UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 360

From Solution into Vacuum - Structural Transitions in Proteins

ALEXANDRA PATRIKSSON

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2007

ISSN 1651-6214
ISBN 978-91-554-7014-2
urn:nbn:se:uu:diva-8300

Dissertation presented at Uppsala University to be publicly examined in B7:113a, BMC, Husargatan 3, Uppsala, Saturday, December 1, 2007 at 13:00 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

Patriksson, A. 2007. From Solution into Vacuum - Structural Transitions in Proteins. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 360. viii, 44 pp. Uppsala. ISBN 978-91-554-7014-2.

Information about protein structures is important in many areas of life sciences, including structure-based drug design. Gas phase methods, like electrospray ionization and mass spectrometry are powerful tools for the analysis of molecular interactions and conformational changes which complement existing solution phase methods. Novel techniques such as single particle imaging with X-ray free electron lasers are emerging as well. A requirement for using gas phase methods is that we understand what happens to proteins when injected into vacuum, and what is the relationship between the vacuum structure and the solution structure.

Molecular dynamics simulations in combination with experiments show that protein structures in the gas phase can be similar to solution structures, and that hydrogen bonding networks and secondary structure elements can be retained. Structural changes near the surface of the protein happen quickly (ns-μs) during transition from solution into vacuum. The native solution structure results in a reasonably well defined gas phase structure, which has high structural similarity to the solution structure.

Native charge locations are in some cases also preserved, and structural changes, due to point mutations in solution, can also be observed in vacuo. Proteins do not refold in vacuo: when a denatured protein is injected into vacuum, the resulting gas phase structure is different from the native structure.

Native structures can be protected in the gas phase by adjusting electrospray conditions to avoid complete evaporation of water. A water layer with a thickness of less than two water molecules seems enough to preserve native conditions.

The results presented in this thesis give confidence in the continued use of gas phase methods for analysis of charge locations, conformational changes and non-covalent interactions, and provide a means to relate gas phase structures and solution structures.

Keywords: molecular dynamics, computer simulations, mass spectrometry, electrospray ionization, free-electron laser, vacuum structure of proteins, solvation, desolvation, single molecule imaging

Alexandra Patriksson, Department of Cell and Molecular Biology, Box 596, Uppsala University, SE-75124 Uppsala, Sweden

ISSN 1651-6214
ISBN 978-91-554-7014-2
urn:nbn:se:uu:diva-8300 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-8300)

List of Publications

This thesis is based on the following publications:

I M. Seibert, **A. Patriksson**, B. Hess and D. van der Spoel, (2005) Reproducible polypeptide folding and structure prediction using molecular dynamics simulations. J. Mol. Biol. 354 pp. 173-183

II **A. Patriksson**, C. Adams, F. Kjeldsen, J. Raber, D. van der Spoel and R. A. Zubarev, (2006) Prediction of N-Cα bond cleavage frequencies in electron capture dissociation of Trp-cage dications by force-field molecular dynamics simulations. Int. J. Mass Spectrom. 248 pp.124-135

III C. M. Adams, F. Kjeldsen, **A. Patriksson**, D. van der Spoel, A. Gräslund, E. Papadopoulos and R. A. Zubarev, (2006) Probing solution- and gas-phase structures of Trp-cage cations by chiral substitution and spectroscopic techniques. Int. J. Mass Spectrom. 253 pp. 263-273

IV **A. Patriksson**, E. Marklund and D. van der Spoel, (2007) Protein structures under electrospray conditions. Biochemistry 46 pp. 933-945

V A. T. Iavarone, **A. Patriksson**, D. van der Spoel and J. H. Parks, (2007) Fluorescence probe of Trp-cage protein conformation in solution and in gas phase. J. Am. Chem. Soc. 129 pp. 6726-6735

VI **A. Patriksson**, C. M. Adams, F. Kjeldsen, R. Zubarev and D. van der Spoel, (2007) A direct comparison of protein structure in the gas and solution phase - the Trp-cage J. Phys. Chem. B *accepted*

VII **A. Patriksson** and D. van der Spoel, A temperature calculator for replica exchange molecular dynamics simulations. *submitted to Phys. Chem. Chem. Phys.*

Reprints were made with permission from the publishers.

Abbreviations

BPTI	bovine pancreatic trypsin inhibitor
CAD	collision activated dissociation
CD	circular dichronism
CRM	charge residue model
CSD	charge state distribution
ctf	C-terminal fragment of ribosomal protein L7/L12
ECD	electron capture dissociation
ESI	electrospray ionization
IEM	ion evaporation mechanism
MALDI	matrix assisted laser desorption ionization
MD	molecular dynamics
MS	mass spectrometry
MS/MS	tandem mass spectrometry
NMR	nuclear magnetic resonance
PDB	protein data bank
REMD	replica exchange molecular dynamics
RMSD	root mean square deviation
XFEL	X-ray free electron laser

Contents

1 Introduction 1
2 Gas phase methods for protein structure analysis 3
2.1 Mass spectrometry 3
2.1.1 Electron capture dissociation (ECD) 4
2.2 Electrospray ionization (ESI) 7
2.3 Single particle imaging with X-ray lasers 9
2.4 Fluorescence probe spectroscopy 11
3 Computers as a complement to experiments 13
3.1 Molecular dynamics (MD) 13
3.2 Replica exchange molecular dynamics (REMD) 15
3.3 Gromacs 18
4 Protein structure and dynamics 19
4.1 Electrostatic and hydrophobic interactions 19
4.2 Structure and charge in the gas phase 20
4.3 Structural waters and protein structure 24
4.4 Folding, unfolding and conformational space 25
5 Conclusions and future perspective 27
6 Sammanfattning på svenska 31
7 Acknowledgements 35
8 References 37

1. Introduction

Proteins play a crucial role in virtually all biological processes and are in many ways critical for our survival and well being. The three-dimensional structure, which is determined by the amino acid sequence encoded in the genes in the DNA, is highly specific and strongly correlated to the function of the protein. Dysfunctional proteins with slightly altered or damaged structures are responsible for a number of diseases, including Alzheimer's, Creutzfeldt-Jacob's disease and several variants of cancers.

Because of their functional significance, proteins are the main target for the pharmaceuticals that we use. Therefore, structure based drug design exploits knowledge of the three-dimensional structure of proteins to develop drugs for specific purposes. Methods like X-ray crystallography and nuclear magnetic resonance spectroscopy (NMR) have produced nearly 50 000 protein structures over the past decades, but there is a large class of proteins and other macromolecular structures that are not amenable for simple structure determination by these techniques. For these molecules, which include misfolded proteins, the majority of membrane proteins, glycoproteins and a range of "floppy" complexes, a means of structure determination would be desirable.

Using mass spectrometric methods, biomolecular interactions in the gas phase can be studied. The combination of different variants of the technique also allows for determination of amino acid sequences of proteins. The use of gas phase methods is, however, always subject to questions whether the results obtained are applicable to proteins in their true environment, i.e. in solution, or not. One of the aims of this thesis was to address these questions by presenting a number of studies, based on computer simulations combined with experiments, mainly of proteins in vacuum. The goal has been to investigate the major differences and similarities between protein structures and dynamics in solution and in the gas phase (**paper I** and **III - VI**), and to find out what we can learn about proteins in their native "wet" environment by studying them in vacuo. **Paper II** goes deeper into one specific

mass spectrometric technique, electron capture dissociation (ECD), and suggests a new possible mechanism for the fragmentation.

The electrospray technique is a widely used method for ionization and injection of molecules within mass spectrometry. It is also considered as injection method for the planned single particle imaging experiments with X-ray free electron lasers (XFEL). Despite its popularity, it is still not fully known how the process, and particularly the evaporation, affects the protein structure. Another area of interest has therefore been to study the behavior of proteins when injected into the gas phase or vacuum. **Paper IV** investigates these questions and gives a hint of what to expect from the experimental output of XFEL sources.

Since my work has been concentrated on computer simulations using molecular dynamics, this thesis also contains some method development. **Paper VII** describes an algorithm developed to generate temperatures for replica exchange molecular dynamics which guarantee good sampling of structures - a requirement for protein folding simulations.

The contents of this thesis has been organized as follows: First, some of the various gas phase methods and techniques that are used for analysis of protein structures and dynamics are explained, second, the use of computer simulations as a complement to experiments is described, and last, protein structure and dynamics in the gas phase and in solution are discussed.

2. Gas phase methods for protein structure analysis

Water is in many ways essential for life, also on a molecular level. Crystal structures of proteins usually contain water molecules which are more organized than the bulk solution and which are believed to be important for the activity and the stability of the structure. The most obvious way to study protein structure and behavior is therefore in an aqueous environment. Gas phase methods nevertheless provide a powerful means for analysis of molecular interactions, and a number of different techniques have emerged through the years. Among the most popular methods to study biomolecules like proteins, are ion mobility spectrometry [43], gas phase spectroscopy [12, 26] and the various mass spectrometric techniques in combination with matrix assisted laser desorption ionization (MALDI) [66, 32] or electrospray ionization (ESI) [15, 75, 76]. The gas phase methods used in the research presented in this thesis include the fragmentation technique, electron capture dissociation (ECD) [79], ESI and fluorescence probe spectroscopy [28].

2.1 Mass spectrometry

The mass spectrometric methods are all based on detection of ionized molecules in vacuo by measurements of mass over charge ratios. Two decades ago, the use of mass spectrometry was generally limited to analysis of low molecular-mass samples and application to proteins was difficult. Since then, the situation has improved radically, and today mass spectrometry is one of several standard methods used to study large protein complexes and non-covalent interactions such as ligand binding in the gas phase [74, 58]. Fragmentation techniques make it possible to produce information not only about the molecular formula of proteins but also about the covalent structure. Tandem mass spectrometry (MS/MS) [42], where the ion of interest first is isolated and then fragmented, has been advantageous for this purpose. The combination of several different fragmentation techniques even allows for sequence determination of unknown proteins and peptides [22].

2.1.1 Electron capture dissociation (ECD)

One of various mass spectrometric fragmentation techniques is the electron capture dissociation (ECD), involving a reaction between multiply-charged gas phase polypeptides and low-energy electrons (< 1 eV) [79]. ECD specifically cleaves the strong N-Cα backbone bonds of proteins, without rupture of weak intra- and intermolecular bonding [19]. This cleavage, schematically shown in Figure 2.1, gives rise to quite different fragments, so called c and z ions, compared to for instance collision activated diffraction (CAD), which in the low energy mode produces b and y fragments from cleavage of the C-N bond [52]. In ECD, cleavage occurs between almost any combination of amino acids, except at the N-terminal side of the cyclic proline residue, and therefore extensive sequence information of proteins and peptides can be obtained from ECD spectra. If this information is compared to the fragmentation pattern obtained using some other method, like CAD, the primary amino acid sequence of unknown proteins can be revealed. For ECD to be effective, the protein ions have to be multiply charged, and the preferred method for production of such polycations is electrospray ionization (section 2.2).

The mechanism of ECD has been widely debated for some time, and a number of plausible hypotheses have been proposed [23, 64, 80, 38]. One of these is the charge solvation model, where cleavage is believed to occur in the vicinity of the protonation site, i.e. the charge produced in the ionization process, due to electron capture on the hydrogen bond created between the extra charge and any other part of the protein. This hypothesis have been tested both experimentally [6] and

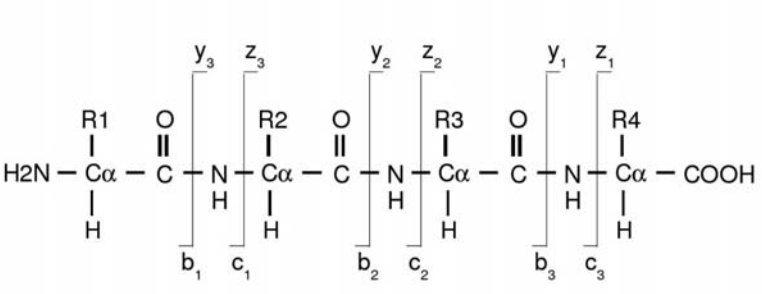

Figure 2.1: Different mass spectrometric techniques give rise to different fragmentation ions. c and z ions are produced from ECD cleavage of the N-Cα bond and CAD produces b and y ions through cleavage between the backbone N and C atoms.

computationally [54] with varying results, but the challenge to explain and predict the relative abundances of the N-Cα cleavages remains.

In **paper II**, the chemistry of the ECD cleavage mechanism is investigated using a combination of experiments and molecular dynamics simulations of stereoisomers of the small 20-residue model protein Trp-cage. A previous study [1] had shown that conversion of the chirality (L to D) of a single amino acid of the peptide gives distinctly different ECD fragmentation, and thus stereoisomers were considered suitable for this study as well. Three amino acids, Tyr^3, Gln^5 and Leu^7 were selected for mutation, yielding D-Tyr^3, D-Gln^5 and D-Leu^7 in addition to the all-L variant, which was identical with the native conformation with all amino acids in L-form. The above mentioned study also showed that, in the gas phase, the protonation of Trp-cage 2+ ions is at Gln^5 and Arg^{16} and not on Lys^8 and Arg^{16} as in solution. Therefore, the protonation of the stereoisomers, including the all-L variant, was selected accordingly. For control, a fourth isomer with the protonation on Lys^8 and Arg^{16} was also constructed. No L to D mutation was done on this isomer, hence, all amino acids were kept in their native L-form.

For each stereoisomer, simulations were run for 200 ns using the replica exchange (REMD) approach (section 3.2) at 16 different temperatures, ranging from 275 to 419 K. From the 305 K trajectories, which were closest in temperature to the experiments (300 K) the following analyses of the hydrogen bonding pattern were done:

I. The total number of time frames (sampled every pico second) in which the protonated side chain of Gln^5 (or Lys^8 in the control system) was found to make a hydrogen bond with a backbone carbonyl of the protein. This parameter represents a quantitative measure of the charge solvation probability.

II. For each time frame, the total number of neutral hydrogen bonds established between backbone carbonyls and any other part of the molecule.

III. The frequency of creation/destruction of neutral hydrogen bonds defined as above.

Initially, the charge solvation hypothesis was tested by comparing analysis I with experimental data generated by Roman Zubarev and coworkers. However, several inconsistencies between simulation and

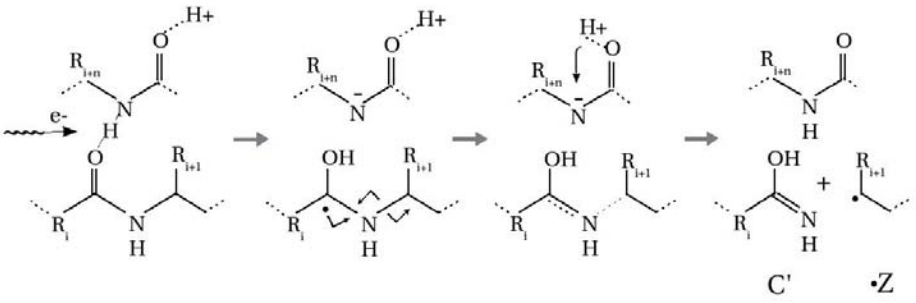

Figure 2.2: The ECD mechanism proposed in **paper II**, suggesting an initial capture of a low-energy electron on a neutral hydrogen bond between a backbone carbonyl and any other part of the protein, for instance a backbone amine. Subsequent hydrogen atom transfer to the carbonyl forms a labile aminoketyl radical that rapidly fragments through N-Cα-cleavage. A hydrogen from one protonation site relocates to the backbone nitrogen to reform the amine group.

experiment were obtained, such as extreme carbonyl solvation of Pro[19] and drastic temperature effects, which were not observed in experiments. Calculations also predicted a two order of magnitude larger frequency of charge solvation in D-Gln[5] and D-Leu[7] compared to the other stereoisomers, whereas the fragment abundance measured in experiments were indistinguishable. No matter how the simulated data was used, it failed broadly to reproduce the observed ECD abundances of N-Cα bond cleavages. Instead, the ECD patterns and the calculated occurrence of neutral hydrogen bonds to carbonyls of the fragmenting amides (analysis II) agreed very well, and as a consequence, a new ECD mechanism was suggested (Figure 2.2). This mechanism is based on the capture of a low-energy electron on a neutral hydrogen bond between a backbone carbonyl and any other part of the protein, for instance a backbone amine. Subsequent hydrogen atom transfer to the carbonyl forms a labile aminoketyl which fragments through N-C_α bond cleavage [80]. The most loosely bound ionization proton will then neutralize the backbone nitrogen via intra-molecular proton transfer [36]. Note: that the schematic mechanism in Figure 2.2 is the correct one - the figure presented in the publication (**paper II**) is missing some details.

Analysis III was done to test whether the N-Cα bond cleavage is due to the hydrogen bond itself or to the energy transfer taking place as the hydrogen bond is created/destructed. The correlation between

the observed N-Cα bond cleavage abundances and the calculated frequencies of hydrogen bond creation/destruction was, however, not as strong as the correlation to the occurrence of hydrogen bonds (analysis II). The conclusion was therefore that it is the hydrogen bonding itself and not the rate of it's establishing, that determines the N-Cα bond cleavages.

The proposal of this new mechanism does not imply a complete dismissal of the charge-solvation hypothesis. It is rather a suggestion that an interplay between the charged and neutral hydrogen bonding is possible. Taking both these theories into account when studying ECD fragment abundances, may help us understand what is going on and thus improve the possibility to predict ECD fragmentation.

2.2 Electrospray ionization (ESI)

A fundamental prerequisite of all mass spectrometric experiments is a method to produce charged species of the molecules that are to be analysed. In particular studies of macromolecules such as proteins require robust methods to generate ion products. One method, which has become very popular because it makes it possible to produce ions directly from solutions at atmospheric pressure, is ESI.

ESI was initially introduced by Dole and co-workers [10] in the late 1960's. Later, the method was developed for use with large molecules by John Fenn [15, 75, 76], an invention that in 2002 was rewarded with the Nobel Prize in Chemistry. The uniqueness of ESI is its ability to form intact ions with a wide variety of charge states under "soft" conditions where weak bonds, such as non-covalent interactions that exist in solution, can be preserved and studied under atmospheric pressure in gas phase or in vacuum. This makes ESI particularly suitable for ionization of proteins which then can be analyzed using, for instance, ECD.

Despite its popularity, the ionization mechanism in ESI is not fully understood, but there are two predominant theories - the charge residue model (CRM), initially proposed by the originator of protein ESI, Dole, and co-workers [10], and the ion evaporation mechanism (IEM), introduced by Iribane and Thomson [29, 67]. Studies suggest that both mechanisms apply, but to species of different sizes. The IEM mechanism seems not to hold for large molecules such as non-denatured

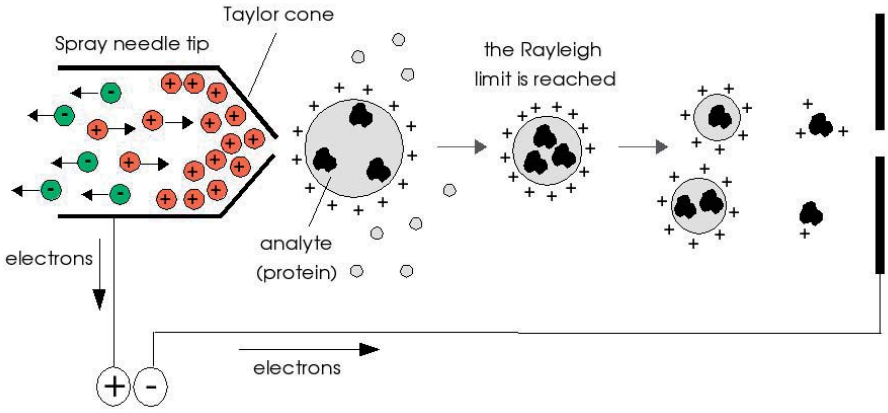

Figure 2.3: An illustration of the electrospray ionization process showing the evaporation and subsequent disintegration of the charged droplets, until the final production of gas phase ions.

globular proteins, whereas CRM does [14, 34, 9, 71, 13, 70]. Therefore, the ESI process will be explained based on the CRM model only.

The ESI process (Figure 2.3) can be divided into three phases: 1) formation of small charged droplets at the tip of the ESI capillary, 2) evaporation of solvent from the droplet and subsequent coulombic fission which produces even smaller, charged, off-spring droplets and 3) formation of charged ions. The first phase is well understood and can today be handled accurately in order to obtain either positively or negatively charged droplets. In its simplest form, a sample solution of the biomolecule is fed through a small capillary and a high electric field at the tip of the capillary pulls positive or negative charges, depending on the direction of the field, towards the front of the liquid. When the electrostatic repulsion becomes stronger than the surface tension of the liquid, a small electrically charged droplet leaves the surface and travels through the surrounding gas or vacuum to the counter electrode. During its flight through the gas, the waters of the droplet starts to evaporate such that the droplet becomes smaller while the charge that it carries remains constant. As an effect, the charge density at the surface is increased and causes electrostatic stress. At a certain point, defined by the the Rayleigh limit [57, 18], the surface coulombic forces exceed the surface tension and the droplet becomes unstable and disintegrates into even smaller off-spring droplets. Further disintegration, of both the

parent droplet and the off-spring droplets, occurs until they become so small that they can bear only a few excess charges. Continued solvent evaporation of these droplets will lead to completely dehydrated ions. The exact details of this process are not fully known, but it seems that the final evaporation brings the ions close enough to the protein surface for protonation to occur to form charged protein ions.

Even though the evaporation-ionization process of the ESI is known to preserve specific non-covalent interactions that exist in solution [58], the influence of ESI on protein conformations has not been possible to determine by experimental means. Varying results have been obtained, ranging from completely unfolded proteins [30, 69, 28] to viruses that are still infective after injection by ESI [68], which suggests that any changes inflicted by dehydration should at least be reversible. The study presented in **paper IV** implies, however, that as long as there are some water molecules bound to the protein, the structure is quite unaffected by evaporation. Proteins that are initially covered with a water layer of 3 Å, show no significant structural changes after 10 ns in vacuum even though between 13 and 23 % of the water has evaporated. The changes in secondary structures are not very much different from the changes recorded in bulk water (Figure 4.1, section 4.2), which indicates that a water layer with a thickness of roughly two molecules is enough to maintain the solution phase structure. For the XFEL experiments, described in section 2.3, this means, that if the ESI process can be fine-tuned to preserve a certain amount of water around the protein [59] as it is hit by the beam, then the likelihood to obtain a diffraction pattern that corresponds to the native structure will be increased. More results from **paper IV**, are presented in section 4.2.

2.3 Single particle imaging with X-ray lasers

Novel single particle imaging experiments have been proposed [8, 16] using X-ray free electron lasers. These experiments differ from traditional X-ray crystallography in that structural information will be retrieved from individual molecules in the gas phase instead of from protein crystals. Molecules will be injected into a vacuum chamber, where they will be hit by an X-ray pulse and scatter onto a detector (Figure 2.4). To obtain diffraction patterns from non-repetitive, amorphous materials, such as individual protein molecules, an X-ray source is needed, which can provide a photon pulse of much higher intensity than the sources that are used for crystallography today. These X-rays can be produced by passing

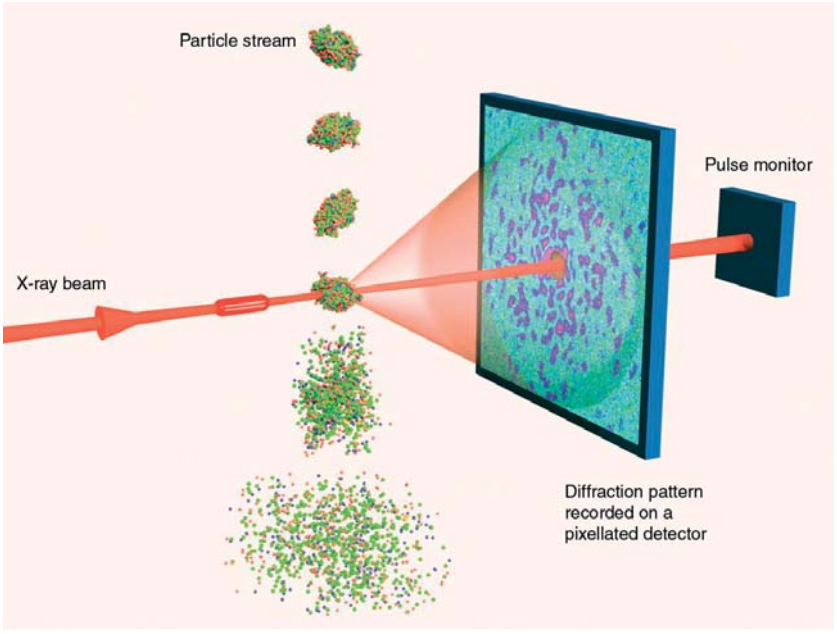

Figure 2.4: Illustrative description of the single particle imaging with an XFEL (figure taken from reference [8] with permission).

an intense and energetic electron pulse, generated in a linear accelerator, through a long periodic magnet - an undulator - which forces the beam to emit radiation. The radiation will interact with the beam pulses within the undulator and then cause a build-up of micro bunches of charge density. The electrons in the micro bunches will eventually radiate coherently, giving rise to extremely intense femtosecond X-ray pulses with wave lengths on the Å-scale. Although biological material is extremely sensitive to radiation damage, recent theoretical studies indicate that, if the experimental parameters are chosen carefully [47, 20, 78, 4], the sample may survive long enough for the scattering pattern to be recorded, thus imaging of small biomolecular aggregates or even large single particles could be possible [8, 16].

A requirement for the XFEL single particle imaging of individual molecules is an injection process that can provide a purified stream of molecules into the vacuum chamber. The use of electrospray ionization (ESI) has been very successful within mass spectrometry, and is therefore considered appropriate for the XFEL experiments as well. The ESI method is described more thoroughly in section 2.2.

2.4 Fluorescence probe spectroscopy

Fluorescence spectroscopy measurements have traditionally been used for conformational studies of proteins in solution, but recently, the method has been extended to applications in the gas phase as well, with very promising results [28, 27, 26]. The method is based on interactions between a photophysical probe and a quencher, which are both incorporated into the protein or peptide of interest. The probe and the quencher can either be natural parts of the protein or peptide [37, 77] or they can be covalently attached [24]. The probe is brought to an excited state by a short light pulse, and quenching, i.e. interaction between the probe and the quencher, will cause a decay of the excited state which can be detected as a change (decrease) in fluorescence intensity or fluorescence lifetime. Since quenching is known to be distance and orientation dependent [49, 48], this decay in fluorescence can be used to probe distances within molecules and interconversion between conformers can be analysed [77]. In the gas phase, the method makes use of modulation of the fluorescence of a covalently attached dye through intramolecular quenching, usually by tryptophan (Trp) (Figure 2.5). Intact unsolvated biomolecule ions are formed by electrospray ionization (ESI) and transferred into a trap. There, the particular ion of interest is isolated through ejection of the unwanted ions which occur at lower or higher mass-to-charge ratios. Fluorescence is induced by a laser. As in the solution experiments, the quenching causes a decrease in the measured fluorescence which is

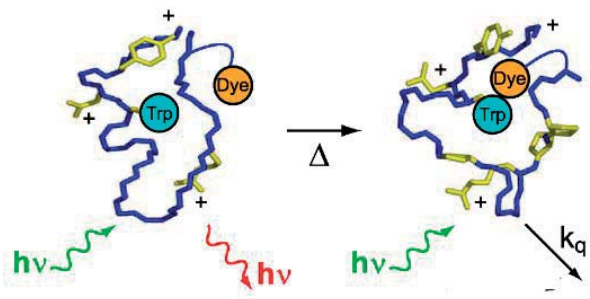

Figure 2.5: The mini protein Trp-cage with the fluorescent dye attached to the C-terminus. When the quencher (Trp) and the dye are separated, fluorescence can be detected. Any conformational changes in the peptide which brings the quencher and the dye into contact will cause a decrease in the detected intensity.

consistent with conformational changes in the protein that causes the tryptophan to get into contact with the dye. The present model [28] assumes that quenching is absent in the folded state.

So far, the gas phase variant of the method has been applied to and tested on the mini protein Trp-cage. In **paper V**, the measurements are used in combination with solution measurements and theoretical calculations, i.e molecular dynamics simulations, in the gas phase, to study protein unfolding and the effects of solvent on protein structure and dynamics. All analyses are done as a function of charge state, charge location and temperature. The study proves that the observed changes in fluorescence are due to dynamics of the polypeptide and not the intrinsic photophysics of the dye. It is found that differently charged peptides produce different melting curves, which indicates that the method is capable to trace temperature-induced conformational changes. Since structures calculated for the different charged states using molecular dynamics, consequently show relatively compact structures, these conformational changes are assumed to be fluctuations in secondary structure rather than large scale changes of the tertiary structure. It is clear that further development of the method is required to explain the local interactions in biomolecules in the gas phase. Still, an essential result of the measurements reported in **paper V** is that fluorescence probe spectroscopy is useful for producing a picture of protein dynamics in the gas phase as a function of temperature.

3. Computers as a complement to experiments

Protein structure determination using experimental methods is not trivial, and there is considerable interest in developing computer based methods to facilitate the process. The grand challenge is to solve the "protein folding problem", i.e. to predict a protein's three-dimensional structure from it's primary amino acid sequence, but the way there is still very long. Even if experiments can not be replaced by theoretical methods yet, there are certain areas where theory and simulation can yield complementary information which is not easily accessible by experimental means. The "static" structures obtained from experiments can be used in computer simulations to generate dynamics trajectories. In this manner we can study events on short timescales with fine spatial resolution.

Computer calculations of molecular systems can be performed at many levels of accuracy, ranging from quantum mechanistic to classical and from atomistic to coarse-grained. When the aim is to simulate protein folding processes, the precision of classical physics is usually good enough, but atomic models are to be preferred over coarser descriptions. A number of approximations can be done, and it is mostly useful to disregard the quantum mechanical properties of the system, not only to save time but also to increase the upper limit of the size of systems that can be simulated. A widely used computational method for studying the folding of small proteins is molecular dynamics (MD) simulations.

3.1 Molecular dynamics (MD)

Molecular dynamics (MD) was first introduced by Alder and Wainwright in the late 1950's [2] as they calculated non-equilibrium and equilibrium properties of hard spheres. The first modelling of a realistic system was done 7 years later by Rahman, as he

simulated liquid argon in 1964 [55]. A decade later, in 1977, the first protein simulations appeared, with the bovine pancreatic trypsin inhibitor (BPTI) as a pioneering system [41]. Since then, research groups all over the world have put a great deal of effort in further developing the MD algorithms [53, 39, 51, 31] and progress is indeed being made. One important milestone within the protein folding field was reached when Duan and Kollman for the first time, in 1998, managed to perform a 1 μs long MD simulation of a small protein in water solution by brute force [11]. Many successful MD studies have been reported and published after that, e.g. [63, 33, 17].

MD is based solely on classical Newtonian physics, which means that the quantum nature of atoms is completely ignored. The theoretical basis for doing so is the Born-Oppenheimer approximation, which states that a separation of the velocities of the atomic nuclei and the electrons is possible. Because of their substantially smaller mass, electrons react instantaneously to changes in the position of the nuclei, thus, the most relevant motional information is obtained by considering the nuclei only.

A simple definition of MD is that it simulates the successive motions of a system of particles - the atoms of a molecule - due to the forces that are present. This means that using Newton's laws of motion, the natural structural changes in a protein can be generated, and from this information, a trajectory can be created which specifies how the positions and velocities of the atoms in the protein vary with time. We know from Newton's second law that "force equals the rate of change of momentum", which mathematically can be expressed as

$$\bar{F}_i = m_i \bar{a}_i = m_i \frac{d^2 \bar{r}_i}{dt^2} \tag{3.1}$$

where $\bar{F}_i$ is the force acting on an atom i with mass m_i and acceleration $\bar{a}_i$. To solve this differential equation, which gives the final trajectory, knowledge of the force, $\bar{F}_i$, is required. This can be computed as the negative derivative of the potential of the simulated protein or system.

$$\bar{F}_i = -\frac{\partial V}{\partial \bar{r}_i} \tag{3.2}$$

This potential, or the total energy of the system, V, and the corresponding parameters, are the very essence of MD called the force field. Usually, this potential is constructed as the sum of a number of individual energy terms, which refers to different bonded and non-bonded atomic interac-

tions such as bond stretching, angle bending, dihedral bond rotation, electrostatic and van der Waals interactions.

$$
\begin{aligned}
V_{tot} &= V_{bonded} + V_{non-bonded} && (3.3)\\
&= V_{bond}(l) + V_{angle}(\theta) + V_{dih}(\phi) + V_{el}(r) + V_{vdw}(r) && (3.4)\\
&= \sum_{bonds} \frac{k_l}{2}(l-l^0)^2 + \sum_{angles} \frac{k_\theta}{2}(\theta-\theta^0)^2 + \sum_{dih} \sum_{n=0}^{N} k_{n\phi}(cos(\phi))^n + \\
&\quad \sum_{i<j} \left[\frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} + \left(\frac{A_{ij}}{r_{ij}}\right)^{12} - \left(\frac{B_{ij}}{r_{ij}}\right)^{6} \right] && (3.5)
\end{aligned}
$$

where k_l, k_θ and $k_{n\phi}$ are force constants, l is the displacement from the equilibrium bond length l^0, θ the displacement from the equilibrium bond angle θ^0, ϕ is the dihedral angle, q_i and q_j are interacting point charges separated by the distance r_{ij} and A_{ij}, and B_{ij} are atom dependent constants.

To capture the dynamic behavior of the simulated system, equation 3.1 has to be solved for every atom i during a period of time. This is done numerically and one of the most commonly used algorithms is the Leap-frog Verlet method [21]. This algorithm uses information about the velocities at time $t - \frac{\Delta t}{2}$ and the positions and forces at time t, to calculate new velocities and positions, according to the following equations:

$$v(t + \frac{\Delta t}{2}) = v(t - \frac{\Delta t}{2}) + \frac{F(t)}{m}\Delta t \tag{3.6}$$

$$r(t + \Delta t) = r(t) + v(t + \frac{\Delta t}{2})\Delta t \tag{3.7}$$

3.2 Replica exchange molecular dynamics (REMD)

Classical MD simulations run at a single constant temperature are often inefficient in sampling of conformational space and therefore require a lot of computer time before any significant results can be observed, especially if the aim is to do protein folding. A better approach, which helps improving the sampling of conformers and thus to produce equilibrium simulations, is to use replica exchange MD (REMD) [65, 25]. REMD involves simultaneous simulation of multiple replicas of the same system at different temperatures (Figure 3.1). At

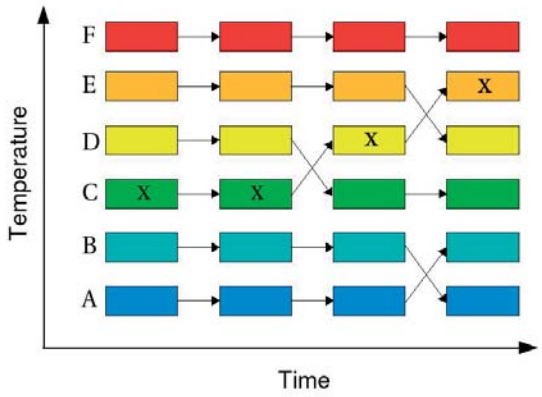

Figure 3.1: REMD involves simultaneous simulation of multiple replicas of the same system at different temperatures. At frequent intervals, conformers of concurrent simulations are allowed to switch temperature.

regular intervals, the N noninteracting replicas are allowed to switch temperature (based on a Metropolis criterion) with each other, thereby making it possible for low temperature replicas to gradually migrate up to higher temperatures and back again. By selecting the temperature range such that it starts from low values, corresponding to native conditions, and goes to high values, where the system is expected to have enough energy to rapidly pass potential barriers, this frequent switching allows for significantly enhanced sampling of conformations as a function of time.

The performance of REMD is tested in **paper I**, where the small 10-residue peptide chignolin is simulated using both REMD and classical MD. It is found that the latter set of simulations is outperformed in terms of sampling - due to the temperature exchanges in the REMD simulations, the native state of the peptide is reached much more quickly than in the single temperature MD. Almost all REMD replicas are found to undergo several folding/unfolding events whereas in the classical MD simulations, the native state occurs only rarely.

REMD simulations require a certain amount of preparatory work. One difficulty is to find an appropriate set of temperatures that will yield satisfying exchange frequencies. It is desirable to have much the same exchange frequency between all adjacent replicas [56]. It should not be too low, because then the simulations will be just as inefficient as

the classical MD, and it should not be too high either, because then the conformations may not have enough time to change between the exchanges, restricting the sampling to a small area in conformational space [72]. Temperatures are often selected using a "trial and error" approach which is to expensive both in terms of computer time and energy. A few schemes for optimal allocation of temperatures have been published [61, 56], but unfortunately they all require pre-calculation or an initial assumption of the energies of the system at a number of initial temperatures. To overcome this, **paper VII** presents an algorithm that, given the number of atoms in the system and an upper and lower temperature limit, determines which temperatures to use to obtain a desired exchange probability of P. No further prior knowledge about energies or temperatures is needed. The algorithm has been tested and validated using four additional REMD simulations which were not included in the parametrization. Three of the simulations contain a protein dissolved in bulk water, and the fourth simulation is a vacuum simulation of a small model protein. For each pair of adjacent replicas, the probability of exchange (P_{pred}) was predicted based on the temperature difference and the number of atoms in the system. The predicted probability, P_{pred}, was then compared to the observed probability, P_{obs}, obtained during simulation, and in all cases the convergence criterion $abs(P_{pred} - P_{obs}) \leq 0.0001$ was fulfilled for every pair. The predictive power seems to be very good for the GROMOS force field [51] even though the algorithm was parameterized using the OPLS force field [31], suggesting that it will work well for other similar force fields as well.

The algorithm has been implemented as a web server at http://folding.bmc.uu.se/remd, which can be freely used to generate appropriate temperatures for any protein based system. This possibility to choose an appropriate exchange rate beforehand, will certainly facilitate the preparation of REMD simulations, resulting in more efficient usage of scarce computational resources.

The parameterization of the algorithm has been done at constant pressure. A comparable set of constant volume simulations should be done to derive constants for NVT-REMD simulations as well, which then will be implemented in the same web server.

3.3 Gromacs

All simulations presented in this thesis have been performed using the MD software Gromacs [3, 39], initially developed by the group of Herman Berendsen at the department of Biophysical Chemistry at Groningen University. The software is primarily designed for biochemical molecules like proteins and lipids, but can be used for simulations of non biological systems like polymers as well. It is freely available to download from the web page http://www.gromacs.org and today the software has more than 2000 users all around the world.

4. Protein structure and dynamics

The three-dimensional structure of proteins is strongly influenced by environmental conditions such as pH, temperature, pressure and the concentration of other chemical substances. Physical and chemical properties, like charge distributions, hydrogen-bonding capacity and chemical reactivity, can be altered due to by small variations in the molecular environment, with changes in the three-dimensional structure as a consequence. Since the gas phase is considerably different from the solution environment in a living cell, the use of gas phase methods must have effects on the protein structure. It is difficult to determine exactly what these effects are because there are no gas phase methods that give direct structural information. Research in the area has, however, improved the general understanding of protein structures in the gas phase and there are indicators that at least aspects of native structures of proteins are retained in the gas phase [60].

4.1 Electrostatic and hydrophobic interactions

Protein structures in solution are stabilized by electrostatic interactions and the hydrophobic effect. Electrostatic interactions refer to a range of interactions between charges and dipoles, including salt bridges and hydrogen bonds. The electrostatic interactions can be either attractive or repulsive, and the atomic groups participating in electrostatic interactions are usually referred to as hydrophilic because they can interact strongly with water as well. In solution, the hydrophilic groups often prefer to be on the outside of the protein, where they can interact with the surrounding water molecules by making hydrogen bonds. The hydrophobic effect refers to the non-polarity of certain atoms or groups of atoms. When exposed to water, these groups tend to behave like oil droplets - they cluster together. The hydrophobic effect can also be considered a consequence of the inability of these atomic groups to compete with the much stronger electrostatic interactions. Therefore, in solution, hydrophobic groups are more frequently found in the inner parts of a

protein, called the core, well hidden from the surrounding water. A third group is the Van der Waals interactions, which are present in all atomic systems. These interactions are neither electrostatic nor hydrophobic but are of non-polar nature.

4.2 Structure and charge in the gas phase

In the gas phase, the effect of the electrostatic and the hydrophobic interactions are reversed: the intramolecular hydrophobic interactions become unimportant, and instead, the relative strength of the electrostatics is increased. The spontaneous assumption would therefore be that, in the gas phase, where the structurally important water is depleted, the stability of the protein structure must be lost and the protein would unfold into something that is significantly different from the native conformation observed in solution [62]. There is even a study stressing that these structural changes should be so drastic that the protein takes on an "inside-out" conformation, with very little resemblance to the native solution conformation [73]. However, according to more recent research [60] and based on the results presented in **paper III - VI**, the "inside-out" hypothesis must be considered unlikely.

During gas phase experiments, like electrospray ionization and mass spectrometry, proteins are transferred from solution into vacuum, a process that will cause evaporation of the surrounding water. Upon evaporation, the proteins adjust their conformations to minimize the internal energy. Side-chains that previously were bonded to surrounding water molecules have to find new sites to interact with and thus a number of new intramolecular hydrogen bonds will be created. This doesn't necessarily mean that the entire conformation has to change. **Paper IV** presents a simulation study of the evaporation process taking place in an electrospray. Five proteins (Trp-cage, insulin, the C-terminal fragment of the ribosomal protein L7/L12 called ctf, ubiquitin and lysozyme) have been simulated, first in bulk water, where they were allowed to equilibrate, and later in vacuum. In the vacuum simulations the proteins were either completely naked, i.e. all water has been depleted, or covered by a water layer of 3 or 6 Å, including any occurrence of structural waters. The native solution phase charges were used on the proteins covered by a water layer whereas the naked proteins were given charges in accordance with published experimental

data on proteins in vacuum [5, 1, 46, 50, 44]. For each protein and water layer (0, 3 and 6 Å), three separate simulations were prepared to obtain some statistics, and the simulations were run for 10 ns at 300 K using classical MD at constant total energy. Analyses were done from the last ns of the simulations, where between 10 and 20 % of the initial water layer had evaporated, causing a temperature decrease of approximately 50 K. It is found that the increase in the number of intramolecular hydrogen bonds, going from a fully solvated protein to the naked, is prominent, 30-50 %. Interestingly, this doesn't have any pronounced effect on the overall size of the protein - only a slight decrease in radius of gyration and total surface area can be observed. However, as the surrounding water layer evaporates, more and more hydrophobic parts are found to be exposed on the surface of the protein. So, does this mean that the "inside-out hypothesis" still holds? The answer is no. Analysis of the secondary structure contents reveals that the conformational changes observed in the simulations where the protein is initially covered by a water layer of 3 Å, are

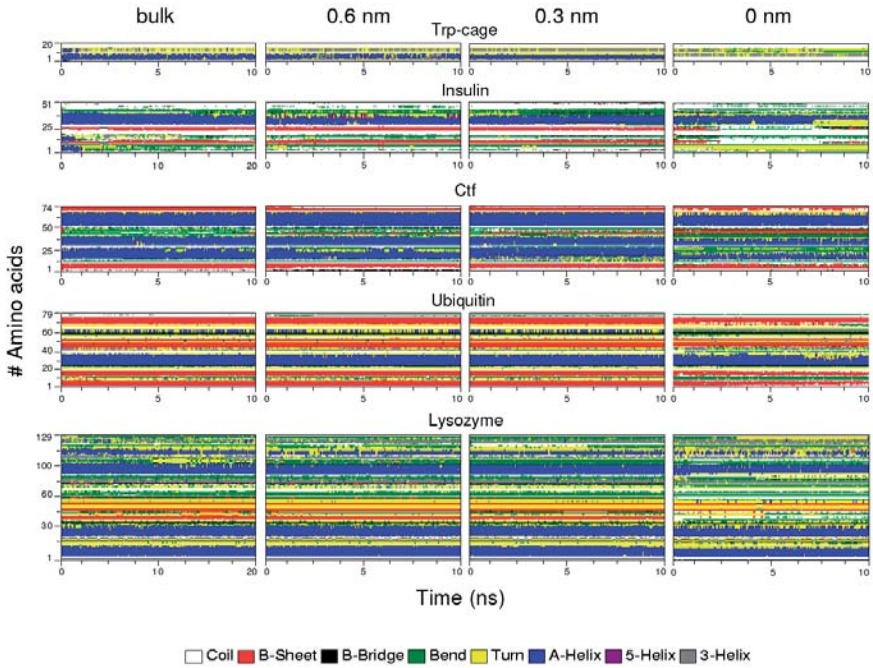

Figure 4.1: Secondary structure content of the five proteins Trp-cage, insulin, ctf, ubiquitin and lysozyme as a function of time, when run in bulk water, with a water layer of 6 Å, 3 Å and with no water at all (0 Å).

similar to the normal fluctuations observed in bulk water (Figure 4.1), hence, a very thin layer, with approximately the size of two water molecules, is enough to emulate a bulk water environment. Also, when all water is gone, as in the 0 Å simulations, the secondary structure is largely intact but dislocated, and β-sheets are somewhat better preserved than α-helices (Figure 4.1). The internal consistency between dehydrated structures (Figure 4.2) obtained from different simulations, suggests a collapse of the solution structure into something that can be defined as a gas phase structure, or an ensemble of structures, during the final step of evaporation. This "dry" ensemble is thereafter restricted to very limited dynamics. A deeper analysis of the hydrogen bonding patterns shows that, depending on the size of the protein, between 30 and 50 % of the intramolecular hydrogen bonds present in the fully solvated conformation can also be found in this ensemble of vacuum structures. This suggests that the structural changes caused by the evaporation are limited as long as the protein is in its native conformation when transferred from the solution into vacuum.

The results presented in **paper IV** are strengthened by results published earlier in **paper III**. Here the quantitative relationship between structural features of the model protein Trp-cage in gas and solution is investigated. In contrast to **paper IV**, this study is based on a number of different methods: circular dichronism (CD) and charge state distribution (CSD) measurements in electrospray are used to track the solution phase properties, while gas phase features are revealed by tandem mass spectrometry (MS/MS), CSD and molecular dynamics simulations. As in **paper II** (section 2.1.1), a number of stereoisomers constructed from a single L to D mutation of each of the first seven N-terminal residues were used for the study, and the native all-L conformer was included as a reference structure. Both CD and CSD data suggest a compact structure with a high α-helical content for the D-Asn^1 isomer in solution, which is exactly what is observed from the simulated structure in the gas phase. Also the D-Leu^7 isomer is found to have high content of 3_{10}-helix and very little α-helix, both in the simulated gas phase structures and in the CD data of the solution conformer. The simulated gas phase structures differ from the native conformation with an RMSD of 3-4 Å, and they all posses an almost twofold increase in intramolecular hydrogen bonds compared to solution structures - results that are in very good agreement with the calculations presented in **paper IV**. Hence, this study confirms that conformational features can be similar in the gas phase and in solution and that effects as subtle as that of a single D-amino acid

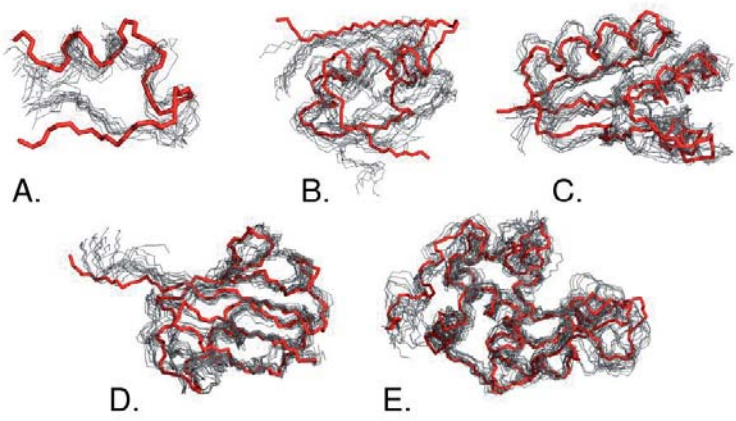

Figure 4.2: Snapshots taken from every second nanosecond of the three 0 Å simulations (gray) superimposed on the PDB structure (red): A) Trp-cage, B) insulin, C) ctf, D) ubiquitin and E) lysozyme.

substitution can be detected in both phases. Despite its small size and the difference in structure in the gas phase, as indicated by the calculated RMSD, most of the structural features of the solution phase conformation of the Trp-cage protein can be observed in vacuum as well.

Proteins in different environments (gas vs. solution) are usually charged differently. For instance, the Trp-cage, is a zwitterion with net charge 1+ spread over five charge groups in solution, while in ESI, a mixture of 2+ and 3+ non-zwitterionic structures is observed, with one of the charges relocated from Lys^8 to Gln^5 [1]. However, it is possible to produce zwitterions with a net charge of 1+ in the gas phase as well, by optimizing the source condition in the ESI [35]. These zwitterions can in principle have a few possible charge configurations, but molecular dynamics simulations indicate strongly that the most stable ion, in terms of free energy, has the same charge locations as in solution. This result, which is presented in **paper VI**, opens up a unique opportunity to compare structures from two phases with no prior need to take differences in charges into consideration. Interestingly, the gas phase ion and the native solution phase structure (i.e. the NMR structure) are very similar (Figure 4.3), with an average violation from the experimentally determined intramolecular distances of 0.14 Å only. Nevertheless, the gas phase ion has twice as many intramolecular hydrogen bonds as the NMR structure, just as is found in **paper III** and **IV**, and four

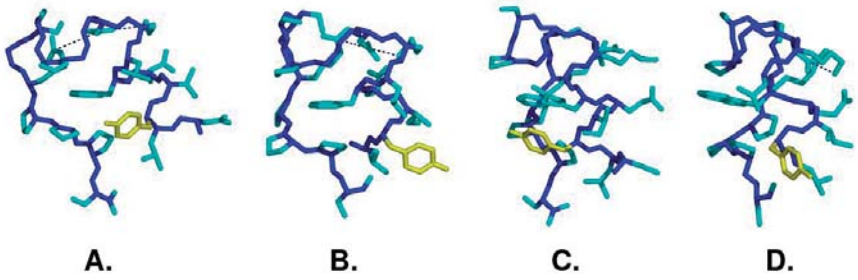

Figure 4.3: Comparison of the solution (NMR) structure (A. and C.) and the most stable 1+ zwitterion in the gas phase (B. and D.), shown from two directions. Backbone atoms are colored blue and side chains cyan. Tyr3 is colored in yellow. The distances between the side chains making up the salt bridge are shown as dark blue dotted lines.

of them are identical to hydrogen bonds observed in solution, i.e. 50 % of the native hydrogen bonds are present also in vacuum. The majority of the violated distances corresponds to atoms belonging to residues making up the hydrophobic core in solution, indicating that they are no longer restrained to the inner part of the protein as they are in solution. The salt-bridge making up the zwitterionic structure is found to be significantly stronger in the gas phase, with much shorter donor-acceptor distances between the three charged residues than in solution. Interestingly, this strong salt-bridge is not believed to be the major contribution to the stability of the structure of the 2+ ion in the gas phase. Laser induced fluorescence measurements in combination with molecular dynamics calculations (**paper V**) of a mutated Trp-cage, in which the salt-bridge has been purposely broken, indicate no change in stability in the gas phase compared to the native structure. On the other hand, a 50 % decrease in fluorescense is observed in solution, which indicates a significant disruption of the solution structure.

4.3 Structural waters and protein structure

The importance of structural (crystal) waters for protecting protein structures in the gas phase has also been investigated. Hydrogen bond analyses of the simulations in **paper IV** show that 25-50 % of the water molecules bound in the PDB structure is also bound to the protein after equilibration in bulk water and 10 ns in vacuum. Although it cannot be expected that the same water molecule will be

bound in the same position forever, this indicates that the structural waters are little affected by the evaporation, at least as long as there are other waters left in the surrounding water layer. New simulations have been performed where structures of ubiquitin and lysozyme have been extracted from the bulk water simulation used in **paper IV**, together with only those equilibrated water molecules which are bound at the same position as the waters in the crystal (PDB) structure. From these protein-crystal water complexes, two sets of simulations have been prepared, one with native solution charges and one with vacuum charges. Preliminary results show varying degrees of structure protection depending on which charges were used - some simulations give RMSD values and secondary structure contents similar to the 3 Å simulations whereas others are comparable to the naked proteins in the 0 Å simulations. Further analysis of these simulations is necessary before we can make conclusive statements about the importance of these particular binding sites for protein structure. It is, however, tempting to speculate about the absolute minimum number of waters necessary to protect the structure from unfolding. The simulations with just structural waters seem to indicate that less than 3 Å worth of water layer might still offer considerable protection.

4.4 Folding, unfolding and conformational space

The folding process of a protein in solution, going from the extended polypeptide chain to the native fold, is believed to be guided to a large extent by the hydrophobic effect [7, 45]. In the gas phase, where the driving force of hydrophobic interactions is absent, other factors are more important. In **paper I**, a number of folding simulations of the decapeptide chignolin are presented. The simulations are performed both in vacuum and in solution, and two different simulation techniques are used - classical molecular dynamics and REMD (section 3.2). It is found that, despite the use of elevated temperatures and simulations that are twice as long as in solution, the peptide never succeeds to reach the native configuration in vacuum. In the solution simulations, in contrast, the peptide undergoes several folding and unfolding events. Structure analysis of the trajectories reveals that in vacuum, the peptide collapses to a certain conformation very rapidly, within a few hundred picoseconds. This conformation, which differs from the native solution phase structure (NMR) by 4.6 Å RMSD, is then maintained

throughout the μs simulation. This finding is in good agreement with the observation of structurally consistent ensembles in the electrospray simulations presented in **paper IV**, and strengthens the hypothesis that such ensembles result from fast conformational changes during the last step of evaporation.

In order to quantify the conformational space sampled by the decapeptide in vacuum and in solution, three dimensional free energy landscapes were constructed (Figure 4.4). The three variables selected were the head-to-tail distance, which measures the degree of compactness of the fold, the distance between the C^{γ} atoms of Tyr^2 and Trp^9, which is indicative of hydrophobic contacts, and the number of main chain hydrogen bonds. It is obvious that the vacuum landscape at 296 K is only a subset of the landscape in solution. In vacuum, the peptide visits almost no extended conformations, and the global minimum is located in a position different from that in solution. The head-to-tail distance is larger in vacuum but the Tyr-Trp interaction is almost the same. The observed initial collapse of the structure in vacuum is found to be guided by backbone hydrogen bond formation rather than side-chain contacts, which supports the notion that hydrophobicity is unimportant in vacuum.

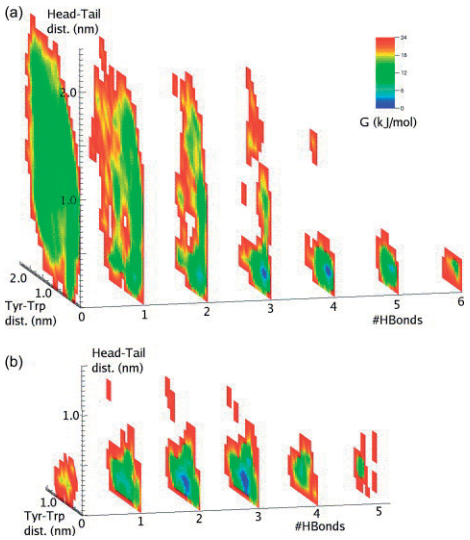

Figure 4.4: Three-dimensional Gibbs free energy landscapes as a function of two intramolecular distances and the number of main-chain hydrogen bonds, at 296 K in a) solution and b) vacuum.

5. Conclusions and future perspective

Development of conventional structural and spectroscopic methods, like X-ray crystallography and NMR spectroscopy, has laid the foundation for understanding protein structure, function and dynamics in crystals and in solution. It is uncertain whether this knowledge can be translated directly to estimate the gas phase structures of macromolecules. It is also unknown what structural rearrangements to expect when moving from a solution to the gas phase, and on what time scale these changes might take place. One theory is that proteins in the gas phase take on an "inside-out" conformation, with very little resemblance to the native solution conformation [73]. More recent research indicate, however, that gas phase structures are "similar but not identical" to proteins in solution [60]. The aim of this thesis has been to investigate the effects of dehydration on protein structures as they are moved from the native aqueous environment into vacuum, and to find out what are the major differences and similarities between protein structures and dynamics in solution and in the gas phase. I have used computer simulations (molecular dynamics) and my results have been validated by extensive comparison to experiments performed by collaborators.

Based on **paper III - VI**, included in this thesis, it can be concluded that gas phase structures differ from solution structures in some respects, mainly because of the absence of the hydrophobic effect in vacuum and the increased importance of electrostatic interactions compared to in solution. In the gas-phase, the number of intramolecular hydrogen bonds increases, salt-bridges become stronger, and hydrophobic parts are more frequently exposed on the surface of the proteins. Nevertheless, much of the significant features are preserved, including specific, structurally important, hydrogen bonds and secondary structure elements.

Our simulations show that structural changes take place over ns-μs during transition from solution to the gas phase. Free energy landscapes (**paper I**) and the simulated electrospray process (**paper IV**) indicate

a collapse of the solution phase structure into a single, or closely related ensemble, of vacuum structures. If the protein is in its native conformation in the solution, just before it is transferred into the gas phase (**paper IV**), then the resulting gas phase structure has high structural similarity to the native solution structure - most of the secondary structure is intact and a significant fraction of the native hydrogen bonding network is preserved. On the other hand, proteins do not refold in vacuo: if the protein is denatured when brought into vacuum (**paper I**), the resulting, compact, gas phase structure is different from both the native structure in solution, and the structure obtained from the native protein in the gas phase.

Results presented in **paper IV** suggest that a proteins native structure can be protected in vacuo by preserving a very thin water layer around the molecule. The layer need not be thicker than approximately two water molecules, including structural waters. This is of particular interest for single particle imaging experiments using an XFEL, where tuning the electrospray conditions could be used to preserve a certain amount of water around the sample.

It may also be possible to regain the "true" native structure from gas phase structures by performing a computational re-solvation procedure [40], where the gas phase structure is allowed to relax in a simulated bulk water environment. Simulations of this kind are running at this very moment.

The importance of structural (crystal) waters for protection of protein structures in the gas phase has also been investigated, and we have found (**paper IV**) that 20-50 % of the water molecules bound in the PDB structure is also bound to the protein after 10 ns of evaporation in vacuum. Whether the structural waters alone are enough to preserve the native structure in the gas phase is, however, still speculative - the preliminary results obtained so far require further analysis before any conclusions can be drawn.

Paper II suggests that the N-C_α bond cleavage obtained from electron capture dissociation (ECD) of peptides and proteins, may be guided by the presence of neutral hydrogen bonding to the corresponding carbonyl. A new mechanism is therefore proposed, where the electron is captured on a neutral hydrogen bond between a backbone carbonyl and any other part of the protein. Subsequent hydrogen atom transfer to the carbonyl gives a labile aminoketyl which fragments through

N-C_α bond cleavage [80]. The most loosely bound ionization proton will then neutralize the backbone nitrogen via intra-molecular proton transfer [36].

The finding that native charge locations are, to some extent, preserved in the gas phase (**paper VI**), supports continued use of mass spectrometric methods to investigate possible charge locations also in solution. Conformational changes and non-covalent interactions in the gas phase can be analysed using either mass spectrometry or fluorescence probe spectroscopy, and we find that structural changes due to point mutations can be observed both in solution and in vacuum (**paper III** and **V**). A requirement for using gas phase methods to draw conclusions about structure and dynamics in solution, is, however, that we understand what structural features are changed in the gas phase, and how to translate these changes into meaningful information about structures in solution. The studies that I have presented in this thesis provide a link between these processes and give ideas about further investigations and analyses.

6. Sammanfattning på svenska

Proteiner spelar en viktig roll i näst intill alla biologiska processer och är därför på många sätt avgörande för vår överlevnad och vårt välmående. För att ett protein ska vara verksamt krävs det att det har en tredimensionell rymdstruktur som är korrekt. Denna rymdstruktur, som bestäms utifrån proteinets aminosyrasekvens, vilken i sin tur kodas av gener i vårt DNA, är unik för varje enskild typ av protein och helt avgörande för dess funktionella förmåga. Man vet idag att ett flertal sjukdomar, så som Alzheimer och Creutzfeldt-Jacobs sjukdom, orsakas av att proteiner aggregerat till följd av att de fått en felaktig struktur. Man vet ocksa att mutationer i ett enda viktigt protein, kallat p53, är den huvudsakliga orsaken till ett flertal cancertyper. Därför är många av de läkemedel som vi använder tillverkade för att interagera med väl utvalda proteiner i vår kropp, för att på så sätt antingen förhindra eller påskynda vissa förutbestämda processer. För att kunna skräddarsy läkemedel krävs det emellertid kunskap om strukturen hos det protein med vilket läkemedlet är tänkt att interagera. Därför är utvecklingen av metoder för strukturbestämning av proteiner viktig för framtagandet av nya läkemedel.

Hittills har man genererat nästan 50 000 proteinstrukturer med hjälp av metoder som röntgenkristallografi och NMR (nuclear magnetic resonance). Dock finns det en stor grupp proteiner som inte går att strukturbestämma med dessa tekniker. Till denna grupp hör membranproteiner, glykoproteiner och diverse flexibla proteinkomplex. Behovet av nya, alternativa, sätt att generera information om proteiners struktur och dynamik är därför stort.

En samling tekniker som under senare år visat sig mycket användbara för studier av molekylära interaktioner är masspektrometri. Teknikerna har inte kapacitet att generera strukturer på det sätt som röntgenkristallografi och NMR kan, men ger information om massa och laddning hos molekyler i vakuum. Genom att kombinera flera olika varianter av tekniken kan man dessutom få fram den primära aminosyrasekvensen hos okända proteiner. Användandet av

masspektrometri bygger på att man kan skapa joniserade, laddade, produkter av de molekyler man vill analysera. En välbeprövad metod för detta ändamål är elektrosprayteknikten. Eftersom den anses vara förhållandevis skonsam och tycks bevara även svaga interaktioner inom molekyler, så har den blivit extra populär för analys av proteiner - så populär att den till och med är tänkt att användas för jonicering och injicering av molekyler vid framtida strukturbestämning med hjälp av frielektronlasrar.

Användandet av gasfasmetoder som elektrospray och masspektrometri är stadigt utsatt för frågor huruvida de resultat som erhålls verkligen är applicerbara på proteiner i deras naturliga miljö, det vill säga i lösning. Det är osäkert exakt vilken inverkan elektrosprayprocessen, och då i synnerhet den avdunstning av omkringliggande vatten som sker i fasövergången från lösning till vakuum, har på proteiners struktur, och det kvarstår en hel del obesvarade frågor kring hur proteiner faktiskt ser ut i gasfas. Lite olika teorier förekommer, några mer osannolika än andra, men enligt den senaste tidens forskning så verkar det mest troligt att proteiner i gas-fas är "lika men inte identiska" med proteiner i lösning.

Syftet med denna avhandling har i huvudsak varit att undersöka hur proteiners struktur påverkas av den förflyttning från sin nativa miljö i lösning till vakuum, som sker i en elektrospray och även inom masspektrometrin. Mina studier bygger helt och hållet på datasimuleringar men mina resultat har validerats och jämförts med omfattande experiment, genomförda av olika samarbetspartners. Baserat på de resultat som finns presenterade i **artikel III - VI**, så är slutsatsen att proteiners struktur i gas-fas skiljer sig från den nativa strukturen i lösning, och att det beror av de förändringar i växelverkan mellan elektrostatiska och hydrofoba interaktioner som sker vid en övergång till gas-fas. Antalet intramolekylära vätebindningar ökar, saltbryggor blir starkare och mer och mer hydrofoba delar exponeras på ytan av proteinet. Trots detta så är en större del av den nativa strukturella informationen väl bevarad - sekundärstrukturer och strukturellt viktiga vätebindningar återfinns även i gasfas. I vissa fall (**artikel VI**) har vi till och med sett att nativa laddningar finns bevarade, vilket stärker fortsatt användande av masspektrometri i syfte att lära oss om laddningsfördelningar och icke-kovalenta interaktioner även i lösning.

Vid en fasövergång från vätska till gas, så tycks strukturella förändringar ske snabbt (ns-μs). Nativa proteiner som förs in i gasfas

resulterar i gasfasstrukturer som är mycket lika den nativa strukturen (**artikel IV**). Denaturerade proteiner däremot, veckar ihop sig till kompakta strukturer som skiljer sig markant både från den nativa konformationen i lösning och den gasfasstruktur som fås från den nativa strukturen (**artikel I**). Resultat (**artikel IV**) antyder dock att det bör kunna gå att skydda den nativa vätskestrukturen hos ett protein även i gasfas genom att bevara en mycket liten mängd vatten runt proteinet - ett vattenlager av mindre än två molekylers tjocklek tycks vara tillräckligt för att bevara proteinet i dess nativa konformation. Detta är intressant, inte minst för framtida försök till strukturbestämning med frielektronlaser - genom att justera elektrosprayprocessen så att en viss mängd vatten bevaras runt proteinet inne i vakuumkammaren så bör man kunna öka sannolikheten för avbildning av nativa strukturer.

Vi hyser även förhoppning om att kunna använda strukturdata från gasfasstrukturer till att återskapa ett proteins nativa struktur, så som det ser ut i dess naturliga miljö i lösning: genom en simulerad återsolvatiseringsprocedur, där gasfasstrukturen får röra sig fritt i en låda fylld med vatten, så bör proteinet kunna återbilda resten av de vätebindningar och sekundärstrukturer som det har i sitt nativa tillstånd och på så vis hitta tillbaka till sin korrekta struktur. Simuleringar av detta slag pågår redan för fullt.

En förutsättning för att vi ska kunna utnyttja gasfastekniker till att öka vårt kunnande om proteiner även i deras naturliga miljö, är att vi lär oss förstå hur strukturella och dynamiska egenskaper förändras i vakuum. Min förhoppning är att de studier som ligger till grund för denna avhandling ska kunna bidra till ökad förståelse för proteiners struktur och dynamik i vakuum, samt ge ideer om nya, fortsatta, studier.

7. Acknowledgements

Thanks to...

David van der Spoel
Janos Hajdu
Roman Zubarev
Erik
Chris Adams, Frank Kjeldsen
Joel Parks
Rich London
Hanna & Abraham
Calle, Magnus, Gösta
Nic, Sara
Per
Marvin, Filipe, Michiel, Johan
Daniel
Karin, Inger, Martin
Florian, Gunilla
Tomas, Biancha
Rosie
Sigrid
Jens C
Talal
Marie-Louise L
Pappa & Majlis
Kerstin
Olle

8. References

[1] C. M. Adams, B. A. Budnik, K. F. Haselmann, F. Kjeldsen, and R. A. Zubarev. Electron capture dissociation distinguishes a single D-amino acid in a protein and probes the tertiary structure. *J. Am. Soc. Mass Spectrom.*, 15:1087–1098, 2004.

[2] B. J. Alder and T. E. Wainwright. Phase transition for a hard sphere system. *J. Chem. Phys.*, 27:1208–1209, 1957.

[3] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen. Gromacs: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.*, 91:43–56, 1995.

[4] M. Bergh, N. Tîmneanu, and D. van der Spoel. A model for the dynamics of a water cluster in a X-ray FEL beam. *Phys. Rev. E*, 70:051904, 2004.

[5] K. Breuker, H. Oh, D. M. Horn, B. A. Cerda, and F. W. McLafferty. Detailed unfolding and folding of gaseous ubiquitin ions characterized by electron capture dissociation. *J. Am. Chem. Soc.*, 124:6407–6420, 2002.

[6] K. Breuker, H. B. Oh, B. K. Carpenter, and F. W. McLafferty. Nonergodic and conformational control of the electron capture dissociation of protein cations. *Proc. Natl. Acad. Sci.*, 101:14011–14016, 2004.

[7] J. W. Bryson, S. F. Betz, H. S. Lu, D. J. Suich, H. X. Zhou, K. T. ONeil, and W. F. DeGrado. Protein design: A hierarchic approach. *Science*, 270:935–941, 1995.

[8] H. N. Chapman, A. Barty, M. J. Bogan, S. Boutet, M. Frank, S. P. Hau-Riege, S. Marchesini, B. W. Woods, S. Bajt, R. A. London, E. Plönjes, M. Kuhlmann, R. Treusch, S. Düsterer, T. Tschentscher, J. R. Schneider, E. Spiller, T. Möller, C. Bostedt, M. Hoener, D. A. Shapiro, K. O. Hodgson, D. van der Spoel, F. Burmeister, M. Bergh, C. Caleman, G. Huldt, M. M. Seibert, F. R. N. C. Maia, R. W. Lee, A. Szőke, N. Tîmneanu, and J. Hajdu. Femtosecond diffractive imaging with a soft-X-ray free-electron laser. *Nat. Phys.*, 12:839–843, 2006.

[9] J. F. de la Mora. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal. Chim. Acta*, 406:93–104, 2000.

[10] M. Dole, L. L. Mach, R. L. Hines, R. C. Mobley, L. P. Ferguson, and M. B. Alice. Molecular beams of macroions. *J. Chem. Phys.*, 49:2240–2249, 1968.

[11] Y. Duan and P. A. Kollman. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science*, 282:740–744, 1998.

[12] R. C. Dunbar. Photodissociation of trapped ions. *Int. J. Mass Spectrom.*, 200:571–589, 2000.

[13] N. Felitsyn, M. Peschke, and P. Kebarle. Origin and number of charges observed on multiply-protonated native proteins produced by ESI. *Int. J. Mass Spectrom.*, 219:39–62, 2002.

[14] J. B. Fenn. Ion formation from charged droplets: Roles of geometry, energy and time. *J. Am. Soc. Mass Spectrom.*, 4:524–535, 1993.

[15] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, and C. M. Whitehouse. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246:64–71, 1989.

[16] K. J. Gaffney and H. N. Chapman. Imaging atomic structure and dynamics with ultrafast X-ray scattering. *Science*, 316:1444–1448, 2007.

[17] A. E. Garcia and J. N. Onuchic. Folding a protein in a computer: An atomic description of the folding/unfolding of protein A. *Proc. Natl. Acad. Sci.*, 100:13898–13903, 2003.

[18] A. Gomez and K. Q. Tang. Charge and fission of droplets in electrostatic sprays. *Physics of Fluids*, 6:404–414, 1994.

[19] K. F. Haselmann, T. J. D. Jorgensen, B. A. Budnik, F. Jensen, and R. A. Zubarev. Electron capture dissociation of weakly bound polypeptide polycationic complexes. *Rapid Comm. Mass Spectrom.*, 16:2260–2265, 2002.

[20] S. P. Hau-Riege, R. A. London, and A. Szőke. Dynamics of biological molecules irradiated by short x-ray pulses. *Phys. Rev. E*, 69:051906, 2004.

[21] R. W. Hockney, S. P. Goel, and J. W. Eastwood. Quiet high-resolution computer models of a plasma. *J. Comp. Phys.*, 14:148–158, 1974.

[22] D. M. Horn, R. A. Zubarev, and F. W. McLafferty. Automated de novo sequencing of proteins by tandem high-resolution mass spectrometry. *Proc. Natl. Acad. Sci.*, 97:10313–10317, 2000.

[23] R. R. Hudgins, K. Hakansson, J. P. Quinn, C. L. Hendrickson, and A.G. Marshall. *Proceedings of the 50th Annual Conference on Mass Spectrometry and Allied Topics, Orlando, FL*, June 2002.

[24] R. R. Hudgins, F. Huang, G. Gramlich, and W. M. Nau. A fluorescence-based method for direct measurement of submicrosecond intramolecular contact formation in biopolymers: An exploratory study with polypeptides. *J. Am. Chem. Soc.*, 124:556–564, 2002.

[25] K. Hukushima and K. Nemoto. Exchange monte carlo method and application to spin glass simulations. *J. Phys. Soc. Jpn.*, 65:1604–1608, 1996.

[26] A. T. Iavarone, D. Duft, and J. H. Parks. Shedding light on biomolecule conformational dynamics using fluorescence measurements of trapped ions. *J. Phys. Chem. A*, 110:12714–12727, 2006.

[27] A. T. Iavarone, J. Meinen, S. Schulze, and J. H. Parks. Fluorescence probe of polypeptide conformational dynamics in gas phase and in solution. *Int. J. Mass Spectrom.*, 253:172–180, 2006.

[28] A. T. Iavarone and J. H. Parks. Conformational change in unsolvated Trp-cage protein probed by fluorescence. *J. Am. Chem. Soc.*, 127:8606–8607, 2005.

[29] J. V. Iribane and B. A. Thomson. On the evaporation of small ions from charged droplets. *J. Chem. Phys.*, 64:2287–2294, 1976.

[30] M. F. Jarrold. Unfolding, refolding and hydration of protein ions in the gas phase. *Acc. Chem. Res.*, 32:360–367, 1999.

[31] W. L. Jorgensen and J. Tirado-Rives. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl. Acad. Sci.*, 102:6665–6670, 2005.

[32] M. Karas and F. Hillenkamp. Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons. *Anal. Chem.*, 60:2299–2301, 1988.

[33] M. Karplus and J. A McCammon. Molecular dynamics simulations of biomolecules. *Nat. Str. Biol.*, 9:646–652, 2002.

[34] P. Kebarle. A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *J. Mass Spectrom.*, 35:804–817, 2000.

[35] F. Kjeldsen, O. A. Silivra, and R. A. Zubarev. Zwitterionic states in gas-phase polypeptide ions revealed by 157-nm ultra-violet photodissociation. *Chem. Eur. J.*, 12:7920–7928, 2006.

[36] A. M. Kuznetsov and J. Ulstrup. Dynamics of low-barrier proton-transfer in polar-solvents and protein media. *Chem. Phys.*, 188:131–141, 1994.

[37] L. J. Lapidus, W. A. Eaton, and J. Hofrichter. Measuring the rate of intramolecular contact formation in polypeptides. *Proc. Natl. Acad. Sci.*, 97:7220–7225, 2000.

[38] N. Leymarie, C. E. Costello, and P. B. O'Connor. Electron capture dissociation initiates a free radical reaction cascade. *J. Am. Chem. Soc.*, 125:8949–8958, 2003.

[39] E. Lindahl, B. A. Hess, and D. van der Spoel. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Mod.*, 7:306–317, 2001.

[40] Y. Mao, M. A. Ratner, and M. F. Jarrold. Molecular dynamics simulations of the rehydration of folded and unfolded cytochrome c ions in the vapor phase. *J. Am. Chem. Soc.*, 123:6503–6507, 2001.

[41] J. A. McCammon, B. R. Gelin, and M. Karplus. Dynamics of folded proteins. *Nature*, 267:585–590, 1977.

[42] F. W. McLafferty. Tandem mass-spectrometry. *Science*, 214:280–287, 1981.

[43] J. A. McLean, B. T. Ruotolo, K. J. Gillig, and D. H. Russell. Ion mobility-mass spectrometry: a new paradigm for proteomics. *Int. J. Mass Spectrom.*, 240:301–315, 2005.

[44] M. Miteva, P. A. Demirev, and A. D. Karshikoff. Multiply-protonated protein ions in the gas phase: Calulation of the electrostatic interactions between charged sites. *J. Phys. Chem. B*, 101:9645–9650, 1997.

[45] V. Munoz, P. A. Thompson, J. Hofrichter, and W. A. Eaton. Folding dynamics and mechanism of β-hairpin formation. *Nature*, 390:196–199, 1997.

[46] E. J. Nettleton, P. Tito, M. Sunde, M. Bouchard, C. M. Dobson, and C. V. Robinson. Characterization of the oligomeric states of Insulin in self-assembly and amyloid fibril formation by mass spectrometry. *Biophys. J.*, 79:1053–1065, 2000.

[47] R. Neutze, G. Huldt, J. Hajdu, and D. van der Spoel. Potential impact of an X-ray free electron laser on structural biology. *Radiat. Phys. Chem.*, 71:905–916, 2004.

[48] H. Neuweiler, S. Doose, and M. Sauer. A microscopic view of miniprotein folding: Enhanced folding efficiency through formation of an intermediate. *Proc. Natl. Acad. Sci.*, 102:16650–16655, 2005.

[49] H. Neuweiler, A. Schulz, M. Bohmer, J. Enderlein, and M. Sauer. Measurements of submicrosecond intramolecular contact formation in peptides at the single-molecule level. *J. Am. Chem. Soc.*, 125:5324–5330, 2003.

[50] H. B. Oh, K. Breuker, S. K. Sze, Y. Ge, B. K. Carpenter, and F. W. McLafferty. Secondary and tertiary structures of gaseous protein ions characterized by electron capture dissociation mass spectrometry and photofragment spectroscopy. *Proc. Natl. Acad. Sci.*, 99:15863–15868, 2002.

[51] C. Oostenbrink, A. Villa, A. E. Mark, and W. F. Van Gunsteren. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comp. Chem.*, 25:1656–1676, 2004.

[52] B. Paizs and S. Suhai. Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.*, 24:508–548, 2005.

[53] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatman, S. Debolt, D. Ferguson, G. Seibel, and P. Kollman. Amber, a package of computer-programs for applying molecular mechanics, normal-mode analysis, molecular dynamics and free-energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Comm.*, 91:1–41, 1995.

[54] N. C. Polfer, K. F. Haselmann, P. R. R. Langridge-Smith, and P. Barran. Structural investigation of naturally occurring peptides by electron capture dissociation and AMBER force field modelling. *Mol. Phys.*, 103:1481–1489, 2005.

[55] A. Rahman. Correlations of the motion of atoms in liquid argon. *Phys. Rev.*, 2:405–411, 1964.

[56] N. Rathore, M. Chopra, and J. J. de Pablo. Optimal allocation of replicas in parallell tempering simulations. *J. Chem. Phys.*, 122:024111, 2005.

[57] R. Rayleigh. On the equilibrium of liquid conducting masses charged with electricity. *Phil. Mag*, 14:184–186, 1882.

[58] C. V. Robinson, E. W. Chung, B. B. Kragelund, J. Knudsen, R. T. Aplin, F. M. Poulsen, and C. M. Dobson. Probing the nature of noncovalent interactions by mass spectrometry. A study of protein-coA ligand binding and assembly. *J. Am. Chem. Soc.*, 118:8646–8653, 1996.

[59] S. E. Rodriguez-Cruz, J. S. Klassen, and E. R. Williams. Hydration of gas-phase ions formed by electrospray ionization. *Am. Soc. Mass Spectrom.*, 10:958–968, 1999.

[60] B. T. Ruotolo and C. V. Robinson. Aspects of native proteins are retained in vacuum. *Curr. Opin. Chem. Biol.*, 10:402–408, 2006.

[61] K. Y. Sanbonmatsu and A. E. Garcia. Structure of Met-enkephalin in explicit aqueous solution using replica exchange molecular dynamics. *Prot. Struc. Func. Gen.*, 46:225–234, 2002.

[62] K. B. Shelimov and M. F. Jerrold. Conformations, unfolding and refolding of a apomyoglobin in vacuum; an activation barrier for gas-phase protein folding. *J. Am. Chem. Soc.*, 119:2987–2994, 1997.

[63] C. Simmerling, B. Strockbine, and A. E. Roitberg. All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc.*, 124:11258–11259, 2002.

[64] M. Sobczyk, I. Anusiewicz, J. Berdys-Kochanska, A. Sawicka, P. Skurski, and J. Simons. Coulomb-assisted dissociative electron attachment: Application to a model peptide. *J. Phys. Chem. A.*, 109:250–258, 2005.

[65] R. H. Swendsen and J. S. Wang. Replica Monte Carlo simulation of spin-glasses. *Phys. Rev. Lett.*, 57:2607–2609, 1986.

[66] K. Tanaka, Y. Ido, S. Akita, Y. Yoshida, and T. Yoshida. Detection of high mass molecules by laser desorption time-of-flight mass spectrometry. *2nd Japan-China joint Symposium on Mass Spectrometry*, pages 185–188, 1987.

[67] B. A. Thomson and J. V. Iribane. Field induced ion evaporation from liquid surfaces at atmospheric pressure. *J. Chem. Phys.*, 71:4451–4463, 1979.

[68] M. A. Tito, K. Tars, K. Valegård, J. Hajdu, and C. V. Robinson. Electrospray time of flight mass spectrometry of the intact MS2 virus capsid. *J. Am. Chem. Soc.*, 122:3550–3551, 2000.

[69] I. Velazquez, C. T. Reimann, and O. Tapia. Proteins in vacuo: Relaxation of unfolded lysozyme leads to folding into native and non-native structures. A molecular dynamics study. *J. Am. Chem. Soc.*, 121:11468–11477, 1999.

[70] U. H. Verkerk, M. Peschke, and P. Kebarle. Effect of buffer cations and of H3O+ on the charge states of native proteins. Significance to determinations of stability constants of protein complexes. *J. Mass Spectrom.*, 38:618–631, 2002.

[71] G. Wang and R. B. Cole. Charged residue versus ion evaporation for formation of alkali metal halid cluster ions in ESI. *Anal. Chim. Acta*, 406:53–65, 2000.

[72] T. Wang and R. C. Wade. On the use of elevated temperature in simulations to study protein unfolding mechanisms. *J. Chem. Theory Comput.*, 3:1476–1483, 2007.

[73] P. G. Wolynes. Biomolecular folding in vacuo. *Proc. Natl. Acad. Sci.*, 92:2426–2427, 1995.

[74] T. Wyttenbach and M. T. Bowers. Intermolecular interactions in biomolecular systems examined by mass spectrometry. *Annu. Rev. Phys. Chem.*, 58:511–533, 2007.

[75] M. Yamashita and J. B. Fenn. Electrospray ion source. Another variation on the free-jet theme. *J. Phys. Chem.*, 88:4451–4459, 1984.

[76] M. Yamashita and J. B. Fenn. Negative ion production with the electrospray ion source. *J. Phys. Chem.*, 88:4671–4675, 1984.

[77] H. Yang, G. Luo, P. Karnchanaphanurach, T.-M. Louie, I. Rech, S. Cova, L. Xun, and X. S. Xie. Protein conformation dynamics probed by single-molecule electron transfer. *Science*, 302:262–266, 2003.

[78] G. Oszlányi Z. Jurek and G. Faigel. Imaging atom clusters by hard x-ray free-electron lasers. *Europhys. Lett.*, 65:491–497, 2004.

[79] R. A. Zubarev, N. L. Kelleher, and F. W. McLafferty. Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc.*, 120:3265–3266, 1998.

[80] R. A. Zubarev, N. A. Kruger, E. K. Fridriksson, M. A. Lewis, D. M. Horn, B. K. Carpenter, and F. W. McLafferty. Electron capture dissociation of gaseous multiply-charged proteins is favored at disulfide bonds and other sites of high hydrogen atom affinity. *J. Am. Chem. Soc.*, 121:2857–2862, 1999.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 360

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-8300

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2007