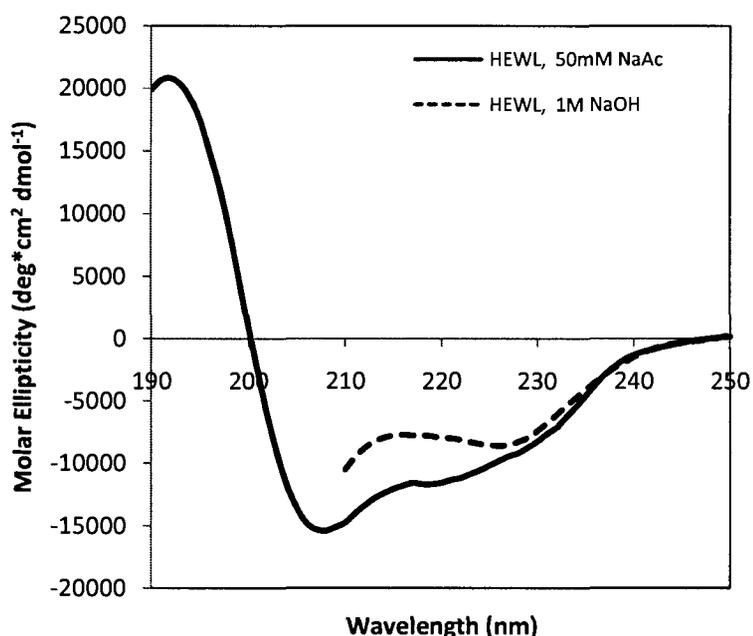
Figure 2.3.21 CD spectra of HEWL recorded at 20°C, pH 5.5 and in 1 M NaOH

Table 2.3.8 Deconvolution results of the CD spectra for HEWL in 1M NaOH

Program	Ref. Set	Helix (%)	Sheet (%)	Turn (%)	Random (%)
CDSSTR	SP43	44	17	17	22
	SDP48	40	18	14	28
	SMP56	43	15	17	25
CONTINLL	SP43	35	7	22	36
	SDP48	32	6	21	41
	SMP56	36	8	23	33
SELCON3	SP43	32	14	19	33
	SDP48	31	7	16	44
	SMP56	32	16	19	30
Average±SD		36±5	12±5	19±3	32±7
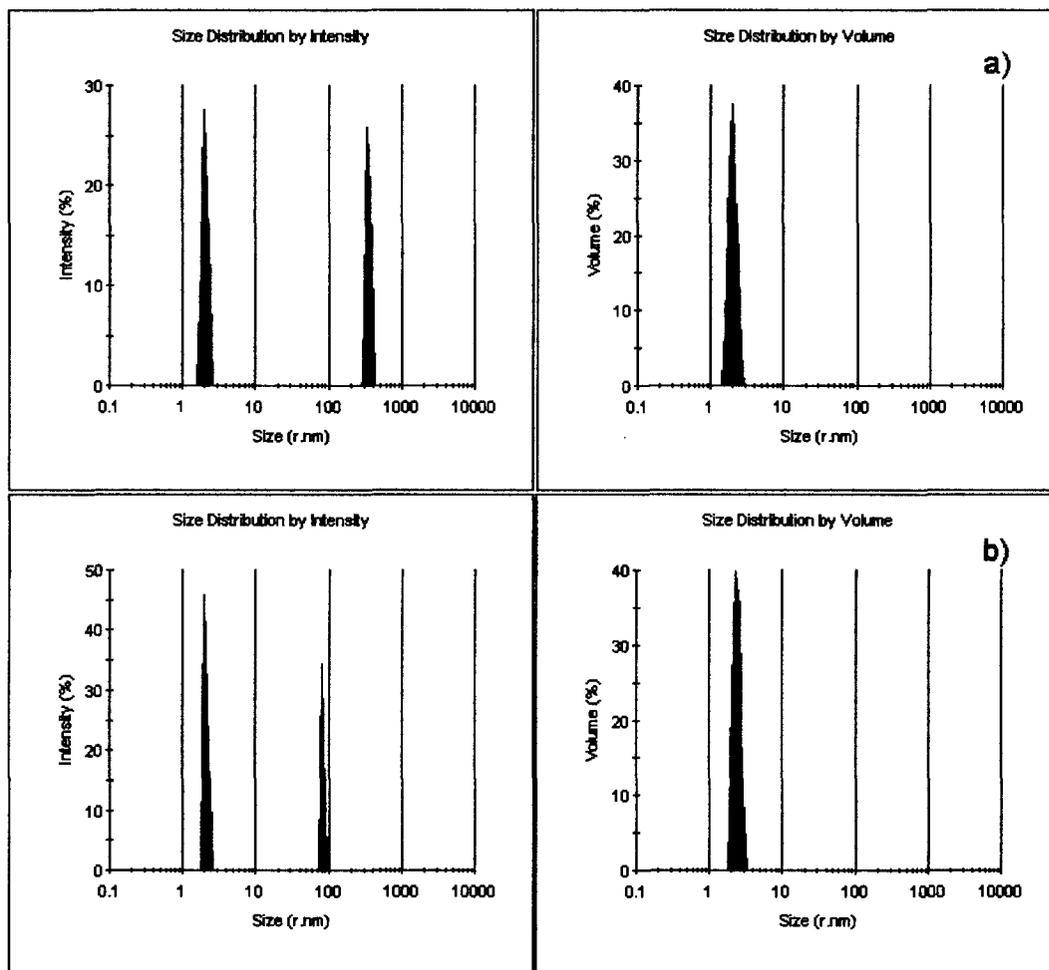
Figure 2.3.22 Light scattering trace of HEWL in: a) 50mM NaAc and b) 1M NaOH

Table 2.3.9 Hydrodynamic radii determined by dynamic light scattering for
HEWL under different solvent conditions

Lysozyme	R$_H$ (nm)	Molecular Weight (kDa)	
		Expected	Estimated
50mM sodium acetate, pH 5.5	1.8	14.3	13.3
1M NaOH	2.1	14.3	18.0

2.3.6. Microwave-assisted base hydrolysis of Horse Myoglobin

Horse Myoglobin, or MYG_HORSE, is a single chain, soluble protein. Muscle tissues have a high concentration of myoglobin where it works as the primary oxygen carrier[12]. The release of myoglobin can be associated with tissue damage. The protein consists of 153 amino acids and contains a heme prosthetic group. According to

Hersleth et al[13] (PDB 1GJN) at pH 5.2 these amino acids are organized in 8 helices containing 114 residues (74%) and 4 turns containing 13 residues (8%). The primary and secondary structure of myoglobin is shown in Figure 2.3.23.


```
       10            20            30            40            50
MGLSDGEWQQ  VLNVWGKVEA  DIAGHGQEVL  IRLFTGHPET  LEKFDKFKHL

       60            70            80            90           100
KTEAEMKASE  DLKKHGTVVL  TALGGILKKK  GHHEAELKPL  AQSHATKHKI

      110           120           130           140           150
PIKYLEFISD  AIIHVLHSKH  PGDFGADAQG  AMTKALELFR  NDIAAKYKEL  GFQG
```

ᴫᴫᴫ α-helix   ➡ β-sheet   ────random coil
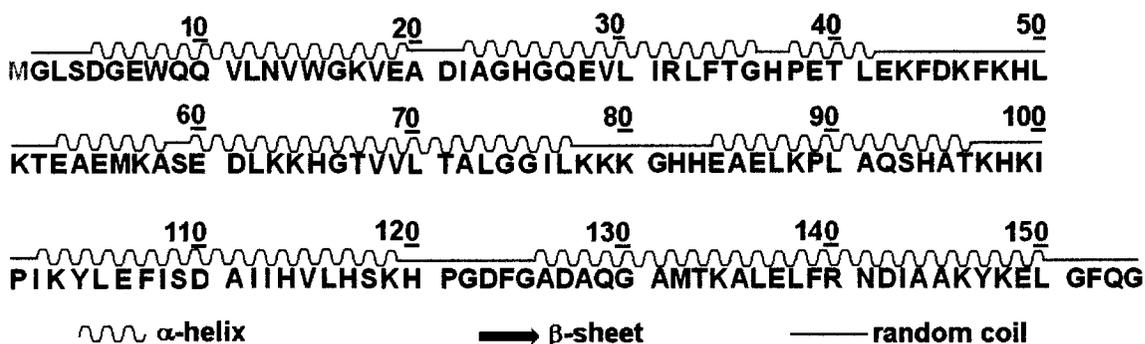
Figure 2.3.23 Primary and secondary structures of MYG_HORSE[13]

Figure 2.3.24 shows the expanded mass spectrum of MYG_HORSE after MAAH in 1M NaOH and 30 seconds irradiation time. The mass spectrum showed peaks along a wide mass range with good signal-to-noise ratios. The lower mass region of the spectrum consists of peaks corresponding to the N- and C-terminal ladder. From the mass spectrometric analysis of the polypeptide ladder a long stretch of sequence can be deduced. From the masses corresponding to the N-terminal ladder it can be determined that the first Met is not present, being cut-off after production, a common post-translational modification. However, in the higher mass region only a limited number of peaks have been identified, some regions of the protein being almost completely missed. Ion suppression[14] and detector saturation[15] can be problems in mass spectrometry, therefore to improve the coverage in the high mass region the mass spectrometer was set to suppress the peaks corresponding to the low mass region. This would allow only higher mass ions than the set minimum to reach the detector. Even with these settings the sequence coverage of the higher mass region could not be improved. Figure 2.3.25 shows the peptide bond hydrolysis map of the MYG_HORSE.
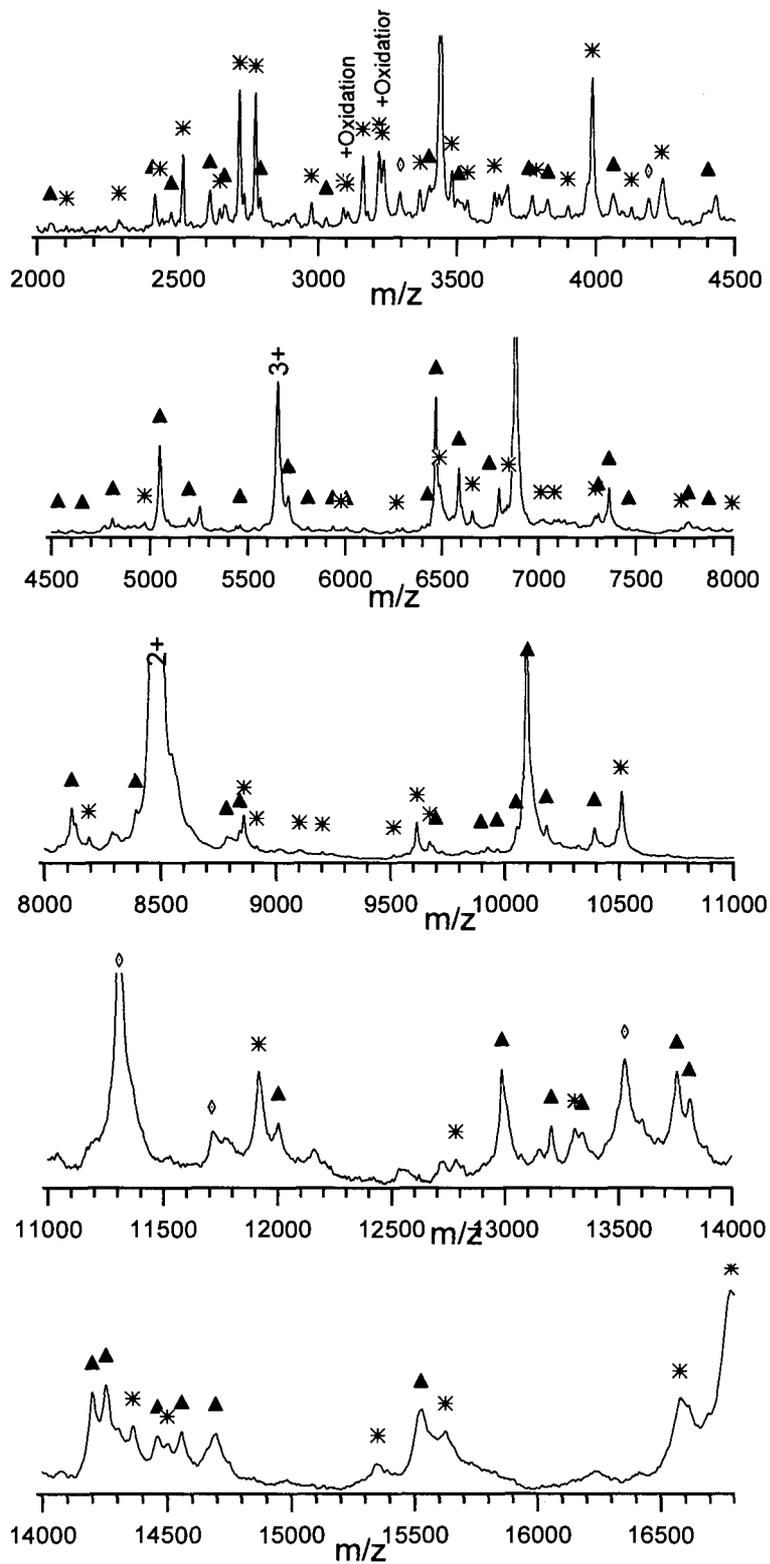
Figure 2.3.24 Expanded MALDI spectra of the MYG_HORSE hydrolysate after 30 seconds irradiation time in 1M HCl

```
        10              20    ↓↓↓↓↓ 30   ↓↓ ↓↓↓↓40 ↓↓↓↓↓↓ 50
MGLSDGEWQQ VLNVWGKVEA DIAGHGQEVL IRLFTGHPET LEKFDKFKHL
    ↑           ↑          ↑          ↑         ↑    ↑

↓↓↓↓ ↓↓↓60 ↓    ↓↓↓ 70↓ ↓   ↓  ↓ 80↓ ↓           ↓90 ↓ ↓↓↓  100
KTEAEMKASE DLKKHGTVVL TALGGILKKK GHHEAELKPL AQSHATKHKI
    ↑          ↑       ↑↑↑   ↑↑ ↑↑    ↑ ↑ ↑      ↑↑↑↑↑↑↑

     110           ↓120↓      ↓↓ ↓130↓  ↓↓↓  140 ↓      150
PIKYLEFISD AIIHVLHSKH PGDFGADAQG AMTKALELFR NDIAAKYKEL GFQG
↑            ↑       ↑↑↑↑↑ ↑↑↑↑↑↑↑↑ ↑ ↑↑↑↑↑ ↑
```
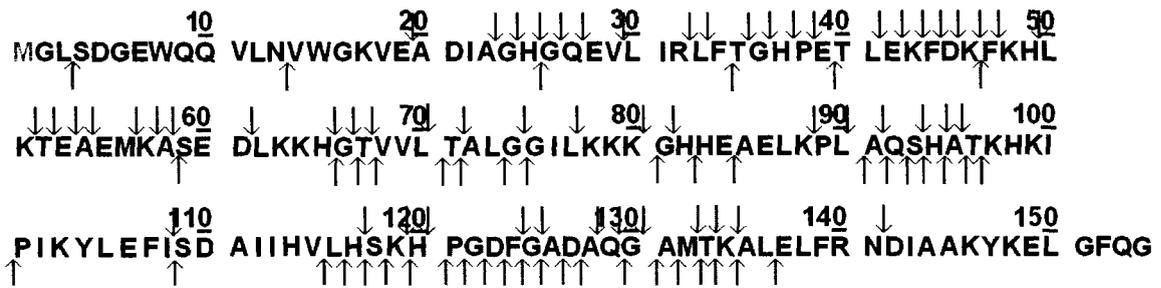
Figure 2.3.25 Peptide bond hydrolysis map of MYG_HORSE in 1M NaOH and 30 seconds irradiation time. ↓ - hydrolysis site identified by the detection of the N-terminal peptide; ↑ - hydrolysis site identified by the detection of the C-terminal peptide. N-terminal purification tag shown in green.

Again, we have tested whether the presence of the base is affecting the overall structure of the protein and whether the protein is in a monomeric state under the conditions used.

For CD experiments a stock solution of 1.2 mg mL^{-1} MYG_HORSE was prepared in 50 mM NaAC, pH 5.5. The sample was then mixed 1 to 1 (v/v) with 2 M NaOH. The CD spectra of MYG_HORSE at pH 5.5 and in 1M NaOH are shown in Figure 2.3.26. In 50 mM NaAc, the spectrum shows very intense signals at both 222nm and 208 nm which show that the protein contains a high percentage of helical content organized in two or more regions. The CDPro results are in a very good agreement with the reference data[13], as shown in Table 2.3.10. No significant differences have been detected after mixing with 2 M NaOH. It seems that the protein is able to maintain its secondary structure under highly basic conditions. Although it is possible that in the presence of microwave radiation and heating of the sample structural alteration occurs. Deconvolution results for MYG_HORSE in 2 M NaOH are shown in Table 2.3.11.

For DLS experiments protein samples of 1 mg mL^{-1} were used. Results are summarized in Table 2.3.12. In 50 mM NaAc, pH 5.5, a hydrodynamic radius of 2.2 nm has been determined. This is corresponding to a molecular weight of 20.7 kDa and shows that the protein is purely monomeric. Under highly basic conditions a hydrodynamic radius of 2.1 nm has been determined which shows no significant differences between the two different conditions. The DLS spectra for MYG_HORSE are presented in Figure 2.3.27.
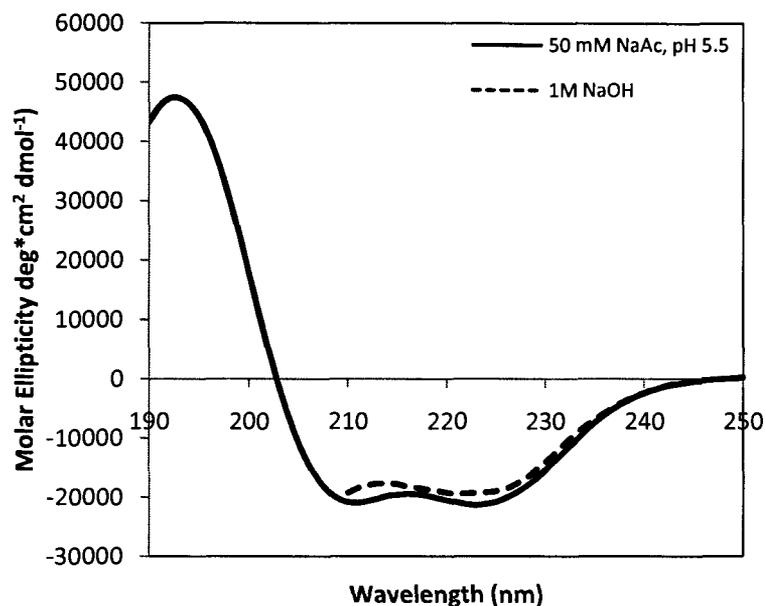
Figure 2.3.26 CD spectra of MYG_HORSE recorded at 20°C, pH 5.5 and in 1 M NaOH

Table 2.3.10 Deconvolution results of the CD spectra for MYG_HORSE in 50 mM NaAc

X-ray	PDB 1GJN	74	0	8	18
Program	Ref. Set	Helix (%)	Sheet (%)	Turn (%)	Random (%)
CDSSTR	SP43	80	4	5	12
	SDP48	79	3	7	10
	SMP56	80	2	5	13
CONTINLL	SP43	73	4	6	17
	SDP48	74	3	8	15
	SMP56	73	2	7	18
SELCON3	SP43	73	2	8	19
	SDP48	73	2	9	17
	SMP56	72	2	10	19
Average±SD		75±3	3±1	7±2	16±3

Table 2.3.11 Deconvolution results of the CD spectra for MYG_HORSE in 1M NaOH

Program	Ref. Set	Helix (%)	Sheet (%)	Turn (%)	Random (%)
CDSSTR	SP43	75	4	7	14
	SDP48	80	7	6	8
	SMP56	77	5	7	11
CONTINLL	SP43	66	5	11	18
	SDP48	66	4	11	19
	SMP56	66	5	12	18
SELCON3	SP43	67	3	12	20
	SDP48	67	4	12	19
	SMP56	65	4	13	21
Average		70±6	4±1	10±3	16±4

Table 2.3.12 Hydrodynamic radii determined by dynamic light scattering for MYG_HORSE under different solvent conditions

Myoglobin	R_H (nm)	Molecular Weight (kDa)	
		Expected	Estimated
50mM NaAc, pH 5.5	2.2	16.9	20.7
1M NaOH	2.1	16.9	19.1
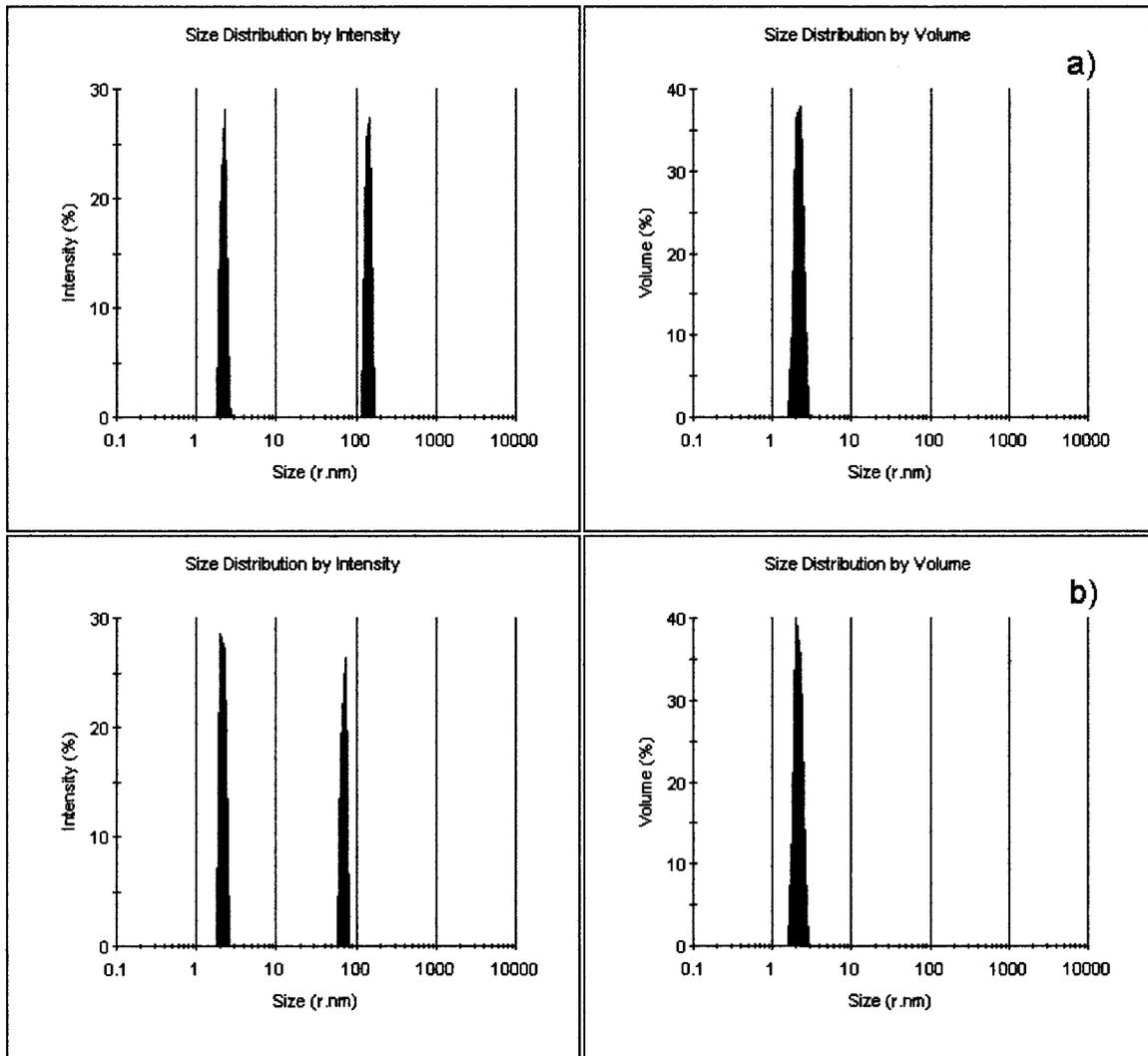
Figure 2.3.27 Light scattering trace of MYG_HORSE in: a) 50mM NaAc and b) 1M NaOH

2.3.7. Microwave-assisted base hydrolysis of Horse Heart Cytochrome C

The smallest standard protein that was used in order to assess the feasibility of MABH combined with MALDI-TOF-MS was Cytochrome C. As already mentioned in the case of MABH of MYG_HORSE, ion suppression can be a problem in mass spectrometry and high molecular weight species might be hard to detect. It was also shown that in the case of CYC_HORSE, MAAH followed by MS analysis can give a mass spectrum from which the complete ladder sequence can be determined. Therefore, if the MABH is able to give the same type of hydrolysis pattern as the MAAH does, the detection of the resulting fragments should be possible (Figure 2.3.4). The mass spectrum of the hydrolysate, Figure 2.3.28, obtained by the MABH of CYC_HORSE in 1M NaOH and 30 seconds irradiation time showed a limited number of peaks with relatively low signal-to-noise ratios, mostly from the N-terminal part of the protein. From the peptide bond hydrolysis map presented in Figure 2.3.29, it can be seen that a fairly good hydrolysis of the N-terminal part was achieved but a very limited number of peaks has been detected for the C-terminal part.
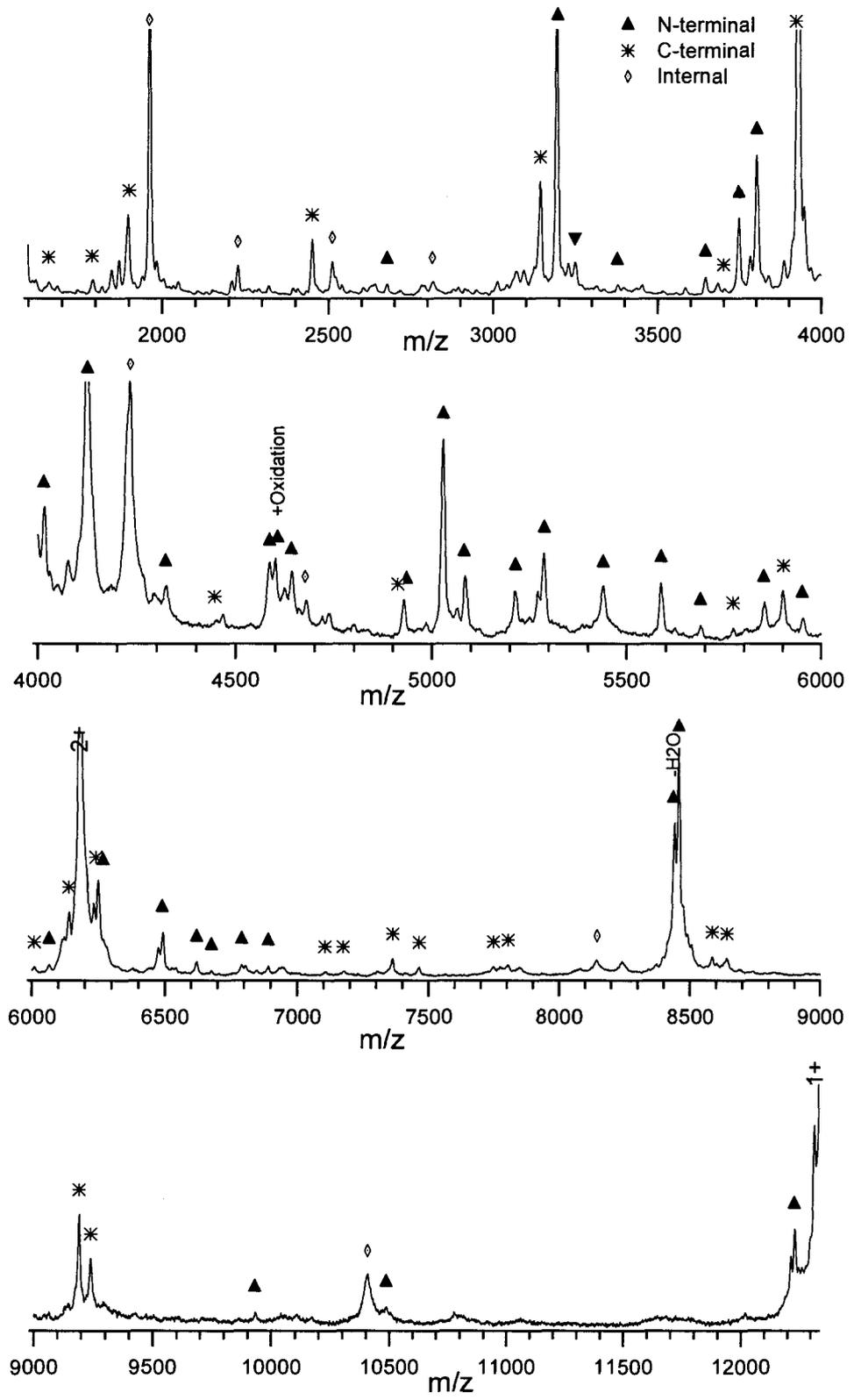
Figure 2.3.28 Expanded MALDI spectra of the CYC_HORSE hydrolysates after 30 seconds
irradiation time in 1M HCl
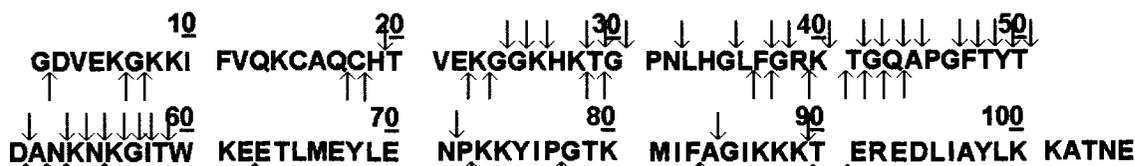
Figure 2.3.29 Peptide bond hydrolysis map of CYC_HORSE in 1M NaOH and 30 seconds irradiation time. ↓ - hydrolysis site identified by the detection of the N-terminal peptide; ↑ - hydrolysis site identified by the detection of the C-terminal peptide. N-terminal purification tag shown in green.

We have also tested how highly basic conditions affect the secondary structure of the protein. For CD experiments a stock solution of 1.0 mg mL^{-1} CYC_HORSE was prepared in 50 mM NaAC, pH 5.5. The sample was than mixed 1 to 1 (v/v) with 2 M NaOH. The CD spectra of CYC_HORSE at pH 5.5 and in 1M NaOH are shown in Figure 2.3.30. The results of the CD deconvolution of the CD spectra for CYC_HORSE at pH 5.5 in 50 mM NaAC were shown and discussed in Section 2.2.2. When mixing with 2 M NaOH a decrease in the signal intensities at 222 nm suggests a decrease of the helical content. From the CD data only it can be concluded that the protein is partially denatured.

The DLS experiments performed in 1M NaOH, Figure 2.3.31, showed a monodisperse peak with a hydrodynamic radius of 2.0 nm. This is slightly higher than the expected radius of 1.8 nm but it shows that the protein stays monomeric under basic conditions. A summary of the DLS results is shown in Table 2.3.14.

It has already been shown that under acidic conditions the protein is partially denatured. It is also possible that after applying the microwave irradiation it becomes fully denatured and hydrolysis of all peptide bonds is possible. Although CYC_HORSE seems to be denatured when performing MAAH, the studies performed using HEWL showed that even if the protein maintains its structural features, hydrolysis of the bonds is still possible. It seems that as long as the peptide bond is solvent accessible hydrolysis of the bond is possible. In the case of the MABH of CYC_HORSE, the protein seems to undergo similar structural changes as shown by the CD and it stays monomeric, meaning that peptides bonds should be accessible for the solvent. Even in these conditions, according to the MALDI-TOF-MS data only a partial hydrolysis has been achieved.
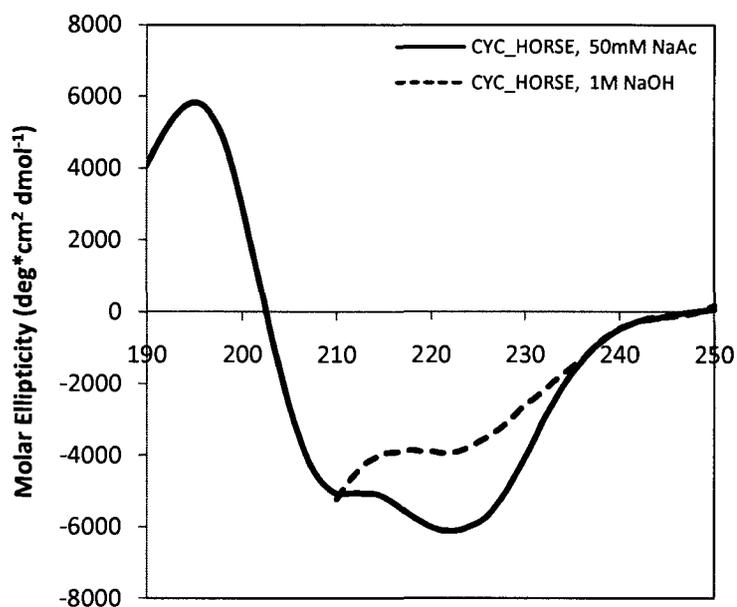
Figure 2.3.30 CD spectra of CYC_HORSE recorded at 20°C, pH 5.5 and in 1 M NaOH

Table 2.3.13 Deconvolution results of the CD spectra for CYC_HORSE in 1M NaOH

Program	Ref. Set	Helix (%)	Sheet (%)	Turn (%)	Random (%)
CDSSTR	SP43	11	27	24	38
	SMP50	7	18	11	64
	SMP56	10	27	25	37
CONTINLL	SP43	19	28	23	31
	SMP50	7	18	11	64
	SMP56	16	27	24	32
SELCON3	SP43	10	35	21	31
	SMP50	5	18	9	65
	SMP56	10	37	21	31
Average		11±4	26±6	19±7	43±16

Table 2.3.14 Hydrodynamic radii determined by dynamic light scattering for CYC_HORSE under different solvent conditions

Cytochrome C	R$_H$ (nm)	Molecular Weight (kDa)	
		Expected	Estimated
50mM NaAc, pH 5.5	1.7	12.3	12.1
1M NaOH	2.0	12.3	17.9
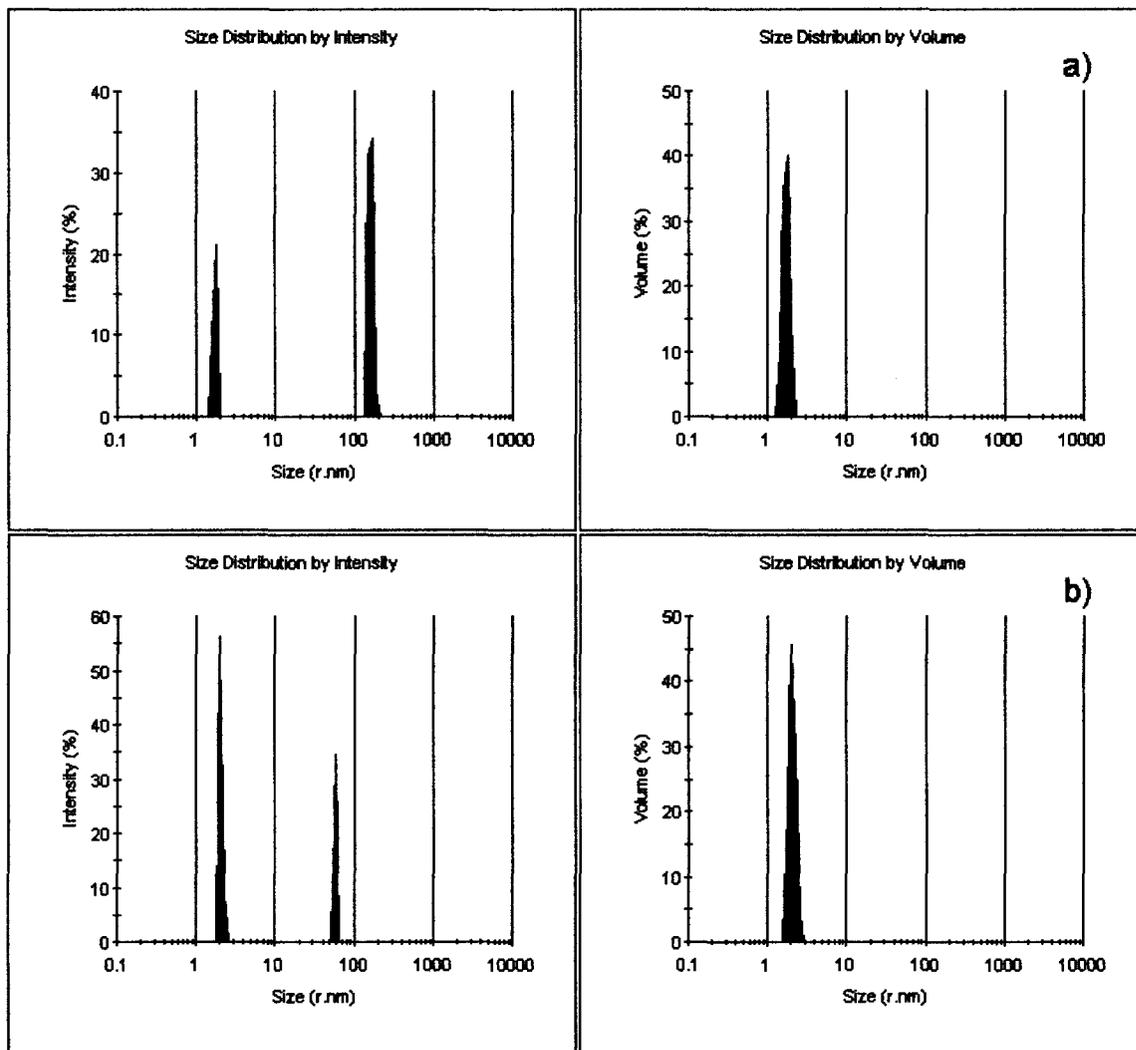
Figure 2.3.31 Light scattering trace of CYC_HORSE in: a) 50mM NaAc and b) 1M NaOH

2.4. Conclusions

In conclusion, we have shown that MAAH is applicable for proteins that contain stabilizing disulphide bonds. According to the CD data the HEWL maintains its secondary

structure under highly acidic conditions. The hydrolysis map shows that hydrolysis of peptide bonds can be observed for all kinds of secondary structures, such as α-helical, β-sheet, turn or random coil, when referring to the original protein structure. We have also tested the feasibility of MABH followed by MALDI-TOF-MS as a means of degrading proteins. It has been found that base hydrolysis generates N- and C-terminal peptides, as MAAH does, although the hydrolysis is not as uniform as in the case of MAAH and the method is not suitable for determination of complete protein sequences and their modification.

2.5. *References*

(1)   Zhong, H.; Zhang, Y.; Wen, Z.; Li, L. *Nat Biotechnol* **2004**, *22*, 1291-1296.

(2)   Wang, N.; Xie, C.; Young, J. B.; Li, L. *Analytical Chemistry*, (in press).

(3)   Dai, Y.; Whittal, R. M.; Li, L. *Anal Chem* **1999**, *71*, 1087-1091.

(4)   Roberts, J. D.; Caserio, M. C. **1964**.

(5)   Qi, P. X.; Beckman, R. A.; Wand, A. J. *Biochemistry* **1996**, *35*, 12275-12286.

(6)   Hirel, P. H.; Schmitter, M. J.; Dessen, P.; Fayat, G.; Blanquet, S. *Proc Natl Acad Sci U S A* **1989**, *86*, 8247-8251.

(7)   Banci, L.; Bertini, I.; Huber, J. G.; Spyroulias, G. A.; Turano, P. *J Biol Inorg Chem* **1999**, *4*, 21-31.

(8)   Wang, J.; Dauter, M.; Alkire, R.; Joachimiak, A.; Dauter, Z. *Acta Crystallogr D Biol Crystallogr* **2007**, *63*, 1254-1268.

(9)   Zhong, H.; Marcus, S. L.; Li, L. *J Am Soc Mass Spectrom* **2005**, *16*, 471-481.

(10)   Zahn, D. *Chem. Phys. Lett.* **2004**, *383*, 134-137.

(11)   Zahn, D. *Eur. J. Org. Chem.* **2004**, *2004*, 4020-4023.

(12)   Ordway, G. A.; Garry, D. J. *J Exp Biol* **2004**, *207*, 3441-3446.

(13)   Hersleth, H. P.; Dalhus, B.; Gorbitz, C. H.; Andersson, K. K. *J Biol Inorg Chem* **2002**, *7*, 299-304.

(14)   Hillenkamp, F.; Karas, M.; Beavis, R. C.; Chait, B. T. *Anal Chem* **1991**, *63*, 1193A-1203A.

(15)   Schriemer, D. C.; Li, L. *Anal Chem* **1997**, *69*, 4169-4175.

Chapter 3: Microwave-assisted acid and base hydrolysis of the rSHPrP90-231 prion construct

3.1. Introduction

Proteinaceous infectious particles, also known as prions, are endogenous proteins that can cause a variety of fatal neurodegenerative diseases, collectively called Transmissible Spongiform Encephalopathy's (TSE's). This disease can affect both humans and animals. In humans, TSE's includes Gerstmann-Straussler-Scheinker (GSS), Fatal familial insomnia (FFI) and Creutzfeldt-Jakob disease (CJD). This group of diseases typically has long incubation periods[1, 2], characteristic of spongiform degeneration of the brain, as well as a failure to induce inflammatory response. Prion diseases are not the only neurodegenerative diseases caused by protein aggregation. There are more than 20 disorders that have been associated with the unfolding and ordered aggregation of a specific protein[3, 4]. The fact that prion diseases are treated as a special category of neurodegenerative diseases is because they can also be infectious.

To map the sequence of an entire protein and to detect any modifications along its sequence, the generation, detection and sequencing of short, sometimes partially overlapping, peptide sequences is needed. The use of enzyme digestion may be difficult to generate peptides that cover the entire sequence of a protein. These problems become even more complex in the case of proteins which have a tendency to aggregate, oligomerize or to form fibrils. An alternative method to enzymatic digestion was developed by Zhong[5] et al. This method uses a high concentration of an acid coupled with microwave irradiation in order to hydrolyse proteins. Using this technique specific cleavage of the amide bond of the intact proteins is achieved and sequence specific ladders containing N- and C- terminal amino acids are produced. The mass spectrometric analysis of the hydrolysate results in a spectrum consisting of peaks of this polypeptide ladder from which the amino acid sequence along with any possible modifications can directly be read.

A simple and efficient digestion of prions could be a key factor in the characterization of their structure, dynamics and kinetics. Enzymes, such as trypsin, chymotrypsin or proteinase K can offer partial sequence coverage of the cellular form of the prion protein, but they are not suitable for the digestion of the insoluble form. Familial transmission of these diseases is also known to be associated with mutations in the PrP gene[6]. Using recombinant prions it is possible to prepare native or chemically modified and selectively mutated prion constructs. These constructs can be modified using different physico-chemical methods, digested and then analyzed using LC-MALDI-MS and MS/MS. The information gained this way can help us to understand the etiology of TSEs and can also serve as basic information for the development of new diagnostic tools and therapeutic products.

To chemically probe the structure, the prion protein in different forms needs to be degraded into peptides for sequencing and determination of the modification sites. In this work, we examine the use of microwave-assisted acid/base hydrolysis (MAAH/MABH) as means of degrading different forms of prion proteins into peptides that can be analyzed by MALDI-MS.

3.2. Experimental

3.2.1. Materials and Reagents

Unless otherwise noted, all chemicals were purchased from Sigma (St. Louis, MO) and were of analytical grade. For HPLC separations, MS analysis, and preparation of digests, Optima grade water and acetonitrile were used (Fisher Scientific, Mississauga, ON). 37% HCl (ACS Grade) was from Merck KGaA, Darmstadt, Germany.

3.2.2. Production and Purification of rSHPrP^{C}90-231

Syrian Hamster prion wild type construct, residues 90-231 (rSHPrP90-231), was over-expressed in *E. coli* as N-terminal histidine-tagged fusion protein and purified using a Ni-NTA agarose resin. The protocol was adapted from the method described by Zahn[7]

et al. Briefly, cells were taken from frozen glycerol stocks and grown in 100 mL LB with Ampicillin selection overnight. This starter culture was used to inoculate 500 mL of M9 media to an O.D. between 0.6 and 1.0. Culture was grown for 1 hour before induction using IPTG. Post induction culture was left to produce rSHPrP90-231 for 18 hours. Cells were harvested by centrifugation at 3,000 rpm for 25 minutes at 4°C, resuspended in lysis buffer (8 M Urea, 10 mM Tris, 100 mM K_2PO_4, 10 mM reduced Glutathione, pH 8.0) and subjected to 5 rounds of freeze/thaw. Lysate was cleared by centrifugation at 12,000 rpm for 1 hour and 10 minutes. Cleared lysate was added to 35 mL of nickel-NTA resin for sufficient binding time. The column was washed with 100 mL of denaturing buffer (8 M Urea, 10 mM Tris, 100 mM K_2PO_4, pH 8.0). The protein was refolded using a 200 mL gradient going from denaturing buffer to refolding buffer (10 mM Tris, 100 mM K_2PO_4, pH 8.0). The column was subsequently washed with 200 mL of refolding buffer and 100 mL of refolding buffer containing 50 mM imidazole. Refolded protein was eluted with 200 mL of a competitive nickel-binding buffer (500 mM imidazole, 10 mM Tris, 100 mM K_2PO_4, pH 5.6). 5 mL fractions were collected and sampled for eluted prions. Samples were dialyzed against 50mM sodium acetate buffer, pH 5.5, for further concentration. Production yields of the target protein varied from 10 to 40 mg L^{-1} of E. coli culture (grown in minimal media).

3.2.3. Enzymatic Digestion

Solution tryptic digests of rSHPrP^{C}90-231 were carried out at a concentration of 1 mg mL^{-1} with an enzyme to substrate ratio of 1:20 at 37°C, overnight. The digests were acidified using TFA and spotted on a MALDI plate using the two-layer sample preparation method.

3.2.4. Microwave-Assisted Acid Hydrolysis

A 1 mg mL^{-1} solution of rSHPrP^{C}90-231 in 50 mM sodium acetate buffer, pH 5.5, was mixed with 1:1 (v/v) with 6M HCl and 1 µL of 450 mM DTT in a 1.5 mL Eppendorf vial. Samples were subjected to microwave irradiation using a simple household microwave oven. Samples were cooled at room temperature for 5 minutes. The

hydrolysates were dried in a vacuum centrifuge (SpeedVac) to remove the acid. Samples were reconstituted in 30 μL of 6M GdnHCl. 2 μL of 450mM DTT was added and the samples were diluted to a final concentration of 3M GdnHCl in 200 mM NH₄HCO₃, pH 8.0. The samples were incubated for 60 minutes at 37°C to reduce the disulphide bond. All samples were acidified and diluted with TFA to a final pH of 2.0.

3.2.5. *Microwave-Assisted Base Hydrolysis*

A 1 mg mL^{-1} solution of rSHPrP90-231 or HEWL in 50 mM sodium acetate buffer pH 5.5 was mixed with 1 μL of 450 mM DTT in a 1.5 mL Eppendorf vial. 2M NaOH was added in a 1:1 ratio (v/v). Samples were subjected to microwave irradiation using a simple household microwave oven. Samples were cooled at room temperature for 5 minutes. GdnHCl prepared in 200mM NH₄HCO₃ was added and concentrated HCl was added until the pH of the samples was adjusted to 8. 450mM DTT was added in a final concentration of 15mM in DTT and also maintaining the concentration of GdnHCl at 3M. The samples were incubated for 60 minutes at 37°C to reduce the disulphide bond. All samples were acidified and diluted with TFA to a final pH of 2.0.

3.2.6. *Desalting*

Peptide desalting was carried out in an Agilent 1100 HPLC system (Palo Alto, CA) using a newly developed method[8]. Briefly, the desalting of the hydrolysates was performed on a 4.6 mm × 5 cm Polaris C18 A column with a particle size of 3 μm and 300 Å pore (Varian, USA). After loading of the polypeptide sample, the column was flushed with 97.5% mobile phase A (0.1% TFA in water) and 2.5% mobile phase B (0.1% TFA in ACN) and the salts were effectively removed. Subsequently, the concentration of phase B in the mobile phase was step-wise increased to 90% to ensure complete elution of the polypeptide fractions from the column.

3.2.7. *MALDI-TOF-MS and MS/MS*

The desalted samples were concentrated in a SpeedVac vacuum centrifuge. The samples were then mixed with matrix solution of α-cyano-4-hydrocynnamic acid (CHCA).

Samples were then deposited on a MALDI plate using the two-layer sample preparation method[9] for matrix-assisted laser desorption ionization (MALDI) analysis.

The acidified tryptic digests were mixed with matrix solution of α-cyano-4-hydrocynnamic acid (CHCA) or 2'6'-dihydroxyacetphenone (DHAP) and spotted on a MALDI plate. A two-layer sample preparation method was used for CHCA and the dried droplet method[10] was used for DHAP.

MALDI-TOF mass spectra were obtained on an Applied Biosystems/MDSSciex 4800 *Plus* MALDI-TOF/TOF™ Analyzer. Ionization was performed with a diode-pumped Nd:YAG laser at 355 nm. The analyzer was used in a reflector or linear mode of operation.

3.2.8. FT-MS

FT-MS spectra were recorded using a Bruker 9.4T Apex-Qe FTICR from Brucker Daltonics (Billerica, MA, USA). Samples were injected by direct infusion using a syringe pump and an ESI source of ionization. Samples were diluted using 20% ACN/0.1% formic or trifluoroacetic acid.

3.2.9. *Circular Dichroism*

CD spectra were recorded using an Olis DSM 17 spectrometer (Olis, Bogart, USA). Scans were performed with a bandwidth of 4 nm and data spacing of 1.0 nm. Each spectrum was recorded at 20°C using a 0.02 cm cuvette on protein samples in the concentration range from 0.5 to 0.7 µg µL^{-1}. All spectra represent the average of five individual scans after subtracting the background spectra. To monitor the conformational changes under acidic and basic conditions, rSHPrP^{C}90-231 samples were mixed 1:1 (v/v) with 6M HCl or 2M NaOH followed by the recording of the CD spectra. Recording of the spectra was possible between 190 nm and 250 nm in 50 mM NaAC, between 200 nm and 250 nm in 3M HCl and between 210nm and 250 nm in 1M NaOH. Deconvolution of the CD spectra was performed using the CDPro software package. Secondary structure was determined using the three programs included in the software package (CDSSTR, CONTINLL and SELCON3) and 3 different Reference Sets. The

reference sets used were SP43, SP48 and SMP56. Three different data sets were chosen because every data set contains the CD spectra for different types of proteins. SP43 contains spectra for 43 different soluble proteins, recorded between 190 – 240 nm. SP48 contains the same spectra as SP43 and an additional 5 spectra for denatured proteins, recorded between 190 – 240 nm. SMP56 contains the 43 spectra contained in SP43 and additional spectra for 13 membrane proteins, recorded between 190 – 240 nm.

3.2.10. *Dynamic Light Scattering*

DLS spectra were recorded with a Malvern Zetasizer Nano-S (Malvern Instruments Ltd., UK). Scans were performed at an angle of 173° in a 50 µL cuvette. Recombinant SHPrP^C were prepared in 50 mM sodium acetate, pH 5.5, having a concentration of 1 mg mL^{-1}. For DLS experiments under acidic and basic conditions samples were mixed with 6M HCl or 1M NaOH. The DLS spectra for all samples were recorded immediately after mixing with the strong acid or base.

3.2.11. *SDS-PAGE and Acetic Acid Nondenaturing Polyacrylamide Gels.*

Samples were analyzed by SDS-PAGE using a 4-12% Ready Gel Tris-HCl precast acrylamide gradient gel (Bio-Rad, Hercules, CA). All samples were diluted with loading buffer to a final volume of 30 µL, boiled for 5 minutes and run at a constant voltage of 100V for 90 minutes. Approximately 10 µg of sample was loaded on each lane. The gel was developed following the Bio-Rad protocol for Bio-Safe Coomasie Blue staining.

Acetic acid nondenaturing polyacrylamide gel electrophoresis was carried out using standard 7.5%, 1.0 mm thick mini-gels stained with Coomasie-blue dye. The gels were prepared as follows: in a 75mL side arm flask 10mL of 30%acrylamide/0.8%bisacrylamide was mixed with 0.4 mL of 300mM sodium sulphite, 10 mL of 4X acetic acid buffer and 19.28 mL of water. Mixture was degassed under vacuum for ~5 minutes in order to speed up polymerization. After degassing, 0.3 mL of 10% (w/v) ammonium persulfate and 0.02 mL of TEMED was added. The mixture was mixed and used immediately.

3.3. Results and Discussion

3.3.1. Characterization of the rSHPrP^{C}90-231

First we have examined the rSHPrP^{C}90-231. The truncated rSHPrP90-231 was subcloned in the pET-15b expression vector. The gene construct used for the over-expression of the target protein is shown in Figure 3.3.1. Based on the sequence of the expression vector the sequence of the recombinant protein should contain a histidine purification tag, a thrombin cleavage site, the PrP^C protein fragment containing residues 90-231 and an extra serine at the end.
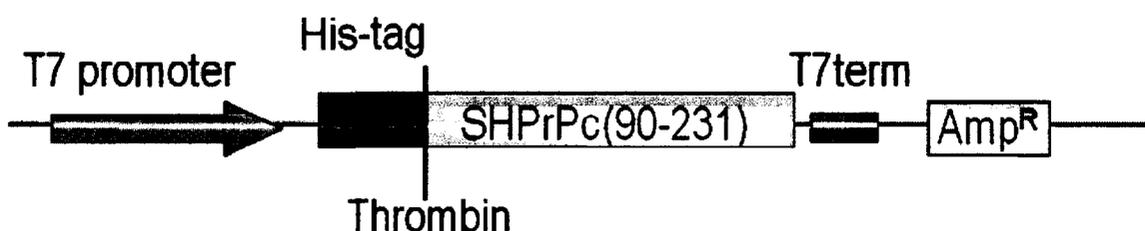
Figure 3.3.1. Syrian hamster prion gene in the pET-15b expression vector.

Figure 3.3.2 shows the theoretical sequence of the construct. The full length of this recombinant protein construct should contain 166 amino acid residues for which the calculated average molecular weight, using the ProtParam tool on the expasy proteomic website (http://expasy.org/tools/protparam.html), should be 18866.9 Da if all Cys residues appear as half cystines or 18864.9 Da if no Cys residues appear as half cystines.

```
MGS SHHHHHHSSG LVPRGSHMLE  ←—— purification tag
90          100        110        120        130        140
G  QGGGTHNQWN KPSKPKTNMK HMAGAAAAGA VVGGLGGYML GSAMSRPMMH

           150        160        170        180        190
   FGNDWEDRYY RENMNRYPNQ VYYRPVDQYN NQNNFVHDCV NITIKQHTVT

           200        210        220        230
   TTTKGENFTE TDIKIMERVV EQMCTTQYQK ESQAYYDGRR SS←—extra residue
```

Figure 3.3.2. Theoretical sequence of the rSHPrP90-231 prion construct.

In order to confirm the primary structure and to check for any possible modifications that could have happened during production and purification of the protein, first, we determined its molecular weight using MALDI-TOF-MS and FT-MS. The MALDI-TOF-MS

81

spectra of the intact protein, shown in Figure 3.3.3, showed an $[M+H]^+$ peak at 18735 Da. The FT-ICR-MS mass spectral analysis, Figure 3.3.4, showed an $[M+H]^+$ peak at 18734.67 Da. The difference between the theoretical mass of the $[M+H]^+$ peak of the oxidized form of the protein (18865.9 Da) and the determined $[M+H]^+$ molecular weight (18734.67Da) is approximately 131.2Da. This mass is closest to the mass of a Met residue (131.0 Da). In conclusion, the collected MS data shows that the first Met has been cleaved form the sequence in the *E. coli* host. This result was confirmed later by the MAAH data (see below).
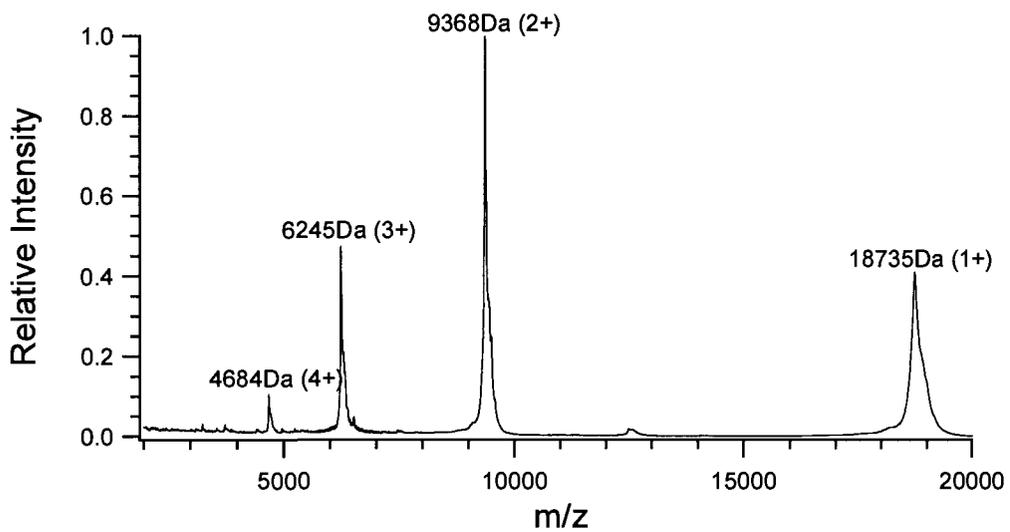
Figure 3.3.3 MALDI-TOF-MS spectra of the rSHPrP^{C}90-231 prion construct.
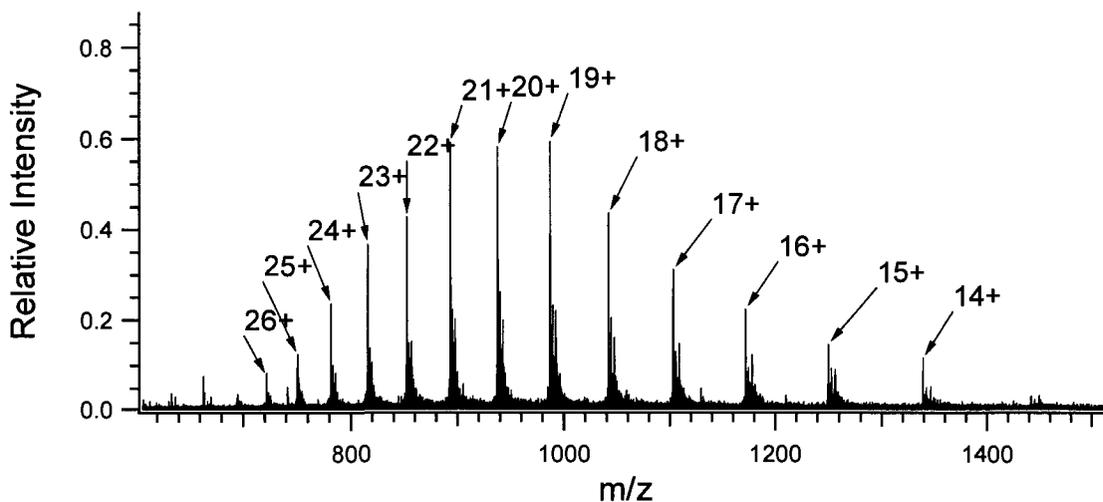
Figure 3.3.4 FT-MS spectra of the rSHPrP^{C}90-231 prion construct

One important post translational modification of the PrP^C is the presence of the disulphide bond between Cys179 and Cys214. The importance of the presence of this disulphide bond in the conversion of the protease sensitive PrP^C to a protease resistant form has been previously studied[11]. It has been found that the presence of reducing agents, such as dithiothreitol, ascorbate and sodium dithionite, inhibits the conversion of the protease resistant PrP^C to a protease resistant conformation in a cell-free conversion system. According to the FT-MS data the protein contains the disulphide bond between its two Cys residues. In order to check the validity of the FT-MS data a trypsin digestion was performed without the addition of any reducing agent in the sample mixture. The theoretical mass of the [M+H]$^+$ tryptic peptide containing the disulphide bond should be 5013.33 Da. The MALDI-TOF-MS spectrum, shown in Figure 3.3.5, showed a peak at 5012.97 Da indicating the presence of the disulphide bond.
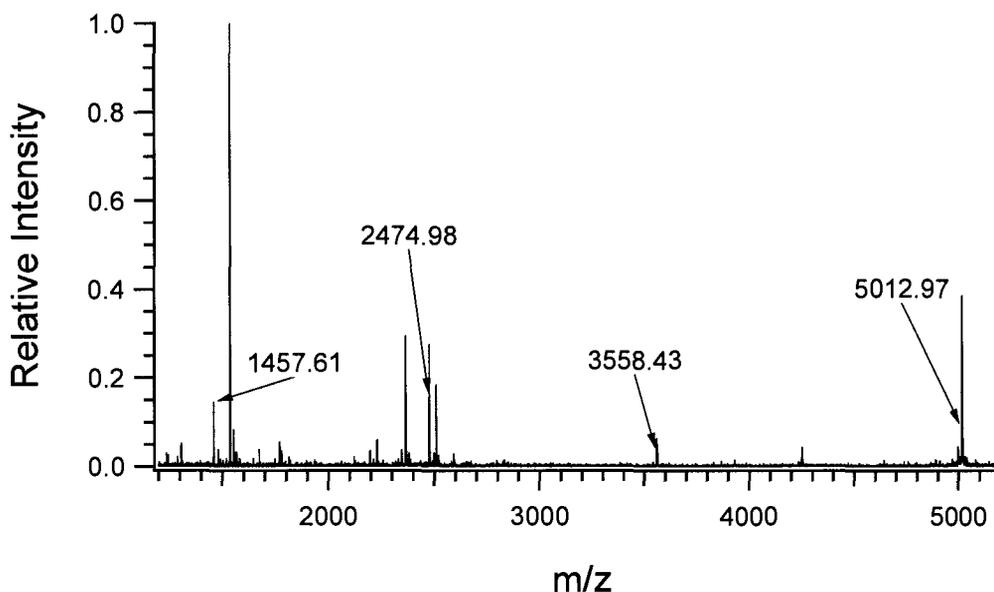
Figure 3.3.5. MALDI-TOF-MS spectra of a the rShPrP^{C}90-231 after trypsin digestion

The sequence of the tryptic peptide that contains the disulphide bond is presented in Figure 3.3.6. In the absence of the disulphide bond two different peaks should be detected in the tryptic digest. These would be detected at 3558.43 Da and 1457.61 Da. In order to confirm the identity of the 5012.97 Da peak, MS/MS analysis was performed. The analysis of the MS/MS spectrum was performed using the MASCOT search engine. The search was performed against a database that contained only the

sequence of the recombinant SHPrP 90-231 and its purification tag. Because MASCOT assumes that disulphide bonds are always reduced and/or blocked using reagents such as IAA, the search was submitted by choosing 3558.43 Da as the mass of the parent ion instead of 5012.97 Da. A high mascot score of 67 for the MS/MS results confirmed that the peptide detected at 5012.97 Da is indeed the disulphide containing peptide. Almost a full ladder of b-ions, Table 3.3.1, has been detected starting after the cross-linked Cys residue. The MS/MS spectrum of the 5012.97 Da tryptic peptide is shown in Figure 3.3.7.

3558.6859 Da ⟵—— YPNQVYYRPVDQYNNQNNFVHDCVNITIK ⎤
                                              | ⎬ 5013.3384 Da
                    1457.6763 Da ⟵—— VVEQMCTTQYQK ⎦
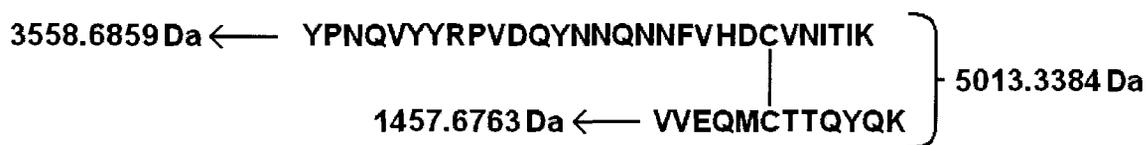
Figure 3.3.6 Sequence of the disulphide containing tryptic peptide
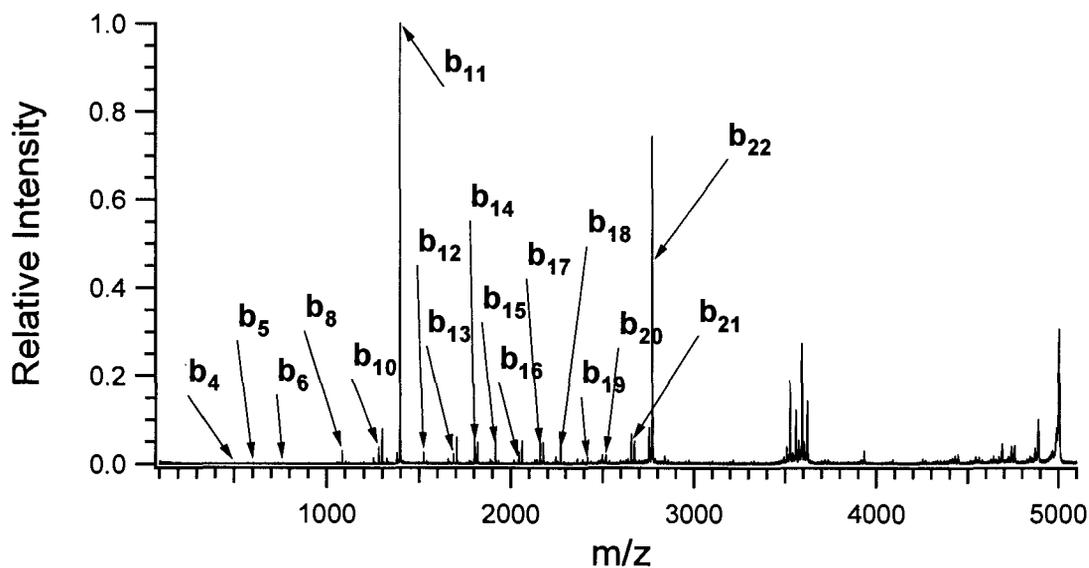
Figure 3.3.7 MS/MS spectrum of the disulphide containing tryptic peptide

Table 3.3.1 b- and y-ion sequence of the 5012.97 Da tryptic peptide. Identified ions in the mass spectrum are shown in bold red.

#	b	Seq.	y	#		#	b	Seq.	y	#
1	164.07	Y		29		15	1914.87	N	1758.85	15
2	261.12	P	3395.62	28		16	2042.93	Q	1644.81	14
3	375.16	N	3298.57	27		17	2156.97	N	1516.75	13
4	503.22	Q	3184.52	26		18	2271.02	N	1402.71	12
5	602.29	V	3056.46	25		19	2418.09	F	1288.67	11
6	765.35	Y	2957.40	24		20	2517.15	V	1141.60	10
7	928.41	Y	2794.33	23		21	2654.21	H	1042.53	9
8	1084.52	R	2631.27	22		22	2769.24	D	905.47	8
9	1181.5	P	2475.17	21		23	2872.25	C	790.44	7
10	1280.64	V	2378.11	20		24	2971.32	V	687.44	6
11	1395.66	D	2279.05	19		25	3085.36	N	588.37	5
12	1523.72	Q	2164.02	18		26	3198.44	I	474.32	4
13	1686.79	Y	2035.96	17		27	3299.49	T	361.24	3
14	1800.83	N	1872.90	16		28	3412.58	I	260.19	2
						29		K	147.11	1

Three other peaks that could possibly contain the reduced Cys residues were also detected. The peaks detected at 1457.61 Da and 3558.43 Da were identified by MS/MS as being Cys containing tryptic peptides. The MS/MS spectra and the corresponding tables with the identified fragment ions are presented in Figure 3.3.8, Table 3.3.2, and Figure 3.3.9, Table 3.3.3, respectively.
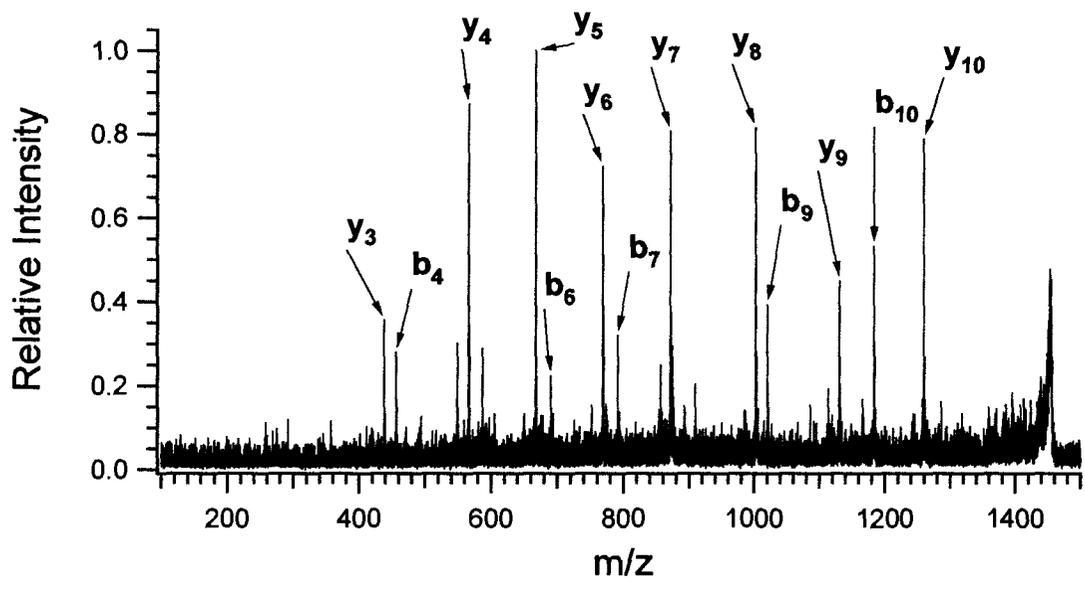
Figure 3.3.8 MS/MS spectrum of VVEQMCTTQYQK tryptic peptide

Table 3.3.2 b- and y-ion sequence for the VVEQMCTTQYQK tryptic peptide. Identified ions in the mass spectrum are shown in bold red

#	b	Seq.	y	#
1	100.07	V		12
2	199.14	V	1358.60	11
3	328.18	E	1259.54	10
4	456.24	Q	1130.49	9
5	587.28	M	1002.43	8
6	690.29	C	871.39	7
7	791.34	T	768.38	6
8	892.39	T	667.34	5
9	1020.44	Q	566.29	4
10	1183.51	Y	438.23	3
11	1311.57	Q	275.17	2
12		K	147.11	1
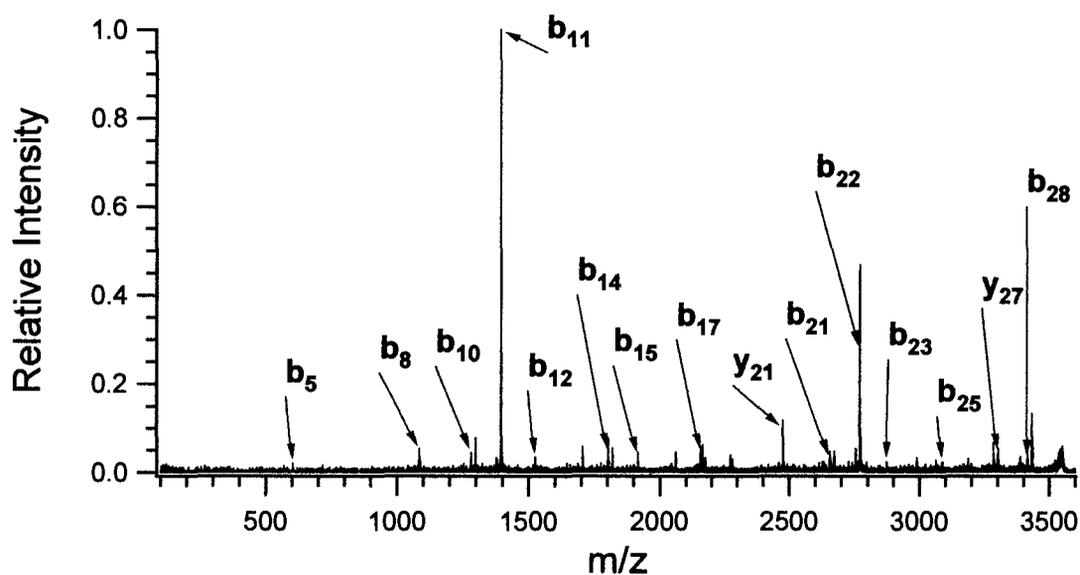
Figure 3.3.9 MS/MS spectrum of YPNQVYYRPVDQYNNQNNFVHDCVNITIK tryptic peptide

Table 3.3.3 b- and y-ion sequence of YPNQVYYRPVDQYNNQNNFVHDCVNITIK tryptic peptide. Identified ions in the mass spectrum are shown in bold red

#	b	Seq.	y	#		#	b	Seq.	y	#
1	164.07	Y		29		16	2042.93	Q	1644.81	14
2	261.12	P	3395.62	28		17	2156.97	N	1516.75	13
3	375.16	N	3298.56	27		18	2271.02	N	1402.71	12
4	503.22	Q	3184.52	26		19	2418.08	F	1288.67	11
5	602.29	V	3056.46	25		20	2517.15	V	1141.60	10
6	765.35	Y	2957.39	24		21	2654.21	H	1042.53	9
7	928.41	Y	2794.33	23		22	2769.24	D	905.47	8
8	1084.52	R	2631.27	22		23	2872.25	C	790.44	7
9	1181.57	P	2475.17	21		24	2971.32	V	687.44	6
10	1280.64	V	2378.11	20		25	3085.36	N	588.37	5
11	1395.66	D	2279.05	19		26	3198.44	I	474.32	4
12	1523.72	Q	2164.02	18		27	3299.49	T	361.24	3
13	1686.79	Y	2035.96	17		28	3412.58	I	260.19	2
14	1800.83	N	1872.90	16		29		K	147.11	1
15	1914.87	N	1758.85	15						

The peak detected at 2474.98 Da could match two different tryptic peptides. The first peptide could be GSHMLEGQGGGTHNQWNKPSKPK and would be the result of a tryptic digestion with no miscleavages. The second peptide could be PVDQYNNQNNFVHDCVNITIK and would be the result of a tryptic digestion with one miscleavage. For these two peptides only the second peptide contains a Cys residue. The MS/MS spectrum shown in Figure 3.3.10 shows an almost complete b- and y-ion

coverage and resulted in a high MASCOT ion score of 145, confirming that the fragments correspond to the tryptic peptide lacking the Cys residue. Table 3.3.4 shows all the theoretical b-and y-ions along with the detected (bold red) ions by MS/MS. The detected b- and y-ions in the MS/MS spectrum of the 2474.98 Da peptide were obtained using the 4800 Plus MALDI-TOF/TOF Analyzer in a reflectron mode of operation.
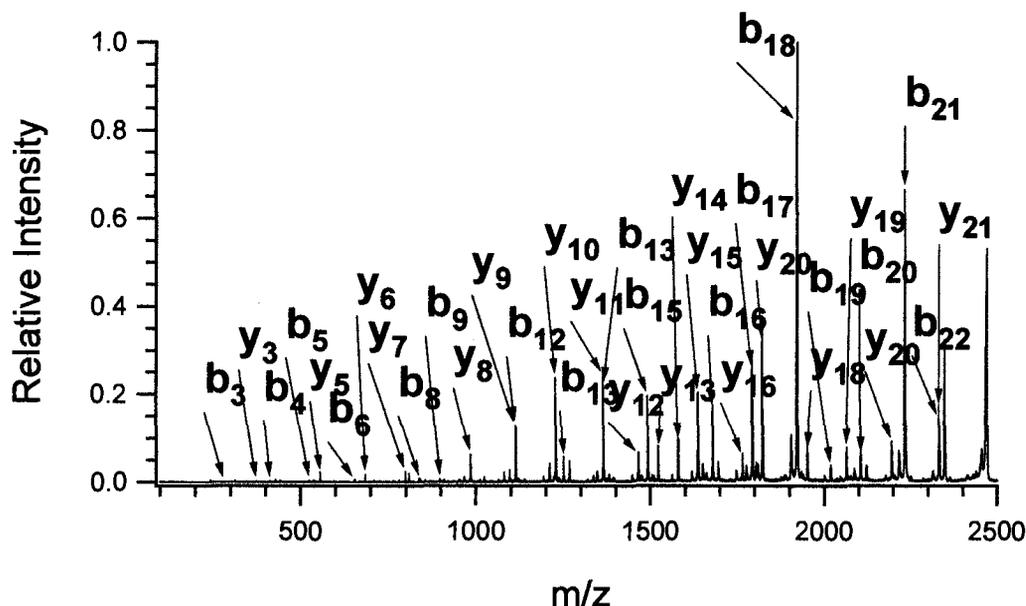
Figure 3.3.10. MS/MS spectrum of GSHMLEGQGGGTHNQWNKPSKPK tryptic peptide.

Table 3.3.4. b- and y-ion sequence of GSHMLEGQGGGTHNQWNKPSKPK tryptic peptide.
Identified ions in the mass spectrum are shown in bold red

#	b	Seq.	y	#		#	b	Seq.	y	#
1	58.02	G		23		13	1249.53	H	1363.72	11
2	145.06	S	2418.17	22		14	1363.58	N	1226.66	10
3	282.11	H	2331.14	21		15	1491.63	Q	1112.62	9
4	413.16	M	2194.08	20		16	1677.71	W	984.56	8
5	526.24	L	2063.04	19		17	1791.76	N	798.48	7
6	655.28	E	1949.95	18		18	1919.85	K	684.44	6
7	712.3083	G	1820.91	17		19	2016.90	P	556.34	5
8	840.36	Q	1763.89	16		20	2103.94	S	459.29	4
9	897.38	G	1635.83	15		21	2232.03	K	372.26	3
10	954.40	G	1578.81	14		22	2329.08	P	244.16	2
11	1011.43	G	1521.79	13		23		K	147.11	1
12	1112.47	T	1464.77	12						

The presence of the two Cys containing peptides suggests that some of the protein could be present in a reduced form. The presence of these peaks could be caused by partial reduction of the disulphide bonds during sample preparation or it can also be an artefact of the MALDI ionization process. The observation of reduction in the mass spectrometric desorption process has been previously reported[12]. However, using DHAP as a matrix only the peak corresponding to the disulphide containing peptide has been detected. No peaks were detected at 1457.61 Da and 3558.43 Da. Figure 3.3.11 shows the regions of the mass spectrum where the Cys containing peaks should be detected.

Based on the above presented data it can be concluded that: 1) the production and purification of the protein was successful; 2) FT-MS and MALDI-TOF-MS of the intact protein indicates that the N-terminal Met has been processed by the *E. coli* host; an expected process when the penultimate residue has a small side chain, such as Gly in this case[13]; and 3) the presence of the disulphide bond has also been detected.
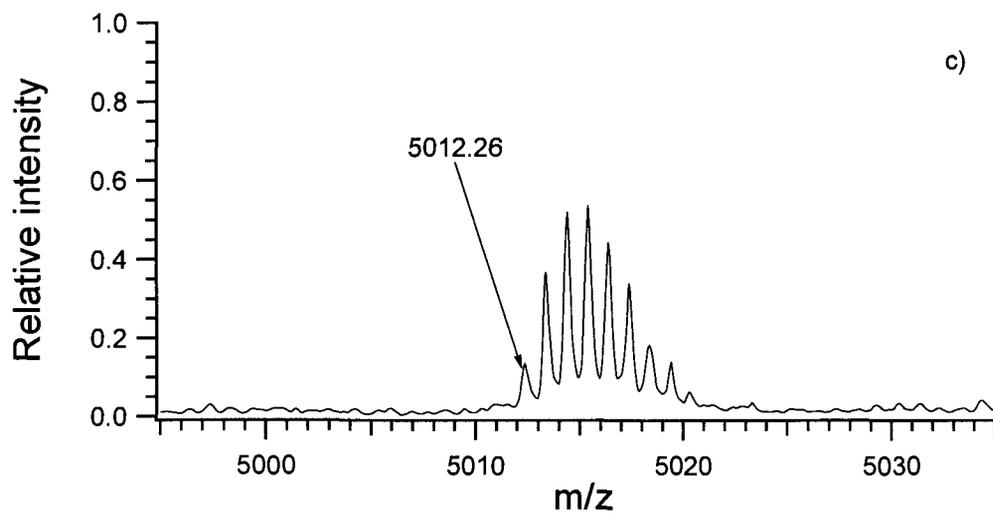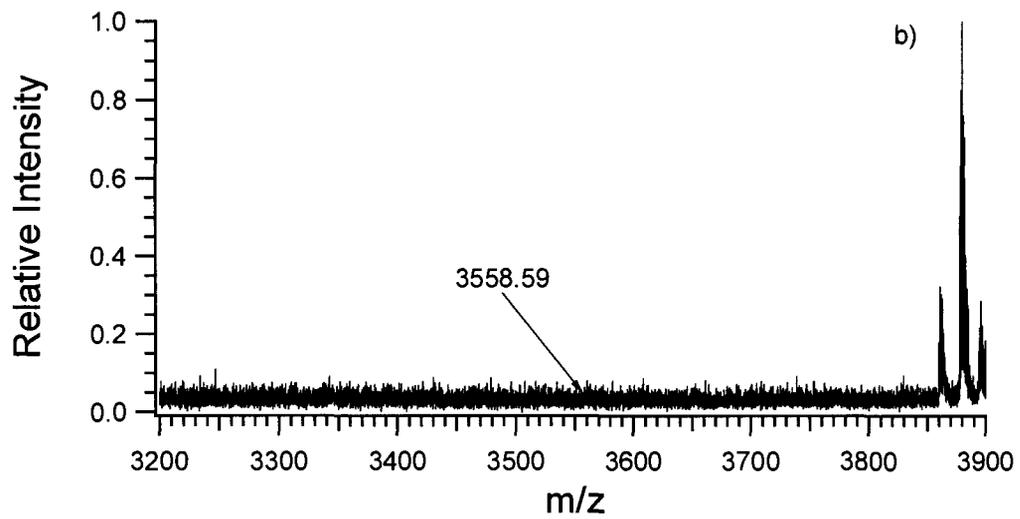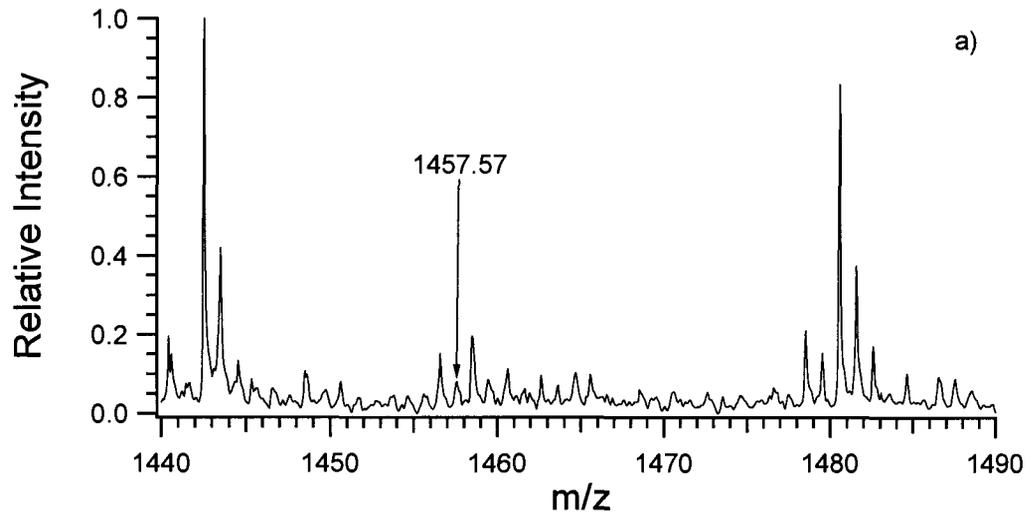
Figure 3.3.11 MS spectra showing the presence or absence of Cys containing peptides: a) 1457.67 Da, b)3558.68 Da and c) 5013.34 Da.

3.3.2. Microwave-Assisted Acid and Base Hydrolysis of the rSHPrP90-231.

MAAH of the rSHPrP90-231 prion construct was performed in 3 M HCl and irradiation times of 30 seconds and 1 minute. For irradiation time of 2 minutes or longer precipitation of the protein has been observed. This was probably caused by long exposure of the protein to a high temperature. Heating of the sample to temperatures higher than 60°C for more than 2 minutes using radiant heat, such as heating the sample on a heating block, has been observed to cause precipitation of the prion protein. Also, exposure of the sample directly to high temperatures (preheated heating block), above 85°C, caused immediate precipitation of the sample in less than 30 seconds. Hydrolysis of the sample by using MAAH under optimum conditions, 3M HCl and 60 seconds irradiation time was possible because of the advantages of this method over conventional methods that use concentrated acid and radiant heat. The advantage is that the hydrolysis in MAAH is assisted by the microwave radiation and not the radiant heat. In the case of the microwave, the radiation increases molecular energy and heating of the sample is performed inside out. Water molecules, because of their dipole, line up with the radiation and start oscillating. The heating of the sample stops immediately when the irradiation is stopped.

For 30 seconds irradiation time, detected peaks were in the lower mass region (less than 7 kDa) and only a few peaks in the higher mass region (7 to 19 kDa). A very intense peak for the molecular ion was also detected for 30 seconds irradiation time which shows that the hydrolysis did not reach the optimal conditions. The mass spectrum of the hydrolysate for the 60 second irradiation time has shown peaks along a very wide mass range with good signal-to-noise ratios. A peak for the molecular ion still could be detected which has been reported[5] to be important for optimum MAAH conditions. This shows that the hydrolysis of the rShPrP90-231 in 3 M HCl and 60 seconds irradiation time is the optimum condition for the hydrolysis of this prion construct. The MALDI-MS spectra for 30 and 60 seconds irradiation times for the region between 10 – 19 kDa are shown in Figure 3.3.12.
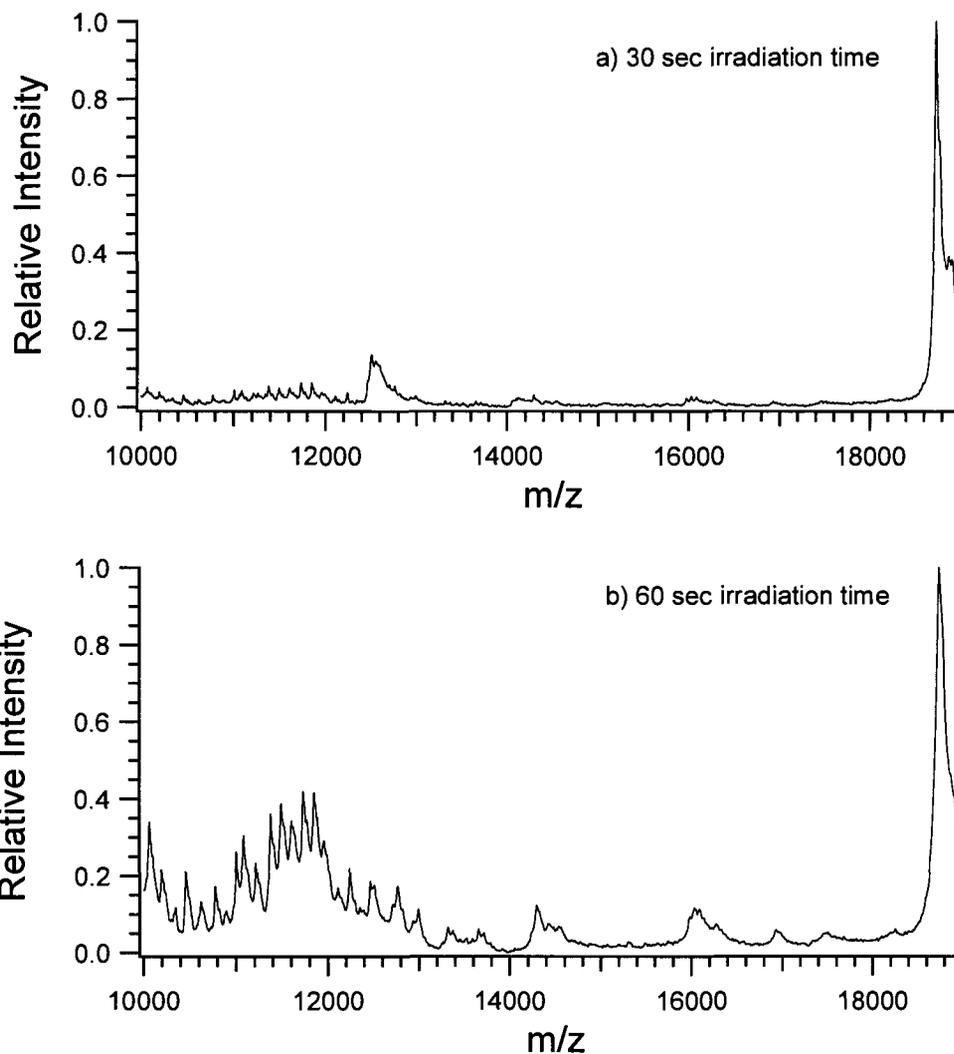
Figure 3.3.12 MALDI-TOF-MS spectra of the rSHPrP90-231 in 3M HCl and a) 30 sec and b)60 sec irradiation times.

Figure 3.3.13 shows the expanded MALDI-TOF-MS spectra of the rSHPrP90-231 hydrolysate after 60 seconds irradiation time in 3M HCl followed by the reduction of the disulphide bond. The theoretical masses along with the identified peptide masses are shown in Table 3.3.5. The mass spectra consist of peaks corresponding to both N- and C-terminal polypeptides. MALDI-TOF-MS and FT-MS of the intact protein indicated that the N-terminal Met has been processed by the *E. coli* host. This was also confirmed by the sequence ladders given that all peaks corresponding to the N-terminal ladder had a mass matching the mass of polypeptides for which the first amino acid would be Gly. Along with the two polypeptide ladders few internal fragments were also identified.
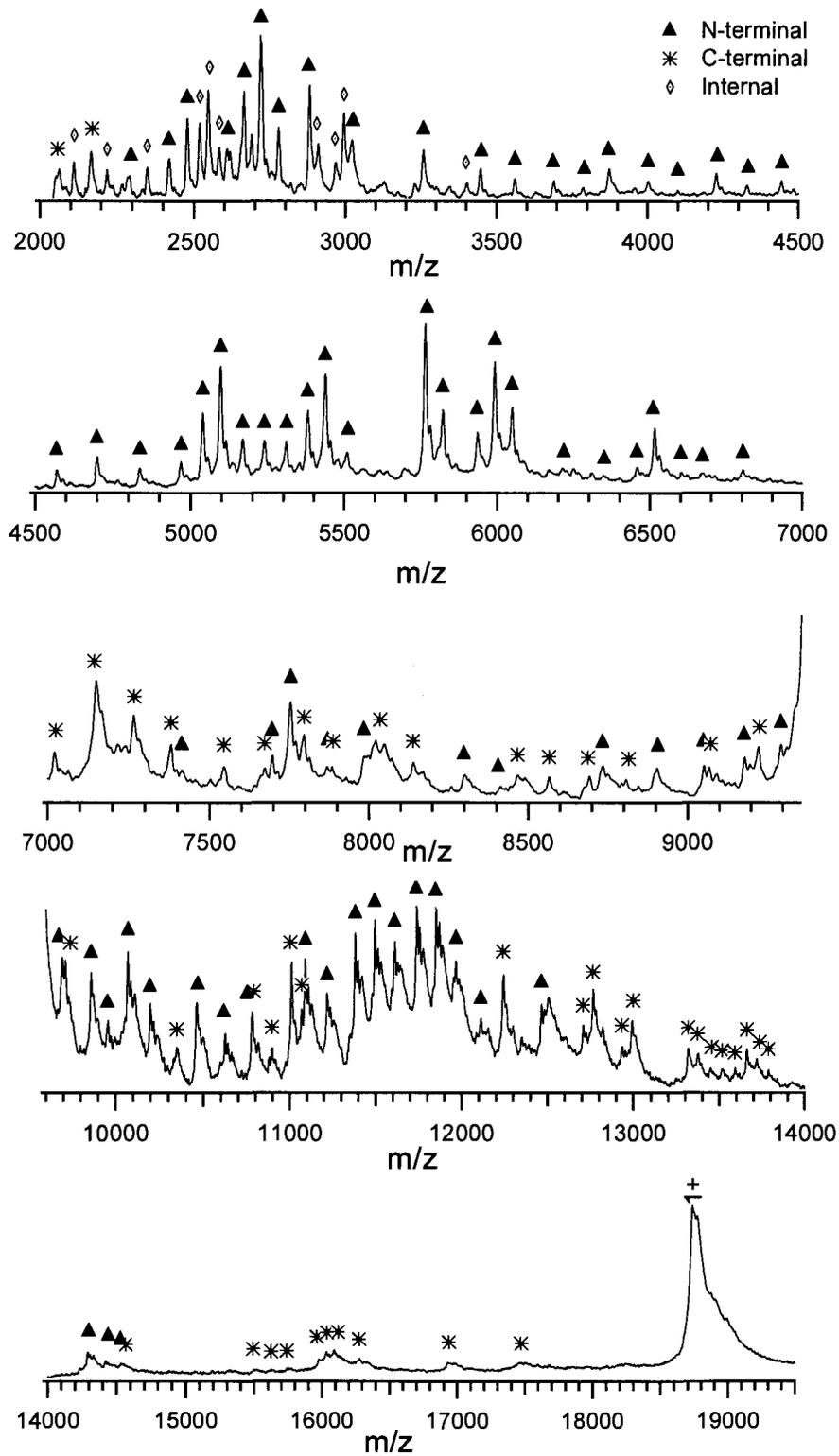
Figure 3.3.13 Expanded MALDI spectra of the rSHPrP90-231 after 60 seconds of irradiation time in 3M HCl followed by DTT reduction.

Table 3.3.5 List of average masses for N- and C-terminal polypeptide ladders detected after MAAH of the rSHPrP90-231. Identified masses shown in bold red. Masses corresponding to the acid resistant core region shown in bold blue.

tag1	76.1	232-tag1	18735.7	tag1-113	5041.3	232-113	13784.9
tag2	163.1	232-tag2	18679.0	tag1-114	5098.3	232-114	13713.9
tag3	250.2	232-tag3	18592.0	tag1-115	5169.4	232-115	13656.8
tag4	387.3	232-tag4	18504.9	tag1-116	5240.5	232-116	13585.7
tag5	524.4	232-tag5	18367.8	tag1-117	5311.6	232-117	13514.7
tag6	661.5	232-tag6	18230.7	tag1-118	5382.6	232-118	13443.6
tag7	798.6	232-tag7	18093.6	tag1-119	5439.7	232-119	13372.5
tag8	935.7	232-tag8	17956.5	tag1-120	5510.8	232-120	13315.4
tag9	1072.8	232-tag9	17819.4	tag1-121	5609.8	232-121	13244.4
tag10	1159.9	232-tag10	17682.3	tag1-122	5708.9	232-122	13145.3
tag11	1247.0	232-tag11	17595.2	tag1-123	5765.9	232-123	13046.2
tag12	1304.0	232-tag12	17508.1	tag1-124	5823.0	232-124	12989.2
tag13	1417.2	232-tag13	17451.1	tag1-125	5936.2	232-125	12932.1
tag14	1516.3	232-tag14	17337.9	tag1-126	5993.2	232-126	12818.9
tag15	1613.4	232-tag15	17238.8	tag1-127	6050.3	232-127	12761.9
tag16	1769.6	232-tag16	17141.7	tag1-128	6213.5	232-128	12704.8
tag17	1826.7	232-tag17	16985.5	tag1-129	6344.7	232-129	12541.6
tag18	1913.8	232-tag18	16928.4	tag1-130	6457.9	232-130	12410.4
tag19	2050.9	232-tag19	16841.4	tag1-131	6514.9	232-131	12297.2
tag20	2182.1	232-tag20	16704.3	tag1-132	6602.0	232-132	12240.2
tag21	2295.3	232-tag21	16573.1	tag1-133	6673.1	232-133	12153.1
tag22	2424.4	232-tag22	16459.9	tag1-134	6804.3	232-134	12082.0
tag1-90	2481.4	232-90	16330.8	tag1-135	6891.4	232-135	11950.8
tag1-91	2609.5	232-91	16273.7	tag1-136	7047.6	232-136	11863.7
tag1-92	2666.6	232-92	16145.6	tag1-137	7144.7	232-137	11707.5
tag1-93	2723.6	232-93	16088.6	tag1-138	7275.9	232-138	11610.4
tag1-94	2780.7	232-94	16031.5	tag1-139	7407.1	232-139	11479.2
tag1-95	2881.8	232-95	15974.5	tag1-140	7544.2	232-140	11348.0
tag1-96	3018.9	232-96	15873.4	tag1-141	7691.4	232-141	11210.9
tag1-97	3133.0	232-97	15736.3	tag1-142	7748.5	232-142	11063.7
tag1-98	3261.1	232-98	15622.2	tag1-143	7862.6	232-143	11006.7
tag1-99	3447.3	232-99	15494.1	tag1-144	7977.7	232-144	10892.6
tag1-100	3561.4	232-100	15307.9	tag1-145	8163.9	232-145	10777.5
tag1-101	3689.6	232-101	15193.8	tag1-146	8293.0	232-146	10591.3
tag1-102	3786.7	232-102	15065.6	tag1-147	8408.1	232-147	10462.2
tag1-103	3873.8	232-103	14968.4	tag1-148	8564.3	232-148	10347.1
tag1-104	4002.0	232-104	14881.4	tag1-149	8727.5	232-149	10190.9
tag1-105	4099.1	232-105	14753.2	tag1-150	8890.7	232-150	10027.7
tag1-106	4227.3	232-106	14656.0	tag1-151	9046.9	232-151	9864.5
tag1-107	4328.4	232-107	14527.8	tag1-152	9176.0	232-152	9708.3
tag1-108	4442.5	232-108	14426.7	tag1-153	9290.1	232-153	9579.2
tag1-109	4573.7	232-109	14312.6	tag1-154	9421.3	232-154	9465.1
tag1-110	4701.9	232-110	14181.4	tag1-155	9535.4	232-155	9333.9
tag1-111	4839.0	232-111	14053.2	tag1-156	9691.6	232-156	9219.8
tag1-112	4970.2	232-112	13916.1	tag1-157	9854.8	232-157	9063.6

tag1-158	9951.9	232-158	8900.4	tag1-195	14288.3	232-195	4523.9
tag1-159	10066.0	232-159	8803.3	tag1-196	14417.3	232-196	4466.8
tag1-160	10194.1	232-160	8689.2	tag1-197	14531.4	232-197	4337.7
tag1-161	10293.1	232-161	8561.1	tag1-198	14678.6	232-198	4223.6
tag1-162	10456.4	232-162	8462.0	tag1-199	14779.7	232-199	4076.4
tag1-163	10619.6	232-163	8298.8	tag1-200	14908.8	232-200	3975.3
tag1-164	10775.8	232-164	8135.6	tag1-201	15009.9	232-201	3846.2
tag1-165	10872.9	232-165	7979.4	tag1-202	15125.0	232-202	3745.1
tag1-166	10971.9	232-166	7882.3	tag1-203	15238.2	232-203	3630.0
tag1-167	11087.0	232-167	7783.2	tag1-204	15366.4	232-204	3516.8
tag1-168	11215.1	232-168	7668.1	tag1-205	15479.6	232-205	3388.6
tag1-169	11378.3	232-169	7540.0	tag1-206	15610.8	232-206	3275.4
tag1-170	11492.4	232-170	7376.8	tag1-207	15739.9	232-207	3144.2
tag1-171	11606.5	232-171	7262.7	tag1-208	15896.1	232-208	3015.1
tag1-172	11734.6	232-172	7148.6	tag1-209	15995.2	232-209	2858.9
tag1-173	11848.7	232-173	7020.5	tag1-210	16094.3	232-210	2759.9
tag1-174	11962.8	232-174	6906.4	tag1-211	16223.4	232-211	2660.8
tag1-175	12110.0	232-175	6792.3	tag1-212	16351.5	232-212	2531.7
tag1-176	12209.1	232-176	6645.1	tag1-213	16482.7	232-213	2403.6
tag1-177	12346.2	232-177	6546.0	tag1-214	16585.8	232-214	2272.4
tag1-178	12461.3	232-178	6408.9	tag1-215	16686.9	232-215	2169.3
tag1-179	12564.4	232-179	6293.8	tag1-216	16788.0	232-216	2068.2
tag1-180	12663.5	232-180	6190.7	tag1-217	16916.1	232-217	1967.1
tag1-181	12777.6	232-181	6091.7	tag1-218	17079.3	232-218	1839.0
tag1-182	12890.8	232-182	5977.6	tag1-219	17207.4	232-219	1675.8
tag1-183	12991.9	232-183	5864.4	tag1-220	17335.6	232-220	1547.7
tag1-184	13105.1	232-184	5763.2	tag1-221	17464.7	232-221	1419.5
tag1-185	13233.3	232-185	5650.1	tag1-222	17551.8	232-222	1290.4
tag1-186	13361.4	232-186	5521.8	tag1-223	17679.9	232-223	1203.3
tag1-187	13498.5	232-187	5393.7	tag1-224	17750.9	232-224	1075.2
tag1-188	13599.6	232-188	5256.6	tag1-225	17914.1	232-225	1004.1
tag1-189	13698.6	232-189	5155.5	tag1-226	18077.3	232-226	840.9
tag1-190	13799.7	232-190	5056.5	tag1-227	18192.4	232-227	677.7
tag1-191	13900.8	232-191	4955.4	tag1-228	18249.5	232-228	562.6
tag1-192	14001.9	232-192	4854.3	tag1-229	18405.7	232-229	505.6
tag1-193	14103.0	232-193	4753.2	tag1-230	18561.9	232-230	349.4
tag1-194	14231.2	232-194	4652.1	tag1-231	18649.0	232-231	193.2
				tag1-232	18735.7	232-232	106.1

Analysis of the spectra shows an N-terminal ladder sequence between 1 kDa up to 12.5 kDa. Two more cleavage sites were detected between 14.2 kDa and 14.5 kDa. No cleavage sites were detected between 12.5 kDa to 14.2 kDa and for the region higher than 14.5 kDa. Peaks corresponding to the C-terminal ladder sequence were seen in the

low mass range only between 1.0 kDa and 2.5 kDa and in the higher mass range starting from 7 kDa. No peaks corresponding to the N-terminal polypeptide ladder were observed between 2.5 kDa and 7 kDa corresponding to the C-terminal ladder. From the overlap of the hydrolysis sites, as seen in Figure 3.3.14, it was observed that the N-terminal part of the construct was readily hydrolysable. There was also a small portion of the C-terminal side of the prion construct that was readily hydrolysable. With the exception of the three cleavage sites between 195G and 198F no other cleavage sites were detected in the region between 179C and 212Q. This data shows a hydrolysis resistant core region in the rSHPrP90-231 prion construct.
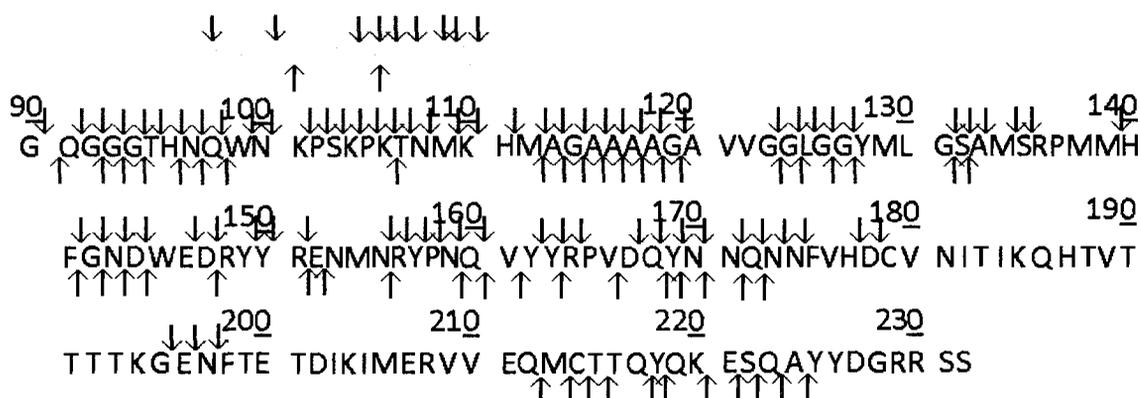
Figure 3.3.14 Peptide bond hydrolysis map of rSHPrP90-231 in 3M HCl and 60 seconds irradiation time followed by DTT reduction of the disulphide bond. ↓ - hydrolysis site identified by the detection of the N-terminal peptide; ↑ - hydrolysis site identified by the detection of the C-terminal peptide. N-terminal purification tag shown in green.

Microwave-assisted base hydrolysis of the rSHPrP90-231 was performed in 1M NaOH for 30, 60 and 90 seconds irradiation times. The optimum irradiation time was found to be 30 to 60 seconds with spectra containing peaks across a wide mass range and good signal-to-noise ratio and the presence of protein molecular ion peak. For 90 seconds irradiation time the mass spectra consisted of mostly lower mass peaks and almost complete disappearance of the intact protein peak. The mass spectra for 30, 60 and 90 seconds irradiation times are shown in Figure 3.3.16.

The peptide bond hydrolysis map for 60 seconds irradiation time is presented in Figure 3.3.15. As already described in Chapter 2 the MABH is not as capable of generating a full polypeptide ladder that is easy to sequence as is the case of MAAH, but

the interesting observation that can be made from this data is that, even though the hydrolysis is "partial", as expected, the base was able to hydrolyze the region that was resistant to acid hydrolysis. This also shows that the lack of peaks in the case of the acid hydrolysis was not caused by the desalting process or ion suppresion effects in MALDI. In other words, if the peptides corresponding to the acid hydrolysis resistant region of the protein were present in the acid hydrolysate, they should be detectable by MALDI.
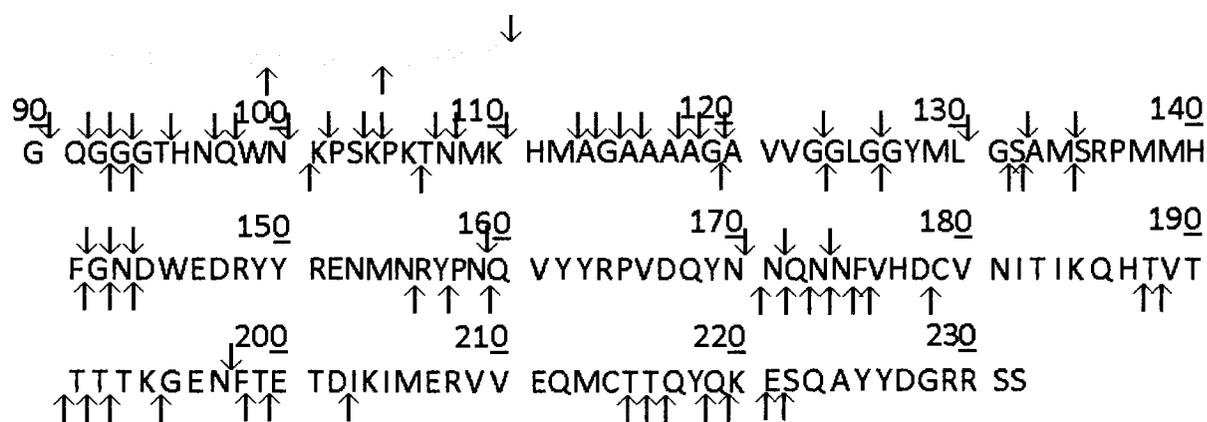
Figure 3.3.15 Peptide bond hydrolysis map of rSHPrP90-231 in 1M NaOH and 60 seconds irradiation time. ↓ - hydrolysis site identified by the detection of the N-terminal peptide; ↑ - hydrolysis site identified by the detection of the C-terminal peptide. N-terminal purification tag shown in green.
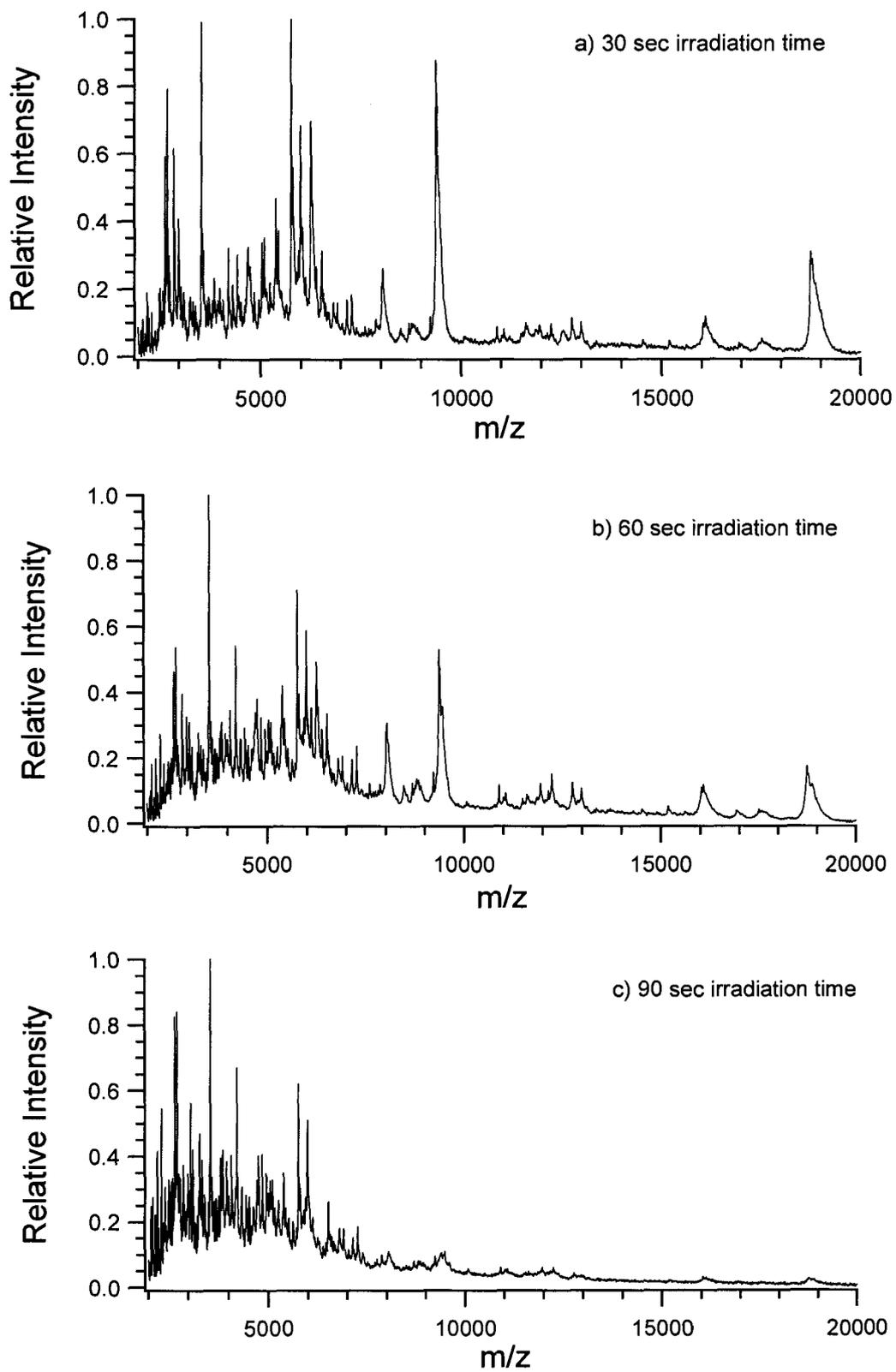
Figure 3.3.16 MALDI-TOF-MS spectra of the rSHPrP90-231 in 1M NaOH and a) 30 sec, b) 60 sec and   c) 90 sec irradiation times

These results are surprising because the rSHPrP^{C}90-231 is easily digested by proteases. It has been previously shown that MAAH is not only capable of hydrolysing soluble proteins but it was successfully used to degrade proteins which are hard to digest using enzymatic methods, such as bacteriorhodopsin or hydrophobic membrane proteins[14]. This acid hydrolysis resistant region was found to contain the sequence between the 2 disulphides that are linked together, although based on the data obtained for the hydrolysis of the HEWL presented in Chapter 2 the presence of the disulphide bond does not seem to interfere with the hydrolysis process. This behaviour of the protein is similar to the behaviour of the misfolded conformer. Based on this data the most probable explanation would be that after mixing with the concentrated HCl the protein undergoes a change, such as conformational change and/or oligomerization, which could give the protein hydrolysis resistance. Therefore we have used circular dichroism, dynamic light scattering and polyacrylamide gel electrophoresis to detect any possible changes under the above described conditions.

3.3.3. Circular Dichroism

Circular dichroism was used to address the question whether the pH change that the protein undergoes when mixed with HCl for MAAH or NaOH for MABH induces any structural alterations of the rSHPrP^{C}90-231 prion construct. The time required mixing the samples with the strong acid or base and the acquisition of the first spectrum was about 6 minutes. No differences were found to be between the first and the next 4 consecutive spectra recorded for each sample which suggests that all structural changes happened within this time frame. Secondary structure of the rSHPrP^C determined from NMR data[15] (PDB 1B10) was used as a reference. Figure 3.3.17 shows the CD spectra of the rSHPrP90-231 under different solvent conditions.

25000
20000
15000
10000
5000
0
-5000
-10000
-15000
-20000

Molar Ellipticity (deg*cm²/dmol)

—— rSHPrP90-231, 50 mM NaAc
·········· rSHPrP90-231, 3M HCl
– – – rSHPrP90-231, 1M NaOH

190   200   210   220   230   240   250

Wavelength (nm)
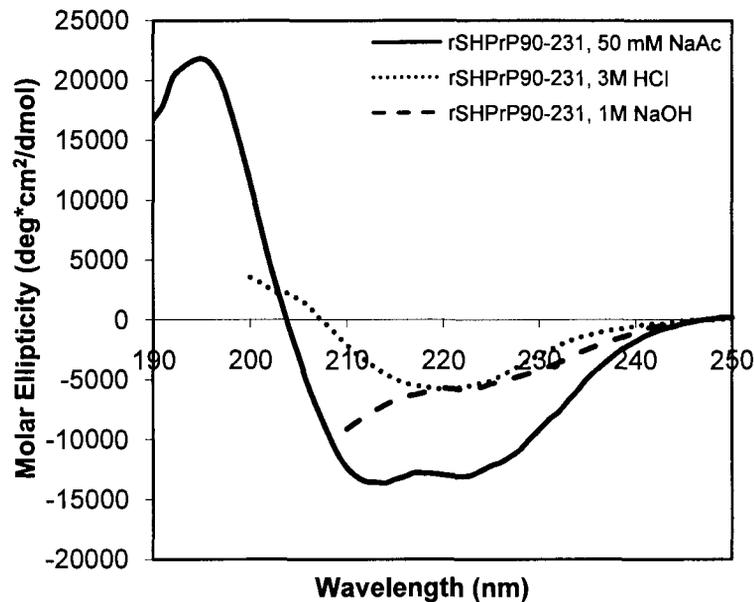
Figure 3.3.17. CD spectra of rSHPrP90-231 prion construct recorded:
(–) at pH 5.5 in 50 mM sodium acetate; (···) 3M HCl and (--) 1M NaOH.

The CD spectrum of the rShPrP^{C}90-231 in 50 mM NaAc, pH 5.5 resembles the spectrum of a typical α-helical protein. It shows a minimum at 208 nm and a shoulder at 222 nm, clearly indicating that the protein contains one or more α-helices. For estimating the α-helical content of proteins, the α-helical content can be determined from the ellipticity at 222 nm. At this wavelength the minimum the rotational strength of an amino acid polymer composed completely of α-helices varies depending on the helix length. The results of the deconvolution for this spectrum and the NMR reference data are in a relatively good agreement. The results obtained by the deconvolution of the circular dichroism spectra are shown in Table 3.3.6. The table shows that all 3 programs, using any of the reference data set, predict a similar secondary structure. The results obtained for the α-helical content were in a very good agreement with the reference NMR data. In the case of the β-sheet content the programs systematically overestimated the value. The overestimation of the β-sheet content by CD of the recombinant SHPrP^{C}90-231 has been previously reported[16]. A possible explanation to this overestimation is that to date all deconvolution programs fail to predict an accurate

value for the β-sheet content in the case of proteins which are mostly helical and contain only a small percentage of the β-sheet conformer.

Table 3.3.6. Results of the deconvolution of CD spectra for rSHPrP^{C}90-231 in 50 mM NaAc, pH 5.5.

NMR	pH=5.2	43	6	15	34
Program	Ref. Set	Helix (%)	Sheet (%)	Turn (%)	Random (%)
CDSSTR	SP43	46	14	18	23
	SDP48	47	14	20	20
	SMP56	45	13	19	23
CONTINLL	SP43	38	18	20	24
	SDP48	39	19	20	23
	SMP56	43	15	21	22
SELCON3	SP43	41	15	19	28
	SDP48	41	15	19	28
	SMP56	44	14	19	27
Average±SD	all	42±3	15±2	19±1	24±3

The CD spectrum for the rSHPrP90-231 under acidic condition (3M HCl) resembles a typical spectrum of a protein in a β-sheet conformation with a single minimum at 218 nm. The deconvolution results, presented in Table 3.3.7, also estimate a significant increase in the β-sheet content from about 15% predicted at pH 5.5 to about 37% under acidic conditions. The increase in the β-sheet content is also accompanied by the loss of the helical content which decreases from 42% at pH 5.5 to a negligible 3% under acidic conditions. This is a clear indication of a conversion of the protease sensitive rSHPrP^C to a PrP^β-form, rich in β-sheet content and lacking helical content, which is a common feature of the protease resistant form of the prion protein[17].

Table 3.3.7 Results of the deconvolution of the CD spectra for rSHPrP^{C}90-231 in 3M HCl.

Program	Ref. Set	Helix	Sheet	Turn	Random
CDSSTR	SP43	3	39	24	33
	SDP48	3	36	23	36
	SMP56	3	40	23	32
CONTINLL	SP43	5	38	23	34
	SDP48	5	37	21	37
	SMP56	6	39	22	33
SELCON3	SP43	5	35	22	39
	SDP48	5	35	22	39
	SMP56	5	36	24	42
Average±SD	all	4±1	37±2	23±1	36±3

The CD spectrum under basic conditions shows a decrease in signal intensity in the region around 222 nm indicating the loss of the helical content. In the presence of 1M NaOH it seems to be an increase in the signal intensity going to lower wavelength. Because of the cut-off in 1M NaOH the acquisition of data at lower wavelength was not possible. Therefore an estimation of the secondary structure using this very limited range can be highly inaccurate. In order to be able to record data at lower wavelength a spectrum at pH 13 has been recorded. A 1 µg µL^{-1} sample was basified using minimal amount of concentrated NaOH until a pH of about 13 was reached. The CD spectra of the rSHPrP90-231 prion construct in 1M NaOH and at pH 13 are shown in Figure 3.3.18. At pH 13 the recording of the CD spectra was possible down to 190 nm. From the shape of the curve it can be estimated that the protein probably maintains part of its helical structure and negative signal intensity at around 198 nm suggests that the protein is in a random coil conformation.
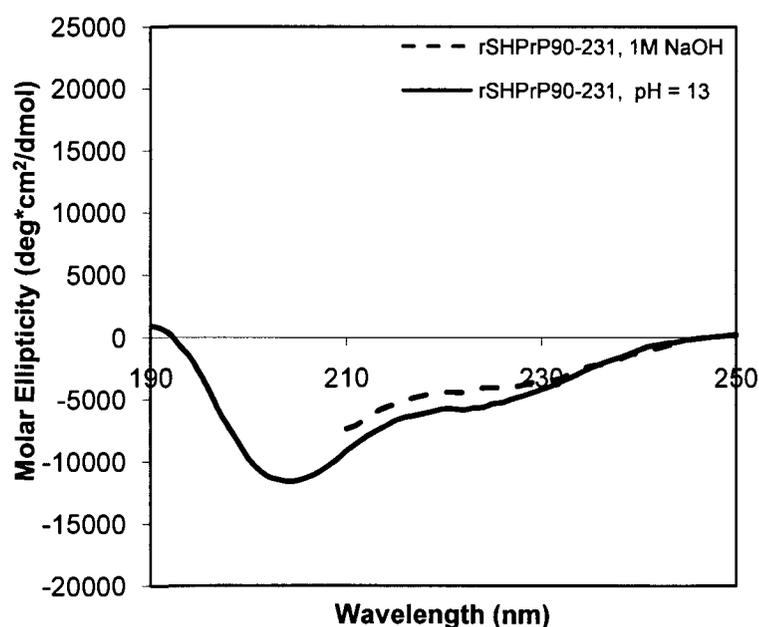
Figure 3.3.18. CD spectra of rSHPrP90-231 prion construct recorded: in (--) 1M NaOH, (–) at pH 13

From Figure 3.3.18, it can be seen that there is no significant difference between the conformation of the protein at pH 13 and in 1M NaOH. Therefore to get a more accurate estimation of the secondary structure the data obtained at pH 13 was used.

The results of the deconvolution of the CD spectrum under basic conditions (pH 13) are shown in Table 3.3.8. It seems that all programs tend to estimate a lot of structural details when using reference sets SP43 and SMP56 even if the spectra shows a negative signal in the 208 nm region which should correspond to random coil. This is most probably caused by the fact that none of these two reference sets contain spectra for denatured protein and the data is fitted to the spectra that has a spectra closest to the one submitted for analysis. When reference set SPD48 is used, which contains spectra for the 5 denatured proteins increased random coil content is predicted. The average value in the table represents the average of only the values predicted using reference set SDP48.

Table 3.3.8. Results of the deconvolution of the CD spectra for rSHPrP^{C}90-231 at pH 13.

Program	Ref. Set	Helix	Sheet	Turn	Random
CDSSTR	SP43	20	24	25	32
	SDP48	**18**	**12**	**14**	**56**
	SMP56	19	25	24	33
CONTINLL	SP43	21	22	24	33
	SDP48	**19**	**11**	**14**	**56**
	SMP56	22	20	24	34
SELCON3	SP43	17	26	23	39
	SDP48	**16**	**17**	**15**	**51**
	SMP56	16	26	22	35
Average±SD	SDP48	**18±2**	**13±3**	**141**	**54±3**

3.3.4. Dynamic Light Scattering

Dynamic light scattering was used to address the question whether the pH change that the protein undergoes when mixed with HCl for MAAH or NaOH for MABH induces any oligomerization of the rSHPrP^{C}90-231 prion construct. Light scattering intensities for 50 mM NaAc buffer, rSHPrP90-231 in 50mM NaAc, pH 5.5 and after mixing with 6M HCl or 2M NaOH are shown in Figure 3.3.19. Samples were not filtered to make sure that in case the sample contains high molecular weight species they would not be filtered out. As shown in this figure a contaminant was present in the 50 mM sodium acetate buffer. It can be seen that in the case of the protein samples only one peak is detected when looking at the size distribution by volume. This shows that the

high molecular weight species detected are most probably contamination from the buffer.

The DLS spectra of the rSHPrP^{C}90-231 at pH 5.5, in 50mM sodium acetate showed a hydrodynamic radius of 2.5nm, for which the estimated molecular weight is 28.7 kDa. This is slightly higher than the expected value of 18.7 kDa but clearly shows that the prion construct is a monomer under this condition. Under acidic conditions a very broad peak with a significantly higher hydrodynamic radius has been detected. This data shows that the sample is fairly polydisperse most probably due to oligomerization or aggregation of the proteins occurred. An average hydrodynamic radius would be about 28 nm. Under basic conditions a significantly smaller increase in the hydrodynamic radius was observed from 2.5 nm at pH 5.5 to 3.9 nm under basic conditions. By assuming a globular structure for the protein the 3.9 nm of the hydrodynamic radius would correspond to a molecular weight of about 81.3 kDa. This suggests that the protein is most probably a tetramer under this condition. Although based on the CD data, which shows a mostly unfolded protein the increase in the hydrodynamic radius can be attributed to the unfolding of the protein, which using Eq. 11 (Chapter 1) would overestimate the MW, as already described in Chapter 2. The hydrodynamic radii of rSHPrP under the different solvent conditions are summarized in Table 3.3.9.

Table 3.3.9. Hydrodynamic radii determined by dynamic light scattering for rSHPrP90-231 in different solvent conditions

rSHPrP90-231	R_H (nm)
25mM sodium acetate, pH 5.5	2.5
3M HCl	28
1M NaOH	3.9
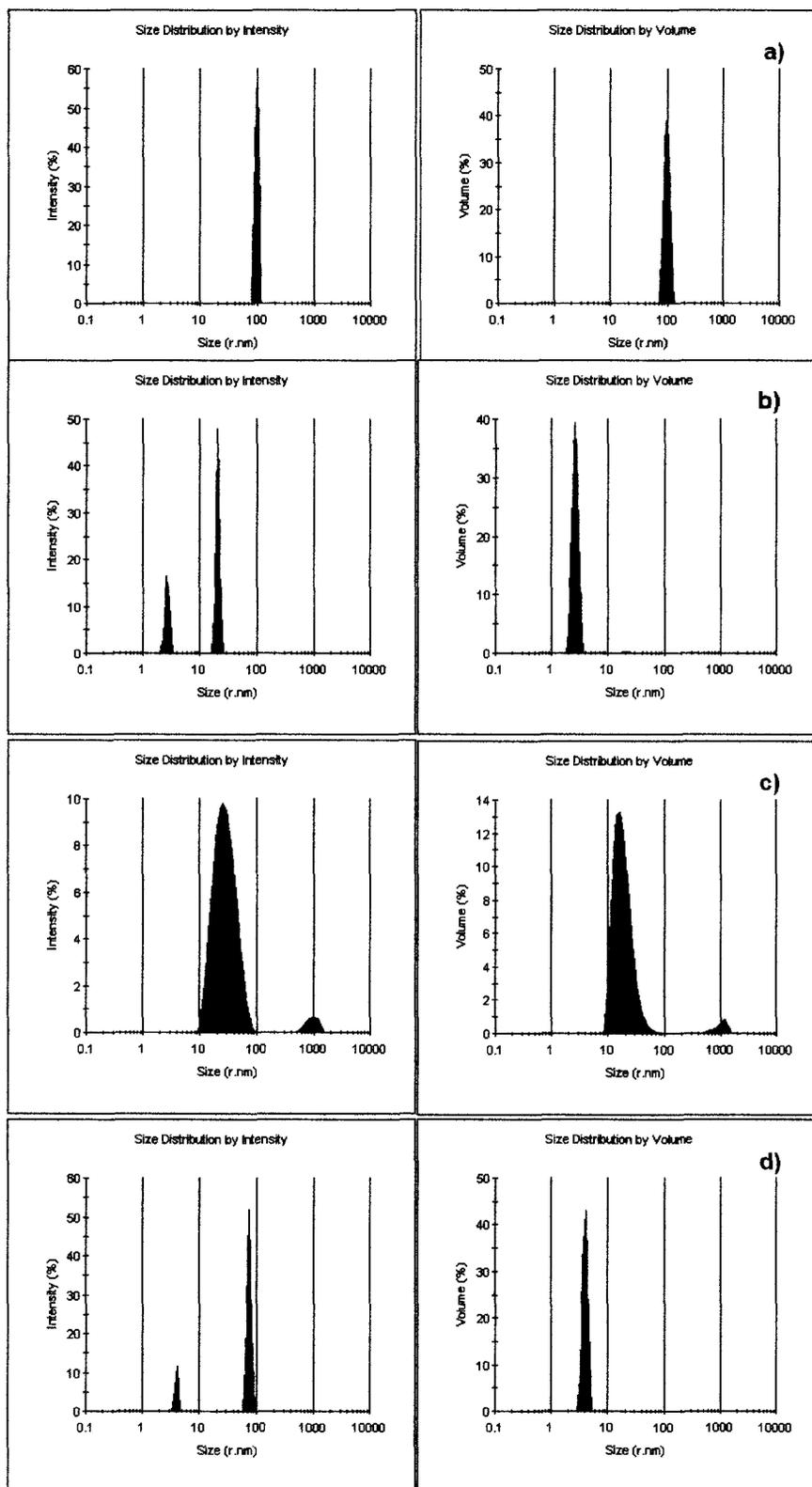
Figure 3.3.19 DLS spectra for: a) 50mM NaAc, pH 5.5, b) rSHPrP^{c}90-231 in 50mM NaAc, pH 5.5, c) rSHPrP^{c}90-231 in 3M HCl and d) rSHPrP^{c}90-231 in 1M NaOH

In summary the above data shows that the truncated form of the Syrian hamster prion protein presents an acid hydrolysis resistant region which suggests a conformational change and an ordered aggregation of the protein. These conclusions were further supported by the CD and DLS data which showed that under acidic condition the rSHPrP^{C}90-231 is undergoing an almost instant conformational change which also leads to or allows the formation of large aggregates or oligomers.

3.3.5. *Determination of the MW of the converted PrP^C*

In order to have a better understanding of the processes that happen between mixing the prion protein with concentrated HCl and the actual microwave digestion we have tried to determine the molecular weight of the oligomers. According to the DLS data the size of the oligomers is very large but in dynamic light scattering the assumption is made that the particles that give scattering are globular. In the case of oligomers with shapes different from a sphere, such as long fibrilar ones the estimation of the molecular weight, and consequently of the number of monomers forming the oligomers, can be inaccurate. Therefore other methods were used in order to estimate the size of the species in the acid converted prion sample.

The first approach that we used to determine the molecular weight of the converted rShPrP90-231 was SDS-PAGE. Figure 3.3.20 shows the SDS-PAGE results obtained for the rSHPrP90-231 kept under different solvent conditions. For the sample loaded on Lane 2 the prion protein was converted by acidification. Before loading, the sample was boiled in SDS containing loading buffer. Bands were been observed between approximately 125 and 170 kDa along with a dimer and monomer band. The high molecular weight band would correspond to a series of oligomers containing 7 to 10 monomeric units. The sample loaded onto Lane 3 was basified using concentrated NaOH and boiled in SDS containing loading buffer. No sign of oligomerization was detected. In this case the same bands were observed as in Lane 1 which corresponds to the rSHPrP^{C}90-231 sample boiled in SDS containing loading buffer.
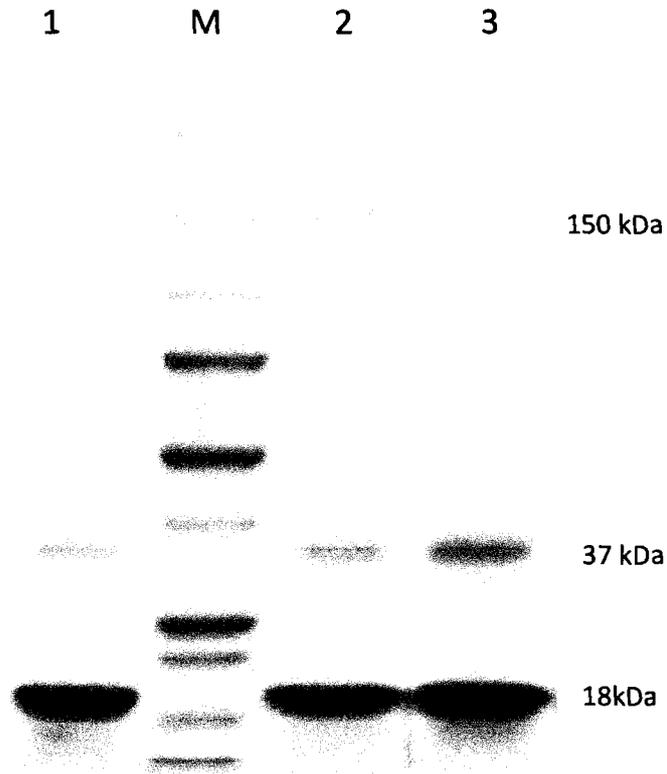
Figure 3.3.20 SDS-PAGE of: Lane 1 rSHPrP^{C}90-231 boiled in sample loading buffer. Lane 2 rSHPrP^{C}90-231 acidified with HCl and boiled in sample loading buffer. Lane 3 rSHPrP^{C}90-231 basified with NaOH and boiled in sample buffer.

While the acid converted prion sample primarily contains oligomeric species, as evidenced by dynamic light scattering (DLS), gel electrophoresis of the sample shows both monomeric and oligomeric forms. The appearance of the strong monomeric bands was attributed to the boiling process of the sample in SDS, which could have caused disaggregation of the oligomers. Disaggregation of fibrilar prion proteins has been previously reported[18]. Silveira *et al.* used fibrilar prion proteins which have been disaggregated by using sodium undecyl sulphate (SUS) in combination with sonication in order to destroy the fibrils. The process generated a wide variety of species ranging from monomers to small and large oligomers.

Because SDS seemed to cause disaggregation of the sample we have tried to run a gel under native conditions. In this case after conversion the sample was mixed with

the native loading buffer, pH 8.3. The pI of our rSHPrP90-231 protein construct is 8.8 which is only 0.5 pH units form the pH of the buffer (TrisHCl, pH 8.3) and it was found that precipitation of the sample occurred under native electrophoresis conditions, the protein sample did not enter the gel and therefore separation was not possible.

In order to overcome the sample precipitation problem we have tried to develop a native gel electrophoresis method under pH conditions under which the protein is stable. The first acidic gel was prepared and run at pH 5.5, the pH at which our protein construct was found to be most stable. At pH 5.5 smearing of the sample have been observed for PrP^C and the acid converted sample did not enter the gel. The pH was further lowered and it was found that the condition under which the prion construct was migrating in the gel was pH 3.8. A series of different proteins were used as molecular weight markers such as cytochrome C, bovine serum albumin, α-macroglobulin, human transferring. Although some proteins had very similar pI's to our prion construct only cytochrome C was found to be migrating under the conditions described above and even cytochrome C showed a very broad band (Lane1), as shown in Figure 3.3.21. Even though it was not possible to estimate the molecular weight from this acidic gel, it can be seen that most of the acidic sample shows up as some forms of oligomers and only a small amount of the sample shows up as an actual monomer (Lane 2). In contrast, the PrP^C showed a single and very intense band for the monomer and no sign of other oligomers have been detected on the gel (Lane 3). A better separation of the high molecular weight species has been achieved by increasing the run time to 120 minutes, as shown in Figure 3.3.21 b).
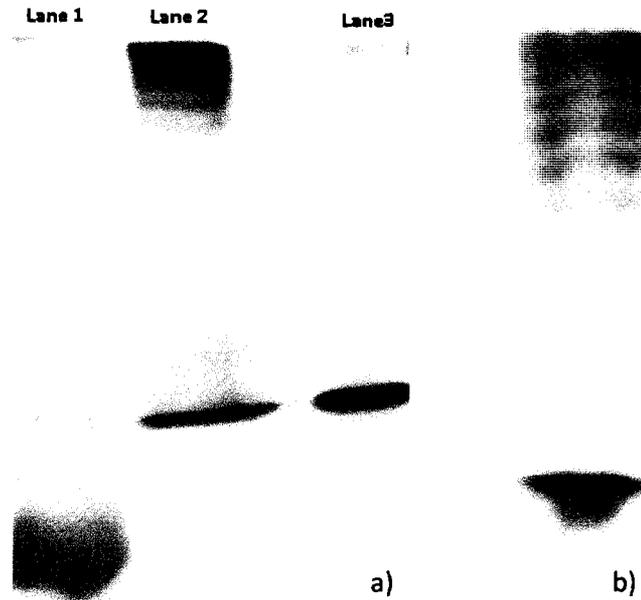
Figure 3.3.21 a) Gel electrophoresis at pH 3.8 and 60 minutes run time of: Lane 1 CYC_HORSE; Lane 2 acid converted rSHPrP^{C}90-231; Lane 3 rSHPrP^{C}90-231. b) Gel electrophoresis at pH 3.8 and 120 minutes run time of the acid converted rSHPrP^{C}90-231.

The last approach to determine the molecular weight of the oligomeric species formed under acidic conditions was by the use of FT-MS. The rSHPrP^{C}90-231 has been converted using concentrated HCl. The sample was dialyzed against 50 mM NaAc buffer, pH 5.5. Figure 3.3.22 shows the CD spectra of the recombinant SHPrP^C at pH 5.5, after acidification and after dialysis. It can be seen that after acidification the prion undergoes a conformational change adopting a high β-sheet containing conformer. After dialysis it seems that the protein is in an intermediate state. Although this might be just an artefact given that CD will show a global measure of all the species present in the sample. Baskakov *et al*[19] has previously shown that mixing different amounts of PrP^C and PrP^β will result in CD spectra similar to the spectrum of homogeneous α-rPrP^C. Therefore it is possible that after dialysis the sample contains a mixture of α-helical monomers and different β-sheet containing oligomers.
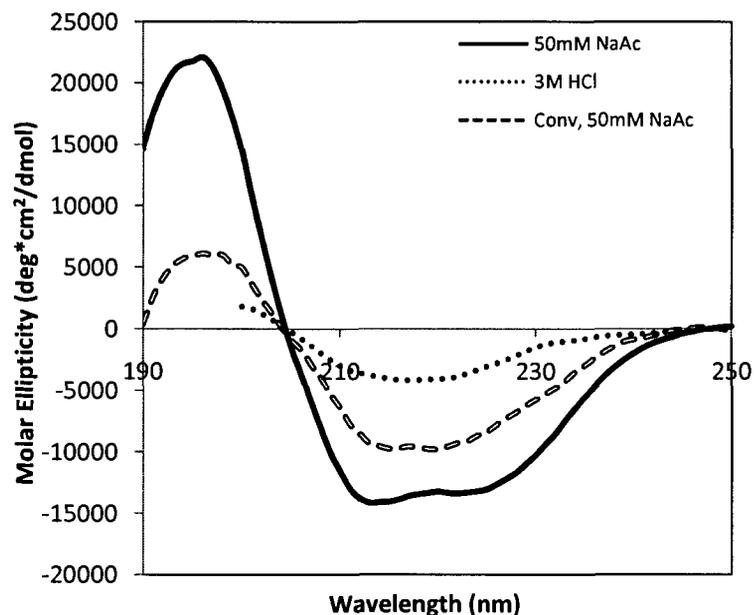
Figure 3.3.22 CD spectra of rSHPrP90-231 prion construct recorded:
(–) at pH 5.5 in 50 mM sodium acetate; (···) after conversion with HCl
and (--) after dialysis into 50mM NaAC, pH 5.5.

In the next step the samples were diluted in 20% acetonitrile/0.1% formic acid and FT-MS was used in order to detect high mass oligomeric species. The mass spectrum of the sample showed only monomeric species. No high mass molecular weight species were detected. The fact that only monomers have been detected could be caused by different processes. One explanation could be that after dialysis all the sample was converted back to monomeric species, although this is unlikely given the stability of these oligomers. A second reason could be that because these oligomers seem to be tightly bound it is unable to pick up enough charge during the ionization process in order to be detected. A third reason could be that the oligomers are disaggregated during the ionization process. Because these oligomers seemed to be very stable at very low pH, we have tried to convert the prion protein using concentrated HCl, followed by the dialysis of the sample against 0.1% TFA. Samples were diluted with 20% ACN/0.1% TFA and injected in the FT-MS. Again, the only detected species corresponded to the monomeric PrP.

3.3.6. MAAH of the acid converted rSHPrP90-231

Conversion of the prion protein under denaturing (3M GdnHCl or 3M urea) and acidic conditions (pH 4.0 and lower) has been previously reported[19-22]. Jackson *et al.*[23] have also reported that aggregation of the sample took place over several hours. The above presented MAAH data was obtained by mixing the recombinant SHPrP90-231 with 6M HCl (3M final) followed immediately by microwave irradiation. In order to check for any differences between the hydrolysis profiles of the rSHPrP and the acid converted rSHPrP due to structural changes or oligomerization, we have performed the MAAH of the acidified rSHPrP90-231 prion construct. Concentrated HCl was added to the sample to lower the pH below 1.0. The sample was kept at low pH for approximately 12 hours. The sample was then mixed with concentrated HCl and MAAH was performed. If any changes would have been detected they would most probably be the result of some slow structural rearrangement and/or stabilization of the oligomers over a period of time. However, an identical hydrolysis profile was observed, with the same acid resistant core region present in both cases. This data suggests that under highly acidic conditions, pH below 1.0, all the structural alteration and oligomerization of the prion protein is happening in a very short time period, immediately after mixing. The peptide bond hydrolysis map for the acid converted rSHPrP90-231s is presented in Figure 3.3.23, along with the expanded MALDI spectrum shown in Figure 3.3.24. The theoretical mass list along with the identified peptide masses are shown in Table 3.3.10.


```
                        ↓↓
           ↑ ↑   ↑↑
90↓ ↓↓↓↓↓↓↓100↓ ↓↓↓↓↓↓↓110↓ ↓↓↓↓↓↓↓120↓ ↓ ↓↓↓↓↓130↓ ↓  ↓↓  140↓↓
G QGGGTHNQWN KPSKPKTNMK HMAGAAAAGA VVGGLGGYML GSAMSRPMMH
↑ ↑↑↑↑↑↑ ↑  ↑   ↑↑     ↑ ↑↑↑↑↑↑↑↑↑ ↑  ↑↑↑↑↑↑   ↑ ↑↑ ↑↑

↓↓↓  ↓↓↓150↓ ↓     160↓  ↓↓↓  ↓↓170↓ ↓↓↓ ↓ ↓↓180      190
FGNDWEDRYY RENMNRYPNQ VYYRPVDQYN NQNNFVHDCV NITIKQHTVT
↑↑↑ ↑ ↑           ↑↑↑↑↑  ↑↑↑    ↑  ↑↑

    ↓↓ 200        210        220        230
TTTKGENFTE TDIKIMERVV EQMCTTQYQK ESQAYYDGRR SS
    ↑↑                  ↑↑↑↑↑    ↑↑↑
```
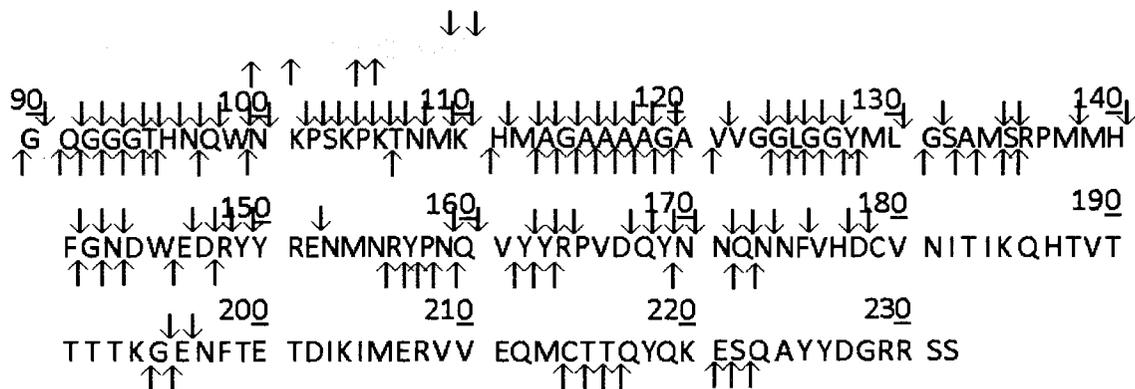
Figure 3.3.23 Peptide bond hydrolysis map of the acid converted rSHPrP90-231 in 3 M HCl and 60 seconds irradiation time followed by DTT reduction of the disulphide bond. ↓ - hydrolysis site identified by the detection of the N-terminal peptide; ↑ - hydrolysis site identified by the detection of the C-terminal peptide. N-terminal purification tag shown in green.

111
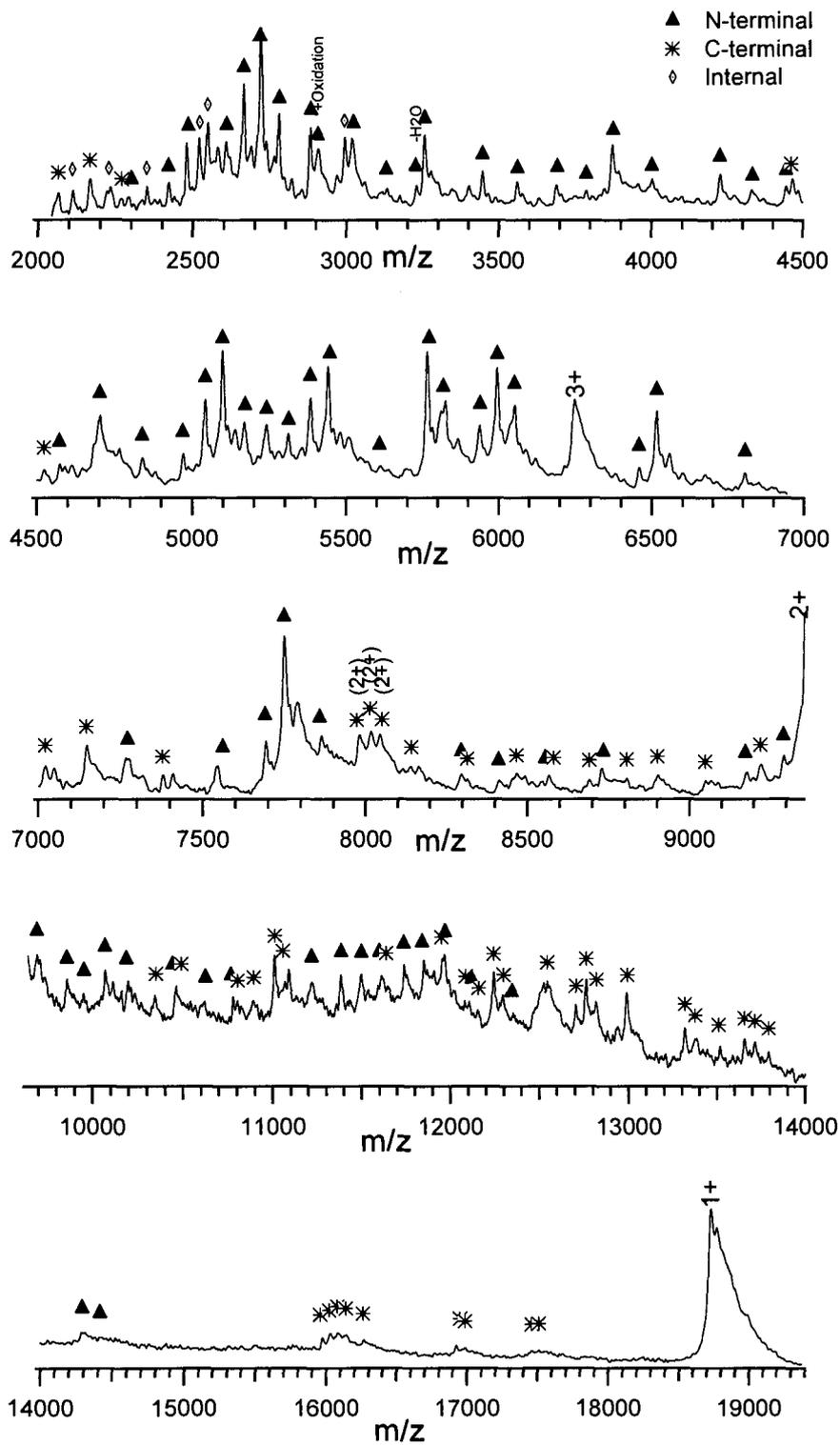
Figure 3.3.24 Expanded MALDI spectra of the acid converted rSHPrP90-231 after 60 seconds of irradiation time in 3M HCL followed by DTT reduction.

Table 3.3.10 List of average masses for N- and C-terminal polypeptide ladders detected after MAAH of the acidified rSHPrP90-231. Identified masses shown in bold red. Masses corresponding to the acid resistant core region shown in bold blue.

tag1	76.1	232-tag1	18735.7	**tag1-113**	**5041.3**	**232-113**	**13784.9**
tag2	163.1	232-tag2	18679.0	**tag1-114**	**5098.3**	**232-114**	**13713.9**
tag3	250.2	232-tag3	18592.0	**tag1-115**	**5169.4**	**232-115**	**13656.8**
tag4	387.3	232-tag4	18504.9	**tag1-116**	**5240.5**	**232-116**	**13585.7**
tag5	524.4	232-tag5	18367.8	**tag1-117**	**5311.6**	**232-117**	**13514.7**
tag6	661.5	232-tag6	18230.7	**tag1-118**	**5382.6**	**232-118**	**13443.6**
tag7	798.6	232-tag7	18093.6	**tag1-119**	**5439.7**	**232-119**	**13372.5**
tag8	935.7	232-tag8	17956.5	**tag1-120**	**5510.8**	**232-120**	**13315.4**
tag9	1072.8	232-tag9	17819.4	**tag1-121**	**5609.8**	**232-121**	**13244.4**
tag10	1159.9	232-tag10	17682.3	**tag1-122**	**5708.9**	232-122	13145.3
tag11	1247.0	232-tag11	17595.2	**tag1-123**	**5765.9**	232-123	13046.2
tag12	1304.0	**232-tag12**	**17508.1**	**tag1-124**	**5823.0**	**232-124**	**12989.2**
tag13	1417.2	**232-tag13**	**17451.1**	**tag1-125**	**5936.2**	**232-125**	**12932.1**
tag14	1516.3	232-tag14	17337.9	**tag1-126**	**5993.2**	**232-126**	**12818.9**
tag15	1613.4	232-tag15	17238.8	**tag1-127**	**6050.3**	**232-127**	**12761.9**
tag16	1769.6	232-tag16	17141.7	tag1-128	6213.5	**232-128**	**12704.8**
tag17	1826.7	**232-tag17**	**16985.5**	tag1-129	6344.7	**232-129**	**12541.6**
tag18	1913.8	**232-tag18**	**16928.4**	**tag1-130**	**6457.9**	232-130	12410.4
tag19	2050.9	232-tag19	16841.4	**tag1-131**	**6514.9**	**232-131**	**12297.2**
tag20	2182.1	232-tag20	16704.3	tag1-132	6602.0	**232-132**	**12240.2**
tag21	**2295.3**	232-tag21	16573.1	tag1-133	6673.1	**232-133**	**12153.1**
tag22	**2424.4**	232-tag22	16459.9	**tag1-134**	**6804.3**	**232-134**	**12082.0**
tag1-90	**2481.4**	**232-90**	**16330.8**	**tag1-135**	**6891.4**	**232-135**	**11950.8**
tag1-91	**2609.5**	**232-91**	**16273.7**	tag1-136	7047.6	232-136	11863.7
tag1-92	**2666.6**	**232-92**	**16145.6**	tag1-137	7144.7	232-137	11707.5
tag1-93	**2723.6**	**232-93**	**16088.6**	**tag1-138**	**7275.9**	232-138	11610.4
tag1-94	**2780.7**	**232-94**	**16031.5**	tag1-139	7407.1	232-139	11479.2
tag1-95	**2881.8**	**232-95**	**15974.5**	**tag1-140**	**7544.2**	232-140	11348.0
tag1-96	**3018.9**	**232-96**	**15873.4**	**tag1-141**	**7691.4**	232-141	11210.9
tag1-97	**3133.0**	232-97	15736.3	**tag1-142**	**7748.5**	**232-142**	**11063.7**
tag1-98	**3261.1**	232-98	15622.2	**tag1-143**	**7862.6**	**232-143**	**11006.7**
tag1-99	**3447.3**	232-99	15494.1	tag1-144	7977.7	**232-144**	**10892.6**
tag1-100	**3561.4**	**232-100**	**15307.9**	tag1-145	8163.9	232-145	10777.5
tag1-101	**3689.6**	232-101	15193.8	**tag1-146**	**8293.0**	**232-146**	**10591.3**
tag1-102	**3786.7**	232-102	15065.6	**tag1-147**	**8408.1**	232-147	10462.2
tag1-103	**3873.8**	232-103	14968.4	**tag1-148**	**8564.3**	**232-148**	**10347.1**
tag1-104	**4002.0**	232-104	14881.4	**tag1-149**	**8727.5**	232-149	10190.9
tag1-105	**4099.1**	232-105	14753.2	tag1-150	8890.7	232-150	10027.7
tag1-106	**4227.3**	**232-106**	**14656.0**	tag1-151	9046.9	232-151	9864.5
tag1-107	**4328.4**	232-107	14527.8	**tag1-152**	**9176.0**	232-152	9708.3
tag1-108	**4442.5**	232-108	14426.7	tag1-153	9290.1	232-153	9579.2
tag1-109	**4573.7**	232-109	14312.6	**tag1-154**	**9421.3**	232-154	9465.1
tag1-110	**4701.9**	232-110	14181.4	tag1-155	9535.4	232-155	9333.9
tag1-111	**4839.0**	**232-111**	**14053.2**	tag1-156	9691.6	**232-156**	**9219.8**
tag1-112	**4970.2**	232-112	13916.1	tag1-157	9854.8	**232-157**	**9063.6**

tag1-158	9951.9	232-158	8900.4	tag1-195	14288.3	232-195	4523.9
tag1-159	10066.0	232-159	8803.3	tag1-196	14417.3	232-196	4466.8
tag1-160	10194.1	232-160	8689.2	tag1-197	14531.4	232-197	4337.7
tag1-161	10293.1	232-161	8561.1	tag1-198	14678.6	232-198	4223.6
tag1-162	10456.4	232-162	8462.0	tag1-199	14779.7	232-199	4076.4
tag1-163	10619.6	232-163	8298.8	tag1-200	14908.8	232-200	3975.3
tag1-164	10775.8	232-164	8135.6	tag1-201	15009.9	232-201	3846.2
tag1-165	10872.9	232-165	7979.4	tag1-202	15125.0	232-202	3745.1
tag1-166	10971.9	232-166	7882.3	tag1-203	15238.2	232-203	3630.0
tag1-167	11087.0	232-167	7783.2	tag1-204	15366.4	232-204	3516.8
tag1-168	11215.1	232-168	7668.1	tag1-205	15479.6	232-205	3388.6
tag1-169	11378.3	232-169	7540.0	tag1-206	15610.8	232-206	3275.4
tag1-170	11492.4	232-170	7376.8	tag1-207	15739.9	232-207	3144.2
tag1-171	11606.5	232-171	7262.7	tag1-208	15896.1	232-208	3015.1
tag1-172	11734.6	232-172	7148.6	tag1-209	15995.2	232-209	2858.9
tag1-173	11848.7	232-173	7020.5	tag1-210	16094.3	232-210	2759.9
tag1-174	11962.8	232-174	6906.4	tag1-211	16223.4	232-211	2660.8
tag1-175	12110.0	232-175	6792.3	tag1-212	16351.5	232-212	2531.7
tag1-176	12209.1	232-176	6645.1	tag1-213	16482.7	232-213	2403.6
tag1-177	12346.2	232-177	6546.0	tag1-214	16585.8	232-214	2272.4
tag1-178	12461.3	232-178	6408.9	tag1-215	16686.9	232-215	2169.3
tag1-179	12564.4	232-179	6293.8	tag1-216	16788.0	232-216	2068.2
tag1-180	12663.5	232-180	6190.7	tag1-217	16916.1	232-217	1967.1
tag1-181	12777.6	232-181	6091.7	tag1-218	17079.3	232-218	1839.0
tag1-182	12890.8	232-182	5977.6	tag1-219	17207.4	232-219	1675.8
tag1-183	12991.9	232-183	5864.4	tag1-220	17335.6	232-220	1547.7
tag1-184	13105.1	232-184	5763.2	tag1-221	17464.7	232-221	1419.5
tag1-185	13233.3	232-185	5650.1	tag1-222	17551.8	232-222	1290.4
tag1-186	13361.4	232-186	5521.8	tag1-223	17679.9	232-223	1203.3
tag1-187	13498.5	232-187	5393.7	tag1-224	17750.9	232-224	1075.2
tag1-188	13599.6	232-188	5256.6	tag1-225	17914.1	232-225	1004.1
tag1-189	13698.6	232-189	5155.5	tag1-226	18077.3	232-226	840.9
tag1-190	13799.7	232-190	5056.5	tag1-227	18192.4	232-227	677.7
tag1-191	13900.8	232-191	4955.4	tag1-228	18249.5	232-228	562.6
tag1-192	14001.9	232-192	4854.3	tag1-229	18405.7	232-229	505.6
tag1-193	14103.0	232-193	4753.2	tag1-230	18561.9	232-230	349.4
tag1-194	14231.2	232-194	4652.1	tag1-231	18649.0	232-231	193.2
				tag1-232	18735.7	232-232	106.1

3.4. *Conclusions*

In the present study, we have investigated the behaviour and hydrolysis profiles of the rSHPrP90-231 prion construct under both acidic and basic conditions. The CD data showed a decrease in the α-helical content accompanied by the increase of the β-sheet content. The DLS data shows that the pH drop induces the oligomerization or aggregation of the protein. Gel electrophoresis run at acidic pH also showed that most of the sample was aggregated or oligomerized. The SDS-PAGE showed that after acidification and boiling in SDS containing loading buffer the sample contained oligomers with a molecular weight range corresponding to the oligomerization of 7 to 10 monomeric units. Therefore the mass spectrum of the hydrolysate generated by the MAAH method corresponds to an oligomerized high β-sheet containing conformer of the prion protein construct. In contrast, under basic conditions no significant decrease accompanied by the increase of sheet was observed by circular dichroism. The DLS data showed the formation of small oligomers under basic conditions. SDS-PAGE showed no significant difference between the rSHPrP^{C}90-231 and the basified sample. This also demonstrated that the small aggregate observed by DLS was only formed by loosely bound monomers.

As already mentioned, the MAAH has been previously reported to be able to generate N- and C-terminal ladders for protein sequencing[5] for a wide variety of proteins with a wide range of pI's and post-translational modifications. In Chapter 2 we showed that the MAAH of the oxidized form of HEWL can take place as long as the peptide bonds are solvent accessible, even in the presence of the stabilizing disulphide bonds. In contrast, the MAAH data of the rSHPrP90-231 prion construct which undergoes a conformational change accompanied by aggregation or oligomerization showed an acid hydrolysis resistant region. A possible explanation of the hydrolysis resistance would be the formation of the self-complementing steric zippers[24]. This motif was recently reported using crystallographic studies on short synthetic peptides. The proposed structure consists of β-sheet pairs stacked on each other by backbone and side-chain hydrogen bonds. However, it has to be noted that the existence of this motif

was only shown on short synthetic peptides consisting of 4 to 7 residues. According to our data, the region corresponding to residues extending from approximately 180V to 212Q which was shown to be resistant to MAAH is most probably responsible for the oligomerization of the protein under acidic conditions. This region covers the region between helix 2 and helix 3 and approximately 70% of helix 2 and 45% of helix 3 in the original monomeric PrP^C. The three cleavage sites observed inside this core region correspond to the loop between helix two and helix three in the monomeric PrP^C. This acid induced conformer seems to share characteristics with the fibrilar form of the prion protein such as high β-sheet content and a similar solvent inaccessible region. Lu et al.[25] have previously reported a similar β-sheet core in human prion protein amyloid fibrils using hydrogen/deuterium exchange followed by mass spectrometric analysis. The HXMS data for the PrP fibrils revealed a highly exchange-protected region, 0-26% deuterium incorporation, between residues 169 and 213. Another recent study[26], performed on human prion protein amyloids, using site-directed spin labelling coupled with EPR spectroscopy, showed the stacking of parallel, in register β-strands, containing residues from approximately 160 – 220.

The MABH, even though it fails to give a complete hydrolysis of all the peptide bonds along the protein sequence, such as the MAAH is capable, it is worth noting the fact that under basic conditions our rSHPrP90-231 prion construct does not convert to a high β-sheet containing conformer. The DLS data shows a slight increase in the hydrodynamic radius suggesting formation of small oligomers, although combined with the CD data which shows unfolding the increase in the hydrodynamic radius is most probably caused by the unfolding of the protein and not oligomerization. Therefore, MABH of the peptide bonds is possible in the region that seems to be responsible for the oligomerization under acidic conditions.

Previous theoretical models[27, 28] showed that the N-terminal part of the prion protein (residues 20-145) could be responsible for the formation of amyloid fibrils and that the 3 α-helices of the C-terminal region are present in the misfolded conformer of the protein. In contrast to these simulated models and in accordance with recent

experimental data, our findings demonstrate that a small region within the C-terminal region of the mature prion protein, involving helix two and three of the original PrP^C, extending approximately from residue 180V to 213Q, is capable of forming a very stable, solvent inaccessible oligomeric structure.

3.5. References

(1)   Bruce, M. E.; Dickinson, A. G. *J Neuropathol Exp Neurol* **1985**, *44*, 285-294.

(2)   Prusiner, S. B.; Cochran, S. P.; Alpers, M. P. *J Infect Dis* **1985**, *152*, 971-978.

(3)   Carrell, R. W.; Lomas, D. A. *Lancet* **1997**, *350*, 134-138.

(4)   Carrell, R. W.; Gooptu, B. *Curr Opin Struct Biol* **1998**, *8*, 799-809.

(5)   Zhong, H.; Zhang, Y.; Wen, Z.; Li, L. *Nat Biotechnol* **2004**, *22*, 1291-1296.

(6)   Chapman, J.; Arlazoroff, A.; Goldfarb, L. G.; Cervenakova, L.; Neufeld, M. Y.; Werber, E.; Herbert, M.; Brown, P.; Gajdusek, D. C.; Korczyn, A. D. *Neurology* **1996**, *46*, 758-761.

(7)   Zahn, R.; Liu, A.; Luhrs, T.; Riek, R.; von Schroetter, C.; Lopez Garcia, F.; Billeter, M.; Calzolai, L.; Wider, G.; Wuthrich, K. *Proc Natl Acad Sci U S A* **2000**, *97*, 145-150.

(8)   Wang, N.; Xie, C.; Young, J. B.; Li, L. *Analytical Chemistry, 0*.

(9)   Dai, Y.; Whittal, R. M.; Li, L. *Anal Chem* **1999**, *71*, 1087-1091.

(10)   Karas, M.; Hillenkamp, F. *Anal Chem* **1988**, *60*, 2299-2301.

(11)   Herrmann, L. M.; Caughey, B. *Neuroreport* **1998**, *9*, 2457-2461.

(12)   Huwiler, K. G.; Mosher, D. F.; Vestling, M. M. *J Biomol Tech* **2003**, *14*, 289-297.

(13)   Hirel, P. H.; Schmitter, M. J.; Dessen, P.; Fayat, G.; Blanquet, S. *Proc Natl Acad Sci U S A* **1989**, *86*, 8247-8251.

(14)   Zhong, H.; Marcus, S. L.; Li, L. *J Am Soc Mass Spectrom* **2005**, *16*, 471-481.

(15)   James, T. L.; Liu, H.; Ulyanov, N. B.; Farr-Jones, S.; Zhang, H.; Donne, D. G.; Kaneko, K.; Groth, D.; Mehlhorn, I.; Prusiner, S. B.; Cohen, F. E. *Proc Natl Acad Sci U S A* **1997**, *94*, 10086-10091.

(16)   Lennon, C. W.; Cox, H. D.; Hennelly, S. P.; Chelmo, S. J.; McGuirl, M. A. *Biochemistry* **2007**, *46*, 4850-4860.

(17)   Stahl, N.; Baldwin, M. A.; Teplow, D. B.; Hood, L.; Gibson, B. W.; Burlingame, A. L.; Prusiner, S. B. *Biochemistry* **1993**, *32*, 1991-2002.

(18)   Silveira, J. R.; Raymond, G. J.; Hughson, A. G.; Race, R. E.; Sim, V. L.; Hayes, S. F.; Caughey, B. *Nature* **2005**, *437*, 257-261.

(19)   Baskakov, I. V.; Legname, G.; Gryczynski, Z.; Prusiner, S. B. *Protein Sci* **2004**, *13*, 586-595.

(20)   Swietnicki, W.; Petersen, R.; Gambetti, P.; Surewicz, W. K. *J Biol Chem* **1997**, *272*, 27517-27520.

(21)   Hornemann, S.; Glockshuber, R. *Proc Natl Acad Sci U S A* **1998**, *95*, 6010-6014.

(22)   Zou, W. Q.; Cashman, N. R. *J Biol Chem* **2002**, *277*, 43942-43947.

(23)   Jackson, G. S.; Hill, A. F.; Joseph, C.; Hosszu, L.; Power, A.; Waltho, J. P.; Clarke, A. R.; Collinge, J. *Biochim Biophys Acta* **1999**, *1431*, 1-13.

(24)   Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D. *Nature* **2005**, *435*, 773-778.

(25)   Lu, X.; Wintrode, P. L.; Surewicz, W. K. *Proc Natl Acad Sci U S A* **2007**, *104*, 1510-1515.

(26)   Cobb, N. J.; Sonnichsen, F. D.; McHaourab, H.; Surewicz, W. K. *Proc Natl Acad Sci U S A* **2007**, *104*, 18946-18951.

(27)   DeMarco, M. L.; Daggett, V. *Proc Natl Acad Sci U S A* **2004**, *101*, 2293-2298.

(28)   Govaerts, C.; Wille, H.; Prusiner, S. B.; Cohen, F. E. *Proc Natl Acad Sci U S A* **2004**, *101*, 8342-8347.

Chapter 4: Conclusions and Future Work

In Chapter 2, we investigated whether protein structure and stability can influence the microwave-assisted acid hydrolysis process of proteins. We have also investigated the applicability of MABH followed by MALDI-TOF-MS as a means of degrading proteins. We have shown that MAAH is applicable for protein hydrolysis followed by MALDI-TOF-MS. We have shown that the hydrolysis was possible even for proteins that contain stabilizing disulphide bonds. The CD data recorded for HEWL showed that the protein was able to maintain its secondary structure under highly acidic conditions. The hydrolysis map obtained for this protein shows that hydrolysis of peptide bonds occurs for all kinds of secondary structures, such as α-helical, β-sheet, turn or random coil, when referring to the original protein structure. We have also tested the feasibility of MABH followed by MALDI-TOF-MS as a means of degrading proteins. It has been found that base hydrolysis is able to generate both N- and C-terminal peptides, just as MAAH does, although the hydrolysis is not as uniform and the method is not suitable for determination of protein sequences and their modifications.

In Chapter 3, we investigated the behaviour and hydrolysis profiles of the wild type rSHPrP90-231 prion construct under acidic and basic conditions. We have shown that under acidic conditions the wild type rSHPrP90-231 prion construct presents a hydrolysis resistant core extending from approximately 180V to 212Q. This data suggested that the protein undergoes a structural change that would lead to a compact and stable structure which would confer solvent inaccessibility in the above specified region. The conclusions drawn from the MALDI-TOF-MS data was completed and confirmed by the CD data that showed a decrease in the helical content of the protein combined with the increase of the sheet content. Further experiments using DLS showed that under acidic conditions the prion protein not only changes its secondary structure but it also undergoes an oligomerization process. All these changes are characteristics of the scrapie form of the prion protein. Therefore, it can be concluded that the presented hydrolysis profile is corresponding not to the wild type rSHPrP^C but to a structure that is the same or similar to the scrapie form of the

prion protein. Even though MABH is not suitable for complete protein sequencing, it has been shown that strong a base can be used to destroy the scrapie form of the prion protein. By using MABH in combination with MALDI-TOF-MS we have found that base was able to cut inside the region that was found to be resistant to acid hydrolysis.

The next step of this project is to investigate the hydrolysis profiles of different mutants of the prion protein. It is well documented that different diseases such as FFI, fCJD, GSS and sCJD, are all linked to different mutations in the Prnp gene. Therefore, MAAH combined with MALDI-TOF-MS can be used to detect whether there are any differences in the hydrolysis profiles of these proteins, which could suggest differences in there structures and/or stabilities. Caughey et al. have shown that the presence of the disulphide bond is crucial for aggregation of the prion protein. The use of a mutated form of the prion protein, such as replacing one or both Cys residues of the protein, for MAAH followed by MALDI-TOF-MS would show whether the structure and structural changes of this acid converted conformer are influenced by the presence of the disulphide bond.

Another application of the techniques described in this thesis work may be involved in the use of different protocols to chemically modify the prion protein. MAAH combined with MALDI-TOF-MS could then be used to detect possible modification sites that would give important information about solvent accessibility of different conformers of the prion protein. MAAH could also be used to detect differences in the hydrolysis profiles of chemically modified prion constructs. This data could be further used by our collaborators to build possible structures of these conformers using sophisticated computer programs. This data could be used to perform computer simulations that would give an insight in the different folding and misfolding aspects of this protein.